THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

On weak and strong convergence of numerical approximations of stochastic partial differential equations

Fredrik Lindgren

CHALMERS | UNIVERSITY OF GOTHENBURG

Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg, Sweden

Gothenburg, 2012

On weak and strong convergence of numerical approximations of stochastic partial differential equations
Fredrik Lindgren
ISBN: 978-91-7385-787-1

Doktorsavhandlingar vid Chalmers Tekniska Högskola
Ny serie nr 3468
ISSN 0346-718X

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Sweden
Telephone +46 (0)31 772 1000

Printed in Gothenburg, Sweden 2012

On weak and strong convergence of numerical approximations of stochastic partial differential equations

Fredrik Lindgren

Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg

Abstract

This thesis is concerned with numerical approximation of linear stochastic partial differential equations driven by additive noise. In the first part, we develop a framework for the analysis of weak convergence and within this framework we analyze the stochastic heat equation, the stochastic wave equation, and the linearized stochastic Cahn-Hilliard, or the linearized Cahn-Hilliard-Cook equation. The general rule of thumb, that the rate of weak convergence is twice the rate of strong convergence, is confirmed.

In the second part, we investigate various ways to approximate the driving noise and analyze its effect on the rate of strong convergence. First, we consider the use of frames to represent the noise. We show that if the frame is chosen in a way that is well suited for the covariance operator, then the number of elements of the frame needed to represent the noise without effecting the overall convergence rate of the numerical method may be quite low. Second, we investigate the use of finite element approximations of the eigenpairs of the covariance operator. It turns out that if the kernel of the operator is smooth, then the number of basis functions needed may be substantially reduced.

Our analysis is done in a framework based on operator semigroups. It is performed in a way that reduces our results to results about approximation of the respective (deterministic) semigroup.

Keywords: Additive noise; Cahn-Hilliard-Cook equation; Error estimate; Finite element; Hyperbolic equation; Parabolic equation; Rational approximation; Stochastic partial differential equation; Strong convergence; Truncation; Wiener process; Weak convergence

Appended papers:

Paper I: Mihály Kovács, Stig Larsson and Fredrik Lindgren, *Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise*, in BIT Numerical Mathematics **52** (2012), 85–108.

Paper II: Mihály Kovács, Stig Larsson and Fredrik Lindgren, *Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise II. Fully discrete schemes*, in BIT Numerical Mathematics (2012) (online first) http://dx.doi.org/10.1007/s10543-012-0405-1.

Paper III: Mihály Kovács, Stig Larsson and Fredrik Lindgren, *Spatial approximation of stochastic convolutions*, in Journal of Computational and Applied Mathematics **235** (2011), (3554–3570).

Paper IV: Mihály Kovács, Stig Larsson and Fredrik Lindgren, *Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise*, in Numerical Algorithms **53** (2010), 309–320.

Contributions to the papers:

Paper I: Took part in the theoretical developments and the writing.

Paper II: Performed the major part of the theoretical developments and the writing.

Paper III: Took part in the theoretical developments and the writing.

Paper IV: Took part in the theoretical developments and wrote the computer code and performed the numerical experiments.

Acknowledgments

I would like to express my gratitude to my supervisors Stig Larsson and Mihály Kovács for their help and guidance. They have generously shared their deep mathematical knowledge during my more than five years as a graduate student. I am also very grateful for the two possibilities I had to visit University of Otago in Dunedin, New Zealand. I thank Mihály and all other I met there for making it a great experience!

I thank Mohammad Asadzadeh for taking the social responsibility of co-supervision seriously, often popping by my office during weekends and late nights to ask me how I am and chat about just everything.

Thomas Ericsson has not only been a superb teacher in various programming matters, always able to suggest a solution to any programming issue, bu he has also been a great colleague.

Bernt Wennberg has, as the supervisor of the departments consultancy business, taught me many things related to mathematical modeling and helped to broaden my mathematical skills in this respect.

I want to thank Adam Andersson and Matteo Molteni for interesting lectures and discussions on SPDE matters. They also read early manuscripts of this thesis and gave valuable comments.

I want to thank all people at the Maths department for making this a great time. Administration and housekeeping have always been excellent. Thanks to all former and current fellow PhD students for making the working environment so pleasant. This goes in particular for those (also seniors!) in the CAM and Optimization groups. Many of you have meant a lot to me. I want to mention Adam Wojciechowski and Frank Eriksson. They have turned out to be not only excellent colleagues but also great friends.

Friends outside the department have, of course, been of great importance. I thank you all for being there. The coffee breaks at Mahogny with Gustav Sjöblom have been crucial for my well-being and Anders Annikas has kept my spirits alive by dragging me around the rainy city late Friday nights, searching for the Holy Grail.

Without the critical but sensitive advice of Per Magnus Johansson this thesis may never have been accomplished.

I would have been nothing without my devoted parents. I owe everything to them. I also want to thank my wonderful sisters. I am very grateful for having them in my life!

Finally, I want to thank my girlfriend Lisa Hallquist for her patience, support and understanding. It is with great expectations and excitement I am moving out of my office and in with you!

Fredrik Lindgren
Gothenburg, November 2012

Contents

1 Introduction

This thesis is concerned with approximations of solutions to stochastic equations of the form

$$\mathrm{d}X(t) + AX(t)\,\mathrm{d}t = B\,\mathrm{d}W(t),\; t > 0; \quad X(0) = X_0, \tag{1.1}$$

where the unknown process $\{X(t)\}_{t\geq 0}$ takes values in a certain separable Hilbert space $\mathcal{H}$. The driving Wiener process $\{W(t)\}_{t\geq 0}$ takes values in another separable Hilbert space $\mathcal{U}$ and B is a bounded linear operator from $\mathcal{U}$ to $\mathcal{H}$. The operator $-A$ will always be an infinitesimal generator of a strongly continuous semigroup of bounded linear operators in $\mathcal{H}$. The solution is given by

$$X(t) = E(t)X_0 + \int_0^t E(t-s)B\,\mathrm{d}W(s), \tag{1.2}$$

where $E(t) = \mathrm{e}^{-tA}$ is the semigroup generated by $-A$. In particular, in two papers we study the so called weak error of numerical schemes for solving the stochastic wave equation, the stochastic heat equation and the linearized Cahn-Hilliard-Cook equation. We define the weak error to be

$$e_{\mathrm{w}}(T) = \mathbb{E}[G(\tilde{X}(T)) - G(X(T))],$$

where G is a functional on $\mathcal{H}$ with two bounded derivatives and $\tilde{X}(T)$ is some approximation of $X(T)$. We start by developing a general framework for analyzing $e_{\mathrm{w}}(T)$ in terms of the error between the semigroup $E(t)$ and some approximation of it. The resulting representation of the error is not at all confined to numerical approximations. The only thing assumed is that there is a well defined process $\{\tilde{Y}(t)\}_{0\leq t\leq T}$ given by

$$\tilde{Y}(t) = \tilde{E}(T)\tilde{X}_0 + \int_0^t \tilde{E}(T-t)\tilde{B}\,\mathrm{d}W(t)$$

such that $\tilde{X}(T) = \tilde{Y}(T)$. This particular form is needed since the treatment of $e_{\mathrm{w}}(T)$ simplifies if one uses the process $Y(t) = E(T-t)X(t) = E(T)X_0 + \int_0^t E(T-s)B\,\mathrm{d}W(s)$, with $X(T) = Y(T)$, instead of $X(t)$. A problem then arises in the case of time discretization with finite differences since that results in a discrete process. Thus, a time interpolation procedure must be performed between the grid points. In the cases we have studied, piecewise constant interpolation of the discrete operator seems to be sufficient.

Once the general framework is there, convergence rates are computed for the linear stochastic heat equation, the linear stochastic wave equation and the linearized stochastic Cahn-Hilliard or, as it also is denoted, the linearized Cahn-Hilliard-Cook equation.

In the other two papers we study the strong error, that is, the error in the mean square norm

$$e_{\mathrm{s}}(T) = \Big(\mathbb{E}\big[\|\tilde{X}(T) - X(T)\|_{\mathcal{H}}^2\big]\Big)^{1/2}.$$

We are mainly concerned with how truncation of the noise affects the numerical solutions. In Paper III the possibility of using frames in general and wavelet bases in particular is investigated. The idea is to write the driving Wiener process $W(t)$ in terms of the frame and then truncate the corresponding sum so that a finite number of terms is used as an approximation. An new equation of the same form but with the truncated driving process may then be formulated and discretized. Discretization is done by the finite element method and the error is analyzed for the stochastic wave and heat equations. Under certain assumptions on the spatial correlation in the noise, we provide *apriori* estimates on the number of elements in the frame needed in order to preserve the convergence rate for the finite element approximation of the original problem.

In paper IV we compute an approximate eigenbasis for Q in the finite element space and use this as a basis for the expansion of the noise. We investigate conditions under which we may truncate this expansion.

There are several reasons for studying numerical approximations of SPDEs. First, they arise in various applications such as phenomenological studies of phase separation in alloys (the Cahn-Hilliard-Cook equation, see [14] and [7]) and modeling of thin fibers in turbulent flow (the stochastic wave equation, see [64] and [16]). Parabolic equations with multiplicative noise (which means that the operator B above depends on the unknown $X(t)$) arises in population genetics and nonlinear parabolic equations are of interest in, for example, neurophysiology [16]. None of the equations in these examples are of the simple form (1.1), but the study of linearized equations with additive noise will help in future attempts to find error estimates for their nonlinear versions and this is indeed the the next step to take (steps in this direction are already taken, see Section 2 for a discussion, but a lot remains to be done).

Another motivation for this is that, most probably, modeling with infinite dimensional stochastic processes in the natural sciences is underdeveloped in comparison to both deterministic models but also in comparison to the mathematical theory of stochastic equations available today. Control of errors and available numerical packages for doing simulations will hopefully produce useful results and thereby drag more attention to the concept of stochastic partial differential equations in the future. Also when it comes to implementation, things need to be done. To increase the interest in stochastic models in applied sciences it would, we assume, be beneficial if fast and scalable programs to assemble, store and factorize covariance matrices were available, programs that are also easy to use. This is however the only add-on to standard PDE-solvers that is needed in order to get running.

A third motivation, taken from the numerical PDE point of view, is that the study of numerical SPDEs strengthens the understanding of non-smooth problems in general and the importance of concepts such as weak convergence may come with new insights and ideas to the study of also deterministic numerical analysis.

As an introduction to the appended papers we try to give a brief description of various mathematical fields that are joined to achieve detailed convergence estimates for numerical approximations of SPDEs. For this reason, in Section 4, we introduce classes of operators and functions of critical importance. This includes compact operators, closed operators and Fréchet derivatives. In Section 5 we discuss operator semigroups and their generators as a powerful framework for the study of initial value problems. We also formulate the deterministic heat, wave and Cahn-Hilliard equations in this setting. This is used in Section 6 to give a rigorous meaning to stochastic partial differential equations and their solutions. In this chapter we also define Wiener processes in infinite dimensional Hilbert spaces as well as infinite dimensional Itô integrals. Further, Itô's formula is stated in this setting. We also state a Theorem that relates the solution of (1.1) to Kolmogorov's equation. This equation is an important tool in our analysis.

In Sections 7–9 we introduce the numerical methods studied in the papers below starting with semidiscrete spatial finite element schemes in Section 7 followed by semidiscrete rational approximation in time in Section 8 and fully discrete schemes in Section 9. In all cases the methods are described, the discrete equations are formulated and error estimates for the deterministic homogeneous Cauchy problems are stated, that is, they are formulated as results about approximations of the respective semigroups. In Section 10 approximation properties of frames and wavelets are described.

Finally, in Section 11, the appended papers are introduced. First the papers concerned with weak convergence are described in Section 11.1. Important steps and constructions are pointed out and the final results are stated: weak error estimates for various numerical schemes for the linear stochastic heat equation, the linearized Cahn-Hilliard-Cook equation and the linear stochastic wave equation. In the latter case also a strong convergence estimate is proved. Following this, in Section 11.2 and 11.3, we give a somewhat briefer introduction to noise truncation. We show how truncation should be perform and state the most important results on this matter.

To start with, we give a short description of the field of numerical SPDE in Section 2 followed by a section where we fix some basic notation used throughout this thesis.

2 The state of the art

During the last two decades the study of numerical methods for stochastic partial differential equations has developed into its own vivid subfield of computational mathematics. The scholar of the field needs to combine knowledge from infinite dimensional probability theory and stochastic processes, infinite dimensional PDEs, numerical PDE, ODE and SDE theory to mention some. These are all by themselves mature sub-fields of mathematics and monographs that give a good introduction to them are available. To mention a few we have the widely used [16] that treats stochastic equations in infinite dimensions and gives an introduction to Hilbert space valued Wiener processes and integrals with respect to such. By the same authors, an equally influential book, [17], on infinite dimensional PDEs is available. The market for books on numerical PDE theory is rich and we only mention the books [10], [59] and [75] that we rely highly on. They are all concerned with finite element methods which is the most important method for spatial discretization studied in this thesis. The literature on numerical methods for ODEs is also plentiful but the treatment in [75] is sufficient for the part of this thesis that is concerned with parabolic equations. For methods for the wave equation we refer to [9]. When it comes to numerical methods for stochastic ordinary differential equations [48] is widely used.

We have chosen the semigroup approach of [16] to formulate the stochastic differential equations. This comes with the need to familiarize oneself with the theory of operator semigroups. The two monographs [27] and [66] are good starting points.

We scratch a bit on the surface of wavelet theory in one of the attached papers. This is a fairly modern field of applied mathematics, in particular when it comes to application to numerical solutions of various kinds of operator equations. The book [19] is the classic book on general wavelet theory. For numerical methods we refer to [76].

When it comes to the focal point of this thesis, numerical methods for stochastic partial differential equations, there seems not to be any monographs yet available and it might be that the field is not yet sufficiently mature for this. Many important questions are still unsolved. Thus we shall try to describe where the "numerical SPDE society" stands today by referring to the most important published works. It is fruitful to distinguish between various classes of challenges that are encountered. Most of the difficulties, not to say all, encountered in the deterministic theory are present also in the stochastic theory. As we are concerned with evolution equations the problems split up in parabolic and hyperbolic, linear and nonlinear equations where the nonlinear problems in turn split up in a wide gallery of subclasses. But we also face problems that are particular for SPDEs. The most important are the division between additive and multiplicative noise, where the first means that the noise is independent of the unknown process. In addition there are various concepts of convergence. Most studied so far is *strong convergence* meaning convergence in the expected value of some power of some norm. Perhaps more important is the concept of *weak convergence* which means convergence of the expected value of some functional applied to the solution. Less studied is convergence in probability. The latter is usually taken to mean that the probability that the difference in some metric is larger than any positive number tends to zero with the stepsize of the numerical method.

The first paper published in the field that we are aware of is [37] that arrived 1995 and where convergence in probability was proved for a quasilinear heat equation driven by space-time white noise and discretized by a finite difference method. The nonlinearity investigated is very general but the result contains no information about the rate of convergence. The year after, and seemingly independently, the paper [30] was published. Again, a quasilinear heat equation in one spatial dimension was studied, but this time driven by multiplicative noise even though the the noise was one dimensional only. However, a rate for the strong error of a spectral Galerkin scheme was proved. Other early contributions are [1], where a linear parabolic equation driven by white noise in one spatial dimension was analyzed and strong convergence rates for both a finite element and a finite difference scheme were proved and in the series [38], [31] and [32] strong error estimates with rates were proved for a quasilinear parabolic equation with multiplicative white noise, still in one spatial dimension. The paper [38] dealt with semidiscretization in time, [38] in space and [31] investigated fully discrete schemes.

All the mentioned papers made use of Green's functions and a "Brownian sheet" approach in the error analysis. In [73] a semigroup approach was used for the first time to analyze numerical approximations of the same type of quasilinear equations as above for explicit finite difference schemes and with similar results. The semigroup approach was perhaps more pronounced in [67], where strong convergence was proved for Lipschitz nonlinearities and convergence in probability for locally Lipschitz nonlinearities. The results were in accordance with the earlier. In [20] these results, with convergence rates of order $\frac{1}{2}$ were proved to be optimal in the multiplicative case but it was also shown that the rate could be doubled if the noise was taken to be additive and if weak convergence was considered. This is the first attempt to investigate weak convergence that we are aware of. In [25] the first discussions of colored noise appeared and they also investigated noise expanded in terms of Haar-wavelets. The analysis was performed in one spatial dimension but the inclusion of more regular noise made it possible to move to higher dimensions which was mentioned.

The same year, 2002, another pioneering work [39] was published. In this paper the semigroup approach was used and convergence for various Euler schemes as well as Crank-Nicolson's scheme was proved for rather general assumptions on the approximation of the generator. The high generality (C_0-semigroups on M-type 2 Banach spaces with abstract assumptions on the approximation of the generator) came, however, to an unfortunate high cost on the transparency. The results in the paper applies also to hyperbolic equations. The paper [40] by the same author is in the same flavor.

The first results on fully nonlinear equations came in [34], where strong convergence was proved for rather general equations followed by [35] and [36], where strong order was achieved under stronger assumptions on the nonlinearities. In [12] and [54] numerical approximations of the Cahn-Hilliard-Cook equation were proved to converge, but no rates were given. The former investigated convergence in probability and the latter strong convergence.

In the middle of the last decade the field gained momentum and a series of papers have appeared. In addition to the above we mention also [49], [55], [60], [68], [77], [78], [82], and [83] that all have treated strong convergence of various SPDEs. The recent paper [56] proves strong rates for a Volterra type evolution equation.

It is notable that, so far, analysis of adaptive methods for SPDEs are absent in the literature. In [69], [70] and [71] though, certain nonuniform timesteps are proved to be highly beneficial. They also prove lower bounds of strong errors. Also other attempts to improve the efficiency has been investigated recently. Milstein schemes have been investigated in for example [4] and [47] and exponential schemes in [13], [43], [44] and [46]. Multilevel methods are treated in [5] and [6].

For weak convergence of SPDEs much less has been done. The first paper devoted to the subject was [41]. The test functions used were however very restrictive. This restriction was removed in [21], where also nonlinear equations were studied. The techniques used rely on the fact that the studied semigroup, the paper treats the Schrödinger equation, is a group. In [23] weak rates for a linear heat equation with additive noise was proved and, under somewhat more restrictive assumptions, in [29]. In [22], weak rates for a temporal discretization of a nonlinear heat equation with multiplicative noise in 1-D is proved by means of Kolmogorov equations and Malliavin calculus. This result is generalized to multiple dimensions and additive noise by [80] and to spatial discretization for additive noise in multiple dimensions as well as multiplicative noise in one dimension by [2]. In the paper [58] an alternative approach to these matters is taken. The use of Kolmogorov equations is avoided but only a linear heat equation is treated. Weak convergence for a stochastic wave equation was proved in one spatial variable for testfunctions depending on the whole trajectory and not only on the final state in [42]. In the paper [63] the results of [23] was generalized to impulsive noise.

The two first papers in this thesis, also found in [52] and [53], treat weak convergence of linear SPDEs with additive noise. The former is concerned with semidiscretization in space using finite elements and proves convergence results for the stochastic heat and wave equations as well as for the linearized CHC equation. In [53] semidiscrete temporal and fully discrete approximations for the wave equation and fully discrete schemes for the linearized CHC equation are treated. In all cases the weak convergence rates are found to be essentially twice the strong rate, a pattern that is well known from the study of numerical SDEs and also have been apparent in other papers treating weak convergence for SPDEs. The results for the wave equation seem to be the first that treat weak convergence for fully discrete schemes in arbitrary spatial dimensions. However, the most important contribution to the field from these two papers is the fact that they offer a uniform approach to treat error analysis of very general linear autonomous stochastic evolution equations. This approach is based on the error formula of Theorem 11.1 in Section 11.1 below. The treatment of the stochastic heat, wave and CH equations examplifies how this formula can be utilized.

The third and fourth papers study a matter that has often been overseen: how the noise should be represented in computations. We know of only one treatment of similar matters, in [61] the possibility of using FFT on rectangular regions is investigated. The generation of noise will be the bottle-neck of all numerical computations if not treated with quite some care. Assume the finite element method with N_h degrees of freedom is used and one simply wants to project the noise on the finite element space using an orthogonal projection operator. If only the covariance operator is known and if this is an integral operator with a kernel that is strictly positive on its domain $\mathcal{D} \times \mathcal{D}$, where $\mathcal{D}$ is the spatial domain considered in the equation, then the assembly complexity is N_h^2 and on top of that a factorization procedure of complexity N_h^3 needs to be performed. When the noise is additive this might be doable but for multiplicative noise, when this needs to be repeated at every timestep, it is not. The two papers take two different approaches to this. In the third paper (see also [51]) it is shown how frames in general, and wavelet basis in particular, might be utilized to decrease the number of frame functions in the representation in comparison to, e.g., the numerical approximation of the eigenbasis of the covariance operator. In the fourth paper it is investigated how the eigenbasis itself may be truncated.

It should come as no surprise that the numerical studies have raised new theoretical questions. We would like to mention [15], where alternatives to the classical Itô formula are invetsigated for cases where the latter can not be used. The "mild Itô formula" proposed there have applications to the analysis of numerical schemes.

To end this section we mention that we are aware of two publications written with the purpose of surveying the field of numerical SPDE. The first, [33], is from 2002 and is written by one of the pioneers, I. Gyöngy. The second, [45], is written by A. Jentzen and P. E. Kloeden and was published in 2009. They contain, of course, a more detailed treatment and further references.

3 Some basic notation

We list some definitions used in the sequel. Some of the concepts mentioned will be rigorously defined when they appear for the first time in the text below whereas others will be thought of as so common that their meaning should be understood directly from the context. The notation in the appended papers differs somewhat from the this introductory text. Now, first, the letter $\mathcal{E}$ will always denote a Banach space and $\mathrm{Bor}(\mathcal{E})$ is the Borel σ-algebra on $\mathcal{E}$. We will by $\mathcal{U}$, $\mathcal{H}$ and H refer to separable Hilbert spaces, real unless otherwise stated. The Hilbert space $\mathcal{H}$ will be the space where a certain equation is posed and the space $\mathcal{U}$ the space where the involved Wiener process $\{W(t)\}_{t\geq 0}$ takes its values whereas H only will be used while developing general theory.

If $\mathcal{E}_1$ and $\mathcal{E}_2$ are Banach spaces and $T\colon \mathcal{E}_1 \to \mathcal{E}_2$ is linear and bounded, then we write $T \in \mathcal{B}(\mathcal{E}_1, \mathcal{E}_2)$ or $T \in \mathcal{B}(\mathcal{E}_1)$ if $\mathcal{E}_1 = \mathcal{E}_2$. We will usually index norms $\|\cdot\|$ and inner products $\langle \cdot, \cdot \rangle$ by the spaces they define but sometimes, when it is clear from the context, we will omit this. In most cases this means that $\|\cdot\| = \|\cdot\|_{\mathcal{H}}$ or $\|\cdot\| = \|\cdot\|_{\mathcal{B}(\mathcal{H})}$. We will repeatedly encounter bounded linear operators belonging to some Schatten class, $\mathcal{L}_p(\mathcal{E}_1, \mathcal{E}_2)$, presented in some detail in Section 4.2. If $\mathcal{E}_1 = \mathcal{E}_2$ we will again write $\mathcal{L}_p(\mathcal{E}_1)$. This short hand notation will be used in all similar cases. For an operator $T \in \mathcal{L}_1(H)$ the trace, $\mathrm{Tr}(T)$, of T is well defined and we will sometimes write $\|T\|_{\mathrm{Tr}}$ instead of $\|T\|_{\mathcal{L}_1(H)}$. If T belongs to $\mathcal{L}_2(H)$ we say that it is a Hilbert-Schmidt operator and write $\|T\|_{\mathrm{HS}} = \|T\|_{\mathcal{L}_2(H)}$.

Several function spaces will be encountered, where the functions domain of definition, $\mathcal{D} \subset \mathbb{R}^n$, almost always will be open and bounded with smooth or polygonal domain, often convex. By $L^p(\mathcal{D})$ we will denote the spaces of functions f such that $|f|^p$ has finite Lebesgue integral. With $H^k(\mathcal{D}), k$ positive integer we denote the space of functions where all derivatives up to and including the k'th belongs to $L^2(\mathcal{D})$. That is, $H^k(\mathcal{D})$ is the usual Sobolev space $W^{2,k}(\mathcal{D})$. We will take $\dot{H}^\alpha(\mathcal{D})$ to be the domain of $(-\Delta)^{\alpha/2}$ with homogeneous Dirichlet boundary conditions and $\tilde{H}^\alpha(\mathcal{D})$ with homogeneous Neumann boundary conditions and zero mean. The space $\dot{H}^1$ plays a somewhat special role equaling $H_0^1(\mathcal{D})$, the completion of $C_0^\infty(\mathcal{D})$ in $H^1(\mathcal{D})$ with respect to its usual norm. In the studies of the wave equation it turns out that the spaces $\mathcal{H}^\alpha := \dot{H}^\alpha \times \dot{H}^{\alpha-1}$ are useful.

When the letter A is used for an operator it will be a densely defined closed operator such that A or $-A$ is the infinitesimal generator of a strongly continuous semigroup $\{T(t)\}_{t\geq 0}$ of bounded linear operators. When A is some of the particular differential operators we study in detail we will use $\{E(t)\}_{t\geq 0}$ instead of $\{T(t)\}_{t\geq 0}$ to denote this particular semigroup. When A is unbounded the domain of A will be denoted by $D(A)$. For the resolvent operator we write $R(\lambda; A) = (\lambda I + A)^{-1}$.

If $G\colon H \to \mathbb{R}$ we denote its Fréchet derivative at a point x in H by $G'(x)$. The space $C_{\mathrm{b}}^2(H, \mathbb{R})$ will play a central role, consisting of continuous functions from H to $\mathbb{R}$ with two bounded continuous Fréchet derivatives but where the function itself not necessarily is bounded.

For the stochastic part a probability space $(\Omega, \mathcal{F}, P)$ is needed, Ω being the event space and $\mathcal{F}$ denoting a σ-algebra (σ-field) of subsets of Ω and P is a probability measure on the measurable space $(\Omega, \mathcal{F})$. If $X\colon (\Omega, \mathcal{F}) \to (\mathcal{E}, \mathrm{Bor}(\mathcal{E}))$ is a measurable function it will be called an $\mathcal{E}$-valued random variable and we denote by $\mathbb{E}[X] := \int_\Omega X(\omega)\,\mathrm{d}P(\omega)$, the expectation value of X. The space of all functions such that $\mathbb{E}[\|X\|_{\mathcal{E}}^p] < \infty$ will be denoted $L_p(\Omega; \mathcal{E})$. When $p = 2$ we will refer to this norm as the mean square norm. If $\mathcal{G} \subset \mathcal{F}$ is a σ-algebra, the conditional expectation of X with respect to $\mathcal{G}$ will be denoted by $\mathbb{E}[X|\mathcal{G}]$. By $\mathcal{N}(m, Q)$ we will denote the Gaussian measure on a Hilbert space H with mean $m \in H$ and covariance operator $Q \in \mathcal{B}(H)$. A filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$ of a σ-algebra $\mathcal{F}$ is a family of σ-algebras such that for $s \leq t$ it holds that $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$. The quadruple $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{0\leq t\leq T})$ is called a filtered probability space.

For the numerics we will use FEM as shorthand for the standard continuous finite element method and we will index with a mesh parameter h when we refer to some approximation on a certain triangulation $\mathcal{T}_h$. We will also use h as the largest triangle side in this triangulation. For time discretization the letter k will be used as the size of the time step in computations and as an index to indicate that a function or operator is an approximation on the corresponding temporal grid.

4 Some facts from functional analysis

We will here present some facts from functional analysis that will be used to build up the framework where stochastic partial differential equations and their numerical solutions will be studied. Compact operators play a crucial role in the analysis of infinite-dimensional stochastic processes and closed operators in semigroup theory. They will be devoted a section each below. Studying weak convergence means applying (nonlinear) functionals to the process under investigation and this comes with the need to impose restrictions on the regularity of these functionals. This will be done by introducing assumptions on their Fréchet derivatives. Fréchet derivatives and the related spaces C_{b}^k will be defined in Section 4.3.

4.1 Closed operators

Assume that $\mathcal{E}_1$ and $\mathcal{E}_2$ are Banach spaces. A closed linear operator is a linear operator $A\colon D(A) \to \mathcal{E}_2$, $D(A) \subset \mathcal{E}_1$ such that the graph of A is a closed subset of $\mathcal{E}_1 \times \mathcal{E}_2$. That is to say that whenever $u_k \in D(A)$, $u_k \to u$ and $Au_k \to f$, then $Au = f$. If $D(A) = \mathcal{E}_1$ and A is closed, then by the the Closed Graph Theorem A is bounded. If $\mathcal{E} := \mathcal{E}_1 = \mathcal{E}_2$ then the resolvent set of a possibly unbounded closed operator A in the set $\rho(A) = \{\lambda \in \mathbb{C} : \lambda I + A \text{ is one-to-one and onto}\}$. The spectrum of A, $\sigma(A)$ is the complement of $\rho(A)$ in $\mathbb{C}$. If $\eta \in \mathbb{C}$ and there exists $x \in \mathcal{E}$ such that $Ax = \eta x$ then η is an eigenvalue of A. An eigenvalue is in the spectrum of A but not all members of the spectrum are necessarily eigenvalues. If A is closed, then the operator family $\{R(\lambda; A)\}_{\lambda \in \rho(A)}$, $R(\lambda; A) = (\lambda I + A)^{-1}$ is a family of bounded operators on $\mathcal{E}$.

We will need closedness of operators to see that they generate operator families of strongly continuous semigroups. The most important case will be $A = -\Delta$ on $\mathcal{E} = L^2(\mathcal{D})$ with $D(A) = H_0^1(\mathcal{D}) \cap H^2(\mathcal{D})$ or on $\mathcal{E} = \{v \in L^2(\mathcal{D}) : \langle v, 1\rangle_{L^2(\mathcal{D})} = 0\}$ with $D(A) = \{v \in \mathcal{E} : \frac{\partial u}{\partial n} = 0\}$. Here $\mathcal{D} \subset \mathbb{R}^n$ is a bounded domain with smooth boundary. In the first case the operator is closed since the graph norm $\|\cdot\|_{L^2(\mathcal{D})} + \|A\cdot\|_{L^2(\mathcal{D})}$ is equivalent to the usual $H^2(\mathcal{D})$-norm on $H_0^1(\mathcal{D}) \cap H^2(\mathcal{D})$. Thus if $u_k \to v$ and $Au_k \to f$, then $\{u_k\}_{k=1}^\infty$ is Cauchy in H^2 and hence converges to an element in $u \in H^2$. Thus u must be v and $Av = f$ must hold. The second case is similar.

4.2 Compact operators

We will see below that certain classes of compact operators play an important role in the study of infinite dimensional random variables and stochastic processes. In particular, the covariance operator is necessarily compact for a Hilbert-space valued Gaussian random variable. Also, through the use of Kolmogorov equations in the case of weak convergence, and the Itô isometry in the case of strong convergence, we will end up in a situation where the following facts about Schatten class operators are crucial.

A *compact linear operator* $T\colon H_1 \to H_2$ is a linear operator that maps bounded sets in H_1 onto relatively compact sets in H_2. That is to say that if $\{x_k\}_{k=1}^{\infty} \subset H_1$ is a bounded sequence, then $\{Tx_k\}_{k=1}^{\infty}$ contains a convergent sub-sequence. Every compact linear operator is bounded. An operator $T \in \mathcal{B}(H_1, H_2)$ is compact if and only if there exists a sequence of finite rank operators $\{T_k\}_{k=1}^{\infty} \subset \mathcal{B}(H_1, H_2)$ that converges to T. That means that T is compact iff it may be approximated arbitrarily well with finite rank operators.

A linear operator T is compact if and only if T^* is and this is in turn equivalent to saying that T^*T is compact. Further, T^*T is positive semi-definite and self-adjoint hence there is a non-negative, decreasing sequence of eigenvalues $\{\gamma_j\}$ of eigenvalues to T^*T and a corresponding eigenbasis $\{f_j\}_{j=1}^{\infty}$. Thus the square root of T^*T may be defined through $(T^*T)^{1/2}x = \sum_{j=1}^{\infty} \gamma_j^{1/2}\langle x, f_j\rangle f_j$. It makes sense to call this operator the absolute value of T and write $|T| = T^*T^{1/2}$ since it is positive semi-definite and there exists a unitary operator U such that $T = U|T|$. The eigenvalues of $|T|$ are the singular values $\{\sigma_j\}_{j=1}^{\infty}$ of T. Both eigenvalues and singular values are counted with multiplicity. We say that a compact operator $T\colon H_1 \to H_2$ is of class $\mathcal{L}_p(H_1, H_2)$, or of the p'th Schatten class, if

$$\|T\|_{\mathcal{L}_p(H_1,H_2)}^p = \sum_{j=1}^{\infty} \sigma_j^p < \infty.$$

The functional $\|\cdot\|_{\mathcal{L}_p(H_1,H_2)}$ is indeed a norm and the spaces $\mathcal{L}_p(H_1, H_2)$ are Banach-spaces for $1 \le p < \infty$. If $T \in \mathcal{L}_p(H_1, H_2)$, then $T^* \in \mathcal{L}_p(H_2, H_1)$ and $\|T^*\|_{\mathcal{L}_p(H_2,H_1)} = \|T\|_{\mathcal{L}_p(H_1,H_2)}$. Take $p = 2$ and let $\{f_j\}$ be an ON-basis for H_1 then $\mathcal{L}_p(H_1, H_2)$ becomes a Hilbert space with the inner product

$$\langle T, S\rangle_{\mathrm{HS}} = \sum_{j=1}^{\infty} \langle Tf_j, Sf_j\rangle_{H_2}. \tag{4.1}$$

The subscript HS stands for Hilbert-Schmidt and the class $\mathcal{L}_2(H_1, H_2)$ is also called the class of Hilbert-Schmidt operators. The Hilbert-Schmidt norm induced by the inner-product in (4.1) equals the $\mathcal{L}_2(H_1, H_2)$-norm. It is an easy task to show that this inner product is independent of the particular choice of ON-basis.

If $T \in \mathcal{L}_1(H_1)$ then the trace of T,

$$\mathrm{Tr}(T) = \sum_{j=1}^{\infty} \langle Tf_j, f_j\rangle_{H_1}$$

is a well defined linear functional on $\mathcal{L}_1(H_1)$, independent of the orthonormal basis. Clearly $\mathrm{Tr}(T) = \mathrm{Tr}(T^*)$. We will repeatedly use the fact that

$$|\,\mathrm{Tr}(T)| \le \|T\|_{\mathcal{L}_1(H_1)}$$

with equality if T is positive definite and self-adjoint.

Both the trace and the Hilbert-Schmidt norms appear often in computations. Moreover, in the case of $p = 1$ there is another way to define the Schatten norm. In addition, this method directly extends to general Banach spaces. For this aim, let E_1 and E_2 be two separable Banach spaces and let $T \in \mathcal{B}(E_1, E_2)$ and $\{f_k\}_{k=1}^{\infty} \subset E_1^*$ and $\{e_k\}_{k=1}^{\infty} \subset E_2$ sequences such that

$$Tx = \sum_{j=1}^{\infty} f_k(x) e_k, \quad x \in E_1, \tag{4.2}$$

then we take the norm

$$\|T\|_{\mathrm{Tr}(E_1,E_2)} = \inf \sum_{j=1}^{\infty} \|f_j\|_{E_1^*} \|e_j\|_{E_2} \tag{4.3}$$

where the infimum is taken over all sequences $\{f_k\}_{k=1}^{\infty}$ and $\{e_k\}_{k=1}^{\infty}$ as in (4.2). The space $\tilde{L}_1(E_1, E_2) = \{T \in \mathcal{B}(E_1, E_2) : \|T\|_{\mathrm{Tr}(E_1,E_2)} < \infty\}$ is a Banach space with this norm and if E_1 and E_2 are Hilbert spaces then $\|T\|_{\tilde{L}_1(E_1,E_2)} = \|T\|_{\mathcal{L}_1(E_1,E_2)}$. Hence we will omit the ˜ in the sequel.

We now collect some facts about various Schatten class operators. First, if $T \in \mathcal{L}_p(H_1, H_2)$ and $B \in \mathcal{B}(H_3, H_1)$ or $B \in \mathcal{B}(H_2, H_3)$ then $TB \in \mathcal{L}_p(H_3, H_2)$ or $BT \in \mathcal{L}_p(H_1, H_3)$ respectively. If $\frac{1}{p} + \frac{1}{q} = 1$ and if $T \in \mathcal{L}_p(H_1, H_2)$ and $S \in \mathcal{L}_q(H_2, H_3)$, then $ST \in \mathcal{L}_1(H_1, H_3)$ and

$$\|ST\|_{\mathcal{L}_1(H_1,H_3)} \leq \|S\|_{\mathcal{L}_q(H_2,H_3)} \|T\|_{\mathcal{L}_p(H_1,H_2)}.$$

The case when $H_1 = H_2 = H_3 = H$ and $p = q = 2$ will be used frequently below. If $S, T \in \mathcal{L}_2(H_1, H_2)$ we have in addition that

$$\langle T, S \rangle_{\mathrm{HS}} = \mathrm{Tr}(S^* T)$$

and as an immediate consequence of this and (4.3)

$$|\langle T, S \rangle_{\mathrm{HS}}| \leq \|S^* T\|_{\mathcal{L}_1(H_1)}.$$

Further, if $T^*T \in \mathcal{L}_1(H_1)$, then $TT^* \in \mathcal{L}_1(H_2)$ and T and T^* are Hilbert-Schmidt operators. Indeed

$$\begin{aligned} \|T^*T\|_{\mathcal{L}_1(H_1)} = \mathrm{Tr}(T^*T) = \langle T, T \rangle_{\mathrm{HS}} = \|T\|_{\mathrm{HS}}^2 = \|T^*\|_{\mathrm{HS}}^2 \\ = \langle T^*, T^* \rangle_{\mathrm{HS}} = \mathrm{Tr}(TT^*) = \|TT^*\|_{\mathcal{L}_1(H_2)}. \end{aligned}$$

If $T, S \in \mathcal{L}_2(H_1, H_2)$ then, as mentioned before T^*T and S^*S are of trace class and it holds that

$$|\langle T, S \rangle_{\mathrm{HS}}| \leq \|T\|_{\mathrm{HS}} \|S\|_{\mathrm{HS}} = \sqrt{\mathrm{Tr}(T^*T)\, \mathrm{Tr}(S^*S)}.$$

4.3 Fréchet derivatives and related function spaces

As mentioned above we will repeatedly encounter functionals on Hilbert spaces, i.e., functions $G\colon H \to \mathbb{R}$ that are non-linear and, notably, unbounded. We will however be forced to impose some restrictions of the class of functionals allowed in order to achieve any useful results. We will do this in terms of Fréchet differentiability.

Definition 4.1. A function $G\colon H \to \mathbb{R}$ is Fréchet differentiable at a point $x \in H$ if there exists a bounded linear functional $F(x, \cdot)$ such that

$$\lim_{\substack{\|h_n\|\to 0\\ n\to\infty}} \frac{|G(x+h_n) - G(x) - F(x,h_n)|}{\|h_n\|} = 0.$$

Since $G'(x,\cdot) \in H^*$ we identify it with an element $F(x)$ in H by $F(x,h) = \langle h, F(x)\rangle$. We shall write $G'(x) = F(x)$ to denote this element whenever it exists. If G is Fréchet differentiable at every x in H, then we say that G is Fréchet differentiable on H. If G is continuous and differentiable on H and if the map $H \to H$, $x \mapsto G'(x)$ is continuous and $|G'|_{C^1_{\mathrm{b}}} := \sup_{x\in H} \|G'(x)\|_H < \infty$ then we shall write $G \in C^1_{\mathrm{b}}(H,\mathbb{R})$. If G is differentiable in a neighborhood of x and if there exists a bounded linear function $\tilde{F}(x,\cdot)\colon H \to H$ such that

$$\lim_{\substack{\|h_n\|\to 0\\ n\to\infty}} \frac{\|G'(x+h_n) - G'(x) - \tilde{F}(x,h_n)\|_{\mathcal{H}}}{\|h_n\|_{\mathcal{H}}}.$$

then we say that G is twice differentiable at x and $\tilde{F}(x,\cdot)$ is the second derivative of G at x. We will in accordance with the notation for the first derivative identify the function $\tilde{F}(x,\cdot)$ with a bounded linear operator $\tilde{F}(x) \in \mathcal{B}(\mathcal{H})$ trough $\tilde{F}(x,y) = \tilde{F}(x)y$ for all $y \in H$ and write $G''(x) = \tilde{F}(x)$. If the mapping $x \to G''(x)$, $H \to \mathcal{B}(H)$ is continuous and the semi-norm

$$|G|_{C^2_{\mathrm{b}}(H,\mathbb{R})} := \sup_{x\in H} \|G'(x)\|_H + \sup_{x\in H} \|G''(x)\|_{\mathcal{B}(H)}$$

is finite, then we write $G \in C^2_{\mathrm{b}}(H,\mathbb{R})$. That is to say that G is a continuous functional on H with two bounded continuous derivatives. We do not assume here that the function itself is bounded, only continuous and with bounded derivatives.

5 Operator semigroups

Our approach to SPDE is the semigroup approach developed in [16]. Also the point of view we take on discretization relies highly on this theory. We give a brief introduction to operator semigroups here and refer to, for example, [27] or [66] for proofs and a more thorough treatment of the subject. A neat presentation of the heat- and wave-equations in a semigroup framework can be found in [28].

A good starting-point here is the structural similarity between the first order scalar problem $\dot{v} + av = 0, t > 0;\ v(0) = v_0$ and the n-dimensional version

$$\dot{u} + \mathbb{A}u = 0,\ t > 0; \quad u(0) = u_0 \tag{5.1}$$

where $\mathbb{A} \in \mathbb{R}^{n\times n}$. The unique solutions are given by with solution $v(t) = e^{-ta}v_0$ and $u(t) = e^{-t\mathbb{A}}u_0$ respectively where the exponential of the matrix is defined by the Taylor series

$$e^{-t\mathbb{A}}x = \sum_{0}^{\infty} \frac{(-t\mathbb{A})^n}{n!}x. \tag{5.2}$$

The series is convergent since it is absolutely convergent. Indeed,

$$\Big\| \sum_{n=0}^{\infty} \frac{(-t\mathbb{A})^n}{n!}x \Big\| \le \sum_{n=0}^{\infty} \Big\| \frac{(-t\mathbb{A})^n}{n!}x \Big\| \le \sum_{n=0}^{\infty} \frac{(t\|\mathbb{A}\|)^n}{n!}\|x\| = e^{t\|\mathbb{A}\|}\|x\|.$$

In the same way the exponential e^{-tA} of a bounded operator $-tA$ on an arbitrary Banach space may be defined and the unique solution of the initial value problem

$$\dot{u} + Au = 0,\ t > 0; \quad u(0) = u_0 \tag{5.3}$$

will still be given by $u(t) = e^{-tA}u_0$. The definition of the exponential function as in (5.2) collapses however, if the operator A is unbounded. It is often possible to construct the exponential function of an unbounded operator by other means, but it turns out that an extraction of the most important properties of the scalar exponential function leads to a powerful and general approach. The main properties in mind are the semigroup properties.

Definition 5.1. A family of bounded linear operators $\{T(t)\}_{t\ge 0}$ on a Banach space $\mathcal{E}$ is called a *semigroup of bounded linear operators* if for every $x \in \mathcal{E}$

$$T(0)x = x, \tag{5.4}$$

$$T(t+s)x = T(t)T(s)x, \quad t, s \ge 0. \tag{5.5}$$

It is easy to see that the real-valued exponential function fulfills these properties and so does the $\mathbb{R}^{n\times n}$-version (see [79, pp. 192f] for a simple proof). We shall soon see cases when a family $\{T(t)\}_{t\ge 0}$ fulfilling (5.4)-(5.5) are solution operators to equations of the form (5.3) and for this these assumptions are very natural. Indeed, the first property makes sure that a function $u(t) = T(t)u_0$ fulfills the initial condition in (5.3) and the second to the fact that the system is autonomous and deterministic. It grants that if $T(t)u_0$ solves (5.3) and $T(t)u_s$ solves $\dot{u} + Au = 0,\ u(s) = u_s$ and if $t = \tau + s$, then $T(t)u_0 = T(\tau)u_s$.

In order to connect semigroups to the studies of abstract initial value problems we will need further assumptions on the semigroups we are studying, in particular on the continuity and differentiability. First we introduce the notion of an infinitesimal generator of a semigroup.

Definition 5.2. An operator A defined by

$$\begin{aligned} D(A) &= \{x \in \mathcal{E} \colon \lim_{t \to 0^+} \frac{T(t)x - x}{t} \text{ exists}\}, \\ Ax &= \lim_{t \to 0^+} \frac{T(t)x - x}{t}, \quad x \in D(A) \end{aligned} \tag{5.6}$$

is called the *infinitesimal generator* of the semigroup $\{T(t)\}$.

We say that $\{T(t)\}_{t\geq 0}$ is a *semigroup of uniformly continuous operators* on $\mathcal{E}$ if $\lim_{t\to 0^+} \|T(t) - I\| = 0$. It turns out that the generator of a uniformly continuous semigroup necessarily must be bounded so the concept doesn't help us much if to study unbounded operators such as differential operators. A natural relaxation is to study semigroups that only converges point-wise to the identity.

Definition 5.3. A family of bounded linear operators $\{T(t)\}_{t\geq 0}$ on $\mathcal{E}$ is called a *strongly continuous semigroup* on $\mathcal{E}$ if, in addition to (5.4) –(5.5),

$$\lim_{t\to 0^+} \|T(t)x - x\| = 0, \quad x \in \mathcal{E}.$$

We will often refer to a strongly continuous semigroup as a C_0-semigroup. It is not difficult to show that for any C_0-semigroup there exists constants $M \geq 1$, $\omega \geq 0$ such that $\|T(t)\| \leq Me^{t\omega}$ for every non-negative t. If M and ω are such numbers, then $\{T(t)\}_{t\geq 0}$ is said to be of *type* (M, ω). If a C_0-semigroup is of type $(M, 0)$ it is said to be a *uniformly bounded C_0-semigroup* and if $M = 1$, then it is a *C_0-semigroup of contractions*.

A weaker concept than uniformly continuous, but stronger than strongly continuous, is the notion of an *analytic semigroup*. To define this, set $\Delta_\phi = \{z = \mathbb{C} \colon |\arg(z)| < \phi,\ \phi > 0\}$, a sector centered around the positive real axis.

Definition 5.4. A family of bounded linear operators $\{T(z)\}_{z\in\Delta_\phi}$ on $\mathcal{E}$ is said to be an *analytic semigroup* in Δ_ϕ if

$$z \mapsto T(z) \text{ is analytic in } \Delta_\phi; \tag{5.7}$$

$$T(0) = I \text{ and } \lim_{\Delta_\phi \ni z \to 0} T(z)x = x, \quad x \in \mathcal{E}; \tag{5.8}$$

$$T(z_1 + z_2) = T(z_1)T(z_2), \quad z_1, z_2 \in \Delta_\phi. \tag{5.9}$$

We will refer also to the restriction of an analytic semigroup to the positive real axis as an analytic semigroup. An analytic semigroup is strongly continuous but not necessarily the other way around.

When studying equations of the form (5.1) the interesting question is really if the operator is an infinitesimal generator of a semigroup of a certain kind. We will thus state a few theorems that gives certain characterizations of generators.

Theorem 5.5 (Hille-Yosida). *A linear, possibly unbounded operator, $-A$ is the infinitesimal generator of a C_0-semigroup of contractions if and only if A is closed and densely defined in $\mathcal{E}$ and the resolvent set of A contains all negative real numbers and*

$$\|R(\lambda; A)\| \leq \frac{1}{\lambda}, \quad \lambda > 0.$$

A general result concerning generators of C_0-semigroups reads as follows.

Theorem 5.6. *A linear operator $-A$ is the infinitesimal generator of a strongly continuous semigroup, or a C_0-semigroup, of type (M,ω) if and only if in the assumptions of Theorem 5.5 holds except that the assumption in* (5.5) *is replaced by*

$$\|R(\lambda;A)^n\| \leq \frac{M}{(\lambda-\omega)^n}, \quad \lambda > \omega,\, n \in \mathbb{N}. \tag{5.10}$$

In the case of analytic semigroups the following is true.

Theorem 5.7. *If $-A$ is the infinitesimal generator of a C_0-semigroup $\{T(t)\}_{t\geq 0}$ of type (M,ω), then this semigroup is analytic if and only if the half-plane $\Re(\lambda) \leq \omega$ belongs to the resolvent set of A and*

$$\|R(\lambda;A)\| \leq \frac{M}{|\lambda-\omega|}, \quad \Re(\lambda) > \omega \tag{5.11}$$

holds. This holds if and only if $T(t)x \in D(A)$ for every $x \in \mathcal{E}$ and if $t > 0$,

$$\|AT(t)x\| \leq \frac{1}{t} C e^{tc} \|x\|. \tag{5.12}$$

If $-A$ generates an analytic semigroup, then fractional powers of A may be defined and for $\alpha \leq \beta \leq \gamma$ and with $x \in D(A^\gamma)$ the following bound on $A^\beta x$ holds.

$$\|A^\beta x\| \leq C \|A^\alpha x\|^{\frac{\gamma-\beta}{\gamma-\alpha}} \|A^\gamma x\|^{\frac{\beta-\alpha}{\gamma-\alpha}}.$$

In our investigation of the wave equation we will encounter also groups of bounded operators.

Definition 5.8. A family $\{T(t)\}_{t\in\mathbb{R}}$ of bounded linear operators on a Banach space $\mathcal{E}$ is a *strongly continuous group* of bounded operators (C_0-group) if

$$T(0) = I, \tag{5.13}$$

$$T(t+s) = T(t)T(s), \quad -\infty < s,t < \infty, \tag{5.14}$$

$$\lim_{t\to 0} T(t)x = x. \tag{5.15}$$

A C_0-group $\{T(t)\}_{t\in\mathbb{R}}$ is of type (M,ω) if $\|T(t)\| \leq Me^{|t|\omega}$ for every $t \in \mathbb{R}$. The infinitesimal generator of a group is defined as in the semigroup case except that the right limit is replaced by the (two-sided) limit in (5.6). A characterization of the generator of a strongly continuous group of contractions is given by the following Hille-Yosida type theorem.

Theorem 5.9. *A linear, possibly unbounded operator A is the infinitesimal generator of a C_0-group of contractions if and only if A is closed and densely defined in $\mathcal{E}$, the resolvent set contains all non-zero real numbers and*

$$\|R(\lambda;A)\| \leq \frac{1}{|\lambda|}, \quad \lambda \in \mathbb{R}.$$

An operator A generates a C_0-group if and only if $-A$ does.

We believe that the connection between operator semigroups and initial value problems is clear by now, but perhaps not fully appreciated. The following theorems from [66, Section 4.1] will strengthen this connection. First we must have a concept of a solution of the equation (5.3) when A is unbounded. We will say that $\{u(t)\}_{t\geq 0}$ solves this equation if $u(t)$ is in the domain of A for every strictly positive t and belongs to $C^0([0,\infty)) \cap C^1((0,\infty))$.

Theorem 5.10. *Let A be a densely defined operator on some Banach space $\mathcal{E}$ and assume that the resolvent set of A is non-empty. The two following statements are equivalent.*

(A) The operator $-A$ is the infinitesimal generator of a C_0-semigroup $\{T(t)\}_{t\geq 0}$.

(B) The initial value problem

$$\begin{aligned} \dot{u}(t) + Au(t) &= 0, \quad t > 0, \\ u(0) &= u_0, \end{aligned} \tag{5.16}$$

has a unique solution $u \in C^1([0,T];\mathcal{E})$ given by $u(t) = T(t)u_0$ for every $u_0 \in D(A)$.

It is worth noting that if $u_0 \notin D(A)$ then the function $u(t) = T(t)u_0$ is still well defined and may be regarded as a weaker type of solution of (5.16). This is usually referred to as a *mild* solution of this equation. We also note that the inhomogeneous equation $\dot{u}(t) + Au(t) = f(t)$ has a unique solution given by

$$u(t) = T(t)u_0 + \int_0^t T(t-s)f(s)\,\mathrm{d}s$$

under some appropriate restrictions on f. We will only encounter Wiener-noise in the right hand side and postpone a somewhat more detailed discussion about inhomogeneous problems until the introduction of stochastic equations and their solutions. Before we introduce the deterministic equations whose stochastic versions are the main focus of this thesis we return to case when strong solutions are guaranteed for all initial values.

Theorem 5.11. *If $-A$ is the generator of an analytic semigroup $\{T(t)\}_{t\geq 0}$, then* (5.16) *has a unique solution $u(t) = T(t)u_0 \in C^\infty((0,T];\mathcal{E})$ for every $u_0 \in \mathcal{E}$. Unless $u_0 \in D(A)$, $u(t)$ is not differentiable at $t = 0$.*

We now introduce the deterministic versions of the stochastic equations we aim to study and show how to formulate them in the semigroup framework. At the same time we define the domains of fractional powers of $-\Delta$ with various boundary values.

5.1 The heat equation

We start by studying the heat equation

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta u &= 0, \quad \text{in } \mathcal{D} \times (0,T], \\ u &= 0, \quad \text{on } \partial\mathcal{D} \times [0,T], \\ u &= u_0 \quad \text{in } \mathcal{D} \times \{0\}, \end{aligned} \tag{5.17}$$

where Δ is the d-dimensional Laplacian and $\mathcal{D} \in \mathbb{R}^d$ is an open bounded domain with smooth or convex polygonal boundary $\partial\mathcal{D}$. We want to put this in the semi group framework in order to be able to define the stochastic heat equation below. To this aim we let $\mathcal{E}$ be the Hilbert space $\mathcal{H} = L^2(\mathcal{D})$ with inner product $(f,g) = \int_{\mathcal{D}} fg \,\mathrm{d}x$ and norm $|f| = (f,f)^{1/2}$. It is a well known fact that $C_0^\infty(\mathcal{D})$ is dense in $L^2(\mathcal{D})$. Further, if we define Λ by

$$\begin{aligned} \Lambda &= -\Delta, \\ D(\Lambda) &= H_0^1(\mathcal{D}) \cap H^2(\mathcal{D}), \end{aligned}$$

where the domain under consideration is the largest subset of $\mathcal{H}$ such that $\Lambda f = v \in H$. Defined like this $D(\Lambda)$ is dense in $\mathcal{H}$. To see this it is enough to recognize that $C_0^\infty(\mathcal{D}) \subset D(\Lambda)$. We saw already in Section 4.1 that Λ is a closed operator. Further, it is well known that the elliptic eigenvalue problem for Λ,

$$\Lambda \phi_j = \lambda_j \phi_j,$$

has solutions $\{\lambda_j, \phi_j\}_{j=1}^\infty$ such that $\{\phi_j\}_{j=1}^\infty$ forms an ON-basis for $\mathcal{H}$ and the sequence $0 < \lambda_1 < \lambda_2 \leq \lambda_3 \leq \dots$ diverges to infinity and such that the set of eigenvalues are the whole spectrum of Λ. It is thus clear that the eigenvalues of $\lambda I + \Lambda$ are the set $\{\lambda + \lambda_j\}_{j=1}^\infty$. They are obviously positive for $\lambda \geq 0$. Hence the largest eigenvalue of $R(\lambda, \Lambda)$ is $\frac{1}{\lambda + \lambda_1} < \frac{1}{\lambda}$. It follows that $-\Lambda$ is the generator of a C_0-semigroup of contractions by the Hille-Yosida theorem. Even more: it easily seen that for $\mathrm{Re}(\lambda) \geq -\lambda_0$ the operator $\lambda I + \Lambda$ is one-to-one and onto; hence the positive real half plane is in the resolvent set of A and (5.11) holds with $\omega = 0$. Hence is $-\Lambda$ the infinitesimal generator of an analytic semigroup $\{E(t)\}_{t \geq 0}$ with the spectral representation

$$E(t)x = \sum_{j=1}^\infty \mathrm{e}^{-t\lambda_j} (x, \phi_j)\phi_j.$$

We are thus guaranteed a solution of the homogeneous heat equation for any initial value $u_0 \in \mathcal{H}$. Since Λ is the generator of an analytic semigroup we know that fractional powers of Λ are defined. Using the positive eigenvalues of of the operator we may write $\Lambda^{\alpha/2}$ for $\alpha \geq 0$ in terms of the eigen expansion

$$\Lambda^{\alpha/2} v = \sum_{j=1}^\infty \lambda_j^{\alpha/2} (v, \phi_j)\phi_j$$

whenever this series converges in $L^2(\mathcal{D})$. The set of all functions $v \in L^2(\mathcal{D})$ such that this holds will be denoted by $\dot{H}^\alpha$ which becomes a Hilbert space with the inner product

$$\langle u, v \rangle_{\dot{H}^\alpha} = \sum_{j=1}^{\infty} \lambda_j^\alpha (u, \phi_j)(v, \phi_j)$$

inducing the norm

$$\|v\|_{\dot{H}^\alpha}^2 = \sum_{j=1}^{\infty} \lambda_j^\alpha (v, \phi_j)^2.$$

For negative α the spaces are defined as the completion of $L^2(\mathcal{D})$ with respect to this norm. It is thus clear that if $\alpha < \beta$ then $\dot{H}^\beta \subset \dot{H}^\alpha$. It is known that $\dot{H}^1 = H_0^1$, $\dot{H}^2 = H^2 \cap H_0^1$ and that for positive α the space $\dot{H}^{-\alpha}$ may be identified with the dual of $\dot{H}^\alpha$. In particular $\dot{H}^{-1}$ equals what is usually denoted H^{-1} in the literature.

The property

$$\int_0^t \|\Lambda^{1/2} E(s)x\| \,\mathrm{d}s \leq \frac{1}{2}\|x\|^2$$

is, together with (5.12), fundamental in our analysis of numerical approximations of the heat equation.

5.2 The linearized Cahn-Hilliard equation

A second parabolic equation of interest is the Cahn-Hilliard equation. We study the linearized stochastic version in the attached papers. Being a parabolic equation it has strong similarities with the heat equation. The deterministic Cauchy-problem reads

$$\begin{aligned} \frac{\partial u}{\partial t} + \Delta^2 u = 0, &\quad \text{in } \mathcal{D} \times (0,T], \\ \frac{\partial u}{\partial n} = \frac{\partial}{\partial n}(\Delta u) = 0 &\quad \text{on } \partial\mathcal{D} \times [0,T], \\ u(0) = u_0 &\quad \text{in } \mathcal{D} \times \{0\}. \end{aligned} \tag{5.18}$$

If we take $\mathcal{H}$ to be functions in $L^2(\mathcal{D})$ with zero mean, i.e.,

$$\mathcal{H} = \{v \in L^2(\mathcal{D}) : (v, 1) = 0\}$$

and Λ as in the previous section, now with domain $D(\Lambda) = \{u \in H^2(\mathcal{D}) \cap \mathcal{H} : \frac{\partial u}{\partial n} = 0\}$ the operator $-A = -\Lambda^2$ is the generator of an analytic semigroup for similar reasons as in the preceding section. Thus we have put equation (5.18) in the the form of (5.16) with $A = \Lambda^2$ and we are guaranteed a unique solution of the equation in terms of an analytic semigroup $\{E(t)\}_{t\geq 0}$ with $E(t)$ given by

$$E(t)x = \sum_{j=1}^{\infty} e^{-t\lambda_j^2} (x, \phi_j)\phi_j$$

with $\{(\lambda_j, \phi_j)\}_{j=1}^{\infty}$ now being the eigen-pairs of Λ with domain as just described.

Further, as in the case of the heat equation, Λ is positive semi-definite and its fractional powers may be defined as in the previous section and we will also be able to define families of fractional spaces with norms and inner products in a similar way as there. We will denote these spaces by $\tilde{H}^{\alpha}$. It can be shown that $\tilde{H}^1 = H^1 \cap \mathcal{H}$.

5.3 The wave equation

A third equation considered below is the stochastic wave equation. Contrary to previous equations the generated C_0-semigroup fails to be analytic. On the other hand it may be extended to a group and, as a matter of fact, a unitary group. To be more precise we are interested in the equation

$$\begin{aligned}
&\frac{\partial^2 u}{\partial t^2} - \Delta u = 0, \quad \text{in } \mathcal{D} \times (0, T], \\
&u = u_0,\ \tfrac{\partial u}{\partial t} = v_0, \quad \text{on } \mathcal{D} \times \{0\}, \\
&u = 0, \quad \text{on } \partial\mathcal{D} \times [0, T].
\end{aligned} \tag{5.19}$$

To achieve a semigroup formulation we need to re-write this equation as a system. To this aim we write $U_1 = u$, $U_2 = \frac{\partial u}{\partial t}$ and $U = [U_1, U_2]^T$ with initial value $U(0) = [u_0, v_0]^T =: U_0$. We may view Λ as a bounded linear operator from $\dot{H}^1$ to $\dot{H}^{-1}$. This is since for $v \in \dot{H}^1$,

$$\|\Lambda v\|_{\tilde{H}^{-1}} = \|\Lambda^{-1/2}\Lambda v\| = \|\Lambda^{1/2} v\| = \|v\|_{\tilde{H}^1} < \infty.$$

So, redefining Λ in this sense (5.19) takes the form

$$\dot{U} = \begin{bmatrix} \dot{U}_1 \\ \dot{U}_2 \end{bmatrix} = \begin{bmatrix} 0 & I \\ -\Lambda & 0 \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}, \quad U(0) \quad = U_0,$$

which is of the form (5.3) with

$$A = \begin{bmatrix} 0 & -I \\ \Lambda & 0 \end{bmatrix}, \quad D(A) = \dot{H}^1 \times \dot{H}^0. \tag{5.20}$$

We shall now see that $-A$ is an infinitesimal generator of a C_0-group on $\mathcal{H} = \dot{H}^0 \times \dot{H}^{-1}$. First, its domain is clearly dense in $\mathcal{H}$ and for closedness we assume that $\{x^n\}_{n=1}^{\infty} \subset D(A)$ such that $[x_1^n, x_2^n]^T = x^n \to x = [x_1, x_2]^T$ in $\mathcal{H}$ and $Ax^n \to f$ in $\mathcal{H}$. We want to show that $Ax = f = [f_1, f_2]$. But $Ax^n = -[x_2^n, \Lambda x_1^n]^T$ hence $-x_2^n \to f_1$ in $\dot{H}^0$ and $-\Lambda x_2^n \to f_2$ in $\dot{H}^{-1}$. It follows immediately that $x_1 = -f_2$. For the second component we note that $\|\Lambda x_1^n\|_{\dot{H}^{-1}} = \|x_1^n\|_{\dot{H}^1}$. Thus the fact that $\{\Delta x_n\}_{n=1}^{\infty}$ converges in $\dot{H}^{-1}$ implies that $\{x_1^n\}_{n=1}^{\infty}$ converges in $\dot{H}^1$. The limiting function must thus equal x_1 and x_1 is in the domain of Λ and $\Lambda x_1 = -f_2$. Thus $Ax = f$. To be able to invoke the group version of the Hille-Yosida theorem,

Theorem 5.9, we therefore need to show that (5.9) holds. To do this we shall start by diagonalizing A. First we extend A and $\mathcal{H}$ by complexification, still denoting the extended operator and the complex Hilbert space by A and $\mathcal{H}$ respectively. Then we consider the eigen-value problem for A on $\mathcal{H}$, that is, we want to find pairs $(\mu_j, \psi_j) \in \mathbb{C} \times \mathcal{H}$ such that

$$A\psi_j = \mu_j \psi_j.$$

Writing $\psi_j = [\psi_{1,j}, \psi_{2,j}]^T$ it follows from (5.20) that this amounts to finding solutions to the the system

$$\begin{aligned} -\mu_j \psi_{1,j} &= \psi_{2,j} \\ \Lambda \psi_{1,j} &= \mu_j \psi_{2,j}. \end{aligned} \tag{5.21}$$

Thus

$$\Lambda \psi_{1,j} = -\mu_j^2 \psi_{1,j}$$

must hold, which means that $\psi_{1,j}$ must be an eigenfunction of Δ, say $\frac{1}{\sqrt{2}}\phi_i$ and $-\mu_j^2 = \lambda_i$, the corresponding eigenvalue. If we for $j \in \mathbb{N}$ write $\psi_{1,j} = \psi_{1,-j} = \frac{1}{\sqrt{2}}\phi_j$, $\mu_j = i\sqrt{\lambda_j}$ and $\mu_{-j} = -i\sqrt{\lambda_j}$ then it follows from (5.21) that $\psi_{2,j} = -\frac{1}{\sqrt{2}} i\sqrt{\lambda_j}\phi_j$ and $\psi_{2,-j} = \frac{1}{\sqrt{2}} i\sqrt{\lambda_j}\phi_j$. Then $\{(\mu_j, \psi_j)\}_{j\in\mathbb{Z}\setminus 0}$ forms an ON-system in $\mathcal{H}$ and that it is complete can be seen by for $j \in \mathbb{N}$ writing $\tilde{\psi}_j = \psi_j + \psi_{-j} = [\sqrt{2}\phi_j, 0]^T$ and $\tilde{\psi}_{-j} = \psi_j - \psi_{-j} = [0, i\sqrt{2\lambda_j}\phi_j]^T$. It is clear that $\operatorname{span}\{\psi_j, \psi_{-j}\} = \operatorname{span}\{\tilde{\psi}_j, \tilde{\psi}_{-j}\}$. Hence the closure of the span of $\{\psi_j\}_{j=-\infty}^{\infty}$ equals the closure of the span of $\{\tilde{\psi}_j\}_{j=-\infty}^{\infty}$ which equals $\mathcal{H}$. Thus $\{\psi_j\}_{j=-\infty}^{\infty}$ diagonalizes the operator A. It follows as in the case of the heat equation that for any $f \in \mathcal{H}$ and $\lambda \in \mathbb{R}$

$$\|R(\lambda; A)f\|^2 = \sum_{0\neq j\in\mathbb{Z}} \frac{1}{\lambda^2 + \mu_j^2} |\langle f, \psi_j\rangle|^2 \le \frac{1}{\lambda^2 + \mu_1^2} \sum_{0\neq j\in\mathbb{Z}} |\langle f, \psi_j\rangle|^2.$$

Thus

$$\|R(\lambda; A)\| \le \frac{1}{|\lambda|}, \quad \lambda \in \mathbb{R}.$$

Therefore, A is an infinitesimal generator of a C_0-group. Also in this case we may define it in terms of the eigenpairs of Λ. For this aim we write

$$S(t)x = \sin(t\Lambda^{1/2}) := \sum_{j=1}^{\infty} \sin(t\sqrt{\lambda_j})(x, \phi_j)\phi_j,$$

$$\Lambda^{1/2} S(t)x := \sum_{j=1}^{\infty} \lambda_j^{1/2} \sin(t\sqrt{\lambda_j})(x, \phi_j)\phi_j,$$

and correspondingly for $C(t) = \cos(t\Lambda^{1/2})$ and $\Lambda^{-1/2}S(t)$. The strongly continuous group $\{E(t)\}_{t\in\mathbb{R}}$ generated by $-A$ as described in (5.20) is then given by

$$E(t) = \begin{bmatrix} C(t) & \Lambda^{-1/2}S(t) \\ -\Lambda^{1/2}S(t) & C(t) \end{bmatrix}.$$

It turns out that the spaces $\mathcal{H}^\alpha = \dot{H}^\alpha \times \dot{H}^{\alpha-1}$ are good when to study convergence rates for approximation of the wave operator. This is related to the fact that $\{E(t)\}_{t\in\mathbb{R}}$ is a unitary group on any $\mathcal{H}^\alpha$ and hence $E(t)$ an isometric isomorphism from $\mathcal{H}^\alpha$ to $\mathcal{H}^\alpha$ for any $t \in \mathbb{R}$.

6 Stochastic equations

Since our aim is to study stochastic equations on Hilbert spaces we need the concept of Hilbert-valued stochastic processes. We work on a probability space $(\Omega, \mathcal{F}, P)$ and a Hilbert space $\mathcal{U}$ and the Borel σ-algebra of $\mathcal{U}$ denoted by $\mathrm{Bor}(\mathcal{U})$. A $\mathcal{U}$-valued random variable is a measurable mapping $X\colon (\Omega, \mathcal{F}) \to (\mathcal{U}, \mathrm{Bor}(\mathcal{U}))$. It will often be the case that we have a probability measure μ on the measurable space $(\mathcal{U}, \mathrm{Bor}(\mathcal{U}))$. For every element $v \in \mathcal{U}$ we define a functional $v'\colon \mathcal{U} \to \mathbb{R}$ by $v'(x) = \langle v, x\rangle_{\mathcal{U}}$. The probability measure μ is then said to be Gaussian if for every $v \in \mathcal{U}$ there exists $m_v \in \mathbb{R}$ and $\sigma_v \geq 0$ such that if $\sigma_v > 0$, then

$$\mu \circ (v')^{-1}(A) = \mu(\{u \in \mathcal{U} : v'(u) \in A\}) = \frac{1}{\sqrt{2\pi\sigma_v^2}} \int_A \mathrm{e}^{-\frac{(r-m_v)^2}{2\sigma_v^2}}\,\mathrm{d}r$$

for all $A \in \mathrm{Bor}(\mathbb{R})$ and if $\sigma_v = 0$, then $\mu \circ (v')^{-1} = \delta_{m_v}$, the Dirac measure at m_v. A $\mathcal{U}$-valued random variable X on $(\Omega, \mathcal{F}, P)$ is Gaussian if the measure $P \circ X^{-1}$ is Gaussian on $(\mathcal{U}, \mathrm{Bor}(\mathcal{U}))$. If X is such a random variable, it follows that there exist a *mean* $m \in \mathcal{U}$ such that $\mathbb{E}[X] = m$ and a positive semidefinite, symmetric operator Q of trace class such that $\mathbb{E}[\langle X - m, v\rangle\langle X - m, u\rangle] = \langle Qv, u\rangle$ for all $u, v \in \mathcal{U}$. This operator is called the *covariance operator* of X. It is also clear from above that for all $u \in \mathcal{U}$, $\langle X, u\rangle$ is a real Gaussian random variable with $\mathbb{E}[\langle X, u\rangle] = \langle m, u\rangle$ and $\mathbb{E}[\langle X - m, u\rangle^2] = \langle Qu, u\rangle = \sigma_u^2$. We say then that X has Gaussian law and write $\mathcal{L}(X) = \mathcal{N}(m, Q)$. Further, a $\mathcal{U}$-valued stochastic process $\{X(t)\}_{t\geq 0}$ is a function from $\mathbb{R}_+ \times \Omega$ to $\mathcal{U}$ such that $X(t)$ is a random variable for all $t \geq 0$. Let $\{\mathcal{F}_t\}_{t\geq 0}$ be a filtration of $\mathcal{F}$ and assume that $X(t)$ is $\mathcal{F}_t$-measurable for every $t \geq 0$. Then $\{X(t)\}_{t\geq 0}$ is *adapted to* $\{\mathcal{F}_t\}_{t\geq 0}$. If, for all $n \in \mathbb{N}$ and arbitrary positive real numbers $\{t_k\}_{k=1}^n$, the $\mathcal{U}^n$-valued random variable $(X(t_1), \dots, X(t_n))$ is Gaussian, then $\{X(t)\}_{t\geq 0}$ is said to be a Gaussian stochastic process. A $\mathcal{U}$-valued Q-Wiener process $\{W(t)\}_{t\geq 0}$ is a stochastic process such that for some bounded positive semi-definite symmetric operator $Q \in \mathcal{L}_1(U)$ it holds that

1. $W(0) = 0$,
2. $\{W(t)\}_{t\geq 0}$ has continuous trajectories, almost surely,
3. $\{W(t)\}_{t\geq 0}$ has independent increments and
4. $\mathcal{L}(W(t) - W(s)) = \mathcal{N}(0, (t-s)Q), \quad t \geq s \geq 0.$

It follows that W is a Gaussian process.

It is a well known fact from spectral theory that a compact positive semidefinite symmetric operator Q on a separable Hilbert space $\mathcal{U}$ has a representation in terms of its eigenpairs $\{\gamma_j, e_j\}_{j=1}^{\infty}$ where $\{\gamma_j\}_{j=1}^{\infty}$ is a positive decreasing sequence and $\{e_j\}_{j=1}^{\infty}$ is a complete ON-basis of $\mathcal{U}$. This means that we may write

$$Qx = \sum_{j=1}^{\infty} \gamma_j \langle x, e_j \rangle e_j, \quad x \in \mathcal{H},$$

and that we can define fractional powers of Q by $Q^{\alpha}x = \sum_{\gamma_j \neq 0} \gamma_j^{\alpha} \langle x, e_j \rangle e_j, \alpha \in \mathbb{R}, x \in \mathcal{U}$. This implies that all Q-Wiener processes W may be be written as

$$W(t) = \sum_{j=1}^{\infty} \gamma_j^{1/2} \beta_j(t) e_j, \tag{6.1}$$

where the realvalued processes $\{\beta_j(t)\}_{t \geq 0}$ are given by $\beta_j(t) = \gamma_j^{-1/2} \langle W(t), e_j \rangle$ if $\gamma_j > 0$ and $\beta_j = 0$ otherwise. It turns out that $\{\beta_j\}_{j=1}^{\infty}$ is a sequence of mutually independent standard Brownian motions. The series in (6.1) converges in $L^2(\Omega, \mathcal{H})$ since $\mathrm{Tr}(Q) < \infty$. Conversely, if $\{\beta_j\}_{j=1}^{\infty}$ are independent standard Brownian motions and $\{\gamma_j\}_{j=1}^{\infty}$ is in l^1, then (6.1) defines a $\mathcal{U}$-valued Wiener-process and hence there exists a Q-Wiener process for any trace class, symmetric, non-negative operator. In order to give a meaning to stochastic partial differential equations we need to integrate operator valued functions from $\mathcal{U}$ to $\mathcal{H}$ with respect to Wiener processes. If the integrand is deterministic the resulting integral is called the Wiener integral and it generalizes to the Itô integral when the integrand is stochastic. The construction is performed in what could be regarded as three steps. First, it is defined for so called elementary processes and, second, the class of integrands is extended to a class of predictable processes to be defined below. A concrete characterization of this extension is then given. The third step is to extend the integral to cylindrical Wiener processes, i.e., Wiener processes where the covariance operator is not of trace class. This is done through the construction of a larger Hilbert space where the process $\{W(t)\}_{t \geq 0}$ becomes a well defined Wiener process. We first give definitions and state facts about martingales.

Definition 6.1. Let $\{M(t)\}_{t \geq 0}$ be an $\mathcal{U}$-valued stochastic process on a probability space $(\Omega, \mathcal{F}, P)$. If $\{\mathcal{F}_t\}_{t \geq 0}$ is a filtration of this space and if $\{M(t)\}_{t \geq 0}$ is adapted then $\{M(t)\}_{t \geq 0}$ is an $\{\mathcal{F}_t\}_{t \geq 0}$-*martingale* if $\mathbb{E}[\|M(t)\|] < \infty$ for $t \geq 0$ and

$$\mathbb{E}[M(t)|\mathcal{F}_s] = M(s), \quad 0 \leq s \leq t < \infty.$$

We write $\{M(t)\}_{0 \leq t \leq T} \in \mathcal{M}_T(\mathcal{U})$ if $\{M(t)\}_{0 \leq t \leq T}$ is an $\{\mathcal{F}_t\}_{t \geq 0}$-martingale, the map $t\colon t \mapsto M(t)$ is continuous P-a.s., and the norm

$$\|M\|_{\mathcal{M}_T(\mathcal{U})} = \mathbb{E}[\|M(T)\|^2]$$

is finite. The space $\mathcal{M}_T$ is a Banach space.

A filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is *normal* if all zero-sets of $\mathcal{F}$ are contained in $\mathcal{F}_0$ and if $\mathcal{F}_t = \bigcap_{t<s} \mathcal{F}_s$ for all $t \geq 0$. If a Q-Wiener process is adapted to a normal filtration $\{\mathcal{F}_t\}_{t\geq 0}$, then it is an $\{\mathcal{F}_t\}_{t\geq 0}$-martingale.

We are ready to define stochastic integrals for *elementary processes*. To this aim, let $0 = t_0 < t_1 \dots t_{n-1} < t_n = T$ define a partition of $[0,T]$. An elementary process $F\colon \Omega \times [0,T] \to \mathcal{U}$ is a process $F(t) = \sum_{j=1}^N F_j \chi_{(t_j,t_{j+1}]}(t), F_j \in \mathcal{B}(\mathcal{U},\mathcal{H})$ such that F_j takes only a finite number of values as a function of ω. We will refer to the class of elementary functions of this kind as $\mathcal{E}_T$. For any function $F \in \mathcal{E}_T$, the stochastic integral of F is the process

$$\int_0^t F(s)\,\mathrm{d}W(s) = \sum_{j=1}^N F_j(t)(W(t\wedge t_j) - W(t \wedge t_{j-1})), \quad t \in [0,T].$$

This integral defines an element in $\mathcal{M}_T^2$ and it is possible to prove that for $F \in \mathcal{E}$

$$\begin{aligned} \Big\| \int_0^{\cdot} F\,\mathrm{d}W \Big\|^2_{\mathcal{M}_T^2} &= \mathbb{E}\Big[\Big\| \int_0^T F(t)\,\mathrm{d}W(t) \Big\|^2 \Big] \\ &= \mathbb{E}\Big[\int_0^T \|F(t)Q^{1/2}\|^2_{\mathcal{L}_2(\mathcal{U},\mathcal{H})}\,\mathrm{d}t \Big] =: \|F\|_T^2. \end{aligned} \tag{6.2}$$

We will refer to the equality (6.2) as the Itô isometry. The functional $\|\cdot\|_T$ only defines a semi-norm on $\mathcal{E}_T$. Clearly, if the range of $Q^{1/2}$ is contained in the null-space of F the operator $FQ^{1/2} \equiv 0$. To get around this we shall re-define $\mathcal{E}_T$. To do this we first introduce the Cameron-Martin space of $\mathcal{U}$ denoted by $\mathcal{U}_0 := Q^{1/2}(\mathcal{U})$ and write $\mathcal{E}_T^0 = \{F \in \mathcal{E}_T : F = 0 \,\text{on}\, \mathcal{U}_0\}$ and let our new space of elementary functions to be the quotient space $\mathcal{E}_T := \mathcal{E}_T \setminus \mathcal{E}_T^0$. At this point, we equip the space $\mathcal{U}_0$ with the inner-product $\langle\cdot,\cdot\rangle_{\mathcal{U}_0} = \langle Q^{-1/2}\cdot, Q^{-1/2}\cdot\rangle_{\mathcal{U}}$ where $Q^{-1/2}$ is the pseudo inverse of $Q^{1/2}$. This makes $\mathcal{U}_0$ a Hilbert space and the space of Hilbert-Schmidt operators from $\mathcal{U}_0$ to $\mathcal{H}$ is thus a well defined Hilbert space which we denote by $\mathcal{L}_2^0 := \mathcal{L}_2(\mathcal{U}_0, \mathcal{H})$.

We now define the operator $\mathrm{Int}\colon (\mathcal{E}_T, \|\cdot\|_T) \to (\mathcal{M}_T^2, \|\cdot\|_{\mathcal{M}_T^2})$ through

$$\mathrm{Int}(F)(t) = \int_0^t F(s)\,\mathrm{d}W(s), \quad t \in [0,T].$$

We may make an abstract completion of $\mathcal{E}_T$ denoted by $\bar{\mathcal{E}}_T$ such that any Cauchy sequence $\{F_n\}_{n=1}^\infty$ in $\mathcal{E}_T$ with respect to the norm $\|\cdot\|_T$ converges to an element $F \in \bar{\mathcal{E}}_T$.[1] Since the integral by (6.2) is an isometric mapping from $(\mathcal{E}_T, \|\cdot\|_T) \to (\mathcal{M}_T^2, \|\cdot\|_{\mathcal{M}_T^2})$ and since $\mathcal{M}_T^2$ is complete we define the stochastic integral of the limiting element to be

$$\mathrm{Int}(F) := \lim_{n\to\infty} \mathrm{Int}(F_n).$$

[1] Rather, the space $\bar{\mathcal{E}}_T$ is the space of Cauchy-sequences of elements in $\mathcal{E}$ and we identify two Cauchy-sequences $\{x_n\}$ and $\{y_n\}$ with each other if $\lim_{n\to\infty} \|x_n - y_n\| = 0$. Every element $f \in \mathcal{E}_T$ may thus be identified with the constant sequence $\{f_n\},\ f_n = f$. This makes $\bar{\mathcal{E}}_T$ complete and $\mathcal{E}_T$ a subset of $\bar{\mathcal{E}}_T$.

It turns out that the abstract completion of $\mathcal{E}_T$ has an explicit characterization. To state this we need a new measurable space where the underlying space is $\Omega_T := \Omega \times [0,T]$. With $\{\mathcal{F}_t\}_{0\le t\le T}$ being the filtration in the definition of $\mathcal{M}_T^2$ above we define the σ-algebra

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

A measurable mapping from $(\Omega_T, \mathcal{P}_T)$ to $(H, \mathrm{Bor}(H))$ where H is a separable Hilbert space is called *H-predictable*.

We are finally ready to state the explicit characterization of $\bar{\mathcal{E}}_T$. It is given by the space

$$\begin{aligned}\mathcal{N}_W^2 &= \mathcal{N}_W^2(0,T;\mathcal{H}) \\ &= \{\Phi\colon [0,T] \times \Omega \to \mathcal{L}_2^0 : \Phi \text{ is } \mathcal{L}_2^0\text{-predictable and } \|\Phi\|_T < \infty\} \\ &= \mathcal{L}_2([0,T] \times \Omega, \mathcal{P}_T, m \times P; \mathcal{L}_2^0)\end{aligned}$$

where m is the Lebesgue measure on $[0,T]$.

Until now we have assumed that $\mathrm{Tr}(Q) < \infty$, but the condition

$$\|\Phi\|_T = \mathbb{E}\Big[\int_0^T \|\Phi(t)Q^{1/2}\|^2_{\mathcal{L}_2(\mathcal{U},\mathcal{H})}\,\mathrm{d}t\Big] < \infty$$

suggests that if Φ is compact, then Q may be less well behaved than of trace class. This turns out to be the case but the construction is not immediate. It is, as already mentioned, not possible to construct a Gaussian random variable in $\mathcal{L}_2(\Omega;\mathcal{H})$ that doesn't have trace class covariance operators. The remedy is to start with a selfadjoint, positive semidefinite operator $Q \in \mathcal{B}(\mathcal{U})$ and $Q = Q^*$, possibly not in $\mathcal{L}_1(\mathcal{U})$, and a corresponding Cameron-Martin space $\mathcal{U}_0$. Then we choose another separable Hilbert space $\tilde{\mathcal{U}}$ with an embedding $J\colon \mathcal{U}_0 \to \tilde{\mathcal{U}}$ such that $J \in \mathcal{L}_2(\mathcal{U}_0, \tilde{\mathcal{U}})$ and hence $\tilde{Q} = JJ^* \in \mathcal{L}_1(\tilde{\mathcal{U}})$. We may then define a $\tilde{Q}$-Wiener process $\{\tilde{W}(t)\}_{t\ge 0}$ on $\tilde{\mathcal{U}}$. If $\Phi \in \mathcal{L}_2(\mathcal{U},\mathcal{H})$, then the operator $\Phi J^{-1} \in \mathcal{L}_2(\tilde{\mathcal{U}}_0, \mathcal{H})$ and the integral of ΦJ^{-1} with respect to $\{\tilde{W}(t)\}_{t\ge 0}$ is therefore well defined. We thus define the integral with respect to the *cylindrical* Q-Wiener process $\{W(t)\}_{t\ge 0}$ to be

$$\int_0^t \Phi(s)\,\mathrm{d}W(s) := \int_0^t \Phi(s)J^{-1}\,\mathrm{d}\tilde{W}(s), t \in [0,T]. \tag{6.3}$$

This definition is independent of $\tilde{U}$ and J.

With the stochastic integral at our disposal, we are in the position to define various solution concepts of stochastic evolution equations of the form

$$\begin{aligned}\mathrm{d}X(t) + AX(t)\mathrm{d}t &= B\mathrm{d}W(t), \quad t \ge 0, \\ X(0) &= X_0\end{aligned} \tag{6.4}$$

where $A\colon D(A) \to \mathcal{H}$ is linear and $B \in \mathcal{B}(\mathcal{U},\mathcal{H})$. Further, we will assume that $-A$ is the generator of a strongly continuous semigroup on $\mathcal{H}$ and that $X_0 \in \mathcal{H}$ is $\mathcal{F}_0$-measurable.

Definition 6.2. An $\mathcal{H}$-valued predictable process $\{X(t)\}_{t\in I}$ is called a *strong* solution of (6.4) if $X \in D(A)$ P_T-a.s., $AX(t)$ is Bochner integrable [2] P-a.s. and

$$X(t) = X_0 - \int_0^t AX(s)\,\mathrm{d}s + BW(t), \quad P\text{-a.s}, t \in [0,T].$$

We note that in order for BW to be well defined we must have $\mathrm{Tr}(BQB^*) < \infty$.

Definition 6.3. A *weak* solution to (6.4) is an $\mathcal{H}$-predictable stochastic process $\{X(t)\}_{0\le t\le T}$ that is Bochner integrable P-a.s. and

$$\langle X(t), v\rangle = \langle X_0, v\rangle - \int_0^T \langle X(s), A^* v\rangle\,\mathrm{d}s + \int_0^t \langle B\,\mathrm{d}W(s), v\rangle,$$
$$P - a.s.,\ v \in D(A^*),\ t \in [0,T].$$

In contrast to $BW(t)$ the process $\int_0^t \langle B\,\mathrm{d}W(s), v\rangle$ is a well-defined process for every $v \in \mathcal{H}$.

Theorem 6.4. *If $-A$ is the infinitesimal generator of a strongly continuous semigroup $\{E(t)\}_{t\ge 0}$, $B \in \mathcal{L}(\mathcal{U},\mathcal{H})$ and if*

$$\int_0^T \|E(s)BQ^{1/2}\|^2_{\mathcal{L}_2(\mathcal{U},\mathcal{H})}\,\mathrm{d}s < \infty, \tag{6.5}$$

then the unique weak solution of (6.4) *is given by the stochastic variation of constants formula*

$$X(t) = E(t)X_0 + \int_0^t E(t-s)B\,\mathrm{d}W(s), \quad t \in [0,T]. \tag{6.6}$$

The second term in the right hand side above is a stochastic convolution which often is denoted by $W_A(t)$. It is, as we have seen, well defined if (6.5) holds. With this theorem we have proved conditions on the noise that will give us existence and uniqueness for the three equations we want to study e.g., the stochastic heat equation (with $\mathcal{U} = \mathcal{H} = L^2(\mathcal{D})$, $A = \Lambda$, $D(\Lambda) = \dot{H}^2$ and $B = I$), the Cahn-Hilliard-Cook equation (with $\mathcal{U} = \mathcal{H} = \{u \in L^2(\mathcal{D}) : (u,1) = 0\}$, $A = \Lambda^2$, $D(A) = \tilde{H}^2$ and $B = I$) and the stochastic wave equation (with $\mathcal{U} = L^2(\mathcal{D})$, $\mathcal{H} = \mathcal{H}^0$, A given by (5.20) with $D(A) = \mathcal{H}^1$ and $B\colon \mathcal{U} \to \mathcal{H}$ given by $B = [0, I]^T$). This is explained in more detail in Sections 5.1–5.3.

We are concerned with numerical approximations of the solutions to various equations of the form (6.4) and their *apriori* error estimates. There are more

[2] The Bochner integral is the extension of the Lebesgue integral to vector valued functions. It allows us to integrate operator families $\{\Phi(t)\}_{0\le t\le T}$ if (and only if) $\int_0^T \|\Phi(t)\|\,\mathrm{d}m(t) < \infty$. The integral is constructed by approximation by simple functions $\{\Phi_n = \sum_{j=1}^n \phi_j \chi_{F_i}(t)\}_{n=1}^\infty$ such that $\int_0^T \|\Phi(t) - \Phi_n(t)\|\,\mathrm{d}m(t) \to 0$ and the integral is then defined as $\int \Phi\,\mathrm{d}m = \lim \int \Phi_n \mathrm{d}m$ where $\int \Phi_n \mathrm{d}m = \sum_{j=1}^n \phi_j m(F_j)$.

notions of convergence in the stochastic case than in the deterministic, such as almost sure convergence, convergence in probability, strong convergence and convergence in measure (weak convergence). We are concerned with strong and weak convergence below. The former is, in general, defined as convergence in $L_p(\Omega, \mathcal{H})$; that is, a sequence $\{X_\epsilon\}$ is said to converge to X in $L_p(\Omega, \mathcal{H})$ if

$$\lim_{\epsilon \longrightarrow 0} \mathbb{E}[\|X_\epsilon - X\|_{\mathcal{H}}^p] = 0, \tag{6.7}$$

but we only study the case $p = 2$ which is referred to as *mean square convergence*. Weak convergence means that the probability law of X_ϵ converges to the law of X but can equivalently be formulated as convergence of $\mathbb{E}[G(X_\epsilon)]$ to the value of $\mathbb{E}[G(X)]$ for all $G \in B_L(\mathcal{H}, \mathbb{R})$ where $B_L(\mathcal{H}, \mathbb{R}) = B_L$ is the space of bounded Lipschitz functions. We do not study convergence with respect to all functions in B_L but with the class $C_\mathrm{b}^2(\mathcal{H}, \mathbb{R})$. The space C_b^2 is not contained in B_L (as functions in C_b^2 are not necessarily bounded), neither does the converse hold. There is, to our knowledge, nothing that says that the class C_b^2 is optimal for getting convergence rates. Rather, it is plausible that for smoother processes rates should be achievable also for non-smooth test-functions, G. The reason for using C_b^2 is that for this class the generalization of Itô's formula to infinite dimensional Hilbert spaces is available and easy to use in the analysis. Furthermore, it is also a condition that in many cases is possible to verify.

Theorem 6.5 (Itô's formula)**.** *Let the process $\{F(t)\}_{t\in[0,T]}$ be an element in $\mathcal{N}_W^2$ and let $\{f(t)\}_{t\in[0,T]}$ be a predictable and Bochner-integrable $\mathcal{H}$-valued stochastic process. Suppose that $\Psi : \mathcal{H} \times [0,T] \rightarrow \mathbb{R}, (x,t) \mapsto \Psi(x,t)$ is differentiable in t and twice differentiable in x with all derivatives continuous on $\mathcal{H} \times [0,T]$.Then, P-a.s., the process*

$$X(t) = X(0) + \int_0^t f(s)\,\mathrm{d}s + \int_0^t F(s)\,\mathrm{d}W(s),$$

where $X(0)$ is $\mathcal{F}_0$-measurable, is well-defined and the process $\Psi(X(t),t)$ can be written, $P-a.s.$, as

$$\begin{aligned}\Psi(X(t),t) = \Psi(X(0),0) &+ \int_0^t \langle \Psi_x(X(s),s), F(s)\,\mathrm{d}W(s)\rangle \\ &+ \int_0^t \Big(\Psi_t(X(s),s) + \langle \Psi_x(X(s),s), f(s)\rangle \\ &\qquad + \tfrac{1}{2}\operatorname{Tr}(\Psi_{xx}(X(s),s)(F(s)Q^{1/2})(F(s)Q^{1/2}))\Big)\,\mathrm{d}s.\end{aligned} \tag{6.8}$$

For a proof, see [11].

The equality (6.8) is called Itô's formula and leads to the possibility of formulating an infinite dimensional Kolmogorov equation where G is used as boundary value at the final time T. Before we proceed, it is worth noting that studying this kind of weak convergence, with functions in C_b^2, is important even though it doesn't imply convergence in law: in many cases we are only interested in a certain function of the solution of (6.4) at some time T. A typical example from the finite dimensional case is when the process $\{X(t)\}_{0\leq t\leq T}$ is the price of a stock and the value of $G(X(T))$ is the amount of money that some derivative pays out at time T. Pricing this derivative at an earlier time t leads to the need to calculate $\mathbb{E}[G(X(T))|\mathcal{F}_t]$. For some derivatives analytical solutions may be found, but in most cases one needs to rely on numerical methods. It should come as no surprise that the need for numerical treatment is even larger in the infinite dimensional case.

Itô's formula is a central theoretical tool in the study of stochastic processes and it may be used to prove the next theorem that relates the solution of an SPDE to a deterministic Kolmogorov's equation. To motivate the form it is stated in we note that if $X(t)$ solves (6.4) and if $0 \leq t \leq T$ then

$$Y(t) = E(T-t)X(t) = E(T)X_0 + \int_0^t E(T-s)B\,\mathrm{d}W(s) \tag{6.9}$$

solves the drift free equation

$$\mathrm{d}Y(t) = E(T-t)B\mathrm{d}W(t), 0 \leq t \leq T, \quad Y(0) = E(T)X_0.$$

We also define the function $Z(t,\tau,x)$ being the solution of the equation

$$\mathrm{d}Z(t) = E(T-t)B\mathrm{d}W(t), \tau \leq t \leq T, \quad Z(\tau) = x. \tag{6.10}$$

It holds that $Z(t, 0, E(T)X_0) = Y(t)$ and in particular, since $Y(T) = X(T)$, that $Z(T, 0, E(T)X_0) = X(T)$. The following theorem can be proved.

Theorem 6.6. *If $g\colon \mathcal{H} \to \mathbb{R}$ is a bounded continuous functional with two bounded Fréchet derivatives and if $Z(T,t,x)$ solves* (6.10)*, then the function*

$$u(x,t) = \mathbb{E}[g(Z(T,t,x))]$$

solves the Backward Kolmogorov equation

$$\begin{aligned} &\frac{\partial u}{\partial t}(x,t) = -\operatorname{Tr}(u_{xx}(x,t)E(T-t)BQ(E(T-t)B)^*), \quad x \in D(A), \\ &u(x,T) = g(x) \end{aligned} \tag{6.11}$$

for $x \in D(A)$.

For a discussion of Kolmogorov equations we refer to [16]. Note that Theorem 6.6 does not say anything about uniqueness.

7 Finite element methods

The finite element method is a certain kind of Galerkin method, which means that one looks for approximate solutions on some finite dimensional subspace $\tilde{S}$ of the Hilbert space V where $V \times V$ is the domain of the bilinear form K

associated with the generator A in (5.3), i.e., a solution to the problem to find $\tilde{u}(t) \in \tilde{S}$ such that

$$\langle \dot{\tilde{u}}, v\rangle + K(\tilde{u}, v) = \langle f, v\rangle, \quad t > 0; \quad \langle u(0), v\rangle = \langle u_0, v\rangle, \quad \forall v \in \tilde{S}. \tag{7.1}$$

Given a basis $\{\psi_j\}_{j=1}^N \subset \tilde{S}$ where N is the dimension $\tilde{S}$ one makes the Ansatz $\tilde{u} = \sum_{j=1}^N \alpha_j(t)\psi_j$, where $\alpha = \{\alpha_j\}_{j=1}^N$ is an R^N-valued function of t. Due to the linearity of the inner product only the basis functions needs to be used as test-functions and (7.1) takes the form

$$\mathbb{M}\dot{\alpha}(t) + \mathbb{K}\alpha(t) = \mathbb{F}(t), t > 0; \quad \mathbb{M}\alpha(0) = \mathbb{G}$$

with $\mathbb{M}_{ij} = \langle \psi_j, \psi_i\rangle$, $\mathbb{K}_{ij} = K(\psi_j, \psi_i)$, $\mathbb{F}_i(t) = \langle f, \psi_i\rangle$ and $\mathbb{G}_i = \langle u_0, \psi_i\rangle$. This is a system of ordinary differential equations that can be solved with methods of Section 8. The question is of course how the subspaces and the basis should be chosen in order to make the solution as fast and as accurate as possible. Finite elements is one answer to this question.

In order to describe the finite element method we first start with a bounded polygonal domain $\mathcal{D} \in \mathbb{R}^n$ and a family of triangulations $\{\mathcal{T}_h\}_{0<h<1}$ thereof. A triangulation is a subdivision of $\mathcal{D}$ into simplexes with non-overlapping interiors and such that no vertex of any simplex lies in the interior of an edge of any other simplex. he largest diameter of any simplex of a triangulation is denoted by h.We restrict our attention to finite element spaces where the space $\tilde{S} = S_h^r(\mathcal{D}) = S_h^r$, corresponding to a certain triangulation $\mathcal{T}_h$, consists of functions that are polynomials of degree $r-1$ on every triangle and globally continuous. The number $r \geq 2$ is the approximation order of the finite element space. When $r = 2$, the space consists of piecewise linear functions and the basis $\{\psi_j\}_{j=1}^{N_h}$ is usually taken to be such that ψ_j equals 1 at triangle node j and vanishes on all other nodes. When $r > 2$ the description is not as easy and we refer to standard literature on finite element methods such as [10] for the details.

To find an approximate solution of the heat equation we define the discrete version of Λ denoted by Λ_h on $S_{h,0}^r := S_h^r \cap H_0^1$ through

$$\langle \Lambda_h x, y\rangle = \langle \nabla x, \nabla y\rangle, \quad x, y \in S_{h,0}^r. \tag{7.2}$$

The operator Λ_h is positive definite and as being defined on a finite dimensional space it is bounded. Thus $-\Lambda_h$ is the generator of an analytic semigroup

$$E_h(t) = \mathrm{e}^{-t\Lambda_h} = \sum_{j=1}^{N_h} \mathrm{e}^{-\lambda_{h,j}t}\langle \cdot, \phi_{h,j}\rangle\phi_{h,j}$$

where $\{\lambda_{h,j}, \phi_{h,j}\}_{j=1}^{N_h}$ are the eigenpairs of Λ_h with N_h the dimension of $S_{h,0}^r$. Further, we define P_h as the orthogonal projection onto $S_{h,0}^r$ (in the sequel P_h will be the projection onto the finite element space where we are looking for a solution, we will not notice on this varying definition). Thus, the equation (7.1) may be rewritten as to find $u_h(t) \in S^r_{h,0}$ such that

$$\dot{u}_h + \Lambda_h u_h = P_h f, t > 0; \quad u_h(0) = P_h u_0$$

with unique weak solution

$$u_h(t) = E_h(t) P_h u_0 + \int_0^t E_h(t-s) P_h f(s) \, \mathrm{d}s. \tag{7.3}$$

The error estimates of the stochastic heat equation relies on the estimate for the homogeneous deterministic heat equation. Before we state this we shall state an assumption of the convergence order of the finite element space. To this aim we consider the elliptic problem $\Lambda u = f$ with solution $u = \Lambda^{-1} f$ and its finite element approximation, $\Lambda_h u_h = P_h f$ with solution $u_h = \Lambda_h^{-1} P_h f$. The space S^r_h has *elliptic convergence order* r if, whenever $f \in \dot{H}^{r-2}$, it holds that

$$\|\Lambda_h^{-1} P_h f - \Lambda^{-1} f\| \leq C h^s \|f\|_{\dot{H}^{r-2}}.$$

This holds for example with $r = 2$ on a convex polygonal domain. For $r > 2$ the situation is more involved. In the sequel, we will always assume that S^r_h has elliptic convergence order r.

Theorem 7.1. *Let $S^r_h(\mathcal{D})$ have elliptic convergence order r, and let $\Lambda = -\Delta$ with $D(\Lambda) = H^2(\mathcal{D}) \cap H^1_0(\mathcal{D})$ and let Λ_h be as in* (7.2). *If $\{E(t)\}_{t\geq 0}$ is the semigroup generated by $-\Lambda$ on $L^2(\mathcal{D})$ and $E_h(t)$ the semigroup generated by $-\Lambda_h$ on $S^r_{h,0}$ and if $v \in \dot{H}^\gamma$ for $0 \leq \gamma \leq \beta \leq r$, then*

$$\|E(t)v - E_h(t) P_h v\| \leq C h^\beta t^{-\frac{\beta-\gamma}{2}} \|v\|_{\dot{H}^\gamma} \tag{7.4}$$

or, equivalently,

$$\|E(t) - E_h(t) P_h\|_{\mathcal{B}(\dot{H}^\gamma, \dot{H}^0(\mathcal{D}))} \leq C h^\beta t^{-\frac{\beta-\gamma}{2}}.$$

This is [75, Theorem 3.5].

To formulate the finite element formulation of the linearized Cahn-Hilliard equation we denote by $S^r_{h,*} = \{v \in S^r_h : \langle v, 1\rangle = 0\}$ a finite element space such that $S^r_{h,*} \subset \{v \in L_2(\mathcal{D}) : \langle v, 1\rangle = 0\} = \mathcal{H}$ and redefine Λ_h accordingly, i.e.,

$$\langle \Lambda_h x, y\rangle = \langle \nabla x, \nabla y\rangle, \quad x, y \in S^r_{h,*}. \tag{7.5}$$

Further, we let P_h be the projection onto $S^r_{h,*}$. Thus, the finite element formulation reads to find u_h and w_h in $S^r_{h,*}$ such that

$$\begin{aligned} \langle \dot{u}_h, \xi\rangle + \langle \nabla w_h, \nabla \xi\rangle &= 0, \ t > 0, \\ \langle w_h, \eta\rangle - \langle \nabla u_h, \nabla \eta\rangle &= 0, \ t > 0, \\ u_h(0) &= P_h u_0, \end{aligned} \tag{7.6}$$

for all $\xi, \eta \in S^r_{h,*}$.

The homogeneous Neumann boundary conditions are imposed through the bi-linear form, they will be fulfilled automatically by any solution to (7.6). Equation (7.6) may be re-written on operator form as to find $u_h(t) \in S^r_{h,*}$ such that

$$\begin{aligned} \dot{u}_h + (\Lambda_h)^2 u_h = 0,\ t > 0, \\ u_h(0) = P_h u_0, \end{aligned} \tag{7.7}$$

with solution of the form (7.3) but with $E_h(t) = e^{-t(\Lambda_h)^2}$. It is notable that the generator of the analytic semigroup $\{E_h(t)\}_{0\le t\le T}$ is $A_h = (\Lambda_h)^2$ and not $-(\Lambda^2)_h$. This avoids the need to use finite elements with two weak derivatives but makes the error analysis more difficult. The error estimate for the homogeneous case is given below without proof, but see [26].

Theorem 7.2. *Assume that u is the solution of the linearized Cahn-Hilliard equation* (5.18) *and that u_h is the solution of* (7.6). *If $v \in \tilde{H}^\beta$, $t > 0$ and $0 \le \gamma \le \beta \le r$, it holds that*

$$\|(E_h(t)P_h - E(t))v\| \le Ch^\beta t^{-\frac{\beta-\gamma}{4}} \|v\|_{\tilde{H}^\gamma}, \quad 0 \le \beta \le \gamma \le r.$$

For the wave equation we start up by noting that $S^r_{h,0} \times S^r_{h,0}$ is a subspace of $\dot{H}^1 \times \dot{H}^0$ and thus, in a similar fashion as above the finite element version of the wave equation reads to find $U_h(t) = [U_{h,1}(t), U_{h,2}(t)]^T \in S^r_{h,0} \times S^r_{h,0}$ such that

$$\dot{U}_h + \begin{bmatrix} 0 & -I \\ \Lambda_h & 0 \end{bmatrix} U_h = 0,\ t > 0; \quad U_h(0) = \begin{bmatrix} P_h U_{0,1} \\ P_h U_{0,2} \end{bmatrix}$$

or with

$$A_h = \begin{bmatrix} 0 & -I \\ \Lambda_h & 0 \end{bmatrix} \tag{7.8}$$

it may be written

$$\begin{aligned} \dot{U}_h(t) + A_h U_h(t) = 0, t > 0, \\ U_h(0) = P_h X_0 \end{aligned}$$

where we, with an abuse of notation, also has written

$$P_h = \begin{bmatrix} P_h & 0 \\ 0 & P_h. \end{bmatrix} \tag{7.9}$$

The operator $-A_h$ may be seen to be the infinitesimal generator of an C_0-group of contractions $\{E_h(t)\}_{t\in\mathbb{R}}$ where

$$E_h(t) = \begin{bmatrix} C_h(t) & \Lambda_h^{-1/2} S_h(t) \\ -\Lambda_h^{1/2} S_h(t) & C_h(t) \end{bmatrix}$$

with

$$C_h(t)x = \sum_{j=1}^{N_h} \cos(\lambda_{h,j}^{1/2} t) \langle x, \psi_{h,j} \rangle \psi_{h,j}$$

and

$$S_h(t)x = \sum_{j=1}^{N_h} \sin(\lambda_{h,j}^{1/2} t) \langle x, \psi_{h,j} \rangle \psi_{h,j}.$$

Thus the unique solution of this problem may be written

$$U_h(t) = E_h(t) P_h U_0 = \begin{bmatrix} C_h(t) P_h U_{0,1} + \Lambda_h^{-1/2} S_h(t) U_{0,2} \\ -\Lambda_h^{1/2} S_h(t) P_h U_{0,1} + C_h(t) P_h U_{0,2} \end{bmatrix}.$$

For the wave operator no error estimate of the approximation of the full semigroup in the $\mathcal{H}^0$-norm is known to us, but proofs of the following theorem that states error bounds for the displacement, or for the operator $P^1(E(t) - E_h(t)P_h)$ where $P^1[X_1, X_2]^T = X_1$, can be found in [3] and [55].

Theorem 7.3. *Let $\{E(t)\}_{t\in\mathbb{R}}$ be the C_0-group generated by the operator $-A$ defined in* (5.20) *and let $\{E_h(t)\}_{t\in\mathbb{R}}$ be the group generated by the operator $-A_h$ in* (7.8). *Let the operator $P^1 \colon \mathcal{H}^0 \to \dot{H}^0$ be defined by $P^1U = U_1$, $U = [U_1, U_2]^T$. If $X \in \mathcal{H}^s$, $0 \le \beta \le r+1$, then*

$$\|P^1(E(t) - E_h(t)P_h)X\|_{L^2(\mathcal{D})} \le C(t) h^{\frac{r}{r+1}\beta} \|X\|_{\mathcal{H}^\beta}.$$

8 Rational approximations

We have already seen how semigroup theory provides an eloquent setting for studying initial value problems and leads to a powerful framework to formulate and analyze stochastic versions of these. In addition to this, once the semigroup theory is there, error estimates for a broad class of finite difference schemes are achievable with a very uniform approach. This relies on realizing that for the scalar version of (5.1) with solution $u(t) = u_0 e^{-ta}$ many difference schemes means multiplying the last computed value u_{n-1} with a rational function of ka, where k is the time step, to get the next iterate, u_n. Thus, for example, the implicit Euler method

$$u_n = (1 + ka)^{-1} u_{n-1}$$

makes use of the rational function $r(x) = 1/(1+x)$ whereas the Crank-Nicolson method means multiplying with $r(ka)$ where $r(x) = (1 - \frac{1}{2}x)/(1 + \frac{1}{2}x)$ and for a general rational function $r(x)$ this means that we approximate the value of $u(t_n)$ (with $t_n = nk$) by $r(ka)^n u_0$ or, stated otherwise, we approximate the exponential function $e^{-kna} = (e^{-ka})^n$ by $r(ka)^n$. The obvious questions that arises are, first, what rational functions one may accept in order to get a good approximation of the exponential function and, second, which of these may be used to approximate a certain unknown semigroup in terms of the known generator. We will not attempt to answer any of these questions fully but state some important theorems that will be enough for our aims. First we define some important concepts in approximation theory.

Definition 8.1. A rational function r is *A-stable* if

$$|r(z)| \leq 1, \quad \text{Re}(z) \geq 0. \tag{8.1}$$

If, in addition, $r(0) = r'(0) = 1$ then r is said to be *A-acceptable*.

Definition 8.2. If r is a rational function and if

$$r(z) - e^{-z} = \mathcal{O}(z^{q+1}), \quad z \to 0, \tag{8.2}$$

then we say that r is *accurate of order* q. If r is accurate of any order $q \geq 1$ we say that it is *consistent* with the exponential function.

The following theorem was proved in [8].

Theorem 8.3. *If r is an A-acceptable rational approximation, accurate of order q, and if $-A$ is the generator of a C_0-semigroup $\{T(t)\}_{t\geq 0}$ of type $(M, \omega), M \geq 1, \omega \geq 0$ on some Banach space $\mathcal{E}$, then, with $t = nk$ and small enough timestep k, there exists $C > 0$ and $om\tilde{e}ga \in \mathbb{R}$ such that*

$$\|r^n(kA)v - T(t)v\| \leq Cte^{\tilde{\omega}t}k^q\|A^{q+1}v\|, \quad v \in D(A^{q+1}). \tag{8.3}$$

We will now make further assumptions on the generator in order to lessen the restrictions on the rational approximation or improving the convergence rates. For this aim we make the following definition.

Definition 8.4. A rational approximation r of the exponential function is said to be *I-acceptable* if it maps the imaginary axis into the unit disc, i.e., if

$$|r(ix)| \leq 1, \quad x \in \mathbb{R}. \tag{8.4}$$

This turns out to be a good property when to approximate C_0-groups. We have the following theorem from [9].

Theorem 8.5. *Let $-A$ be the generator of a C_0-group $\{T(t)\}_{t\in\mathbb{R}}$ and let r be an I-stable rational approximation of the exponential function, accurate of order q. Then* (8.3) *still holds.*

If $-A$, on the other hand, generates an analytic semigroup then the regularity assumptions on the initial function might be relaxed . More precisely the following holds.

Theorem 8.6. *If, in addition to the assumption of Theorem 8.3, $-A$ is the infinitesimal generator of an analytic semigroup $\{T(t)\}_{t\geq 0}$, then*

$$\|r^n(kA)v - T(t)\| \leq Cte^{\tilde{\omega}t}k^q\|A^q v\|, \quad v \in D(A^q) \tag{8.5}$$

and if $|r(\infty)| < 1$ then

$$\|r^n(kA)v - T(t)v\| \leq Ct^{-q}e^{\tilde{\omega}t}k^q\|v\| \tag{8.6}$$

for all positive t if $v \in \mathcal{H}$.

A proof may be found in [75].

9 Fully discrete schemes

We have above presented error estimates for semi-discrete schemes in both space and time. In most cases, of course, one needs to discretize in both time and space simultaneously. It is not always immediate that convergence will be achieved in k and h regardless of how they tend to zero. The schemes introduced above are unconditionally stable in Hilbert spaces though and the time-step and the spatial discretization may be chosen independent of each other. We will first sketch the discretization procedure and then state error estimates for the Cahn-Hilliard equation and the wave equation. We start by a spatial discretization as outlined in Section 7. This yields as we have seen the equation

$$\dot{U}_h + A_h U_h = 0, \quad U_h(0) = P_h U_0 \tag{9.1}$$

on a finite-dimensional subspace of $\mathcal{H}$. Further, the operator A_h is bounded, hence is $-A_h$ the infinitesimal generator of a C_0-semigroup $\{E_h(t)\}_{t\geq 0}$. Thus the solution of (9.1) may be approximated with the help of rational functions as in Section 8 and we get

$$U^n_{h,k} = r(kA_h)U^{n-1}_{h,k}, \quad U^0_{k,h} = P_h U_0. \tag{9.2}$$

Naively, we could try to analyze the error $U(T) - U^N_{h,k}$ by adding and subtracting $U_h(T)$ and using the triangle inequality and the results from Section 7 and Section 8 to get

$$\begin{aligned} &\|(E(T) - r(kA_h)^N P_h)U_0\| \\ &\quad\leq \|(E(T) - E_h(T)P_h)U_0\| + \|(E_h(T) - r(kA_h)^N P_h)U_0\| \\ &\quad\leq C_1 h^r \|A^{\tilde{r}} U_0\| + C_2 k^p \|A_h^{\tilde{p}} P_h U_0\| \end{aligned} \tag{9.3}$$

for some $\tilde{r}$ and some $\tilde{p}$ depending on the problem under consideration. If we would be able to show that $C_2 k^p \|A_h^{\tilde{p}} P_h U_0\| \leq Ck^p \|A^{\tilde{p}} U_0\|$ we would be done. This is not so simple in most cases and the way to prove a bound on the full discretization error differs between different equations. We omit the sometimes technical details of this and refer to [3] in the case of the wave equation, to Paper II in the case of the CHC equation and to [75] for a proof of fully discrete schemes for the heat equation. The latter case is not analyzed in this thesis, but see Remark 5.2 in Paper II.

The results that we will use below are the following.

Theorem 9.1. *Let* $A = \Lambda^2$ *where* $D(\Lambda) = \{v \in \mathcal{H} \cap H^2(\mathcal{D}) : \frac{\partial v}{\partial} n = 0\}$ *and let* $\{E(t)\}_{0\leq t\leq T}$ *be the analytic semigroup generated by* $-A$. *Further, let* Λ_h *be given by* (7.5) *with the order of the finite element space* $q = 2$, *and let* $A_h = (\Lambda_h)^2$. *If* $r(x) = (1+x)^{-1}$ *and if* $T = kN$, *then*

$$\|(E(T) - r(kA_h)^N P_h)U_0\|_{\mathcal{H}} \leq CT^{-\alpha/4}\big(h^\alpha + k^{\alpha/4}\big)\|U_0\|. \tag{9.4}$$

As mentioned already the operator $-\Lambda^2$ in the Cahn-Hilliard equation generates an analytic semigroup whereas the wave operator is the generator of a (non-analytic) group. Thus the result in the latter case looks a bit different as in the following theorem.

Theorem 9.2. *Let A be given by* (5.20) *with $D(A) = \mathcal{H}^1$ and let $\{E(t)\}_{0\le t\le T}$ be the group generated by $-A$ on $\mathcal{H}$. Let also A_h given by* (7.8) *and let the finite element space be of order q. If r is an I-stable rational approximation of order p, if $T = kN$ and if $U_0 \in \mathcal{H}^s$, $s = \max(q, p)$, then*

$$\left\| P^1\big(E(T) - r(kA_h)^N P_h\big)U_0 \right\|_{L^2(\mathcal{D})} \le C(T)(h^q \|U_0\|_{\mathcal{H}^q} + k^p \|U_0\|_{\mathcal{H}^p}). \tag{9.5}$$

We note here that Theorem 9.2 only gives us results for the first component, the displacement, of the wave equation. No convergence results for the full wave semigroup when discretized in both time and space is known to us. However, in the light of Theorem 8.5, Theorem 9.2 and the findings in [55] we offer the following conjecture.

Conjecture 9.3. With the fully discrete scheme described above the error estimate for the wave semigroup is given by

$$\|(E(T) - r(kA_h)^N P_h)U_0\| \le CT(h^r \|U_0\|_{\mathcal{H}^{r+1}} + k^p \|U_0\|_{\mathcal{H}^{p+1}}). \tag{9.6}$$

10 Frames and wavelets

If we assume that $\mathcal{U} = L_2(\mathcal{D})$, a typical covariance operator Q on $\mathcal{U}$ being of trace class has a representation as a integral operator with kernel $q\colon \mathcal{D}\times\mathcal{D} \to \mathbb{R}$, i.e.,

$$\big(Qu\big)(x) = \int_{\mathcal{D}} q(x,y)u(y)\,\mathrm{d}y.$$

Discretizing this with, say, finite elements will lead to a full matrix $\mathbb{Q}$ where

$$\mathbb{Q}_{ij} = \langle Q\psi_i, \psi_j\rangle = \int_{\mathcal{D}\times\mathcal{D}} q(x,y)\psi_i(x)\psi_j(y)\,\mathrm{d}x\mathrm{d}y, \tag{10.1}$$

ψ_i and ψ_j being finite element basis functions, that needs to be factorized into $\mathbb{Q} = \mathbf{L}\mathbf{L}^T$ in order to be able to simulate the noise.[3] If the kernel has some decay properties the matrix will however be quasisparse since if the two basis functions ψ_i and ψ_j in (10.1) have supports that are distant from each other, then $\mathbb{Q}_{ij}$ will be much smaller than when the supports are less distant. However, the quasi

[3] Theoretically, Cholesky factorization would be the method of choice here if Q is strictly positive definite. However, if the eigenvalues of Q decay rather rapidly, Cholesky factorization methods will fail due to round-off errors and methods for symmetric indefinite matrices must be used. This has two drawbacks. First, they are slightly slower in theory and, second, they are not available in standard parallel linear algebra packages as ScaLAPack, [72].

sparsity is captured rather badly by finite elements and thus, in Paper III, approximations of the noise by using a wavelet basis is investigated since wavelets are known to have better approximation properties of integral operators than finite elements.

To give a brief description of the most useful mathematical properties of wavelets we need the concept of a frame.

Definition 10.1. A countable subset $\{\phi_j\}_{j\in\mathcal{J}}$ of a Hilbert space H is a *frame* for H if there exists constants $a, b > 0$ such that for every $f \in H$ it holds that

$$a\|f\|^2 \leq \sum_{j\in\mathcal{J}} |\langle f, \phi_j\rangle|^2 \leq b\|f\|^2. \tag{10.2}$$

The numbers a and b will be denoted *lower* and *upper frame constants*, respectively. If $a = b = 1$, then the frame is said to be *tight*.

For every frame $\{\phi_j\}_{j\in\mathcal{J}}$ for H there exists a *dual frame* $\{\tilde{\phi}_j\}_{j\in\mathcal{J}}$, a frame such that $\langle \phi_i, \tilde{\phi}_j\rangle = \delta_{ij}$, i.e., they are biorthogonal. The dual frame is also a frame for H having lower frame constant $1/b$ and upper frame constant $1/a$. Note that every ON-basis of H is a tight frame. A frame is not necessarily a basis (it may be redundant) but when it is, it is denoted a *Riesz basis*.

It turns out that the trace of a trace class operator T may be evaluated using an arbitrary frame and its dual. To be precise we have that

$$\mathrm{Tr}(T) = \sum_{j\in\mathcal{J}} \langle T\phi_j, \tilde{\phi}_j\rangle$$

and

$$a|\,\mathrm{Tr}(T)| \leq \Big|\sum_{j\in\mathcal{J}} \langle T\phi_j, \phi_j\rangle\Big| \leq b|\,\mathrm{Tr}(T)|.$$

The concept of a frame assures some minimal approximation properties but these will not suffice in most cases. In order to make fruitful assumptions we need some more notation. From now on, we will solely be concerned with the case $\mathcal{U} = L_2(\mathcal{D})$. First, we will assume that there is a hierarchical description of the frame, i.e.,

$$\mathcal{J} = \{(i,k)\colon i \in \mathbb{N}, i \geq i_0, k \in J_i\}, \quad i_0 \in \mathbb{N}.$$

The variable J_i is an index set that depends on the spatial dimension d and the *level*, i. For two elements $\phi_{i,k}$ and $\phi_{j,l}$ in a frame we denote their common support by Δ_{ikjl}, that is,

$$\Delta_{ikjl} = \operatorname{supp}\phi_{i,k} \cap \operatorname{supp}\phi_{j,l}.$$

We will enforce the following assumptions on the frames.

1. The size of the support of the elements of the frame decreases with the level, i.e.,

$$\mathrm{diam}(\operatorname{supp}\phi_{i,k}) \sim \mathrm{diam}(\operatorname{supp}\tilde{\phi}_{i,k}) \sim 2^{-i}.$$

2. There is a limit of the number of basis functions on each level. More precisely
$$\#J_i \leq C2^i.$$

3. There is also a limit of the number of basis functions on two levels with overlapping support. We assume that for $i \geq j$ it holds that
$$\#\{k \in J_i \colon \Delta_{ikjl} \neq 0\} \leq C2^{i-j}.$$

4. The elements of the frame has $\tilde{m}$ *vanishing moments* and the dual frame has m vanishing moments. This means that
$$|\langle f, \phi_{i,k}\rangle| \leq C2^{-j(s+d/2)}|f|_{W^{s,\infty}(\operatorname{supp}\phi_{i,k})}, \quad s \leq \tilde{m}, j \geq j_0 \text{ and}$$
$$|\langle f, \tilde{\phi}_{i,k}\rangle| \leq C2^{-j(s+d/2)}|f|_{W^{s,\infty}(\operatorname{supp}\tilde{\phi}_{i,k})}, \quad s \leq m, j \geq j_0.$$

5. Finally we assume that the frame and its dual fulfill the inverse estimates
$$\|\phi_{i,k}\|_{H^s(\mathcal{D})} \leq C2^{si}\|\phi_{i,k}\|_{L_2(\mathcal{D})}, \quad 0 \leq s \leq \gamma, \text{ and}$$
$$\|\tilde{\phi}_{i,k}\|_{H^s(\mathcal{D})} \leq C2^{si}\|\tilde{\phi}_{i,k}\|_{L_2(\mathcal{D})}, \quad 0 \leq s \leq \tilde{\gamma}$$
where $\gamma = m - \frac{1}{2}$ and $\tilde{\gamma}$ is an increasing function of $\tilde{m}$.

A Riesz basis that fulfills assumptions 1–5 will be referred to as a *wavelet basis* and its dual as the dual wavelet basis. This is not in full agreement with how the concept of a wavelet basis is usually used but is still motivated by the fact that we have a basis of localized oscillating functions. Bases of this kind does exist and one explicit construction is carried out in for example [18].

We are thus ready with our description of the theoretical framework in what we have worked and are ready to discuss the attached papers.

11 Introduction to the papers

The appended papers split up in two categories. In the first two papers, a framework for analyzing weak convergence of numerical schemes for linear stochastic partial differential equations is developed and applied to the stochastic heat and wave equations and the Cahn-Hilliard-Cook equation. In the first paper spatial discretization is investigated whereas the second treats semidiscretization in time as well as full (both space and time) discretization. To avoid repetition, we will treat the two as one and state the results of the both together. Indeed, the results of the first paper can be seen as a special case of the results in the second by taking the time step $k = 0$.

The third and fourth papers are quite different from the preceding ones. First, strong convergence is investigated and, second, they are focused on noise representation. In Paper III it is shown how the driving Wiener process may be represented with an arbitrary frame. This representation comes as an infinite sum. By truncating this sum we get a new, approximate Wiener process. A new equation of the form (1.1) may thus be formulated, using the approximate noise instead, and then solved. An error formula for the solution of the truncated version is thus given.The truncated equation is then approximated in its turn by the finite element method and the error of this further approximation is thus analyzed. Finally, the frame is taken to be a wavelet basis with standard properties and convergence rates are computed when the number of basis function in the truncated representation is coupled to the mesh size in the finite element mesh.

The last paper takes a different approach to truncation of the noise. Here the starting point is the eigenfunction representation (6.1). If the eigenpairs of Q are unknown, one may attempt to find them with the finite element method. This is computationally expensive. But in certain cases one does not need all N_h eigenpairs. If the eigenvalues decay sufficiently fast, then the discrete series may be truncated more or less severely. Conditions on the kernel of Q that imply a certain decay rate, and therefore a truncation level, are stated.

11.1 Paper I and II – Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise

As mentioned above, the first two appended papers are indeed very similar, both concerned with weak error analysis of perturbations of stochastic processes that solve a stochastic partial differential equation. In Paper I, spatially semidiscrete schemes are investigated. First a general and exact error formula is stated that involves differences between the semigroups stemming from the original problem and its finite element approximation, respectively. In the proof of this formula it is only assumed that the operator $-A$ in (1.1) is the generator of a C_0-semigroup. In the second part of the paper the error formula is used to prove convergence rates for the stochastic heat, wave and Cahn-Hilliard equations. In the second paper it is pointed out that the error formula holds for more general perturbations than the finite element approximations. Most importantly, also time discrete schemes may be analyzed with the same methods.

In all cases, we study approximations of the problem (1.1). We study the spatially semidiscrete finite element problem

$$\mathrm{d}X_h(t) + A_h X_h(t)\,\mathrm{d}t = B_h\,\mathrm{d}W(t),\ t > 0; \quad X_h(0) = P_h X_0, \tag{11.1}$$

with solution

$$X_h(t) = E_h(t)P_h X_0 + \int_0^t E_h(t-s)B_h\,\mathrm{d}W(s).$$

We study the temporally semidiscrete problem

$$X_k^n = r(kA)(X_k^{n-1} + B\Delta W^n),\ n \geq 1; \quad X_k^0 = X_0,$$

where $\Delta W^n = W(t_n) - W(t_{n-1})$ with solution

$$X_k^n = r(kA)^n X_0 + \sum_{j=1}^{n} r(kA)^{n-j+1} B\Delta W^j,$$

and we study the fully discrete scheme

$$X_{h,k}^n = r(kA_h)(X_{h,k}^{n-1} + B_h \Delta W^n), \ n \geq 1; \quad X_k^0 = P_h X_0.$$

with the solution

$$X_{h,k}^n = r(kA)^n P_h X_0 + \sum_{j=1}^{n} r(kA_h)^{n-j+1} B_h \Delta W^j.$$

The general error formula is given by the following theorem.

Theorem 11.1. *Let* $\{X(t)\}_{0\leq t\leq T}$ *be the unique weak solution of* (1.1) *with*

$$\mathrm{Tr}\Big(\int_0^T E(t)BQ(E(t)B)^* \,\mathrm{d}t\Big) < \infty$$

and let $\tilde{X}(T)$ *be an approximation of* $X(T)$ *such that there exists a process* $\{\tilde{Y}(t)\}_{0\leq t\leq T}$ *with* $\tilde{Y}(T) = \tilde{X}(T)$ *such that*

$$\tilde{Y}(t) = \tilde{Y}(0) + \int_0^t \tilde{E}(T-r)\tilde{B}\,\mathrm{d}W(r), \tag{11.2}$$

where

$$\mathrm{Tr}\Big(\int_0^T \tilde{E}(t)\tilde{B}Q(\tilde{E}(t)\tilde{B})^* \,\mathrm{d}t\Big) < \infty.$$

If $G \in C_{\mathrm{b}}^2(\mathcal{H}, \mathbb{R})$, *then*

$$\begin{aligned}
&\mathbb{E}\Big[G(\tilde{X}(T)) - G(X(T))\Big] \\
&\quad = \int_0^t \Big\langle u_x\Big(\big(\tilde{E}(T)\tilde{X}_0 - E(T)X_0\big)r + E(T)X_0, 0\Big), \tilde{E}(T)\tilde{X}_0 - E(T)X_0\Big\rangle \,\mathrm{d}r \\
&\quad + \frac{1}{2}\int_0^T \mathrm{Tr}\Big(u_{xx}(\tilde{Y}(r), r)\big(\tilde{E}(T-r)\tilde{B} - E(T-r)B\big) \\
&\qquad\qquad \times Q\big(\tilde{E}(T-r)\tilde{B} + E(T-r)B\big)^*\Big) \,\mathrm{d}r.
\end{aligned}$$

The significance of this formula is that all terms in the left hand side contain, essentially, the difference between the original semigroup $\{E(t)\}_{0\leq t\leq T}$ and the perturbed family $\{\tilde{E}(t)\}_{0\leq t\leq T}$. Since this deterministic error is usually known from the literature, we may estimate the weak error by additional manipulations.

This is rather tricky and has to be done differently depending on the type of semigroup.

To prove Theorem 11.1 is not straightforward. The proof makes use of Itô's formula (6.8) and Kolmogorov's backward equation (6.11) as well as the auxiliary functions Y and Z defined in Section 6 above Theorem 6.6.

Due to the use of Itô's formula in the proof the form (11.2) is important. To get there from a time discrete process $\{X_k^j\}_{j=0}^N$ one defines the new process $\{Y_k^j\}_{j=0}^N$ in a discrete analogy to how the function Y was defined in (6.9). To arrive at (11.2) one then uses a standard stochastic interpolation. This is described in detail in Section 4.2 of Paper II.

11.1.1 Weak convergence of numerical schemes for the stochastic heat equation

In Paper I a semidiscrete finite element method is investigated for the heat equation in arbitrary spatial dimensions. We gather here Theorems 4.1 and 4.2 from Paper I. They differ only in the assumption made on the operator Q.

Theorem 11.2. *Assume that X solves* (1.1) *with $A = \Lambda$, $D(A) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$ and $B = I$, where $\mathcal{U} = \mathcal{H} = L^2(\mathcal{D})$ and that X_h solves* (11.1) *with $A_h = \Lambda_h$ and $B_h = P_h$ on $S_{h,0}^r$. If $G \in C_{\mathrm{b}}^2(\mathcal{H}, \mathbb{R})$ and X_0 is sufficiently smooth and if*

$$\|\Lambda^{\beta-1} Q\|_{\mathrm{Tr}} < \infty \tag{11.3}$$

for some $\beta \in (0, \frac{r}{2}]$, where r is the order of the finite element method, then

$$|\mathbb{E}[G(X_h(T)) - G(X(T))]| \leq C h^{2\beta} |\log(h)|. \tag{11.4}$$

If $\beta \in (0, 1]$, then (11.4) *holds with the condition* (11.3) *replaced by*

$$\|\Lambda^{(\beta-1)/2} Q^{1/2}\|_{\mathrm{HS}} < \infty. \tag{11.5}$$

The first condition (11.3), which is used with a larger range of β, is less symmetric than the second one (11.5). This is due to difficulties associated with manipulating the terms in the error representation formula. We know that (11.3) implies (11.5) for $\beta \geq 0$. Moreover, under assumption (11.3) we have strong convergence of order h^β for $0 \leq \beta \leq r$. Therefore, the weak rate is essentially twice the strong rate.

For the linear Cahn-Hilliard-Cook equation, the findings are in agreement with the results for the heat equation: The weak rate is essentially twice the strong rate. Since the analysis is similar as for the heat equation, only a bit more complicated and with only slightly different assumptions on the covariance, we omit to state the results here and refer to the papers.

11.1.2 Weak and strong convergence of numerical schemes for the stochastic wave equation

In papers I and II we also apply our methodology to semidiscrete and completely discrete approximations of the stochastic wave equation. Also here we find that weak convergence rates are twice the rates of strong convergence up to certain maximal levels. The findings of Theorem 5.1 from Paper I and Theorems 4.3 and 4.5 from Paper II are summarized below. Note that the test function g acts only on the first component X_1, while the test function G acts on the whole vector X.

Theorem 11.3. *Let $X(T) = [X_1(T), X_2(Y)]^T$ be the solution of the stochastic wave equation at time T, $X_h(T) = [X_{h,1}(T), X_{h,2}(T)]$ the (spatially) semi-discrete finite element approximation thereof, $X_k^N = [X_{k,1}^N, X_{k,2}^N]^T$ a (temporally) semi-discrete rational approximation of $X(T)$ and $X_{h,k}^N = [X_{h,k,1}^N, X_{h,k,2}^N]^T$ a fully discrete approximation. Let further $G \in C_{\mathrm{b}}^2(\mathcal{H}, \mathbb{R})$ and $g \in C_{\mathrm{b}}^2(\dot{H}^0, \mathbb{R})$ and assume that*

$$\|\Lambda^{\beta-1/2} Q \Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty \tag{11.6}$$

for some $\beta \geq 0$. If $r \geq 1$ is the order of the finite element approximation and $p \geq 1$ of the rational approximation and if the initial value is sufficiently smooth, then

$$\mathbb{E}[g(X_{h,1}(T)) - g(X_1(T))] \leq C h^{\min(\frac{r}{r+1} 2\beta, r)}, \tag{11.7}$$

$$\mathbb{E}[G(X_k^N) - G(X(T))] \leq C k^{\min(\frac{p}{p+1} 2\beta, 1)}$$

and

$$\mathbb{E}[g(X_{h,k,1}^N) - g(X_1(T))] \leq C(h^{\min(\frac{r}{r+1} 2\beta, r)} + k^{\min(\frac{p}{p+1} 2\beta, 1)}).$$

We remark that if Conjecture 9.3 is true then by inspection of the proof of (11.3) in Theorem 4.5 of Paper II it is immediate that the same convergence rates will be achieved if $\mathbb{E}[g(X_{h,k,1}^N) - g(X_1(T))]$ is replaced by $\mathbb{E}[G(X_{h,k}^N) - g(X(T))]$. This applies to the spatially semidiscrete case as well.

Since, in contrast to the heat equation, strong convergence results for fully discrete schemes with finite elements in arbitrary spatial dimension were unknown to us, we also investigated this in Paper II.

The rate of strong convergence for the displacements is given by the following theorem (Theorem 4.6 in Paper II). Again we know that the less symmetric condition (11.6) implies condition (11.8).

Theorem 11.4. *If, in Theorem 11.3, the assumption* (11.6) *is replaced by*

$$\|\Lambda^{\frac{\beta-1}{2}} Q^{1/2}\|_{\mathrm{HS}} < \infty, \tag{11.8}$$

then the mean-square error of the fully discrete scheme is given by

$$\left(\mathbb{E}[\|X_{h,k,1}^N - X_1(T)\|^2]\right)^{1/2} \leq C(k^{\min(\beta \frac{p}{p+1}, 1)} + h^{\min(\beta \frac{r}{r+1}, r)}).$$

11.2 Paper III – Spatial approximation of stochastic convolutions

The focus of this paper is on the solution to (1.1) when $X_0 = 0$. Thus we study the stochastic convolution

$$W_A(t) = \int_0^t E(t-s)B\,\mathrm{d}W(s). \tag{11.9}$$

The Ansatz is to represent the driving Q-Wiener process $W(t)$ in terms of a frame for $\mathcal{H}$ and then truncating it so that (formally if Q is not of trace class)

$$W(t) = \sum_{j\in\mathcal{J}} \langle W(t), \tilde{\phi}_j\rangle \phi_j \approx \sum_{j\in J} \langle W(t), \tilde{\phi}_j\rangle \phi_j =: W^J(t),$$

where $J \subset \mathcal{J}$ is finite. The process is then a $P_J Q P_J^*$-Wiener process, where P_J is the projection defined by

$$P_J f = \sum_{j\in J} \langle f, \tilde{\phi}_j\rangle \phi_j.$$

It is a fairly immediate consequence that the error formula

$$\mathbb{E}\big[\|W_A(t) - W_A^J(t)\|^2\big] = \int_0^t \|E(s)B(I-P_J)Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}s$$

holds. Moreover, both in the case of the stochastic wave and heat equations one can show that

$$\begin{aligned}\mathbb{E}\big[\|W_A(t) - W_A^J(t)\|^2\big] &\le C(t)\|\Lambda^{-1/2}(I-P_J)Q^{1/2}\|_{\mathrm{HS}}^2 \\ &= C(t) \sum_{j,k\in\mathcal{J}\setminus J} \langle \Lambda^{-1}\phi_j, \phi_k\rangle\langle Q\tilde{\phi}_j, \tilde{\phi}_k\rangle.\end{aligned}$$

Since Λ^{-1} and often also Q have representations in terms of integral operators this is rigged up for using wavelets of the kind described in Section 10.

The next step is to solve the truncated equation

$$\mathrm{d}X(t) + AX(t)\,\mathrm{d}t = B\,\mathrm{d}W^J(t) \tag{11.10}$$

by the finite element method. Thus, we need to assume that the underlying Hilbert spaces $\mathcal{U}$ and $\mathcal{H}$ are function spaces where the finite element method may be defined. Doing so our problem is to solve the equation

$$\mathrm{d}X_h(t) + A_h X_h(t)\mathrm{d}t = P_h B\,\mathrm{d}W^J(t). \tag{11.11}$$

As (11.10) has the solution

$$W_A^J(t) = \int_0^t E(t-s)BP_J\,\mathrm{d}W(s)$$

and (11.11) is solved by

$$W^J_{A_h}(t) = \int_0^t E_h(t-s)P_h B P_J \,\mathrm{d}W(s)$$

we have that

$$W^J_{A_h}(t) - W^J_A(t) = \int_0^t (E_h(t-s)P_h - E(t-s))BP^J \,\mathrm{d}W(s)$$

and by the Itô isometry

$$\mathbb{E}\big[\|W^J_{A_h}(t) - W^J_A(t)\|^2\big] = \int_0^t \|(E_h(t-s)P_h - E(t-s))BP^J\|^2_{\mathrm{HS}} \,\mathrm{d}s.$$

If the elements in the frame $\{\phi_j\}_{j\in\mathcal{J}}$ have enough regularity ($\phi_j \in \dot{H}^{\beta-1}$) it holds for the heat equation that

$$\begin{aligned}\mathbb{E}\|W^J_{A_h}(t) - W^J_A(t)\|^2 &\leq Ch^{2\beta}\|\Lambda^{(\beta-1)/2}P_J Q^{1/2}\| \\ &\leq Ch^{2\beta}\sum_{j,k\in\mathcal{J}}\langle\Lambda^{(\beta-1)/2}\phi_j, \Lambda^{(\beta-1)/2}\phi_k\rangle\langle Q\tilde{\phi}_j, \tilde{\phi}_k\rangle,\end{aligned} \tag{11.12}$$

if $0 \leq \beta \leq r$, where r is the order of the finite element method. For the first component of the wave equation a similar result is proved but with $h^{2\beta}$ replaced by $h^{\frac{2r\beta}{r+1}}$ with $0 \leq \beta \leq r+1$ and C replaced by a time dependent constant $C(t)$.

The obvious question arising is what kind of frames that could, and should, be used. Eigenfunctions for either the generator or the covariance operator are obvious candidates if they are known and it is shown in the paper that in one spatial dimension one achieves the optimal convergence rate $h^{1/2}$ for white noise if the eigenfunctions of Λ are used. In order to get this convergence rate the number of eigenfunctions N must be chosen such that $N = 1/h$, which means that the numbers of eigenfunctions and finite element functions shall be the same. Choosing instead the Haar-basis, that is, a wavelet basis of the type described in Section 10 with $m = \tilde{m} = 1$, consisting of piecewise constant functions, [76, Section 2.1], leads to the same convergence rate if one uses instead only $N = -\log_2(h)$ basis functions. This is good news also since quadrature with piecewise constant functions usually is less expensive than with trigonometric functions.

The reason for this result is not so surprising. That the noise is white means that it is spatially uncorrelated and using a basis with as small support as possible can therefore be expected to be good. The Haar-basis is a wavelet with minimal support whereas the eigenfunctions are non-zero over almost the entire spatial domain, a fact that is reflected in the mentioned result.

Yet another thing investigated is when, for the heat equation with piece-wise linear finite elements, the optimal rate of h^2 is achieved. It is enough that Q has an integral kernel $q \in W^{3,\infty}(\mathcal{D}\times\mathcal{D})$ and to use a wavelet basis with $m, \tilde{m} \geq 2$. This convergence rate is then achieved if $N = -\log_2(h)$. The proof is based on the fact that the Green's function of the Laplace operator $g \in W^{1,\infty}(\mathcal{D}\times\mathcal{D})$ and the assumptions of the wavelets in Section 10 used in (11.2) and (11.12).

11.3 Paper IV – Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise

In this paper we expand the noise in terms of the computed eigenvalues of the approximate covariance operator $Q_h = P_h Q P_h$ and study whether the resulting expansion may be truncated and to what extent. The equation under investigation is a semilinear parabolic SPDE, i.e., of the form

$$\mathrm{d}X(t) + AX(t)\,\mathrm{d}t + f(X(t))\,\mathrm{d}t = \mathrm{d}W(t)$$

where f is Lipschitz continuous on $\mathcal{H}$.

For a finite element space S_h of dimension N_h the eigenspace of the discretized covariance matrix $\mathbb{Q}_h$ has an eigenspace of dimension N_h and this is also the maximal number of eigenfunctions that could be found. Using them all would be equivalent to using some other (possibly faster) factorization of $\mathbb{Q}_h$. If the decay of the eigenvalues of Q and thus $\mathbb{Q}_h$ is fast, then only a few terms in (6.1) needs to be taken into account and so also with its discrete counterpart, i.e., the sum

$$W_h(t) = \sum_{j=1}^{N_h} \gamma_{h,j}^{1/2} \beta_j(t) e_{h,j}$$

where the eigenpairs $(\gamma_{h,j}, e_{h,j})$ solve the eigenvalue problem

$$Q_h e_{h,j} = \gamma_{h,j} e_{h,j}. \tag{11.13}$$

We also assume that the sequence $\{\gamma_{h,j}\}_{j=1}^{N_h}$ is sorted in decreasing order. The operator Q_h is defined as

$$\langle Q_h x, y\rangle = \langle Qx, y\rangle, \quad x, y \in S_h.$$

The way the analysis is performed is rather standard. An error estimate of the non-truncated solution X_h is computed and then the truncated finite element solution $X_h^N(t)$ is compared to the solution of the non-truncated problem. Under certain regularity assumptions the error is proved to be

$$\mathbb{E}[\|X_h(t) - X_h^N(t)\|^2] \leq C\|A^{-1/2}\|^2 \sum_{j=N}^{N_h} \gamma_k.$$

Thus the problem is to find certain characterizations of the covariance kernel q such that $N = N(N_h)$ can be taken as small as possible and keep the convergence rate. Three different criteria where analyzed. First, if $q\colon \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ is analytic, then one may take only $N = c(\log(N_h))^d$ terms in (11.13). Second, if q is only smooth, then one may take $N = N_h^{\frac{2\beta}{sd}}$ for any $s > \max(1, 2\beta/d)$. Third, if $q \in H^p \otimes H^0$, then one has to take $N = N_h^{\frac{2\beta}{p}}$ if $p > \max(d, 2\beta)$. If p is larger, then no possibilities for truncation have been proved.

It is worth mentioning that this is in agreement with the findings in the previous section. Truncation of eigenfunction expansions work well for smooth noise but it works less satisfactorily when the noise is nonsmooth. To get a more complete description of this would be beneficial.

12 Corrections to the appended papers

We have not discovered any errors in the two first appended papers but, unfortunately, there are a few mistakes in the subsequent ones. We point these out below and show how to correct the errors. All numbers below refer to the paper discussed.

12.1 Errors in Paper III

- In the proof of Lemma 2.3 we claim that the eigenvalues of TT^* and T^*T coincide. It is not true. What is true that their nonzero eigenvalues coincide but that is exactly what we need for the trace equality.
- In Theorem 5.1 and in Theorem 6.1 the assumption

$$\|\Lambda^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$$

can be completely removed. This is since W_A^J and $W_{A_h}^J$ are well defined whatever Q is because J is finite. Since we assume that $\phi_j \in \dot{H}^{\beta-1}$ it follows that $\Lambda^{\frac{\beta-1}{2}} P_J Q^{1/2}$ is Hilbert-Schmidt automatically again as long as Q is bounded. Indeed,

$$\Lambda^{\frac{\beta-1}{2}} P_J Q^{1/2} x = \sum_{j \in J} \langle x, Q^{1/2} \tilde{\phi}_j \rangle \Lambda^{\frac{\beta-1}{2}} \phi_j$$

which is a trace class operator by (4.3) and hence also Hilbert-Schmidt.

- From the first paragraph in the proof of Theorem 7.1 it is clear that the primal frame needs to be in the domain of $\Lambda^{1/2}$, i.e., in $\dot{H}^1$. But this is not a problem; such wavelets can be constructed, see [76] and references therein. It is also important here that the assumption $\|\Lambda^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$ is not required in Theorems 5.1 and 6.1 because this would impose an additional requirement on the kernel q in Theorem 7.1.

12.2 Error in Paper IV

The inequality in (2.4) does not hold. It holds, however, with exponent $-1/2$ instead of -1. It turns out that this is what is needed anyway. We used (2.4) to get (3.8) from (3.7) in the paper. But using now (2.4) with exponent $-1/2$ instead of -1 and that $\gamma_{h,k} \leq \gamma_k$ we get

$$\sum_{k=M}^{N_h} \gamma_{h,k} \|A_h^{-\frac{1}{2}} e_{h,k}\|^2 \leq C \sum_{k=M}^{N_h} \gamma_k \|A^{-\frac{1}{2}} e_{h,k}\|^2 \leq C \|A^{-\frac{1}{2}}\|^2 \sum_{k=M}^{N_h} \gamma_k.$$

Thus, Theorem 3.4 holds.

References

[1] E. J. Allen, S. J. Novosel, and Z. Zhang, *Finite element and difference approximation of some linear stochastic partial differential equations*, Stochastics Stochastics Rep. **64** (1998), 117–142.

[2] A. Andersson and S. Larsson, *Weak error for the approximation in space of the non-linear stochastic heat equation*, preprint 2012.

[3] G.A Baker and J.H. Bramble, *Semidiscrete and single step fully discrete approximations for second order hyperbolic equations*, RAIRO Numer. Anal. **13** (1979), 76–100.

[4] A. Barth and A. Lang, *Milstein approximation for advection-diffusion equations driven by multiplicative noncontinuous martingale noises*, Applied mathematics and optimization **66** (2012), 387–413.

[5] ——, *Multilevel Monte Carlo method with applications to stochastic partial differential equations*, International Journal of Computer Mathematics, doi:10.1080/00207160.2012.701735.

[6] A. Barth, A. Lang and C. Schwab, *Multilevel Monte Carlo method for parabolic stochastic partial differential equations*, BIT (2012) (online first).

[7] D. Blömker, S. Maier-Paape, and T. Wanner, *Second phase spinodal decomposition for the Cahn-Hilliard-Cook equation*, Trans. Amer. Math. Soc. **360** (2008), 449–489.

[8] P. Brenner and V. Thomée, *On rational approximations of semigroups*, SIAM J. Numer. Anal. **16** (1979), 683–694.

[9] P. Brenner and V. Thomée, *On rational approximations of groups of operators*, SIAM J. Numer. Anal. **17** (1980), 119–125.

[10] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Elements*, second edition, Texts in Applied Mathematics **15**, Springer, 2002.

[11] Z. Brzeźniak, *Some remarks on Itô and Stratonovich integration in 2-smooth Banach spaces*, Probabilistic methods in fluids, World Sci. Publ., 48–69, 2003.

[12] C. Cardon-Weber, *Implicit approximation scheme for the Cahn-Hilliard stochastic equation*, Preprint 2001, www.proba.jussieu.fr/mathdoc/textes/PMA-613bis.ps.

[13] D. Cohen, S. Larsson and M. Sigg, *A trigonometric method for the linear stochastic wave equation*, arXiv:1203.3668.

[14] H. E. Cook, *Brownian motion in spinodal decomposition*, Acta Metallurgica **18** (1970), 297–306.

[15] G. Da Prato, A. Jentzen and M. Röckner, A mild Ito formula for SPDEs, arXiv:1009.3526v4 [math.PR].

[16] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications **44**, Cambridge University Press, 1992.

[17] ——, *Second Order Partial Differential Equations in Hilbert Spaces*, London Mathematical Society Lecture Note Series **293**, Cambridge University Press, 2002.

[18] W. Dahmen, A. Kunoth, and K. Urban, *Biorthogonal spline wavelets on the interval—stability and moment conditions*, Appl. Comput. Harmon. Anal. **6** (1999), 132–196.

[19] I. Daubechies, *Ten Lectures on Wavelets*, CBMS–NSF Regional conference series in applied mathematics **61**, Society of industrial and applied mathematics, 1992.

[20] A. M. Davie and J. G. Gaines, *Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations*, Math. Comp. **70** (2001), 121–134, (electronic).

[21] 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.

[22] A. Debussche, *Weak approximation of stochastic partial differential equations: the non linear case*, Math. Comp. **80** (2011), 89–117.

[23] A. Debussche and J. Printems, *Weak order for the discretization of the stochastic heat equation*, Math. Comp. **78** (2009), 845–863.

[24] J. Demmel, *Applied Numerical Linear Algebra*, SIAM, 1997.

[25] Q. Du and T. Zhang, *Numerical approximation of some linear stochastic partial differential equations driven by special additive noises*, SIAM J. Numer. Anal. **40** (2002), 1421–1445 (electronic).

[26] C. M. Elliott and S. Larsson, *Error estimates with smooth and nonsmooth data for the finite element method for the Cahn-Hilliard equation*, Math. Comp. **58** (1992), 603–630.

[27] K. J. Engel and R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Graduate Texts in Mathematics **194**, Springer, 1991.

[28] L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics **19**, American Mathematical Society, 1998.

[29] 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.

[30] W. Grecksch, P. E. Kloeden, *Time-discretised Galerkin approximations of parabolic stochastic PDEs*, Bull. Aust. Math. Soc. **54** (1996), 79–85.

[31] I. Gyöngy, *Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. I*, Potential Anal. **9** (1998), 1–25.

[32] ——, *Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. II*, Potential Anal. **11** (1999), 1–37.

[33] ——, *Approximations of stochastic partial differential equations*, in Stochastic partial differential equations and applications (Trento, 2002), 287–307, Lecture Notes in Pure and Appl. Math. **227**, Dekker, New York, 2002.

[34] I. Gyöngy and A. Millet, *On discretization schemes for stochastic evolution equations*, Potential Anal. **23** (2005), 99–134.

[35] ——, *Rate of convergence of implicit approximations for stochastic evolution equations*, In P. Baxendale and S. Lototsky (eds.), Stochastic differential equations: Theory and Applications (A volume in honor of Boris L. Rosovskii) **2** (281–310), World scientific interdisciplinary mathematical sciences. World Scientific, 2007.

[36] ——, *Rate of convergence of space time approximations for stochastic evolution equations*, Potential Anal. **30** (2009), 29–64.

[37] I. Gyöngy and D. Nualart, *Implicit scheme for quasi-linear parabolic partial differential equations perturbed by space-time white noise*, Stochastic Process. Appl. **58** (1995), 57–72.

[38] ——, *Implicit scheme for stochastic parabolic partial differential equations driven by space-time white noise*, Potential Anal. **7** (1997), 725–752.

[39] E. Hausenblas, *Numerical analysis of semilinear stochastic evolution equations in Banach spaces*, J. Comput. Appl. Math. **147** (2002), 485–516.

[40] ——, *Approximation for semilinear stochastic evolution equations*, Potential Anal., **18** (2003), 141–186.

[41] ——, *Weak approximation for semilinear stochastic evolution equations*, Stochastic Analysis and Related Topics VIII, Progr. Probab. **53**, 111–128, Birkhäuser, 2003.

[42] ——, *Weak approximation of the stochastic wave equation*, J. Comput. Appl. Math. **235** (2010), 33–58.

[43] A. Jentzen, *Higher order pathwise numerical approximations of SPDEs with additive noise* , SIAM J. Numer. Anal. **49** (2011), 642–667.

[44] A. Jentzen and P. E. Kloeden, *Overcoming the barrier in the numerical approximation of stochastic partial differential equations with additive space-time noise*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **465** (2009), 649–667.

[45] ——, *The numerical approximation of stochastic partial differential equations*, Milan J. Math. **77** (2009), 205–244.

[46] A. Jentzen, P. E. Kloeden and G. Winkel, *Efficient simulation of nonlinear parabolic SPDEs with additive noise*, Ann. Appl. Probab. **21** (2011), 908–950.

[47] A. Jentzen and M. Röckner, *A Milstein scheme for SPDEs*, arXiv:1001.2751v4.

[48] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Applications of Mathematics **23**, Springer-Verlag, 1992.

[49] G. T. Kossioris and G. E. Zouraris, *Fully-discrete finite element approximations for a fourth-order linear stochastic parabolic equation with additive space-time white noise* ESAIM: Mathematical Modelling and Numerical Analysis **44** (2010), 289–322.

[50] 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), 309–320.

[51] ——, *Spatial approximation of stochastic convolutions*, J. Comput. Appl. Math. **235** (2011), (3554–3570).

[52] ——, *Weak convergence of finite element approximations of stochastic evolution equations with additive noise*, BIT **52** (2012), 85–108.

[53] ——, *Weak convergence of finite element approximations of stochastic evolution equations with additive noise II, Fully discrete schemes*, BIT online first (2012).

[54] M. Kovács, S. Larsson and A. Mesforush, *Finite element approximation of the Cahn-Hilliard-Cook equation*, SIAM J. Numer. Anal. 49 (2011), 2407-2429.

[55] M. Kovács, S. Larsson and F. Saedpanah, *Finite element approximation of the linear stochastic wave equation with additive noise*, SIAM J. Numer. Anal. **48** (2010), 408–427.

[56] M. Kovács and J. Printems, *Strong order of convergence of a fully discrete approximation of a linear stochastic Volterra type evolution equation*, preprint, `www.maths.otago.ac.nz/~mkovacs/KovacsPrintemsArxiv.pdf`.

[57] R. Kruse, *Optimal error estimates of galerkin finite element methods for stochastic partial differential equations with multiplicative noise*, `arXiv:1103.4504`.

[58] ——, *Strong and Weak Approximation of Semilinear Stochastic Evolution Equations*, PhD thesis, University of Bielefeld, 2012.

[59] S. Larsson and V. Thomée, *Partial Differential Equations with Numerical Methods*, Texts in applied mathematics **45**, Springer-Verlag, 2003.

[60] S. Larsson and A. Mesforush, *Finite element approximation of the linear stochastic Cahn-Hilliard equation*, IMA J. Numer. Anal. **31** (2011), 1315–1333.

[61] A. Lang and J. Potthoff, *Fast simulation of Gaussian random fields*, Monte Carlo Meth. Appl. **13** (2011), 195–214.

[62] P. D. Lax, *Functional Analysis*, Wiley, 2002.

[63] F. Lindner, R. L. Shilling, *Weak order for the discretization of the stochastic heat equation driven by impulsive noise*, Potential Analysis. doi:10.1007/s11118-012-9276-y.

[64] N. Marheineke and R. Wegener, *Modeling and application of a stochastic drag for fiber dynamics in turbulent flows*, Int. J. Multiph. Flow **37** (2011), 136–148.

[65] S. P. Nørsett, G. Wanner, *The real-pole sandwich for rational approximations and oscillation equations*, BIT **19** (1979), 79–94.

[66] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied mathematical sciences **44**, Springer-Verlag, 1983.

[67] J. Printems, *On the discretization in time of parabolic stochastic partial differential equations*, Math. Model. Numer. Anal. **35** (2001), 1055–1078.

[68] L. Quer-Sardanyons and M. Sanz-Solé, *Space semi-discretisations for a stochastic wave equation*, Potential Anal. **24** (2006), 303–332.

[69] T. Müller-Gronbach and K. Ritter, *Lower bounds and nonuniform time discretization for approximation of stochastic heat equations*, Found. Comput. Math. **7** (2007), 135–181.

[70] ——, *An implicit Euler scheme with non-uniform time discretization for heat equations with multiplicative noise*, BIT **47** (2007), 393 – 418.

[71] T. Müller-Gronbach, K. Ritter and T. Wagner, *Optimal pointwise approximation of infinite-dimensional Ornstein-Uhlenbeck processes*, Stoch. Dyn. **8** (2008), 519–514.

[72] ScaLAPack homepage: www.netlib.org/scalapack/.

[73] T. Shardlow, *Numerical methods for stochastic parabolic PDEs*, Numer. Funct. Anal. Optim. **20** (1999), 121–145.

[74] ——, *Weak convergence of a numerical method for a stochastic heat equation*, BIT **43** (2003), 179–193.

[75] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer Series in Computational Mathematics **25**, Springer Verlag, 2006.

[76] K. Urban, *Wavelet Methods for Elliptic Partial Differential Equations*, Numerical Mathematics and Scientific Computation, Oxford University Press, 2009.

[77] J. B. Walsh, *Finite element methods for parabolic stochastic PDEs*, Potential Anal. **23** (2005), 1–43.

[78] J. B. Walsh, *On numerical solutions of the stochastic wave equation*, Illinois J. Math. **50** (2006), 991–1018.

[79] W. Walter, *Ordinary Differential Equations*, Readings in Mathematics **182**, Springer, 1998.

[80] X. Wang and S. Gan, *Weak convergence analysis of the linear implicit euler method for semilinear stochastic partial differential equations with additive noise*, Journal of Mathematical Analysis and Applications **398** (2012), 151–169.

[81] J. Weidmann, *Linear Operators in Hilbert Spaces*, Graduate Texts in Mathematics **68**, Springer, 1980.

[82] Y. Yan, *Semidiscrete Galerkin approximation for a linear stochastic parabolic partial differential equation driven by additive noise*, BIT **44** (2004), 829–847.

[83] Y. Yan, *Galerkin finite element methods for stochastic parabolic partial differential equations*, SIAM J. Numer. Anal. **43** (2005), 1363–1384.