Approximation of Evolution Equations with Random Data

Vom Promotionsausschuss der
Technischen Universität Hamburg

zur Erlangung des akademischen Grades

Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation (Monografie)

von
Katharina Kliobа

aus
Hamburg

2024

1. Gutachter: PD Dr. rer. nat. habil. Christian Seifert
2. Gutachter: Prof. Dr. ir. Mark Veraar

Tag der mündlichen Prüfung: 17. September 2024

doi:https://doi.org/10.15480/882.13663
ORCiD: https://orcid.org/0009-0002-7946-917X

Summary

Evolution equations are partial differential equations (PDEs) describing the evolution over time of, e.g., heat, waves, or epidemics. To account for random perturbations, random coefficients or noise terms are added to the model, resulting in a random or stochastic evolution equation, respectively. Adding random data often renders a numerical solution the only possibility. This entails taking the random data into account during approximation. For evolution equations with random coefficients, this is done by using polynomial chaos approximation (PCE) to reduce to a deterministic problem. For stochastic PDEs (SPDEs), the noise has to be incorporated into the space and time discretisation. Quantifying the speed of convergence provides helpful insights for numerical simulations.

The contributions of this thesis are twofold. First, a joint rate of convergence in randomness, space, and time is obtained for the full discretisation error under certain regularity conditions. Second, stability, qualitative, and quantitative convergence are obtained for the pathwise uniform error in time for semilinear SPDEs in the hyperbolic Kato setting.

Regarding random evolution equations, our focus lies on C_0-semigroups induced by a form, for which a novel quantified version of the Trotter–Kato theorem yields spatial convergence rates. PCE error estimates are extended from the one-dimensional scalar-valued to the multivariate vector-valued case. To obtain a deterministic coupled system of abstract Cauchy problem structure, PCE has to be applied in a non-straightforward manner, leading to a different semi-discretisation in randomness than the PCE of the mild solution. Using semigroup theory, we derive pathwise estimates and lift them via a composition estimate. A modified Trotter–Kato argument then yields polynomial decay of the semi-discretisation error of arbitrary order, depending on the regularity of the initial values in the random parameter and, in analogy to the deterministic case, also the spatial regularity. Further approximating in space and time, we derive regularity conditions on the form under which convergence rates are conserved for the full discretisation. Applications cover a parabolic equation with random anisotropic diffusion.

For semilinear SPDEs with Gaussian noise, optimal bounds on the *uniform strong error* $\mathrm{E}_k^\infty := (\mathbb{E}\sup_{j\in\{0,\ldots,N_k\}} \|U(t_j) - U^j\|_X^p)^{1/p}$ are derived, where $p \in [2,\infty)$, X is a 2-smooth Banach space, U is the mild solution, U^j is the temporal semi-discretisation, k is the step size, and $N_k = T/k$ for some final time $T > 0$. Under conditions on the nonlinearity and the noise, for a large class of time discretisation schemes, we show

- $\mathrm{E}_k^\infty \lesssim k\sqrt{\log(T/k)}$ (linear equation, additive noise, general S);
- $\mathrm{E}_k^\infty \lesssim \sqrt{k}\sqrt{\log(T/k)}$ (nonlinear equation, multiplicative noise, contractive S);
- $\mathrm{E}_k^\infty \lesssim k\sqrt{\log(T/k)}$ (nonlinear wave equation, multiplicative noise).

The bounds obtained coincide with the optimal bounds for SDEs. The logarithmic factor can be removed if the exponential Euler method is used. Error estimates are extended to the full time interval $[0,T]$ via novel path regularity results. Stability in the pathwise uniform sense is proven by martingale techniques. By virtue of a regularisation argument, convergence, albeit without rate, is extended to rough initial values and irregular Lipschitz nonlinearities and noise. For stochastic Maxwell equations, Schrödinger equations, and wave equations, these results improve and reprove several existing results with a unified method and provide the first results known for implicit Euler and Crank–Nicolson.

Acknowledgements

Writing my PhD thesis would certainly have been a much more tedious, if not impossible, task without the support of several important people around me. First of all, I am deeply thankful to my supervisor Christian Seifert who gave me excellent support and encouragement in every aspect I could have wished for, and this already since being my mentor in my first semester ten years ago. Many valuable insights in mathematics and beyond originate from discussions with him.

Two years ago, I was very fortunate to encounter who feels like a second PhD supervisor, albeit not on paper. I am extremely grateful to Mark Veraar for introducing me to the world of SPDEs, which I hopefully get to enjoy for a little bit longer, as well as to De Lelie ice cream. Thank you for welcoming me in Delft with open arms.

Besides, I want to thank my colleagues from the Institute of Mathematics at TUHH for creating an enjoyable and supportive work environment. A distinguished mention deserve my office mate Riko Ukena, Dennis Schmeckpeper, Marko Lindner, Vanessa Trapp, Judith Angel, Lina Fesefeldt, Thorben Abel, Yannick Mogge, Michela Ascolese, and Marco Wolkner for many helpful discussions, coffee breaks, and the cake circle.

Equally, I would like to thank my colleagues from Delft for the warm welcome I received during my time there. The special atmosphere in the analysis group really impressed me and I sincerely enjoyed my stay. Especially, I am grateful to Jan van Neerven for great discussions and creative solutions to housing challenges, Max Sauerbrey for letting me steal his office spot, as well as Floris Roodenburg, Esmée Theewis, Joris van Winden, and Joshua Willems.

Scientific and personal exchange have been a cornerstone of my PhD. I would like to thank the DAAD for their financial support to spend six months at TU Delft. Moreover, I would like to thank the GAMM, the GAMM Juniors, the internet seminar community, and everybody else I had great discussions with on conferences or summer schools.

Furthermore, I thank my family and friends for their invaluable love and support through both the highs and the lows of the last four years. Special thanks go to Skadi for being such a great friend and to Alessa for being my favourite (and only) sister.

Finally, my heartfelt thanks go to a wonderful person without whom my PhD journey would only have been half as fun and twice as complicated. Thank you for everything, Philippe.

Contents

Chapter 1

Introduction

1.1 Why study Evolution Equations with Random Data?

The functional analytic treatment of a time-dependent partial differential equation (PDE) results in a description of it as an ordinary differential equation in a Banach space; that is, the PDE is formulated as an abstract Cauchy problem; see, e.g., the classical monographs on this topic [45, 66, 111]. From a practical point of view, this allows us to model a variety of different physical phenomena, such as heat conduction or diffusion phenomena [4, 110], wave propagation [23, 132], how a quantum mechanical state evolves over time [2], or electromagnetic fields [28, 107].

In applications, the coefficients of these PDEs often originate from material laws involving material parameters, which may be unknown or can only be determined up to some uncertainty. It can also be that a random forcing is present in the system under investigation or there are randomly appearing fluctuations, which disturb our model. As an example, one could think of random particle movements (described by Brownian motion) or measurement noise. Moreover, this is an important factor in order to take unforeseen events and external influences into account when modeling, say, the spreading of a disease in the population [145], which has undeniably been of significant societal importance during the preparation of this thesis.

When modeling such situations with uncertain data, from a mathematical point of view, one has to decide which class of equations to consider. In this thesis, we distinguish between two different types of incorporating uncertainty in the model: *evolution equations with random coefficients* and *stochastic evolution equations* with (additive or multiplicative) noise. In both cases, the equations can typically not be solved analytically. Supposing well-posedness of the model thus obtained, in order to then approximate the solution, one has to perform suitable discretisations, taking the randomness into account. Naturally, the question arises of how the randomness affects the convergence behaviour compared to the deterministic setting. For applications, it is paramount to quantify this convergence behaviour depending on the regularity of the equation and the numerical scheme employed in order to make reliable predictions on the accuracy of simulations.

The first possibility of modeling uncertainty consists of choosing suitable random variables (or stochastic processes, random fields) to describe the coefficients in the PDE. This approach results in a *random time-dependent PDE*, where the coefficients are assumed to be determined by *finite-dimensional noise.* Evolution equations with parametric or random coefficients have been treated extensively in the past decades; see [53, 101, 119, 138] and references therein, as well as [8, 9, 10, 26, 27, 44, 47, 56, 125, 139, 140, 141, 142, 143] (just to mention a few), with an emphasis on stochastic computation as well as approximation. From a numerical analysis point of view, a large portion is devoted to elliptic PDEs. To obtain a numerical approximation, one has to perform suitable discretisations with respect to the randomness, the spatial, as well as the temporal variables. One of the central aims of this thesis is to provide such an approximation, including convergence results, in the framework of random evolution equations given by abstract Cauchy problems with random generators. We work mainly from a functional analytic point of view and, for this type of equations, solely work in the Hilbert space setting.

While in the first model the randomness can be treated separately for the discretisation, eventually resulting in a deterministic system as an approximation, this is not the case for the second model considered: introducing a random forcing term in the equation, which might even depend on the solution itself, leads to the notion of *stochastic partial differential equations* (SPDEs). This can be viewed as *infinite-dimensional noise*, which is classified as *multiplicative* or *additive*, depending on whether it depends on the solution itself or not. Mathematically, this noise is modeled by a (cylindrical) Brownian motion. Consequently, and in contrast to the first model, a wide range of tools from stochastic analysis are required, first and foremost, stochastic integration. The structure of the equation forces us to incorporate the discretisation of the noise into the space and time discretisation by approximating the action of the Brownian motion. Since the presence of the Brownian motion leads to significantly lower regularity in time of solutions of the SPDE, many challenges arise for the temporal discretisation. In general, we cannot expect similarly high convergence rates as in the deterministic case. The main contribution of this thesis to the temporal approximation of SPDEs consists of a general framework for hyperbolic SPDEs that yields optimal convergence rates for the pathwise uniform error. This improves and extends various results for specific hyperbolic SPDEs from recent years (see, e.g., [2, 28, 132]). Our framework allows us to consider such equations not only in Hilbert spaces, but in a special class of Banach spaces, to be specified later.

1.2 Evolution Equations with Random Coefficients

As an illustrative example of an evolution equation with random coefficients, let us consider the heat equation with random coefficients given by

$$\begin{aligned} \partial_t\widetilde{u}(t,x,\omega) &= \operatorname{div}\widetilde{a}^\omega(x)\operatorname{grad}\widetilde{u}(t,x,\omega) && (t>0, x\in G, \omega\in\Omega),\\ \widetilde{u}(t,x,\omega) &= 0 && (t>0, x\in\partial G, \omega\in\Omega),\\ \widetilde{u}(0,x,\omega) &= \widetilde{u}_0(x,\omega) && (x\in G, \omega\in\Omega), \end{aligned}$$

where $G\subseteq\mathbb{R}^d$ is a bounded domain, $(\Omega,\mathcal{F},\mathbb{P})$ is a probability space, and, for $\omega\in\Omega$, $\widetilde{a}^\omega$ maps to $\mathbb{K}^{d\times d}$ and is $\mathbb{P}$-almost surely uniformly elliptic and bounded, grad acts on the x-coordinate, and $\widetilde{u}_0$ is the initial condition. Here, $\mathbb{K}\in\{\mathbb{R},\mathbb{C}\}$ denotes the scalar field of all vector spaces appearing throughout. We assume that the random fields $(x,\omega)\mapsto\widetilde{a}^\omega(x)$ and $(x,\omega)\mapsto u_0(x,\omega)$ are determined by *finite-dimensional noise*; that is, there exists a random vector $Z\colon\Omega\to\mathbb{R}^N$ such that $\widetilde{a}^\omega(x)=a^{Z(\omega)}(x)$ for all $x\in G$ and $\omega\in\Omega$ for some $a\colon G\times\mathbb{R}^N\to\mathbb{K}^{d\times d}$. Such finite-dimensional noise can be obtained by, e.g., a truncated Karhunen–Loève expansion [79, 82, 103]. Pushing the random heat equation forward w.r.t. Z, we obtain another random heat equation, where now the random parameter comes from the probability space $(\mathbb{R}^N,\mathcal{B}(\mathbb{R}^N),\mathbb{P}_Z)$:

$$\begin{aligned} \partial_t u(t,x,z) &= \operatorname{div}a^z(x)\operatorname{grad}u(t,x,z) && (t>0, x\in G, z\in\mathbb{R}^N),\\ u(t,x,z) &= 0 && (t>0, x\in\partial G, z\in\mathbb{R}^N),\\ u(0,x,z) &= u_0(x,z) && (x\in G, z\in\mathbb{R}^N). \end{aligned}$$

Taking the functional analytic point of view, we set $H:=L^2(G)$ and consider the unbounded operator $A_z:=-\operatorname{div}a^z\operatorname{grad}$ on H for $z\in\mathbb{R}^N$. Note that A_z is the operator associated with some form a_z for $z\in\mathbb{R}^N$. Thus, we obtain the family of random abstract Cauchy problems

$$u_z'(t) = -A_z u_z(t) \quad (t>0), \quad u_z(0)=u_{0,z}$$

for $z \in \mathbb{R}^N$. Considering $(A_z)_{z\in\mathbb{R}^N}$ as a multiplication operator $\mathbf{A}$ on the Hilbert space $\mathbf{H} := L^2(\mathbb{R}^N, \mathbb{P}_Z; H)$ results in the abstract Cauchy problem in $\mathbf{H}$

$$\mathbf{u}'(t) = -\mathbf{A}\mathbf{u}(t) \quad (t > 0), \quad \mathbf{u}(0) = \mathbf{u}_0.$$

In order to solve this problem numerically, we perform three independent discretisations:

(i) discretisation in randomness: here we choose the (generalised) Polynomial Chaos Expansion (PCE); cf. [136, 140],

(ii) discretisation in space: here we choose an abstract Galerkin method (a finite element method for the concrete example above); cf. [22, 51],

(iii) discretisation in time: here we use an A-stable method; cf. [21, 41].

Clearly, in the deterministic case, approximation results and convergence rates are classical and well-known; see, e.g., [49, 124] and references therein. As it turns out, there is an intricate relationship between spatial and temporal discretisation. In particular, when first discretising in space, spatial regularity is also required for the estimation of the semi-discretisation error in time. We will exploit the corresponding relationship for the discretisation in randomness and in space-time, and find a similar behaviour for the full discretisation in the random setting.

Using PCE, we can obtain a deterministic coupled system as an approximation, which again has the form of an abstract Cauchy problem, now in $\mathfrak{H}_n := H^{d_n}$, where d_n is the number of H-valued equations determined by the discretisation parameter n and the dimension N of the noise. This system can then be approximated in space and time, resulting in a fully discretised approximation. What we will trace explicitly is the interplay between these two discretisation steps (randomness vs. space-time). One has to be careful to use PCE in a meaningful way to maintain the abstract Cauchy problem structure when deriving the deterministic coupled system. In particular, the semi-discretisation $\mathbf{u}_n(t)$ of $\mathbf{u}(t)$ in randomness does not agree with the PCE approximation of $\mathbf{u}(t)$ in most cases. Therefore, the estimation of the semi-discretisation error $\|\mathbf{u}(t) - \mathbf{u}_n(t)\|_{\mathbf{H}}$ is more than a simple application of a PCE error bound. Instead, in Theorem 3.70, we employ a modified Trotter–Kato argument based on pointwise in z estimates of A_z and its resolvent, which are then lifted to Sobolev subspaces of $\mathbf{H}$ via a composition estimate. In analogy to the deterministic case, spatial regularity is required for the estimation of the part of the full error due to semi-discretisation in randomness.

Altogether, we then obtain an estimate of the form

$$\|\mathbf{u}(t) - \mathbf{J}_{n,m}\mathbf{u}_{n,m,k}(t)\|_{\mathbf{H}} \leq C\big(n^{-\ell} + m^{-p_{\mathrm{x}}} + \tau^{p_{\mathrm{t}}}\big)\|\mathbf{u}_0\|_{H^{2\ell}_{\rho}(\mathbb{R}^N, \mathbb{P}_Z; D^{\max\{\alpha+1, p_{\mathrm{t}}\}})}$$

in the case of symmetric forms a_z (see our main result in Theorem 3.74). Here, $\mathbf{J}_{n,m}$ is an embedding, n is the PCE parameter, m the spatial discretisation parameter, k the temporal discretisation parameter, τ the maximal step size for the time discretisation, and t a temporal grid point. Moreover, $\ell \in \mathbb{N}_0$, ρ is a suitable weight, p_{x} and p_{t} are the rates for the spatial and the temporal methods, respectively, and $\alpha > 0$ is such that the spatial method converges with rate p_{x} on D^α, where D is the by assumption common domain of the A_z's. If the forms a_z are not $\mathbb{P}_Z$-almost surely symmetric, additional spatial regularity is required, and we only consider specific time discretisation schemes. In this case, for all $\varepsilon > 0$, in Theorem 3.75, we show

$$\|\mathbf{u}(t) - \mathbf{J}_{n,m}\mathbf{u}_{n,m,k}(t)\|_{\mathbf{H}} \leq C_\varepsilon\big(n^{-\ell} + m^{-p_\mathrm{x}} + \tau^{p_\mathrm{t}}\big)\|\mathbf{u}_0\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\max\{\alpha,p_\mathrm{t}\}+\varepsilon})}.$$

In the time-independent elliptic case, we note that we can use the tensor product structure $L^2(\mathbb{R}^N, \mathbb{P}_Z; L^2(G)) = L^2(\mathbb{R}^N, \mathbb{P}_Z) \otimes L^2(G)$, so discretisations in randomness and space can be performed simultaneously; see, e.g., [9]. Moreover, the case when the diffusion coefficient is affine in the random inputs has been thoroughly studied; cf. [27, 47, 125]. In contrast, we allow for fairly general dependence of the coefficients on the random inputs; in particular, they need not be holomorphic as in [26].

1.3 Stochastic Evolution Equations

The second main contribution of this thesis concerns the temporal discretisation of nonlinear stochastic PDEs driven by additive or multiplicative Gaussian noise. The equations we consider can be written as abstract stochastic evolution equations on a Hilbert space X of the form

$$\begin{cases} \mathrm{d}U + AU\,\mathrm{d}t = F(U)\,\mathrm{d}t + G(U)\,\mathrm{d}W_H & \text{on } [0,T], \\ \qquad U(0) = u_0 \in L^p(\Omega; X). \end{cases} \tag{1.3.1}$$

Here, $-A$ is the generator of a C_0-semigroup $(S(t))_{t\geq 0}$, W_H is a cylindrical Brownian motion, F and G are globally Lipschitz, u_0 is the initial data, and $p \in [2, \infty)$.

Our aim is to obtain strong convergence rates for temporal discretisation schemes that cover the hyperbolic setting. The hyperbolic setting has been extensively studied in recent years (see [2, 3, 11, 14, 20, 23, 28, 29, 30, 31, 33, 36, 37, 62, 70, 75, 91, 92, 93, 96, 132, 133] and references therein). In the parabolic setting, (i.e., $(S(t))_{t\geq 0}$ being an analytic semigroup) regularisation phenomena occur, which make it possible to prove very different convergence results. In the non-parabolic case, new methods to show convergence rates are needed and related to a way to obtain regularity. Kato's setting for the hyperbolic case from his seminal work [84] creates a way to obtain this regularity, which has proven to be very useful in the analysis of quasilinear equations as well as their numerical treatment [43, 67, 68, 89, 116].

The main idea in Kato's setting is to consider two spaces X and Y with $Y \hookrightarrow X$ (or sometimes even three spaces) on which the operator A and the nonlinearities F and G can be analysed. In this way, one can obtain regularity of U and better mapping properties of the nonlinearities. In numerical approximations, the obtained regularity can be used to obtain convergence rates, as illustrated for the deterministic case in the references above.

The above setting often also applies to the parabolic case, in which, however, the required mapping properties of F on Y can often be avoided due to the regularising effect of the convolution with the analytic semigroup S. For these equations, it does not seem necessary to work with the Kato setting, as regularisation phenomena can be exploited. For details on the parabolic case, the reader is referred to [4, 12, 13, 16, 34, 42, 59, 76, 77, 78, 81, 94, 95, 97, 104, 131] and references therein. Consequently, our focus lies on the hyperbolic setting.

Leaving the Kato setting and merely assuming global Lipschitz continuity and no further regularity assumptions on the nonlinearity F and noise G, we cannot expect to obtain a convergence rate. Rather, our goal is to show pathwise uniform convergence of contractive time discretisation schemes for such irregular nonlinearities and rough initial data.

1.3.1 The setting

In the above-mentioned literature on the hyperbolic case (and often in the parabolic case), the error considered usually is the *pointwise strong error*

$$\sup_{j\in\{0,\dots,N_k\}} \mathbb{E}\|U(t_j) - U^j\|_X^p, \tag{1.3.2}$$

where U denotes the mild solution to (1.3.1) and $(U^j)_{j=0}^{N_k}$ an approximation thereof. The latter is recursively determined by $U^0 = u_0$,

$$U^j = R_k U^{j-1} + kR_k F(U^{j-1}) + R_k G(U^{j-1})\Delta W^j, \quad j = 1, \dots, N_k, \tag{1.3.3}$$

for some time discretisation scheme $R_k \in \mathcal{L}(X)$ that is an approximation of the semigroup S at time $k > 0$. Here, $N_k = T/k$ is the number of time grid points, $k = t_j - t_{j-1}$ is the uniform step size, $t_j = jk$, and $\Delta W^j = W_H(t_j) - W_H(t_{j-1})$.

When performing numerical simulations to approximate the solution of a stochastic equation, one naturally wants the simulation to be close to the solution of (1.3.1). However, (1.3.2) being small does not provide enough information to conclude this, see Example 1.1. Also, from a probabilistic point of view, (1.3.2) contains no information on the convergence of the path. Instead, it is a more meaningful question to find convergence rates for the *uniform strong error*

$$\mathbb{E} \sup_{j\in\{0,\dots,N_k\}} \|U(t_j) - U^j\|_X^p, \tag{1.3.4}$$

where now the supremum over j is inside the expectation. Hence, it describes moments of the maximal error in time rather than the maximum in time of moments of the error. In the deterministic setting, there is no difference between (1.3.2) and (1.3.4). In the presence of noise, it can, however, not be controlled by the simpler pointwise strong error (1.3.2) without a loss of rate. Indeed, if the pointwise strong error converges at rate $\alpha > 0$, i.e., (1.3.2) is bounded by $CN^{-\alpha p}$ for some $C > 0$, then

$$\begin{aligned}\mathbb{E} \sup_{j\in\{1,\dots,N\}} \|U(t_j) - U^j\|_X^p &\le \mathbb{E} \sum_{j=1}^{N} \|U(t_j) - U^j\|_X^p = \sum_{j=1}^{N} \mathbb{E}\|U(t_j) - U^j\|_X^p \\ &\le N \sup_{j\in\{1,\dots,N\}} \mathbb{E}\|U(t_j) - U^j\|_X^p \le CN^{1-\alpha p}.\end{aligned} \tag{1.3.5}$$

Taking p-th roots, convergence at rate $\alpha - \frac{1}{p}$ is obtained. Since we have arbitrarily slow rates $\alpha \in (0, \frac{1}{2}]$ and are also interested in the case $p = 2$, this estimate is not strong enough for our purposes. Still, convergence of the whole path can be expected, as numerical simulations suggest [2, 28, 132].

It is a widely known open problem in the field to find optimal estimates for (1.3.4). Estimates where the supremum is inside the expectation are usually referred to as *maximal* *estimates*, and ample literature is available on them for general stochastic processes [121]. However, for processes that do not have any Gaussian or martingale structure, it can be quite complicated to prove (sharp) maximal estimates. Even maximal estimates for the mild solution U to (1.3.1) with $F = 0$ and $G(u)$ replaced by a progressively measurable $g \in L^2(\Omega \times (0,T); X)$, are unknown in general (see the survey [127, Section 4] for details). The difference between the errors (1.3.2) and (1.3.4) is illustrated in the following simple example.

Example 1.1. *Let $\Omega = [0,1]$ and let $\mathbb{P}$ denote the Lebesgue measure. For $\gamma \in (0,1]$, let $v_N\colon \Omega \times [0,1] \to \mathbb{R}$ be given by $v_N(\omega,t) = 1$ if $|t-\omega| < 1/(2N^\gamma)$, and zero otherwise. Then one can check that the following error estimates hold:*

$$\sup_{t\in[0,1]} \mathbb{E}|v_N(t)|^p \leq \frac{1}{N^\gamma} \quad \text{and} \quad \mathbb{E} \sup_{t\in[0,1]} |v_N(t)|^p = 1.$$

One even has $\sup_{t\in[0,1]} |v_N(\omega,t)| = 1$ for any $\omega \in \Omega$. This shows the discrepancy between having the supremum inside the expectation or not. Continuity of v_N plays no role here. Indeed, one can easily replace the indicator function by a continuous function without influencing the above error estimates.

In the case where S generates a C_0-*group*, it is known how to estimate the uniform strong error (1.3.4) for the *exponential Euler method* (i.e., $R_k = S(k)$). In this case, one can use the group structure in the following way

$$\int_0^t S(t-s)g(s)\,\mathrm{d}W_H(s) = S(t)\int_0^t S(-s)g(s)\,\mathrm{d}W_H(s),$$

and, similarly, for the discrete approximation. This makes it possible to avoid maximal estimates for stochastic convolutions and use martingale techniques instead. This technique was first applied in [132] to obtain optimal convergence rates for the uniform strong error of the exponential Euler method for abstract wave equations. Later, this technique was extended to other settings (see [2, 14, 28, 38]), and, in particular, applied to stochastic Schrödinger and Maxwell equations. However, if S is not a group, this technique is no longer applicable. Equations in which S is not a group include transport equations, equations with dissipation (e.g. damped wave equations), parabolic equations, etc. Of course, there are also many important systems where groups are unavailable (e.g. if a parabolic equation is coupled to a wave or transport equation). Even more importantly, for schemes involving rational approximations (e.g. implicit Euler, Crank–Nicolson), it is unclear how to use the C_0-group structure to estimate the uniform strong error, since the group does not appear in the scheme.

On the other hand, for other discretisation schemes estimates for the simpler pointwise strong error (1.3.2) are available (see e.g. the above-mentioned papers in the hyperbolic case). Moreover, simulations suggest that optimal rates of convergence for the uniform strong error (1.3.4) hold as well. The main goal of Chapter 4 is to prove such optimal bounds for (1.3.4) for more general semigroups and more general schemes. In particular, proofs of such bounds are presented under the condition that S and R are contractive. This solves the open problem on optimal rates for (1.3.4) for this class of semigroups and numerical schemes.

Structural assumptions on the nonlinearity F and the noise G have to be made to obtain convergence rates for (1.3.4). Namely, we suppose that the spatial regularity of the argument is preserved in the sense of the mapping property $F\colon Y \to Y$, and likewise for G. Moreover, F and G are assumed to be of linear growth on Y and the initial data shall be of the same additional regularity. Naturally, the question arises whether these assumptions can be relaxed. Clearly, we cannot expect to preserve the convergence rate because this already fails in the linear deterministic case. The convergence can be arbitrarily slow if the initial values alone are rough. However, it is an open question whether qualitatively, pathwise uniform convergence holds under weaker assumptions on F and G as well as for rough initial data from $L^p(\Omega; X)$. The second main goal of Chapter 4 is to prove that this question can be answered positively, merely assuming global Lipschitz continuity of F and G on the full space X. This allows us to treat stochastic evolution equations with significantly more irregular nonlinearities and noise.

1.3.2 Main results

As in Kato's setting for the hyperbolic case, let X and Y be Hilbert spaces with $Y \hookrightarrow X$. For $\alpha \in (0,1]$ we say that a time discretisation scheme $R\colon [0,T] \to \mathcal{L}(X), k \mapsto R_k$ *approximates S to order α on Y* if there is a constant $C_\alpha \geq 0$ such that for all $x \in Y$, $k > 0$, and $j \in \{0, \ldots, N_k\}$

$$\|(S(t_j) - R_k^j)x\|_X \leq C_\alpha k^\alpha \|x\|_Y,$$

where $R_k^j = (R_k)^j$ denotes the j-th power of the scheme at time step k. Such a time discretisation scheme is called *contractive* if $\|R_k\|_{\mathcal{L}(X)} \leq 1$. Denote the set of Hilbert–Schmidt operators from a Hilbert space H to X by $\mathcal{L}_2(H, X)$. Our main result on convergence rates for (1.3.4) is as follows.

Theorem 1.2. *Let X and Y be Hilbert spaces such that $Y \hookrightarrow X$. Let $-A$ be the generator of a C_0-contraction semigroup $(S(t))_{t\geq 0}$ on X and Y. Suppose that $(R_k)_{k>0}$ is a time discretisation scheme which is contractive on both X and Y, that R approximates S to order $\alpha \in (0, 1/2]$ on Y, and that $Y \hookrightarrow \mathrm{D}(A^\alpha)$. Suppose that $F\colon X \to X$ and $G\colon X \to \mathcal{L}_2(H, X)$ are Lipschitz continuous, and that $F\colon Y \to Y$ and $G\colon Y \to \mathcal{L}_2(H, Y)$ are of linear growth. Let $p \in [2, \infty)$, $u_0 \in L^p(\Omega; Y)$, and U be the mild solution to (1.3.1). Let $k \in (0, T/(2p)]$ and let $(U^j)_{j=0}^{N_k}$ be given by (1.3.3). Then there is a constant $C_T \geq 0$ not depending on u_0 and k such that*

$$\Big\| \max_{0\leq j\leq N_k} \|U(t_j) - U^j\|_X \Big\|_{L^p(\Omega)} \leq C_T(1 + \|u_0\|_{L^p(\Omega;Y)}) k^\alpha \sqrt{\log(T/k)}. \tag{1.3.6}$$

In particular, the approximations $(U^j)_j$ converge at rate α as $k \to 0$ up to a logarithmic factor.

Explicit formulas for C_T are included in the main text. Our second main result concerning the convergence for irregular nonlinearities and noise is the following.

Theorem 1.3. *Let X be a Hilbert space. Suppose that $-A$ generates a C_0-contraction semigroup $(S(t))_{t\geq 0}$ on both X and $\mathrm{D}(A)$. Let $(R_k)_{k>0}$ be a contractive time discretisation scheme on both X and $\mathrm{D}(A)$ that approximates S to some order $\alpha \in (0, 1]$ on $\mathrm{D}(A)$.*

Suppose that $F\colon X \to X$ and $G\colon X \to \mathcal{L}_2(H,X)$ are Lipschitz continuous. Let $T > 0$, $p \in [2,\infty)$, and $u_0 \in L^p(\Omega;X)$. Denote by U the mild solution to (1.3.1) *on* $[0,T]$. *Let $k \in (0,T/2]$ and $(U^j)_{j=0,\dots,N_k}$ be given by* (1.3.3). *Define the piecewise constant extension $\tilde{U}\colon [0,T] \to L^p(\Omega;X)$ by $\tilde{U}(t) := U^j$ for $t \in [t_j,t_{j+1})$, $0 \le j \le N_k - 1$, and $\tilde{U}(T) := U^{N_k}$. Then*

$$\lim_{k\to 0} \Big\| \sup_{t\in[0,T]} \|U(t) - \tilde{U}(t)\|_X \Big\|_{L^p(\Omega)} = 0.$$

Theorems 1.2 and 1.3 follow from Theorems 4.24 and 4.57 in the main body of this thesis, respectively. They hold beyond the Hilbert space setting. More precisely, X and Y can be arbitrary 2-smooth Banach spaces, including, among others, L^p-spaces for $p \in [2,\infty)$. The noise then has to be considered as mapping into the space $\gamma(H,X)$ of γ-radonifying operators (cf. Section 2.5), which coincides with the space of Hilbert–Schmidt operators if X is Hilbert.

Among the schemes covered by Theorems 1.2 and 1.3 are

- exponential Euler (EE): $R_k = S(k)$;
- implicit Euler (IE): $R_k = (1 + kA)^{-1}$;
- Crank–Nicolson (CN): $R_k = (2 - kA)(2 + kA)^{-1}$.

Higher-order implicit Runge–Kutta methods such as Radau methods, BDF(2), Lobatto IIA, IIB, and IIC, and some DIRK schemes are covered as well. Contractivity of the scheme R on X and $\mathrm{D}(A)$ in the case of (EE) and (IE) is an immediate consequence of the contractivity of the semigroup S. For other rational schemes, the contractivity of $R_k = r(kA)$ follows from the holomorphy of the corresponding rational function $r\colon \mathbb{C}_- \to \mathbb{C}$ and $|r(z)| \le 1$ for all $z \in \mathbb{C}_-$, which, in particular, is satisfied for A-acceptable or A-stable schemes. These assertions follow from functional calculus (see Proposition 2.28).

In the above, one usually takes Y to be a suitable intermediate space between X and $\mathrm{D}(A)$. In the special and important case that $Y = \mathrm{D}(A)$, one can take $\alpha = \frac{1}{2}$ for all aforementioned schemes. More general convergence rates can be found in Table 1.1.

	Exponential Euler	Implicit Euler	Crank–Nicolson
α	$\beta \wedge \frac{1}{2}$	$\frac{\beta}{2} \wedge \frac{1}{2}$	$\frac{2\beta}{3} \wedge \frac{1}{2}$

Table 1.1: Convergence rates α in case $Y = \mathrm{D}(A^\beta)$ in Theorem 1.2

Up to the logarithmic factor, the estimate (1.3.6) is optimal in the sense that the rate is the same as the rate for the initial value term on its own (i.e. with $F = 0$ and $G = 0$). Theorem 1.2 follows from Theorem 4.24. In the case of exponential Euler, we show that the logarithmic factor can be omitted, see Corollary 4.26. In the case of additive noise, a similar result is obtained in Theorem 4.1 for the range $\alpha \in (0,1]$ for semigroups and schemes which are not necessarily contractive.

The error estimate (1.3.6) can be extended from the grid points to the full time interval $[0,T]$ assuming higher integrability of the initial values. Provided that $u_0 \in L^{p_0}(\Omega;Y)$ holds for some $p_0 \in (2,\infty)$ in addition to the assumptions of Theorem 1.2, the pathwise uniform error on the full time interval can be estimated as (see Theorem 4.32 below)

$$\Big\| \sup_{t\in[0,T]} \|U(t) - \tilde{U}(t)\|_X \Big\|_{L^p(\Omega)} \leq C_T(1 + \|u_0\|_{L^{p_0}(\Omega;Y)}) k^\alpha \sqrt{\log(T/k)} \tag{1.3.7}$$

for all $p \in [2, p_0)$ and the piecewise constant extension $\tilde{U}$ of $(U_j)_{j=0,\dots,N_k}$ to $[0,T]$. This rate is known to be optimal already for scalar SDEs, implying optimal convergence rates on the full time interval $[0,T]$. In practice, this implies that the rate of convergence in the grid points is maintained already for a piecewise constant interpolation to other times. The error estimate relies on new optimal path regularity estimates of stochastic convolutions in suitable log-Hölder spaces, which will be presented in Proposition 4.31.

Applications to Schrödinger and Maxwell equations are included in the main text (see Subsections 4.1.3, 4.4.4, and 4.4.5). Our results improve several results from the literature to more general schemes and general rates α. In Section 4.5, we include a setting for abstract wave equations, which was considered in [132] only for the exponential Euler method. We prove similar higher-order convergence rates for more general schemes and, in particular, recover [132] as a special case. A Schrödinger equation with irregular nonlinearities for which previous results [2, Thm. 4.3] or Theorem 1.2 are not applicable is investigated in Subsection 4.6.2 and convergence is shown.

Let us emphasise that schemes involving rational approximations, such as implicit Euler or Crank–Nicolson, are in the focus of our work. While we improve existing results for exponential Euler, the main novelty of our work lies in the possibility to treat other schemes with a semigroup approach. To the best of the author's knowledge, the present work is the first contribution to pathwise uniform convergence (rates) for hyperbolic problems from a theoretical point of view, both in the generality and for the concrete examples listed above. The main innovations are:

- first optimal pathwise uniform convergence rates for implicit Euler, Crank–Nicolson, and any other contractive time discretisation scheme for hyperbolic SPDEs
- first use of Kato's framework for SPDEs to systematically treat hyperbolic problems
- maximal estimates for the convergence rate rather than pointwise estimates
- path regularity results allowing to consider the error on the full time interval
- novel pathwise uniform stability estimates
- convergence up to order 1 for abstract wave equations for any contractive scheme
- pathwise uniform convergence for irregular nonlinearities and noise as well as rough initial values which do not fall into the Kato setting

It was recently shown in [32] that one can transfer (1.3.2) to (1.3.4) using some of the Hölder continuity in the p-th moment at the price of decreasing the convergence rate via the Kolmogorov–Chentsov theorem. The strength of this lies in the generality of possible applications. However, to get practically useful bounds in concrete cases, there are limitations. A more detailed comparison is made in Remark 4.25.

To make the above results applicable to implementable numerical schemes for SPDEs, one would additionally need a space discretisation. Since the main novelty of our work for SPDEs lies in the treatment of temporal discretisations, we will only consider the latter in Chapter 4. Space discretisation for SPDEs is usually performed by means of spectral Galerkin methods [75, 78, 81, 133], finite differences [4, 31, 59], or finite elements [3, 30, 91, 92, 93, 96, 97], sometimes combined with a discontinuous Galerkin approach [11, 70], or other methods in space or space-time [12, 37, 38, 42, 62, 100].

A detailed understanding of the globally Lipschitz setting is a quintessential step towards the treatment of locally Lipschitz nonlinearities, which occur more frequently in practice. Our result should be seen as a first step towards a convergence analysis in the locally Lipschitz setting.

1.3.3 Method of proof

For the proof of the convergence rate, we need several ingredients. Firstly, we need to prove that the mild solution actually is continuous with values in the subspace Y. This can be seen as the replacement of the usual regularisation one has for parabolic equations in spirit of the Kato setting explained before. Surprisingly, we do not need any Lipschitz assumptions on F and G as mappings from Y to Y, but linear growth conditions suffice. This is crucial since Lipschitz estimates typically fail for Nemytskij mappings on Sobolev spaces of higher order (see [40] and Remark 4.15).

A key estimate in the proof is a new maximal inequality for discrete convolutions. In particular, this inequality will be used to prove the stability of schemes such as (1.3.3), i.e.,

$$\mathbb{E} \sup_{j\in\{0,\dots,N_k\}} \|U^j\|_Y^p \le C,$$

where C is independent of the step size k. But it also plays a role in further estimates for the convergence.

A second key ingredient is an improved version of an estimate recently proven in [128], which allows estimating stochastic integral processes that contain a supremum

$$\mathbb{E} \sup_{i\in\{1,\dots,n\}} \sup_{t\ge 0} \Big\| \int_0^t \Phi_i(s)dW_H(s)\Big\|_X^p \tag{1.3.8}$$

by certain square functions with a square-root-logarithmic dependency on n (see Proposition 2.62 below).

Finally, to prove the desired convergence rate of Theorem 1.2 we need to split the error obtained in (1.3.3) into

1 (initial value part) + 4 (deterministic terms) + 5 (stochastic terms) = 10 terms.

To estimate these terms we require precise estimates for $\|S(t_j) - R_k^j\|_{\mathcal{L}(Y,X)}$, $\mathbb{E}\|U(t) - U(s)\|_X^p$, stability estimates, and maximal estimates for continuous and discrete convolutions. Stability of the scheme can be obtained by a dilation argument in Hilbert spaces X and Y. This is no longer possible in general 2-smooth Banach spaces. Instead, the key ingredient of the stability proof is a novel maximal inequality for discrete convolutions based on a martingale argument. Lemma 4.18 illustrates how martingale (difference) sequences are used in this argument, resulting in stability as stated in Proposition 4.19.

In the end, we derive an estimate for the error in terms of itself, and we apply a standard discrete Gronwall argument to deduce the desired error bound. In the case of exponential Euler, some terms disappear since $S(t_j) = R_k^j$, which makes it possible to omit the logarithmic terms originating from terms such as (1.3.8).

Passing to irregular nonlinearities and noise, linear growth on more regular subspaces of X no longer holds true and the initial values are not pathwise more regular. To circumvent this problem, we regularise the nonlinearity, the noise, and the initial values occurring in (1.3.1) according to

$$ {}_mF := mR(m,-A)F, \quad {}_mG := mR(m,-A)G, \quad {}_mu_0 := mR(m,-A)u_0 $$

for some regularisation parameter $m \in \mathbb{N}$, where $R(m,-A) := (m+A)^{-1}$ denotes the resolvent. By construction, ${}_mF$ maps to $\mathrm{D}(A)$, ${}_mG$ maps to $\mathcal{L}_2(H,\mathrm{D}(A))$, and ${}_mu_0 \in L^p(\Omega;\mathrm{D}(A))$, giving the desired additional regularity in structure with the more regular space being $Y = \mathrm{D}(A)$. Hence, this enables us to apply the quantified convergence results of Theorem 1.2 for the regularised discretisation given by

$$ {}_mU^j := R_k({}_mU^{j-1}) + kR_k({}_mF({}_mU^{j-1})) + R_k({}_mG({}_mU^{j-1}))\Delta W_j, \quad {}_mU^0 := {}_mu_0, \quad (1.3.9) $$

for $1 \le j \le N_k$. They approximate the mild solution ${}_mU$ of the regularised evolution equation

$$ {}_mU = (-A_mU + {}_mF({}_mU))\,\mathrm{d}t + {}_mG({}_mU)\,\mathrm{d}W_H(t), \quad {}_mU(0) = {}_mu_0 \in X. \qquad (1.3.10) $$

Theorem 1.2 yields convergence of the pathwise uniform discretisation error (1.3.4) of the regularised problem as $k \to 0$. Note that this convergence is not uniform in the regularisation parameter $m \in \mathbb{N}$. The main task consists in proving convergence of the regularisation error, both for the mild solutions to (1.3.10) and for the discretisations (1.3.9), as $m \to \infty$ uniformly in the number of time steps N_k. For the continuous regularisation error, this is achieved by a combination of a maximal inequality for stochastic convolutions, continuity of paths of the nonlinearities evaluated at the mild solution, and a classical continuous Gronwall argument.

An analogous straightforward estimate of the discrete regularisation error fails. Instead, the maximal inequality for discrete convolutions already used in the stability proof in 2-smooth Banach spaces proves helpful. In addition, a clever splitting of the error is required so that it can be rewritten in terms of the continuous regularisation error, the discretisation error of the regularised problem, as well as the full error, i.e., the discretisation error of the original problem (1.3.1). The regularisation parameter $m \in \mathbb{N}$ can then be fixed large enough such that the first error becomes small, and we already showed uniform convergence of the second as $k \to 0$. Finally, we thus derive an estimate for the full error in terms of itself, and we apply a standard discrete Gronwall argument, finally resulting in the desired convergence statement for irregular nonlinearities.

1.4 Overview

Let us outline the rest of the thesis. Chapter 2 contains the preliminaries for the rest of the thesis. Evolution equations with random coefficients are investigated in Chapter 3, corresponding to Section 1.2 of this introduction. The results presented in this chapter are contained in the publication [85] by the author of the thesis. Chapter 4 is devoted to the temporal approximation of stochastic evolution equations, as outlined in Section 1.3. The results stated in Sections 4.1-4.5 and Section 4.6 can be found in the author's publications [86] and [87], respectively. Please note that, contrary to [86], all results are formulated in 2-smooth Banach spaces using the extension discussed in [87, Section 3]. Moreover, a complete proof of the well-posedness result in Section 4.2 has been added in this thesis.

We give a more detailed overview of the content of the thesis, highlighting novel results obtained in the respective sections.

- The analysis of evolution equations with random coefficients commences with a formal introduction of the random evolution equations considered and a reminder of properties of the corresponding semigroups in Section 3.1.

- Section 3.2 reviews and collects the facts needed for the three discretisations performed in space, time, and randomness separately. In space, the focus lies on the approximation of form-induced semigroups and includes a novel quantified version of the Trotter–Kato theorem for form-induced semigroups in Theorem 3.8. Corollary 3.36 contains an extension of PCE error estimates from the one-dimensional scalar-valued to the multivariate vector-valued case.

- Results on joint convergence rates for space-time discretisation in the deterministic case are recalled from the literature in Section 3.3.

- The main part of Chapter 3 is Section 3.4, where we prove a joint convergence rate for the full discretisation (see Theorems 3.74 and 3.75).

- An application to a parabolic equation with random anisotropic diffusion is contained in Section 3.5.

- Section 4.1 discusses temporal approximation of SPDEs with *additive noise* and semigroups that are not necessarily contractive. We prove convergence at rate α up to order one, in case the noise and data are regular enough. This is proved under the assumption that the numerical scheme R_k approximates the semigroup at rate α. Results are illustrated for the Schrödinger equation, for which the obtained results improve several bounds from the literature for the exponential Euler scheme, and provide the first uniform bounds for a large class of other numerical methods including implicit Euler and Crank–Nicolson.

- In Section 4.2, we introduce the nonlinear evolution equation with *multiplicative noise* that we consider in the rest of the thesis. After providing a detailed proof of a standard well-posedness result, we introduce a special case of the Kato setting and prove that the solution has regularity in the subspace Y in case of linear growth in the Y-setting (see Theorem 4.14). This novel well-posedness result holds without assuming Lipschitz continuity on Y.

- Section 4.3 is concerned with the stability of the discretisation schemes for the nonlinear evolution equation introduced in Section 4.2. The main stability result can be found in Proposition 4.19 and only requires linear growth. Hence, it is applicable on both X and Y. A simpler proof in Hilbert spaces resulting in better constants is given in Proposition 4.16.

- Section 4.4 is central in the thesis, and here we prove the main result stated in Theorem 1.2 on convergence rates for the pathwise uniform error in time for the nonlinear evolution equation introduced in Section 4.2 (see Theorem 4.24 for the extended version). Moreover, we prove the error bound (1.3.7) on the full time interval in Theorem 4.32. For this, we first establish a new optimal path regularity result for the solution in Proposition 4.31, which is of independent interest. In Subsections 4.4.4 and 4.4.5, we present applications to the Schrödinger equation as well as the Maxwell equation.

- In Section 4.5, we consider abstract stochastic wave equations, and obtain convergence rates up to order one (see Theorem 4.48). Although we are not in the setting of Section 4.4, an inspection of the proofs given there shows that certain terms behave better for abstract wave equations by virtue of their second-order nature. Again, convergence rates are obtained for a large class of numerical schemes, and versions of (1.3.7) are obtained. Examples with trace class, space-time white noise, and smooth noise are included and can be found in Subsections 4.5.4, 4.5.5, and 4.5.6, respectively. All these results are new for schemes different from exponential Euler. Most notably, for smooth noise, we can explain the numerical convergence rates one sees in [132, Fig. 6.1] for implicit Euler and Crank–Nicolson.

- Section 4.6 contains the second main result for SPDEs: We state and prove convergence for irregular Lipschitz nonlinearity and noise as well as rough initial data, which leads to Theorem 1.3. The result is illustrated for a version of the nonlinear stochastic Schrödinger equation that does not fall into the Y-setting.

Chapter 2

Preliminaries

In this chapter, we recall fundamental results from the different areas of mathematics required in order to follow the subsequent chapters: functional analysis, more specifically semigroup and interpolation theory, numerical analysis, and stochastic analysis. The reader interested in the approximation of evolution equations with random coefficients presented in Chapter 3 is advised to read Sections 2.1, 2.2, and 2.3. Sections 2.4 - 2.10 only become relevant for the analysis of stochastic PDEs in Chapter 4.

Semigroups, generators, resolvents, and relations inbetween them are introduced in Section 2.1 alongside evolution equations and forms. Approximation of semigroups both in space and time is reviewed in Section 2.2, before some concepts from interpolation theory are recalled in Section 2.3. Section 2.4 introduces a special class of Banach spaces, the 2-smooth Banach spaces, on which a generalisation of Hilbert–Schmidt operators can be defined, the so-called γ-radonifying operators studied in Section 2.5. These operators become relevant when extending the stochastic integral introduced in Section 2.6 alongside other fundamental notions from stochastic analysis to the realm of 2-smooth Banach spaces in Section 2.7. A useful martingale inequality from stochastic analysis is discussed in Section 2.8. We introduce Besov–Orlicz spaces in Section 2.9 and give estimates for stochastic integrals in these spaces before finishing the preliminaries with different versions of Gronwall's lemma in Section 2.10.

2.1 Fundamentals of Semigroup Theory

The mathematical centrepiece of this thesis are C_0-semigroups. They provide a particularly useful functional analytic tool to analyse partial differential equations. We review central concepts such as the relation to infinitesimal generators and resolvents, generation theorems, and how to solve evolution equations using semigroup theory, based on one of the standard textbooks in the field [45].

Definition 2.1. A *(one-parameter) semigroup* on X is a family $(S(t))_{t\geq 0}$ of bounded linear operators on a Banach space X satisfying the functional equation $S(t+s) = S(t)S(s)$ for all $t, s \geq 0$ and $S(0) = I$. It is a *strongly continuous semigroup* or *C_0-semigroup* if $S(t)f \to f$ as $t \searrow 0$ for all $f \in X$. If the functional equation holds for all $t, s \in \mathbb{R}$ and strong continuity holds as $t \to 0$ then $S\colon \mathbb{R} \to \mathcal{L}(X)$ is a *C_0-group*.

Henceforth, all semigroups studied are strongly continuous ones. Consequently, we use the notions "semigroup" and "C_0-semigroup" equivalently in the following.

Every semigroup is exponentially bounded. Different types of semigroups are classified by the value of the exponential boundedness constants.

Proposition 2.2 (Proposition I.5.5 in [45])**.** *For every C_0-semigroup $(S(t))_{t\geq 0}$, there exist constants $w \in \mathbb{R}$, $M \geq 1$ such that $\|S(t)\| \leq M\mathrm{e}^{wt}$ for all $t \geq 0$.*

Definition 2.3. For a given semigroup $(S(t))_{t\geq 0}$, let $w \in \mathbb{R}$ and $M \geq 1$ be the constants from Proposition 2.2. Then (M, w) is called the *type* of S or, equivalently, S is said to be *of type* (M, w). It is called *quasi-contractive* if it is of type $(1, w)$ for some $w \in \mathbb{R}$ and *contractive* or *contraction semigroup* if it is of type $(1, 0)$, i.e., $\|S(t)\| \leq 1$ for all $t \geq 0$.

Another central building block of semigroup theory consists of the generator of a semigroup. It is obtained by differentiating the *orbit maps* $\xi_x\colon [0,\infty) \to X$, $t \mapsto S(t)x$ for those $x \in X$ for which ξ_x is differentiable.

Definition 2.4. The *(infinitesimal) generator* $-A\colon \mathrm{D}(A) \subseteq X \to X$ of a strongly continuous semigroup $(S(t))_{t\geq 0}$ on a Banach space X is the (unbounded) operator

$$-Ax := \lim_{t\searrow 0} \frac{S(t)x - x}{t}$$

with the *domain* $\mathrm{D}(A)$ consisting of all $x \in X$ for which the limit exists.

By definition, $-A$ is a linear operator. Moreover, it is closed, densely defined, and determines the semigroup uniquely. If A is bounded, the corresponding semigroup is given by the matrix exponential $S(t) = \mathrm{e}^{-tA}$ for $t \geq 0$. C_0-semigroups can be interpreted as an extension of the exponential function to unbounded operators. The following lemma establishes a relation between semigroup and generator, reminiscent of the fundamental theorem of calculus.

Lemma 2.5 (Lemma II.1.3 in [45])**.** *Let $-A$ be the generator of a C_0-semigroup $(S(t))_{t\geq 0}$. If $x \in \mathrm{D}(A)$, then $S(t)x \in \mathrm{D}(A)$ and $\frac{\mathrm{d}}{\mathrm{d}t}S(t)x = -AS(t)x = -S(t)Ax$ for all $t \geq 0$. Moreover, for every $x \in X$ and $t \geq 0$, we have $\int_0^t S(s)x\,\mathrm{d}s \in \mathrm{D}(A)$. Furthermore, for every $t \geq 0$, we have*

$$S(t)x - x = -A\int_0^t S(s)x\,\mathrm{d}s = -\int_0^t S(s)Ax\,\mathrm{d}s,$$

where the first equality holds for all $x \in X$ and the second one only for $x \in \mathrm{D}(A)$.

As, for some instances, it is difficult to determine the domain of the generator precisely, one works with (often more regular) subspaces that are "large enough" in the following sense.

Definition 2.6. A subspace $D \subseteq X$ is called a *core* for $-A$ if $\overline{-A|_D} = -A$. Equivalently, D is a core for $-A$ if $D \subseteq \mathrm{D}(A)$ is dense for the *graph norm* defined via $\|x\|_{\mathrm{D}(A)}^2 := \|x\|_X^2 + \|Ax\|_X^2$ for $x \in \mathrm{D}(A)$.

Typical examples of cores for $-A$ are given by $\mathrm{D}(A^\ell)$ for $\ell \in \mathbb{N}$.

The third central element besides semigroups and generators is given by resolvents. The *resolvent of* A *in* $\lambda \in \mathbb{C}$ is given by $R(\lambda, A) := (\lambda - A)^{-1}$. It is defined for all λ in the *resolvent set* $\rho(A)$, which contains all $\lambda \in \mathbb{C}$ such that $R(\lambda, A) \in \mathcal{L}(X)$, i.e., such that $\lambda - A$ is boundedly invertible. A simple calculation shows the following useful identities for resolvents.

Proposition 2.7 (First and second resolvent identity)**.** *Let A, B be closed operators on a Banach space X. For all $\lambda, \mu \in \rho(A)$, we then have*

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda)R(\lambda, A)R(\mu, A)$$

and for all $\lambda \in \rho(A) \cap \rho(B)$,

$$R(\lambda, A) - R(\lambda, B) = R(\lambda, A)(A - B)R(\lambda, B).$$

Since we will frequently use fractional powers of the negative generator, we want the negative generator to be positive, and thus let $-A$ be the generator throughout. Consequently, we are considering the resolvents $R(\lambda, -A) = (\lambda + A)^{-1}$. The resolvent can be directly obtained from the semigroup via the *Laplace transform* as illustrated in part (a) of the following theorem.

Theorem 2.8 (Theorem II.1.10 in [45])**.** *Let $(S(t))_{t\geq 0}$ be a C_0-semigroup of type (M, w) on a Banach space X. For the generator $-A$ of $(S(t))_{t\geq 0}$, the following properties hold.*

(a) If there is $\lambda \in \mathbb{C}$ such that $R(\lambda)x := \int_0^\infty \mathrm{e}^{-\lambda s} S(s)x \,\mathrm{d}s$ exists for all $x \in X$, then $\lambda \in \rho(-A)$ and $R(\lambda) = R(\lambda, -A)$.

(b) If $\operatorname{Re}\lambda > w$, then $\lambda \in \rho(-A)$ and the resolvent is given by the integral in (a).

(c) $\|R(\lambda, -A)\| \leq \frac{M}{\operatorname{Re}\lambda - w}$ for all $\operatorname{Re}\lambda > w$.

The integral representation of the resolvent $R(\lambda, -A)$ in (a) has to be understood as an improper Riemann integral in the strong sense.

Provided that $(\lambda R(\lambda, -A))_{\lambda \geq w}$ is bounded (which does not follow from Part (c) in general), strong convergence to the identity is obtained as $\lambda \to \infty$.

Lemma 2.9 (Lemma II.3.4 in [45])**.** *Let $A\colon \mathrm{D}(A) \subseteq X \to X$ be a closed, densely defined operator on a Banach space X. Suppose there exist $w \in \mathbb{R}$ and $M > 0$ such that $[w, \infty) \subseteq \rho(-A)$ and $\|\lambda R(\lambda, -A)\| \leq M$ for all $\lambda \geq w$. Then the following convergence statements hold for $\lambda \to \infty$.*

(a) $\lambda R(\lambda, -A)x \to x$ *for all* $x \in X$.

(b) $\lambda AR(\lambda, -A)x = \lambda R(\lambda, -A)Ax \to Ax$ *for all* $x \in \mathrm{D}(A)$.

Knowing how to pass from the semigroup to the generator and the resolvent, naturally, the question arises of when a given unbounded operator generates a semigroup. The answer is given by the following celebrated generation theorem due to Hille and Yosida. We state the version for semigroups of arbitrary type due to Feller, Miyadera, and Phillips rather than the original version formulated for contraction semigroups.

Theorem 2.10 (Hille–Yosida generation theorem, Theorem II.3.8 in [45])**.** *Let* $A\colon \mathrm{D}(A) \subseteq X \to X$ *be a linear operator on a Banach space* X *and let* $w \in \mathbb{R}, M \geq 1$. *Then the following properties are equivalent.*

(i) $-A$ *generates a* C_0*-semigroup* $(S(t))_{t\geq 0}$ *of type* (M, w).

(ii) A *is closed, densely defined, and for every* $\lambda > w$, *we have* $\lambda \in \rho(-A)$ *and for all* $n \in \mathbb{N}$ *it holds that* $\|[(\lambda - w)R(\lambda, -A)]^n\| \leq M$.

(iii) A *is closed, densely defined, and for every* $\lambda \in \mathbb{C}$ *with* $\operatorname{Re}\lambda > w$, *we have* $\lambda \in \rho(-A)$ *and* $\|R(\lambda, -A)^n\| \leq M(\operatorname{Re}\lambda - w)^{-n}$ *for all* $n \in \mathbb{N}$.

In practice, verifying the bounds for the n-th power of the generator is often not feasible due to a lacking explicit formula for powers of the resolvent. This challenge is overcome by applying the generation theorem of Lumer and Phillips instead. This requires the notion of dissipativity, which is easier to verify for a given operator.

Definition 2.11. We call a linear operator $A\colon \mathrm{D}(A) \subseteq X \to X$ on a Banach space X *dissipative* if $\|(\lambda - A)x\| \geq \lambda\|x\|$ for all $\lambda > 0$ and $x \in \mathrm{D}(A)$.

Theorem 2.12 (Lumer–Phillips generation theorem, Theorem II.3.15 in [45])**.** *For a densely defined, dissipative operator* $-A\colon \mathrm{D}(A) \subseteq X \to X$ *on a Banach space* X *the following statements are equivalent.*

(i) The closure $\overline{-A}$ *of* $-A$ *generates a contraction semigroup.*

(ii) The range of $\lambda + A$ *is dense in* X *for some (hence all)* $\lambda > 0$.

Via the *inverse Laplace transform*, also known as *complex inversion formula*, the semigroup can be obtained directly from the resolvent on the domain of the generator.

Theorem 2.13 (Corollary III.5.15 in [45])**.** *Let* $(S(t))_{t\geq 0}$ *be a* C_0*-semigroup on a Banach space* X. *Then*

$$S(t)x = \lim_{n\to\infty} \frac{1}{2\pi \mathrm{i}} \int_{w-\mathrm{i}n}^{w+\mathrm{i}n} \mathrm{e}^{t\lambda} R(\lambda, -A)x \,\mathrm{d}\lambda \quad (x \in \mathrm{D}(A), t \geq 0, w > \omega_0(S)),$$

where the convergence is uniform on compact time intervals. Here, $\omega_0(S)$ *denotes the* growth bound *of* S *defined as the infimum over all* $w \in \mathbb{R}$ *such that* S *is of type* (M_w, w) *for some* $M_w \geq 1$.

Our main motivation to study semigroups is to solve an important class of partial differential equations (PDEs), namely, *evolution equations*. They describe the evolution of a given quantity over time. The simplest equation to describe an evolution over time is the ordinary differential equation (ODE) $u'(t) = -\alpha u(t)$ for all $t \geq 0$, some $\alpha \in \mathbb{C}$, and $u(0) = u_0 \in \mathbb{C}$. Clearly, the solution to this equation is given by $u(t) = \mathrm{e}^{-\alpha t}u_0$. Keeping this simple equation in mind, we introduce *abstract Cauchy problems*, which can be interpreted as reformulating a PDE as an ODE in another space. The price one has to pay is passing from a finite-dimensional space on which the PDE is formulated to a typically infinite-dimensional space for the ODE. The operator determining the differential equation becomes an unbounded operator, but as in the scalar example above, the solution is determined by the initial values and the exponential function, i.e., the C_0-semigroup in the general (infinite-dimensional) case. Due to the infinite-dimensional nature of the problem, we distinguish between different types of solutions.

Definition 2.14. The *abstract Cauchy problem* associated with $-A$ and the initial value $u_0 \in X$ is given by the initial value problem

$$u'(t) = -Au(t) \quad (t \geq 0) \quad u(0) = u_0 \in X. \tag{ACP}$$

A function $u\colon [0,\infty) \to X$ is called a *classical solution* of (ACP) if u is continuously differentiable, $u(t) \in \mathrm{D}(A)$ for all $t \geq 0$, and (ACP) holds. It is called a *mild solution* of (ACP) if

$$u(t) = S(t)u_0, \quad (t \geq 0).$$

One can show that every classical solution to (ACP) is also a mild solution.

So far, we have considered the *homogeneous* problem (ACP). The semigroup approach also allows us to treat the *inhomogeneous* problem with some time-dependent right-hand side f. More precisely, given a function $f\colon [0,\infty) \to X$ and an initial value $u_0 \in X$, the *inhomogeneous abstract Cauchy problem* consists of finding a function $u\colon [0,\infty) \to X$ that satisfies

$$u'(t) = -Au(t) + f(t) \quad (t \geq 0), \quad u(0) = x. \tag{iACP}$$

Here and in the following, we write $L^p(a,b;X)$ for $L^p((a,b);X)$ with X a Banach space and $a,b \in [0,\infty)$ such that $a < b$.

Definition 2.15. Let X be a Banach space, $u_0 \in X$, and $f \in L^1(0,\infty;X)$. Suppose that $-A$ generates a C_0-semigroup $(S(t))_{t\geq 0}$ on X. Then the function $u\colon [0,\infty) \to X$ defined by

$$u(t) \coloneqq S(t)x + \int_0^t S(t-s)f(s)\,\mathrm{d}s \quad (t \geq 0)$$

is called the *mild solution of* (iACP). If a function $u\colon [0,\infty) \to X$ that fulfils (iACP) is, in addition, continuously differentiable and $u(t) \in \mathrm{D}(A)$ for $t \geq 0$, we call it a *classical solution.*

As for the homogeneous case ($f = 0$), every classical solution of (iACP) is also a mild solution of (iACP), whence classical solutions of (iACP) are unique.

Another useful classical result from semigroup theory will prove useful to prove a stability result in Section 4.3. The following famous dilation theorem [120, Thm. I.4.2] is stated in the version of [115].

Theorem 2.16 (Sz.-Nagy dilation theorem)**.** *Let H be a Hilbert space. Every contraction $S\colon H \to H$ has a unitary dilation $U\colon \tilde{H} \to \tilde{H}$ to a Hilbert space $\tilde{H}$ having H as a subspace. That is, there exist a contractive projection $P\colon \tilde{H} \to H$, a contractive injection $J\colon H \to \tilde{H}$, and a unitary mapping $U\colon \tilde{H} \to \tilde{H}$ such that $S^n = PU^nJ$ for all $n \in \mathbb{N}_0$.*

In particular, choosing the C_0-semigroup at time t as S in the above, this ensures the existence of a C_0-group U on a larger Hilbert space such that $S(nt) = PU(nt)J$.

In Hilbert spaces, an interesting class of generators are those associated with a form. Let V be a Hilbert space. A *form* on V is a sesquilinear mapping $a\colon V \times V \to \mathbb{K}$, and we write $a(u) := a(u, u)$ for the corresponding quadratic form. For a given form a we call $a^*\colon V \times V \to \mathbb{K}$, $a^*(u, v) := \overline{a(v, u)}$ the *adjoint form*. A form a is called *sectorial* if there exists $c \geq 0$ such that $|\operatorname{Im} a(u)| \leq c \operatorname{Re} a(u)$ for all $u \in V$ or, put differently, the *numerical range* $\mathcal{N}(a) := \{a(u) : \ u \in V, \|u\|_V = 1\}$ is contained in the sector $\Sigma_\theta := \{z \in \mathbb{C} \setminus \{0\} : \ |\arg z| \leq \theta\} \cup \{0\}$ with $\theta := \arctan c$. We call the form a *symmetric* if $a^* = a$ and *bounded* if there exists $M \geq 0$ such that $|a(u, v)| \leq M\|u\|_V\|v\|_V$ for all $u, v \in V$. Moreover, a is called *coercive* if there exists $\kappa > 0$ such that $\operatorname{Re} a(u) \geq \kappa\|u\|_V^2$ for all $u \in V$. Note that bounded coercive forms are sectorial of angle less than $\frac{\pi}{2}$. Weakening the coercivity condition slightly results in *accretive* forms, which merely satisfy $\operatorname{Re} a(u) \geq 0$ for all $u \in V$.

Let H be a Hilbert space such that $V \hookrightarrow H$ densely. We can associate an unbounded operator A with $\mathrm{D}(A) \subseteq H$ to a form a acting on V such that $(Au \,|\, v)_H = a(u, v)$ for all $u \in \mathrm{D}(A), v \in V$ under suitable conditions on the form. The construction of A for bounded and coercive forms is explained in detail in Subsection 3.2.1. We call an operator A an *accretive* operator if $(Au \,|\, u)_H \geq 0$ for all $u \in \mathrm{D}(A)$. Clearly, A is accretive if it is associated with an accretive form. For some results to hold, a stronger notion than accretivity is required. An accretive operator A is said to be *m-accretive* or, equivalently, *maximal accretive* if, in addition, $1 \in \rho(-A)$. Operators associated with forms that are accretive, bounded, and closed are m-accretive [110, Prop. 1.22].

We finish the section by recalling a standard result from functional analysis with applications to semigroup theory.

Proposition 2.17 (Proposition A.3 in [45])**.** *Let X be a Banach space. On bounded subsets of $\mathcal{L}(X)$, the following topologies coincide.*

(i) The strong operator topology.

(ii) The topology of pointwise convergence on a dense subset of X.

(iii) The topology of uniform convergence on relatively compact subsets of X.

2.2 Approximation of Semigroups

For many evolution equations of interest, the corresponding semigroup cannot be computed analytically. Hence, approximating the semigroup often provides the only way to compute a solution of some evolution equation. Approximation can be performed in space and in time. We review central definitions and results concerning both in this subsection, starting with spatial approximation.

Let X be a Banach space and $-A$ the generator of a C_0-semigroup $(S(t))_{t\geq 0}$ on X. The idea of spatial approximation is to choose suitable approximations A_m of A such that $-A_m$ also generates a C_0-semigroup $(S_m(t))_{t\geq 0}$ on X for $m \in \mathbb{N}$. The celebrated Trotter–Kato theorem provides a characterisation of the strong convergence of these approximating semigroups, which directly implies the desired convergence of mild solutions. More precisely, it states that convergence of the generators yields strong convergence of the associated semigroups under a uniform exponential boundedness condition of the approximating semigroups. We state it in the version of [45, Thm. III.4.8]. Recall that a subspace $D \subseteq X$ is called a *core* for A if $\overline{A|_D} = A$.

Theorem 2.18 (Trotter–Kato Approximation Theorem)**.** *Let $(S(t))_{t\geq 0}$ and $(S_m(t))_{m\in\mathbb{N}}$ be C_0-semigroups on a Banach space X with generators $-A$ and $-A_m$, respectively. Let D be a core for $-A$. Assume that there are $M \geq 1$ and $w \in \mathbb{R}$ such that for all $m \in \mathbb{N}$,*

$$\|S(t)\|_{\mathcal{L}(X)}, \|S_m(t)\|_{\mathcal{L}(X)} \leq M\mathrm{e}^{wt} \quad (t \geq 0).$$

Then the following are equivalent.

(i) For each $x \in D$, there exists $x_m \in \mathrm{D}(A_m)$ such that $x_m \to x$ and $A_m x_m \to Ax$.

(ii) $R(\lambda, -A_m)x \to R(\lambda, -A)x$ for all $x \in X$ and some/all $\lambda > w$.

(iii) $S_m(t)x \to S(t)x$ for all $x \in X$, uniformly for t in compact intervals.

In particular, these assertions hold true if $D \subseteq \mathrm{D}(A_m)$ for all $m \in \mathbb{N}$ and $A_m x \to Ax$ for all $x \in D$.

Often, the approximating generators A_m are obtained by restricting A to suitably chosen subspaces X_m of X. It is also possible to consider A_m on Banach spaces X_m which are not contained in X, using so-called non-conforming methods. Thus, we suppose that $-A_m$ is generating a semigroup $(S_m(t))_{t\geq 0}$ on some Banach space X_m for $m \in \mathbb{N}$. To quantify the convergence of the approximating semigroups $(S_m(t))_{t\geq 0}$ on X_m to $(S(t))_{\geq 0}$ on X, we need the notion of *spatial convergence rates*. To this end, for $m \in \mathbb{N}$ we consider bounded linear operators $P_m\colon X \to X_m$ and $J_m\colon X_m \to X$ satisfying

(A1) $\exists M_1, M_2 \geq 0\, \forall m \in \mathbb{N} : \|P_m\|_{\mathcal{L}(X,X_m)} \leq M_1, \|J_m\|_{\mathcal{L}(X_m,X)} \leq M_2$,

(A2) $\|J_m P_m x - x\|_X \to 0$ as $m \to \infty$ for all $x \in X$,

(A3) $P_m J_m = I_m$ for all $m \in \mathbb{N}$, where I_m is the identity operator on X_m.

Definition 2.19. Let $Y_{\mathrm{x}} \hookrightarrow X$ be a Banach space and $p_{\mathrm{x}} > 0$.

(a) A space discretisation $(A_m)_{m\in\mathbb{N}}$ of A is said to be *convergent of order* p_{x} on Y_{x} (w.r.t. the norm of X) *for the stationary problem* if there exists a constant $C_{\mathrm{x}} \geq 0$ and a $\lambda_0 \in \rho(-A) \cap \bigcap_{m\in\mathbb{N}} \rho(-A_m)$ such that

$$\|J_m R(\lambda_0, -A_m) P_m u_0 - R(\lambda_0, -A) u_0\|_X \leq C_{\mathrm{x}} \frac{\|u_0\|_{Y_{\mathrm{x}}}}{m^{p_{\mathrm{x}}}}$$

for all $u_0 \in Y_{\mathrm{x}}$ and $m \in \mathbb{N}$.

(b) A space discretisation $(A_m)_{m\in\mathbb{N}}$ of A is said to be *convergent of order* p_x on Y_x (w.r.t. the norm of X) *for the evolution problem* if for all $T \geq 0$ there exists a constant $C_\mathrm{x} = C_\mathrm{x}(T) \geq 0$ such that

$$\|J_m S_m(t) P_m u_0 - S(t)u_0\|_X \leq C_\mathrm{x} \frac{\|u_0\|_{Y_\mathrm{x}}}{m^{p_\mathrm{x}}}$$

for all $t \in [0,T]$, $u_0 \in Y_\mathrm{x}$, and $m \in \mathbb{N}$.

We can now present a quantified version of the Trotter–Kato theorem due to Kappel and Ito [74].

Proposition 2.20 (Propositions 2.2 and 2.3 in [74])**.** *Assume that (A1) and (A3) are satisfied. Assume that there are $M \geq 1$ and $w \in \mathbb{R}$ such that for all $m \in \mathbb{N}$,*

$$\|S(t)\|_{\mathcal{L}(X)}, \|S_m(t)\|_{\mathcal{L}(X_m)} \leq M\mathrm{e}^{wt} \quad (t \geq 0).$$

Further, suppose that for some $\alpha > 0$ the space discretisation $(A_m)_{m\in\mathbb{N}}$ of A is convergent of order p_x on $\mathrm{D}(A^\alpha)$ for the stationary problem. Then for any $T > 0$, there exists a constant $C = C(T,\alpha) \geq 0$ such that

$$\|J_m S_m(t) P_m x - S(t)x\|_X \leq \frac{C}{m^{p_\mathrm{x}}} \|x\|_{\mathrm{D}(A^\beta)} \tag{2.2.1}$$

for all $t \in [0,T]$, $\beta = \alpha + 2$, $x \in \mathrm{D}(A^\beta)$, and $m \in \mathbb{N}$.
If, in addition, the semigroup S is analytic, for any $\varepsilon > 0$, there is $C = C(T,\alpha,\varepsilon) \geq 0$ such that (2.2.1) *holds with $\beta = \alpha + 1 + \varepsilon$.*
Suppose that in addition to analyticity of S, X and X_m are Hilbert spaces for all $m \in \mathbb{N}$, $-A_m$ are self-adjoint bounded operators on X_m, and for some $\lambda_0 \in \rho(-A) \cap \bigcap_{m\in\mathbb{N}} \rho(-A_m)$, we have $\|S_m(t)\|_{\mathcal{L}(X_m)} \leq \mathrm{e}^{\lambda_0 t}$ for all $m \in \mathbb{N}$. Then (2.2.1) *holds with $\beta = \alpha + 1$.*

Lower regularity assumptions ($\beta = \alpha + \frac{1}{2}$) can only be achieved on time intervals of the form $[\frac{1}{\delta}, \delta]$ for fixed $\delta > 0$ due to singularities at $t = 0$.

A commonly used approach to choose the operators $(A_m)_{m\in\mathbb{N}}$ consists of restricting the form associated with A to subspaces and considering the operators associated with the restricted form. This idea lies at the heart of finite element methods. We restrict our considerations to Hilbert spaces X to make sense of the notion of associated forms, and thus write H instead of X. Let $a\colon V \times V \to \mathbb{K}$ be a form on a Hilbert space V densely embedded into H.

Definition 2.21. An *approximating sequence* of V is a sequence $(V_m)_{m\in\mathbb{N}}$ of closed subspaces of V such that $\mathrm{dist}_V(v, V_m) \to 0$ as $m \to \infty$ for all $v \in V$.

Given an approximating sequence $(V_m)_{m\in\mathbb{N}}$ of V and a form $a\colon V \times V \to \mathbb{K}$, we say that the *uniform Banach–Nečas–Babuška condition (BNB)* is satisfied if

$$\exists\, \beta > 0\, \forall\, m \in \mathbb{N}\, \forall\, u \in V_m : \sup_{v\in V_m,\, \|v\|_V = 1} |a(u,v)| \geq \beta \|u\|_V. \tag{BNB}$$

It ensures the convergence of Galerkin approximations obtained by considering the forms $a_m\colon V_m \times V_m \to \mathbb{K}$, $a_m \coloneqq a|_{V_m\times V_m}$; cf. [6, Prop. 2.4]. Clearly, if a is coercive, then (BNB) is trivially satisfied for every approximating sequence $(V_m)_{m\in\mathbb{N}}$ of V.

Since a_m acts on V but the semigroup S to be approximated acts on H, approximating generators A_m have to be constructed on a subspace of H. For $m \in \mathbb{N}$, denote by $H_m \subseteq H$ the closure of V_m in H. Then $V_m \hookrightarrow H_m$, and trivially, V_m is dense in H_m for all $m \in \mathbb{N}$.

Definition 2.22. A *space discretisation* of A associated with a form a consists of an approximating sequence $(V_m)_{m\in\mathbb{N}}$ of V and a sequence $(A_m)_{m\in\mathbb{N}}$ of operators A_m in H_m such that $-A_m$ generates a C_0-semigroup $(S_m(t))_{t\geq 0}$ on H_m for $m \in \mathbb{N}$. If the approximating sequence used is clear from the context, we also refer to $(A_m)_{m\in\mathbb{N}}$ as a space discretisation of A.

The detailed construction of a space discretisation from the approximating forms a_m is contained in Subsection 3.2.1, alongside another quantified version of the Trotter–Kato theorem with conditions on a and $(V_m)_{m\in\mathbb{N}}$ directly instead of the resolvents as in Proposition 2.20.

Another fundamental part of the approximation of solutions to a (deterministic or stochastic) evolution equation entails the temporal approximation of a semigroup by a time discretisation method. Let X be a Banach space and $(S(t))_{t\geq 0}$ a C_0-semigroup on X with generator $-A$.

Definition 2.23. A *time discretisation method* for $(S(t))_{t\geq 0}$ is a function $F\colon [0,\infty) \to \mathcal{L}(X)$ satisfying $F(0) = I$. It is also referred to as *($\mathcal{L}(X)$-valued) scheme* or *($\mathcal{L}(X)$-valued) time discretisation scheme.*

Let $F\colon [0,\infty) \to \mathcal{L}(X)$ be a time discretisation method for $(S(t))_{t\geq 0}$. Then F is called *contractive* if $\|F(\tau)\|_{\mathcal{L}(X)} \leq 1$ for all $\tau \geq 0$ and *stable* if there exist $M \geq 1$ and $\lambda \in \mathbb{R}$ such that for all $N \in \mathbb{N}$ and $(\tau^i)_{1\leq i\leq N}$ in $[0,\infty)$, we have

$$\left\| \prod_{i=1}^{N} F(\tau^i) \right\|_{\mathcal{L}(X)} \leq M \mathrm{e}^{\lambda \sum_{i=1}^{N} \tau^i}.$$

We say that F is *consistent* if there exists a core $D \subseteq X$ for $-A$ such that

$$\lim_{\tau\to 0} \frac{1}{\tau}(F(\tau)f - f) = -Af \quad (f \in D).$$

Moreover, F is called *convergent* if for all $t \geq 0$, $N \in \mathbb{N}$, and all $(\tau_N^i)_{1\leq i\leq N}$ such that $\tau_N^i = \tau_N^i(t,N) \in [0,t]$ for all $1 \leq i \leq N$ and $\max_{i=1,\dots,N} \tau_N^i \to 0$ as well as $\sum_{i=1}^{N} \tau_N^i \to t$ as $N \to \infty$, we have

$$\lim_{N\to\infty} \prod_{i=1}^{N} F\big(\tau_N^i\big) f = S(t)f \quad (f \in X).$$

The close connection between the concepts of consistency, stability, and convergence is elaborated in Subsection 3.2.2. The following definition allows us to quantify the temporal approximation behaviour.

Definition 2.24. Let $F\colon [0,\infty) \to \mathcal{L}(X)$ be a time discretisation method. Let $Y_{\mathrm{t}} \hookrightarrow X$ with $Y_{\mathrm{t}} \subseteq \mathrm{D}(A)$ be a core for $-A$, and $p_{\mathrm{t}} > 0$. Then F is called *consistent of order* p_{t} on Y_{t} if there exist $C \geq 0$ and $\tau_0 > 0$ such that for all $\tau \in [0,\tau_0]$, we have

$$\|F(\tau)f - S(\tau)f\|_X \leq C\tau^{p_{\mathrm{t}}+1} \|f\|_{Y_{\mathrm{t}}} \quad (f \in Y_{\mathrm{t}}).$$

Moreover, F is called *convergent of order* p_{t} on Y_{t} if for all $T \geq 0$ there exist $C_T \geq 0$ and $\tau_0 = \tau_0(T) \in (0,T]$ such that for all $t \in [0,T]$, $N \in \mathbb{N}$, and $(\tau^i)_{1\leq i\leq N}$ with $\tau^i = \tau^i(t,N) \in$

$[0, \tau_0]$ for $1 \leq i \leq N$ and $\sum_{i=1}^N \tau^i = t$, we have

$$\left\| \prod_{i=1}^N F(\tau^i)f - S(t)f \right\|_X \leq C_T \Big(\max_{i=1,\dots,N} \tau^i \Big)^{p_t} \|f\|_{Y_t} \quad (f \in Y_t).$$

Equivalently, we say that F *approximates* S *to order* p_t *on* Y_t or that F *converges of order* p_t *on* Y_t.

In the case of uniform time steps $\tau^i = \tau > 0$ for $1 \leq i \leq N$ and $N\tau \leq T$, convergence of F of order $p_t > 0$ on Y_t means

$$\|(F(\tau)^j - S(j\tau))f\|_X \leq C_T \tau^{p_t} \|f\|_{Y_t}$$

for all $1 \leq j \leq N$ and f, T, C_T as above. Subsequently, we will omit the index for norms in the space X.

Remark 2.25. Let $Y_t \hookrightarrow X$ with Y_t dense in X such that $Y_t \subseteq \mathrm{D}(A)$ and Y_t is invariant under $(S(t))_{t\geq 0}$. Then Y_t is a core for $-A$; cf. [45, Prop. II.1.7]. Note that every core for $-A$ is dense in X.

Furthermore, consistency and convergence of a certain positive order imply consistency and convergence as in Definition 2.23, respectively, where stability is needed for convergence. We include a proof of the respective statements.

Firstly, if $F\colon [0,\infty) \to \mathcal{L}(X)$ is a consistent time discretisation method of order $p_t > 0$ on Y_t then F is consistent. Indeed, for $f \in Y_t$ and $\tau > 0$, we have

$$\frac{1}{\tau}(F(\tau)f - f) + Af = \frac{1}{\tau}(F(\tau)f - S(\tau)f) + \frac{1}{\tau}(S(\tau)f - f) + Af,$$

and therefore for sufficiently small $\tau > 0$, we observe

$$\left\| \frac{1}{\tau}(F(\tau)f - f) + Af \right\| \leq C\tau^{p_t} \|f\|_{Y_t} + \left\| \frac{1}{\tau}(S(\tau)f - f) + Af \right\| \to 0.$$

Moreover, if F is stable and convergent of order $p_t > 0$ on Y_t then F is convergent. Indeed, let $t \geq 0$, $N \in \mathbb{N}$, and $(\tau_N^i)_{1\leq i\leq N}$ with $\tau_N^i \in [0,t]$ as well as $\max_{i=1,\dots,N} \tau_N^i \to 0$ and $\sum_{i=1}^N \tau_N^i \to t$ as $N \to \infty$, and $f \in Y_t$. Then

$$\prod_{i=1}^N F(\tau_N^i)f - S(t)f = \left(\prod_{i=1}^N F(\tau_N^i)f - S\Big(\sum_{i=1}^N \tau_N^i\Big)f \right) + \left(S\Big(\sum_{i=1}^N \tau_N^i\Big)f - S(t)f \right),$$

and therefore $\prod_{i=1}^N F(\tau_N^i)f - S(t)f \to 0$ as $N \to \infty$ by convergence of order p_t on Y_t and strong continuity of $(S(t))_{t\geq 0}$. Furthermore, $\big(\prod_{i=1}^N F(\tau_N^i) - S(t)\big)_{N\in\mathbb{N}}$ is uniformly bounded by stability of F and exponential boundedness of $(S(t))_{t\geq 0}$. Thus, F is convergent.

Example 2.26. *We review some classical examples of time discretisation methods, which approximate the semigroup* $(S(t))_{t\geq 0}$ *generated by* $-A$ *to different orders.*

(a) *Clearly, the trivial choice* $F = S$ *yields a stable and consistent and therefore also convergent time discretisation method of arbitrary order* $p_t > 0$ *on* X*. In the context of nonlinear evolution equations, this is referred to as the* exponential Euler method.

(b) *Let* $\lambda \geq 0$ *such that* $[\lambda, \infty) \subseteq \rho(-A)$*, and define* $F\colon [0,\infty) \to \mathcal{L}(X)$ *by*

$$F(\tau) := \begin{cases} I, & \tau = 0, \\ (I + \tau A)^{-1} = \frac{1}{\tau} R\big(\frac{1}{\tau}, -A\big), & \tau \in (0, \frac{1}{\lambda}), \\ \big(I + \frac{1}{\lambda} A\big)^{-1}, & \tau \geq \frac{1}{\lambda}. \end{cases}$$

Then F *is called* implicit Euler method (IE) *and is stable and consistent of order* 1 *on* $\mathrm{D}(A^2)$*. More generally, it converges of order* $p_{\mathrm{t}} \in (0,1]$ *on* $\mathrm{D}(A^{2p_{\mathrm{t}}})$ *if* $(S(t))_{t\geq 0}$ *is bounded, see [55, Thm. 1.3] or [90, Cor. 4.4]. If, additionally,* S *is analytic, the convergence order* $p_{\mathrm{t}} \in (0,1]$ *is attained already on* $\mathrm{D}(A^{p_{\mathrm{t}}})$*. If* $-A$ *generates a contractive* C_0*-semigroup, then* $\lambda = 0$ *in the above and* F *is contractive.*

(c) *Let* $\lambda \geq 0$ *such that* $[\lambda, \infty) \subseteq \rho(-A)$*, and define* $F\colon [0,\infty) \to \mathcal{L}(X)$ *by*

$$F(\tau) := \begin{cases} I, & \tau = 0, \\ \big(I - \frac{\tau}{2} A\big)\big(I + \frac{\tau}{2} A\big)^{-1} = (2I - \tau A)(2I + \tau A)^{-1}, & \tau \in (0, \frac{2}{\lambda}), \\ \big(I - \frac{1}{\lambda} A\big)\big(I + \frac{1}{\lambda} A\big)^{-1}, & \tau \geq \frac{2}{\lambda}. \end{cases}$$

Then F *is called* Crank–Nicolson scheme (CN) *and is consistent, but may not always be stable, as the left shift semigroup on* $C_0(\mathbb{R})$ *illustrates. Provided that* F *is stable, it converges of order* $p_{\mathrm{t}} \in (0,2]$ *on* $\mathrm{D}(A^{3p_{\mathrm{t}}/2})$ *for bounded semigroups [90, Cor. 4.4]. If, moreover,* S *is analytic, Crank–Nicolson converges of order* $p_{\mathrm{t}} \in (0,2]$ *already on* $\mathrm{D}(A^{p_{\mathrm{t}}})$*. In the absence of stability, the convergence order is lower; it is given by* p_{t} *on* $\mathrm{D}(A^{\beta})$*, where* (β, p_{t}) *is on the graph of the piecewise linear interpolation of the points* $(\frac{1}{2}, 0), (1, \frac{1}{2}), (2, \frac{4}{3})$*, and* $(3,2)$*; cf. [90, Thm. 1.1 and 4.1] and [128, Ex. 5.10].*

Implicit Euler and Crank–Nicolson are instances of a larger, highly relevant class of methods, so-called *rational schemes.*

Definition 2.27. Let $r\colon \mathbb{C} \to \mathbb{C} \cup \{\infty\}$ be a rational function. We say that a time discretisation method F for $(S(t))_{t\geq 0}$ is *induced by* r if $F(\tau) = r(\tau\cdot(-A))$ for all sufficiently small $\tau > 0$, where $-A$ denotes the generator of $(S(t))_{t\geq 0}$. Such schemes F are called *rational schemes.*

If $|r(z)| \leq 1$ for all $z \in \mathbb{C}$ with $\operatorname{Re} z \leq 0$ and $r(z) - \mathrm{e}^z = o(z)$ as $z \to 0$ then r is called *A-acceptable.* If $|r(z)| < 1$ for all $z \in \mathbb{C}$ with $\operatorname{Re} z < 0$, it is called *A-stable.*

In particular, considering $r(z) = (1-z)^{-1}$ and $r(z) = (1+\frac{z}{2})(1-\frac{z}{2})^{-1}$ we recover (IE) and (CN) from Example 2.26, respectively.

The notions of A-acceptability and A-stability originate from [21] and [41], respectively. Sometimes, the schemes induced by A-acceptable or A-stable r are also referred to as such. For schemes induced by A-acceptable r, or, simply put, A-acceptable schemes, for which there exists $p > 0$ such that $r(z) - \mathrm{e}^z = \mathcal{O}(|z|^{p+1})$ as $z \to 0$, it was shown in [21, Thm. 3] that F is a convergent time discretisation method of order p on $\mathrm{D}(A^{p+1})$. If X is a Hilbert space and A is self-adjoint, then F even converges of order p on $\mathrm{D}(A^p)$; see, e.g., [124, Thm. 7.1]. For a more general convergence result, see Subsection 3.2.2.

Contractivity of the semigroup and the approximating scheme play a central role in Chapter 4 for the approximation of SPDEs. While the contractivity of the exponential

Euler method is immediate from the contractivity of the semigroup, we state a useful sufficient condition to verify the contractivity of rational schemes via functional calculus. A standard assumption is that the scheme F is induced by a rational function $r\colon \mathbb{C}_- \to \mathbb{C}$ with $|r(z)| \leq 1$ for all z in the negative open half plane $\mathbb{C}_-$. Clearly, this condition holds for both A-acceptable and A-stable schemes, i.e., also for implicit Euler and Crank–Nicolson.

Proposition 2.28. *Let $-A$ be the generator of a C_0-semigroup of contractions on a Banach space X. Suppose that $r\colon \mathbb{C}_- \to \mathbb{C}$ is holomorphic, $|r(z)| \leq 1$ for $z \in \mathbb{C}_-$, and let F be induced by r. Then F is contractive.*

Proof. This is a consequence of the properties of the bounded H^∞-calculus of A as the negative generator of a contraction semigroup, since $F(\tau) = r(-\tau A)$ is defined via H^∞-calculus. Note that contractivity of $(S(t))_{t\geq 0}$ implies sectoriality of A and thus the fractional powers A^β exist for $\beta > 0$. The underlying theorem can be found in [72, Thm. 10.2.24]. □

As a consequence of this proposition, contractive schemes include (IE), (CN), and some higher-order implicit Runge–Kutta methods such as Radau methods, BDF(2), Lobatto IIA, IIB, and IIC as well as some DIRK schemes.

2.3 Fundamentals of Interpolation Theory

A common choice for the spaces Y_x and Y_t in Definitions 2.19 and 2.24 on which a given space or time discretisation scheme, respectively, approximates S are domains of fractional powers of the negative of the generator $-A$. An important property of these spaces is their embedding into the real interpolation spaces with parameter ∞, i.e., for $0 < \alpha < 1$,

$$\mathrm{D}(A^\alpha) \hookrightarrow D_A(\alpha, \infty). \tag{2.3.1}$$

Here, $D_A(\alpha,\infty)$ denotes the real interpolation space $(X, \mathrm{D}(A))_{\alpha,\infty}$, which we will encounter frequently in the convergence analysis of time discretisations of stochastic PDEs. On later occasions, also the real interpolation spaces $(X, \mathrm{D}(A))_{\alpha,2}$ will be used. Hence, we include a brief survey of real interpolation theory.

Interpolation theory is concerned with finding suitable spaces X_α, $0 < \alpha < 1$, which interpolate between two given Banach spaces $X_0, X_1 \hookrightarrow Z$, which embed into the same Hausdorff topological space Z, in the sense of continuous inclusions

$$X_0 \cap X_1 \hookrightarrow X_\alpha \hookrightarrow X_0 + X_1. \tag{2.3.2}$$

Depending on the desired properties of the intermediate space X_α, different interpolation methods are used. The two primary interpolation methods are complex and real interpolation. See [105, 126] or [71, Appendix C] for details on both types of interpolation spaces. Here, we introduce real interpolation spaces briefly and state some of their elementary properties based on the aforementioned literature. Real interpolation spaces can be constructed via the *K-method* or the *J-method.* Since they agree for $\alpha \in (0,1)$, we solely present the K-method. Consider the *K-functional*

$$K(t, x; X_0, X_1) := \inf\{\|x_0\|_{X_0} + t\|x_1\|_{X_1} : x_0 \in X_0, x_1 \in X_1, x = x_0 + x_1\}$$

for $x \in X_0 + X_1$ and $t > 0$.

Definition 2.29. Let $0 < \alpha < 1$, $1 \leq p \leq \infty$, and X_0, X_1 be Banach spaces over the real numbers. Define the *real interpolation space of X_0 and X_1 with parameters α and p* by

$$(X_0, X_1)_{\alpha,p} \coloneqq \{x \in X_0 + X_1 : \|x\|_{\alpha,p} < \infty\},$$

where $\|\cdot\|_{\alpha,p}$ is defined as

$$\|x\|_{\alpha,p} \coloneqq \begin{cases} \Big(\int_0^\infty [t^{-\alpha} K(t, x; X_0, X_1)]^p \, \frac{\mathrm{d}t}{t} \Big)^{1/p}, & 1 \leq p < \infty, \\ \operatorname{ess\,sup}_{t>0} t^{-\alpha} K(t, x; X_0, X_1), & p = \infty. \end{cases}$$

This indeed defines a Banach space with the norm $\|\cdot\|_{\alpha,p}$, which admits the desired inclusions (2.3.2). In addition, the interpolation spaces can be ordered w.r.t. the parameter p: For $1 \leq p_0 \leq p_1 \leq \infty$, the embeddings

$$X_0 \cap X_1 \hookrightarrow (X_0, X_1)_{\alpha,p_0} \hookrightarrow (X_0, X_1)_{\alpha,p_1} \hookrightarrow X_0 + X_1.$$

are continuous. If X_0 and X_1 are Hilbert spaces, the complex interpolation space is obtained for $p = 2$, i.e., $(X_0, X_1)_{\alpha,2} = [X_0, X_1]_\alpha$ with equivalent norms. For sectorial A, the domains of fractional powers of the operator relate to the real interpolation spaces [105, Cor. 2.2.3]. More precisely, for $0 < \alpha < 1$ and $1 \leq p < \infty$, we have

$$(X, \mathrm{D}(A))_{\alpha,p} \hookrightarrow D_A(\alpha) \hookrightarrow (X, \mathrm{D}(A))_{\alpha,\infty} = D_A(\alpha, \infty) \hookrightarrow \overline{\mathrm{D}(A)},$$

where $D_A(\alpha) \coloneqq \{x \in D_A(\alpha, \infty) : \lim_{t \searrow 0} t^{1-\alpha} A e^{-tA} x = 0\}$. Real interpolation spaces allow us to interpolate operators.

Theorem 2.30 (Theorem C.3.3 in [71])**.** *Let X_0, X_1, Y_0, Y_1 be Banach spaces and $T \colon X_0 + X_1 \to Y_0 + Y_1$ a linear operator such that $T(X_j) \subseteq Y_j$ for $j = 1, 2$. Suppose that $\|T\|_{\mathcal{L}(X_j,Y_j)} = C_j$ for $C_j \geq 0$ and $j = 1, 2$. Then for all $0 < \alpha < 1$ and $1 \leq p \leq \infty$, it holds that T maps $(X_0, X_1)_{\alpha,p}$ into $(Y_0, Y_1)_{\alpha,p}$ and its operator norm is bounded by*

$$\|T\|_{\mathcal{L}((X_0,X_1)_{\alpha,p},(Y_0,Y_1)_{\alpha,p})} \leq C_0^{1-\alpha} C_1^\alpha.$$

The above theorem is also helpful to interpolate between operators with the same domain or codomain noting that $(X, X)_{\alpha,p} = X$.

Embeddings of the form (2.3.1) and properties of $D_A(\alpha, \infty)$ allow us to obtain decay rates for semigroup differences as follows. Let $(S(t))_{t\geq 0}$ be a C_0-semigroup with generator $-A$ such that $\|S(t)\| \leq M\mathrm{e}^{wt}$ for some $M \geq 1$ and $w \geq 0$ for all $t \geq 0$. We recall that such M and w exist for every C_0-semigroup; cf. Proposition 2.2. Then $\|S(t) - S(s)\|_{\mathcal{L}(X)} \leq 2M\mathrm{e}^{wT}$ for $0 \leq s \leq t \leq T$. Since

$$\big\|[S(t) - S(s)]x\big\|_X = \Big\| \int_s^t (-S(r)Ax) \, \mathrm{d}r \Big\|_X \leq M\mathrm{e}^{wT}(t - s)\|x\|_{\mathrm{D}(A)}$$

for $x \in \mathrm{D}(A)$, we have $\|S(t) - S(s)\|_{\mathcal{L}(\mathrm{D}(A),X)} \leq 2M\mathrm{e}^{wT}(t - s)$. By interpolation as in Theorem 2.30,

$$\|S(t) - S(s)\|_{\mathcal{L}(D_A(\alpha,\infty),X)} \leq 2^{1-\alpha} M\mathrm{e}^{wT}(t - s)^\alpha \leq 2M\mathrm{e}^{wT}(t - s)^\alpha$$

for $\alpha \in (0,1)$. Let Y be another Banach space such that $Y \hookrightarrow X$. Under the assumption that $Y \hookrightarrow D_A(\alpha,\infty)$ continuously for some $\alpha \in (0,1)$ or $Y \hookrightarrow \mathrm{D}(A)$ continuously, in which case we set $\alpha = 1$, this implies

$$\|S(t) - S(s)\|_{\mathcal{L}(Y,X)} \leq 2M\mathrm{e}^{wT} C_Y (t-s)^{\alpha}, \tag{2.3.3}$$

where C_Y denotes the embedding constant of Y into $D_A(\alpha,\infty)$ or $\mathrm{D}(A)$. This interpolation estimate will prove helpful in estimating convergence rates, e.g. in Section 4.4.

Interpolating between two Sobolev spaces can lead to more general function spaces, such as Besov spaces $B^s_{p,q}$ or Bessel potential spaces $H^{p,s}$, where $s \in \mathbb{R}$ and $p,q \in [1,\infty]$. See [15, Chapter 6.2] for their definition. They are generalisations of Sobolev spaces in the sense that $H^{p,N} = W^{p,N}$ as well as $B^k_{2,2} = H^k$ for $N \in \mathbb{N}$. Similarly to Sobolev embeddings, certain embeddings hold for these spaces depending on their parameters and the dimension of the underlying space.

Theorem 2.31 (Theorem 6.5.1 in [15])**.** *Let $k, \ell \in \mathbb{R}$, $d \in \mathbb{N}$, and $1 \leq p_1 \leq p_2 \leq \infty$. Assume that $k - \frac{d}{p_1} \geq \ell - \frac{d}{p_2}$. Then for all $1 \leq q_1 \leq q_2 \leq \infty$,*

$$B^k_{p_1,q_1}(\mathbb{R}^d) \hookrightarrow B^\ell_{p_2,q_2}(\mathbb{R}^d).$$

If, moreover, $p_1 > 1$ and $p_2 < \infty$, then

$$H^{k,p_1}(\mathbb{R}^d) \hookrightarrow H^{\ell,p_2}(\mathbb{R}^d).$$

The term $k - \frac{d}{p_1}$ appearing in Theorem 2.31 measures the smoothness of the corresponding space. The smaller it is, the less regular the space. This term is referred to as the *Sobolev index*.

Bessel potential spaces also arise in the estimation of products in Sobolev spaces that are not Banach algebras. Likewise, they play an important role in the estimation of compositions with nonlinear functions, as the following three propositions illustrate. They will be particularly helpful in the study of stochastic Schrödinger equations in Subsections 4.1.3 and 4.4.4.

Proposition 2.32 (Proposition 2.1.1 in [122])**.** *Let $\sigma \geq 0$, $1 < p < \infty$, $q_1, r_1 \in (1,\infty]$, $q_2, r_2 \in (1,\infty)$ such that $\frac{1}{p} = \frac{1}{q_1} + \frac{1}{q_2} = \frac{1}{r_1} + \frac{1}{r_2}$. Then there is a constant $C \geq 0$ such that*

$$\|fg\|_{H^{\sigma,p}} \leq C\|f\|_{L^{q_1}}\|g\|_{H^{\sigma,q_2}} + C\|f\|_{H^{\sigma,r_2}}\|g\|_{L^{r_1}}$$

for all $f \in L^{q_1} \cap H^{\sigma,r_2}$ and $g \in L^{r_1} \cap H^{\sigma,q_2}$.

Proof. We only need to comment on the case $\sigma = 0$ not included in [122]. The estimate reduces to Hölder's inequality in this case, and any of the two terms on the right-hand side could be omitted. □

Proposition 2.33 (Proposition 2.4.1 in [122])**.** *Let $\phi\colon \mathbb{C} \to \mathbb{C}$ be holomorphic. Assume $\phi(0) = 0$ and there is a constant $K \geq 0$ such that $|\phi'| \leq K$. Then for $\sigma \in [0,1)$ and $p \in (1,\infty)$, there is a constant $C \geq 0$ such that for all $u \in H^{\sigma,p}$*

$$\|\phi \circ u\|_{H^{\sigma,p}} \leq CK\|u\|_{H^{\sigma,p}}.$$

Proof. We only need to comment on the case $\sigma = 0$. Since ϕ has bounded derivatives, it is Lipschitz continuous with constant K. Hence,

$$\|\phi \circ u\|_{L^p}^p = \int |\phi(u(s)) - \phi(0)|^p \, \mathrm{d}s \le \int K^p |u(s)|^p \, \mathrm{d}s = K^p \|u\|_{L^p}^p. \qquad \square$$

Proposition 2.34 (Proposition 2.6.1 in [122])**.** *Let $\phi\colon \mathbb{C} \to \mathbb{C}$ be Lipschitz continuous, $\phi(0) = 0$, $p \in (1, \infty)$, and $d \in \mathbb{N}$. If $u \in H^{1,p}(\mathbb{R}^d; \mathbb{R})$, then $\phi(u) \in H^{1,p}(\mathbb{R}^d; \mathbb{C})$.*

2.4 2-smooth Banach Spaces

Our convergence results for the temporal approximation of stochastic evolution equations in Chapter 4 are stated in 2-smooth Banach spaces. This class of spaces is a generalisation of Hilbert spaces, which is characterised by a parallelogram inequality instead of a parallelogram identity.

Definition 2.35. For $D \ge 1$, a $(2, D)$*-smooth Banach space* is a Banach space X for which

$$\|x + y\|^2 + \|x - y\|^2 \le 2\|x\|^2 + 2D^2\|y\|^2$$

holds for all $x, y \in X$. We call a Banach space 2*-smooth* if it is $(2, D)$-smooth for some $D \ge 1$.

Henceforth, the constant D is reserved for the smoothness constant, and we shall always implicitly assume $D \ge 1$ when working in $(2, D)$-smooth Banach spaces.

In the realm of stochastic analysis, this class of spaces plays an important role. As a consequence of the parallelogram identity, it includes all Hilbert spaces with $D = 1$. Furthermore, the spaces $L^p(\mu)$ are contained in this class for $2 \le p < \infty$ with $D = \sqrt{p-1}$, see [113, Prop. 2.1]. Moreover, if X is $(2, D)$-smooth and A is a closed linear operator, then $\mathrm{D}(A)$ equipped with the graph norm $\|x\|_{\mathrm{D}(A)} = (\|x\|^2 + \|Ax\|^2)^{1/2}$ is again $(2, D)$-smooth.

2.5 γ-Radonifying Operators

To give sense to stochastic integrals in Banach spaces X that are non-Hilbert, the space of γ-radonifying operators $\gamma(H, X)$ is required, where H denotes a Hilbert space. It is obtained as the closure of a subset of the space of γ-summing operators. In this section only, let $(\gamma_n)_{n \in \mathbb{N}}$ denote a Gaussian sequence, i.e., a sequence of i.i.d. standard Gaussian random variables. We introduce γ-summing operators depending on a parameter $p \in [1, \infty)$, but we shall see that the definition is actually independent of p.

Definition 2.36. We call a linear operator $R\colon H \to X$ γ*-summing* if for some $1 \le p < \infty$,

$$\|R\|_{\gamma_p^\infty(H,X)} \coloneqq \sup \Big(\mathbb{E} \Big\| \sum_{n=1}^{N} \gamma_n R h_n \Big\|_X^p \Big)^{1/p} < \infty, \tag{2.5.1}$$

with the supremum being taken over all finite orthonormal systems $\{h_1, \dots, h_N\}$ in H for $N \in \mathbb{N}$. The space of all γ-summing operators is denoted by $\gamma_p^\infty(H, X)$. We set $\gamma^\infty(H, X) \coloneqq \gamma_2^\infty(H, X)$.

The thus-obtained $(\gamma_p^\infty(H,X), \|\cdot\|_{\gamma_p^\infty(H,X)})$ is a normed space, which is contained in the space of linear bounded operators $\mathcal{L}(H,X)$. The inclusion follows immediately from considering orthonormal systems $\{h\}$ consisting of a single element. Furthermore, $\gamma^\infty(H,X)$ is a Banach space. If (2.5.1) holds for some $1 \le p < \infty$, it automatically holds for all $1 \le p < \infty$ and by the Kahane–Khintchine inequalities [72, Thm. 6.2.6], the γ-summing norms are equivalent for all $1 \le p < \infty$. This justifies omitting the index $p = 2$ in $\gamma^\infty(H,X)$.

The space of γ-radonifying operators from H to X is now obtained as the closure of finite rank operators in the space of γ-summing operators. In Section 2.7, we shall encounter it again as the "correct" space in the context of stochastic integration.

Definition 2.37. Let $N \in \mathbb{N}$, $h_n \in H$, and $x_n \in X$ for $1 \le n \le N$ and define the operator $h_n \otimes x_n \in \mathcal{L}(H,X)$ by $h \mapsto (h_n \otimes x_n)h := \langle h_n, h\rangle_H x_n$. Operators of the form $R = \sum_{n=1}^N h_n \otimes x_n$ are called *finite rank operators* and the space of all such operators is denoted by $\mathrm{FR}(H,X)$. We define the space $\gamma(H,X)$ of all *γ-radonifying operators* as the closure of $\mathrm{FR}(H,X)$ in $\gamma^\infty(H,X)$.

As a closed subspace of $\gamma^\infty(H,X)$, $\gamma(H,X)$ is a Banach space with the norm $\|\cdot\|_{\gamma(H,X)} := \|\cdot\|_{\gamma^\infty(H,X)}$.

Example 2.38. *To provide some intuition for the notion of γ-radonifying operators, we collect some basic properties and common examples.*

(a) Trivially, $\mathrm{FR}(H,X) \subseteq \gamma^\infty(H,X)$. For a finite rank operator $R = \sum_{n=1}^N h_n \otimes x_n \in \gamma(H,X)$ with orthonormal $\{h_1, \dots, h_N\} \subseteq H$, the norm simplifies to

$$\|R\|_{\gamma(H,X)}^2 = \mathbb{E}\Big\| \sum_{n=1}^N \gamma_n x_n \Big\|^2.$$

(b) In case X is a Hilbert space, the norm of R further simplifies to $\sum_{n=1}^N \|x_n\|^2$. Hence, by taking the completion, we see that $\gamma(H,X)$ coincides with the space $\mathcal{L}_2(H,X)$ of Hilbert–Schmidt operators for Hilbert spaces X and H. We recall that the space of Hilbert–Schmidt operators consists of all bounded operators $R\colon H \to X$ such that

$$\|R\|_{\mathcal{L}_2(H,X)}^2 := \sum_{i\in I} \|Rh_i\|_X^2 < \infty,$$

where $(h_i)_{i\in I}$ is an orthonormal basis of H and the norm is independent of the choice of the orthonormal basis.

(c) An example for a γ-radonifying operator on a non-Hilbert space is given by the indefinite integration operator $I_T\colon L^2(0,T) \to C[0,T]$ defined by $f \mapsto [t \mapsto \int_0^t f(s)\,\mathrm{d}s]$ for $t \in [0,T]$.

A property of $\gamma(H,X)$ frequently used in the following is the (left) ideal property.

Proposition 2.39 (Theorem 9.1.10 in [72])**.** *Let $R \in \gamma^\infty(H,X)$ and let $\tilde{H}$ and $\tilde{X}$ be another Hilbert and Banach space, respectively. Then for all $T \in \mathcal{L}(\tilde{H}, H)$ and $S \in \mathcal{L}(X, \tilde{X})$, we have $SRT \in \gamma^\infty(\tilde{H}, \tilde{X})$ and*

$$\|SRT\|_{\gamma^\infty(\tilde{H},\tilde{X})} \le \|S\|_{\mathcal{L}(X,\tilde{X})} \|R\|_{\gamma^\infty(H,X)} \|T\|_{\mathcal{L}(\tilde{H},H)}.$$

If, moreover, $R \in \gamma(H,X)$, then $SRT \in \gamma(\tilde{H}, \tilde{X})$.

For details on the aforementioned and further properties of γ-radoniyfing operators, the reader is referred to [72, Section 9.1].

2.6 Fundamentals of Stochastic Analysis

In this section, some central concepts of stochastic analysis are reviewed, such as filtrations, different measurability concepts, Brownian motions, conditional expectations, and martingales. Stochastic integrals with respect to (one-dimensional) Brownian motions are briefly introduced in preparation for the Banach space-valued stochastic integral in Section 2.7. The selection of topics covered here is by no means exhaustive. Good introductions to the topic are provided in [71, 114, 137]. We collect selected definitions and statements from these references, which form the basis of the subsequent sections. The results being widely known in the context of stochastic analysis, we have referenced the corresponding statements in the standard textbooks for the convenience of the reader.

Definition 2.40. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and I some ordered index set. A family of sub-σ-algebras $(\mathscr{F}_i)_{i\in I}$ is called a *filtration* on this probability space if $\mathscr{F}_i \subseteq \mathscr{F}_j$ for all $i \leq j$ in I. Then $(\Omega, \mathscr{F}, (\mathscr{F}_i)_{i\in I}, \mathbb{P})$ is referred to as a *filtered space.*

In the following, we will restrict our considerations to two cases: $I = \{0, \dots, N\} \cap \mathbb{N}_0$ for some $N \in \mathbb{N}_0 \cup \{\infty\}$ or $I = [0, T] \cap [0, \infty)$ for some $T \in [0, \infty]$, so that the index can be interpreted as (discrete or continuous) time. Intuitively, the filtration determines which information about some $\omega \in \Omega$ is available at a certain time $i \in I$: All the information is given by the values of $Z(\omega)$ for all $\mathscr{F}_i$-measurable functions Z. For Banach spaces X and $1 \leq p < \infty$, we denote by $L^p_{\mathscr{F}_0}(\Omega; X)$ the subspace of $L^p(\Omega; X)$ consisting of all $\mathscr{F}_0$-measurable functions. In the case $I = [0, \infty)$, the following is a common assumption on the filtration.

Definition 2.41. Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered space. Then $(\mathscr{F}_t)_{t\geq 0}$ is a *normal filtration* if all $\mathbb{P}$-null sets are contained in $\mathscr{F}_0$ and the filtration is *right continuous.* That is, $(\mathscr{F}_t)_{t\geq 0} = (\mathscr{F}_t^+)_{t\geq 0}$, where $\mathscr{F}_t^+ := \bigcap_{r>t} \mathscr{F}_r$ for $t \in [0, \infty)$.

For functions on probability spaces, several distinct notions of measurability exist. In the following, let X always denote a Banach space. In addition to the classical notion of $\mathscr{F}$-measurability of some function $f\colon \Omega \to X$ defined via preimages of Borel-measurable subsets of X, *strong $\mathscr{F}$-measurability* plays an important role, since it allows us to employ approximation arguments.

Definition 2.42. We call a function $f\colon \Omega \to X$ *$\mathscr{F}$-simple* or *simple* if for some $N \in \mathbb{N}$ it can be written as

$$f = \sum_{n=1}^{N} \mathbf{1}_{A_n} x_n$$

for sets $A_n \in \mathscr{F}$ and elements $x_n \in X$. The function $f\colon \Omega \to X$ is said to be *strongly $\mathscr{F}$-measurable* or *simply measurable* for short if there exists a sequence $(f_n)_{n\in\mathbb{N}}$ of $\mathscr{F}$-simple functions $f_n\colon \Omega \to X$ such that $\lim_{n\to\infty} f_n = f$ pointwise on Ω.

Strong measurability is equivalent to $\mathscr{F}$-measurability if f only takes values in a separable closed subspace $Y \subseteq X$. Identifying functions that are equal almost everywhere in the usual L^p way, a vector space is obtained from the set of all strongly measurable functions $f\colon \Omega \to X$. This vector space is denoted by $L^0(\Omega; X)$, which is a completely metrisable space [71, Prop. A.2.4], although not a normed one.

Remark 2.43. We would like to draw the reader's attention to the difference in the use of the term *strong* in the context of stochastic and functional analysis. In functional analysis, particularly in operator theory, "strong" refers to requiring a certain property to hold with respect to the strong operator topology. For instance, strong continuity of a semigroup $(S(t))_{t\geq 0}$ on X, i.e. $S(t) \to I$ strongly as $t \searrow 0$, means that $S(t)f \to f$ as $t \searrow 0$ for all $f \in X$. While this notion is only meaningful for operators, strong measurability in the stochastic sense (see Definition 2.42) can also be used for functions. As a rule of thumb, strong measurability shall be understood in the sense of Definition 2.42 in this thesis, unless stated otherwise.

Two main concepts of measurability are to be distinguished for stochastic processes. We recall that stochastic processes are families $(f_i)_{i\in I}$ of X-valued random variables defined on the same underlying probability space indexed over a variable $i \in I$ often representing time, i.e., $f_i \in L^0(\Omega; X)$ for all $i \in I$. Depending on the context, it can be helpful to consider stochastic processes as mappings $f\colon \Omega\times I \to X$ instead of $(f_i)_{i\in I}$ with *trajectories* $f_i\colon \Omega \to X$. Taking another perspective, one can study the *paths* $f(\omega,\cdot)\colon I \to X$, which describe the evolution in time for a specific realisation of the random variable. It is important to distinguish between *adapted* stochastic processes and *progressively measurable* ones.

Definition 2.44. Let $(\Omega, \mathscr{F}, (\mathscr{F}_i)_{i\in I}, \mathbb{P})$ be a filtered space and $(f_i)_{i\in I}$ a stochastic process with $f_i \in L^0(\Omega; X)$. Then $(f_i)_{i\in I}$ is called *adapted* to the filtration $(\mathscr{F}_i)_{i\in I}$ if each f_i is strongly $\mathscr{F}_i$-measurable.

Definition 2.45. Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,\infty)}, \mathbb{P})$ be a filtered space. The *progressive* σ*-algebra* $\mathcal{P}$ is given by

$$\{A \subseteq \Omega \times [0,\infty) : A \cap (\Omega \times [0,t]) \in \mathscr{F}_t \otimes \mathcal{B}([0,t]) \text{ for all } t \geq 0\}.$$

An X-valued stochastic process $(f_t)_{t\geq 0}$ is called *progressively measurable* if it is measurable with respect to the progressive σ-algebra.

Progressive measurability of $f\colon \Omega\times[0,\infty) \to X$ thus means that $\Omega\times[0,t] \to X$, $(\omega, s) \mapsto (f|_{\Omega\times[0,t]})_s(\omega)$ is $\mathscr{F}_t\otimes\mathcal{B}([0,t])$-measurable for all $t \geq 0$. The definition extends to functions from $\Omega \times [0,T]$ for some $T > 0$ in the straightforward way. Henceforth, we will use the index $\mathcal{P}$ to denote the progressively measurable subspace of a given space, as, e.g., in $L^p_{\mathcal{P}}(\Omega; L^q(0,T;X))$ for $p \in [0,\infty)$, $q \in [1,\infty]$ and $T > 0$. Progressive measurability is the stronger notion in the sense that every progressively measurable process is also adapted as an immediate consequence of the definition, but not vice versa. In practice, one often makes use of the following property to show progressive measurability.

Proposition 2.46 (Proposition 4.8 in [114])**.** *Let $f\colon \Omega \times [0,\infty) \to X$ be adapted and have left- or right-continuous paths $f(\omega,\cdot)$ for each $\omega \in \Omega$. Then f is progressively measurable.*

Proof. Without loss of generality, assume f has right-continuous paths. Fix $t \geq 0$. Define $f_n\colon \Omega \times [0,t] \to X$ by $f_n(\omega, 0) := f(\omega, 0)$ and

$$f_n(\omega, s) := f\left(\omega, \frac{(k_n+1)t}{2^n}\right) \text{ for } s \in \left(\frac{k_n t}{2^n}, \frac{(k_n+1)t}{2^n}\right],$$

where $k_n = 0, \dots, 2^n - 1$. For $s \in (0,t]$, $s_n := 2^{-n}(k_n+1)t \searrow s$ as $n \to \infty$. Thus, by construction of f_n and right-continuity of paths of f, we conclude $\lim_{n\to\infty} f_n(\omega, s) = \lim_{n\to\infty} f(\omega, s_n) = f(\omega, s)$ for all $\omega \in \Omega$. Moreover, $\lim_{n\to\infty} f_n(\omega, 0) = f(\omega, 0)$. Hence, $\lim_{n\to\infty} f_n(\omega, s) = f(\omega, s)$ for all $(\omega, s) \in \Omega \times [0,t]$. It remains to show that f_n is progressively measurable for all $n \in \mathbb{N}$, as pointwise limits of progressively measurable functions are progressively measurable. Indeed, we can write f_n as a step function in time

$$f_n(\omega, s) = \sum_{k=0}^{2^n-1} \mathbf{1}_{\left(\frac{kt}{2^n}, \frac{(k+1)t}{2^n}\right]}(s) f\left(\omega, \frac{(k+1)t}{2^n}\right) + \mathbf{1}_{\{0\}}(s) f(\omega, 0),$$

where the random variables $\Omega \to X, \omega \mapsto f(\omega, 2^{-n}(k+1)t)$ are $(\mathscr{F}_{2^{-n}(k+1)t})$- measurable for all $0 \le k \le 2^n - 1$ due to adaptedness of f. $\square$

A particularly well-studied class of stochastic processes are Brownian motions.

Definition 2.47. Let $T > 0$. A real-valued stochastic process $(W(t))_{t\in[0,T]}$ is called a *Brownian motion* if the following three properties hold.

(a) $W(0) = 0$ almost surely.

(b) $W(t) - W(s)$ is normally distributed with mean 0 and variance $t - s$ for all $0 \le s \le t \le T$.

(c) $W(t) - W(s)$ is independent of $\{W(r) : 0 \le r \le s\}$ for all $0 \le s \le t \le T$.

Such processes indeed exist [114, Thm. 1.9]. Often, pathwise continuity of $(W(t))_{t\in[0,T]}$ is included in the above definition. However, it is already a consequence of the three properties stated above and Kolmogorov's continuity criterion [114, Thm. 1.8] that a Brownian motion has a modification with continuous paths [114, Thm. 1.9], even Hölder continuous with exponent below $\frac{1}{2}$. That is, there exists a stochastic process $(\tilde{W}(t))_{t\in[0,T]}$ such that for each $t \in [0,T]$, $W(t) = \tilde{W}(t)$ almost surely and $t \mapsto \tilde{W}(t)$ is (Hölder) continuous.

In order to formulate, analyse, and eventually approximate stochastic (partial) differential equations, stochastic integrals are indispensable. While Chapter 4 will be concerned with infinite-dimensional noise corresponding to the so-called Itô integrals w.r.t. H-cylindrical Brownian motions, we start by introducing the less technical Wiener integral w.r.t. a Brownian motion. The main ideas carry over to the infinite-dimensional case treated in Section 2.7. In particular, a certain isometry already plays a central role.

Definition 2.48. Let $(W(t))_{t\ge0}$ be a standard Brownian motion on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. The *Wiener integral* of an indicator function $\mathbf{1}_{(a,b)}$ with $0 \le a < b$ is defined as

$$\int_0^\infty \mathbf{1}_{(a,b)} \, \mathrm{d}W := W(b) - W(a) \in L^2(\Omega).$$

This definition extends to step functions $\Phi\colon (0,\infty) \to \mathbb{R}$ by linearity. We make a very useful observation: On the one hand, the definition of the Wiener integral yields

$$\left\| \int_0^\infty \mathbf{1}_{(a,b)} \, \mathrm{d}W \right\|_{L^2(\Omega)}^2 = \int_\Omega (W(b) - W(a))^2 \, \mathrm{d}\mathbb{P} = \mathbb{E}|W(b) - W(a)|^2 = b - a$$

recalling that $W(b)-W(a)$ is normally distributed with variance $b-a$. On the other hand, of course, we also have

$$\|\mathbf{1}_{(a,b)}\|^2_{L^2(0,\infty)} = \int_0^\infty \mathbf{1}^2_{(a,b)}(t)\,\mathrm{d}t = \int_a^b 1\,\mathrm{d}t = b-a.$$

Thus, the L^2-norms in time of the integrand and in randomness of the stochastic integral agree. Hence, the Wiener integral acts as an isometry from the space of step functions on $[0,\infty)$ to $L^2(\Omega)$. Consequently, we can extend it uniquely to an isometry from $L^2(0,\infty)$ to $L^2(\Omega)$ by density of step functions. The isometry thus obtained is the famous *Wiener–Itô isometry*

$$J_W\colon L^2(0,\infty)\to L^2(\Omega), \quad J_W\Phi =: \int_0^\infty \Phi\,\mathrm{d}W \tag{2.6.1}$$

defining the Wiener integral for integrands $\Phi\in L^2(0,\infty)$ which need not be step functions. This constitutes the simplest, $\mathbb{R}$-valued case. Extensions to Banach space-valued functions and integrals w.r.t infinite-dimensional analogues of the Brownian motion are discussed in the next section.

Another central concept in stochastic analysis are conditional expectations, a special type of random variables. In particular, they are needed to introduce martingales. For $1\le p\le\infty$ and a sub-σ-algebra $\mathscr{G}\subseteq\mathscr{F}$, we denote by $L^p(\Omega,\mathscr{G};X)$ the closed subspace of all $f\in L^p(\Omega;X)$ that have a $\mathscr{G}$-measurable representative. Clearly, $L^p(\Omega,\mathscr{F};X)=L^p(\Omega;X)$.

Definition 2.49. Let $(\Omega,\mathscr{F},\mathbb{P})$ be a probability space and $\mathscr{G}\subseteq\mathscr{F}$ a sub-σ-algebra. Let X be a Banach space and $f\in L^1(\Omega;X)$. Then we call $g\in L^1(\Omega,\mathscr{G};X)$ the *conditional expectation w.r.t.* $\mathscr{G}$ *of* f if

$$\forall A\in\mathscr{G}: \int_A g\,\mathrm{d}\mathbb{P} = \int_A f\,\mathrm{d}\mathbb{P}.$$

We denote $\mathbb{E}(f|\mathscr{G}) := g$.

The use of the word "the" is justified since existence and uniqueness (up to versions which agree almost everywhere) of the conditional expectation hold true in $L^1(\Omega;X)$ [71, Thm. 2.6.23]. It is common to write $\mathbb{E}(f|h)$ as an abbreviation for $\mathbb{E}(f|\sigma(h))$, where $h\colon\Omega\to X$ is another random variable and $\sigma(h)$ the σ-algebra generated by it. We list some useful properties that can be found in [71, Sec. 2.6].

Proposition 2.50. *Let $(\Omega,\mathscr{F},\mathbb{P})$ be a probability space, $\mathscr{G},\mathscr{H}\subseteq\mathscr{F}$ sub-σ-algebras, and $f,g\in L^1(\Omega;X)$ random variables. Then the following statements hold true.*

(a) If g is any version of $\mathbb{E}(f|\mathscr{G})$, then $\mathbb{E}(g)=\mathbb{E}(f)$.

(b) If f is $\mathscr{G}$-measurable, then $\mathbb{E}(f|\mathscr{G})=f$ almost surely.

(c) If f is independent of $\mathscr{G}$, then $\mathbb{E}(f|\mathscr{G})=\mathbb{E}(f)$ is constant.

(d) Taking conditional expectations w.r.t. $\mathscr{G}$ is linear.

(e) (conditional Jensen's inequality) If $\phi\colon X\to\mathbb{R}$ is convex and lower semicontinuous and $\phi(f)\in L^1(\Omega,\mathscr{G})$, then $\mathbb{E}(\phi(f)|\mathscr{G})\ge\phi(\mathbb{E}(f|\mathscr{G}))$ almost surely. In particular, $\|\mathbb{E}(f|\mathscr{G})\|_{L^p(\Omega;X)}\le\|f\|_{L^p(\Omega;X)}$ for all $p\in[1,\infty)$.

(f) *If g is $\mathscr{G}$-measurable and bounded, then $\mathbb{E}(gf|\mathscr{G}) = g\mathbb{E}(f|\mathscr{G})$ almost surely. The same holds true if $f \in L^p(\Omega; X)$ and $g \in L^q(\Omega, \mathscr{G}; X)$ for $p \in (1, \infty)$ and q such that $\frac{1}{p} + \frac{1}{q} = 1$.*

Conditional expectations are, among others, required to introduce the notion of martingales. Martingales will play a crucial role in later stability and convergence proofs; cf. Sections 4.3 and 4.6.

Definition 2.51. Let $I = \{0, \dots, N\} \cap \mathbb{N}_0$ for some $N \in \mathbb{N}_0 \cup \{\infty\}$ or $I = [0, T] \cap [0, \infty)$ for some $T \in [0, \infty]$, and let X be a Banach space. A *martingale* is a family of functions $(M_i)_{i \in I}$ in $L^1(\Omega; X)$ with respect to a filtration $(\mathscr{F}_i)_{i \in I}$ if it is adapted to $(\mathscr{F}_i)_{i \in I}$ and

$$\mathbb{E}(M_i|\mathscr{F}_j) = M_j \quad (j \leq i,\ i, j \in I).$$

Every martingale indexed over $I = \{0, \dots, N\}$ has an associated *(martingale) difference sequence* $(d_n)_{n=1,\dots,N}$ given by $d_n \coloneqq M_n - M_{n-1}$.

For discrete martingales, moments of the maximal value of a martingale up to a certain time $N \in \mathbb{N}_0$ can be estimated by a multiple of the moments of the single martingale at this time N. This estimate known as *Doob's maximal inequality* is particularly helpful to estimate moments of maximal functions, see Section 2.8.

Theorem 2.52 (Doob's maximal inequality, Theorem 3.2.2 in [71])**.** *Let X be a Banach space, $(M_n)_{n \in \mathbb{N}_0}$ be a martingale in $L^p(\Omega; X)$, $p \in (1, \infty)$, and $N \in \mathbb{N}_0$. Define the maximal function $M_N^* \coloneqq \max_{0 \leq n \leq N} \|M_n\|_X$. Then*

$$\|M_N^\star\|_p \leq \frac{p}{p-1} \|M_N\|_p.$$

Lastly, we present a useful connection between Brownian motions, conditional expectations, and martingales.

Example 2.53. *Let $(W(t))_{t \in [0,T]}$ be a Brownian motion and $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. Then $(W(t))_{t \in [0,T]}$ is a martingale with respect to the filtration $(\mathscr{F}_t^W)_{t \in [0,T]}$ defined by*

$$\mathscr{F}_t^W \coloneqq \sigma(W(s) : 0 \leq s \leq t).$$

Indeed, adaptedness is an immediate consequence of the definition of $\mathscr{F}_t^W$ and $W(t) \in L^1(\Omega)$ follows from the definition of a Brownian motion. Linearity of the conditional expectation and $W(s) \in \mathscr{F}_s^W$ allow us to write

$$\mathbb{E}\big(W(t)|\mathscr{F}_s^W\big) = \mathbb{E}\big(W(s)|\mathscr{F}_s^W\big) + \mathbb{E}\big(W(t) - W(s)|\mathscr{F}_s^W\big) = W(s) + \mathbb{E}\big(W(t) - W(s)|\mathscr{F}_s^W\big)$$

for $0 \leq s \leq t \leq T$. Since $W(t) - W(s)$ is independent of $\{W(r) : 0 \leq r \leq s\}$, it is also independent of the σ-algebra generated by it, that is, $\mathscr{F}_s^W$. Hence, Proposition 2.50(c) yields $\mathbb{E}(W(t) - W(s)|\mathscr{F}_s^W) = \mathbb{E}(W(t) - W(s)) = 0$ because $W(t) - W(s)$ has mean zero. In conclusion, $\mathbb{E}(W(t)|\mathscr{F}_s^W) = W(s)$ and thus the Brownian motion is a martingale w.r.t. $(\mathscr{F}_t^W)_{t \in [0,T]}$. From the proof, we can deduce another useful property for $0 \leq s \leq t \leq T$ and $r \in [0, T] \setminus (s, t)$:

$$\mathbb{E}(W(t) - W(s)|\mathscr{F}_r) = \begin{cases} W(t) - W(s), & r \geq t \\ 0, & r \leq s \end{cases}.$$

Recalling that $W(t) - W(s)$ is normally distributed with variance $t - s$, we further obtain

$$\mathbb{E}(|W(t) - W(s)|^2 | \mathscr{F}_s) = \mathbb{E}(|W(t) - W(s)|^2) = t - s.$$

Analogous statements hold true for H-cylindrical Brownian motions, which will be introduced in the next section.

2.7 Stochastic Integration in Banach Spaces

Let H denote a Hilbert space, X a 2-smooth Banach space and $R \in \gamma(H, X)$ a γ-radonifying operator. Since γ-radonifying operators are separably valued, we can assume separability of H without loss of generality. For a sequence $\gamma = (\gamma_n)_{n\in\mathbb{N}}$ of centred i.i.d. normally distributed random variables, we write

$$R\gamma \coloneqq \sum_{n\in\mathbb{N}} \gamma_n R h_n, \tag{2.7.1}$$

where $\{h_n\}_{n\in\mathbb{N}}$ is an orthonormal basis of H and the convergence is in $L^p(\Omega; X)$ for $p < \infty$ and almost surely (see [72, Cor. 6.4.12]). While the convergence is independent of the choice of $(h_n)_{n\in\mathbb{N}}$, $R\gamma$ depends on it in general.

In stochastic integrals appearing in expressions such as (1.3.8), the integrator is an H-cylindrical Brownian motion rather than a (standard) Brownian motion to take operator-valued integrands into account. For $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, we call a $\mathbb{K}$-valued random variable *Gaussian* if it is normally distributed with mean zero.

Definition 2.54. An *H-cylindrical Brownian motion* is a mapping $W_H \colon L^2(0, T; H) \to L^2(\Omega)$ such that

(a) $W_H b$ is Gaussian for all $b \in L^2(0, T; H)$,

(b) $W_H b$ is $\mathcal{F}_t$-measurable for all $b \in L^2(0, T; H)$ with support in $[0, t]$,

(c) $W_H b$ is independent of $\mathcal{F}_s$ for all $b \in L^2(0, T; H)$ with support in $[s, T]$,

(d) $\mathbb{E}(W_H b_1 \cdot W_H b_2) = (b_1 \,|\, b_2)_{L^2(0,T;H)}$ for all $b_1, b_2 \in L^2(0, T; H)$,

where we include a complex conjugate on $W_H b_2$ in case we want to use a complex H-cylindrical Brownian motion. As a shorthand notation, we write $W_H(t)h \coloneqq W_H(\mathbf{1}_{(0,t)} \otimes h)$ for $h \in H$ and $t \in [0, T]$.

Consequently, an H-cylindrical Brownian motion can be interpreted as the infinite-dimensional analogue of a Brownian motion in the sense that $(W_H(t)h)_{t\in[0,T]}$ is a Brownian motion for each fixed $h \in H$. It is standard if and only if $\|h\|_H = 1$. In the special case $H = \mathbb{R}$, this notion coincides with Brownian motions as introduced in Definition 2.47.

A notion closely related to H-cylindrical Brownian motions are so-called Q-Wiener processes. Subsequently, we will encounter them in the applications to concrete stochastic evolution equations, such as the stochastic Schrödinger equation. The purpose thereof is to make our results comparable with existing results in the literature formulated in terms of Q-Wiener processes.

Definition 2.55. For a Hilbert space H with orthonormal basis $(e_m)_{m\in I}$, we say that a random variable $\gamma\colon H \to L^2(\Omega)$ is *Gaussian* or *normally distributed with covariance* $Q \in \mathcal{L}(H)$ if γh_m is normally distributed with mean 0 for all $m \in I$ and $\mathbb{E}(\gamma h_m \overline{\gamma h_n}) = (Qh_m \mid h_n)_H$ for all $m, n \in I$. If $Q = I$, γ is a *standard Gaussian random variable.* We refer to an H-valued stochastic process $(W(t))_{t\geq 0}$ with $W(t)\colon H \to L^2(\Omega)$ as a *Q-Wiener process* for an operator $Q \in \mathcal{L}(H)$ if $W(0) = 0$, W has continuous trajectories and independent increments, and $W(t) - W(s)$ is normally distributed with mean 0 and covariance $(t-s)Q$ for $t \geq s \geq 0$.

The operator Q is in $\mathcal{L}(H)$, positive, self-adjoint, and of trace class. One can show that W is a Q-Wiener process if and only if there exists an H-cylindrical Brownian motion W_H such that $Q^{1/2}W_H(t) := \sum_{n\geq 1} Q^{1/2}h_n W_H(t)h_n = W(t)$ for an orthonormal basis $(h_n)_{n\geq 1}$ of H (cf. (2.7.1)). Note that $Q^{1/2}W_H(t)$ is independent of the choice of $(h_n)_{n\in\mathbb{N}}$. To consider a stochastic evolution equation with a Q-Wiener process W instead of a cylindrical Brownian motion, one can replace the multiplicative noise $G\colon \Omega \times [0,T] \times X \to \gamma(H,X)$ by $GQ^{1/2}$ and reduce to the cylindrical case.

Stochastic integrals of suitable functions $\Phi\colon (0,T) \to \mathcal{L}(H,X)$ with respect to H-cylindrical Brownian motions (or Q-Wiener processes) can be defined in the sense of Wiener integrals. We present the construction for H-cylindrical Brownian motions, but the construction for Q-Wiener processes is similar, see [39]. Compared to the Wiener integral introduced in Section 2.6, we have to take the possibly infinite dimension of H into account, which makes a second approximation argument by finite rank step functions necessary. The Wiener integral from Definition 2.48 is recovered for $H = X = \mathbb{R}$.

Definition 2.56. Let $0 \leq a < b \leq T < \infty$, $h \in H$, $x \in X$, and $(W_H(t))_{t\in[0,T]}$ be an H-cylindrical Brownian motion. We define the *Wiener integral* of the finite rank step function $\Phi = \mathbf{1}_{(a,b)}(h \otimes x)\colon (0,T) \to \mathcal{L}(H,X)$ as the random variable

$$\int_0^T \Phi \,\mathrm{d}W_H := W_H(\mathbf{1}_{(a,b)} \otimes h) \otimes x = (W_H(b) - W_H(a))h \otimes x \in L^2(\Omega; X).$$

We can extend the above definition to $\mathrm{FR}(H,X)$-valued step functions, the so-called *finite rank step functions*, by linearity. As in the scalar-valued case, an isometry allows us to extend the stochastic integral to more general functions.

Theorem 2.57 (Itô isometry)**.** *Let X be a Banach space, H a Hilbert space, $\Phi\colon (0,T) \to \mathcal{L}(H,X)$ be a finite rank step function, and define*

$$R_\Phi\colon L^2(0,T;H) \to X, \quad f \mapsto \int_0^T \Phi(t)f(t)\,\mathrm{d}t.$$

Then $R_\Phi \in \gamma(L^2(0,T;H),X)$ is γ-radonifying, $\int_0^T \Phi\,\mathrm{d}W_H$ is a Gaussian random variable, and we have

$$\mathbb{E}\Big\| \int_0^T \Phi\,\mathrm{d}W_H \Big\|^2 = \|R_\Phi\|^2_{\gamma(L^2(0,T;H),X)},$$

which defines an isometric embedding between $\gamma(L^2(0,T;H),X)$ and $L^2(\Omega;X)$ acting as $R_\Phi \mapsto \int_0^T \Phi\,\mathrm{d}W_H$.

Itô's isometry enables us to define the stochastic integral of any $R_\Psi \in \gamma(L^2(0,T;H),X)$. It remains to identify those Φ which can be represented as R_Ψ for some $\Psi\colon (0,T) \to \mathcal{L}(H,X)$. This leads to the notion of stochastic integrability.

Definition 2.58. Let X be a Banach space, H a Hilbert space, $(W_H(t))_{t\in[0,T]}$ a Brownian motion, and write $(\Phi h)(t) := \Phi(t)h$ for $\Phi\colon (0,T) \to \mathcal{L}(H,X)$ and $h \in H$. We call an operator-valued function $\Phi\colon (0,T) \to \mathcal{L}(H,X)$ *stochastically integrable w.r.t.* W_H if there exists a sequence of finite rank step functions $\Phi_n\colon (0,T) \to \mathcal{L}(H,X)$ such that $\lim_{n\to\infty} \Phi_n h = \Phi h$ in measure for all $h \in H$ and $\int_0^T \Phi_n \,\mathrm{d}W_H$ converges in probability to an X-valued random variable as $n \to \infty$. This limit in probability then defines the *Wiener integral*, or, simply, *stochastic integral*, of Φ.

This definition is independent of the choice of approximating sequence $(\Phi_n)_{n\in\mathbb{N}}$ and the stochastic integrals converge in probability if and only if they converge in $L^p(\Omega;X)$ for some (all) $1 \le p < \infty$.

One final extension of the stochastic integration theory presented thus far is in order to treat stochastic PDEs on 2-smooth Banach spaces: We want to pass from time-dependent deterministic integrands to stochastic processes, i.e., time-dependent random variables $\Phi\colon \Omega \times (0,T) \to \mathcal{L}(H,X)$ as integrands w.r.t. an H-cylindrical Brownian motion. This leads to the notion of *Itô integrals.* The extension from Wiener to Itô integrals is not possible in arbitrary Banach spaces X but requires further structure. We shall only be concerned with 2-smooth Banach spaces, since central inequalities are only available in these spaces (see Theorem 2.61 further below). A further extension to so-called *UMD Banach spaces* is possible [129] but exceeds the scope of this thesis. To define Itô integrals, we start by defining it for suitably measurable step processes.

Definition 2.59. Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered space, X a 2-smooth Banach space, and H a Hilbert space. Then $\Phi\colon \Omega\times(0,T) \to \mathcal{L}(H,X)$ is a *finite rank adapted step process* if it can be written as

$$\Phi(\omega,t) = \sum_{m=1}^{M}\sum_{n=1}^{N} \mathbf{1}_{A_{mn}}(\omega)\mathbf{1}_{(t_{n-1},t_n)}(t)\sum_{j=1}^{J} h_j \otimes x_{jmn} \tag{2.7.2}$$

for $M,N,J \in \mathbb{N}$, where for all $1 \le n \le N$, the sets $A_{mn} \in \mathscr{F}_{t_{n-1}}$ are disjoint for $m = 1,\dots,M$, $(t_n)_{n=0,\dots,N} \subseteq [0,T]$ is strictly increasing, $h_j \in H$ are orthonormal for $1 \le j \le J$, and $x_{jmn} \in X$ for all j,m,n in the respective index sets.

Let $(W_H(t))_{t\in[0,T]}$ be an H-cylindrical Brownian motion that is adapted to $(\mathscr{F}_t)_{t\in[0,T]}$. Then the *Itô integral* of a finite rank adapted step process Φ w.r.t. W_H is defined as

$$\int_0^T \Phi(t)\,\mathrm{d}W_H(t) := \sum_{m=1}^{M}\sum_{n=1}^{N}\mathbf{1}_{A_{mn}}\sum_{j=1}^{J}(W_H(t_n)h_j - W_H(t_{n-1})h_j)x_{jmn}.$$

A possible choice of filtration such that $(W_H(t))_{t\in[0,T]}$ is automatically adapted is given by the filtration $(\mathscr{F}_t^{W_H})_{t\in[0,T]}$ defined analogously to $(\mathscr{F}_t^{W})_{t\in[0,T]}$ (see Example 2.53). Unless otherwise stated, we use this filtration henceforth. Furthermore, it is useful to note that the Itô integral from Definition 2.59 is a random variable with mean zero and belongs to $L^p(\Omega, \mathscr{F}_T; X)$ for all $p \in [1,\infty)$.

Considering the trajectories $t \mapsto \Phi(\omega, t)$ of a process Φ as in (2.7.2) for each $\omega \in \Omega$ yields finite rank step functions, for which we have seen the Itô isometry for the corresponding operators $R_{\Phi(\omega,\cdot)} \in \gamma(L^2(0,T;H),X)$. As a function of ω, this yields a simple random variable $R_\Phi\colon \Omega \to \gamma(L^2(0,T;H),X)$. Having seen the construction of Wiener integrals in the scalar and the Banach space-valued cases, we would now expect an isometry. Indeed, for scalar adapted processes and Brownian motions, this is the Itô isometry

$$\mathbb{E}\Big|\int_0^1 \Phi \,\mathrm{d}W\Big|^2 = \|\Phi\|^2_{L^2(\Omega\times[0,1],\mathbb{R})}$$

in its original form due to Itô [73]. For infinite-dimensional X, the norm equality of Itô's formula is lost, but bounds in both directions still hold with constants depending only on X and the order of moments considered.

Theorem 2.60. *Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered space, X a 2-smooth Banach space, H a Hilbert space, and let $(W_H(t))_{t\in[0,T]}$ be an H-cylindrical Brownian motion that is adapted to $(\mathscr{F}_t)_{t\in[0,T]}$. Then all finite rank adapted step processes $\Phi\colon \Omega\times(0,T) \to \mathcal{L}(H,X)$ satisfy*

$$\mathbb{E}\Big\|\int_0^T \Phi(t)\,\mathrm{d}W_H(t)\Big\|^p \eqsim_{p,X} \mathbb{E}\|R_\Phi\|^p_{\gamma(L^2(0,T;H),X)}$$

for all $p \in (1,\infty)$, with constants depending only on p and X.

Hence, the Itô integral extends uniquely to an isomorphic embedding from the adapted subspace of $L^p(\Omega;\gamma(L^2(0,T;H),X))$ to $L^p(\Omega;X)$. This allows us to define the Itô integral for those stochastic processes Φ that can be strongly approximated in measure by a sequence $(\Phi_n)_{n\in\mathbb{N}}$ of finite rank adapted step processes whose Itô integrals converge to some X-valued random variable in $L^p(\Omega;X)$ as $n \to \infty$ analogous to Definition 2.58.

Further properties of H-cylindrical Brownian motions, Q-Wiener processes, and details on the Itô integral in Hilbert spaces can be found in [39]. An overview of stochastic integration in Banach spaces is contained in [130].

To estimate Itô integrals w.r.t. such H-cylindrical Brownian motions, the Burkholder–Davis–Gundy inequalities are particularly helpful. In case X is a Hilbert space and $p \in [2,\infty)$, they imply that

$$\Big(\mathbb{E}\sup_{t\in[0,T]}\Big\|\int_0^t g(s)\,\mathrm{d}W_H(s)\Big\|_X^p\Big)^{1/p} \le B_p\|g\|_{L^p(\Omega;L^2(0,T;\gamma(H,X)))} \tag{2.7.3}$$

for some constant $B_p > 0$ for all $g \in L^p(\Omega;L^2(0,T;\gamma(H,X)))$. In particular, one can take $B_2 = 2$ (by Doob's maximal inequality [71, Thm. 3.2.2] and the Itô isometry) and $B_p = 4\sqrt{p}$ for $p > 2$. Indeed, this follows by combining the scalar result of [24, Thm. A] with the reduction technique in [80, Thm. 3.1] and the simple estimate $\|(\xi^2+\eta^2)^{1/2}\|_p \le (\|\xi\|_p^2 + \|\eta\|_p^2)^{1/2}$ valid for real-valued random variables ξ and η and $p \in [2,\infty)$. In general $(2,D)$-smooth Banach spaces, (2.7.3) holds with $B_p = 10D\sqrt{p}$. This follows as a special case of the following maximal inequality for stochastic convolutions from [128] based on earlier works on the contractive case in Hilbert spaces [63], which can be extended to the quasi-contractive case via a scaling argument.

We recall that a C_0-semigroup $(S(t))_{t\geq 0}$ is said to be *quasi-contractive* with parameter $\lambda \geq 0$ if $\|S(t)\| \leq \mathrm{e}^{\lambda t}$ for all $t \geq 0$. The following maximal inequality plays a central role in the estimation of error terms due to noise in Chapter 4.

Theorem 2.61 (Theorem 4.1 in [128])**.** *Let $(S(t))_{t\geq 0}$ be a quasi-contractive C_0-semigroup with parameter $\lambda \geq 0$ on a $(2, D)$-smooth Banach space X.*
Then for every $g \in L^0_{\mathcal{P}}(\Omega; L^2(0, T; \gamma(H, X)))$ the process $(\int_0^t S(t-s)g(s)\,\mathrm{d}W_H(s))_{t\in[0,T]}$ has a continuous modification, which satisfies, for all $0 < p < \infty$,

$$\mathbb{E} \sup_{t\in[0,T]} \left\| \int_0^t S(t-s)g(s)\,\mathrm{d}W_H(s) \right\|_X^p \leq C_{p,D}^p \|g\|^p_{L^p(\Omega;L^2(0,T;\gamma(H,X)))}, \tag{2.7.4}$$

with a constant $C_{p,D}$ depending only on p and D. For $2 \leq p < \infty$ the inequality holds with $C_{p,D} = 10\mathrm{e}^{\lambda T} D\sqrt{p}$. If, in addition, X is a Hilbert space, (2.7.4) holds with $C_{p,D} = B_p$, where B_p denotes the constant from (2.7.3). In particular, one can take $B_2 = 2$ and $B_p = 4\sqrt{p}$ for $2 < p < \infty$.

Recall that for $p < 1$, the expression on the right is only a seminorm of g. Considering S as the trivial semigroup, we recover continuity of Itô's isomorphism. For a further discussion of stochastic integration in $\mathbb{R}$, the reader is referred to the previous section.

Next, we state a special maximal inequality, which will be needed to estimate stochastic integral terms not involving semigroups. It is the essential ingredient allowing us to consider arbitrary contractive time discretisation schemes instead of solely the exponential Euler method in, among others, Subsections 4.1.1 and 4.4.1. similar result with constant of order $\log(N)$ can be found in [128, Proposition 2.7].

Proposition 2.62. *Let X be a 2-smooth Banach space and let $0 < p < \infty$. Let $\Phi := (\Phi^{(j)})_{j=1}^N$ be a finite sequence in $L^p_{\mathcal{P}}(\Omega; L^2(0, T; \gamma(H, X)))$ and set*

$$I_N^\Phi(p) := \left(\mathbb{E} \sup_{t\in[0,T], j\in\{1,\dots,N\}} \left\| \int_0^t \Phi_s^{(j)}\,\mathrm{d}W_H(s) \right\|_X^p \right)^{1/p}.$$

Then for some $K_{p,D} \geq 0$,

$$I_N^\Phi(p) \leq K_{p,D} \max\left\{ \sqrt{\log(N)}, \sqrt{p} \right\} \|\Phi\|_{L^p(\Omega;\ell^\infty_N(L^2(0,T;\gamma(H,X))))} \qquad \text{if } N \geq 2.$$

If $2 \leq p < \infty$, this estimate holds with $K_{p,D} = K_D := 10D\exp(1 + \frac{1}{2\mathrm{e}})$, which is p-independent. If additionally, X is a Hilbert space, this can be improved to $K_{p,D} = K := 4\exp(1 + \frac{1}{2\mathrm{e}}) \approx 13.07$.

The above result was pointed out to Mark Veraar and the author by Sonja Cox and its short proof below by Emiel Lorist.

Proof. To prove the result, by approximation, we may assume that each $\Phi^{(j)}$ is contained in $L^\infty(\Omega; L^2(0, T; \gamma(H, X)))$. We start by showing the statement with explicit constants for Hilbert spaces X. First, consider $p_N = \log(N)$ with $N \geq 8$. Then using $\ell^{p_N} \hookrightarrow \ell^\infty$ contractively, and the Burkholder–Davis–Gundy inequalities with $B_p \leq 4\sqrt{p}$ in X (see (2.7.3)), we find

$$I_N^\Phi(p_N) \leq \left(\sum_{j=1}^N \mathbb{E} \sup_{t\in[0,T]} \left\| \int_0^t \Phi_s^{(j)}\,\mathrm{d}W_H(s) \right\|_X^{p_N} \right)^{1/p_N} \leq 4\sqrt{p_N} \left(\sum_{j=1}^N \mathbb{E} \|\Phi^{(j)}\|^{p_N}_{L^2(0,T;\gamma(H,X))} \right)^{1/p_N}$$

$$\leq 4\sqrt{p_N} N^{1/p_N} \|\Phi\|_{L^{p_N}(\Omega;\ell_N^\infty(L^2(0,T;\gamma(H,X))))}.$$

Since $\sqrt{p_N} N^{1/p} = \mathrm{e}\sqrt{\log(N)}$, this proves the result for $p = p_N$. To deduce the result for arbitrary $p \in (0, p_N)$ note that by Lenglart's inequality for increasing functions [52, Theorem 2.2] and with $r = p/p_N \in (0,1)$

$$\begin{aligned} I_N^\Phi(p)^p = I_N^\Phi(rp_N)^{rp_N} &\leq r^{-r}\big(4\mathrm{e}\sqrt{\log(N)}\big)^p \mathbb{E}\|\Phi\|_{\ell_N^\infty(L^2(0,T;\gamma(H,X)))}^{rp_N} \\ &= r^{-r}\big(4\mathrm{e}\sqrt{\log(N)}\big)^p \|\Phi\|_{L^p(\Omega;\ell_N^\infty(L^2(0,T;\gamma(H,X))))}^p. \end{aligned}$$

Taking $1/p$-th powers, the result follows. Moreover, for $p \in [2, p_N)$ the result with the stated constant follows after using $r^{-r/p} = (\frac{p_N}{p})^{1/p_N} \leq (\frac{p_N}{2})^{1/p_N} \leq \exp(\frac{1}{2\mathrm{e}})$.

If $p \in (p_N, \infty)$, then using Minkowski's inequality, we obtain

$$\begin{aligned} I_N^\Phi(p)^p &\leq \mathbb{E}\bigg|\sum_{j=1}^N \sup_{t\in[0,T]}\bigg\|\int_0^t \Phi_s^{(j)}\,\mathrm{d}W_H(s)\bigg\|_X^{p_N}\bigg|^{p/p_N} \leq \bigg(\sum_{j=1}^N \bigg|\mathbb{E}\sup_{t\in[0,T]}\bigg\|\int_0^t \Phi_s^{(j)}\,\mathrm{d}W_H(s)\bigg\|_X^p\bigg|^{p_N/p}\bigg)^{p/p_N} \\ &\leq N^{p/p_N} \sup_{j\in\{1,\dots N\}} \mathbb{E}\sup_{t\in[0,T]}\bigg\|\int_0^t \Phi_s^{(j)}\,\mathrm{d}W_H(s)\bigg\|_X^p \leq (4\mathrm{e}\sqrt{p})^p \sup_{j\in\{1,\dots N\}} \mathbb{E}\|\Phi^{(j)}\|_{L^2(0,T;\gamma(H,X))}^p, \end{aligned}$$

where we used (2.7.3) once more. Taking $1/p$-th powers and pulling the supremum over j inside the expectation, the required estimate follows.

For general 2-smooth Banach spaces X, an analogous argument with 4 replaced by $10D$ due to the different constant value in (2.7.3) yields the desired estimate and constant.

It remains to comment on the case $2 \leq N \leq 7$. Again by Lenglart's inequality, it suffices to consider $p \in [2, \infty)$. In this case, the triangle inequality and (2.7.3) give

$$\begin{aligned} I_N^\Phi &\leq \bigg(\sum_{j=1}^N \mathbb{E}\sup_{t\in[0,T]}\bigg\|\int_0^t \Phi_s^{(j)}\,\mathrm{d}W_H(s)\bigg\|_X^p\bigg)^{1/p} \leq B_p\bigg(\sum_{j=1}^N \|\Phi^{(j)}\|_{L^p(\Omega;L^2(0,T;\gamma(H,X)))}^p\bigg)^{1/p} \\ &\leq 4\sqrt{p}N^{1/p}\|\Phi\|_{L^p(\Omega;\ell_N^\infty(L^2(0,T;\gamma(H,X))))} \\ &\leq 4\exp\Big(1+\frac{1}{2\mathrm{e}}\Big)\max\{\sqrt{\log(N)},\sqrt{p}\}\|\Phi\|_{L^p(\Omega;\ell_N^\infty(L^2(0,T;\gamma(H,X))))}, \end{aligned}$$

where the last estimate follows from $N^{1/p} \leq \sqrt{7} \leq \exp(1+\frac{1}{2\mathrm{e}})$ for $2 \leq N \leq 7$. □

2.8 A Version of the Burkholder–Rosenthal Inequality

On the fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we consider a finite filtration $(\mathcal{F}_j)_{j=0}^\ell$, $\ell \in \mathbb{N}$, and denote by $\mathbb{E}_{\mathcal{F}_j} := \mathbb{E}(\cdot \mid \mathcal{F}_j)$ the conditional expectation with respect to $\mathcal{F}_j$. For an X-valued martingale $(M_j)_{j=0}^\ell$ with respect to $(\mathcal{F}_j)_{j=0}^\ell$, we denote by $(d_j)_{j=1}^\ell$ its difference sequence defined by $d_j := M_j - M_{j-1}$. Furthermore, let the non-negative random variables $M_j^\star$ (for $0 \leq j \leq \ell$) and $d_j^\star$ and $s_j(M)$ (for $1 \leq j \leq \ell$) be given by

$$M_j^\star := \max_{0\leq i\leq j}\|M_i\|, \quad d_j^\star := \max_{1\leq i\leq j}\|d_i\|, \quad s_j(M) := \Big(\sum_{i=1}^j \mathbb{E}_{\mathcal{F}_{i-1}}\|d_i\|^2\Big)^{1/2},$$

and set $M^\star := M_\ell^\star$, $d^\star := d_\ell^\star$, and $s(M) := s_\ell(M)$.

We call a mapping $V\colon \Omega \to \mathcal{L}(X)$ such that $\omega \mapsto V(\omega)x$ is strongly measurable for all $x \in X$ a *random operator* on X and a *random contraction* on X if, additionally, its range consists of contractions. A sequence of random operators $(V_j)_{j\in\mathbb{N}}$ on X is said to be *strongly predictable* in case each $V_j x$ is strongly $\mathscr{F}_{j-1}$-measurable for all $x \in X$.

An adapted X-valued sequence $(\xi_j)_{j=1}^{\ell}$ is called *conditionally symmetric given* $(\mathcal{F}_j)_{j=0}^{\ell}$ if for all $1 \le j \le \ell$ the random variables ξ_j and $-\xi_j$ are conditionally equidistributed given $\mathcal{F}_{j-1}$, i.e., for all Borel sets $B \in \mathcal{B}(X)$ it holds that

$$\mathbb{E}_{\mathcal{F}_{j-1}} \mathbf{1}_{\{\xi_j \in B\}} = \mathbb{E}_{\mathcal{F}_{j-1}} \mathbf{1}_{\{-\xi_j \in B\}}.$$

Recently, in [128, Thm. 3.1] an extended version of Pinelis's version of the Burkholder–Rosenthal inequality (see [113]) was proven. An alternative approach based on Bellman function techniques was found in [146].

Theorem 2.63 (Theorem 3.1 in [128])**.** *Let X be a $(2,D)$-smooth Banach space. Suppose that $(\widetilde{M}_j)_{j=0}^{\ell}$ is an adapted sequence of X-valued random variables, $(M_j)_{j=0}^{\ell}$ is an X-valued martingale with difference sequence $(d_j)_{j=1}^{\ell}$, $(V_j)_{j=1}^{\ell}$ is a sequence of random contractions on X that is strongly predictable, and assume that we have $\widetilde{M}_0 = M_0 = 0$ and*

$$\widetilde{M}_j = V_j \widetilde{M}_{j-1} + d_j, \quad j = 1, \dots, \ell.$$

Then for all $2 \le p < \infty$ we have

$$\|(\widetilde{M})^*\|_p \le 30p\|d^*\|_p + 40D\sqrt{p}\|s(M)\|_p.$$

If, moreover, $(M_j)_{j=0}^{\ell}$ has conditionally symmetric increments, then

$$\|(\widetilde{M})^*\|_p \le 5p\|d^*\|_p + 10D\sqrt{p}\|s(M)\|_p.$$

This martingale estimate will play a central role in the stability analysis in 2-smooth Banach spaces in Section 4.3, in particular Lemma 4.18, as well as the convergence analysis for irregular nonlinearities in Section 4.6.

2.9 Generalised Hölder, Orlicz, and Besov–Orlicz Spaces

Showing optimal and quasi-optimal error estimates on the full time interval in Subsections 4.4.3 and 4.5.3 via maximal regularity estimates involves three types of spaces commonly used in harmonic analysis: generalised Hölder, Orlicz, and Besov–Orlicz spaces.

We start by recalling the definition of *generalised Hölder spaces*. Given a non-decreasing function $\Phi\colon [0,T] \to [0,\infty)$ such that $\Phi \neq 0$ on $(0,T]$ we say that $u \in C^{\Phi}([0,T];X)$ if $u\colon [0,T] \to X$ is continuous and

$$[u]_{C^{\Phi}([0,T];X)} \coloneqq \sup_{0\le s<t\le T} \frac{\|u(t)-u(s)\|}{\Phi(t-s)} < \infty.$$

Moreover, we set $\|u\|_{C^{\Phi}([0,T];X)} \coloneqq \|u\|_{L^{\infty}(0,T;X)} + [u]_{C^{\Phi}([0,T];X)}$, which makes $C^{\Phi}([0,T];X)$ a Banach space. We further need generalisations of both L^p- and Besov spaces.

Definition 2.64. Let X be a Banach space and $(U, \mathcal{U}, \mu)$ a σ-finite measure space. A function $\mathcal{N} \colon [0,\infty) \to [0,\infty)$ is said to be a *Young function* if it is non-decreasing, left-continuous, convex, and satisfies $\mathcal{N}(0) = 0$ and $\lim_{x\to\infty} \mathcal{N}(x) = \infty$. The *Orlicz space* $L^{\mathcal{N}}(U,\mu;X)$ consists of all $f \colon U \to X$ with finite *Luxemburg norm* $\|\cdot\|_{L^{\mathcal{N}}(U,\mu;X)}$ given by

$$\|f\|_{L^{\mathcal{N}}(U,\mu;X)} := \inf\Big\{\lambda > 0 : \int_U \mathcal{N}\Big(\frac{\|f\|_X}{\lambda}\Big)\,\mathrm{d}\mu \le 1\Big\}.$$

The measure can be omitted from the notation if the measure used is the Lebesgue measure. Furthermore, let $\alpha \in (0,1)$, $q \in [1,\infty]$, $I \subseteq \mathbb{R}$ be an interval, and $\Phi_2 \colon [0,\infty) \to [0,\infty)$, $\Phi_2(x) := \exp(x^2)-1$. We define $\omega_{\Phi_2,I}(f,t) := \sup\{\|f(\cdot+h)-f\|_{L^{\Phi_2}(I(h);X)} : |h| \le t\}$ with $I(h) := \{s \in I : s+h \in I\}$ for $h \in \mathbb{R}$ and $f \in L^{\Phi_2}(I;X)$. The *Besov–Orlicz space* $B^\alpha_{\Phi_2,q}(I;X)$ consists of all $f \in L^{\Phi_2}(I;X)$ such that

$$\|f\|_{B^\alpha_{\Phi_2,q}(I;X)} := \|f\|_{L^{\Phi_2}(I;X)} + \|t^{-\alpha}\omega_{\Phi_2,I}(f,t)\|_{L^q(0,\infty;t^{-1}\,\mathrm{d}t)} < \infty.$$

Orlicz spaces are Banach spaces and can be seen as generalised L^p-spaces, since the L^p-spaces are recovered for $\mathcal{N}(t) = t^p$ for $1 < p < \infty$. Besov–Orlicz spaces are Banach spaces as well. The Besov–Orlicz space $B^\alpha_{\Phi_2,q}(I;X)$ is continuously embedded into $C^{r^\alpha|\log r|^{1/2}}(I;X)$ for $q \in [1,\infty]$ and $\alpha \in (0,1)$ [109, Formula (2.12)].

Theorem 2.65 (Theorem 3.2(vi) in [109])**.** *Let X be a separable 2-smooth Banach space, H a Hilbert space, and W_H an H-cylindrical Brownian motion. Then for all $T \ge 0$ there exists a constant $C \ge 0$ depending on T such that*

$$\Big\|t \mapsto \int_0^t f(s)\,\mathrm{d}W_H(s)\Big\|_{L^p(\Omega;B^\alpha_{\Phi_2,\infty}(0,T;X))} \le Cp^{1/2}\|f\|_{L^{N_p}(\Omega;L^q(0,T;\gamma(H,X)))} \tag{2.9.1}$$

holds for all $p \in [1,\infty)$, $q \in (2,\infty]$, and $f \in L^{N_p}(\Omega;L^q(0,T;\gamma(H,X)))$, where $\alpha = \frac{1}{2} - \frac{1}{q}$ and $N_p(t) = t^p\log^{p/2}(t+1)$. In particular, if $f \in L^{p_0}(\Omega;L^q(0,T;\gamma(H,X)))$ for some $p_0 \in (p,\infty)$, (2.9.1) *holds and the norm of f can be replaced by the norm in $L^{p_0}(\Omega;L^q(0,T;\gamma(H,X)))$.*

Proof. We only comment on the statement concerning p_0 and refer to [109] for the proof of the main statement. Note that for all $\varepsilon > 0$ there is a constant $c > 0$ such that $\log^{p/2}(t+1) \le ct^\varepsilon$ for all $t \in [0,T]$ due to the growth behaviour of the logarithm. Since the Orlicz space with Young function $[t \mapsto t^{p+\varepsilon}]$ coincides with $L^{p+\epsilon}$, the last statement follows with $\varepsilon = p_0 - p$. □

2.10 Gronwall-type Lemmas

If an estimate of a quantity is only available in terms of an integral or a sum over said quantity, Gronwall-type lemmas provide a useful technical tool to obtain a direct estimate. The classical Gronwall inequality concerns estimates involving integrals of continuous functions.

Lemma 2.66 (Gronwall's inequality [58])**.** *Let $\phi \colon [0,T] \to [0,\infty)$ be a continuous function and let $\alpha, \beta \in [0,\infty)$ be constants. Suppose that for $t \in [0,T]$,*

$$\phi(t) \le \alpha + \beta\int_0^t \phi(s)\,\mathrm{d}s.$$

Then for $t \in [0,T]$,

$$\phi(t) \le \alpha e^{\beta t}.$$

We need the following two variants of the classical Gronwall inequality.

Lemma 2.67. *Let* $\phi\colon [0,T] \to [0,\infty)$ *be a continuous function and let* $\alpha, \beta \ge 0$ *be constants. Suppose that for* $t \in [0,T]$,

$$\phi(t) \le \alpha + \beta\Big(\int_0^t \phi(s)^2 ds\Big)^{1/2}.$$

Then for $t \in [0,T]$,

$$\phi(t) \le \alpha(1+\beta^2 t)^{1/2} \exp\Big(\frac{1}{2} + \frac{1}{2}\beta^2 t\Big).$$

Proof. Using the variant $(a+b)^2 \le (1+\theta)a^2 + (1+\theta^{-1})b^2$ of Young's inequality for $a, b \ge 0$ and $\theta > 0$, we can write

$$\phi(t)^2 \le (1+\theta)\alpha^2 + \beta^2(1+\theta^{-1}) \int_0^t \phi(s)^2 ds$$

for $t \in [0,T]$. Therefore, applying Gronwall's inequality from Lemma 2.66 to ϕ^2, we see that

$$\phi(t)^2 \le (1+\theta)\alpha^2 \exp\big(\beta^2(1+\theta^{-1})t\big).$$

Taking $\theta = \beta^2 t$ we obtain

$$\phi(t)^2 \le (1+\beta^2 t)\alpha^2 \exp(\beta^2 t + 1),$$

which gives the desired estimate. □

In the same way, one can prove the following discrete analogue by using the discrete version of Gronwall's inequality instead (see [69, Prop. 5]).

Lemma 2.68. *Let* $(\varphi_j)_{j\ge 0}$ *be a non-negative sequence and let* $\alpha, \beta \ge 0$ *be constants. Suppose that for* $j \in \mathbb{N}_0$,

$$\varphi_j \le \alpha + \beta\Big(\sum_{i=0}^{j-1} \varphi_i^2\Big)^{1/2}.$$

Then for $j \in \mathbb{N}_0$,

$$\varphi_j \le \alpha(1+\beta^2 j)^{1/2} \exp\Big(\frac{1}{2} + \frac{1}{2}\beta^2 j\Big).$$

Chapter 3

Approximation of Evolution Equations with Random Coefficients

Approximating a random evolution equation rather than a deterministic one raises the question of how to take the randomness into account when discretising and under which regularity assumptions the separate rates in space, time, and randomness are preserved for the full discretisation error. In the case of random coefficients considered here, a possible answer to the first question consists of using polynomial chaos expansions. Depending on the smoothness of the argument, we prove polynomial convergence of arbitrary order for the discretisation in randomness alone. The answer to the second question is more involved and is given by the main result of this chapter. Suppose that the space discretisation converges of some rate $p_{\mathrm{x}} > 0$ on $\mathrm{D}(A_z^{\alpha})$ for some $\alpha > 0$, and an A-stable rational time discretisation method is employed that converges of order $p_{\mathrm{t}} > 0$. Here, $-A_z$ is the generator of a C_0-semigroup and is given by a weak formulation involving a form a_z. Both depend on the realisation z of a random variable Z with distribution $\mathbb{P}_Z$ and taking values in $\mathbb{R}^N$, $N \in \mathbb{N}$. Under suitable assumptions on the smoothness of the forms $(a_z)_{z\in\mathbb{R}^N}$ and the initial values, for a large class of admissible random variables Z, we obtain convergence of the full discretisation with rates p_{x}, p_{t}, and any $\ell \in \mathbb{N}$ in space, time, and randomness, respectively. Provided that, in addition, the form is $\mathbb{P}_Z$-almost surely symmetric, this holds for initial values in the intersection of $L^2(\mathbb{R}^N, \mathbb{P}_Z; \mathrm{D}(A_z^{\max\{\alpha+1,p_{\mathrm{t}}\}}))$ and $H^{2\ell}(\mathbb{R}^N, \mathbb{P}_Z; \mathrm{D}(A_z))$. Without the symmetry assumption, additional spatial regularity is required for the joint convergence rate to hold, and further assumptions are imposed on the time discretisation schemes. For more details, the interested reader is invited to read Section 1.2 of the introduction.

Let us outline the content of this chapter. In Section 3.1, we formally introduce the random evolution equations we consider and recall properties of the corresponding semigroups. Central convergence results required for the separate approximation in space, time, and randomness are reviewed in Section 3.2, which also contains a novel quantified version of the Trotter–Kato theorem for form-induced semigroups and quantified error estimates for multivariate polynomial chaos; cf. Theorems 3.8 and 3.25. We recall space-time discretisation in the deterministic setting in Section 3.3. The main part is Section 3.4, where we prove the convergence theorem for the full discretisation of a random evolution equation, including rates. The main results can be found in Theorems 3.74 and 3.75. Section 3.5 illustrates these results for a random parabolic equation with anisotropic diffusion.

3.1 Random Evolution Equations

Let H be a separable Hilbert space and $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space. For $\omega \in \Omega$ let A_ω be a closed and densely defined operator in H such that $(A_\omega)_{\omega\in\Omega}$ is measurable, i.e., $\bigcap_{\omega\in\Omega} \rho(A_\omega) \neq \varnothing$ and $\Omega \ni \omega \mapsto (\lambda - A_\omega)^{-1} \in \mathcal{L}(H)$ is weakly measurable for one, and hence all, $\lambda \in \bigcap_{\omega\in\Omega} \rho(A_\omega) \in \mathcal{L}(H)$. Note that by Pettis' measurability theorem [112, Cor. 1.11] and separability of H this is equivalent to strong measurability of $\Omega \ni \omega \mapsto (\lambda - A_\omega)^{-1}$ for all $\lambda \in \bigcap_{\omega\in\Omega} \rho(A_\omega)$.

We consider the random evolution equation

$$u_\omega'(t) = -A_\omega u_\omega(t) \quad (t > 0), \quad u_\omega(0) = u_{0,\omega} \in H.$$

We will study this random family of evolution equations in the following way: Let $\mathbf{A}$ in $\mathbf{H} := L^2(\Omega, \mathbb{P}) \otimes H \cong L^2(\Omega, \mathbb{P}; H)$ be the multiplication operator defined by

$$\begin{aligned} \mathrm{D}(\mathbf{A}) &:= \Big\{ f \in \mathbf{H};\ f(\omega) \in \mathrm{D}(A_\omega) \text{ for } \mathbb{P}\text{-a.e. } \omega \in \Omega,\ \int_\Omega \|A_\omega f(\omega)\|_H^2 \,\mathrm{d}\mathbb{P}(\omega) < \infty \Big\}, \\ \mathbf{A}f &:= \big[\Omega \ni \omega \mapsto A_\omega f(\omega) \in H\big] \end{aligned}$$

Let $\mathcal{N}_\Omega \subseteq \Omega$ be a $\mathbb{P}$-null set such that for all $\omega \in \Omega \setminus \mathcal{N}_\Omega$ the operator $-A_\omega$ generates a strongly continuous semigroup $(S_\omega(t))_{t\geq 0}$ and there exists $\sigma \in \mathbb{R}$ such that $(\sigma, \infty) \subseteq \rho(-A_\omega)$ for all $\omega \in \mathcal{N}_\Omega$. Note that the exponential formula then yields that $\Omega \ni \omega \mapsto \mathbf{1}_{\Omega\setminus\mathcal{N}_\Omega}(\omega) S_\omega(t)$ is strongly measurable for all $t \geq 0$. As a consequence of [123, Prop. 2.3.11], we obtain the following proposition.

Proposition 3.1. *Let $M \geq 1$, $\sigma \in \mathbb{R}$ such that $\|S_\omega(t)\| \leq M\mathrm{e}^{\sigma t}$ for all $t \geq 0$ and $\omega \in \Omega \setminus \mathcal{N}_\Omega$. Then $-\mathbf{A}$ is the generator of a C_0-semigroup $(\mathbf{S}(t))_{t\geq 0}$ satisfying $\|\mathbf{S}(t)\| \leq M\mathrm{e}^{\sigma t}$ for all $t \geq 0$, and for $f \in \mathbf{H}$, $t \geq 0$ we have $(\mathbf{S}(t)f)(\omega) = \mathbf{1}_{\Omega\setminus\mathcal{N}_\Omega}(\omega) S_\omega(t) f(\omega)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$.*

Remark 3.2. In view of Proposition 3.1, we may also adjust $A_\omega := 0$ for $\omega \in \mathcal{N}_\Omega$ to get $S_\omega(t) = I$ for all $t \geq 0$ and $\omega \in \mathcal{N}_\Omega$. For $t \geq 0$ and $f \in \mathbf{H}$ we then obtain $(\mathbf{S}(t)f)(\omega) = S_\omega(t) f(\omega)$ for $\mathbb{P}$-a.e. $\omega \in \Omega$.

We then consider

$$u'(t) = -\mathbf{A}u(t) \quad (t > 0), \quad u(0) = u_0 \in \mathbf{H}.$$

In order to approximate the solution u, we have to combine three types of approximations, namely *spatial* approximations taking care of approximation w.r.t. the space H, *temporal* approximations for the time t (typically considered on bounded intervals), and *randomness* approximations taking care of the randomness in $L^2(\Omega, \mathbb{P})$.

3.2 Approximation Methods

In this section, we review the three types of approximations separately. We start with spatial approximations, then turn to temporal approximations, and conclude with randomness approximations.

3.2.1 Spatial approximation

Let V, H be separable Hilbert spaces, $V \hookrightarrow H$ with V dense in H so that we obtain a Gelfand triple $V \hookrightarrow H \hookrightarrow V^*$, where V^* denotes the antidual of V. Let a be a bounded form. Then we set $\mathcal{A}\colon V \to V^*$, $\mathcal{A}u := a(u,\cdot)$ for $u \in V$, and define the operator A in H associated with a via $\mathrm{D}(A) := \mathcal{A}^{-1}(H) \subseteq H$ and $A := \mathcal{A}|_{\mathrm{D}(A)}$, i.e.,

$$a(u,v) = (\mathcal{A}u)(v) = (Au \,|\, v)_H\,, \quad (u \in \mathrm{D}(A), v \in V). \tag{3.2.1}$$

Additionally, suppose that a is coercive. Then the Lax–Milgram lemma yields that $\mathcal{A}$ is an isomorphism, whence the stationary problem is well-posed, i.e.,

$$\forall\, F \in V^*\, \exists!\, u \in V : a(u,v) = F(v) \quad (v \in V) \quad \text{and} \quad \|u\|_V \le \|\mathcal{A}^{-1}\| \|F\|_{V^*}.$$

Moreover, A is m-accretive and therefore $-A$ generates a contractive and (in case $\mathbb{K} = \mathbb{C}$) holomorphic C_0-semigroup $S\colon [0,\infty) \to \mathcal{L}(H)$. Thus, $u\colon [0,\infty) \to H$ is a solution of

$$u'(t) = -Au(t) \quad (t > 0), \quad u(0) = u_0 \in H$$

if and only if $u(t) = S(t)u_0$ for $t \ge 0$. Furthermore, A admits fractional powers and $0 \in \rho(A)$. We include a proof of this folklore statement for the reader's convenience.

Lemma 3.3. *Let a be a bounded and coercive form and A the associated operator. Then $0 \in \rho(A)$.*

Proof. Denote by $C_{V\hookrightarrow H}$ the embedding constant of the embedding $V \hookrightarrow H$. Let $\kappa > 0$ be the coercivity constant of a. Coercivity of a yields

$$\|Au\|_H \|u\|_H \ge \mathrm{Re}\,(Au \,|\, u)_H = \mathrm{Re}\, a(u) \ge \kappa \|u\|_V^2 \ge \frac{\kappa}{C_{V\hookrightarrow H}^2} \|u\|_H^2$$

for all $u \in \mathrm{D}(A)$. Hence, A is injective and has a bounded inverse.

Let $f \in H$. Since $H \hookrightarrow V^*$ and $\mathcal{A}$ is an isomorphism, there exists $u \in V$ such that $\mathcal{A}u = f$. Thus, by definition, $u = \mathcal{A}^{-1}f \in \mathrm{D}(A)$ and $Au = f$, so A is surjective. □

Remark 3.4. Let a be a bounded and coercive form and A the associated operator. Then, for $\alpha > 0$, also $0 \in \rho(A^\alpha)$ (see e.g. [60, Prop. 3.1.1(e)]). Thus, the graph norm $\|\cdot\|_{A^\alpha}$ of A^α and the norm $\|A^\alpha \cdot\|_H$ are equivalent on $\mathrm{D}(A^\alpha)$. Indeed, for $u \in \mathrm{D}(A^\alpha)$, we have

$$\|A^\alpha u\|_H \le \|u\|_{A^\alpha} = \|u\|_H + \|A^\alpha u\|_H \le \big(\|(A^\alpha)^{-1}\|_{\mathcal{L}(H)} + 1\big)\|A^\alpha u\|_H.$$

Let us now turn to approximation. We recall the notions of approximating sequences $(V_m)_{m\in\mathbb{N}}$ of V and space discretisations $(A_m)_{m\in\mathbb{N}}$ of A from Definitions 2.21 and 2.22, respectively. For an approximating sequence $(V_m)_{m\in\mathbb{N}}$ of V let $H_m \subseteq H$ be the closure of V_m in H for $m \in \mathbb{N}$. Then $V_m \hookrightarrow H_m$, and trivially, V_m is dense in H_m for all $m \in \mathbb{N}$. Note that typically (e.g., in applications), approximating sequences consist of finite-dimensional subspaces, and then $H_m = V_m$ as vector spaces.

Further, let $P_m\colon H \to H_m \subseteq H$ be the H-orthogonal projection of H onto H_m and $J_m\colon H_m \to H$ the canoncial embedding for $m \in \mathbb{N}$. Then $J_m P_m \to I$ strongly. Indeed, $(J_m P_m)_{m\in\mathbb{N}}$ is uniformly bounded and strong convergence for $f \in V$ holds by virtue of

$\|J_m P_m f - f\|_H = \|P_m f - f\|_H = \operatorname{dist}_H(f, H_m) \leq C_{V\hookrightarrow H} \operatorname{dist}_V(f, V_m) \to 0$. Strong convergence on H now follows from the uniform boundedness principle.

For $m \in \mathbb{N}$, we define the approximating forms $a_m := a|_{V_m \times V_m}$. Then a_m is trivially bounded and coercive for all $m \in \mathbb{N}$ with the same constants as a. Let $\mathcal{A}_m : V_m \to V_m^*$, $\mathcal{A}_m u := a_m(u, \cdot)$, and let A_m be the m-accretive operator in H_m associated with a_m for $m \in \mathbb{N}$. Further, let $S_m : [0, \infty) \to \mathcal{L}(H_m)$ be the contractive and (in case $\mathbb{K} = \mathbb{C}$) holomorphic C_0-semigroup generated by $-A_m$. Then for given initial data $u_{0,m} \in H_m$, the function $u_m : [0, \infty) \to H_m$ is a solution of

$$u_m'(t) = -A_m u_m(t) \quad (t > 0), \quad u_m(0) = u_{0,m}$$

if and only if $u_m(t) = S_m(t) u_{0,m}$ for $t \geq 0$. Thus, we obtain a space discretisation $(A_m)_{m\in\mathbb{N}}$ of A, which is called the *(Bubnov-)Galerkin discretisation.*

Coercivity of a implies that the uniform Banach–Nečas–Babuška condition is satisfied, which in turn yields convergence of the Galerkin discretisation; see also (BNB) in Section 2.2. Our aim is to quantify the order of convergence. Contrary to the quantified version of the Trotter–Kato theorem in Proposition 2.20, where conditions are formulated in terms of the resolvent, we derive conditions directly on the form a and the approximating sequence $(V_m)_{m\in\mathbb{N}}$. This quantifies the form convergence statement of [25, Thm. 3.7].

As a preparation for a convergence theorem, we make use of the following lemma, which can be found in [144, Thm. 2].

Lemma 3.5. *Let $a : V \times V \to \mathbb{K}$ be a bounded and coercive form with constants $M \geq 0$ and $\kappa > 0$. Let $(V_m)_{m\in\mathbb{N}}$ be an approximating sequence of V and let $a_m := a|_{V_m \times V_m}$ for $m \in \mathbb{N}$. For $F \in V^*$ let $u := \mathcal{A}^{-1} F \in V$ and $u_m := \mathcal{A}_m^{-1}(F|_{V_m}) \in V_m$ for all $m \in \mathbb{N}$. Then*

$$\|u - u_m\|_V \leq \frac{M}{\kappa} \operatorname{dist}_V(u, V_m) \quad (m \in \mathbb{N}).$$

In particular, $u_m \to u$ in V.

We consider two quantities relating the approximation speed of the approximating sequence and the properties of the operators $\mathcal{A}$ or $\mathcal{A}^*$.

Definition 3.6 (based on (6.6) in [6])**.** Let $\mathcal{X} \hookrightarrow V^*$ be a Banach space and $(V_m)_{m\in\mathbb{N}}$ an approximating sequence of V. For $m \in \mathbb{N}$, define

$$\gamma_m(\mathcal{X}) := \sup_{f\in\mathcal{X},\, \|f\|_{\mathcal{X}}=1} \operatorname{dist}_V(\mathcal{A}^{-1} f, V_m) \quad \text{and} \quad \gamma_m^*(\mathcal{X}) := \sup_{f\in\mathcal{X},\, \|f\|_{\mathcal{X}}=1} \operatorname{dist}_V((\mathcal{A}^*)^{-1} f, V_m).$$

We say that $(\gamma_m(\mathcal{X}))_{m\in\mathbb{N}}$ and $(\gamma_m^*(\mathcal{X}))_{m\in\mathbb{N}}$ *decay at rate* $p_1, p_2 \geq 0$, respectively, if there exist constants $C_{\gamma,\mathcal{X}}, C_{\gamma,\mathcal{X}}^* \geq 0$ such that

$$\gamma_m(\mathcal{X}) \leq \frac{C_{\gamma,\mathcal{X}}}{m^{p_1}} \quad \text{and} \quad \gamma_m^*(\mathcal{X}) \leq \frac{C_{\gamma,\mathcal{X}}^*}{m^{p_2}} \quad \text{for all} \quad m \in \mathbb{N}.$$

As an immediate consequence of this definition,

$$\operatorname{dist}_V(u, V_m) \leq \gamma_m(\mathcal{X}) \|\mathcal{A}u\|_{\mathcal{X}} \quad \text{for all} \quad u \in \mathcal{A}^{-1}\mathcal{X}.$$

We are now in a position to prove a convergence theorem including rate of convergence, starting with the stationary problem.

Theorem 3.7. *Let $a\colon V\times V\to\mathbb{C}$ be a bounded and coercive form and $(V_m)_{m\in\mathbb{N}}$ an approximating sequence of V. Let $\alpha>0$, $p_1,p_2\geq 0$ such that $(\gamma_m(\mathrm{D}(A^\alpha)))_{m\in\mathbb{N}}$ decays at rate $p_1=p_1(\alpha)$ and $(\gamma_m^*(H))_{m\in\mathbb{N}}$ decays at rate p_2. Then there exists a constant $C_\mathrm{x}=C_\mathrm{x}(\alpha)\geq 0$ such that*

$$\left\|J_mA_m^{-1}P_mf-A^{-1}f\right\|_H\leq C_\mathrm{x}\frac{\|f\|_{A^\alpha}}{m^{p_1+p_2}}.$$

for all $f\in\mathrm{D}(A^\alpha)$ and $m\in\mathbb{N}$. Moreover, if $p_1+p_2>0$, the space discretisation $(A_m)_{m\in\mathbb{N}}$ converges of order p_1+p_2 on $\mathrm{D}(A^\alpha)$ for the stationary problem.

Note that p_1 depends on α, and thus the order of convergence scales with the smoothness of the initial value encoded in the domain of powers of A.

The proof is essentially contained in [6, Thm. 6.3]. Note that [6] requires finite-dimensional subspaces for approximating sequences, which is not needed for [6, Thm. 6.3].

Proof. (i) Let $M,\kappa>0$ be the boundedness and coercivity constants of a. By Lemma 3.3, we have $0\in\rho(A)$. Analogously, $0\in\rho(A_m)$ for all $m\in\mathbb{N}$.

(ii) Fix $f\in\mathrm{D}(A^\alpha)$ and let $F\in V^*$ be defined by $F(v):=(f\,|\,v)_H$ $(v\in V)$. Let $u:=\mathcal{A}^{-1}F$ and $u_m:=\mathcal{A}_m^{-1}(F|_{V_m})$ for $m\in\mathbb{N}$. Then u is the unique solution of

$$a(u,v)=F(v)\quad(v\in V)$$

and, for $m\in\mathbb{N}$, u_m is the unique solution of

$$a_m(u_m,v)=F(v)=(F|_{V_m})(v)\quad(v\in V_m).$$

Then [6, Thm. 6.3] yields

$$\|u-u_m\|_H\leq\frac{M^2}{\kappa}\gamma_m(\mathrm{D}(A^\alpha))\gamma_m^*(H)\|f\|_{A^\alpha}\quad(m\in\mathbb{N}).$$

By assumption, there exist $C_{\gamma,\mathrm{D}(A^\alpha)},C_{\gamma,H}^*\geq 0$ such that

$$\gamma_m(\mathrm{D}(A^\alpha))\leq\frac{C_{\gamma,\mathrm{D}(A^\alpha)}}{m^{p_1}}\quad\text{and}\quad\gamma_m^*(H)\leq\frac{C_{\gamma,H}^*}{m^{p_2}}\quad\text{for all } m\in\mathbb{N}.$$

Thus,

$$\|u-u_m\|_H\leq\frac{M^2}{\kappa}C_{\gamma,\mathrm{D}(A^\alpha)}C_{\gamma,H}^*\frac{\|f\|_{A^\alpha}}{m^{p_1+p_2}}\quad(m\in\mathbb{N}).\tag{3.2.2}$$

For $m\in\mathbb{N}$, we define the resolvent difference Δ_m in H (restricted to $\mathrm{D}(A^\alpha)$) by

$$\mathrm{D}(\Delta_m):=\mathrm{D}(A^\alpha),\quad\Delta_mg:=J_mA_m^{-1}P_mg-A^{-1}g\quad(g\in\mathrm{D}(\Delta_m)).\tag{3.2.3}$$

Note that $A^{-1}f=\mathcal{A}^{-1}F=u$ and for $m\in\mathbb{N}$, we have $A_m^{-1}P_mf=u_m$. Indeed, given an orthonormal basis $\mathcal{B}_m$ of H_m, we observe

$$\begin{aligned}A_mu_m&=\sum_{\varphi\in\mathcal{B}_m}(A_mu_m\,|\,\varphi)_H\,\varphi=\sum_{\varphi\in\mathcal{B}_m}a_m(u_m,\varphi)\varphi=\sum_{\varphi\in\mathcal{B}_m}F(\varphi)\varphi\\&=\sum_{\varphi\in\mathcal{B}_m}(f\,|\,\varphi)_H\,\varphi=P_mf.\end{aligned}$$

In conclusion, $J_m A_m^{-1} P_m f = J_m u_m = u_m$ for all $m \in \mathbb{N}$ and thus $\Delta_m f = u_m - u$. By (3.2.2), we conclude

$$\|\Delta_m f\|_H = \|u - u_m\|_H \le \frac{M^2}{\kappa} C_{\gamma,\mathrm{D}(A^\alpha)} C^*_{\gamma,H} \cdot \frac{1}{m^{p_1+p_2}} \|f\|_{A^\alpha}. \qquad \square$$

We now turn to the evolution problem and make use of the quantified versions of the Trotter–Kato theorem in [74] (see also Proposition 2.20).

Theorem 3.8. *Let $a \colon V \times V \to \mathbb{C}$ be a bounded and coercive form and $(V_m)_{m\in\mathbb{N}}$ an approximating sequence of V. Let $\alpha > 0$, $p_1, p_2 \ge 0$ such that $(\gamma_m(\mathrm{D}(A^\alpha)))_{m\in\mathbb{N}}$ decays at rate $p_1 = p_1(\alpha)$ and $(\gamma_m^*(H))_{m\in\mathbb{N}}$ decays at rate p_2. Then for all $T \ge 0$ and $\varepsilon > 0$ there exists a constant $C_\mathrm{x} = C_\mathrm{x}(T,\varepsilon,\alpha) \ge 0$ such that*

$$\|J_m S_m(t) P_m u_0 - S(t)u_0\|_H \le C_\mathrm{x} \frac{\|u_0\|_{A^{\alpha+1+\varepsilon}}}{m^{p_1+p_2}}.$$

for all $t \in [0,T]$, $u_0 \in \mathrm{D}(A^{\alpha+1+\varepsilon})$, and $m \in \mathbb{N}$. Moreover, if $p_1 + p_2 > 0$, the space discretisation $(A_m)_{m\in\mathbb{N}}$ converges of order p_1+p_2 on $\mathrm{D}(A^{\alpha+1+\varepsilon})$ for the evolution problem.

Proof. (i) Let $M, \kappa > 0$ be the boundedness and coercivity constants of a. By Theorem 3.7, we have

$$\left\|(J_m A_m^{-1} P_m - A^{-1})f\right\|_H \le \frac{M^2}{\kappa} C_{\gamma,\mathrm{D}(A^\alpha)} C^*_{\gamma,H} \cdot \frac{\|f\|_{A^\alpha}}{m^{p_1+p_2}}$$

for all $f \in \mathrm{D}(A^\alpha)$.

(ii) Let $T \ge 0$, $\varepsilon > 0$, and Δ_m as defined in (3.2.3). Then by [74, Prop. 2.2b] (see also Proposition 2.20), there exists $C = C(T,\varepsilon,\alpha) \ge 0$ such that for all $t \in [0,T]$, $u_0 \in \mathrm{D}(A^{\alpha+1+\varepsilon})$, and $m \in \mathbb{N}$,

$$\begin{aligned}\|J_m S_m(t) P_m u_0 - S(t)u_0\|_H &\le C\|\Delta_m\|_{\mathcal{L}(\mathrm{D}(A^\alpha),H)} \|u_0\|_{A^{\alpha+1+\varepsilon}} \\ &\le \frac{M^2}{\kappa} C C_{\gamma,\mathrm{D}(A^\alpha)} C^*_{\gamma,H} \frac{\|u_0\|_{A^{\alpha+1+\varepsilon}}}{m^{p_1+p_2}},\end{aligned}$$

where we used (i) to estimate the norm of Δ_m. $\square$

Convergence under more general smoothness assumptions on the initial value u_0 can be obtained provided that a is symmetric.

Corollary 3.9. *Suppose that the assumptions from Theorem 3.8 hold, and additionally assume that a is symmetric. Then:*

(a) For all $T \ge 0$ there exists $C = C(T,\alpha)$ such that

$$\|J_m S_m(t) P_m u_0 - S(t)u_0\|_H \le C \frac{\|u_0\|_{A^{\alpha+1}}}{m^{p_1+p_2}}$$

for all $t \in [0,T]$, $u_0 \in \mathrm{D}(A^{\alpha+1})$, and $m \in \mathbb{N}$. Thus, the space discretisation $(A_m)_{m\in\mathbb{N}}$ converges of order $p_1 + p_2$ on $\mathrm{D}(A^{\alpha+1})$ for the evolution problem if $p_1 + p_2 > 0$.

(b) *For all $T \geq 1$ there exists $C = C(T, \alpha)$ such that*

$$\|J_m S_m(t) P_m u_0 - S(t) u_0\|_H \leq C \frac{\|u_0\|_{A^{\alpha+1/2}}}{m^{p_1+p_2}}$$

for all $t \in [(T)^{-1}, T]$, $u_0 \in \mathrm{D}(A^{\alpha+1/2})$, and $m \in \mathbb{N}$.

Proof. Note that symmetry of a implies symmetry of a_m and therefore self-adjointness of A_m for all $m \in \mathbb{N}$. The proof then uses [74, Prop. 2.3] instead of [74, Prop. 2.2b] in step (ii) of the proof of Theorem 3.8, see also Proposition 2.20. □

Remark 3.10. Theorem 3.8 and Corollary 3.9 actually state that convergence of the spatial discretisation for the stationary problem can be turned into convergence of the spatial discretisation for the evolution problem by slightly increasing the regularity required for the initial value.

3.2.2 Temporal approximation

Let X be a Banach space and $(S(t))_{t\geq 0}$ a C_0-semigroup on X with generator $-A$. Let $F\colon [0, \infty) \to \mathcal{L}(X)$ be a time discretisation method for $(S(t))_{t\geq 0}$ as introduced in Definition 2.23. We use the setting and notation of Section 2.2.

We first state a version of the well-known Chernoff product formula (cf. [118, Prop. 2.1]), which states that a stable and consistent time discretisation method is convergent.

Theorem 3.11 (Chernoff Product Fomula). *Let $F\colon [0, \infty) \to \mathcal{L}(X)$ be a stable and consistent time discretisation method for $(S(t))_{t\geq 0}$. Then F is convergent, and the convergence is uniform on compact time intervals.*

Note that the Chernoff product formula yields the qualitative statement of convergence. The statement can be quantified in the sense of consistency and convergence of some order $p_{\mathrm{t}} > 0$ as introduced in Definition 2.24. More precisely, stability and consistency of some order $p_{\mathrm{t}} > 0$ imply convergence of order p_{t} provided that the semigroup $(S(t))_{\geq 0}$ is exponentially bounded on Y_{t}. We include a proof of this folklore statement for the reader's convenience.

Proposition 3.12. *Let $Y_{\mathrm{t}} \hookrightarrow X$ with Y_{t} dense in X such that $Y_{\mathrm{t}} \subseteq \mathrm{D}(A)$ and Y_{t} is invariant under $(S(t))_{t\geq 0}$, and assume that S is exponentially bounded on Y_{t}. Let $F\colon [0, \infty) \to \mathcal{L}(X)$ be a stable time discretisation method that is consistent of order $p_{\mathrm{t}} > 0$ on Y_{t}. Then F is convergent of order p_{t} on Y_{t}.*

Proof. Let $M \geq 1$ and $\lambda \in \mathbb{R}$ such that for all $(\tau^i)_{1\leq i\leq N}$ in $[0, \infty)$ with $N \in \mathbb{N}$, we have $\|\prod_{i=1}^N F(\tau^i)\| \leq M \exp(\lambda \sum_{i=1}^N \tau^i)$ by stability of F and $\|S(t)\|_{\mathcal{L}(Y_{\mathrm{t}})} \leq M \mathrm{e}^{\lambda t}$ for all $t \geq 0$ by exponential boundedness of S.

Let $T \geq 0$. By consistency of order p_{t}, there exist $C \geq 0$ and $\tau_0 > 0$ such that for all $\tau \in [0, \tau_0]$ and $f \in Y_{\mathrm{t}}$, we have

$$\|F(\tau)f - S(\tau)f\| \leq C\tau^{p_{\mathrm{t}}+1}\|f\|_{Y_{\mathrm{t}}}.$$

Let $t \in [0, T]$, $N \in \mathbb{N}$, and $(\tau^i)_{1\leq i\leq N}$ in $[0, \min\{\tau_0, T\}]$ with $\sum_{i=1}^N \tau^i = t$. Let $f \in Y_{\mathrm{t}}$. Since

$$\prod_{i=1}^N F(\tau^i)f - S(t)f = F(\tau^N)\cdots F(\tau^1)f - S(\tau^N)\cdots S(\tau^1)f$$

$$= \sum_{i=1}^{N} \Big(\prod_{j=i+1}^{N} F(\tau^j) \Big) \big(F(\tau^i) - S(\tau^i) \big) \Big(\prod_{j=1}^{i-1} S(\tau^j) \Big) f,$$

we can estimate

$$\begin{aligned} \Big\| \prod_{i=1}^{N} F(\tau^i) f - S(t) f \Big\| &\leq \sum_{i=1}^{N} \Big\| \prod_{j=i+1}^{N} F(\tau^j) \Big\| \Big\| \big(F(\tau^i) - S(\tau^i) \big) \Big(\prod_{j=1}^{i-1} S(\tau^j) \Big) f \Big\| \\ &\leq \sum_{i=1}^{N} \Big\| \prod_{j=i+1}^{N} F(\tau^j) \Big\| C (\tau^i)^{p_{\mathrm{t}}+1} \Big\| \Big(\prod_{j=1}^{i-1} S(\tau^j) \Big) f \Big\|_{Y_{\mathrm{t}}} \\ &\leq M^2 C \Big(\max_{i=1,\dots,N} \tau^i \Big)^{p_{\mathrm{t}}} \Big(\sum_{i=1}^{N} \mathrm{e}^{\lambda \sum_{j=i+1}^{N} \tau^j} \cdot \tau^i \cdot \mathrm{e}^{\lambda \sum_{j=1}^{i-1} \tau^j} \Big) \|f\|_{Y_{\mathrm{t}}} \\ &\leq M^2 C t \mathrm{e}^{\max\{\lambda, 0\} t} \Big(\max_{i=1,\dots,N} \tau^i \Big)^{p_{\mathrm{t}}} \|f\|_{Y_{\mathrm{t}}}. \end{aligned}$$

□

So far, the time steps τ^i may depend on the particular value $t \in [0, T]$. In applications, one typically works with one family of $N_k \in \mathbb{N}$ time steps $\tau_k = (\tau_k^i)_{1 \leq i \leq N_k}$ and only considers times generated by the grid of these time steps. To obtain a finer discretisation, N_k is increased, which corresponds to larger parameters $k \in \mathbb{N}$. That is, for $T \geq 0$ and $k, N_k \in \mathbb{N}$ a family of *time steps* is a family $\tau_k = (\tau_k^i)_{1 \leq i \leq N_k} \in [0, T]^{N_k}$ such that $\tau_k \to 0$ and $N_k \to \infty$ as $k \to \infty$ as well as $\sum_{i=1}^{N_k} \tau_k^i = T$. Then $\mathcal{T}_k \coloneqq \{ \sum_{i=1}^{j} \tau_k^i : j \in \{0, \dots, N_k\} \}$ is called the *grid* associated with τ_k. Moreover, for a time discretisation method $F \colon [0, \infty) \to \mathcal{L}(X)$ and $t \in \mathcal{T}_k$, we write $F_k(t) \coloneqq \prod_{i=1}^{N_{t,k}} F(\tau_k^i)$, where $N_{t,k} \coloneqq \max\{ j \in \{0, \dots, N_k\} : \sum_{i=1}^{j} \tau_k^i = t \}$.

Proposition 3.13. *Let $Y_{\mathrm{t}} \hookrightarrow X$ and $F \colon [0, \infty) \to \mathcal{L}(X)$ be a convergent time discretisation method of order $p_{\mathrm{t}} > 0$ on Y_{t}. Let $T \geq 0$ and for $k, N_k \in \mathbb{N}$ let $\mathcal{T}_k$ be the grid associated with time steps $\tau_k = (\tau_k^i)_{1 \leq i \leq N_k}$. Then for all $k \in \mathbb{N}$, there exist $C_T \geq 0$ and $\tau_0 = \tau_0(T) > 0$ such that if $\max_{i=1,\dots,N_k} \tau_k^i \leq \tau_0$ then*

$$\max_{t \in \mathcal{T}_k} \|F_k(t) f - S(t) f\| \leq C_T \Big(\max_{i=1,\dots,N_k} \tau_k^i \Big)^{p_{\mathrm{t}}} \|f\|_{Y_{\mathrm{t}}} \quad (f \in Y_{\mathrm{t}}).$$

Proof. By convergence of F, for all $T > 0$, there exist $C_T \geq 0$ and $\tau_0 \in (0, T]$ such that for all $t \in [0, \tau_0]$, $N \in \mathbb{N}$, and all $(\tilde{\tau}^i)_{1 \leq i \leq N}$ in $[0, \tau_0]$ with $\sum_{i=1}^{N} \tilde{\tau}^i = t$, we have

$$\Big\| \prod_{i=1}^{N} F(\tilde{\tau}^i) f - S(t) f \Big\| \leq C_T \Big(\max_{i=1,\dots,N} \tilde{\tau}^i \Big)^{p_{\mathrm{t}}} \|f\|_{Y_{\mathrm{t}}} \quad (f \in Y_{\mathrm{t}}).$$

Since $t \in \mathcal{T}_k$, we have $t = \sum_{i=1}^{N_{t,k}} \tau_k^i$ and clearly, $\max_{i=1,\dots,N_{t,k}} \tau_k^i \leq \max_{i=1,\dots,N_k} \tau_k^i$. Thus, the assertion follows with from the above with $N = N_{t,k}$ and $\tilde{\tau}^i = \tau_k^i$ for $1 \leq i \leq N_{t,k}$. □

Lastly, we recall from Definition 2.27 the class of rational methods that we are considering in the following. Examples of rational schemes, including the implicit Euler method and Crank–Nicolson, are given in the preliminaries; cf. Example 2.26.

Definition 3.14. Let $r\colon \mathbb{C} \to \mathbb{C} \cup \{\infty\}$ be a rational function. We say that a time discretisation method F for $(S(t))_{t\geq 0}$ is *induced by* r if $F(\tau) = r(\tau\cdot(-A))$ for all sufficiently small $\tau > 0$. Such schemes F are called *rational schemes*. If $|r(z)| \leq 1$ for all $z \in \mathbb{C}$ with $\operatorname{Re} z \leq 0$ and $r(z) - \mathrm{e}^z = o(z)$ as $z \to 0$ then r is called *A-acceptable*.

Numerous classical references have discussed the approximation of semigroups by rational schemes, such as [21, 35, 65], which has later been treated by a unified approach in [54, 55]. We collect some key results in the following theorem taken from [128, Thm. 5.6] based on results from [21, Thm. 4] and interpolating [99, Thm. 4.2] using the stability result [35, Thm. 5].

Theorem 3.15 (Theorem 5.6 in [128])**.** *Let $r\colon \mathbb{C} \to \mathbb{C} \cup \{\infty\}$ be an A-acceptable rational function and assume that there exists an integer $p \geq 1$ such that $|r(z) - \mathrm{e}^z| = \mathcal{O}(z^{p+1})$ as $z \to 0$. Let $-A$ be the generator of a bounded C_0-semigroup $(S(t))_{t\geq 0}$ on a Banach space X and let F be induced by r. Then F approximates S to order $\eta(p,k)$ on $\mathrm{D}(A^k)$ for all integers $k \in \{1, \dots, p+1\}$ with $k \neq \frac{p+1}{2}$, where*

$$\eta(p,k) := \begin{cases} k - \frac{1}{2}, & k < \frac{p+1}{2}; \\ \frac{kp}{p+1}, & \frac{p+1}{2} < k \leq p+1. \end{cases}$$

If the semigroup is analytic and bounded on a sector, then F approximates S to order α on $\mathrm{D}(A^\alpha)$ for all $\alpha \in (0, p]$.

An extension to domains of fractional powers and interpolation spaces can be found in [90], see also the rates for implicit Euler and Crank–Nicolson in Example 2.26.

3.2.3 Randomness approximation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $N \in \mathbb{N}$, and $Z = (Z_0, \dots, Z_{N-1})\colon \Omega \to \mathbb{R}^N$ a random variable. For $n \in \mathbb{N}_0$, let

$$\mathcal{P}_n^N := \left\{ p\colon \mathbb{R}^N \to \mathbb{K} :\ \exists (c_\alpha)_{|\alpha|\leq n} \subseteq \mathbb{K} :\ p(x) = \sum_{\alpha\in\mathbb{N}_0^N, |\alpha|\leq n} c_\alpha x^\alpha \quad (x \in \mathbb{R}^N) \right\},$$

be the vector space of polynomials in N variables of order at most n. Note that $\dim \mathcal{P}_n^N = \binom{n+N}{n}$. Moreover, let $\mathcal{P}^N := \bigcup_{n\in\mathbb{N}_0} \mathcal{P}_n^N$ be the vector space of polynomials. We will assume that $\mathcal{P}^N \subseteq L^2(\mathbb{R}^N, \mathbb{P}_Z)$ is dense, i.e., $\mathbb{P}_Z$ has moments of all orders. Then $(\mathcal{P}_n^N)_{n\in\mathbb{N}_0}$ is an approximating sequence of $L^2(\mathbb{R}^N, \mathbb{P}_Z)$.

Definition 3.16. For $n \in \mathbb{N}_0$, let $R_n\colon L^2(\mathbb{R}^N, \mathbb{P}_Z) \to \mathcal{P}_n^N \subseteq L^2(\mathbb{R}^N, \mathbb{P}_Z)$ be the orthogonal projection. Then $R_n f$ is called the *polynomial chaos approximation*, or, for short, *PC approximation*, of $f \in L^2(\mathbb{R}^N, \mathbb{P}_Z)$ of *order* n.

Remark 3.17. Let $f \in L^2(\mathbb{R}^N, \mathbb{P}_Z)$ and $n \in \mathbb{N}_0$. Then $R_n f$ is the best approximation of f in $\mathcal{P}_n^N$. Therefore, $\|f - R_n f\|_{L^2(\mathbb{R}^N, \mathbb{P}_Z)} = \operatorname{dist}(f, \mathcal{P}_n^N) \to 0$ as $n \to \infty$.

Let $(Z_0, \dots, Z_{N-1})$ be independent. Then

$$L^2(\mathbb{R}^N, \mathbb{P}_Z) = \bigotimes_{j=0}^{N-1} L^2(\mathbb{R}, \mathbb{P}_{Z_j}).$$

By assumption, $\mathcal{P}^1$ is dense in $L^2(\mathbb{R}, \mathbb{P}_{Z_j})$, so we can construct an orthogonal basis $(\Phi_n^j)_{n\in\mathcal{N}_j}$ of polynomials $\Phi_n^j \in \mathcal{P}_n^1$ of degree $\deg(\Phi_n^j) = n$. Here, the index set $\mathcal{N}_j = \{0, \dots, L_j\}$ for some $L_j \in \mathbb{N}_0$ in the case of a discrete probability measure with finite support and $\mathcal{N}_j = \mathbb{N}_0$ otherwise. Due to the assumptions we will impose on the distribution of each Z_j in the following, we can restrict our considerations to the case $\mathcal{N}_j = \mathbb{N}_0$ for all $0 \leq j \leq N-1$. Let $\mathcal{N} := \mathcal{N}_0 \times \ldots \times \mathcal{N}_{N-1} = \mathbb{N}_0^N$. For $\alpha \in \mathcal{N}$, let $\Phi_\alpha := \bigotimes_{j=0}^{N-1} \Phi_{\alpha_j}^j$. Then $(\Phi_\alpha)_{\alpha\in\mathcal{N}}$ is an orthogonal basis of $L^2(\mathbb{R}^N, \mathbb{P}_Z)$. Thus, it suffices to consider the one-dimensional case $N = 1$ and then use a tensor product construction.

Remark 3.18. Let $(\Phi_\alpha)_{\alpha\in\mathcal{N}}$ be an orthogonal basis of $L^2(\mathbb{R}^N, \mathbb{P}_Z)$ such that $\Phi_\alpha \in \mathcal{P}_n^N$ for $|\alpha| \leq n$, $n \in \mathbb{N}_0$. Then $(\Phi_\alpha)_{|\alpha|\leq n}$ is an orthogonal basis of the closed subspace $\mathcal{P}_n^N \subseteq L^2(\mathbb{R}^N, \mathbb{P}_Z)$, and hence for $f \in L^2(\mathbb{R}^N, \mathbb{P}_Z)$ and $n \in \mathbb{N}_0$, we have

$$R_n f = \sum_{|\alpha|\leq n} \frac{1}{\|\Phi_\alpha\|^2_{L^2(\mathbb{R}^N,\mathbb{P}_Z)}} \left(f \mid \Phi_\alpha\right)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)} \Phi_\alpha = \sum_{|\alpha|\leq n} \hat{f}_\alpha \Phi_\alpha,$$

where $\hat{f}_\alpha := \|\Phi_\alpha\|^{-2}_{L_2(\mathbb{R}^N,\mathbb{P}_Z)} \left(f \mid \Phi_\alpha\right)_{L_2(\mathbb{R}^N,\mathbb{P}_Z)}$ is the generalised Fourier coefficient of f for $\alpha \in \mathcal{N}$, sometimes also called *PCE coefficient*. The series representation

$$f = \sum_{\alpha\in\mathcal{N}} \frac{1}{\|\Phi_\alpha\|^2_{L^2(\mathbb{R}^N,\mathbb{P}_Z)}} \left(f \mid \Phi_\alpha\right)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)} \Phi_\alpha = \sum_{\alpha\in\mathcal{N}} \hat{f}_\alpha \Phi_\alpha$$

converges in $L^2(\mathbb{R}^N, \mathbb{P}_Z)$ and is referred to as the *polynomial chaos expansion (PCE).*

We list the orthogonal polynomials corresponding to common distributions, such as the standard normal distribution.

Example 3.19 (Classical orthogonal polynomials)**.** *Let $N = 1$.*

(a) Let Z be standard normally distributed, i.e. $\mathbb{P}_Z(B) = \int_B (2\pi)^{-1/2} \mathrm{e}^{-z^2/2} \mathrm{d}z$ for $B \in \mathcal{B}(\mathbb{R})$. The (one-dimensional) Hermite polynomials *are defined as*

$$H_k(z) := (-1)^k \mathrm{e}^{\frac{z^2}{2}} \frac{\mathrm{d}^k}{\mathrm{d}z^k} \mathrm{e}^{-\frac{z^2}{2}} \quad (z \in \mathbb{R}, k \in \mathbb{N}_0).$$

Then $(H_k)_{k\in\mathbb{N}_0}$ is an orthogonal basis of $L^2(\mathbb{R}, \mathbb{P}_Z)$ and $\operatorname{supp} \mathbb{P}_Z = \mathbb{R}$.

(b) Let $\alpha, \beta > -1$ and Z be Beta-distributed with parameters α and β, i.e.,

$$\mathbb{P}_Z(B) = \int_B \mathbf{1}_{(-1,1)}(z) \frac{\Gamma(\alpha+\beta+2)}{2^{\alpha+\beta+1}\Gamma(\alpha+1)\Gamma(\beta+1)} (1-z)^\alpha (1+z)^\beta \mathrm{d}z$$

for $B \in \mathcal{B}(\mathbb{R})$. The (one-dimensional) Jacobi polynomials *are defined as*

$$J_k(z) := J_k^{(\alpha,\beta)}(z) := \frac{(-1)^k}{2^k k!} (1-z)^{-\alpha} (1+z)^{-\beta} \frac{\mathrm{d}^k}{\mathrm{d}z^k} \left[(1-z)^{k+\alpha} (1+z)^{k+\beta}\right]$$

for $z \in (-1,1)$ and $k \in \mathbb{N}_0$. Then $(J_k)_{k\in\mathbb{N}_0}$ is an orthogonal basis of $L^2(\mathbb{R}, \mathbb{P}_Z)$ and $\operatorname{supp} \mathbb{P}_Z = [-1, 1]$. *Note that $\alpha = \beta$ yields* Gegenbauer polynomials, *$\alpha = \beta = 0$ yields* Legendre polynomials *corresponding to Z being uniformly distributed on* $[-1, 1]$, *and $\alpha = \beta = -\frac{1}{2}$ yields* Chebyshev polynomials.

(c) *Let $\alpha > -1$ and Z be Gamma-distributed with shape parameter $\alpha + 1$ and rate 1, i.e. $\mathbb{P}_Z(B) = \int_B \mathbf{1}_{(0,\infty)}(z)\Gamma(\alpha+1)^{-1} z^\alpha \mathrm{e}^{-z} \mathrm{d}z$ for $B \in \mathcal{B}(\mathbb{R})$. The (one-dimensional)* Laguerre polynomials *are defined as*

$$L_k(z) \coloneqq L_k^{(\alpha)}(z) \coloneqq \frac{z^{-\alpha}\mathrm{e}^z}{k!}\frac{\mathrm{d}^k}{\mathrm{d}z^k}(z^{k+\alpha}\mathrm{e}^{-z}) \quad (z > 0, k \in \mathbb{N}_0).$$

Then $(L_k)_{k\in\mathbb{N}_0}$ is an orthogonal basis of $L^2(\mathbb{R}, \mathbb{P}_Z)$ and $\operatorname{supp}\mathbb{P}_Z = [0,\infty)$. Note that the special case $\alpha = 0$ corresponds to Z being exponentially distributed with rate 1.

Note that, by [102], up to affine transformations, these three distributions are exactly the probability distributions whose corresponding orthogonal polynomials yield an orthogonal basis of eigenfunctions of a Sturm–Liouville differential operator. As these operators are essential for the proof of error estimates for the PC approximation, we assume the following.

Assumption 3.20. *Let $N \in \mathbb{N}$ and $Z = (Z_0, \ldots, Z_{N-1})\colon \Omega \to \mathbb{R}^N$ be a random variable. Let $(Z_0, \ldots, Z_{N-1})$ be independent and assume that each of them is either standard normally distributed, Gamma-distributed (with rate 1), or Beta-distributed, explicitly allowing for different Z_j to have a different distribution.*

The one-dimensional Sturm–Liouville operators for the respective orthogonal polynomials associated with the distributions are the following.

Example 3.21. *Let $N = 1$.*

(a) *For $k \in \mathbb{N}_0$, the Hermite polynomial H_k is an eigenfunction of*

$$Q = -\frac{\mathrm{d}^2}{\mathrm{d}z^2} + z\frac{\mathrm{d}}{\mathrm{d}z}$$

to the eigenvalue $\lambda_k = k$.

(b) *For $k \in \mathbb{N}_0$, the Jacobi polynomial $J_k = J_k^{(\alpha,\beta)}$ is an eigenfunction of*

$$Q = Q^{(\alpha,\beta)} = -(1-z^2)\frac{\mathrm{d}^2}{\mathrm{d}z^2} + (\alpha - \beta + (\alpha+\beta+2)z)\frac{\mathrm{d}}{\mathrm{d}z}$$

to the eigenvalue $\lambda_k = k(k+\alpha+\beta+1)$.

(c) *For $k \in \mathbb{N}_0$, the Laguerre polynomial $L_k = L_k^{(\alpha)}$ is an eigenfunction of*

$$Q = Q^{(\alpha)} = -z\frac{\mathrm{d}^2}{\mathrm{d}z^2} + (z - \alpha - 1)\frac{\mathrm{d}}{\mathrm{d}z}$$

to the eigenvalue $\lambda_k = k$.

Definition 3.22. Let $Y_\mathrm{z} \hookrightarrow L^2(\mathbb{R}^N, \mathbb{P}_Z)$, $p_\mathrm{z} > 0$. The polynomial chaos approximation is said to be *convergent of order* p_z on Y_z if there exists a constant $C_\mathrm{z} \geq 0$ such that

$$\|R_n f - f\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z)} \leq C_\mathrm{z} n^{-p_\mathrm{z}} \|f\|_{Y_\mathrm{z}}$$

for all $f \in Y_\mathrm{z}$ and $n \in \mathbb{N}_0$.

In order to obtain convergence orders, we need sufficient regularity of the corresponding function. For $\ell \in \mathbb{N}_0$, we recall the Sobolev spaces $H^\ell(\mathbb{R}^N, \mathbb{P}_Z)$ consisting of all $f \in L^2(\mathbb{R}^N, \mathbb{P}_Z)$ such that the distributional derivatives up to order ℓ belong to $L^2(\mathbb{R}^N, \mathbb{P}_Z)$. We recall the one-dimensional convergence orders first and write $\mathbf{z}$ for the function $z \mapsto z$.

Proposition 3.23 (Theorems 6.2.4-6 in [50]). *Let $N = 1$ and $\ell \in \mathbb{N}_0$.*

(a) Let Z be standard normally distributed. Then there exists $C \geq 0$ such that

$$\|f - R_n f\|_{L^2(\mathbb{R},\mathbb{P}_Z)} \leq C n^{-\ell} \left\| \frac{\mathrm{d}^{2\ell}}{\mathrm{d}z^{2\ell}} f \right\|_{L^2(\mathbb{R},\mathbb{P}_Z)}$$

for all $f \in H^{2\ell}(\mathbb{R}, \mathbb{P}_Z)$ and $n > 2\ell$.

(b) Let Z be Beta-distributed. Then there exists $C \geq 0$ such that

$$\|f - R_n f\|_{L^2(\mathbb{R},\mathbb{P}_Z)} \leq C n^{-\ell} \left\| (1-\mathbf{z}^2)^{\ell/2} \frac{\mathrm{d}^{\ell}}{\mathrm{d}z^{\ell}} f \right\|_{L^2(\mathbb{R},\mathbb{P}_Z)} \leq C n^{-\ell} \left\| \frac{\mathrm{d}^{\ell}}{\mathrm{d}z^{\ell}} f \right\|_{L^2(\mathbb{R},\mathbb{P}_Z)}$$

for all $f \in H^{\ell}(\mathbb{R}, \mathbb{P}_Z)$ and $n > \ell$.

(c) Let Z be Gamma-distributed (with rate 1). Then there exists $C \geq 0$ such that

$$\|f - R_n f\|_{L^2(\mathbb{R},\mathbb{P}_Z)} \leq C n^{-\ell} \left\| \mathbf{z}^{\ell} \frac{\mathrm{d}^{2\ell}}{\mathrm{d}z^{2\ell}} f \right\|_{L^2(\mathbb{R},\mathbb{P}_Z)}$$

for all $f \in L^2(\mathbb{R}, \mathbb{P}_Z)$ being weakly differentiable up to order 2ℓ such that $\mathbf{z}^{k/2} \frac{\mathrm{d}^k}{\mathrm{d}z^k} f \in L^2(\mathbb{R}, \mathbb{P}_Z)$ for all $k \in \{0, \dots, 2\ell\}$, and $n > 2\ell$.

Remark 3.24. Note that the statements in Proposition 3.23 can be formulated for all $n \in \mathbb{N}_0$ as well by adjusting the constant C, if necessary.

We can now prove an error estimate yielding an order of convergence for the PC approximation. Denote by $J_{\text{Normal}}, J_{\text{Gamma}}, J_{\text{Beta}}$ the set of all those $j \in \{0, \dots, N-1\}$ such that Z_j is standard normally distributed, Gamma-distributed, or Beta-distributed, respectively. Motivated by the norms appearing in the error estimates of Proposition 3.23, let $\rho \colon \mathbb{R}^N \to \mathbb{R}^N$ be defined by

$$\rho(z)_j := \begin{cases} 1 & j \in J_{\text{Normal}} \cup J_{\text{Beta}}, \\ z_j & j \in J_{\text{Gamma}}. \end{cases} \tag{3.2.4}$$

For $\ell \in \mathbb{N}_0$, we define $H^\ell_\rho(\mathbb{R}^N, \mathbb{P}_Z)$ to be the set of all $f \in L^2(\mathbb{R}^N, \mathbb{P}_Z)$ such that f is weakly differentiable up to order ℓ and $\rho(\cdot)^{\alpha/2} \partial^\alpha f \in L^2(\mathbb{R}^N, \mathbb{P}_Z)$ for all $\alpha \in \mathcal{N}$ with $|\alpha| \leq \ell$. Here, $\rho(z)^{\alpha/2} = \prod_{j=0}^{N-1} \rho(z)_j^{\alpha_j/2} \in \mathbb{R}$ for all $z \in \mathbb{R}^N$, $\alpha \in \mathcal{N}$. We equip $H^\ell_\rho(\mathbb{R}^N, \mathbb{P}_Z)$ with the norm $\|\cdot\|_{H^\ell_\rho(\mathbb{R}^N,\mathbb{P}_Z)}$ given by

$$\|f\|^2_{H^\ell_\rho(\mathbb{R}^N,\mathbb{P}_Z)} := \sum_{|\alpha| \leq \ell} \left\| \rho(\cdot)^{\alpha/2} \partial^\alpha f \right\|^2_{L_2(\mathbb{R}^N,\mathbb{P}_Z)},$$

which makes $H_\rho^\ell(\mathbb{R}^N, \mathbb{P}_Z)$ a Hilbert space. Here and in the following, when summing over $|\alpha| \leq \ell$, the index α is assumed to be in $\mathcal{N} = \mathbb{N}_0^N$. In case there are no Gamma-distributed components, the usual Sobolev spaces $H_\rho^\ell(\mathbb{R}^N, \mathbb{P}_Z) = H^\ell(\mathbb{R}^N, \mathbb{P}_Z)$ are recovered. The weighted Sobolev spaces are those spaces $Y_\mathbf{z}$ on which the PC approximation converges of a certain order, as stated in the following main result of this subsection for the scalar-valued case. It extends [138, Thm. 3.6] to the multi-dimensional case (i.e., from $\mathbb{R}$ to $\mathbb{R}^N$) with explicit constants and additionally covers Beta- and Gamma-distribution.

Theorem 3.25. *Suppose that Z satisfies Assumption 3.20. Let $\ell \in \mathbb{N}_0$. Then*

$$\|R_n f - f\|_{L^2(\mathbb{R}^N, \mathbb{P}_Z)} \leq C_{\ell,N} n^{-\ell} \|f\|_{H_\rho^{2\ell}(\mathbb{R}^N, \mathbb{P}_Z)}$$

for all $f \in H_\rho^{2\ell}(\mathbb{R}^N, \mathbb{P}_Z)$ and $n \in \mathbb{N}_0$ with constant

$$C_{\ell,N} := N^{\ell/2} \prod_{r=0}^{\ell-1} \Big(\max_{0 \leq j \leq N-1} C_j(2r) \Big),$$

where $C_j(2r)$ is as defined in Proposition 3.29. In particular, the polynomial chaos approximation is convergent of order ℓ on $H_\rho^{2\ell}(\mathbb{R}^N, \mathbb{P}_Z)$ given that $\ell > 0$.

For the proof of Theorem 3.25 we need some auxiliary statements. Let Q_j be the Sturm–Liouville differential operator associated with the orthogonal polynomials corresponding to $\mathbb{P}_{Z_j}$ for $j \in \{0, \ldots, N-1\}$ and denote its k-th eigenvalue by λ_k^j for $k \in \mathbb{N}_0$; cf. Example 3.21. In higher dimension, we consider $Q = \sum_{j=0}^{N-1} \widetilde{Q}_j$ with $\widetilde{Q}_j := I \otimes \ldots \otimes I \otimes Q_j \otimes I \otimes \ldots \otimes I$, where Q_j acts on the j-th component.

Proposition 3.29 establishes a quantitative estimate of the ℓ-th Sobolev norm of Qf in terms of the $(\ell+2)$-th Sobolev norm of f based on the corresponding one-dimensional estimates of Lemmas 3.26-3.28. Having rewritten the scalar product of some function f with the orthogonal polynomial Φ_α in terms of the scalar product of $Q^\ell f$ with $\Phi^\alpha f$ in Lemma 3.30, we can use the Sobolev norm estimates to finish the proof of the PC error bound. The reader mostly interested in the order of convergence rather than explicit constants is advised to continue reading at Lemma 3.30.

First, Sobolev norms of Qf are estimated in dimension $N = 1$ for standard normal distribution, Beta distribution, and Gamma distribution in Lemmas 3.26, 3.27, and 3.28, respectively.

Lemma 3.26. *Let $w(z) := \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-z^2/2}$ for $z \in \mathbb{R}$. Let $Q := -\partial^2 + z\partial$ and $\ell \in \mathbb{N}_0$. Then for $f \in H^{\ell+2}(\mathbb{R}, w)$, we have $Qf \in H^\ell(\mathbb{R}, w)$ and*

$$\|Qf\|_{H^\ell(\mathbb{R},w)} \leq \sqrt{21 + 3\ell^2} \|f\|_{H^{\ell+2}(\mathbb{R},w)}.$$

Note that if Z is standard normally distributed then w is the density of $\mathbb{P}_Z$.

Proof. Note that for $g \in H^1(\mathbb{R}, w)$, we have $\mathbf{z}g \in L^2(\mathbb{R}, w)$ and

$$\|\mathbf{z}g\|_{L^2(\mathbb{R},w)}^2 \leq 2\|g\|_{L^2(\mathbb{R},w)}^2 + 4\|g'\|_{L^2(\mathbb{R},w)}^2,$$

see, e.g., [106, Lem. 2.1]. Thus, for $f \in H^{\ell+2}(\mathbb{R},w)$, we have $Qf = -f'' + \mathbf{z}f' \in H^{\ell}(\mathbb{R},w)$. A simple calculation shows that $\partial^k Qf = -\partial^{k+2}f + \mathbf{z}\partial^{k+1}f + k\partial^k f$ for $k \in \mathbb{N}_0$. For $f \in H^{\ell+2}(\mathbb{R},w)$, we deduce from the above estimate that

$$\begin{aligned}\|Qf\|^2_{H^\ell(\mathbb{R},w)} &= \sum_{k=0}^{\ell} \big\|\partial^{k+2}f - \mathbf{z}\partial^{k+1}f - k\partial^k f\big\|^2_{L^2(\mathbb{R},w)} \\ &\le \sum_{k=0}^{\ell} 3\Big(5\big\|\partial^{k+2}f\big\|^2_{L^2(\mathbb{R},w)} + 2\big\|\partial^{k+1}f\big\|^2_{L^2(\mathbb{R},w)} + k^2\big\|\partial^k f\big\|^2_{L^2(\mathbb{R},w)}\Big) \\ &\le \big(21+3\ell^2\big)\|f\|^2_{H^{\ell+2}(\mathbb{R},w)}. \qquad \square\end{aligned}$$

Lemma 3.27. *Let $\alpha,\beta > -1$ and*

$$w(z) := \mathbf{1}_{(-1,1)}(z)\frac{\Gamma(\alpha+\beta+2)}{2^{\alpha+\beta+1}\Gamma(\alpha+1)\Gamma(\beta+1)}(1-z)^\alpha(1+z)^\beta \quad (z\in\mathbb{R}).$$

Let $\ell \in \mathbb{N}_0$, $Q := -(1-z^2)\partial^2 + (\alpha-\beta+(\alpha+\beta+2)z)\partial$, $\ell\in\mathbb{N}_0$, and $C_{\alpha,\beta} := \alpha+\beta+1$. Then for $f \in H^{\ell+2}(\mathbb{R},w)$, we have $Qf \in H^\ell(\mathbb{R},w)$ and

$$\|Qf\|_{H^\ell(\mathbb{R},w)} \le \sqrt{3\big(1+4(\ell+1+\max\{\alpha,\beta\})^2+\ell^2(\ell+C_{\alpha,\beta})^2\big)}\,\|f\|_{H^{\ell+2}(\mathbb{R},w)}.$$

Note that if Z is Beta-distributed then w is the density of $\mathbb{P}_Z$.

Proof. As in the proof of Lemma 3.26, we calculate the k-th derivatives of Q. Taking the bounded interval $\operatorname{supp} w = [-1,1]$ and the eigenvalues $\lambda_k = k(k+\alpha+\beta+1)$ of Q given in Example 3.21(b) into account, for $f\in H^{\ell+2}(\mathbb{R},w)$ we deduce

$$\begin{aligned}\|Qf\|^2_{H^\ell(\mathbb{R},w)} &= \sum_{k=0}^{\ell}\big\|(\mathbf{z}^2-1)\partial^{k+2}f + (\alpha-\beta+C_{\alpha,\beta,k}\mathbf{z})\partial^{k+1}f + \lambda_k\partial^k f\big\|^2_{L^2(\mathbb{R},w)} \\ &\le \sum_{k=0}^{\ell} 3\Big(\|\partial^{k+2}f\|^2_{L^2(\mathbb{R},w)} + 4(k+1+M_{\alpha,\beta})^2\|\partial^{k+1}f\|^2_{L^2(\mathbb{R},w)} + \lambda_k^2\|\partial^k f\|^2_{L^2(\mathbb{R},w)}\Big) \\ &\le \big(3+12(\ell+1+M_{\alpha,\beta})^2+3\ell^2(\ell+\alpha+\beta+1)^2\big)\|f\|^2_{H^{\ell+2}(\mathbb{R},w)},\end{aligned}$$

where we have abbreviated $C_{\alpha,\beta,k} := \alpha+\beta+2k+2$ and $M_{\alpha,\beta} := \max\{\alpha,\beta\}$. $\square$

Lemma 3.28. *Let $\alpha > -1$, $w(z) := \mathbf{1}_{(0,\infty)}(z)\frac{1}{\Gamma(\alpha+1)}z^\alpha \mathrm{e}^{-z}$, and $\rho(z) := \mathbf{1}_{(0,\infty)}(z)z$ for $z\in\mathbb{R}$. Let $Q := -z\partial^2 + (z-\alpha-1)\partial$ and $\ell\in\mathbb{N}_0$. Then for $f\in H^{\ell+2}_\rho(\mathbb{R},w)$, we have $Qf \in H^\ell_\rho(\mathbb{R},w)$ and*

$$\|Qf\|_{H^\ell_\rho(\mathbb{R},w)} \le \sqrt{24\alpha+87+24\ell+3\ell^2}\,\|f\|_{H^{\ell+2}_\rho(\mathbb{R},w)}.$$

Note that if Z is Gamma-distributed then w is the density of $\mathbb{P}_Z$.

Proof. Let $f \in H^{\ell+2}_\rho(\mathbb{R},w)$. The k-th derivative of Qf contains, among others, terms of the form $\rho^{k/2}\mathbf{z}\partial^{k+1}f$ as well as $\rho^{k/2}\mathbf{z}\partial^{k+1}f$, which we have to estimate by $\|f\|_{H^{\ell+2}_\rho(\mathbb{R},w)}$. First, for $k\in\{0,\dots,\ell\}$, $\rho^{k/2}\mathbf{z}\partial^{k+1}f \in L^2(\mathbb{R},w)$ and

$$\big\|\rho^{k/2}\mathbf{z}\partial^{k+1}f\big\|^2_{L^2(\mathbb{R},w)} \le 2C_{\alpha,k}\big\|\rho^{(k+1)/2}\partial^{k+1}f\big\|^2_{L^2(\mathbb{R},w)} + 4\big\|\rho^{(k+2)/2}\partial^{k+2}f\big\|^2_{L^2(\mathbb{R},w)}$$

with $C_{\alpha,k} \coloneqq (\alpha+2+k)$. Indeed, we compute by integration by parts

$$\begin{aligned}\int_0^\infty z^k z^2 \big|\partial^{k+1} f(z)\big|^2 z^\alpha \mathrm{e}^{-z}\,\mathrm{d}z &= \int_0^\infty (\alpha+2+k) z^{k+1} \big|\partial^{k+1} f(z)\big|^2 z^\alpha \mathrm{e}^{-z}\,\mathrm{d}z \\ &\quad + \int_0^\infty z^k z^2 2\,\mathrm{Re}\,\big(\partial^{k+1} f(z)\big)\partial^{k+2} f(z) z^\alpha \mathrm{e}^{-z}\,\mathrm{d}z \\ &\le C_{\alpha,k}\Gamma(\alpha+1)\big\|\rho^{(k+1)/2}\partial^{k+1} f\big\|^2_{L^2(\mathbb{R},w)} \\ &\quad + \frac{1}{2}\int_0^\infty z^k z^2 \big|\partial^{k+1} f(z)\big|^2 z^\alpha \mathrm{e}^{-z}\,\mathrm{d}z + 2\Gamma(\alpha+1)\big\|\rho^{(k+2)/2}\partial^{k+2} f\big\|^2_{L^2(\mathbb{R},w)}.\end{aligned}$$

Second, $\rho^{k/2}\partial^{k+1} f \in L^2(\mathbb{R}, w)$ and with $M_{k,\alpha} \coloneqq k+1+\alpha$, we have

$$\big\|\rho^{k/2}\partial^{k+1} f\big\|^2_{L^2(\mathbb{R},w)} \le \frac{2}{M_{k,\alpha}}\big\|\rho^{(k+1)/2}\partial^{k+1} f\big\|^2_{L^2(\mathbb{R},w)} + \frac{4}{M_{k,\alpha}^2}\big\|\rho^{(k+2)/2}\partial^{k+2} f\big\|^2_{L^2(\mathbb{R},w)}.$$

Indeed, we integrate by parts to compute

$$\begin{aligned}\int_0^\infty z^k \big|\partial^{k+1} f(z)\big|^2 z^\alpha \mathrm{e}^{-z}\,\mathrm{d}z &= -\int_0^\infty \frac{1}{k+1+\alpha} z^{k+1} 2\,\mathrm{Re}\,\big(\partial^{k+1} f(z)\big)\partial^{k+2} f(z) z^\alpha \mathrm{e}^{-z}\,\mathrm{d}z \\ &\quad + \frac{1}{k+1+\alpha}\int_0^\infty z^{k+1}\big|\partial^{k+1} f(z)\big|^2 z^\alpha \mathrm{e}^{-z}\,\mathrm{d}z \\ &\le \frac{1}{2}\int_0^\infty z^k \big|\partial^{k+1} f(z)\big|^2 z^\alpha \mathrm{e}^{-z}\,\mathrm{d}z + \frac{2\Gamma(\alpha+1)}{(k+1+\alpha)^2}\big\|\rho^{(k+2)/2}\partial^{k+2} f\big\|^2_{L^2(\mathbb{R},w)} \\ &\quad + \frac{\Gamma(\alpha+1)}{k+1+\alpha}\big\|\rho^{(k+1)/2}\partial^{k+1} f\big\|^2_{L^2(\mathbb{R},w)}.\end{aligned}$$

Thus, for $f \in H_\rho^{\ell+2}(\mathbb{R}, w)$ we have $Qf = -\mathbf{z}f'' + (\mathbf{z}-\alpha-1)f' \in H_\rho^\ell(\mathbb{R}, w)$ and

$$\begin{aligned}\|Qf\|^2_{H_\rho^\ell(\mathbb{R},w)} &= \sum_{k=0}^{\ell} \big\|\rho^{k/2}\big(-\mathbf{z}\partial^{k+2} f - k\partial^{k+1} f + (\mathbf{z}-\alpha-1)\partial^{k+1} f + k\partial^k f\big)\big\|^2_{L^2(\mathbb{R},w)} \\ &= \sum_{k=0}^{\ell} \big\| -\rho^{(k+2)/2}\partial^{k+2} f + \mathbf{z}^{k/2}(\mathbf{z}-\alpha-1-k)\partial^{k+1} f + \rho^{k/2} k\partial^k f\big\|^2_{L^2(\mathbb{R},w)} \\ &\le \sum_{k=0}^{\ell} 3\Big(\big\|\rho^{(k+2)/2}\partial^{k+2} f\big\|^2_{L^2(\mathbb{R},w)} + 4(2\alpha+3+2k)\big\|\rho^{(k+1)/2}\partial^{k+1} f\big\|^2_{L^2(\mathbb{R},w)} \\ &\qquad + 16\big\|\rho^{(k+2)/2}\partial^{k+2} f\big\|^2_{L^2(\mathbb{R},w)} + k^2\big\|\rho^{k/2}\partial^k f\big\|^2_{L^2(\mathbb{R},w)}\Big) \\ &\le (24\alpha + 87 + 24\ell + 3\ell^2)\|f\|^2_{H_\rho^{\ell+2}(\mathbb{R},w)}.\end{aligned}$$

□

We now lift these Sobolev norm estimates to the multi-dimensional case.

Proposition 3.29. *Suppose that Z satisfies Assumption 3.20. Let $\ell \in \mathbb{N}_0$ and $f \in H_\rho^{2\ell+2}(\mathbb{R}^N, \mathbb{P}_Z)$. Then $Qf \in H_\rho^{2\ell}(\mathbb{R}^N, \mathbb{P}_Z)$ and*

$$\|Qf\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z)} \le \sqrt{N}\Big(\max_{j\in\{0,\dots,N-1\}} C_j(2\ell)\Big)\|f\|_{H_\rho^{2\ell+2}(\mathbb{R}^N,\mathbb{P}_Z)},$$

where

$$C_j(\ell) := \begin{cases} \sqrt{21+3\ell^2}, & j \in J_{\text{Normal}}, \\ \sqrt{24\alpha+87+24\ell+3\ell^2}, & j \in J_{\text{Gamma}}, \\ \sqrt{3\big(1+4(\ell+1+\max\{\alpha,\beta\})^2+\ell^2(\ell+\alpha+\beta+1)^2\big)}, & j \in J_{\text{Beta}}. \end{cases}$$

Proof. Note that $H^{2\ell+2}_\rho(\mathbb{R}^N,\mathbb{P}_Z) = \bigotimes_{j=0}^{N-1} H^{2\ell+2}_{\rho_j}(\mathbb{R},\mathbb{P}_{Z_j})$. Hence, it suffices to consider $f = \bigotimes_{j=0}^{N-1} f_j$ with $f_j \in H^{2\ell+2}_{\rho_j}(\mathbb{R},\mathbb{P}_{Z_j}) = H^{2\ell+2}_{\rho_j}(\mathbb{R},w_j)$ for $j \in \{0,\dots,N-1\}$. From the Lemmas 3.26-3.28, we deduce $\|Qf_j\|_{H^{2\ell}_{\rho_j}(\mathbb{R},w_j)} \le C_j(2\ell)\|f_j\|_{H^{2\ell+2}_{\rho_j}(\mathbb{R},w_j)}$ for $j \in \{0,\dots,N-1\}$. Then

$$\begin{aligned} \|Qf\|^2_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z)} &= \Big\| \sum_{j=0}^{N-1} f_0 \otimes \ldots \otimes f_{j-1} \otimes Q_j f_j \otimes f_{j+1} \otimes \ldots \otimes f_{N-1} \Big\|^2_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z)} \\ &\le N \sum_{j=0}^{N-1} \|Q_j f_j\|^2_{H^{2\ell}_{\rho_j}(\mathbb{R},w_j)} \prod_{k=0,k\neq j}^{N-1} \|f_k\|^2_{H^{2\ell}_{\rho_k}(\mathbb{R},w_k)} \\ &\le N \sum_{j=0}^{N-1} C_j(2\ell)^2 \|f_j\|^2_{H^{2\ell+2}_{\rho_j}(\mathbb{R},w_j)} \prod_{k=0,k\neq j}^{N-1} \|f_k\|^2_{H^{2\ell}_{\rho_k}(\mathbb{R},w_k)} \\ &\le N \Big(\max_{j\in\{0,\dots,N-1\}} C_j(2\ell) \Big)^2 \|f\|^2_{H^{2\ell+2}_\rho(\mathbb{R}^N,\mathbb{P}_Z)}. \end{aligned}$$

□

The second main ingredient for the proof of the error estimate is the relation between scalar products of f and $Q^\ell f$ with Φ_α.

Lemma 3.30. *Suppose that Z satisfies Assumption 3.20. Let $\ell \in \mathbb{N}_0$, $\alpha \in \mathcal{N}$, and $f \in H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z)$. Then*

$$(f \mid \Phi_\alpha)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)} = \Big(\sum_{j=0}^{N-1} \lambda^j_{\alpha_j} \Big)^{-\ell} \big(Q^\ell f \big| \Phi_\alpha\big)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)}.$$

Proof. We have $\Phi_\alpha = \otimes_{j=0}^{N-1} \Phi^j_{\alpha_j}$ with the orthogonal polynomial $\Phi^j_{\alpha_j}$ of degree α_j to $\mathbb{P}_{Z_j}$ for $j \in \{0,\dots,N-1\}$. Then $Q_j \Phi^j_{\alpha_j} = \lambda^j_{\alpha_j}\Phi^j_{\alpha_j}$ for $j \in \{0,\dots,N-1\}$. Thus, for $\ell \in \mathbb{N}_0$ we observe

$$Q\Phi_\alpha = \sum_{j=0}^{N-1} \widetilde{Q}_j \Phi_\alpha = \sum_{j=0}^{N-1} \lambda^j_{\alpha_j} \Phi_\alpha \quad \Rightarrow \quad Q^\ell \Phi_\alpha = \Big(\sum_{j=0}^{N-1} \lambda^j_{\alpha_j} \Big)^\ell \Phi_\alpha.$$

Since $f \in H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z)$ yields $Q^\ell f \in L^2(\mathbb{R}^N,\mathbb{P}_Z)$ (see Proposition 3.29) and Q is symmetric in $L^2(\mathbb{R}^N,\mathbb{P}_Z)$, we deduce that

$$(f \mid \Phi_\alpha)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)} = \Big(\sum_{j=0}^{N-1} \lambda^j_{\alpha_j} \Big)^{-\ell} (f | Q^\ell \Phi_\alpha)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)} = \Big(\sum_{j=0}^{N-1} \lambda^j_{\alpha_j} \Big)^{-\ell} (Q^\ell f | \Phi_\alpha)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)}.$$

□

We are now in a position to prove the multi-dimensional PC error estimate for scalar-valued functions.

Proof of Theorem 3.25. We follow the proof for the one-dimensional case in [138, Thm. 3.6]. Let $f \in H^{2\ell}_\rho(\mathbb{R}^N, \mathbb{P}_Z)$. By Parseval's identity and Lemma 3.30, the approximation error is given by

$$\begin{aligned}\|f - R_n f\|^2_{L_2(\mathbb{R}^N,\mathbb{P}_Z)} &= \sum_{|\alpha|\geq n+1} \hat{f}^2_\alpha \|\Phi_\alpha\|^2_{L_2(\mathbb{R}^N,\mathbb{P}_Z)} = \sum_{|\alpha|\geq n+1} \frac{(f \mid \Phi_\alpha)^2_{L_2(\mathbb{R}^N,\mathbb{P}_Z)}}{\|\Phi_\alpha\|^2_{L_2(\mathbb{R}^N,\mathbb{P}_Z)}} \\ &= \sum_{|\alpha|\geq n+1} \left(\sum_{j=0}^{N-1} \lambda^j_{\alpha_j}\right)^{-2\ell} \frac{1}{\|\Phi_\alpha\|^2_{L_2(\mathbb{R}^N,\mathbb{P}_Z)}} \left(Q^\ell f \,\middle|\, \Phi_\alpha\right)^2_{L_2(\mathbb{R}^N,\mathbb{P}_Z)}.\end{aligned}$$

Further define the lower eigenvalue sum bound

$$d(n) := \min_{|\alpha|\geq n} \sum_{j=0}^{N-1} \lambda^j_{\alpha_j} \quad (n \in \mathbb{N}_0).$$

Then

$$\begin{aligned}\|f - R_n f\|^2_{L_2(\mathbb{R}^N,\mathbb{P}_Z)} &\leq d(n+1)^{-2\ell} \sum_{\alpha\in\mathcal{N}} \frac{1}{\|\Phi_\alpha\|^2_{L_2(\mathbb{R}^N,\mathbb{P}_Z)}} \left(Q^\ell f \,\middle|\, \Phi_\alpha\right)^2_{L_2(\mathbb{R}^N,\mathbb{P}_Z)} \\ &= d(n+1)^{-2\ell} \sum_{\alpha\in\mathcal{N}} \left(\widehat{(Q^\ell f)}_\alpha\right)^2 \|\Phi_\alpha\|^2_{L_2(\mathbb{R}^N,\mathbb{P}_Z)} \\ &= d(n+1)^{-2\ell} \left\|Q^\ell f\right\|^2_{L_2(\mathbb{R}^N,\mathbb{P}_Z)}.\end{aligned}$$

A repeated application of Proposition 3.29 yields that

$$\begin{aligned}\|f - R_n f\|_{L_2(\mathbb{R}^N,\mathbb{P}_Z)} &\leq d(n+1)^{-\ell}\sqrt{N}\Big(\max_{0\leq j\leq N-1} C_j(0)\Big)\|Q^{\ell-1} f\|_{H^2_\rho(\mathbb{R}^N,\mathbb{P}_Z)} \\ &\leq \ldots \leq C_{\ell,N}\, d(n+1)^{-\ell}\|f\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z)}.\end{aligned}$$

Since the eigenvalues $\lambda^j_{\alpha_j}$ are known; cf. Example 3.21, we have $d(n+1) \geq n$ for all $n \in \mathbb{N}_0$. Hence, the assertion follows. □

Remark 3.31. (a) We remark that we can slightly improve the result in Theorem 3.25 by making use of the lower eigenvalue sum bound as in the proof.

(b) Note that for the Beta-distributed components of $(Z_0, \ldots, Z_{N-1})$, for the corresponding coordinates, it suffices to have weak derivatives up to order ℓ only; see also Proposition 3.23.

The error bound for the PC approximation can easily be extended from the scalar- to the vector-valued case. This extension is required for the joint convergence rate in space, time, and randomness of Theorem 3.74 in Section 3.4. Note that for a Hilbert space H, we have $L^2(\mathbb{R}^N, \mathbb{P}_Z; H) = L^2(\mathbb{R}^N, \mathbb{P}_Z) \otimes H$.

Definition 3.32. Let H be a separable Hilbert space with an orthonormal basis $\mathcal{B}$ and let $n \in \mathbb{N}_0$. Then the *(vector-valued) polynomial chaos approximation* $\mathbf{R}_n\mathbf{f}$ of $\mathbf{f} = \sum_{\varphi\in\mathcal{B}} f_\varphi \varphi \in L^2(\mathbb{R}^N, \mathbb{P}_Z; H)$ of *order* n is given by

$$\mathbf{R}_n\mathbf{f} := \sum_{\varphi\in\mathcal{B}} (R_n f_\varphi)\varphi \in \mathcal{P}^N_n \otimes H.$$

Remark 3.33. Note that $\mathbf{R}_n$ is well-defined. Indeed, let $\mathcal{B}$ and $\mathcal{C}$ be orthonormal bases of H, $n \in \mathbb{N}_0$, and $\mathbf{f} = \sum_{\varphi\in\mathcal{B}} f_\varphi \varphi = \sum_{\psi\in\mathcal{C}} f_\psi \psi \in L^2(\mathbb{R}^N, \mathbb{P}_Z; H)$. Then $f_\psi = (\mathbf{f}(\cdot) \,|\, \psi)_H$ with the overloaded notation $(\mathbf{f}(\cdot) \,|\, \psi)_H := [z \mapsto (\mathbf{f}(z) \,|\, \psi)_H] \in L_2(\mathbb{R}^N, \mathbb{P}_Z)$ and $\psi = \sum_{\varphi\in\mathcal{B}} (\psi \,|\, \varphi)_H \,\varphi$ for all $\psi \in \mathcal{C}$, and therefore

$$\begin{aligned} R_n f_\psi = R_n \left(\mathbf{f}(\cdot) \,|\, \psi\right)_H &= R_n\Big(\mathbf{f}(\cdot)\Big|\sum_{\varphi\in\mathcal{B}} (\psi \,|\, \varphi)_H \,\varphi\Big)_H \\ &= R_n \sum_{\varphi\in\mathcal{B}} (\mathbf{f}(\cdot) \,|\, \varphi)_H \,\overline{(\psi \,|\, \varphi)_H} = \sum_{\varphi\in\mathcal{B}} R_n \left(\mathbf{f}(\cdot) \,|\, \varphi\right)_H (\varphi \,|\, \psi)_H \end{aligned}$$

by continuity of R_n. Since also $f_\varphi = (\mathbf{f}(\cdot) \,|\, \varphi)_H$ and $\varphi = \sum_{\psi\in\mathcal{C}} (\varphi \,|\, \psi)_H \,\psi$ for all $\varphi \in \mathcal{B}$, we conclude well-posedness from

$$\sum_{\psi\in\mathcal{C}} (R_n f_\psi)\psi = \sum_{\psi\in\mathcal{C}} \sum_{\varphi\in\mathcal{B}} R_n \left(\mathbf{f}(\cdot) \,|\, \varphi\right)_H (\varphi \,|\, \psi)_H \,\psi = \sum_{\varphi\in\mathcal{B}} R_n \left(\mathbf{f}(\cdot) \,|\, \varphi\right)_H \varphi = \sum_{\varphi\in\mathcal{B}} (R_n f_\varphi)\varphi.$$

Remark 3.34. Let H be a separable Hilbert space with a subspace $Y \subseteq H$, and $\mathbf{f} \in L^2(\mathbb{R}^N, \mathbb{P}_Z; Y)$. Then $\mathbf{R}_n \mathbf{f} \in \mathcal{P}_n^N \otimes Y$ for all $n \in \mathbb{N}_0$.

Analogously to the scalar-valued case, we define convergence orders.

Definition 3.35. Let H be a separable Hilbert space, $Y_\mathrm{z} \hookrightarrow L^2(\mathbb{R}^N, \mathbb{P}_Z; H)$, $p_\mathrm{z} > 0$. The polynomial chaos approximation is said to be *convergent of order* p_z on Y_z if there exists a constant $C_\mathrm{z} \geq 0$ such that

$$\|\mathbf{R}_n \mathbf{f} - \mathbf{f}\|_{L^2(\mathbb{R}^N, \mathbb{P}_Z; H)} \leq C_\mathrm{z} n^{-p_\mathrm{z}} \|\mathbf{f}\|_{Y_\mathrm{z}}$$

for all $\mathbf{f} \in Y_\mathrm{z}$ and $n \in \mathbb{N}_0$.

From the scalar-valued case in Theorem 3.25, we deduce the main result of this subsection, an error estimate for the vector-valued PC approximation.

Corollary 3.36. *Suppose that Z satisfies Assumption 3.20. Let $\ell \in \mathbb{N}_0$. Then there exists $C \geq 0$ such that for all separable Hilbert spaces H, we have*

$$\|\mathbf{R}_n \mathbf{f} - \mathbf{f}\|_{L^2(\mathbb{R}^N, \mathbb{P}_Z; H)} \leq C n^{-\ell} \|\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N, \mathbb{P}_Z; H)}$$

for all $\mathbf{f} \in H_\rho^{2\ell}(\mathbb{R}^N, \mathbb{P}_Z; H)$ and $n \in \mathbb{N}_0$. In particular, the polynomial chaos approximation is convergent of order ℓ on $H_\rho^{2\ell}(\mathbb{R}^N, \mathbb{P}_Z; H)$ if $\ell > 0$.

Proof. Let H be a separable Hilbert space and $\mathcal{B}$ an orthonormal basis of H. Let $\mathbf{f} \in H_\rho^{2\ell}(\mathbb{R}^N, \mathbb{P}_Z; H)$. Then $\mathbf{f} = \sum_{\psi\in\mathcal{B}} (\mathbf{f}(\cdot) \,|\, \psi)_H \,\psi$ and thus

$$(\mathbf{R}_n \mathbf{f})(z) = \sum_{\psi\in\mathcal{B}} (R_n f_\psi)(z)\psi = \sum_{\psi\in\mathcal{B}} \big(R_n \left(\mathbf{f}(\cdot) \,|\, \psi\right)_H (z)\big)\psi = \big(R_n \left(\mathbf{f}(\cdot) \,|\, \varphi\right)_H\big)(z)$$

for $\mathbb{P}_Z$-a.e. $z \in \mathbb{R}^N$. Parseval's identity, the monotone convergence theorem, and Theorem 3.25 with constant $C \geq 0$ then imply

$$\|\mathbf{R}_n \mathbf{f} - \mathbf{f}\|^2_{L^2(\mathbb{R}^N, \mathbb{P}_Z; H)} = \sum_{\varphi\in\mathcal{B}} \int_{\mathbb{R}^N} \left| \left((\mathbf{R}_n \mathbf{f} - \mathbf{f})(z) \,|\, \varphi\right)_H \right|^2 \mathrm{d}\mathbb{P}_Z(z)$$

$$
\begin{aligned}
&= \sum_{\varphi\in\mathcal{B}} \int_{\mathbb{R}^N} \left|(R_n\,(\mathbf{f}(\cdot)\,|\,\varphi)_H)(z) - (\mathbf{f}(z)\,|\,\varphi)_H\right|^2 \,\mathrm{d}\mathbb{P}_Z(z)\\
&= \sum_{\varphi\in\mathcal{B}} \|R_n\,(\mathbf{f}(\cdot)\,|\,\varphi)_H - (\mathbf{f}(\cdot)\,|\,\varphi)_H\,\|^2_{L^2(\mathbb{R}^N,\mathbb{P}_Z)}\\
&\le C^2 n^{-2\ell} \sum_{\varphi\in\mathcal{B}} \|\,(\mathbf{f}(\cdot)\,|\,\varphi)_H\,\|^2_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z)} = C^2 n^{-2\ell}\|\mathbf{f}\|^2_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)}. \qquad \square
\end{aligned}
$$

3.3 Approximation of Deterministic Evolution Equations

As a preparation for the approximation of random evolution equations, we review results on joint convergence rates in space and time for deterministic equations. Symmetry of the underlying form allows for the treatment of less regular initial values and more general time discretisation schemes in Subsection 3.3.1 compared to the general case presented in Subsection 3.3.2. For the convenience of the reader, we sketch the proof of the main deterministic result in Theorem 3.37, even though the result is standard in the literature; see [124].

Let V, H be separable Hilbert spaces with $V \hookrightarrow H$ dense. Let $a\colon V\times V\to\mathbb{K}$ be a bounded and coervice form, i.e., there exist $M\ge 0$, $\kappa>0$ such that

$$|a(u,v)| \le M\|u\|_V\|v\|_V \quad\text{and}\quad \operatorname{Re} a(u) \ge \kappa\|u\|_V^2$$

for all $u, v\in V$.

Let A be the operator associated with a in H and $(S(t))_{t\ge0}$ the contractive C_0-semigroup generated by $-A$. Denote the graph norm in H by $\|\cdot\|_A$.

Let $u_0\in H$. Consider the abstract Cauchy problem

$$u'(t) = -Au(t) \quad (t>0), \quad u(0)=u_0.$$

Its mild solution is given by

$$u\colon [0,\infty)\to H, \quad u(t) := S(t)u_0 \quad (t\ge 0).$$

Let $(V_m)_{m\in\mathbb{N}}$ be an approximating sequence of V and denote the closure of V_m in $\|\cdot\|_H$ by H_m. The spatial approximation of a as discussed in Subsection 3.2.1 yields approximating forms $(a_m)_{m\in\mathbb{N}}$ associated with operators $(A_m)_{m\in\mathbb{N}}$. Moreover, for $m\in\mathbb{N}$ let $(S_m(t))_{t\ge0}$ be the C_0-semigroup generated by $-A_m$.

For $m\in\mathbb{N}$, we recall the H-orthogonal projection $P_m\colon H\to H_m\subseteq H$ and the corresponding embedding $J_m\colon H_m\to H$ from Subsection 3.2.1. This yields approximate solutions

$$u_m\colon [0,\infty)\to H_m, \quad u_m(t) := S_m(t)P_m u_0 \quad (t\ge0).$$

For $m\in\mathbb{N}$, let $F_m\colon[0,\infty)\to\mathcal{L}(H_m)$ be a time discretisation method for $(S_m(t))_{t\ge0}$ as in Subsection 3.2.2. Fix a final time $T\ge0$. For $k\in\mathbb{N}$ and $N_k\in\mathbb{N}$ let $\tau_k=(\tau_k^i)_{1\le i\le N_k}\in[0,T]^{N_k}$ be time steps with associated grid $\mathcal{T}_k$, $F_{m,k}(t) := \prod_{i=1}^{N_{t,k}} F_m(\tau_k^i)$, and $N_{t,k}$ as defined in Subsection 3.2.2. Temporal approximation with $F_{m,k}$ then yields the approximate solution

$$u_{m,k}\colon \mathcal{T}_k\to H_m, \quad u_{m,k}(t) := F_{m,k}(t)P_m u_0 \quad (t\in\mathcal{T}_k).$$

3.3.1 The case of symmetric forms

In this subsection, we will consider the case where a is symmetric.

Theorem 3.37 (symmetric case)**.** *Let a be symmetric, bounded, and coercive, and assume that the embedding $V \hookrightarrow H$ is compact. Let $\alpha > 0$ and assume that the space discretisation converges with order $p_{\mathrm{x}} > 0$ on $\mathrm{D}(A^{\alpha})$ for the stationary problem. Let r be an A-stable rational function as in Definition 3.14, suppose that the time discretisation methods F_m are induced by r for all $m \in \mathbb{N}$, and let $p_{\mathrm{t}} > 0$ be the order of convergence of the time discretisations.*
Then for all $T > 0$, there exist $C_T \geq 0$ and $\tau_0 = \tau_0(T) > 0$ such that for $\max_{i=1,\dots,N_k} \tau_k^i \leq \tau_0$, we have

$$\|J_m u_{m,k}(t) - u(t)\|_H \leq C_T \Big(m^{-p_{\mathrm{x}}} + \big(\max_{i=1,\dots,N_k} \tau_k^i \big)^{p_{\mathrm{t}}} \Big) \|u_0\|_{A^{\max\{\alpha+1, p_{\mathrm{t}}\}}}$$

for all $t \in \mathcal{T}_k$, $u_0 \in \mathrm{D}(A^{\max\{\alpha+1,p_{\mathrm{t}}\}})$, and $m, k \in \mathbb{N}$.

Note that the single convergence rates for the discretisation in space and time, respectively, appear exactly in the joint rate given sufficient regularity.

Remark 3.38. By the choice of admissible time discretisation methods, we only consider A-stable time discretisations, i.e., those induced by a rational function r satisfying $|r(\lambda)| < 1$ for all $\lambda < 0$. The theorem does not cover time discretisation methods for which a CFL-type condition arises.

We will make use of the following lemma (which does not require any symmetry of the form) to prove Theorem 3.37.

Lemma 3.39 (cf. [124, Lem. 7.1])**.** *For $p, m \in \mathbb{N}$, and $v \in \mathrm{D}(A^p)$, we have*

$$v = (A^{-1} - J_m A_m^{-1} P_m) A v + \sum_{j=1}^{p-1} J_m A_m^{-j} P_m (A^{-1} - J_m A_m^{-1} P_m) A^{j+1} v + J_m A_m^{-p} P_m A^p v.$$

Proof. This merely is a telescopic sum argument, noting that $P_m J_m = I_m$. □

The following lemma is a version of [124, Lem. 7.2] for not necessarily equidistant time steps, with an analogous proof.

Lemma 3.40 (cf. [124, Lem. 7.2])**.** *Let a be symmetric. Let r be an A-stable rational function as in Definition 3.14 and suppose that the time discretisation methods F_m are induced by r for all $m \in \mathbb{N}$. Then there exist $C \geq 0$ and $\tau_0 > 0$ such that*

$$\big\|(F_{m,k}(t) - S_m(t)) A_m^{-j} P_m v\big\|_{H_m} \leq C \big(\max_{i=1,\dots,N_k} \tau_k^i \big)^j \|v\|_H$$

for all $T > 0$, $m, k \in \mathbb{N}$ with $\max_{i=1,\dots,N_k} \tau_k^i \leq \tau_0$, $t \in \mathcal{T}_k$, $v \in H$, and $j \in \{0, \dots, p_{\mathrm{t}}\}$.

Proof of Theorem 3.37. We follow the proof of [124, Thm. 7.8].

By symmetry of a, A is self-adjoint, and coercivity of a implies that A is (strictly) positive definite. Compactness of the embedding $V \hookrightarrow H$ yields that A has compact resolvent.

For $m, k \in \mathbb{N}$, $t \in \mathcal{T}_k$, and $u_0 \in Y$, we estimate

$$\|J_m u_{m,k}(t) - u(t)\|_H \leq \|J_m(u_{m,k}(t) - u_m(t))\|_H + \|J_m u_m(t) - u(t)\|_H.$$

Thus, we have to estimate the single error terms corresponding to discretisation in time and space.

To estimate the first error term, we observe that $\|J_m\|_{\mathcal{L}(H_m,H)} = 1$, so

$$\|J_m(u_{m,k}(t) - u_m(t))\|_H \leq \|u_{m,k}(t) - u_m(t)\|_{H_m} = \|(F_{m,k}(t) - S_m(t))P_m u_0\|_{H_m},$$

which we have to estimate controlling the dependence on m uniformly.

Let $C \geq 0$ and $\tau_0 > 0$ as in Lemma 3.40. Let $(\varphi_\ell)_{\ell\in\mathbb{N}}$ be an orthonormal basis of eigenvectors of A with a corresponding sequence of eigenvalues $(\lambda_\ell)_{\ell\in\mathbb{N}}$. Clearly, $\varphi_\ell \in \bigcap_{n\in\mathbb{N}} \mathrm{D}(A^n)$ for all $\ell \in \mathbb{N}$.

Fix $k \in \mathbb{N}$ such that $\tau \leq \tau_0$ for $\tau := \max_{i=1,\ldots,N_k} \tau_k^i$. Define $u_\tau := u_\tau(\tau, u_0) := \sum_{\ell\in\mathbb{N}:\tau\lambda_\ell\leq 1} (u_0 \,|\, \varphi_\ell)\, \varphi_\ell$. Note that $u_\tau \in \bigcap_{n\in\mathbb{N}} \mathrm{D}(A^n)$. Then straightforward calculations using the expansions of u_0 and u_τ w.r.t. $(\varphi_\ell)_\ell$ yield

$$\|u_0 - u_\tau\|_H \leq \tau^{p_\mathrm{t}} \|A^{p_\mathrm{t}} u_0\|_H, \tag{3.3.1}$$

$$\|A^{p_\mathrm{t}} u_\tau\|_H \leq \|A^{p_\mathrm{t}} u_0\|_H, \tag{3.3.2}$$

$$\left\|A^{\alpha+1+j} u_\tau\right\|_H \leq \tau^{-j} \left\|A^{\alpha+1} u_0\right\|_H \quad (j \in \{0, \ldots, p_\mathrm{t} - 1\}). \tag{3.3.3}$$

By Lemma 3.39 applied to u_τ, we have

$$\begin{aligned}(F_{m,k}(t) - S_m(t))P_m u_\tau &= (F_{m,k}(t) - S_m(t))A_m^{-p_\mathrm{t}} P_m A^{p_\mathrm{t}} u_\tau \\ &+ \sum_{j=0}^{p_\mathrm{t}-1} (F_{m,k}(t) - S_m(t))A_m^{-j} P_m(A^{-1} - J_m A_m^{-1} P_m) A^{j+1} u_\tau\end{aligned} \tag{3.3.4}$$

for all $m \in \mathbb{N}$, since $P_m J_m = I_m$.

Lemma 3.40 implies that for all $v \in H$ and all $j \in \{0, \ldots, p_\mathrm{t}\}$,

$$\left\|(F_{m,k}(t) - S_m(t))A_m^{-j} P_m v\right\|_{H_m} \leq C\tau^j \|v\|_H. \tag{3.3.5}$$

For $j \in \{0, \ldots, p_\mathrm{t} - 1\}$ we apply this with $v := (A^{-1} - J_m A_m^{-1} P_m)A^{j+1} u_\tau$. Together with the convergence assumption for the stationary problem and (3.3.3), we conclude that for some $C_\mathrm{x} \geq 0$,

$$\begin{aligned}\left\|(F_{m,k}(t) - S_m(t))A_m^{-j} P_m v\right\|_{H_m} &\leq C\tau^j \left\|(A^{-1} - J_m A_m^{-1} P_m)A^{j+1} u_\tau\right\|_H \\ &\leq CC_\mathrm{x} \tau^j m^{-p_\mathrm{x}} \left\|A^\alpha A^{j+1} u_\tau\right\|_H \\ &\leq CC_\mathrm{x} m^{-p_\mathrm{x}} \left\|A^{\alpha+1} u_0\right\|_H.\end{aligned}$$

For $j = p_\mathrm{t}$ we apply (3.3.5) with $v := A^{p_\mathrm{t}} u_\tau$ followed by (3.3.2) to obtain

$$\|(F_{m,k}(t) - S_m(t))A_m^{-p_\mathrm{t}} P_m A^{p_\mathrm{t}} u_\tau\|_{H_m} \leq C\tau^{p_\mathrm{t}} \|A^{p_\mathrm{t}} u_\tau\|_H \leq C\tau^{p_\mathrm{t}} \|A^{p_\mathrm{t}} u_0\|_H.$$

Inserting these estimates in (3.3.4), we conclude

$$\|(F_{m,k}(t) - S_m(t))P_m u_\tau\|_{H_m} \leq C\tau^{p_\mathrm{t}} \|A^{p_\mathrm{t}} u_0\|_H + p_\mathrm{t} CC_\mathrm{x} m^{-p_\mathrm{x}} \left\|A^{\alpha+1} u_0\right\|_H.$$

In view of [21, Thm. 1], the F_m's are uniformly stable, i.e., there exist $M \geq 1$ and $\sigma \in \mathbb{R}$ such that $\|F_{m,k}(t)\| \leq M\mathrm{e}^{\sigma t}$ for all $t \in \mathcal{T}_k$ and $m, k \in \mathbb{N}$. Moreover, $\|S_m(t)\| \leq 1$ for all $t \geq 0$ and $m \in \mathbb{N}$. Invoking (3.3.1), this yields

$$\begin{aligned}\|(F_{m,k}(t) - S_m(t))P_m(u_0 - u_\tau)\|_{H_m} &\leq (M\mathrm{e}^{\sigma^+ t} + 1)\|u_0 - u_\tau\|_H \\ &\leq (M\mathrm{e}^{\sigma^+ t} + 1)\tau^{p_\mathrm{t}}\|A^{p_\mathrm{t}}u_0\|_H,\end{aligned}$$

where $\sigma^+ := \max\{\sigma, 0\}$. Since $\mathrm{e}^{\sigma^+ t} \leq \mathrm{e}^{\sigma^+ T}$, there exists $C_T \geq 0$ such that

$$\|u_{m,k}(t) - u_m(t)\|_{H_m} \leq C_T\big(\tau^{p_\mathrm{t}}\|A^{p_\mathrm{t}}u_0\|_H + m^{-p_\mathrm{x}}\big\|A^{\alpha+1}u_0\big\|_H\big)$$

for all $t \in \mathcal{T}_k$, $u_0 \in \mathrm{D}(A^{\max\{\alpha+1, p_\mathrm{t}\}})$, and $m, k \in \mathbb{N}$.

To estimate the term for the spatial discretisation of the evolution problem, we note that there exists $C_\mathrm{x} \geq 0$ such that

$$\|J_m u_m(t) - u(t)\|_H \leq C_\mathrm{x} m^{-p_\mathrm{x}}\big\|A^{\alpha+1}u_0\big\|_H$$

for all $t \in \mathcal{T}_k$, $u_0 \in \mathrm{D}(A^{\alpha+1})$ and $m \in \mathbb{N}$; cf. Corollary 3.9 in combination with Remarks 3.10 and 3.4 to pass from the graph norm $\|\cdot\|_{A^{\alpha+1}}$ to $\|A^{\alpha+1}\cdot\|$. Thus, we obtain the assertion. ☐

3.3.2 The general case

In this subsection, we do not assume that the form a is symmetric. Higher regularity of the initial values is then required to obtain the same rate of convergence compared to the symmetric case in Theorem 3.37.

Theorem 3.41. *Let a be bounded and coercive. Let $\alpha > 0$ and assume that the space discretisation converges with order $p_\mathrm{x} > 0$ on $\mathrm{D}(A^\alpha)$ for the stationary problem. Let the time discretisation methods F_m be either the implicit Euler method, where $p_\mathrm{t} := 1$, or the Crank–Nicolson method, where $p_\mathrm{t} := 2$, for all $m \in \mathbb{N}$. Then for all $T > 0$ and $\varepsilon > 0$, there exist $\tau_0 = \tau_0(T) > 0$ and $C_{T,\varepsilon} \geq 0$ such that for $\max_{i=1,\ldots,N_k} \tau_k^i \leq \tau_0$, we have*

$$\|J_m u_{m,k}(t) - u(t)\|_H \leq C_{T,\varepsilon}\left(m^{-p_\mathrm{x}} + \Big(\max_{i=1,\ldots,N_k} \tau_k^i\Big)^{p_\mathrm{t}}\right)\|u_0\|_{A^{1+\max\{\alpha+\varepsilon, p_\mathrm{t}\}}}$$

for all $t \in \mathcal{T}_k$, $u_0 \in \mathrm{D}(A^{1+\max\{\alpha+\varepsilon, p_\mathrm{t}\}})$, and $m, k \in \mathbb{N}$.

Proof. This result can be obtained by adjusting the proofs in [49, Thm. 2.7 and 2.14]. ☐

3.4 Approximation of Random Evolution Equations

In view of Section 3.1, we consider the random evolution equation

$$\mathbf{u}'(t) = -\mathbf{A}\mathbf{u}(t) \quad (t > 0), \quad \mathbf{u}(0) = \mathbf{u}_0 \in \mathbf{H}. \tag{3.4.1}$$

with multiplication operator $\mathbf{A}$ associated with a random family of operators $(A_z)_{z\in\mathbb{R}^N}$. We are interested in the approximation of (3.4.1), particularly its approximation in randomness. First, we introduce the setting and the standing assumptions.

Throughout, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $Z\colon \Omega \to \mathbb{R}^N$ a random variable with independent components, whose distributions are either standard normal, Gamma (with rate 1), or Beta. Furthermore, let V, H be separable Hilbert spaces, $V \hookrightarrow H$ densely. For $z \in \mathbb{R}^N$ let $a_z\colon V \times V \to \mathbb{K}$ be a form and assume that $\mathbb{R}^N \ni z \mapsto a_z(u)$ is measurable for all $u \in V$.

Assumption 3.42. *Assume that $(a_z)_{z\in\mathbb{R}^N}$ is $\mathbb{P}_Z$-almost surely uniformly bounded and $\mathbb{P}_Z$-almost surely uniformly coercive, i.e., there exist a $\mathbb{P}_Z$-null set $\mathcal{N}_Z \subseteq \mathbb{R}^N$ and $M \geq 0$, $\kappa > 0$ such that*

$$|a_z(u,v)| \leq M\|u\|_V\|v\|_V, \quad \operatorname{Re} a_z(u) \geq \kappa\|u\|_V^2$$

for all $u, v \in V$ and $z \in \mathbb{R}^N \setminus \mathcal{N}_Z$.

For $z \in \mathbb{R}^N \setminus \mathcal{N}_Z$ let A_z in H be the operator associated with a_z, while for $z \in \mathcal{N}_Z$ we set $A_z = 0$. For $z \in \mathbb{R}^N \setminus \mathcal{N}_Z$ let $(S_z(t))_{t\geq 0}$ be the contractive C_0-semigroup generated by $-A_z$.

As introduced in Section 3.1, we consider the multiplication operator $\mathbf{A}$ in $\mathbf{H} = L^2(\mathbb{R}^N, \mathbb{P}_Z; H)$ associated with $(A_z)_{z\in\mathbb{R}^N}$. Then Proposition 3.1 yields that $-\mathbf{A}$ generates a contractive C_0-semigroup $(\mathbf{S}(t))_{t\geq 0}$ on $\mathbf{H}$. The mild solution of (3.4.1) is given by

$$\mathbf{u}\colon [0,\infty) \to \mathbf{H}, \quad \mathbf{u}(t) \coloneqq \mathbf{S}(t)\mathbf{u}_0 \quad (t \geq 0).$$

Remark 3.43. Let $\mathbf{V} \coloneqq L^2(\mathbb{R}^N, \mathbb{P}_Z; V)$. Then it is easy to see that $\mathbf{V} \hookrightarrow \mathbf{H}$ densely. Define $\mathbf{a}\colon \mathbf{V} \times \mathbf{V} \to \mathbb{K}$ by

$$\mathbf{a}(\mathbf{u},\mathbf{v}) \coloneqq \int_{\mathbb{R}^N} a_z(\mathbf{u}(z), \mathbf{v}(z))\, \mathrm{d}\mathbb{P}_Z(z),$$

noting that $z \mapsto a_z(\mathbf{u}(z), \mathbf{v}(z))$ is measurable for all $\mathbf{u}, \mathbf{v} \in \mathbf{V}$. Then $\mathbf{a}$ is bounded and coercive, hence sectorial, and $\mathbf{A}$ is the operator associated with $\mathbf{a}$.

The form $\mathbf{a}$ gives access to the weak formulation of (3.4.1)

$$\big(\mathbf{u}'(t) \,|\, \mathbf{v}\big)_{\mathbf{H}} = -\left(\mathbf{A}\mathbf{u}(t) \,|\, \mathbf{v}\right)_{\mathbf{H}} = -\mathbf{a}(\mathbf{u}(t), \mathbf{v}) \quad (t > 0, \mathbf{v} \in \mathbf{V}), \quad \mathbf{u}(0) = \mathbf{u}_0.$$

To approximate the random evolution equation (3.4.1), an approximation in randomness is performed by means of the polynomial chaos approximation $\mathbf{R}_n$ of order $n \in \mathbb{N}_0$ (see Definition 3.32).

Remark 3.44. The straightforward approach of applying $\mathbf{R}_n$ to (3.4.1) fails, since

$$\mathbf{R}_n\mathbf{u}'(t) = (\mathbf{R}_n\mathbf{u})'(t) = -\mathbf{R}_n\mathbf{A}\mathbf{u}(t) \quad (t > 0), \quad \mathbf{R}_n\mathbf{u}(0) = (\mathbf{R}_n\mathbf{u})(0) = \mathbf{R}_n\mathbf{u}_0$$

cannot be written as an abstract Cauchy problem for the PC approximation $\tilde{\mathbf{u}}_n \coloneqq \mathbf{R}_n\mathbf{u}$ of $\mathbf{u}$. Indeed, in general, $-\mathbf{R}_n\mathbf{A}\mathbf{u}(t) \neq -\mathbf{R}_n\mathbf{A}\tilde{\mathbf{u}}_n(t) = -\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{u}(t)$ for the right-hand side. However, an abstract Cauchy problem is required in order to apply the space and time approximation results from Section 3.3, requiring a different choice of approximation $\mathbf{u}_n \neq \mathbf{R}_n\mathbf{u}$.

In Subsection 3.4.1, we show how to approximate the mild solution of (3.4.1) with PCE such that the approximate problem is an abstract Cauchy problem that can equivalently be rewritten as a coupled system of deterministic equations. Subsection 3.4.2 illustrates how a joint convergence rate for the semi-discretisation in space and time can be obtained therefrom. It remains to estimate the semi-discretisation error in randomness, which is the subject of Subsection 3.4.3, composed of auxiliary estimates in Subsections 3.4.3.1-3.4.3.3 and the main error estimate in randomness in Subsection 3.4.3.4. This allows us to state the main result on a joint convergence rate for the full discretisation in randomness, space, and time in Subsection 3.4.4. The main results are Theorems 3.74 and 3.75.

3.4.1 Derivation of the approximate problem

Our aim is to find an approximate problem to (3.4.1) of abstract Cauchy problem structure whose solution $\mathbf{u}_n$ is contained in $\mathcal{P}_n^N \otimes H \subseteq \mathbf{H}$ and can thus be obtained from a finite number of elements in H depending on the polynomial degree $n \in \mathbb{N}_0$. In other words, we want to find a (potentially coupled) system of deterministic evolution equations such that the approximate solution $\mathbf{u}_n$ can be reconstructed from the solution of the deterministic system.

To circumvent the issue of the missing abstract Cauchy problem structure, which arises from a straightforward PC approximation (see Remark 3.44), we start by restricting the form $\mathbf{a}$ to polynomial spaces w.r.t. the random parameters. Consider the restrictions $\mathbf{a}_n := \mathbf{a}|_{(\mathcal{P}_n^N \otimes V)\times(\mathcal{P}_n^N \otimes V)}$. Since $\mathbf{R}_n$ is an $\mathbf{H}$-orthogonal projection, $\mathbf{R}_n\mathbf{v}_n = \mathbf{v}_n$ for any $\mathbf{v}_n \in \mathcal{P}_n^N \otimes V \subseteq \mathcal{P}_n^N \otimes H$. It follows that for all $\mathbf{w}_n, \mathbf{v}_n \in \mathcal{P}_n^N \otimes V$, $n \in \mathbb{N}_0$,

$$\begin{aligned}\mathbf{a}_n(\mathbf{w}_n, \mathbf{v}_n) = \mathbf{a}(\mathbf{w}_n, \mathbf{v}_n) = \mathbf{a}(\mathbf{R}_n\mathbf{w}_n, \mathbf{R}_n\mathbf{v}_n) &= (\mathbf{A}\mathbf{R}_n\mathbf{w}_n \,|\, \mathbf{R}_n\mathbf{v}_n)_{\mathbf{H}} \\ &= (\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{w}_n \,|\, \mathbf{v}_n)_{\mathbf{H}} = (\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{w}_n \,|\, \mathbf{v}_n)_{\mathcal{P}_n^N \otimes H} .\end{aligned}$$

Hence, the restricted form $\mathbf{a}_n$ is associated with $\mathbf{R}_n\mathbf{A}\mathbf{R}_n$. (To be precise, $\mathbf{R}_n\mathbf{A}\mathbf{R}_n$ is the operator associated with $\mathbf{a}_n$ on the closed subspace $\mathcal{P}_n^N \otimes V \subseteq \mathbf{V}$. We can and will extend $\mathbf{R}_n\mathbf{A}\mathbf{R}_n$ by zero to $\mathbf{H}$.) Thus, we postulate that $\mathbf{u}_n$ arises as the solution of the abstract Cauchy problem associated with $\mathbf{R}_n\mathbf{A}\mathbf{R}_n$.

Definition 3.45. Let $n \in \mathbb{N}_0$. Define $\mathbf{u}_n\colon [0,\infty) \to \mathcal{L}(\mathbf{H})$ as the mild solution of the abstract Cauchy problem

$$\mathbf{u}_n'(t) = -\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{u}_n(t) \quad (t > 0), \quad \mathbf{u}_n(0) = \mathbf{u}_{0n} := \mathbf{R}_n\mathbf{u}_0. \tag{3.4.2}$$

Note that (3.4.2) is not the PC approximation of the original abstract Cauchy problem, since we additionally truncate $\mathbf{u}_n$ before applying $\mathbf{A}$ on the right-hand side. Naturally, it remains to verify the well-posedness of (3.4.2). The mild solution exists uniquely, since $-\mathbf{R}_n\mathbf{A}\mathbf{R}_n$ generates a C_0-semigroup. This is a consequence of the following resolvent bound, which will also be needed for future estimates.

Lemma 3.46. *Suppose Assumption 3.42 holds. Then for all $n \in \mathbb{N}_0$ and $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > 0$, we have*

$$\|\lambda R(\lambda, -\mathbf{R}_n\mathbf{A}\mathbf{R}_n)\|_{\mathcal{L}(\mathbf{H})} \le 1.$$

In particular, $-\mathbf{R}_n\mathbf{A}\mathbf{R}_n$ generates a contractive C_0-semigroup $(\mathbf{S}_n(t))_{t\ge 0}$ on $\mathbf{H}$ for all $n \in \mathbb{N}_0$.

Proof. Since $\mathbf{a}$ is bounded and coercive as a consequence of Assumption 3.42, also the restricted forms $\mathbf{a}_n$ are bounded and coercive with uniform parameters. Hence, $(\mathbf{a}_n)_{n\in\mathbb{N}_0}$ are uniformly sectorial of angle less than $\frac{\pi}{2}$. Consequently, the negative associated operators $-\mathbf{R}_n\mathbf{A}\mathbf{R}_n$ are (uniformly) m-sectorial, i.e., satisfy the resolvent estimate stated. In particular, they are generators of contractive C_0-semigroups. □

Abbreviate the index set by $\mathcal{N} := \mathbb{N}_0^N$ and let $(\Phi_\alpha)_{\alpha\in\mathcal{N}}$ be an orthonormal basis of $L^2(\mathbb{R}^N, \mathbb{P}_Z)$ such that $\Phi_\alpha \in \mathcal{P}_n^N$ for $|\alpha| \leq n$, and $n \in \mathbb{N}_0$. In contrast to Subsection 3.2.3, we consider Φ_α of unit norm in the following. It follows from (3.4.2) that $\mathbf{u}_n(t) \in \mathcal{P}_n^N \otimes H$, since $\mathbf{u}_n'(t) \in \mathcal{P}_n^N \otimes H$. Hence, $\mathbf{u}_n(t)$ can be completely described by finitely many H-valued coefficients. More precisely, $\mathbf{u}_n(t) = \sum_{|\beta|\leq n} \Phi_\beta \otimes u_\beta$ for $u_\beta \in H$ to be determined. Let $d_n := \sharp\{\alpha \in \mathcal{N} : \ |\alpha| \leq n\}$. Note that $d_n = \binom{n+N}{n}$. Hence, we need d_n H-valued coefficients, and consequently, we can rewrite the random abstract Cauchy problem (3.4.2) for $\mathbf{u}_n(t)$ as a system of d_n deterministic equations, from whose solution $\mathbf{u}_n(t)$ can be obtained. To derive this deterministic system, we first observe that the restricted forms can be represented by a finite linear combination using the following notation and a simple but useful observation in Remark 3.48.

Notation 3.47. *To abbreviate, for* $\mathbf{w} \in \mathbf{H}$ *and* $\alpha \in \mathcal{N}$ *let us write*

$$\widehat{\mathbf{w}}_\alpha := (\mathbf{w} \,|\, \Phi_\alpha)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)} := \int_{\mathbb{R}^N} \mathbf{w}(z)\Phi_\alpha(z)\,\mathrm{d}\mathbb{P}_Z(z) \in H$$

with a slight abuse of notation. For $u, v \in V$*, denote the* α*-th generalised Fourier coefficient of* $a.(u, v)$ *by* $\hat{a}_\alpha(u, v)$*, i.e.,* $\hat{a}_\alpha(u, v) := (a.(u, v) \,|\, \Phi_\alpha)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)}$ *for* $\alpha \in \mathcal{N}$*. Further, for* $\alpha, \beta, \gamma \in \mathcal{N}$*, denote*

$$\varepsilon_{\alpha,\beta,\gamma} := (\Phi_\alpha \,|\, \Phi_\beta\Phi_\gamma)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)} = \int_{\mathbb{R}^N} \Phi_\alpha\Phi_\beta\Phi_\gamma \,\mathrm{d}\mathbb{P}_Z.$$

Remark 3.48. Note that $\varepsilon_{\alpha,\beta,\gamma} = 0$ for all $\alpha \in \mathcal{N}$ with $|\alpha| > |\beta| + |\gamma|$. Indeed, since $(\Phi_\alpha)_{\alpha\in\mathcal{N}}$ is an orthonormal basis, $\Phi_\alpha \perp \mathcal{P}_\ell^N$ for all $\ell < |\alpha|$. The statement then follows from $\Phi_\beta\Phi_\gamma \in \mathcal{P}_{|\beta|+|\gamma|}^N$. Moreover, Φ_α is $\mathbb{R}$-valued (see Subsection 3.2.3, which is why the complex conjugates can be omitted in the definition of $\varepsilon_{\alpha,\beta,\gamma}$ and the indices can be interchanged.

Let $\mathbf{w}_n = \sum_{|\beta|\leq n} \Phi_\beta \otimes u_\beta, \mathbf{v}_n = \sum_{|\gamma|\leq n} \Phi_\gamma \otimes v_\gamma \in \mathcal{P}_n^N \otimes V$ for $n \in \mathbb{N}_0$. Then a PC approximation of $a.(u_\beta, v_\gamma)$ combined with Remark 3.48 results in

$$\begin{aligned}
\mathbf{a}_n(\mathbf{w}_n, \mathbf{v}_n) = \mathbf{a}(\mathbf{w}_n, \mathbf{v}_n) &= \sum_{|\beta|\leq n}\sum_{|\gamma|\leq n} \int_{\mathbb{R}^N} a_z(u_\beta, v_\gamma)(\Phi_\beta\Phi_\gamma)(z)\,\mathrm{d}\mathbb{P}_Z(z) \\
&= \sum_{|\beta|\leq n}\sum_{|\gamma|\leq n} \int_{\mathbb{R}^N} \left(\sum_{\alpha\in\mathcal{N}} \hat{a}_\alpha(u_\beta, v_\gamma)\Phi_\alpha(z)\right)(\Phi_\beta\Phi_\gamma)(z)\,\mathrm{d}\mathbb{P}_Z(z) \\
&\overset{(*)}{=} \sum_{|\beta|\leq n}\sum_{|\gamma|\leq n}\sum_{\alpha\in\mathcal{N}} \hat{a}_\alpha(u_\beta, v_\gamma)\varepsilon_{\alpha,\beta,\gamma} \\
&= \sum_{|\beta|\leq n}\sum_{|\gamma|\leq n}\sum_{|\alpha|\leq 2n} \hat{a}_\alpha(u_\beta, v_\gamma)\varepsilon_{\alpha,\beta,\gamma},
\end{aligned} \tag{3.4.3}$$

where in $(*)$ we were allowed to interchange integration and summation since the polynomial chaos expansion converges in $L^2(\mathbb{R}^N, \mathbb{P}_Z)$ and thus the integrand also converges in $L^1(\mathbb{R}^N, \mathbb{P}_Z)$. Motivated by the above observation that $\mathbf{a}_n$ can be evaluated by a finite linear combination, we define the following deterministic forms.

Definition 3.49. For $n \in \mathbb{N}_0$ and $\beta, \gamma \in \mathcal{N}$ we define $a_{n,\beta,\gamma} \colon V \times V \to \mathbb{K}$ by

$$a_{n,\beta,\gamma} \coloneqq \sum_{|\alpha| \leq n} \varepsilon_{\alpha,\beta,\gamma} \widehat{a}_\alpha.$$

Let $\mathfrak{V}_n \coloneqq V^{d_n}$ be equipped with the norm $\|(u_\beta)_{|\beta \leq n}\|^2_{\mathfrak{V}_n} \coloneqq \sum_{|\beta| \leq n} \|u_\beta\|^2_V$ and define $\mathfrak{a}_n \colon \mathfrak{V}_n \times \mathfrak{V}_n \to \mathbb{K}$ by

$$\mathfrak{a}_n((u_\beta)_{|\beta| \leq n}, (v_\gamma)_{|\gamma| \leq n}) \coloneqq \sum_{|\gamma| \leq n} \sum_{|\beta| \leq n} a_{2n,\beta,\gamma}(u_\beta, v_\gamma).$$

The corresponding quadratic form $\mathfrak{a}_n \colon \mathfrak{V}_n \to \mathbb{K}$ is denoted by $\mathfrak{a}_n$ as well.

Proposition 3.50. *Suppose that Assumption 3.42 holds. Let $n \in \mathbb{N}_0$. Then*

$$\begin{aligned} \left|\mathfrak{a}_n((u_\beta)_{|\beta| \leq n}, (v_\gamma)_{|\gamma| \leq n})\right| &\leq M \left\|(u_\beta)_{|\beta| \leq n}\right\|_{\mathfrak{V}_n} \left\|(v_\gamma)_{|\gamma| \leq n}\right\|_{\mathfrak{V}_n}, \\ \operatorname{Re} \mathfrak{a}_n((u_\beta)_{|\beta| \leq n}) &\geq \kappa \left\|(u_\beta)_{|\beta| \leq n}\right\|^2_{\mathfrak{V}_n} \end{aligned}$$

for all $(u_\beta)_{|\beta| \leq n}, (v_\gamma)_{|\gamma| \leq n} \in \mathfrak{V}_n$.

Proof. Let $\mathbf{w}_n \coloneqq \sum_{|\beta| \leq n} \Phi_\beta \otimes u_\beta$, $\mathbf{v}_n \coloneqq \sum_{|\gamma| \leq n} \Phi_\gamma \otimes v_\gamma$. From (3.4.3), we deduce

$$\mathfrak{a}_n\big((u_\beta)_{|\beta| \leq n}, (v_\gamma)_{|\gamma| \leq n}\big) = \sum_{|\beta| \leq n} \sum_{|\gamma| \leq n} \sum_{|\alpha| \leq 2n} \hat{a}_\alpha(u_\beta, v_\gamma) \varepsilon_{\alpha,\beta,\gamma} = \mathbf{a}(\mathbf{w}_n, \mathbf{v}_n). \tag{3.4.4}$$

Boundedness and coercivity of $\mathbf{a}$ allow us to estimate $\mathfrak{a}_n$ in terms of $\|\mathbf{w}_n\|_{\mathbf{V}}$ and $\|\mathbf{v}_n\|_{\mathbf{V}}$. Parseval's identity yields the desired norm identity

$$\|\mathbf{w}_n\|^2_{\mathbf{V}} = \sum_{\beta \in \mathcal{N}} \|(\widehat{\mathbf{w}_n})_\beta\|^2_V = \sum_{\beta \in \mathcal{N}} \left\| (\mathbf{w}_n \mid \Phi_\beta)_{L_2(\mathbb{R}^N, \mathbb{P}_Z)} \right\|^2_V = \sum_{|\beta| \leq n} \|u_\beta\|^2_V = \left\|(u_\beta)_{|\beta| \leq n}\right\|^2_{\mathfrak{V}_n}. \square$$

Note that the above proposition implies sectoriality of $\mathfrak{a}_n$. Having established the connection between the random form $\mathbf{a}$ and the deterministic forms $\mathfrak{a}_n$ acting on the finite-dimensional Cartesian product $\mathfrak{V}_n$ of the (typically infinite-dimensional) space V, the question of associated operators arises. Being a bounded and coercive form, we can associate an operator $\mathfrak{A}_n$ with $\mathfrak{a}_n$. To get a representation for this operator, we start by considering the operators associated with the building blocks $a_{n,\beta,\gamma}$ of $\mathfrak{a}_n$.

Lemma 3.51. *Let $n \in \mathbb{N}_0$ and $\beta, \gamma \in \mathcal{N}$. Suppose that Assumption 3.42 holds. Then $a_{n,\beta,\gamma}$ is bounded.*

Proof. By Bessel's inequality, $\widehat{a}_\alpha$ is bounded for $\alpha \in \mathcal{N}$. As the sum defining $a_{n,\beta,\gamma}$ is finite, we obtain the assertion. □

By Lemma 3.51, we can associate an operator in H to $a_{n,\beta,\gamma}$ for $n \in \mathbb{N}_0$, $\beta, \gamma \in \mathcal{N}$. Denote this operator by $A_{n,\beta,\gamma}$. An assumption is necessary in order to ensure these operators are densely defined.

Assumption 3.52. *Suppose that Assumption 3.42 holds. Further, assume that there exists $\overline{\alpha} > 0$ and a subspace $D^{\overline{\alpha}} \subseteq H$ such that $\mathrm{D}(A_z^{\overline{\alpha}}) = D^{\overline{\alpha}}$ for $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$ and the associated graph norms are $\mathbb{P}_Z$-almost surely equivalent.*

Remark 3.53. Suppose that Assumption 3.52 holds for some $\overline{\alpha} > 0$, and let $0 < \alpha < \overline{\alpha}$. Then, by interpolation theory, we have that $D^\alpha \coloneqq [H, D^{\overline{\alpha}}]_{\alpha/\overline{\alpha}} = \mathrm{D}(A_z^\alpha)$ for $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$ and the associated graph norms are $\mathbb{P}_Z$-almost surely equivalent. If $\overline{\alpha} \geq 1$ we abbreviate $D \coloneqq D^1$.

Lemma 3.54. *Suppose that Assumption 3.52 holds for $\overline{\alpha} = 1$. Then $\mathrm{D}(A_{n,\beta,\gamma}) = D$ for all $n \in \mathbb{N}_0$ and $\beta, \gamma \in \mathcal{N}$.*

Proof. Let $\alpha \in \mathcal{N}$ and let $\widehat{A}_\alpha$ be the operator in H associated with $\widehat{a}_\alpha$. We show that $\mathrm{D}(\widehat{A}_\alpha) = D$. This suffices to yield the assertion by definition of $A_{n,\beta,\gamma}$ as a linear combination.

Let $u \in D$ and $v \in V$. Then $1 \otimes u \in \mathrm{D}(\mathbf{A})$ and

$$\begin{aligned}\widehat{a}_\alpha(u,v) &= (a.(u,v) \,|\, \Phi_\alpha)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)} = \int_{\mathbb{R}^N} a_z(1 \cdot u, \Phi_\alpha(z)v)\,\mathrm{d}\mathbb{P}_Z(z)\\ &= \mathbf{a}(1 \otimes u, \Phi_\alpha \otimes v) = (\mathbf{A}(1 \otimes u) \,|\, \Phi_\alpha \otimes v)_{\mathbf{H}}\\ &= \Big((\mathbf{A}(1 \otimes u) \,|\, \Phi_\alpha)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)} \,\Big|\, v \Big)_H .\end{aligned}$$

Thus, $u \in \mathrm{D}(\widehat{A}_\alpha)$, so $D \subseteq \mathrm{D}(\widehat{A}_\alpha)$.

Now, let $u \in \mathrm{D}(\widehat{A}_\alpha)$. Then $u \in V$ and therefore $1 \otimes u \in \mathbf{V}$. Thus, for $\mathbf{v} \in \mathbf{V}$

$$\begin{aligned}\mathbf{a}(1 \otimes u, \mathbf{v}) &= \int_{\mathbb{R}^N} a_z(u, \mathbf{v}(z))\,\mathrm{d}\mathbb{P}_Z(z) = \int_{\mathbb{R}^N} \sum_{\alpha \in \mathcal{N}} \widehat{a}_\alpha(u, \mathbf{v}(z))\Phi_\alpha(z)\,\mathrm{d}\mathbb{P}_Z(z)\\ &= \int_{\mathbb{R}^N} \sum_{\alpha \in \mathcal{N}} \Big(\widehat{A}_\alpha u \,\Big|\, \mathbf{v}(z) \Big)_H \Phi_\alpha(z)\,\mathrm{d}\mathbb{P}_Z(z)\\ &\overset{(*)}{=} \int_{\mathbb{R}^N} \left(\sum_{\alpha \in \mathcal{N}} \Phi_\alpha(z)\widehat{A}_\alpha u \,\middle|\, \mathbf{v}(z) \right)_H \mathrm{d}\mathbb{P}_Z(z) = \left(\sum_{\alpha \in \mathcal{N}} \Phi_\alpha \otimes \widehat{A}_\alpha u \,\middle|\, \mathbf{v} \right)_{\mathbf{H}} .\end{aligned}$$

Since the polynomial chaos expansion converges in $L^2(\mathbb{R}^N, \mathbb{P}_Z)$ and thus the integrand converges also in $L^1(\mathbb{R}^N, \mathbb{P}_Z)$, and since taking scalar products is a linear functional, we were allowed to interchange series and scalar product in $(*)$. Thus, $1 \otimes u \in \mathrm{D}(\mathbf{A})$, i.e., $u \in D$. □

This ensures that $\mathfrak{a}_n$ is associated with a densely defined operator.

Definition 3.55. Let $\mathfrak{H}_n \coloneqq \oplus_{|\beta| \leq n} H = H^{d_n}$ be equipped with the scalar product

$$\big((u_\beta)_{|\beta| \leq n} \,\big|\, (v_\gamma)_{|\gamma| \leq n} \big)_{\mathfrak{H}_n} \coloneqq \sum_{|\beta| \leq n} (u_\beta \,|\, v_\gamma)_H$$

and the corresponding induced norm for $n \in \mathbb{N}_0$. Define $\mathfrak{A}_n$ in $\mathfrak{H}_n$ via

$$\mathrm{D}(\mathfrak{A}_n) \coloneqq \bigoplus_{|\beta| \leq n} D = D^{d_n}, \quad \mathfrak{A}_n((u_\beta)_{|\beta| \leq n}) \coloneqq \left(\sum_{|\beta| \leq n} A_{2n,\beta,\gamma} u_\beta \right)_{|\gamma| \leq n} \in \mathfrak{H}_n.$$

Propositions 3.50 implies sectoriality of $\mathfrak{A}_n$. Under the additional assumption of symmetric forms a_z, self-adjointness of $\mathfrak{A}_n$ is obtained.

Proposition 3.56. *Suppose that Assumption 3.52 holds for* $\overline{\alpha} = 1$*, and let* a_z *be symmetric for* $\mathbb{P}_Z$*-almost every* $z \in \mathbb{R}^N$*,* $n \in \mathbb{N}_0$*. Then* $\mathfrak{a}_n$ *is symmetric, and* $\mathfrak{A}_n$ *is self-adjoint.*

Proof. Let $\alpha \in \mathcal{N}$. It is easy to see that $\widehat{a}_\alpha$ is symmetric. Thus, $a_{n,\beta,\gamma}$ is also symmetric. Moreover, $\varepsilon_{\alpha,\beta,\gamma} = \varepsilon_{\alpha,\gamma,\beta}$ for all $\beta, \gamma \in \mathcal{N}$; cf. Remark 3.48. Hence, $a_{n,\beta,\gamma} = a_{n,\gamma,\beta}$. Thus, we easily observe the symmetry of $\mathfrak{a}_n$, which directly implies the self-adjointness of $\mathfrak{A}_n$. □

Finally, we can now relate the solution $\mathbf{u}_n$ of the random abstract Cauchy problem to the solution of the coupled deterministic system given by $\mathfrak{A}_n$. We find that the solution vector of the deterministic system contains exactly the PCE coefficients of order up to n of $\mathbf{u}_n$. We recall Notation 3.47.

Proposition 3.57. *Suppose that Assumption 3.52 holds for* $\overline{\alpha} = 1$*. For* $n \in \mathbb{N}_0$*, let* $\mathbf{u}_n$ *be the mild solution of* (3.4.2)*. Define the approximate initial condition*

$$\mathfrak{u}_{0n} := \big(\widehat{(\mathbf{u}_0)}_\alpha\big)_{|\alpha|\le n} = \big((\mathbf{u}_0 \,|\, \Phi_\alpha)_{L_2(\mathbb{R}^N,\mathbb{P}_Z)} \big)_{|\alpha|\le n}. \tag{3.4.5}$$

Then the coupled deterministic system

$$\mathfrak{u}_n'(t) = -\mathfrak{A}_n \mathfrak{u}_n(t) \quad (t > 0), \quad \mathfrak{u}_n(0) = \mathfrak{u}_{0n} \tag{3.4.6}$$

is solved by $\mathfrak{u}_n \colon [0,\infty) \to \mathfrak{H}_n$*,*

$$\mathfrak{u}_n(t) := \big((\mathbf{u}_n(t) \,|\, \Phi_\beta)_{L^2(\mathbb{R}^N,\mathbb{P}_Z)} \big)_{|\beta|\le n} \in \mathfrak{H}_n.$$

Proof. We show the initial condition and $\mathfrak{u}_n'(t) = -\mathfrak{A}_n\mathfrak{u}_n(t)$ componentwise. Let $\gamma \in \mathcal{N}$ with $|\gamma| \le n$. Since the γ-th Fourier coefficients of $\mathbf{R}_n\mathbf{u}_0$ and $\mathbf{u}_0$ agree for $|\gamma| \le n$,

$$(\mathfrak{u}_n(0))_\gamma = (\mathbf{u}_n(0) \,|\, \Phi_\gamma)_{L_2(\mathbb{R}^N,\mathbb{P}_Z)} = \widehat{(\mathbf{R}_n\mathbf{u}_0)}_\gamma = \widehat{(\mathbf{u}_0)}_\gamma = (\mathfrak{u}_{0n})_\gamma.$$

Abbreviate the PCE coefficients of $\mathbf{u}_n(t)$ by $u_\beta := \widehat{(\mathbf{u}_n(t))}_\beta$. Testing the negative right-hand side of the system with $v \in V$ yields

$$((\mathfrak{A}_n\mathfrak{u}_n(t))_\gamma \,|\, v)_H = \sum_{|\beta|\le n} a_{2n,\beta,\gamma}(u_\beta, v) = \sum_{|\beta|\le n}\sum_{|\alpha|\le 2n} \varepsilon_{\alpha\beta\gamma}\hat{a}_\alpha(u_\beta, v) \tag{3.4.7}$$

by definition of $\mathfrak{u}_n(t), \mathfrak{A}_n, a_{n,\beta,\gamma}$, and using that $a_{2n,\beta,\gamma}$ is the form corresponding to $A_{2n,\beta,\gamma}$. For the negative left-hand side, using the definition of $\mathbf{u}_n$, we rewrite

$$\begin{aligned}(-\mathfrak{u}_n'(t))_\gamma &= -\frac{\mathrm{d}}{\mathrm{d}t}\,(\mathbf{u}_n(t) \,|\, \Phi_\gamma)_{L_2(\mathbb{R}^N,\mathbb{P}_Z)} = \big(-\mathbf{u}_n'(t) \,\big|\, \Phi_\gamma\big)_{L_2(\mathbb{R}^N,\mathbb{P}_Z)} \\ &= (\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{u}_n(t) \,|\, \Phi_\gamma)_{L_2(\mathbb{R}^N,\mathbb{P}_Z)}.\end{aligned}$$

Testing with $v \in V$, we can use that $\mathbf{a}_n$ corresponds to $\mathbf{R}_n\mathbf{A}\mathbf{R}_n$ and (3.4.3) holds to conclude

$$\begin{aligned}\big((-\mathfrak{u}_n'(t))_\gamma \,\big|\, v\big)_H &= \Big((\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{u}_n(t) \,|\, \Phi_\gamma)_{L_2(\mathbb{R}^N,\mathbb{P}_Z)} \,\Big|\, v\Big)_H \\ &= (\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{u}_n(t) \,|\, \Phi_\gamma \otimes v)_{\mathbf{H}} = \mathbf{a}_n(\mathbf{u}_n(t), \Phi_\gamma \otimes v) \\ &= \sum_{|\beta|\le n}\sum_{|\alpha|\le 2n} \hat{a}_\alpha(u_\beta, v)\varepsilon_{\alpha,\beta,\gamma},\end{aligned}$$

which agrees with (3.4.7) and thus finishes the proof. □

3.4.2 Semi-discretisation in space and time

In Proposition 3.57, we have derived the deterministic abstract Cauchy problem (3.4.6). Due to its structure, we can now discretise in space and time in the same manner as in Section 3.3.

Let $n \in \mathbb{N}_0$ and $(V_m)_{m\in\mathbb{N}}$ be an approximating sequence of V. Then, $(\mathfrak{V}_{n,m})_{m\in\mathbb{N}}$ with $\mathfrak{V}_{n,m} := V_m^{d_n}$ is an approximating sequence of $\mathfrak{V}_n$. The spatial approximation of $\mathfrak{a}_n$ as discussed in Subsection 3.2.1 yields approximating forms $(\mathfrak{a}_{n,m})_{m\in\mathbb{N}}$ on $\mathfrak{V}_{n,m}$ and associated operators $(\mathfrak{A}_{n,m})_{m\in\mathbb{N}}$ on $\mathfrak{H}_{n,m} := H_m^{d_n}$, where $H_m := \overline{V_m}^{\|\cdot\|_H}$. Moreover, for $m \in \mathbb{N}$, let $(\mathfrak{S}_{n,m}(t))_{t\geq 0}$ be the C_0-semigroup generated by $-\mathfrak{A}_{n,m}$.

For $m \in \mathbb{N}$ we recall the H-orthogonal projection $P_m\colon H \to H_m \subseteq H$ and the corresponding embedding $J_m\colon H_m \to H$ from Subsection 3.2.1, which can be naturally lifted to $\mathfrak{P}_{n,m}\colon \mathfrak{H}_n \to \mathfrak{H}_{n,m} \subseteq \mathfrak{H}_n$ and $\mathfrak{J}_{n,m}\colon \mathfrak{H}_{n,m} \to \mathfrak{H}_n$, as well as

$$\mathbf{P}_{n,m}\colon \mathcal{P}_n^N \otimes H \to \mathcal{P}_n^N \otimes H_m, \quad \mathbf{J}_{n,m}\colon \mathcal{P}_n^N \otimes H_m \to \mathcal{P}_n^N \otimes H \tag{3.4.8}$$

for $n \in \mathbb{N}_0$. Then $\mathbf{P}_{n,m}(\sum_{|\beta|\leq n} \Phi_\beta \otimes u_\beta) = \sum_{|\beta|\leq n} \Phi_\beta \otimes (P_m u_\beta)$. This yields semi-discretisations

$$\mathfrak{u}_{n,m}\colon [0,\infty) \to \mathfrak{H}_{n,m}, \quad \mathfrak{u}_{n,m}(t) := \mathfrak{S}_{n,m}(t)\mathfrak{P}_{n,m}\mathfrak{u}_{0n} \quad (t \geq 0).$$

For $n \in \mathbb{N}_0$ and $m \in \mathbb{N}$, let $\mathfrak{F}_{n,m}\colon [0,\infty) \to \mathcal{L}(\mathfrak{H}_{n,m})$ be a time discretisation method for $(\mathfrak{S}_{n,m}(t))_{t\geq 0}$ as in Subsection 3.2.2. Fix a final time $T \geq 0$. For $k \in \mathbb{N}$ and $N_k \in \mathbb{N}$ let $\tau_k = (\tau_k^i)_{1\leq i\leq N_k} \in [0,T]^{N_k}$ be a vector of N_k time steps such that $\tau_k \to 0$ and $N_k \to \infty$ as $k \to \infty$ as well as $\sum_{i=1}^{N_k} \tau_k^i = T$ for all $k \in \mathbb{N}$. For $k \in \mathbb{N}$ let $\mathcal{T}_k$ be the grid associated with τ_k. For $t \in \mathcal{T}_k$ let $\mathfrak{F}_{n,m,k}(t) := \prod_{i=1}^{N_{t,k}} \mathfrak{F}_{n,m}(\tau_k^i)$, where $N_{t,k} := \max\{j \in \{0,\dots,N_k\} : \sum_{i=1}^{j} \tau_k^i = t\}$. Temporal approximation with $\mathfrak{F}_{n,m,k}$ then yields the discretisations

$$\mathfrak{u}_{n,m,k}\colon \mathcal{T}_k \to \mathfrak{H}_{n,m}, \quad \mathfrak{u}_{n,m,k}(t) := \mathfrak{F}_{n,m,k}(t)\mathfrak{P}_{n,m}\mathfrak{u}_{0n} \quad (t \in \mathcal{T}_k).$$

Following the idea of Proposition 3.57, we consider the corresponding element in $\mathbf{H}_m := L^2(\mathbb{R}^N, \mathbb{P}_Z; H_m)$ with PCE coefficients $\mathfrak{u}_{n,m,k}(t)$. That is, we let

$$\mathbf{u}_{n,m,k}\colon \mathcal{T}_k \to \mathbf{H}_m, \quad \mathbf{u}_{n,m,k}(t) := \sum_{|\beta|\leq n} \Phi_\beta \otimes (\mathfrak{u}_{n,m,k}(t))_\beta \quad (t \in \mathcal{T}_k). \tag{3.4.9}$$

Combining the results from Section 3.3 and Subsection 3.4.1, we are in a position to estimate the semi-discretisation error in space and time. We focus on the symmetric case corresponding to Subsection 3.3.1 until we can estimate the full discretisation error in Subsection 3.4.4. We note that also the general case will be treated there (taking into account Subsection 3.3.2).

Proposition 3.58. *Suppose that Assumption 3.52 holds for some $\overline{\alpha} \geq 1$. Further, assume that a_z is symmetric for $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$ and that $V \hookrightarrow H$ is compact. Suppose that the space discretisation converges with order $p_\mathrm{x} > 0$ on D^α for some $0 < \alpha < \overline{\alpha}$ for the stationary problem. Let r be an A-stable rational function as in Definition 3.14, suppose that the time discretisation methods $\mathfrak{F}_{n,m}$ are induced by r for all $n, m \in \mathbb{N}$, and let $p_\mathrm{t} > 0$ be the order of convergence of the time discretisations. Let $\mathbf{J}_{n,m}, \mathbf{u}_{n,m,k}, \mathbf{u}_n$, and $\mathfrak{u}_{0n}$ as in* (3.4.8), (3.4.9), (3.4.2), *and* (3.4.5), *respectively.*

Then for all $T > 0$ *there exist* $C_T \geq 0$ *and* $\tau_0 = \tau_0(T) > 0$ *such that for* $\max_{i=1,\ldots,N_k} \tau_k^i \leq \tau_0$*, we have*

$$\|\mathbf{J}_{n,m}\mathbf{u}_{n,m,k}(t) - \mathbf{u}_n(t)\|_{\mathbf{H}} \leq C_T\Big(m^{-p_{\mathrm{x}}} + \Big(\max_{i=1,\ldots,N_k} \tau_k^i\Big)^{p_{\mathrm{t}}}\Big)\|\mathfrak{u}_{0n}\|_{\mathfrak{A}_n^{\max\{\alpha+1,p_{\mathrm{t}}\}}}$$

for all $t \in \mathcal{T}_k$*,* $\mathbf{u}_0 \in \mathbf{H}$ *such that* $\mathfrak{u}_{0n} \in \mathrm{D}(\mathfrak{A}_n^{\max\{\alpha+1,p_{\mathrm{t}}\}})$*,* $n \in \mathbb{N}_0$*, and* $m, k \in \mathbb{N}$*.*

Proof. Note that by Parseval's identity, we have

$$\|\mathbf{J}_{n,m}\mathbf{u}_{n,m,k}(t) - \mathbf{u}_n(t)\|_{\mathbf{H}} = \|\mathfrak{J}_{n,m}\mathfrak{u}_{n,m,k}(t) - \mathfrak{u}_n(t)\|_{\mathfrak{H}_n},$$

as $(\Phi_\alpha)_\alpha$ is an orthonormal basis. By Theorem 3.37, we conclude

$$\|\mathbf{J}_{n,m}\mathbf{u}_{n,m,k}(t) - \mathbf{u}_n(t)\|_{\mathbf{H}} \leq C\left(m^{-p_{\mathrm{x}}} + \Big(\max_{i=1,\ldots,N_k} \tau_k^i\Big)^{p_{\mathrm{t}}}\right)\|\mathfrak{u}_{0n}\|_{\mathfrak{A}_n^{\max\{\alpha+1,p_{\mathrm{t}}\}}},$$

since $\mathfrak{a}_n$ is symmetric by Proposition 3.56. □

Rather than estimating the error in $\mathbf{H}$ in terms of $\|\mathfrak{u}_{0n}\|_{\mathfrak{A}_n^{\max\{\alpha+1,p_{\mathrm{t}}\}}}$, we would prefer to estimate it in terms of a suitable norm of $\mathbf{u}_{0n}$ or $\mathbf{u}_0$ directly. However, a further condition is required to link the graph norm of $\mathfrak{u}_{0n}$ to a Sobolev norm of $\mathbf{u}_0$.

Moreover, we are only considering the semi-discretisation error in space and time in Proposition 3.58. For an estimate of the full discretisation error, it remains to estimate the randomness semi-discretisation error $\|\mathbf{u}_n(t) - \mathbf{u}(t)\|_{\mathbf{H}}$ in terms of a suitable norm of $\mathbf{u}_0$ with decay at some rate in n.

3.4.3 Semi-discretisation in randomness

Since our approximation $\mathbf{u}_n(t) \neq \mathbf{R}_n\mathbf{u}(t)$ (see Remark 3.44), estimating the semidiscretisation error in randomness is more intricate than a simple application of the polynomial chaos error estimate from Corollary 3.36. This subsection is devoted to establishing an estimate for $\|\mathbf{u}_n(t) - \mathbf{u}(t)\|_{\mathbf{H}}$, allowing us to estimate the full error in Subsection 3.4.4.

The estimate is established in several steps. In Subsection 3.4.3.1, we present an assumption, which implies a bound for $\mathbf{A}$ and its resolvent in suitable norms, as we will illustrate in Subsection 3.4.3.2. Obtaining these estimates on (subspaces of) $\mathbf{H}$ requires pointwise in z estimates presented in this subsection, which are then lifted to Sobolev subspaces of $\mathbf{H}$ via a composition estimate. From the estimates thus obtained, we deduce a resolvent difference estimate in Subsection 3.4.3.3. Via a modified Trotter–Kato argument in Subsection 3.4.3.4, this yields a convergence rate for the corresponding semigroups, i.e., a convergence rate for the semi-discretisation in randomness.

In the non-symmetric case, the Trotter–Kato argument requires working with fractional powers of the generator. Within each subsection, we first present the estimates required for the symmetric case, followed by those for the non-symmetric case.

Henceforth, C, C_ℓ, $C_{\ell,r}$, etc. denote generic constants, whose values can vary from line to line and depend on the quantities indexed.

3.4.3.1 Pointwise (in z) estimates

We recall Assumption 3.20 on the distribution of the components of Z and the definition of ρ from (3.2.4). In the case where Z has only normally or Beta-distributed components, ρ can be omitted in the following due to $\rho \equiv 1$. It is, however, required in the presence of Gamma-distributed components of Z.

Assumption 3.59. *Suppose that Assumption 3.52 holds for some $\overline{\alpha} \geq 1$, and let Assumption 3.20 hold for Z. Assume that for some $\ell \in \mathbb{N}_0$, $[z \mapsto a_z(u,v)] \in C^\ell(\mathbb{R}^N)$ for all $u, v \in V$ and there exists a constant $C_\ell \geq 0$ such that*

$$|\rho(z)^{\alpha/2} \partial_z^\alpha a_z(u,v)| \leq C_\ell \|u\|_D \|v\|_H$$

for $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$ and for all $u \in D$, $v \in V$, and $\alpha \in \mathcal{N}$ with $|\alpha| \leq \ell$.

Recall that then $(A_z)_{z\in\mathbb{R}^N}$ are $\mathbb{P}_Z$-almost surely uniformly sectorial as a consequence of $\mathbb{P}_Z$-almost sure boundedness and coercivity of $(a_z)_{z\in\mathbb{R}^N}$.

Definition 3.60. Let $\theta_0 \in (0, \frac{\pi}{2})$ be the $\mathbb{P}_Z$-almost surely uniform angle of sectoriality of $(A_z)_{z\in\mathbb{R}^N}$ such that $\sigma(A_z) \subseteq \Sigma_{\theta_0}$, where $\Sigma_{\theta_0} := \{\lambda \in \mathbb{C} : \lambda \neq 0, |\arg(\lambda)| < \theta_0\}$ is the open sector of angle θ_0. Let $\eta_0 := \pi - \theta_0$ and $\eta \in (\frac{\pi}{2}, \eta_0)$. For $r > 0$, define the curve γ_r as the union of the three curves γ_r^1, γ_r^2, and γ_r^3, where $\gamma_r^1 \colon (-\infty, -r) \to \mathbb{C}, \gamma_r^1(\rho) := -\rho e^{-i\eta}$, $\gamma_r^2 \colon (-\eta, \eta) \to \mathbb{C}, \gamma_r^2(\varphi) := r e^{i\varphi}$, and $\gamma_r^3 \colon (r, \infty) \to \mathbb{C}, \gamma_r^3(\rho) := \rho e^{i\eta}$.

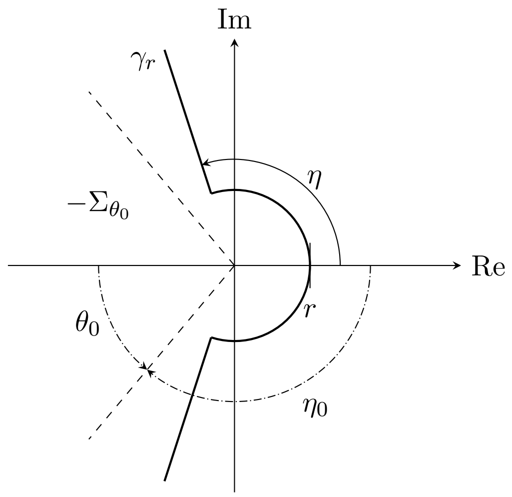

Figure 3.1: The curve γ_r.

We summarise some consequences of Assumption 3.59 in the following lemma.

Lemma 3.61. *Suppose that Assumption 3.59 holds for some $\ell \in \mathbb{N}_0$. Then all the following statements hold for all $\alpha \in \mathcal{N}$ with $|\alpha| \leq \ell$ and $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$, with all constants independent of α and z.*

(a) *We have $\|\rho(z)^{\alpha/2} \partial_z^\alpha A_z\|_{\mathcal{L}(D,H)} \leq C_\ell$ for some $C_\ell \geq 0$.*

(b) *For all $r > 0$ there exists $C_{\ell,r} \geq 0$ such that $\|\rho(z)^{\alpha/2} \partial_z^\alpha (\lambda + A_z)^{-1}\|_{\mathcal{L}(H,D)} \leq C_{\ell,r}$ for all $\lambda \in \gamma_r$. For $\alpha = 0$, the estimate holds with $C_{0,r} = 2 + \frac{1}{r}$.*

(c) *For all $T > 0$, there exists $C_{T,\ell} \geq 0$ such that for all $\lambda \in \gamma_{\frac{1}{t}}$ with $t \in (0,T]$, $\|\rho(z)^{\alpha/2}\partial_z^\alpha(\lambda + A_z)^{-1}\|_{\mathcal{L}(H)} \leq \frac{C_{T,\ell}}{|\lambda|}$.*

(d) *For all $T > 0$ there exists $C_{T,\ell} \geq 0$ such that $\left\|\rho(z)^{\alpha/2}\partial_z^\alpha S_z(t)\right\|_{\mathcal{L}(H)} \leq C_{T,\ell}$ for all $t \in [0,T]$.*

(e) *For all $T > 0$ there exists $C_{T,\ell} \geq 0$ such that for all $t \in [0,T]$,*

$$\sup_{0\leq\tau\leq t} \|\rho(z)^{\alpha/2}\partial_z^\alpha(\tau A_z S_z(\tau))\|_{\mathcal{L}(H)} \leq C_{T,\ell}.$$

Proof. We start by showing (a). For $u \in D$ and $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$, we estimate

$$\begin{aligned}\|\rho(z)^{\alpha/2}\partial_z^\alpha A_z u\|_H &= \sup_{v\in V, \|v\|_H=1} |\rho(z)^{\alpha/2}\,(\partial_z^\alpha A_z u\,|\,v)_H\,| = \sup_{v\in V,\|v\|_H=1} |\rho(z)^{\alpha/2}\partial_z^\alpha a_z(u,v)| \\ &\leq \sup_{v\in V,\|v\|_H=1} C_\ell\|u\|_D\|v\|_H = C_\ell\|u\|_D.\end{aligned}$$

By uniform sectoriality of $(A_z)_{z\in\mathbb{R}^N}$,

$$\|\lambda(\lambda + A_z)^{-1}\|_{\mathcal{L}(H)} \leq 1 \tag{3.4.10}$$

for all $\lambda \in \Sigma_\eta \supseteq \gamma_r$ with $\eta < \eta_0$. This implies (b) for $\alpha = 0$ via $\rho(z)^{0/2} = 1$ and

$$\begin{aligned}\|(\lambda + A_z)^{-1}u\|_D &= \sqrt{\|[I - \lambda(\lambda + A_z)^{-1}]u\|_H^2 + \|(\lambda + A_z)^{-1}u\|_H^2} \\ &\leq \sqrt{2^2 + \frac{1}{|\lambda|^2}}\|u\|_H \leq \sqrt{4 + \frac{1}{r^2}}\|u\|_H \leq \left(2 + \frac{1}{r}\right)\|u\|_H\end{aligned}$$

for all $\lambda \in \gamma_r$, where we have used that $|\lambda| \geq r$ on γ_r and $A_z(\lambda + A_z)^{-1} = I - \lambda(\lambda + A_z)^{-1}$. To show (b) for general $|\alpha| \leq \ell$, we assume without loss of generality that $\alpha - e_1 \in \mathcal{N}$. We note that by Leibniz' rule and iteratively rewriting the derivatives of the resolvent, for $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$, we obtain

$$\begin{aligned}\partial_z^\alpha(\lambda + A_z)^{-1} &= \partial_z^{\alpha-e_1}(-(\lambda + A_z)^{-1}(\partial_z^{e_1}A_z)(\lambda + A_z)^{-1}) \\ &= -\sum_{\beta\leq\alpha-e_1}\binom{\alpha - e_1}{\beta}\partial_z^{\alpha-e_1-\beta}(\lambda + A_z)^{-1}\sum_{\gamma\leq\beta}\binom{\beta}{\gamma}(\partial_z^{\beta-\gamma+e_1}A_z)\partial_z^\gamma(\lambda + A_z)^{-1} \\ &= \ldots = (\lambda + A_z)^{-1}\sum_{P_j}\Big[\prod_{i=1}^{p_j} C_{\alpha,i,j}\Big(\partial_z^{\beta_i^j}A_z\Big)(\lambda + A_z)^{-1}\Big]\end{aligned}$$

for some $C_{\alpha,i,j} \in \mathbb{R}$. The sum is taken over all partitions $P_j = (\beta_i^j)_{1\leq i\leq p_j}$ of α such that $\alpha = \beta_1^j + \ldots + \beta_{p_j}^j$ for some $1 \leq p_j \leq |\alpha|$ and with $\beta_i^j \neq 0$ for all i,j. Consequently, also

$$\rho(z)^{\alpha/2}\partial_z^\alpha(\lambda + A_z)^{-1} = (\lambda + A_z)^{-1}\sum_{P_j}\Big[\prod_{i=1}^{p_j} C_{\alpha,i,j}\Big(\rho(z)^{\beta_i^j/2}\partial_z^{\beta_i^j}A_z\Big)(\lambda + A_z)^{-1}\Big]. \tag{3.4.11}$$

Since as many derivatives (of some order) of A_z appear (weighted with the respective power of $\rho(z)$) as resolvents of A_z, we can estimate the $\mathcal{L}(H)$-norm of each factor in the product by $|C_{\alpha,i,j}|C_\ell(2+\frac{1}{r})$ by part (a) and the sectoriality estimate (3.4.10). Since sum and product are finite, the estimate $\alpha = 0$ applied to the first factor in (3.4.11) yields the claim.

We continue with (c). Let $r = \frac{1}{t}$. For $\alpha = 0$, (3.4.10) yields the claim with $C_{T,\ell} = 1$. For general α, we note that the $\mathcal{L}(H)$-norm of each factor of the product in (3.4.11) is bounded by $C_\ell(2+\frac{1}{r}) = C_\ell(2+t) \le C_\ell(2+T)$. The $\mathcal{L}(H)$-norm of (3.4.11) can thus be estimated by $C_{T,\ell}\|(\lambda + A_z)^{-1}\|_{\mathcal{L}(H)}$, where the norm is bounded by $\frac{1}{|\lambda|}$ due to (3.4.10).

We pass to the proof of (d). Observe that (d) and (e) are trivially satisfied for $t = 0$. Hence, let $t > 0$ and $r = \frac{1}{t}$. The Laplace inversion formula (see Theorem 2.13) yields the representation

$$S_z(t)u = \frac{1}{2\pi\mathrm{i}} \int_{\gamma_r} \mathrm{e}^{t\lambda}(\lambda + A_z)^{-1}u \,\mathrm{d}\lambda$$

of the semigroup for all $t > 0$, $u \in \mathrm{D}(A)$, and $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$. Interchanging differentiation and integration yields that $z \mapsto S_z(t)u$ is continuously differentiable up to order ℓ. Furthermore, by part (c) there exists $C_{T,\ell} \ge 0$ with

$$\big\|\rho(z)^{\alpha/2}\partial_z^\alpha S_z(t)u\big\|_H \le \frac{1}{2\pi}\int_{\gamma_r} |\mathrm{e}^{t\lambda}|\|\rho(z)^{\alpha/2}\partial_z^\alpha(\lambda + A_z)^{-1}u\|_H \,\mathrm{d}\lambda \le \frac{C_{T,\ell}}{2\pi}\|u\|_H \int_{\gamma_r} \frac{|\mathrm{e}^{\lambda t}|}{|\lambda|}\,\mathrm{d}\lambda$$

for all $t \in [0, T]$, $u \in H$, $|\alpha| \le \ell$, and $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$. It remains to estimate the curve integral uniformly in t. On γ_r^3, due to $r = \frac{1}{t}$, we obtain

$$\int_{\gamma_r^3} \frac{|\mathrm{e}^{t\lambda}|}{|\lambda|}\,\mathrm{d}\lambda = \int_r^\infty \frac{\mathrm{e}^{t\rho\cos(\eta)}}{\rho}\,\mathrm{d}\rho \le \frac{1}{r}\int_r^\infty \mathrm{e}^{t\rho\cos(\eta)}\,\mathrm{d}\rho = \frac{1}{rt}\frac{\mathrm{e}^{rt\cos(\eta)}}{|\cos(\eta)|} = \frac{\mathrm{e}^{\cos(\eta)}}{|\cos(\eta)|}.$$

An analogous estimate holds on γ_r^1. The proof of (d) is finished by estimating

$$\int_{\gamma_r^2} \frac{|\mathrm{e}^{t\lambda}|}{|\lambda|}\,\mathrm{d}\lambda = \int_{-\eta}^{\eta} \frac{\mathrm{e}^{rt\cos(\varphi)}}{r}\cdot r\,\mathrm{d}\varphi \le \int_{-\eta}^{\eta} \mathrm{e}^1\,\mathrm{d}\varphi = 2\mathrm{e}\eta.$$

Lastly, we show (e). Leibniz' rule, multiplying by 1, and part (a) yield

$$\begin{aligned}\big\|\rho(z)^{\alpha/2}\partial_z^\alpha(\tau A_z S_z(\tau))u\big\|_H &= \Bigg\|\sum_{\beta\le\alpha}\binom{\alpha}{\beta}(\rho(z)^{(\alpha-\beta)/2}\partial_z^{\alpha-\beta}A_z)A_z^{-1}\tau A_z(\rho(z)^{\beta/2}\partial_z^\beta S_z(\tau))u\Bigg\|_H \\ &\le C_{|\alpha|}C_\ell\|A_z^{-1}\|_{\mathcal{L}(H,D)}\max_{\beta\le\alpha}\|\tau A_z(\rho(z)^{\beta/2}\partial_z^\beta S_z(\tau))u\|_H.\end{aligned}$$

The coercivity assumed implies $\|A_z u\|_H \ge \kappa\|u\|_H$ for $u \in \mathrm{D}(A_z)$ via the embedding $V \hookrightarrow H$, and thus $\|A_z^{-1}\|_{\mathcal{L}(H)} \le \frac{1}{\kappa}$. Consequently, $\|A_z^{-1}u\|_D^2 = \|u\|_H^2 + \|A_z^{-1}u\|_H^2 \le (1+\frac{1}{\kappa^2})\|u\|_H^2$. Hence, $\|A_z^{-1}\|^2_{\mathcal{L}(H,D)} \le 1 + \frac{1}{\kappa^2}$. It remains to estimate the last factor. Again using the Laplace inversion formula, for $r = T^{-1}$, (3.4.11) with all partitions P_j of β instead of α allows us to rewrite

$$\tau A_z(\rho(z)^{\beta/2}\partial_z^\beta S_z(\tau))u = \frac{1}{2\pi\mathrm{i}}\int_{\gamma_r} \tau\mathrm{e}^{\tau\lambda}A_z(\rho(z)^{\beta/2}\partial_z^\beta(\lambda + A_z)^{-1})u\,\mathrm{d}\lambda$$

$$= \frac{C_\ell}{2\pi \mathrm{i}} \int_{\gamma_r} \tau \mathrm{e}^{\tau\lambda} A_z(\lambda + A_z)^{-1} \sum_{P_j} \Big[\prod_{i=1}^{p_j} \Big(\rho(z)^{\beta_i^j/2} \partial_z^{\beta_i^j} A_z \Big) (\lambda + A_z)^{-1} \Big] u \,\mathrm{d}\lambda.$$

In the proof of (c), we have already shown that the $\mathcal{L}(H)$-norm of the sum over all partitions P_j is bounded by some $C_{T,\ell} \geq 0$. We further recall that

$$\|A_z(\lambda + A_z)^{-1}\|_{\mathcal{L}(H)} = \|I - \lambda(\lambda + A_z)^{-1}\|_{\mathcal{L}(H)} \leq 2$$

by the sectoriality estimate (3.4.10). Hence,

$$\|\rho(z)^{\beta/2} \tau A_z \partial_z^\beta S_z(\tau) u\|_H \leq \frac{C_\ell C_{T,\ell}}{\pi} \|u\|_H \int_{\gamma_r} \tau |\mathrm{e}^{\tau\lambda}| \,\mathrm{d}\lambda \leq \frac{C_\ell C_{T,\ell}}{\pi} \left(\frac{2}{|\cos(\eta)|} + 2\mathrm{e}\eta \right) \|u\|_H,$$

where the last step is obtained from an estimation of the curve integral. □

We continue with the non-symmetric case, first collecting some auxiliary statements. Denote the commutator of two operators A and B by $[A, B] := AB - BA$ on its natural domain.

Lemma 3.62. *Suppose that Assumption 3.59 holds for some $\ell \in \mathbb{N}_0$. Then the following statements hold for all $\alpha \in \mathcal{N}$ with $|\alpha| \leq \ell$ and $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$, with all constants independent of α and z.*

(a) $\|(\lambda + A_z)^{-1}\|_{\mathcal{L}(H,D)} \leq C_\kappa$ for some $C_\kappa \geq 0$ for all $\lambda \in \gamma_r \cup \{0\}$ with $r > 0$.

(b) For all $\nu \in (0, 1)$, there is $C_{\kappa,\nu} \geq 0$ such that $\|(\lambda + A_z)^{-1}\|_{\mathcal{L}(H,D^{1-\nu})} \leq \frac{C_{\kappa,\nu}}{|\lambda|^\nu}$ for all $\lambda \in \gamma_r$ with $r > 0$.

(c) For all $\varepsilon \in (0, 1]$, there is $C_{\kappa,\varepsilon} \geq 0$ such that $\|(\lambda + A_z)^{-1}\|_{\mathcal{L}(D^\varepsilon,D)} \leq \frac{C_{\kappa,\varepsilon}}{|\lambda|^\varepsilon}$ for all $\lambda \in \gamma_r$ with $r > 0$.

(d) $\|[\rho(z)^{\beta/2} \partial_z^\beta A_z, A_z^{-1}]\|_{\mathcal{L}(D,H)} \leq C_{\ell,\kappa}$ for some $C_{\ell,\kappa} \geq 0$ for all $\beta \leq \alpha$.

Proof. (a) We start with the case $\lambda = 0$. Note that $\|A_z u\|_H \geq \operatorname{Re} a_z(u, \|u\|_H^{-1} u) \geq \kappa \|u\|_H$ by coercivity, which in turn implies $\|A_z^{-1}\|_{\mathcal{L}(H)} \leq \kappa^{-1}$. Thus, $\|A_z^{-1} u\|_D^2 = \|u\|_H^2 + \|A_z^{-1} u\|_H^2 \leq (1 + \kappa^{-2}) \|u\|_H^2$. For $\lambda \in \gamma_r$ for some $r > 0$, by (3.4.10) we rewrite and estimate

$$\|(\lambda + A_z)^{-1} u\|_H = \|[I - \lambda(\lambda + A_z)^{-1}] A_z^{-1} u\|_H \leq \frac{2}{\kappa} \|u\|_H.$$

Together with $\|A_z(\lambda + A_z)^{-1} u\|_H = \|[I - \lambda(\lambda + A_z)^{-1}] u\|_H \leq 2\|u\|_H$, this implies the desired bound with $C_\kappa = 2\sqrt{1 + \kappa^{-2}}$.

(b) This follows by interpolation from the sectoriality estimate (3.4.10) and part (a).

(c) Note that $\|(\lambda + A_z)^{-1} u\|_D^2 = \|(\lambda + A_z)^{-1} A_z u\|_H^2 + \|(\lambda + A_z)^{-1} u\|_H^2 \leq \frac{1}{|\lambda|} \|u\|_D^2$ by (3.4.10) and $\|(\lambda + A_z)^{-1} u\|_D \leq C_\kappa \|u\|_H$ by part (a). Interpolation yields the desired inequality with constant $C_{\kappa,\varepsilon} = C_\kappa^{1-\varepsilon}$.

(d) From Lemma 3.61(a), $\|A_z^{-1}\|_{\mathcal{L}(H)} \leq \kappa^{-1}$, and part (a), we deduce

$$\begin{aligned}\|[\rho(z)^{\beta/2}\partial_z^\beta A_z, A_z^{-1}]u\|_H &\leq \|\rho(z)^{\beta/2}(\partial_z^\beta A_z)A_z^{-1}u\|_H + \|A_z^{-1}\rho(z)^{\beta/2}(\partial_z^\beta A_z)u\|_H \\ &\leq C_\ell C_\kappa\|u\|_H + \frac{1}{\kappa}C_\ell\|u\|_D \leq \max\{C_\kappa, \kappa^{-1}\}C_\ell\|u\|_D. \qquad \square\end{aligned}$$

For the non-symmetric case, a stronger assumption is made, allowing to consider derivatives of the generator in stronger norms than the H-norm. It allows us to show the pointwise estimates required in the non-symmetric case.

Assumption 3.63. *Suppose that Assumption 3.59 holds for some $\ell \in \mathbb{N}_0$ and $\overline{\alpha} \geq 2$. Further assume that there exists a constant $C_{\ell,D} \geq 0$ such that*

$$\|\rho(z)^{\alpha/2}\partial_z^\alpha A_z u\|_D \leq C_{\ell,D}\|u\|_{D^2}$$

for $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$ and for all $u \in D$ and $\alpha \in \mathcal{N}$ with $|\alpha| \leq \ell$.

Lemma 3.64. *Suppose that Assumption 3.59 holds for some $\ell \in \mathbb{N}_0$. Then the following statements hold for all $\alpha \in \mathcal{N}$ with $|\alpha| \leq \ell$ and $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$, with all constants independent of α and z.*

(a) We have $\|\rho(z)^{\alpha/2}\partial_z^\alpha A_z^\theta\|_{\mathcal{L}(D^\varepsilon,H)} \leq C_{\ell,\theta,\varepsilon}$ for all $0 < \theta < \varepsilon < 1$ for some $C_{\ell,\theta,\varepsilon} \geq 0$.

(b) For all $T > 0$ and $0 < \nu < \theta < 1$ there exists $C_{T,\ell,\nu,\theta} \geq 0$ such that $\sup_{0\leq\tau\leq t}\|\rho(z)^{\alpha/2}\partial_z^\alpha(\tau^{1-\nu}A_z^{1-\theta}S_z(\tau))\|_{\mathcal{L}(H)} \leq C_{T,\ell,\nu,\theta}$ for all $t \in [0,T]$.

(c) If, additionally, Assumption 3.63 is satisfied for ℓ, then for all $\varepsilon \in (0,1)$ there is $C_{\ell,\varepsilon} \geq 0$ such that $\|\rho(z)^{\alpha/2}\partial_z^\alpha A_z u\|_{\mathcal{L}(D^{1+\varepsilon},D^\varepsilon)} \leq C_{\ell,\varepsilon}$.

Proof. For the sake of readability, we omit the weights ρ in this proof. Splitting $\rho(z)^{\alpha/2}$ according to the order of the derivatives yields the correct prefactors to apply previous statements. We start by showing part (a). For $u \in D$, we make use of Balakrishnan's formula for fractional powers [60, Proposition 3.1.12]

$$A_z^\theta u = \frac{\sin(\pi\theta)}{\pi}\int_0^\infty A_z(t+A_z)^{-1}u\frac{\mathrm{d}t}{t^{1-\theta}}. \tag{3.4.12}$$

Step 1: Show that $K := \partial_z^\alpha A_z^\theta|_D \in \mathcal{L}(D,H)$. Let $u \in D$. Differentiating (3.4.12), Leibniz' rule, the formula for derivatives of the resolvent from the proof of Lemma 3.61(c), and applying (3.4.12) a second time yield

$$\begin{aligned}\partial_z^\alpha A_z^\theta u &= \frac{\sin(\pi\theta)}{\pi}\int_0^\infty\sum_{\beta\leq\alpha}\binom{\alpha}{\beta}(\partial_z^{\alpha-\beta}A_z)(t+A_z)^{-1}\sum_{P_j}\Big[\prod_{i=1}^{p_j}C_{\beta,i,j}(\partial_z^{\gamma_i^j}A_z)(t+A_z)^{-1}\Big]u\,\frac{\mathrm{d}t}{t^{1-\theta}} \\ &= (\partial_z^\alpha A_z)A_z^{-1}A_z^\theta u + \frac{\sin(\pi\theta)}{\pi}\Big(\int_0^1 + \int_1^\infty\Big)\sum_{0<\beta\leq\alpha}\cdots\frac{\mathrm{d}t}{t^{1-\theta}}. \end{aligned} \tag{3.4.13}$$

Here, we sum over all partitions $P_j = (\gamma_i^j)_{1\leq i\leq p_j}$ of β with $\beta = \gamma_1^j + \ldots + \gamma_{p_j}^j$ for all j and some $1 \leq p_j \leq |\beta|$ as well as $\gamma_i^j \neq 0$ for all i,j. Via Lemmas 3.61(a) and 3.62(a), we estimate the norm of the first term by $C_\ell C_\kappa \|u\|_{D^\theta}$. Proceeding analogously for the integral from 0 to 1 results in the bound

$$\frac{\sin(\pi\theta)}{\pi}\cdot\frac{1}{\theta}\sum_{0<\beta\le\alpha}\binom{\alpha}{\beta}2C_{|\alpha-\beta|}C_\kappa\sum_{P_j}\Big[\prod_{i=1}^{p_j}2C_{\beta,i,j}C_{|\gamma_i^j|}C_\kappa\Big]\|u\|_H$$

For the integral from 1 to ∞, we estimate analogously apart from the last factor for $i = p_j$. Instead, we estimate $\|(t+A_z)^{-1}u\|_D \le \frac{C_{\kappa,1}}{t}\|u\|_D$ by Lemma 3.62(c) for $\varepsilon = 1$. This yields the additional decay in t for the integral to converge. Altogether, for a suitable constant $C_{\ell,\kappa,\theta} \ge 0$, we find

$$\|Ku\|_H = \|\partial_z^\alpha A_z^\theta u\|_H \le C_{\ell,\kappa,\theta}(\|u\|_{D^\theta} + \|u\|_D) \quad (u \in D)$$

and thus $K \in \mathcal{L}(D, H)$.

Step 2: Let $\varepsilon \in (\theta, 1)$. Show that

$$KA_z^{-1}u = A_z^{-1}Ku + R_z u \quad (u \in D) \tag{3.4.14}$$

for a remainder term $R_z \in \mathcal{L}(D^\varepsilon, H)$ $\mathbb{P}_Z$-almost surely. From (3.4.13), we deduce

$$\begin{aligned} R_z u = \frac{\sin(\pi\theta)}{\pi}\int_0^\infty &\sum_{\beta\le\alpha}\binom{\alpha}{\beta}[\partial_z^{\alpha-\beta}A_z, A_z^{-1}](t+A_z)^{-1}\sum_{P_j}\Big[\prod_{i=1}^{p_j}C_{\beta,i,j}(\partial_z^{\gamma_i^j}A_z)(t+A_z)^{-1}\Big] \\ &+ (\partial_z^{\alpha-\beta}A_z)(t+A_z)^{-1}\sum_{P_j}\sum_{r=1}^{p_j}\Big[\prod_{i=1}^{r-1}C_{\beta,i,j}(\partial_z^{\gamma_i^j}A_z)(t+A_z)^{-1}\Big] \\ &\cdot C_{\beta,r,j}[\partial_z^{\gamma_r^j}A_z, A_z^{-1}](t+A_z)^{-1}\Big[\prod_{i=r+1}^{p_j}C_{\beta,i,j}(\partial_z^{\gamma_i^j}A_z)(t+A_z)^{-1}\Big]u\,\frac{\mathrm{d}t}{t^{1-\theta}} \end{aligned} \tag{3.4.15}$$

To estimate the norm in H of the integral from 0 to 1, we use Lemma 3.62(d) for the commutators and estimate the remaining terms as in Step 1. This results in an upper bound in terms of $\|u\|_H$. For the integral from 1 to ∞, we use Lemma 3.62(c) to estimate the last factor by

$$\|(t+A_z)^{-1}u\|_D \le C_{\kappa,\varepsilon}\frac{\|u\|_{D^\varepsilon}}{t^\varepsilon}.$$

Since $\varepsilon > \theta$, we have $\int_1^\infty t^{-1+\theta-\varepsilon}\,\mathrm{d}t < \infty$ and thus the integral from 1 to ∞ in (3.4.15) can be estimated by $C_{\ell,\kappa,\theta}C_{\kappa,\varepsilon}(\varepsilon-\theta)^{-1}\|u\|_{D^\varepsilon}$. In conclusion, $\|R_z u\|_H \le C_{\ell,\kappa,\theta,\varepsilon}\|u\|_{D^\varepsilon}$ for a suitable constant for all $u \in D^\varepsilon$. That is, $R_z \in \mathcal{L}(D^\varepsilon, H)$.

Step 3: We show that the closure of K yields an operator $\overline{K} \in \mathcal{L}(D^\varepsilon, H)$. Consider a sequence $(u_n)_{n\in\mathbb{N}}$ in D such that $u_n \to 0$ in D^ε and suppose that $Ku_n \to y$ in H for some $y \in H$. By Step 2, $u_n \to 0$ in D^ε implies $R_z u_n \to 0$ in H. Since $A_z^{-1} \in \mathcal{L}(H, D)$ by Lemma 3.62(a), also $A_z^{-1} \in \mathcal{L}(D^\varepsilon, D)$, and thus $A_z^{-1}u_n \to 0$ in D. By Step 1, this implies $KA_z^{-1}u_n \to 0$ in H. Rewriting (3.4.14), we deduce

$$A_z^{-1}Ku_n = KA_z^{-1}u_n - R_z u_n \to 0 \text{ in } H.$$

Moreover, the term on the left converges to $A_z^{-1}y$ in H by assumption. Consequently, $A_z^{-1}y = 0$ and thus $y = 0$, meaning that $Ku_n \to 0$ in H. Hence, the closure of $K = \partial_z^\alpha A_z^\theta|_D$

is a linear and bounded operator from D^ε to H. The norm bounds from Steps 1 and 2 further imply the desired norm bound $\|\partial_z^\alpha A_z^\theta u\|_H \le C_{\ell,\theta,\varepsilon}\|u\|_{D^\varepsilon}$, where we have omitted the dependence on κ in the constant.

We continue with the proof of part (b). By Leibniz' formula and part (a),

$$\sup_{0\le\tau\le t}\|\partial_z^\alpha(\tau^{1-\nu}A_z^{1-\theta}S_z(\tau)u)\|_H \le \sum_{\gamma\le\alpha}\binom{\alpha}{\gamma}C_{|\alpha-\gamma|,1-\theta,1-\nu}\sup_{0\le\tau\le t}\|\tau^{1-\nu}\partial_z^\gamma S_z(\tau)u\|_{D^{1-\nu}},$$

since $\|\partial_z^{\alpha-\gamma}A_z^{1-\theta}\|_{\mathcal{L}(D^{1-\nu},H)} \le C_{|\alpha-\gamma|,1-\theta,1-\nu}$ by Lemma 3.61(a). It thus suffices to consider the supremum of the graph norm in time. By Lemma 3.61(d), the H-norm of this term is bounded by $\sup_{0\le\tau\le t}\tau^{1-\nu}C_{T,|\gamma|} \le \max\{t^{1-\nu},1\}C_{T,\ell}$. To rewrite the remaining term, we again use the Laplace inversion formula and the formula for derivatives of the resolvent. Because $\|A_z^{1-\nu}\|_{\mathcal{L}(D^{1-\nu},H)} \le 1$ by definition of the graph norm, Lemma 3.62(a) and (b), and Lemma 3.61(a) result in

$$\begin{aligned}
&\sup_{0\le\tau\le t}\|\tau^{1-\nu}A_z^{1-\nu}\partial_z^\gamma S_z(\tau)u\|_H \\
&\le \sup_{0\le\tau\le t}\Big\|\frac{1}{2\pi\mathrm{i}}\int_{\gamma_{1/T}}\tau^{1-\nu}\mathrm{e}^{\tau\lambda}A_z^{1-\nu}(\lambda+A_z)^{-1}\sum_{P_j}\Big[\prod_{i=1}^{p_j}C_{\gamma,i,j}(\partial_z^{\delta_i^j}A_z)(\lambda+A_z)^{-1}\Big]u\,\mathrm{d}\lambda\Big\|_H \\
&\le \frac{C_\kappa}{2\pi}\sum_{P_j}\Big[\prod_{i=1}^{p_j}C_{\gamma,i,j}C_{|\delta_i^j|}C_\kappa\Big]\|u\|_H\sup_{0\le\tau\le t}\int_{\gamma_{1/T}}\frac{\tau^{1-\nu}|\mathrm{e}^{\tau\lambda}|}{|\lambda|^\nu}\,\mathrm{d}\lambda,
\end{aligned}$$

where P_j are partitions of γ. This yields the desired statement, provided that the curve integrals are uniformly bounded in τ. Indeed, we calculate that on $\gamma_{1/T}^3$,

$$\sup_{0\le\tau\le t}\int_{\gamma_{1/T}^3}\tau^{1-\nu}\frac{|\mathrm{e}^{\tau\lambda}|}{|\lambda|^\nu}\,\mathrm{d}\lambda = \sup_{0\le\tau\le t}\int_{\tau/T}^\infty\frac{\mathrm{e}^{\zeta\cos(\eta)}}{\zeta^\nu}\,\mathrm{d}\zeta \le \int_0^1\frac{\mathrm{e}^{\cos(\eta)}}{\zeta^\nu}\,\mathrm{d}\zeta + \int_1^\infty \mathrm{e}^{\zeta\cos(\eta)}\,\mathrm{d}\zeta < \infty.$$

The estimate on $\gamma_{1/T}^1$ is analogous. Lastly, on $\gamma_{1/T}^2$, we estimate

$$\sup_{0\le\tau\le t}\int_{\gamma_{1/T}^2}\tau^{1-\nu}\frac{|\mathrm{e}^{\tau\lambda}|}{|\lambda|^\nu}\,\mathrm{d}\lambda = \sup_{0\le\tau\le t}\int_{-\eta}^{\eta}\tau^{1-\nu}\frac{\mathrm{e}^{\cos(\varphi)\tau/T}}{(1/T)^\nu}\frac{\mathrm{d}\varphi}{T} \le 2\mathrm{e}\eta\sup_{0\le\tau\le t}\Big(\frac{\tau}{T}\Big)^{1-\nu} \le 2\mathrm{e}\eta < \infty.$$

Finally, part (c) follows by interpolation of $\|\partial_z^\alpha A_z u\|_H \le C_\ell\|u\|_D$ by Lemma 3.61(a) and $\|\partial_z^\alpha A_z u\|_D \le C_{\ell,D}\|u\|_{D^2}$ by Assumption 3.63 with the constant $C_{\ell,\varepsilon} = C_\ell^{1-\varepsilon}C_{\ell,D}^\varepsilon$. □

3.4.3.2 Mapping properties of A and its resolvent

From the pointwise estimates for A_z, we now derive estimates on $\mathbf{A}$, its resolvent, and the semigroup generated by $-\mathbf{A}$ in suitable Sobolev spaces.

Proposition 3.65. *Suppose that Assumption 3.59 holds for some $\ell \in \mathbb{N}_0$. Then for all $q \in \{0,\dots,\ell\}$, $r > 0$, and $T > 0$ there are constants $C_q, C_{q,r}, C_{T,q} \ge 0$ such that the following estimates hold for all $\mathbf{f} \in H_\rho^q(\mathbb{R}^N,\mathbb{P}_Z;H)$.*

(a) $\|\mathbf{Af}\|_{H_\rho^q(\mathbb{R}^N,\mathbb{P}_Z;H)} \le C_q\|\mathbf{f}\|_{H_\rho^q(\mathbb{R}^N,\mathbb{P}_Z;D)}$ *if additionally* $\mathbf{f} \in H_\rho^q(\mathbb{R}^N,\mathbb{P}_Z;D)$. *In particular,* $\|\mathbf{Af}\|_{\mathbf{H}} \le C_0\|\mathbf{f}\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D)}$.

(b) $\|(\lambda+\mathbf{A})^{-1}\mathbf{f}\|_{H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)} \le C_{q,r}\|\mathbf{f}\|_{H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)}$ *for all* $\lambda\in\gamma_r$. *In particular,* $\|(\lambda+\mathbf{A})^{-1}\mathbf{f}\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D)} \le C_{0,r}\|\mathbf{f}\|_{\mathbf{H}}$.

(c) $\|\mathbf{S}(t)\mathbf{f}\|_{\mathcal{L}(H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;H))} \le C_{T,q}\|\mathbf{f}\|_{H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)}$ *for all* $t\in[0,T]$.

(d) $\sup_{0\le\tau\le t}\|\tau\mathbf{A}\mathbf{S}(\tau)\mathbf{f}\|_{H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)} \le C_{T,q}\|\mathbf{f}\|_{H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)}$ *for all* $t\in[0,T]$.

We need a lemma to lift the pointwise estimates to subspaces of $\mathbf{H}$.

Lemma 3.66. *Let* H_1 *and* H_2 *be Hilbert spaces,* $\ell\in\mathbb{N}_0$*, and* $F\colon\mathbb{R}^N\times H_1\to H_2$ *such that* $F(z,\cdot)\in\mathcal{L}(H_1,H_2)$ *for* $\mathbb{P}_Z$*-almost every* $z\in\mathbb{R}^N$ *and* $\mathbb{P}_Z\text{-ess sup}_{z\in\mathbb{R}^N}\|F(z,\cdot)\|_{\mathcal{L}(H_1,H_2)}<\infty$*. Further, suppose that* $F(\cdot,u)\in C^\ell(\mathbb{R}^N;H_2)$ *and for some* $c_\ell\ge0$*,* $\left\|\rho(z)^{\alpha/2}\partial_z^\alpha F(z,u)\right\|_{H_2}\le c_\ell\|u\|_{H_1}$ *for all* $u\in H_1$*,* $|\alpha|\le\ell$*, and* $\mathbb{P}_Z$*-almost every* $z\in\mathbb{R}^N$*. Let* $G\in H^\ell_\rho(\mathbb{R}^N,\mathbb{P}_Z;H_1)$*. Then* $F(\cdot,G(\cdot))\in H^\ell_\rho(\mathbb{R}^N,\mathbb{P}_Z;H_2)$ *and there exists* $C_\ell\ge0$ *such that*

$$\|F(\cdot,G(\cdot))\|_{H^\ell_\rho(\mathbb{R}^N,\mathbb{P}_Z;H_2)} \le C_\ell\|G\|_{H^\ell_\rho(\mathbb{R}^N,\mathbb{P}_Z;H_1)}.$$

Proof. By induction, the product rule and the chain rule, or more precisely, their generalisations Leibniz' rule and Faà di Bruno's formula, we obtain the assertion. We sketch the proof for $\ell=1$, since the induction step $\ell\to\ell+1$ follows analogously using the two rules. The induction start is an immediate consequence of $\|F(z,u)\|_{H_2}\le c_0\|u\|_{H_1}$ for $\mathbb{P}_Z$-almost every $z\in\mathbb{R}^N$ taking $u=G(z)\in H_1$. Since by assumption, $F(z,\cdot)$ is linear for $\mathbb{P}_Z$-almost every $z\in\mathbb{R}^N$, we obtain $\partial_2F(z,\cdot)=F(z,\cdot)\in\mathcal{L}(H_1,H_2)$ for $\mathbb{P}_Z$-almost every $z\in\mathbb{R}^N$ and higher derivatives vanish. For $\mathbb{P}_Z$-almost every $z\in\mathbb{R}^N$, we can thus find $C\ge0$ such that

$$\begin{aligned}\|\partial_z^{e_j}F(z,G(z))\|_{H_2} &= \|\partial_1^{e_j}F(z,G(z))+[\partial_2F(z,G(z))](\partial_z^{e_j}G(z))\|_{H_2}\\ &\le \|\partial_1^{e_j}F(z,G(z))\|_{H_2}+\Big(\mathbb{P}_Z\text{-}\operatorname*{ess\,sup}_{z\in\mathbb{R}^N}\|F(z,G(z))\|_{\mathcal{L}(H_1,H_2)}\Big)\|\partial_z^{e_j}G(z)\|_{H_1}.\end{aligned}$$

Multiplying with $\rho(z)^{e_j/2}=(\rho(z))_j^{1/2}\in\mathbb{R}$, from the assumptions we can deduce

$$\|\rho(z)^{e_j/2}\partial_z^{e_j}F(z,G(z))\|_{H_2}\le c_\ell\|G(z)\|_{H_1}+C\|\rho(z)^{e_j/2}\partial_z^{e_j}G(z)\|_{H_1}$$

Integrating in z, noting that $\rho(z)^0=1$ and summing over $j\in\{1,\dots,N\}$ yields that, for some $C_\ell\ge0$,

$$[F(\cdot,G(\cdot))]^2_{H^1_\rho(\mathbb{R}^N,\mathbb{P}_Z;H_2)}\le C_\ell\|G\|^2_{H^1_\rho(\mathbb{R}^N,\mathbb{P}_Z;H_1)}$$

for the seminorm $[\cdot]_{H^1_\rho(\mathbb{R}^N,\mathbb{P}_Z;H_2)}$ in $H^1_\rho(\mathbb{R}^N,\mathbb{P}_Z;H_2)$. The norm estimate follows from the induction assumption, i.e., for $\ell=1$ from the induction start. □

Proof of Proposition 3.65. All statements are consequences of Lemma 3.66 with F chosen as below and $G=\mathbf{f}$, which can be applied due to Lemma 3.61(a), (b), (d), and (e), respectively.

(a) $F\colon\mathbb{R}^N\times D\to H$, $F(z,u):=A_zu$

(b) $F\colon\mathbb{R}^N\times H\to D$, $F(z,u):=(\lambda+A_z)^{-1}u$

(c) $F\colon\mathbb{R}^N\times H\to H$, $F(z,u):=S_z(t)u$

(d) $F\colon \mathbb{R}^N \times H \to H$, $F(z,u) := \sup_{0\le\tau\le t} \tau A_z S_z(\tau)u$ □

Analogously, we deduce the estimates in subspaces of $\mathbf{H}$ for the non-symmetric case.

Proposition 3.67. *Suppose that Assumption 3.63 holds for some $\ell \in \mathbb{N}_0$. Then for all $q \in \{0,\dots,\ell\}$, $0<\nu<\theta<\varepsilon<1$, and $T>0$ there are constants $C_{q,\varepsilon}, C_{q,\theta,\varepsilon}, C_{T,q,\nu,\theta} \ge 0$ such that the following estimates hold for all $\mathbf{f} \in H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)$.*

(a) $\|\mathbf{A}^\theta \mathbf{f}\|_{H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)} \le C_{q,\theta,\varepsilon}\|\mathbf{f}\|_{H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^\varepsilon)}$ *if additionally* $\mathbf{f} \in H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^\varepsilon)$.

(b) $\sup_{0\le\tau\le t} \|\tau^{1-\nu}\mathbf{A}^{1-\theta}\mathbf{S}(\tau)\mathbf{f}\|_{H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)} \le C_{T,q,\nu,\theta}\|\mathbf{f}\|_{H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)}$ *for all* $t\in[0,T]$.

(c) $\|\mathbf{A}\mathbf{f}\|_{H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^\varepsilon)} \le C_{q,\varepsilon}\|\mathbf{f}\|_{H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\varepsilon})}$ *if additionally* $\mathbf{f} \in H^q_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\varepsilon})$.

Proof. Due to Lemma 3.64 (a), (b), and (c), respectively, Lemma 3.66 can be applied and yields the statement of the proposition. □

3.4.3.3 Estimate of the difference of resolvents

We now estimate the difference of the resolvents of $-\mathbf{A}$ and $-\mathbf{R}_n\mathbf{A}\mathbf{R}_n$. The convergence rate obtained for the resolvents will be the key ingredient for the convergence rate in randomness in the following subsection. To be able to speak of convergence, we henceforth consider approximation orders $n\in\mathbb{N}$ rather than $n\in\mathbb{N}_0$.

Proposition 3.68. *Let $\ell\in\mathbb{N}$ and $r>0$. Suppose that Assumption 3.59 holds for 2ℓ. Then there exists $C_{\ell,r}\ge 0$ such that for all $\lambda\in\gamma_r$, we have*

$$\|[R(\lambda,-\mathbf{R}_n\mathbf{A}\mathbf{R}_n)\mathbf{R}_n - R(\lambda,-\mathbf{A})]\mathbf{f}\|_{\mathbf{H}} \le C_{\ell,r} n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)}$$

for all $n\in\mathbb{N}$ and $\mathbf{f}\in H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)$.

Note that convergence of order ℓ requires regularity of order 2ℓ in the sense of $\mathbf{f} \in H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)$. Given sufficient regularity in z, the convergence thus is of arbitrarily high (polynomial) order. We will continue to write C_ℓ for constants depending on the value of 2ℓ rather than ℓ itself. For the proof, we need a lemma on the difference of the generators as preparation.

Lemma 3.69. *Let $\ell\in\mathbb{N}$. Suppose that Assumption 3.59 holds for 2ℓ. Then there exists $C_\ell\ge 0$ such that for all $n\in\mathbb{N}$ and $\mathbf{f}\in H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)$, we have*

$$\|(\mathbf{R}_n\mathbf{A}\mathbf{R}_n - \mathbf{A})\mathbf{f}\|_{\mathbf{H}} \le C_\ell n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)}.$$

Proof. Note that $\mathbf{R}_n$ is contractive on $\mathbf{H}$ (being an orthogonal projection). Proposition 3.65(a) applied to $r\in\{0,2\ell\}$ and the PCE error estimate from Corollary 3.36 applied to both D- and H-valued functions yield

$$\begin{aligned}
\|\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{f} - \mathbf{A}\mathbf{f}\|_{\mathbf{H}} &\le \|\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{f} - \mathbf{R}_n\mathbf{A}\mathbf{f}\|_{\mathbf{H}} + \|\mathbf{R}_n\mathbf{A}\mathbf{f} - \mathbf{A}\mathbf{f}\|_{\mathbf{H}}\\
&\le \|\mathbf{A}(\mathbf{R}_n - I)\mathbf{f}\|_{\mathbf{H}} + \|(\mathbf{R}_n - I)\mathbf{A}\mathbf{f}\|_{\mathbf{H}}\\
&\le C_0\|(\mathbf{R}_n - I)\mathbf{f}\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D)} + \|(\mathbf{R}_n - I)\mathbf{A}\mathbf{f}\|_{\mathbf{H}}\\
&\le C_{\mathrm{z}} n^{-\ell}\big(C_0\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)} + \|\mathbf{A}\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)}\big)\\
&\le C_{\mathrm{z}}(C_0 + C_\ell) n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)}.
\end{aligned}$$

□

Proof of Proposition 3.68. The second resolvent identity and $\mathbf{R}_n^2 = \mathbf{R}_n$ imply

$$R(\lambda, -\mathbf{R}_n\mathbf{A}\mathbf{R}_n)\mathbf{R}_n - R(\lambda, -\mathbf{A}) = R(\lambda, -\mathbf{R}_n\mathbf{A}\mathbf{R}_n)\mathbf{R}_n(\mathbf{A} - \mathbf{R}_n\mathbf{A}\mathbf{R}_n)R(\lambda, -\mathbf{A}).$$

Applying, in this order, Lemma 3.69, Lemma 3.46, and Proposition 3.65(b) for $q = 2\ell$ results in

$$\begin{aligned}\big\|[R(\lambda, -\mathbf{R}_n\mathbf{A}\mathbf{R}_n)\mathbf{R}_n - R(\lambda, -\mathbf{A})]\mathbf{f}\big\|_{\mathbf{H}} &\le \frac{C}{|\lambda|}\big\|(\mathbf{A} - \mathbf{R}_n\mathbf{A}\mathbf{R}_n)R(\lambda, -\mathbf{A})\mathbf{f}\big\|_{\mathbf{H}} \\ \le \frac{C}{r}C_\ell n^{-\ell}\|R(\lambda, -\mathbf{A})\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)} &\le \frac{C}{r}C_\ell C_{\ell,r} n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)}. \qquad \square\end{aligned}$$

3.4.3.4 Error estimate in randomness: A Trotter–Kato argument

Finally, we conclude convergence of the approximate mild solutions via strong convergence of the associated semigroups. In order to show convergence of the associated semigroups, we prove a suitable adaptation of the quantified version of the Trotter–Kato theorem. It relates the convergence rate for the resolvents from the last subsection to a convergence rate for the semigroups. Higher regularity of the initial data is required in the case of non-symmetric forms. Recall that for $n \in \mathbb{N}$ the semigroup $(\mathbf{S}_n(t))_{t\ge 0}$ is generated by $-\mathbf{R}_n\mathbf{A}\mathbf{R}_n$.

Theorem 3.70 (symmetric case)**.** *Let $\ell \in \mathbb{N}$. Suppose that Assumption 3.59 holds for 2ℓ and that $(a_z)_{z\in\mathbb{R}^n}$ is $\mathbb{P}_Z$-almost surely symmetric. Then, for any $T > 0$, there exists a constant $C_{T,\ell} > 0$ such that*

$$\|\mathbf{S}_n(t)\mathbf{R}_n\mathbf{f} - \mathbf{S}(t)\mathbf{f}\|_{\mathbf{H}} \le C_{T,\ell}n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)}$$

for all $\mathbf{f} \in H^{2\ell}_\rho(\mathbb{R}^N, \mathbb{P}_Z; D)$, $n \in \mathbb{N}$, and $t \in [0, T]$. In particular, for the mild solutions $\mathbf{u}_n$ and $\mathbf{u}$ of (3.4.2) *and* (3.4.1) *with $\mathbf{u}_0 \in H^{2\ell}_\rho(\mathbb{R}^N, \mathbb{P}_Z; D)$, respectively, we have*

$$\|\mathbf{u}_n(t) - \mathbf{u}(t)\|_{\mathbf{H}} \le C_{T,\ell}n^{-\ell}\|\mathbf{u}_0\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)} \quad (t \in [0, T]).$$

Proof. The proof is inspired by [74, Prop. 2.3a] for the deterministic case. Let $\mathbf{f} \in H^{2\ell}_\rho(\mathbb{R}^N, \mathbb{P}_Z; D) \subseteq \mathrm{D}(\mathbf{A})$. Consider the error $\mathbf{e}_n(t) \coloneqq [\mathbf{S}_n(t)\mathbf{R}_n - \mathbf{R}_n\mathbf{S}(t)]\mathbf{f}$. Then, the scaled error $\mathbf{s}_n(t) \coloneqq t\mathbf{e}_n(t)$ satisfies

$$\mathbf{s}_n'(t) = -\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{s}_n(t) + \mathbf{e}_n(t) - t\mathbf{R}_n\mathbf{A}\mathbf{R}_n\Delta_n\mathbf{A}\mathbf{S}(t)\mathbf{f} + t\,\square_n\mathbf{S}(t)\mathbf{f} \quad (t > 0), \quad \mathbf{s}_n(0) = 0,$$

where we have used $\mathbf{R}_n^2 = \mathbf{R}_n$ and set

$$\begin{aligned}\Delta_n &\coloneqq \Delta_n(\kappa) \coloneqq R\left(\frac{\kappa}{2}, -\mathbf{A}\right) - R\left(\frac{\kappa}{2}, -\mathbf{R}_n\mathbf{A}\mathbf{R}_n\right)\mathbf{R}_n, \\ \square_n &\coloneqq \square_n(\kappa) \coloneqq \frac{\kappa}{2}\Big[R\Big(\frac{\kappa}{2}, -\mathbf{R}_n\mathbf{A}\mathbf{R}_n\Big)\mathbf{R}_n\mathbf{A} - \mathbf{R}_n\mathbf{A}\mathbf{R}_nR\Big(\frac{\kappa}{2}, -\mathbf{A}\Big)\Big],\end{aligned}$$

noting that $\frac{\kappa}{2}$ is in the respective resolvent sets by coercivity of $\mathbf{a}$ and $\mathbf{a}_n$. Thus, $\mathbf{s}_n$ is given by the variation-of-constants formula

$$\mathbf{s}_n(t) = \int_0^t \mathbf{S}_n(t-\tau)\mathbf{e}_n(\tau)\,\mathrm{d}\tau - \int_0^t \mathbf{S}_n(t-\tau)\mathbf{R}_n\mathbf{A}\mathbf{R}_n\Delta_n\tau\mathbf{A}\mathbf{S}(\tau)\mathbf{f}\,\mathrm{d}\tau$$

$$+ \int_0^t \mathbf{S}_n(t-\tau)\tau\,\Box_n\mathbf{S}(\tau)\mathbf{f}\,\mathrm{d}\tau.$$

Integration by parts in the second integral and division by t result in

$$\begin{aligned}\mathbf{e}_n(t) &= \frac{1}{t}\int_0^t \mathbf{S}_n(t-\tau)\mathbf{e}_n(\tau)\,\mathrm{d}\tau - \Delta_n\mathbf{A}\mathbf{S}(t)\mathbf{f} + \frac{1}{t}\int_0^t \mathbf{S}_n(t-\tau)\Delta_n\mathbf{A}\mathbf{S}(\tau)\mathbf{f}\,\mathrm{d}\tau \\ &\quad - \frac{1}{t}\int_0^t \mathbf{S}_n(t-\tau)\Delta_n\tau\mathbf{A}^2\mathbf{S}(\tau)\mathbf{f}\,\mathrm{d}\tau + \frac{1}{t}\int_0^t \mathbf{S}_n(t-\tau)\tau\,\Box_n\mathbf{S}(\tau)\mathbf{f}\,\mathrm{d}\tau \\ &=: E_1 + E_2 + E_3 + E_4 + E_5,\end{aligned}$$

where $E_j \coloneqq E_j(t,n)$ for $j \in \{1,\dots,5\}$. We bound the norm of all five terms separately, starting with the second one. The resolvent difference estimate of Proposition 3.68 for $r = \frac{\kappa}{2}$, the boundedness of the semigroup on $H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;H)$ from Proposition 3.65(c), and the generator bound from Proposition 3.65(a) imply

$$\begin{aligned}\|E_2\|_{\mathbf{H}} &\le \|\Delta_n\|_{\mathcal{L}(H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;H),\mathbf{H})}\|\mathbf{A}\mathbf{S}(t)\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;H)} \\ &\le C_{\ell,\kappa}n^{-\ell}\|\mathbf{S}(t)\mathbf{A}\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;H)} \le C_{\ell,\kappa}C_{T,\ell}n^{-\ell}\|\mathbf{A}\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;H)} \\ &\le C_{\ell,\kappa}C_{T,\ell}C_\ell n^{-\ell}\|\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;D)}, \end{aligned} \tag{3.4.16}$$

where we have written $C_{\ell,\kappa}$ for $C_{\ell,r}$ with $r = \frac{\kappa}{2}$. Using the contractivity of $(\mathbf{S}_n(t))_{t\ge 0}$ from Lemma 3.46 allows us to proceed as for E_2 to obtain

$$\begin{aligned}\|E_3\|_{\mathbf{H}} &\le \frac{1}{t}\int_0^t \|\Delta_n\|_{\mathcal{L}(H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;H),\mathbf{H})}\|\mathbf{A}\mathbf{S}(\tau)\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;H)}\,\mathrm{d}\tau \\ &\le C_{\ell,\kappa}C_{T,\ell}C_\ell n^{-\ell}\|\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;D)}. \end{aligned} \tag{3.4.17}$$

To estimate the fourth error term, combining the arguments issued for E_3 with the analyticity estimate from Proposition 3.65(d) yields

$$\begin{aligned}\|E_4\|_{\mathbf{H}} &\le C_{\ell,\kappa}n^{-\ell}\cdot\frac{1}{t}\int_0^t \|\tau\mathbf{A}\mathbf{S}(\tau)\mathbf{A}\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;H)}\,\mathrm{d}\tau \\ &\le C_{\ell,\kappa}C_{T,\ell}n^{-\ell}\|\mathbf{A}\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;H)} \le C_{\ell,\kappa}C_{T,\ell}C_\ell n^{-\ell}\|\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;D)}. \end{aligned} \tag{3.4.18}$$

Let $0 \le \tau \le T$. Applying, in this order, the triangle inequality, Lemma 3.69, Proposition 3.68, and Proposition 3.65(b), (c), and (a), we obtain the estimate

$$\begin{aligned}\|\Box_n\mathbf{S}(\tau)\mathbf{f}\|_{\mathbf{H}} &\le \frac{\kappa}{2}\Big\|[\mathbf{R}_n\mathbf{A}\mathbf{R}_n - \mathbf{A}]R\Big(\frac{\kappa}{2},-\mathbf{A}\Big)\mathbf{S}(\tau)\mathbf{f}\Big\|_{\mathbf{H}} + \frac{\kappa}{2}\Big\|\Delta_n\mathbf{S}(\tau)\mathbf{A}\mathbf{f}\Big\|_{\mathbf{H}} \\ &\le \frac{\kappa}{2}C_{\ell,\kappa}C_{T,\ell}C_\ell n^{-\ell}\big(\|\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;H)} + \|\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;D)}\big) \\ &\le \kappa C_{\ell,\kappa}C_{T,\ell}C_\ell n^{-\ell}\|\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;D)}. \end{aligned} \tag{3.4.19}$$

Consequently, we can bound the fifth term by

$$\|E_5\|_{\mathbf{H}} \le \frac{1}{t}\int_0^t \|\tau\,\Box_n\mathbf{S}(\tau)\mathbf{f}\|_{\mathbf{H}}\,\mathrm{d}\tau \le \frac{T}{2}\kappa C_{\ell,\kappa}C_{T,\ell}C_\ell n^{-\ell}\|\mathbf{f}\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;D)}. \tag{3.4.20}$$

To estimate the remaining term E_1, we first note that $\mathbf{e}_n$ satisfies the ODE $\mathbf{e}_n(0) = 0$,

$$\mathbf{e}_n'(t) = -\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{e}_n(t) - \mathbf{R}_n\mathbf{A}\mathbf{R}_n\Delta_n\mathbf{A}\mathbf{S}(t)\mathbf{f} + \Box_n\mathbf{S}(t)\mathbf{f} \quad (0 < t \leq T), \tag{3.4.21}$$

thus giving rise to the error representation for $0 < \tau \leq T$

$$\mathbf{e}_n(\tau) = -(\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^{-1}\mathbf{e}_n'(\tau) - \Delta_n\mathbf{A}\mathbf{S}(\tau)\mathbf{f} + (\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^{-1}\Box_n\mathbf{S}(\tau)\mathbf{f}. \tag{3.4.22}$$

Furthermore, the $\mathbb{P}_Z$-almost sure symmetry of $(a_z)_{z\in\mathbb{R}^N}$ implies that $-(\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^{-1}$ is self-adjoint. Thus, for $0 < \tau \leq T$,

$$\operatorname{Re}\,\big(\mathbf{e}_n(\tau)\,\big|\,-(\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^{-1}\mathbf{e}_n'(\tau)\big)_{\mathbf{H}} = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\tau}\big(\mathbf{e}_n(\tau)\,\big|\,-(\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^{-1}\mathbf{e}_n(\tau)\big)_{\mathbf{H}}. \tag{3.4.23}$$

As a consequence of the coercivity of $\mathbf{a}_n$, $-\mathbf{R}_n\mathbf{A}\mathbf{R}_n$ is dissipative and thus also its inverse. Integrating the scalar product of $\mathbf{e}_n(\tau)$ with both sides of (3.4.22), using (3.4.23), the dissipativity of $-(\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^{-1}$, the Cauchy–Schwarz inequality, and Hölder's inequality, we deduce for $t \in (0,T]$

$$\begin{aligned}\|\mathbf{e}_n\|^2_{L^2(0,t;\mathbf{H})} &= \frac{1}{2}\langle\mathbf{e}_n(t), -(\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^{-1}\mathbf{e}_n(t)\rangle_{\mathbf{H}} + \int_0^t \langle\mathbf{e}_n(\tau), -\Delta_n\mathbf{A}\mathbf{S}(\tau)\mathbf{f}\rangle_{\mathbf{H}}\,\mathrm{d}\tau \\ &\quad + \int_0^t \langle\mathbf{e}_n(\tau), (\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^{-1}\Box_n\mathbf{S}(\tau)\mathbf{f}\rangle_{\mathbf{H}}\,\mathrm{d}\tau \\ &\leq \|\mathbf{e}_n\|_{L^2(0,t;\mathbf{H})}\big(\|\Delta_n\mathbf{A}\mathbf{S}(\cdot)\mathbf{f}\|_{L^2(0,t;\mathbf{H})} + \|(\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^{-1}\Box_n\mathbf{S}(\cdot)\mathbf{f}\|_{L^2(0,t;\mathbf{H})}\big).\end{aligned}$$

Dividing by the norm of $\mathbf{e}_n$ and using the estimates (3.4.16) and (3.4.19) yields

$$\|\mathbf{e}_n\|_{L^2(0,t;\mathbf{H})} \leq 2\sqrt{t}C_{\ell,\kappa}C_{T,\ell}C_\ell n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)},$$

where we have used that $\|(\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^{-1}\|_{\mathcal{L}(\mathbf{H})} \leq \frac{1}{\kappa}$, which is an immediate consequence of the coercivity of $\mathbf{a}_n$. Finally, with the Cauchy–Schwarz inequality and uniform contractivity of $\mathbf{S}_n(t)$, this implies

$$\begin{aligned}\|E_1\|_{\mathbf{H}} &= \left\|\frac{1}{t}\int_0^t \mathbf{S}_n(t-\tau)\mathbf{e}_n(\tau)\,\mathrm{d}\tau\right\|_{\mathbf{H}} \leq \frac{1}{t}\|\mathbf{S}_n\|_{L^2(0,t;\mathcal{L}(\mathbf{H}))}\|\mathbf{e}_n\|_{L^2(0,t;\mathbf{H})} \\ &\leq 2C_{\ell,\kappa}C_{T,\ell}C_\ell n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)}.\end{aligned} \tag{3.4.24}$$

Combining (3.4.24) with (3.4.16)-(3.4.20) and omitting the dependence on κ, we deduce the estimate

$$\|\mathbf{e}_n(t)\|_{\mathbf{H}} \leq C_{T,\ell}n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)}.$$

Finally, the PCE estimate of Corollary 3.36 and Proposition 3.65(c) give the desired estimate

$$\begin{aligned}\|\mathbf{S}_n(t)\mathbf{R}_n\mathbf{f} - \mathbf{S}(t)\mathbf{f}\|_{\mathbf{H}} &\leq \|\mathbf{e}_n(t)\|_{\mathbf{H}} + \|\mathbf{R}_n - I\|_{\mathcal{L}(H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H),\mathbf{H})}\|\mathbf{S}(t)\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)} \\ &\leq C_{T,\ell}n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)} + C_{T,\ell}n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)},\end{aligned}$$

from which the second inequality is obtained for $\mathbf{f} = \mathbf{u}_0$. □

Remark 3.71. In (3.4.20), the estimate of E_5 can be improved to an estimate in terms of the weaker norm $\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H)}$ by using Proposition 3.65(d) to estimate τ times the second term in (3.4.19). However, this improvement is not possible for the remaining error terms E_1 to E_4 and thus not included for the sake of readability.

Theorem 3.72 (non-symmetric case)**.** *Let $\ell \in \mathbb{N}$. Suppose that Assumption 3.63 holds for 2ℓ. Then, for any $T > 0$ and $\varepsilon \in (0,1)$, there exists a constant $C_{T,\ell,\varepsilon} > 0$ such that*

$$\|\mathbf{S}_n(t)\mathbf{R}_n\mathbf{f} - \mathbf{S}(t)\mathbf{f}\|_{\mathbf{H}} \le C_{T,\ell,\varepsilon} n^{-\ell} \|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\varepsilon})}$$

for all $\mathbf{f} \in H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\varepsilon})$, $n \in \mathbb{N}$, and $t \in [0,T]$. In particular, for the mild solutions $\mathbf{u}_n(t)$ and $\mathbf{u}(t)$ of (3.4.2) *and* (3.4.1) *with $\mathbf{u}_0 \in H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\varepsilon})$, respectively, we have*

$$\|\mathbf{u}_n(t) - \mathbf{u}(t)\|_{\mathbf{H}} \le C_{T,\ell,\varepsilon} n^{-\ell} \|\mathbf{u}_0\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\varepsilon})} \quad (t \in [0,T]).$$

Proof. Adopt the notation of the proof of Theorem 3.70. Then $\mathbf{u}_n(t) \coloneqq (\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^{-1}\mathbf{e}_n(t)$ solves

$$\mathbf{u}_n'(t) = -\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{u}_n(t) - \Delta_n\mathbf{A}\mathbf{S}(t)\mathbf{f} + (\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^{-1}\Box_n\mathbf{S}(t)\mathbf{f}.$$

We multiply the corresponding mild solution formula by $\mathbf{R}_n\mathbf{A}\mathbf{R}_n$ and integrate by parts in the first integral, resulting in

$$\begin{aligned}\mathbf{e}_n(t) = -\Delta_n\mathbf{A}\mathbf{S}(t)\mathbf{f} + \mathbf{S}_n(t)\Delta_n\mathbf{A}\mathbf{f} - \int_0^t \mathbf{S}_n(t-\tau)\Delta_n\mathbf{A}^{1-\theta}\mathbf{S}(\tau)\mathbf{A}^{1+\theta}\mathbf{f}\,\mathrm{d}\tau \\ + \int_0^t \mathbf{S}_n(t-\tau)\Box_n\mathbf{S}(\tau)\mathbf{f}\,\mathrm{d}\tau \eqqcolon E_1 + E_2 + E_3 + E_4,\end{aligned}$$

where $E_j \coloneqq E_j(t,n)$ for $j \in \{1,\dots,4\}$. Observe that symmetry was only used in the estimate of the first term in the proof of Theorem 3.70. Hence, analogous arguments yield

$$E_1 + E_2 + E_4 \le C_{\ell,\kappa}C_\ell(C_{T,\ell} + 1 + \kappa T C_{T,\ell})n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)}. \tag{3.4.25}$$

Let $\varepsilon \in (0,1)$, $0 < \nu < \theta < \varepsilon$. For the remaining term, using contractivity of $\mathbf{S}_n$ and Proposition 3.68 as before, followed by Proposition 3.67(b), (a), and (c) results in

$$\begin{aligned}\|E_3(t)\|_{\mathbf{H}} &\le \|\Delta_n\|_{\mathcal{L}(H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H),\mathbf{H})} \sup_{0\le\tau\le t}\|\tau^{1-\nu}\mathbf{A}^{1-\theta}\mathbf{S}(\tau)\|_{\mathcal{L}(H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H))} \\ &\quad \cdot \|\mathbf{A}^\theta\|_{\mathcal{L}(H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^\varepsilon),H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;H))}\|\mathbf{A}\|_{\mathcal{L}(H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\varepsilon}),H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^\varepsilon))} \\ &\quad \cdot \|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\varepsilon})}\int_0^t \tau^{-(1-\nu)}\,\mathrm{d}\tau \\ &\le \frac{T^\nu}{\nu}C_{\ell,\kappa}C_{T,\ell,\nu,\theta}C_{\ell,\theta,\varepsilon}C_{\ell,\varepsilon}n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\varepsilon})}\end{aligned}$$

Altogether, omitting the dependence on κ and setting $\nu = \varepsilon/4$ and $\theta = \varepsilon/2$, we deduce

$$\|\mathbf{e}_n(t)\|_{\mathbf{H}} \le C_{T,\ell,\varepsilon}n^{-\ell}\|\mathbf{f}\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\varepsilon})},$$

from which the desired estimates are deduced as in the proof of Theorem 3.70. □

3.4.4 Joint convergence rate for the full discretisation error

The last step required for an estimate of the full discretisation error is to relate the graph norm of $\mathfrak{u}_{0n}$ to a suitable Sobolev norm of $\mathbf{u}_0$ in order to make use of the semi-discretisation error estimate of Subsection 3.4.2.

Lemma 3.73. *Suppose that Assumption 3.59 holds for some $\overline{\alpha} \geq 1$ and $\ell = 0$. Let $0 < \beta \leq \overline{\alpha}$, $n \in \mathbb{N}_0$, $\mathbf{w} \in \mathbf{H}$, and $\mathfrak{w}_n := (\widehat{\mathbf{w}}_\gamma)_{|\gamma|\leq n}$. Set $D^\beta := \mathrm{D}(A_z^\beta)$. We distinguish two cases.*

(a) Then for $\mathbf{w} \in L^2(\mathbb{R}^N, \mathbb{P}_Z; D^\beta)$, we have $\mathfrak{w}_n \in \mathrm{D}(\mathfrak{A}_n^\beta)$ and there exists $C_\beta \geq 0$ independent of $\mathbf{w}$ and n such that

$$\|\mathfrak{w}_n\|_{\mathfrak{A}_n^\beta} \leq C_\beta \|\mathbf{w}\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D^\beta)}.$$

(b) For general a_z satisfying Assumption 3.59, we have $\mathfrak{w}_n \in \mathrm{D}(\mathfrak{A}_n^\beta)$ provided that $\mathbf{w} \in L^2(\mathbb{R}^N, \mathbb{P}_Z; D^{\beta+\varepsilon})$ for some $0 < \varepsilon \leq \overline{\alpha} - \beta$ and $\beta < \overline{\alpha}$. Moreover, there exists $C_{\beta,\varepsilon} \geq 0$ independent of $\mathbf{w}$ and n such that

$$\|\mathfrak{w}_n\|_{\mathfrak{A}_n^\beta} \leq C_{\beta,\varepsilon} \|\mathbf{w}\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D^{\beta+\varepsilon})}.$$

Proof. (i) We first let $\beta = 1$. From $\mathrm{D}(\mathbf{A}) = L^2(\mathbb{R}^N, \mathbb{P}_Z; D)$, we conclude $\mathbf{R}_n\mathbf{w} \in \mathrm{D}(\mathbf{A}) \subseteq \mathbf{V}$ and $\mathfrak{w}_n \in \mathfrak{V}_n$. Let $\mathfrak{v}_n = (v_\gamma)_{|\gamma|\leq n} \in \mathfrak{V}_n$ and define $\mathbf{v}_n := \sum_{|\gamma|\leq n} v_\gamma \Phi_\gamma \in \mathcal{P}_n^N \otimes V$. Then, from (3.4.4), $\mathbf{a}_n \sim \mathbf{R}_n\mathbf{A}\mathbf{R}_n$, and $\mathbf{R}_n^2 = \mathbf{R}_n$ we conclude

$$\begin{aligned} \mathfrak{a}_n(\mathfrak{w}_n, \mathfrak{v}_n) &= \mathbf{a}_n(\mathbf{R}_n\mathbf{w}, \mathbf{v}_n) = \big(\mathbf{R}_n\mathbf{A}\mathbf{R}_n^2\mathbf{w} \,\big|\, \mathbf{v}_n\big)_{\mathbf{H}} \\ &= (\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{w} \,|\, \mathbf{v}_n)_{\mathcal{P}_n^N\otimes H} = \Big(((\widehat{\mathbf{R}_n\mathbf{A}\mathbf{R}_n}\mathbf{w})_\gamma)_{|\gamma|\leq n}\Big|\mathfrak{v}_n\Big)_{\mathfrak{H}_n} \end{aligned}$$

by Plancherel's theorem. Hence, $\mathfrak{w}_n \in \mathrm{D}(\mathfrak{A}_n)$ and $\mathfrak{A}_n\mathfrak{w}_n = ((\widehat{\mathbf{R}_n\mathbf{A}\mathbf{R}_n}\mathbf{w})_\gamma)_{|\gamma|\leq n}$. By Parseval's identity twice and Proposition 3.65(a) with $q = 0$, we conclude

$$\|\mathfrak{A}_n\mathfrak{w}_n\|_{\mathfrak{H}_n} = \|\mathbf{R}_n\mathbf{A}\mathbf{R}_n\mathbf{w}\|_{\mathbf{H}} \leq \|\mathbf{A}\mathbf{R}_n\mathbf{w}\|_{\mathbf{H}} \leq C\|\mathbf{R}_n\mathbf{w}\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D)} \leq C\|\mathbf{w}\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D)}.$$

(ii) Let $\mathbb{N} \ni m \leq \overline{\alpha}$. Repeating the argument of step (i), we obtain $\mathfrak{w}_n \in \mathrm{D}(\mathfrak{A}_n^m)$ for all $\mathbf{w} \in \mathrm{D}(\mathbf{A}^m)$ and for some $C_m \geq 0$,

$$\|\mathfrak{A}_n^m\mathfrak{w}_n\|_{\mathfrak{H}_n} = \|(\mathbf{R}_n\mathbf{A}\mathbf{R}_n)^m\mathbf{w}\|_{\mathbf{H}} \leq C_m\|\mathbf{w}\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D^m)}. \tag{3.4.26}$$

(iii) We prove part (a). Let $0 < \beta \leq \overline{\alpha}$ and choose $\mathbb{N} \ni m \geq \beta$. Due to symmetry of the form, $\mathfrak{A}_n$ is self-adjoint and thus $\mathfrak{A}_n^m$ is m-accretive. A generalisation of the Heinz inequality [83] then yields the statement for $\mathfrak{A}_n^\beta$ via interpolation. Note that $\mathrm{D}(\mathbf{A}^\beta) = L^2(\mathbb{R}^N, \mathbb{P}_Z; D^\beta)$.
(iv) Now, consider the setting of part (b). Let $0 < \beta < \overline{\alpha}$ and choose $\mathbb{N} \ni m \geq \beta$. To pass from (3.4.26) to exponent β, additional spatial regularity is required since $\mathfrak{A}_n^m$ is no longer m-accretive in general. A version of the Heinz inequality for non-selfadjoint operators [88, Prop. 9.3] implies that for all $0 < \varepsilon \leq \overline{\alpha} - \beta$

$$\big\|\mathfrak{A}_n^\beta\mathfrak{w}_n\big\|_{\mathfrak{H}_n} \leq C_{\beta,\varepsilon}\|\mathbf{R}_n\mathbf{w}\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D^{\beta+\varepsilon})} \leq C_{\beta,\varepsilon}\|\mathbf{w}\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D^{\beta+\varepsilon})}$$

for some $C_{\beta,\varepsilon} \geq 0$ for all $\mathbf{w} \in \mathrm{D}(\mathbf{A}^{\beta+\varepsilon}) = L^2(\mathbb{R}^N, \mathbb{P}_Z; D^{\beta+\varepsilon})$ and $n \in \mathbb{N}_0$. □

We can now state our main result on the joint convergence rate for the approximation of random evolution equations.

Theorem 3.74 (symmetric case)**.** *Let $\ell \in \mathbb{N}$ and $T > 0$. Suppose that Assumption 3.59 holds for some $\overline{\alpha} \geq 1$ and 2ℓ. Further, assume that $(a_z)_{z\in\mathbb{R}^N}$ is $\mathbb{P}_Z$-almost surely symmetric and that $V \hookrightarrow H$ compactly. Suppose that the space discretisation converges with order $p_{\mathrm{x}} > 0$ on $D^\alpha \coloneqq \mathrm{D}(A_z^\alpha)$ for some $0 < \alpha \leq \overline{\alpha}$ for the stationary problem. Further, suppose that the time discretisation methods $\mathfrak{F}_{n,m}$ are induced by an A-stable rational function r for all $n \in \mathbb{N}_0$, $m \in \mathbb{N}$, and let $p_{\mathrm{t}} > 0$ be the order of convergence of the time discretisations. Let $\mathbf{u}$ be the mild solution of (3.4.1) with $\mathbf{u}_0 \in H^{2\ell}_\rho(\mathbb{R}^N, \mathbb{P}_Z; D) \cap L^2(\mathbb{R}^N, \mathbb{P}_Z; D^{\max\{\alpha+1,p_{\mathrm{t}}\}})$ where $\max\{\alpha+1, p_{\mathrm{t}}\} \leq \overline{\alpha}$ and, for $n, m, k \in \mathbb{N}$, let $\mathbf{J}_{n,m}$ and $\mathbf{u}_{n,m,k}$ be as in (3.4.8) and (3.4.9) via time discretisation methods $\mathfrak{F}_{n,m}$, respectively.*
Then there exist $C_{T,\ell}, C_{T,\alpha,p_{\mathrm{t}}} \geq 0$, $\tau_0 > 0$ such that for $\max_{i=1,\dots,N_k} \tau_k^i \leq \tau_0$,

$$\|\mathbf{J}_{n,m}\mathbf{u}_{n,m,k}(t) - \mathbf{u}(t)\|_{\mathbf{H}} \leq C_{T,\ell} n^{-\ell} \|\mathbf{u}_0\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)} + C_{T,\alpha,p_{\mathrm{t}}} \left(m^{-p_{\mathrm{x}}} + \Big(\max_{i=1,\dots,N_k} \tau_k^i \Big)^{p_{\mathrm{t}}} \right) \|\mathbf{u}_0\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D^{\max\{\alpha+1,p_{\mathrm{t}}\}})}$$

for all $n, m, k \in \mathbb{N}$ and $t \in \mathcal{T}_k$.

Proof. Abbreviate $\tau_{\max,k} \coloneqq \max_{i=1,\dots,N_k} \tau_k^i$. Then Proposition 3.58 and Theorem 3.70 imply that for some C_T, we have

$$\begin{aligned}\|\mathbf{J}_{n,m}\mathbf{u}_{n,m,k}(t) - \mathbf{u}(t)\|_{\mathbf{H}} &\leq \|\mathbf{J}_{n,m}\mathbf{u}_{n,m,k}(t) - \mathbf{u}_n(t)\|_{\mathbf{H}} + \|\mathbf{u}_n(t) - \mathbf{u}(t)\|_{\mathbf{H}} \\ &\leq C_{T,\alpha,p_{\mathrm{t}}} \big(m^{-p_{\mathrm{x}}} + \tau_{\max,k}^{p_{\mathrm{t}}}\big) \|\mathfrak{u}_{0n}\|_{\mathfrak{A}_n^{\max\{\alpha+1,p_{\mathrm{t}}\}}} + C_{T,\ell} n^{-\ell} \|\mathbf{u}_0\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D)}.\end{aligned}$$

The statement then follows from Lemma 3.73 applied to $\mathbf{w} = \mathbf{u}_0$ and $\beta = \max\{\alpha+1, p_{\mathrm{t}}\}$, which gives

$$\|\mathfrak{u}_{0n}\|_{\mathfrak{A}_n^{\max\{\alpha+1,p_{\mathrm{t}}\}}} \leq C_\beta \|\mathbf{u}_0\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D^{\max\{\alpha+1,p_{\mathrm{t}}\}})}.$$ ☐

Repeating the argument in Proposition 3.58 using Theorem 3.41 instead of 3.37, Theorem 3.72 instead of 3.70 as well as Lemma 3.73(b) rather than (a), we obtain a joint convergence rate for the non-symmetric case, proceeding analogously to the proof of Theorem 3.74. As in the deterministic setting, additional spatial regularity is required.

Theorem 3.75 (non-symmetric case)**.** *Suppose that all assumptions of Theorem 3.74 hold true apart from $\mathbb{P}_Z$-almost sure symmetry of $(a_z)_{z\in\mathbb{R}^N}$ and the compactness of the embedding $V \hookrightarrow H$. Let the time discretisation methods $\mathfrak{F}_{n,m}$ be either the implicit Euler method, where $p_{\mathrm{t}} \coloneqq 1$, or the Crank–Nicolson method, where $p_{\mathrm{t}} \coloneqq 2$, for all $n, m \in \mathbb{N}$. Further, let $\varepsilon_1 \in (0,1)$ such that $\mathbf{u}_0 \in H^{2\ell}_\rho(\mathbb{R}^N, \mathbb{P}_Z; D^{1+\varepsilon_1})$. Moreover, let $\varepsilon_2 > 0$ such that $\mathbf{u}_0 \in L^2(\mathbb{R}^N, \mathbb{P}_Z; D^{1+\varepsilon_2+\max\{\alpha,p_{\mathrm{t}}\}})$, where $1 + \varepsilon_2 + \max\{\alpha, p_{\mathrm{t}}\} \leq \overline{\alpha}$.*

Then there exist $\tau_0 > 0$ and $C_{T,\ell,\varepsilon_1}, C_{T,\alpha,p_{\mathrm{t}},\varepsilon_2} \geq 0$ such that for $\max_{i=1,\dots,N_k} \tau_k^i \leq \tau_0$,

$$\|\mathbf{J}_{n,m}\mathbf{u}_{n,m,k}(t) - \mathbf{u}(t)\|_H \leq C_{T,\ell,\varepsilon_1} n^{-\ell} \|\mathbf{u}_0\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\varepsilon_1})} + C_{T,\alpha,p_{\mathrm{t}},\varepsilon_2} \left(m^{-p_{\mathrm{x}}} + \Big(\max_{i=1,\dots,N_k} \tau_k^i \Big)^{p_{\mathrm{t}}} \right) \|\mathbf{u}_0\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D^{1+\varepsilon_2+\max\{\alpha,p_{\mathrm{t}}\}})}$$

for all $n, m, k \in \mathbb{N}$ and $t \in \mathcal{T}_k$.

Remark 3.76. Let us comment on the case of deterministic initial values, i.e., $u_0 \in D^\beta$ for some $\beta > 0$. Then $\mathbf{u}_0 := \mathbf{1}_{\mathbb{R}^N} \otimes u_0 \in H^{2\ell}_\rho(\mathbb{R}^N, \mathbb{P}_Z; D^\beta)$ for all $\ell \in \mathbb{N}_0$ and

$$\|\mathbf{u}_0\|_{H^{2\ell}_\rho(\mathbb{R}^N,\mathbb{P}_Z;D^\beta)} = \|u_0\|_{D^\beta},$$

whence we naturally recover the results of Section 3.3.

3.5 Application to Random Anisotropic Diffusion

Let $G \subseteq \mathbb{R}^2$ be open, bounded, convex and polygonal, and let $H := L^2(G)$ and $V := H^1_0(G)$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $Z\colon \Omega \to \mathbb{R}^N$ a random variable with independent components and distribution $\mathbb{P}_Z$. Assume that each of the components is standard normally distributed, Beta-distributed, or Gamma-distributed; cf. Assumption 3.20. Let $\mathbb{R}^N \times G \ni (z, x) \mapsto a^z(x) \in \mathbb{K}^{2\times 2}$ such that $a^z(x)$ is Hermitian for all $x \in G$ and $\mathbb{P}_Z$-a.e. $z \in \mathbb{R}^N$, and assume there exist $\kappa, M > 0$ such that

$$\kappa \le a^z(x) \le M \quad (x \in G, \mathbb{P}_Z\text{-a.e. } z \in \mathbb{R}^N). \tag{3.5.1}$$

Moreover, let $z \mapsto \int_G a^z(x)_{jk} h(x)\,\mathrm{d}x$ be measurable for all $j, k \in \{1, 2\}$ and $h \in L^1(G)$.

For those $z \in \mathbb{R}^N$ for which the above bound holds, define $a_z\colon V \times V \to \mathbb{K}$ by

$$a_z(u, v) := \int_G a^z(x) \operatorname{grad} u(x) \cdot \overline{\operatorname{grad} v(x)}\,\mathrm{d}x \quad (u, v \in H^1_0(G)),$$

and let $a_z(u, v) = 0$ otherwise. Then $z \mapsto a_z(u)$ is measurable for all $u \in V$ and

$$\kappa\|u\|_V^2 \le a_z(u), \quad |a_z(u, v)| \le M\|u\|_V\|v\|_V \quad (u, v \in V, \mathbb{P}_Z\text{-a.e. } z \in \mathbb{R}^N).$$

That is, $(a_z)_{z\in\mathbb{R}^N}$ is $\mathbb{P}_Z$-almost surely uniformly bounded and coercive, as required in Assumption 3.42.

For $z \in \mathbb{R}^N$ let A_z be the self-adjoint operator in H associated with a_z. Note that $[0, \infty) \subseteq \rho(-A_z)$ for $\mathbb{P}_Z$-a.e. $z \in \mathbb{R}^N$.

Remark 3.77. Provided that $a^z \in W^1_\infty(G)$ for $\mathbb{P}_Z$-a.e. $z \in \mathbb{R}^N$, we $\mathbb{P}_Z$-almost surely have

$$\begin{aligned} \mathrm{D}(A_z) &= H^1_0(G) \cap H^2(G), \\ A_z u &= -\operatorname{div} a^z(x) \operatorname{grad} u. \end{aligned}$$

Hence, we can set $D := H^1_0(G) \cap H^2(G)$ equipped with the H^2-norm.

To establish the estimate of derivatives of the form in z required for Assumption 3.59 to hold, smoothness of the coefficient matrix is required. We endow $\mathbb{K}^{2\times 2}$ with the spectral norm $\|\cdot\|_2$ and $\mathbb{K}^2$ with the Euclidean norm. Denote by $\operatorname{Div}\colon \mathbb{K}^{2\times 2} \to \mathbb{K}^2$ the row-wise divergence operator so that $\operatorname{Div}(b(\cdot)^T)$ calculates the divergence of a matrix-valued function $b \in C^1(G; \mathbb{K}^{2\times 2})$ column-wise.

Definition 3.78. For $\ell \in \mathbb{N}_0$, define $C^{\ell,1}_\rho(\mathbb{R}^N \times G; \mathbb{K}^{2\times 2})$ as the space of all $f\colon \mathbb{R}^N \times G \to \mathbb{K}^{2\times 2}$ such that $f(\cdot, x) \in C^\ell(\mathbb{R}^N; \mathbb{K}^{2\times 2})$ for all $x \in G$, $f(z, \cdot) \in C^1(G; \mathbb{K}^{2\times 2})$ for $\mathbb{P}_Z$-a.e. $z \in \mathbb{R}^N$ and there is $C_\ell \ge 0$ such that for all $\alpha \in \mathbb{N}_0^N$ with $|\alpha| \le \ell$, we have

$$\mathbb{P}_Z\text{-}\operatorname*{ess\,sup}_{z\in\mathbb{R}^N} \big\|\rho(z)^{\alpha/2} \operatorname{Div}\big((\partial_z^\alpha f(z,\cdot))^T\big)\big\|_{L^\infty(G;\mathbb{K}^2)} \le C_\ell \quad \text{and}$$

$$\mathbb{P}_Z\text{- }\operatorname*{ess\,sup}_{z\in\mathbb{R}^N} \|\rho(z)^{\alpha/2}\partial_z^\alpha f(z,\cdot)\|_{L^\infty(G;\mathbb{K}^{2\times 2})} \le C_\ell.$$

Let $C_\rho^{\ell,2}(\mathbb{R}^N\times G;\mathbb{K}^{2\times 2})$ denote the space of all $f \in C_\rho^{\ell,1}(\mathbb{R}^N\times G;\mathbb{K}^{2\times 2})$ such that $f(z,\cdot)\in C^2(G;\mathbb{K}^{2\times 2})$ for $\mathbb{P}_Z$-a.e. $z\in\mathbb{R}^N$ and $[(z,x)\mapsto \partial_x^{e_j} f(z,x)] \in C_\rho^{\ell,1}(\mathbb{R}^N\times G;\mathbb{K}^{2\times 2})$ for $j=1,2$. Analogously, $C_\rho^{\ell,3}(\mathbb{R}^N\times G;\mathbb{K}^{2\times 2})$ denotes the space of all $f\in C_\rho^{\ell,2}(\mathbb{R}^N\times G;\mathbb{K}^{2\times 2})$ such that $f(z,\cdot)\in C^3(G;\mathbb{K}^{2\times 2})$ for $\mathbb{P}_Z$-a.e. $z\in\mathbb{R}^N$ and $[(z,x)\mapsto \partial_x^{e_j} f(z,x)] \in C_\rho^{\ell,2}(\mathbb{R}^N\times G;\mathbb{K}^{2\times 2})$ for $j=1,2$.

Assume that $[(z,x)\mapsto a^z(x)]\in C_\rho^{\ell,1}(\mathbb{R}^N\times G;\mathbb{K}^{2\times 2})$. Then the graph norms of A_z are equivalent for $\mathbb{P}_Z$-a.e. $z\in\mathbb{R}^N$ since there are $C,c\ge 0$ such that

$$\|u\|_{H^2} \le c\|A_z u\|_{L^2} \le C\|u\|_{H^2} \tag{3.5.2}$$

for all $u\in \mathrm{D}(A_z)$ and $\mathbb{P}_Z$-a.e. $z\in\mathbb{R}^N$. Indeed, the first inequality follows from [57, Thm. 3.2.1.2 and 3.1.3.1] because G is convex. Since $[(z,x)\mapsto a^z(x)]$ is continuously differentiable in x, for $u\in\mathrm{D}(A_z)$, we can rewrite

$$A_z u = -\operatorname{div}(a^z(\cdot)\operatorname{grad} u) = -\operatorname{Div}\left(a^z(\cdot)^T\right)\cdot \operatorname{grad} u - a^z(\cdot) : \operatorname{Hess} u,$$

where $\operatorname{Hess} u$ is the Hessian matrix of u and $B : C = \sum_{i,j=1}^2 b_{i,j}c_{i,j}$ is the sum of the entry-wise product of two matrices $B,C\in\mathbb{K}^{2\times 2}$. By assumption for $\alpha=0$ and due to the definition of $C_\rho^{\ell,1}(\mathbb{R}^N\times G;\mathbb{K}^{2\times 2})$, we can thus estimate

$$\begin{aligned}
\|A_z u\|_{L^2} &\le \left\|\operatorname{Div}\left(a^z(\cdot)^T\right)\cdot\operatorname{grad} u\right\|_{L^2} + \|a^z(\cdot):\operatorname{Hess} u\|_{L^2}\\
&\le \sup_{x\in G}\left\|\operatorname{Div}\left(a^z(x)^T\right)\right\|_\infty \|\operatorname{grad} u\|_{L^2} + \left(\sup_{x\in G}\max_{i,j\in\{1,2\}} |(a^z(x))_{i,j}|\right)\left\|\sum_{i,j=1}^2 \partial_{ij}u\right\|_{L^2}\\
&\le C_0\|\operatorname{grad} u\|_{L^2} + C_0\sum_{i,j=1}^2 \|\partial_{ij}u\|_{L^2} \le \sqrt{6}C_0\|u\|_{H^2}
\end{aligned} \tag{3.5.3}$$

for all $u\in\mathrm{D}(A_z)$ and $\mathbb{P}_Z$-a.e. $z\in\mathbb{R}^N$, where $\|\cdot\|_\infty$ denotes the maximum norm in $\mathbb{K}^2$. This yields the second inequality. Consequently, Assumption 3.52 is satisfied for $\overline{\alpha}=1$. However, we would like to have the assumption satisfied for some $\overline{\alpha}>1$. This can be shown under additional assumptions on the diffusion coefficients. In case we want $1<\overline{\alpha}<\frac{5}{4}$, reasoning as in [61] we observe that for $[(z,x)\mapsto a^z(x)]\in C_\rho^{\ell,2}(\mathbb{R}^N\times G;\mathbb{K}^{2\times 2})$ we have $D^{\overline{\alpha}} = H_0^1(G)\cap H^{2\overline{\alpha}}(G)$ with corresponding graph norms of $A_z^{\overline{\alpha}}$ being equivalent, i.e., Assumption 3.52 is satisfied. In case we want $\overline{\alpha}\ge\frac{5}{4}$, further regularity on the coefficients and, due to the Dirichlet boundary conditions, further assumptions on the traces of the coefficients may be needed. We will consider the case $\overline{\alpha}=2$. In order to obtain that the domains $\mathrm{D}(A_z^2)$ are $\mathbb{P}_Z$-almost surely constant, we may assume that $[(z,x)\mapsto a^z(x)]\in C_\rho^{\ell,3}(\mathbb{R}^N\times G;\mathbb{K}^{2\times 2})$ as well as $\operatorname{tr} a^z$ and $\operatorname{tr}\operatorname{Div}(a^z(\cdot)^T)$ are constant for $\mathbb{P}_Z$-almost every $z\in\mathbb{R}^N$, where tr denotes the trace operator for G. Then, a similar reasoning as above yields that Assumption 3.52 is satisfied for $\overline{\alpha}=2$, and the graph norms of A_z^2 can be compared to the $H^4(G)$-norm.

To verify the form estimate from Assumption 3.59, it suffices to show that there is $C_\ell\ge 0$ such that

$$\left\|\rho(z)^{\alpha/2}\partial_z^\alpha A_z u\right\|_{L^2} \le C_\ell\|u\|_{H^2} \tag{3.5.4}$$

for all $\alpha \in \mathbb{N}_0^N$ with $|\alpha| \le \ell$ and $\mathbb{P}_Z$-a.e. $z \in \mathbb{R}^N$. Indeed, for $u \in \mathrm{D}(\partial_z^\alpha A_z) \cap H^2(G)$ and $v \in H_0^1(G)$, (3.5.4) implies

$$\begin{aligned}\left|\rho(z)^{\alpha/2}\partial_z^\alpha a_z(u,v)\right| &= \left|\left(\rho(z)^{\alpha/2}\partial_z^\alpha A_z \middle| v\right)\right| \le \left\|\rho(z)^{\alpha/2}\partial_z^\alpha A_z\right\|_{L^2}\|v\|_{L^2}\\ &\le C_\ell\|u\|_{H^2}\|v\|_{L^2} = C_\ell\|u\|_D\|v\|_H.\end{aligned}$$

Now, we show (3.5.4). By assumption, $[(z,x) \mapsto a^z(x)]$ is continuously differentiable in x (and ℓ times continuously differentiable in z) so that

$$\partial_z^\alpha A_z u = -\operatorname{div}(\partial_z^\alpha a^z(\cdot)\operatorname{grad} u) = -\operatorname{Div}\left(\partial_z^\alpha a^z(\cdot)^T\right)\cdot\operatorname{grad} u - \left(\partial_z^\alpha a^z(\cdot)\right) : \operatorname{Hess} u.$$

Analogously to the proof of the second inequality in (3.5.2), we infer that for $[(z,x) \mapsto a^z(x)] \in C_\rho^{\ell,1}(\mathbb{R}^N \times G; \mathbb{K}^{2\times 2})$,

$$\begin{aligned}\left\|\rho(z)^{\alpha/2}\partial_z^\alpha A_z u\right\|_{L^2} &\le \left\|\rho(z)^{\alpha/2}\operatorname{Div}\left((\partial_z^\alpha a^z(\cdot))^T\right)\cdot\operatorname{grad} u\right\|_{L^2} + \left\|\rho(z)^{\alpha/2}\left(\partial_z^\alpha a^z(\cdot)\right) : \operatorname{Hess} u\right\|_{L^2}\\ &\le \sqrt{6}C_\ell\|u\|_{H^2}.\end{aligned}$$

In conclusion, Assumption 3.59 is satisfied for $\overline{\alpha}$ and $\ell \in \mathbb{N}$ in any of the following cases:

(i) $[(z,x) \mapsto a^z(x)] \in C_\rho^{\ell,1}(\mathbb{R}^N \times G; \mathbb{K}^{2\times 2})$ and $\overline{\alpha} = 1$,

(ii) $[(z,x) \mapsto a^z(x)] \in C_\rho^{\ell,2}(\mathbb{R}^N \times G; \mathbb{K}^{2\times 2})$ and $1 < \overline{\alpha} < \frac{5}{4}$,

(iii) $[(z,x) \mapsto a^z(x)] \in C_\rho^{\ell,3}(\mathbb{R}^N \times G; \mathbb{K}^{2\times 2})$ and $\frac{5}{4} \le \overline{\alpha} \le 2$ as well as $\operatorname{tr} a^z$ and $\operatorname{tr}\operatorname{Div}(a^z(\cdot)^T)$ are $\mathbb{P}_Z$-almost surely constant.

As space discretisation, we employ the quadratic finite element method (FEM). More precisely, we consider quasi-uniform triangulations G_h of G, $h > 0$, consisting of triangles with a circumference no larger than $h = \frac{2}{m}$. Let

$$V_m := \{u \in H_0^1(G) : \ u|_K \in \mathcal{P}_2^2 \ \ \forall K \in G_h\}$$

be the corresponding quadratic triangular finite element (FE) space. Due to the finite dimension of V_m, the spaces H_m and V_m coincide. For $m \in \mathbb{N}$, P_m is the $L^2(G)$-orthogonal projection from $L^2(G)$ onto V_m. These are not to be confused with the V-orthogonal projections from $V = H_0^1(G)$ onto V_m yielding the FE approximation u_h of u from V_m. Then the error estimate [19, Satz 6.4]

$$\|u - u_h\|_{H^r(G)} \le Ch^{R-r}|u|_{H^R(G)} \quad (r \in \{0,\dots,R\}) \tag{3.5.5}$$

holds true for some $C \ge 0$ for $R = 2, 3$ and all $u \in H^R(G)$, where $|\cdot|_{H^R(G)}$ denotes the standard $H^R(G)$-seminorm. Setting $r = 1$ and varying R yields

$$\|u - u_h\|_{H^1(G)} \le Ch|u|_{H^2(G)} \quad (u \in H^2(G)) \tag{3.5.6}$$

for $R = 2$ and, for $R = 3$,

$$\|u - u_h\|_{H^1(G)} \le Ch^2|u|_{H^3(G)} \quad (u \in H^3(G)). \tag{3.5.7}$$

Since $h = \frac{2}{m}$, this implies linear and quadratic decay in m, respectively.

From these FEM estimates for the stationary problem, we can deduce a spatial convergence rate for the time-dependent problem via Corollary 3.9(a). To this end, we determine the decay rates $p_1(\alpha)$, $0 < \alpha \le \overline{\alpha} - 1$, and p_2 of $(\gamma_m(D^\alpha))_{m\in\mathbb{N}}$ and $(\gamma_m^*(H))_{m\in\mathbb{N}}$, respectively. We let $\overline{\alpha} > 1$ and start with the latter.

Firstly, a_z is symmetric $\mathbb{P}_Z$-almost surely since we assumed $a^z(x)$ to be Hermitian for all $x \in G$ and $\mathbb{P}_Z$-a.e. $z \in \mathbb{R}^N$. Secondly, $\mathcal{A}_z^{-1}H = \mathrm{D}(A_z)$ $\mathbb{P}_Z$-a.s. by definition of A_z as the restriction of $\mathcal{A}_z$ to the preimage of H under $\mathcal{A}_z$. From the FE estimate (3.5.6) and the almost sure equivalence (3.5.2) of the H^2-norm and $\|A_z \cdot\|_{L^2}$, we deduce that $\mathbb{P}_Z$-almost surely

$$\inf_{v\in V_m} \|u-v\|_{H^1} \le \|u-u_h\|_{H^1} \le \frac{C}{m}|u|_{H^2} \le \frac{cC}{m}\|A_z u\|_{L^2}$$

provided that $[(z,x)\mapsto a^z(x)] \in C_\rho^{\ell,1}(\mathbb{R}^N \times G;\mathbb{K}^{2\times 2})$. Therefore, the decay rate $p_2 = 1$ is obtained.

Next, we calculate $p_1(\alpha)$ depending on the value of $0 < \alpha \le \overline{\alpha} - 1$. Since this implies $\overline{\alpha} > 1$, we are only interested in the cases (ii) and (iii) listed above. Calculating $p_1(\alpha)$ requires establishing an estimate of the form

$$\inf_{u\in V_m} \|u-v\|_{H^1} \le \frac{C_{\gamma,\mathrm{D}(A^\alpha)}}{m^{p_1(\alpha)}}\|A_z u\|_{\mathrm{D}(A_z^\alpha)} \quad (u \in \mathrm{D}(A_z) \text{ with } A_z u \in \mathrm{D}(A_z^\alpha) = D^\alpha)$$

for $\mathbb{P}_Z$-a.e. $z \in \mathbb{R}^N$. Repeating the estimate performed for p_2, we obtain $p_1(\alpha) = 1$ for all $0 < \alpha \le \overline{\alpha} - 1$ since $D^\alpha \hookrightarrow L^2(G)$. However, this estimate does not take the increased smoothness of u for larger α into account. Provided that $\alpha \ge \frac{1}{2}$, we can make use of the smoothness of $u \in \mathrm{D}(A_z)$ such that $A_z u \in \mathrm{D}(A_z^\alpha)$ almost surely to obtain a higher decay rate.

Suppose that $\overline{\alpha} \ge \frac{3}{2}$ and $\alpha \ge \frac{1}{2}$. Then $u \in \mathrm{D}(A_z) = H^2(G) \cap H_0^1(G)$ and $A_z u \in D^{1/2}$ $\mathbb{P}_Z$-almost surely. The higher-order FEM estimate (3.5.7) and the equivalence of the H^2-norm and the graph norm; cf. (3.5.2), imply

$$\begin{aligned}\inf_{v\in V_m} \|u-v\|_{H^1}^2 &\le \|u-u_h\|_{H^1}^2 \le \Big(\frac{C}{m^2}\Big)^2 |u|_{H^3}^2 = \Big(\frac{C}{m^2}\Big)^2 \big(|\partial_1 u|_{H^2}^2 + |\partial_2 u|_{H^2}^2\big) \\ &\le c^2\Big(\frac{C}{m^2}\Big)^2 \big(\|A_z\partial_1 u\|_{L^2}^2 + \|A_z\partial_2 u\|_{L^2}^2\big) \quad (\mathbb{P}_Z\text{-a.e. } z \in \mathbb{R}^N). \qquad (3.5.8)\end{aligned}$$

In order to estimate $\|A_z\partial_j u\|_{L^2}^2$ for $j = 1,2$, we need $[(z,x)\mapsto a^z(x)] \in C_\rho^{\ell,2}(G;\mathbb{K}^{2\times 2})$. In particular, $a^z(\cdot) \in C_b^2(G;\mathbb{K}^{2\times 2})$ for $\mathbb{P}_Z$-a.e. $z \in \mathbb{R}^N$. This smoothness allows us to explicitly compute

$$\partial_j(\mathrm{Div}(b)\,\mathrm{grad}\,u) = \mathrm{Div}(b)\,\mathrm{grad}(\partial_j u) + \mathrm{Div}(\partial_j b)\,\mathrm{grad}\,u$$

for $j = 1,2$, $u \in H^3(G)$, and coefficients $b \in C^2(G;\mathbb{K}^{2\times 2})$. Hence, $\mathbb{P}_Z$-almost surely

$$\begin{aligned}A_z\partial_j u &= -\operatorname{div}(a^z(\cdot)\,\mathrm{grad}(\partial_j u)) = -\operatorname{Div}(a^z(\cdot)^T)\,\mathrm{grad}\,\partial_j u - a^z(\cdot) : \mathrm{Hess}\,\partial_j u \\ &= -\partial_j(\mathrm{Div}(a^z(\cdot)^T)\,\mathrm{grad}\,u) + \mathrm{Div}(\partial_j(a^z(\cdot)^T))\,\mathrm{grad}\,u \\ &\quad - \partial_j(a^z(\cdot) : \mathrm{Hess}\,u) + (\partial_j a^z(\cdot)) : \mathrm{Hess}\,u \\ &= \partial_j(A_z u) + \mathrm{Div}(\partial_j(a^z(\cdot)^T))\,\mathrm{grad}\,u + (\partial_j a^z(\cdot)) : \mathrm{Hess}\,u.\end{aligned}$$

By assumption, $\partial_j(a^z(\cdot)^T)$ and $\partial_j a^z(\cdot)$ have bounded derivatives, so that an analogous estimate to (3.5.3) results in

$$\|A_z\partial_j u\|_{L^2} \le \|\partial_j(A_z u)\|_{L^2} + C\|u\|_{H^2} \quad (u \in H^3(G),\ j = 1,2).$$

Thus, with C denoting a generic constant taking different values,

$$\sum_{j=1}^{2} \|A_z \partial_j u\|_{L^2}^2 \le C \sum_{j=1}^{2} \|\partial_j (A_z u)\|_{L^2}^2 + C\|u\|_{H^2}^2 = C\| \operatorname{grad}(A_z u)\|_{L^2}^2 + C\|u\|_{H^2}^2$$
$$\le C\|A_z^{1/2} A_z u\|_{L^2}^2 + C\|A_z u\|_{L^2}^2 \le C\|A_z u\|_{\mathrm{D}(A_z^{1/2})}^2,$$

where in the penultimate inequality we have used the Kato square root property for self-adjoint elliptic operators [7, Thm. 1] for the first term and (3.5.2) for the second one. Combining this estimate with (3.5.8) yields $p_1(\frac{1}{2}) = 2$, and thus $p_1(\alpha) = 2$ for all $\alpha \ge \frac{1}{2}$.

As a consequence of Corollary 3.9(a), we obtain a spatial convergence rate depending on the smoothness of the space Y_{x} by summing $p_1(\alpha)$ and p_2.

Proposition 3.79. *(a) Suppose that $[(z, x) \mapsto a^z(x)]$ satisfies (ii). Then the space discretisation defined here above converges on $Y_{\mathrm{x}} = D^{\alpha+1}$ with order $p_{\mathrm{x}} = 2$ for $0 < \alpha < \frac{1}{4}$.*

(b) Suppose that $[(z, x) \mapsto a^z(x)]$ satisfies (iii). Then the space discretisation defined here above converges on $Y_{\mathrm{x}} = D^{\alpha+1}$ with order $p_{\mathrm{x}} = 2$ for $0 < \alpha < \frac{1}{2}$ and with order $p_{\mathrm{x}} = 3$ for $\frac{1}{2} \le \alpha \le 1$.

Remark 3.80. Higher values of $\alpha > \frac{1}{2}$ do not result in higher values of $p_1(\alpha)$ or p_{x}. This restriction is due to quadratic FE spaces being used. To further increase $p_1(\alpha)$ for large α, higher-order FE spaces, such as piecewise cubic polynomials on a quasi-uniform triangulation, have to be used.

Having verified Assumption 3.59 and calculated the spatial convergence rate for our example, we can apply Theorem 3.74 to obtain a joint convergence rate for the full discretisation of the abstract Cauchy problem associated with A_z.

Theorem 3.81. *Let $Z : \Omega \to \mathbb{R}^N$ be a random variable satisfying Assumption 3.20. Suppose that $\mathbb{R}^N \times G \ni (z, x) \mapsto a^z(x) \in \mathbb{K}^{2\times 2}$ satisfies (3.5.1) and is contained in $C_\rho^{2\ell,2}(\mathbb{R}^N \times G; \mathbb{K}^{2\times 2})$ for some $\ell \in \mathbb{N}$. Let $0 < \alpha < \frac{1}{4}$ and $\mathbf{u} = (u_z)_{z\in\mathbb{R}^N}$ be the mild solution of*

$$u_z'(t) = -\operatorname{div} a^z(x) \operatorname{grad} u_z(t) \quad (t > 0), \quad u_z(0) = u_{0,z}$$

with $\mathbf{u}_0 = (u_{0,z})_{z\in\mathbb{R}^N} \in H_\rho^{2\ell}(\mathbb{R}^N, \mathbb{P}_Z; D) \cap L^2(\mathbb{R}^N, \mathbb{P}_Z; D^{\alpha+1})$ for $D = H_0^1(G) \cap H^2(G)$. For $n, m, k \in \mathbb{N}$, let $\mathbf{u}_{n,m,k}$ as in (3.4.9), where a quadratic finite element method and implicit Euler are used for discretisation in space and time, respectively.
Then there exist $C_{T,\ell}, C_{T,\alpha} \ge 0$, $\tau_0 > 0$ such that for $\max_{i=1,\dots,N_k} \tau_k^i \le \tau_0$,

$$\|\mathbf{u}_{n,m,k}(t) - \mathbf{u}(t)\|_{\mathbf{H}} \le C_{T,\ell} n^{-\ell} \|\mathbf{u}_0\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;D)} + C_{T,\alpha}\left(m^{-2} + \max_{i=1,\dots,N_k} \tau_k^i\right)\|\mathbf{u}_0\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D^{\alpha+1})}$$

for all $n, m, k \in \mathbb{N}$ and $t \in \mathcal{T}_k$.

Theorem 3.82. *Let $Z : \Omega \to \mathbb{R}^N$ be a random variable satisfying Assumption 3.20. Suppose that $\mathbb{R}^N \times G \ni (z, x) \mapsto a^z(x) \in \mathbb{K}^{2\times 2}$ satisfies (3.5.1) and is contained in*

$C_\rho^{2\ell,3}(\mathbb{R}^N \times G; \mathbb{K}^{2\times 2})$ for some $\ell \in \mathbb{N}$ such that $\operatorname{tr} a^z$ and $\operatorname{tr}\operatorname{Div}(a^z(\cdot)^T)$ are constant for $\mathbb{P}_Z$-almost every $z \in \mathbb{R}^N$. Let $0 < \alpha \leq 1$ and $\mathbf{u} = (u_z)_{z\in\mathbb{R}^N}$ be the mild solution of

$$u_z'(t) = -\operatorname{div} a^z(x) \operatorname{grad} u_z(t) \quad (t > 0), \quad u_z(0) = u_{0,z}$$

with $\mathbf{u}_0 = (u_{0,z})_{z\in\mathbb{R}^N} \in H_\rho^{2\ell}(\mathbb{R}^N, \mathbb{P}_Z; D) \cap L^2(\mathbb{R}^N, \mathbb{P}_Z; D^{\alpha+1})$. For $n, m, k \in \mathbb{N}$, let $\mathbf{u}_{n,m,k}$ as in (3.4.9), where a quadratic finite element method and implicit Euler are used for discretisation in space and time, respectively.
Then there exist $C_{T,\ell}, C_{T,\alpha} \geq 0$, $\tau_0 > 0$ such that for $\max_{i=1,\dots,N_k} \tau_k^i \leq \tau_0$,

$$\begin{aligned}\|\mathbf{u}_{n,m,k}(t) - \mathbf{u}(t)\|_{\mathbf{H}} \leq{}& C_{T,\ell} n^{-\ell} \|\mathbf{u}_0\|_{H_\rho^{2\ell}(\mathbb{R}^N,\mathbb{P}_Z;D)} \\ &+ C_{T,\alpha}\left(m^{-p_{\mathrm{x}}} + \max_{i=1,\dots,N_k} \tau_k^i\right) \|\mathbf{u}_0\|_{L^2(\mathbb{R}^N,\mathbb{P}_Z;D^{\alpha+1})}\end{aligned}$$

for all $n, m, k \in \mathbb{N}$ and $t \in \mathcal{T}_k$ with $p_{\mathrm{x}} = 3$ for $\alpha \geq \frac{1}{2}$ and $p_{\mathrm{x}} = 2$ for $\alpha \in (0, \frac{1}{2})$.

As a concrete example, we consider a random anisotropic diffusion. Let $Z\colon \Omega \to \mathbb{R}$ be a standard normally distributed random variable (i.e., $N = 1$). Let $G = (0, 1)^2$ and for $z \in \mathbb{R}$ and $x \in G$, consider the coefficient matrix given by

$$\begin{aligned} a^z(x) &= f(z) \cdot \begin{pmatrix} g_1(x) & 0 \\ 0 & g_2(x) \end{pmatrix}, & f(z) &= \frac{1}{1 + e^{-z}} + 1, \\ g_1(x) &= 1 + \|x\|_2^2, & g_2(x) &= 3 - \|x\|_2^2. \end{aligned} \tag{3.5.9}$$

We observe that $f(z) \in [1, 2]$ for all $z \in \mathbb{R}$, $g_1(x), g_2(x) \in [1, 3]$ for all $x \in (0, 1)^2$. Hence, $\kappa \leq a^z(x) \leq M$ is satisfied for $\kappa = 1$ and $M = 6$. It remains to verify that $(z, x) \mapsto a^z(x)$ is in $C_\rho^{2\ell,2}(\mathbb{R}\times(0, 1)^2; \mathbb{R}^{2\times 2})$ in order to apply Theorem 3.81. Note that $f \in C^\infty(\mathbb{R}; \mathbb{R}^{2\times 2})$ and thus $a^\cdot(x) \in C^{2\ell}(\mathbb{R}; \mathbb{R}^{2\times 2})$ for any $\ell \in \mathbb{N}$. Also, clearly, $a^z(\cdot) \in C^2((0, 1)^2; \mathbb{R}^{2\times 2})$. Likewise, the product structure of $a^z(x)$ implies $\partial_x^{e_j} a^\cdot(x) \in C^{2\ell}(\mathbb{R}; \mathbb{R}^{2\times 2})$. To establish the bounds stated in Definition 3.78, we first observe that derivatives of f are bounded in z. Indeed, as a logistic function, the derivatives of $F := f - 1$ can be expressed as a polynomial of F. Since all derivatives of f and F agree, $F(z) \in [0, 1]$ for $z \in \mathbb{R}$, and continuous functions are bounded on compact intervals,

$$c_\ell := \sup_{z\in\mathbb{R}} \max_{0\leq\alpha\leq\ell} |(\partial_z^\alpha f)(z)| < \infty$$

for all $\ell \in \mathbb{N}$. Consequently, for all $0 \leq \alpha \leq 2\ell$,

$$\mathbb{P}_Z\text{-}\operatorname*{ess\,sup}_{z\in\mathbb{R}} \|\partial_z^\alpha a^z(\cdot)\|_{L^\infty((0,1)^2;\mathbb{R}^{2\times 2})} \leq \sup_{z\in\mathbb{R}} |(\partial_z^\alpha f)(z)| \sup_{x\in(0,1)^2} \left\| \begin{pmatrix} g_1(x) & 0 \\ 0 & g_2(x) \end{pmatrix} \right\|_2 \leq 3c_{2\ell},$$

where $\rho \equiv 1$ by choice of distribution of Z. Since for all $z \in \mathbb{R}$,

$$\left\|\operatorname{Div}((\partial_z^\alpha a^z(\cdot))^T)\right\|_{L^\infty(G;\mathbb{R}^2)} = |\partial_z^\alpha f(z)| \cdot \sup_{x\in(0,1)^2} \left\| \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} \right\|_2 = 2\sqrt{2} |\partial_z^\alpha f(z)|$$

it also holds that for $0 \leq \alpha \leq 2\ell$

$$\mathbb{P}_Z\text{-}\operatorname*{ess\,sup}_{z\in\mathbb{R}} \left\|\operatorname{Div}((\partial_z^\alpha a^z(\cdot))^T)\right\|_{L^\infty(G;\mathbb{R}^2)} \leq 2\sqrt{2} c_{2\ell}.$$

Proceeding likewise, we conclude analogous statements to the last two ones for $a^z(\cdot)$ replaced by $\partial_x^{e_j} a^z(\cdot)$, $j = 1, 2$, with constant $2c_{2\ell}$ in both cases. In conclusion, the coefficients belong to $C_\rho^{2\ell,2}(\mathbb{R}\times(0,1)^2;\mathbb{R}^{2\times 2})$. Hence, we conclude convergence of arbitrary polynomial order in randomness for the random anisotropic diffusion problem. For simplicity, we fix $0 < \alpha < \frac{1}{4}$.

Corollary 3.83. *Adopt the assumptions and notation of Theorem 3.81. Let $Z\colon \Omega \to \mathbb{R}$ be standard normally distributed and consider $a^z(x)$ as in* (3.5.9)*. Then for all $\ell \in \mathbb{N}$ there exist $C_{T,\ell} \geq 0$ and $\tau_0 > 0$ such that for $\max_{i=1,\dots,N_k} \tau_k^i \leq \tau_0$,*

$$\|\mathbf{u}_{n,m,k}(t) - \mathbf{u}(t)\|_{\mathbf{H}} \leq C_{T,\ell}\Big(n^{-\ell} + m^{-2} + \max_{i=1,\dots,N_k} \tau_k^i\Big)\|\mathbf{u}_0\|_{H^{2\ell}(\mathbb{R},\mathbb{P}_Z;D^{1+\alpha})}$$

for all $n, m, k \in \mathbb{N}$, $t \in \mathcal{T}_k$, $0 < \alpha < \frac{1}{4}$, and $\mathbf{u}_0 \in H^{2\ell}(\mathbb{R},\mathbb{P}_Z;D^{1+\alpha})$.

Chapter 4

Temporal Approximation of Stochastic Evolution Equations

At the heart of this chapter are convergence rates for the pathwise uniform error arising from a temporal semi-discretisation of a stochastic evolution equation. Depending on the kind of noise, scheme, semigroup, and whether the error is considered in the grid points only or on the full time interval $[0,T]$, different rates of convergence are obtained. We summarise them in Tables 4.1-4.3. The error decay stated is the best one that can be attained in case of sufficiently regular nonlinearities, noise, and initial values, as well as schemes of sufficiently high order. For details on the regularity conditions, the reader is referred to the full statement of the theorem indicated in brackets. In Tables 4.2 and 4.3, contractivity of the semigroup is assumed and for the rates on $[0,T]$ that the underlying space is Hilbert. For a more general introduction, we kindly ask the reader to read Section 1.3 beforehand.

	general S, general schemes	quasi-contractive S, exponential Euler
error decay	$(\log(T/k))^{1/2}k$ (Theorem 4.1)	k (Corollary 4.4)

Table 4.1: Maximal convergence rates for the linear equation with additive noise

	general contractive schemes	exponential Euler
error decay	$(\log(T/k))^{1/2}k^{1/2}$ (Theorem 4.24)	$k^{1/2}$ (Corollary 4.26)
on $[0,T]$	$(\log(T/k))^{1/2}k^{1/2}$ (Theorem 4.32)	$(\log(T/k))^{1/2}k^{1/2}$ (Theorem 4.32)

Table 4.2: Maximal convergence rates for the nonlinear equation with multiplicative noise

	general contractive schemes	exponential Euler
error decay	$(\log(T/k))^{1/2}k$ (Theorem 4.48)	k (Corollary 4.49)
on $[0,T]$	$(\log(T/k))^{1/2}k$ (Corollary 4.51)	$(\log(T/k))^{1/2}k$ (Corollary 4.51)

Table 4.3: Maximal convergence rates for abstract wave equations with multiplicative noise

This chapter is organised as follows. First, we study the linear additive noise case in Section 4.1 and semigroups that are not necessarily contractive. Results are illustrated for the linear Schrödinger equation. The nonlinear evolution equation with multiplicative noise is introduced in Section 4.2, and well-posedness is shown on the whole space and, using the Kato setting, on a more regular subspace Y. Stability of contractive time discretisation schemes is obtained in Section 4.3, with two different techniques in Hilbert and 2-smooth Banach spaces. The first main result on pathwise uniform convergence rates is stated in Theorem 4.24 in Section 4.4. It also contains optimal error estimates on the full time interval in Theorem 4.32 based on the novel optimal path regularity result for the solution in Proposition 4.31. Applications to nonlinear Schrödinger equations and Maxwell's equations are included. Convergence rates up to order one are obtained for abstract wave equations in Section 4.5, as illustrated for examples with trace class, space-time white noise, and smooth noise. Lastly, Section 4.6 contains the second main result on convergence of temporal discretisations for equations with irregular Lipschitz nonlinearity and noise as well as rough initial data. This result can be found in Theorem 4.57 and is illustrated for an irregular nonlinear stochastic Schrödinger equation.

Notation

Throughout this chapter, we fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with filtration $(\mathscr{F}_t)_{t\in[0,T]}$ and write $\mathcal{P}$ for the progressive σ-algebra (cf. Definitions 2.41 and 2.45 for the respective definitions). Unless otherwise stated, all random variables and stochastic processes considered are defined on $(\Omega, \mathscr{F}, \mathbb{P})$. Moreover, X and Y denote 2-smooth Banach spaces (cf. Subsection 2.4) and H a Hilbert space, which is used to define the $(\mathscr{F}_t)_{t\in[0,T]}$-cylindrical Brownian motion W_H. Subsequently, the space of γ-radonifying operators from H to X is denoted by $\gamma(H, X)$ (cf. Section 2.5) and the Borel σ-algebra of X by $\mathcal{B}(X)$.

Contrary to Chapter 3, we work on a uniform time grid with time step $k > 0$ and the letter τ is reserved for stopping times. Let the final time $T > 0$ be fixed and consider the uniform time grid with $t_j = jk$ for $j = 0, \dots, N_k$ with $N_k = T/k \in \mathbb{N}$, and define $\lfloor t \rfloor \coloneqq \max\{t_j : t_j \le t\}$ for $t \in [0, T]$. An extension of the results presented to non-uniform grids is possible. Working on a uniform grid allows us to highlight the main results and techniques used to prove them, while keeping the notation concise. Since polynomial chaos does not enter the stage in this chapter, henceforth, we denote by $R\colon [0,\infty) \to \mathcal{L}(X)$ a time discretisation scheme approximating the semigroup $(S(t))_{t\ge 0}$. We write $R_k \coloneqq R(k)$ for $k \ge 0$, and reserve the letter F for nonlinearities. For a given evolution equation, $(U(t))_{t\in[0,T]}$ denotes the exact solution and U^j the numerical solution approximating U at time t_j for $j = 0, \dots, N_k$. For f and g in the respective spaces and Z a Banach space, we write $\|f\|_{p,q,Z} \coloneqq \|f\|_{L^p(\Omega;L^q(0,T;Z))}$ and $\|\!|g|\!\|_{p,q,Z} \coloneqq \|g\|_{L^p(\Omega;L^q(0,T;\mathcal{L}_2(H,Z)))}$ for $p \in [2,\infty)$ and $q \in [1,\infty]$. Additionally, we write $\|\cdot\|_p$ to denote the norm in $L^p(\Omega)$. We use the notation $f \lesssim g$ to denote that there is a constant $C \ge 0$ such that for all x in the respective set, $f(x) \le Cg(x)$. To indicate both $f \lesssim g$ and $g \lesssim f$, we write $f \eqsim g$ for short.

4.1 Convergence Rates for Additive Noise

In this section, we present several results on convergence rates for linear equations with additive noise. The reason to start with this case is twofold. Higher convergence rates can be proved in this case. Moreover, it allows us to explain the new techniques in a simpler setting before passing to the more advanced multiplicative setting of Section 4.4.

Our subject of interest are stochastic evolution equations with additive noise of the form

$$\mathrm{d}U + AU\,\mathrm{d}t = g(t)\,\mathrm{d}W_H(t) \text{ on } [0,T], \quad U(0) = u_0 \in L^p_{\mathcal{F}_0}(\Omega;X), \tag{4.1.1}$$

where $-A$ generates a C_0-semigroup $(S(t))_{t\geq 0}$ on a $(2,D)$-smooth Banach space X with norm $\|\cdot\|$, W_H is an H-cylindrical Brownian motion for some Hilbert space H, and $p \in [2,\infty)$. For Hölder continuous noise $g \in L^p_{\mathcal{P}}(\Omega; C^\alpha([0,T];\gamma(H,X)))$, $\alpha \in (0,1]$, mapping into a more regular space $Y \hookrightarrow X$, we prove rates of convergence for time discretisation schemes. An improvement of the rate is shown for the exponential Euler method for quasi-contractive semigroups, motivating the focus on the (quasi-)contractive case in subsequent sections. Results are illustrated for the linear Schrödinger equation in Subsection 4.1.3.

The mild solution to (4.1.1) is uniquely given by

$$U(t) = S(t)u_0 + \int_0^t S(t-s)g(s)\,\mathrm{d}W_H(s) \tag{4.1.2}$$

for $t \in [0,T]$. To approximate it, we employ a time discretisation scheme $R\colon [0,\infty) \to \mathcal{L}(X)$ with time step $k > 0$ on a uniform grid $\{t_j = jk : \ j = 0,\dots,N_k\} \subseteq [0,T]$ with final time $T = t_{N_k} > 0$ and $N_k = \frac{T}{k} \in \mathbb{N}$ being the number of time steps. The discrete approximation is given by $U^0 \coloneqq u_0$ and

$$U^j \coloneqq R_k U^{j-1} + R_k g(t_{j-1})\Delta W_j = R_k^j u_0 + \sum_{i=0}^{j-1} R_k^{j-i} g(t_i)\Delta W_{i+1}, \quad j = 1,\dots,N_k, \tag{4.1.3}$$

with Wiener increments $\Delta W_j \coloneqq W_H(t_j) - W_H(t_{j-1})$, where we used the shorthand notation (2.7.1).

4.1.1 General semigroups

Our first result concerns general C_0-semigroups $(S(t))_{t\geq 0}$. An improvement under further conditions on S is discussed in Subsection 4.1.2. Below we denote the Hölder seminorm in $C^\alpha([0,T];\gamma(H,X))$ by $[\cdot]_\alpha$ for $\alpha \in (0,1]$ and, for $p \in [2,\infty)$ and $Y \hookrightarrow X$, let

$$\|g\|_{p,\infty,Y} \coloneqq \|g\|_{L^p(\Omega;C([0,T];\gamma(H,Y)))}, \quad (g \in L^p(\Omega; C([0,T];\gamma(H,Y)))). \tag{4.1.4}$$

Theorem 4.1. *Let X and Y be $(2,D)$-smooth Banach spaces such that $Y \hookrightarrow X$. Let $-A$ be the generator of a C_0-semigroup $(S(t))_{t\geq 0}$ on X with $\|S(t)\| \leq M\mathrm{e}^{\lambda t}$ for some $M \geq 1$ and $\lambda \geq 0$. Let $(R_k)_{k>0}$ be a time discretisation scheme and assume that R approximates S to order $\alpha \in (0,1]$ on Y. Suppose that $Y \hookrightarrow D_A(\alpha,\infty)$ continuously if $\alpha \in (0,1)$ or $Y \hookrightarrow \mathrm{D}(A)$ continuously if $\alpha = 1$. Let $p \in [2,\infty)$, $u_0 \in L^p_{\mathcal{F}_0}(\Omega;Y)$, and $g \in L^p_{\mathcal{P}}(\Omega; C([0,T];\gamma(H,Y)))$ as well as $g \in L^p_{\mathcal{P}}(\Omega;C^\alpha([0,T];\gamma(H,X)))$. Denote by U the mild solution of (4.1.1) and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in (4.1.3). Then for $N_k \geq 2$*

$$\left\| \max_{0\leq j\leq N_k} \|U(t_j) - U^j\| \right\|_p \leq \left(C_1 + C_2\sqrt{\max\{\log(T/k), p\}}\right)k^\alpha$$

with constants $C_1 := C_\alpha \|u_0\|_{L^p(\Omega;Y)}$ *and*

$$C_2 := \frac{K_{p,D}\sqrt{T}}{\sqrt{2\alpha+1}}\Big(M\mathrm{e}^{\lambda T}\big\|[g]_\alpha\big\|_p + (2M\mathrm{e}^{\lambda T}C_Y + C_\alpha)|||g|||_{p,\infty,Y}\Big),$$

where $K_{p,D}$ *is as in Proposition 2.62,* C_Y *denotes the embedding constant of* Y *into* $D_A(\alpha,\infty)$ *or* $\mathrm{D}(A)$*, and* $C_\alpha = C_\alpha(T)$ *denotes the constant as in Definition 2.24.*

In particular, the approximations $(U^j)_j$ *converge at rate* $\min\{\alpha, 1\}$ *up to a logarithmic correction factor as* $k \to 0$*.*

Proof. Define $S^k(t) := R_k^j$ for $t \in (t_{j-1}, t_j]$ and let $\lfloor t \rfloor := \max\{t_j : t_j \le t\}$ for $t \in [0,T]$. Then the discrete approximations are given by the integral representation

$$U^j = R_k^j u_0 + \int_0^{t_j} S^k(t_j - s)g(\lfloor s \rfloor)\,\mathrm{d}W_H(s).$$

Combining this representation with the mild solution formula (4.1.2), the error can be bounded by

$$\begin{aligned}
E := \Big\| \max_{0\le j\le N_k} \|U(t_j) - U^j\|\Big\|_p &\le \Big\| \max_{0\le j\le N_k} \|[S(t_j) - R_k^j]u_0\|\Big\|_p \\
&+ \Big\| \max_{0\le j\le N_k} \Big\| \int_0^{t_j} S(t_j - s)[g(s) - g(\lfloor s \rfloor)]\,\mathrm{d}W_H(s)\Big\|\Big\|_p \\
&+ \Big\| \max_{0\le j\le N_k} \Big\| \int_0^{t_j} [S(t_j - s) - S(t_j - \lfloor s \rfloor)]g(\lfloor s \rfloor)\,\mathrm{d}W_H(s)\Big\|\Big\|_p \\
&+ \Big\| \max_{0\le j\le N_k} \Big\| \int_0^{t_j} [S(t_j - \lfloor s \rfloor) - S^k(t_j - s)]g(\lfloor s \rfloor)\,\mathrm{d}W_H(s)\Big\|\Big\|_p \\
&=: E_1 + E_2 + E_3 + E_4.
\end{aligned} \tag{4.1.5}$$

We proceed to estimate all terms individually. Since R approximates S to order α on Y,

$$E_1 \le C_\alpha \|u_0\|_{L^p(\Omega;Y)} k^\alpha. \tag{4.1.6}$$

For the second term, we note that for $s \in [t_\ell, t_{\ell+1})$ for some $0 \le \ell \le N_k$, the definition of the Hölder seminorm $[\cdot]_\alpha$ implies that almost surely

$$\begin{aligned}
\Big\| \sum_{i=0}^{j-1} \mathbf{1}_{[t_i,t_{i+1})}(s)S(t_j - s)[g(s) - g(t_i)]\Big\|_{\gamma(H,X)} &\le \|S(t_j - s)\|_{\mathcal{L}(X)}\|g(s) - g(t_\ell)\|_{\gamma(H,X)} \\
&\le M\mathrm{e}^{\lambda T}[g]_\alpha(s - t_\ell)^\alpha.
\end{aligned}$$

Applying Proposition 2.62 with $\Phi_s^{(j)} = \sum_{i=0}^{j-1} \mathbf{1}_{[t_i,t_{i+1})}(s)S(t_j - s)[g(s) - g(t_i)]$ followed by the above estimate then yields

$$\begin{aligned}
E_2 &= \Big\| \max_{0\le j\le N_k} \Big\| \int_0^{t_j} \sum_{i=0}^{j-1} \mathbf{1}_{[t_i,t_{i+1})}(s)S(t_j - s)[g(s) - g(t_i)]\,\mathrm{d}W_H(s)\Big\|\Big\|_p \\
&\le K_{p,D}\sqrt{\max\{\log(N_k), p\}}\Big\| \Big(\int_0^T \max_{1\le j\le N_k} \|\Phi_s^{(j)}\|_{\gamma(H,X)}^2\,\mathrm{d}s\Big)^{1/2}\Big\|_p
\end{aligned}$$

$$
\begin{aligned}
&\leq K_{p,D} M \mathrm{e}^{\lambda T} \sqrt{\max\{\log(N_k), p\}} \Big\| \Big(\sum_{\ell=0}^{N_k-1} \int_{t_\ell}^{t_{\ell+1}} [g]_\alpha^2 (s-t_\ell)^{2\alpha} \, \mathrm{d}s \Big)^{1/2} \Big\|_p \\
&\leq K_{p,D} M \mathrm{e}^{\lambda T} \frac{1}{\sqrt{2\alpha+1}} \sqrt{\max\{\log(N_k), p\}} k^{\alpha+1/2} \Big\| \Big(\sum_{\ell=0}^{N_k-1} [g]_\alpha^2 \Big)^{1/2} \Big\|_p \\
&= K_{p,D} M \mathrm{e}^{\lambda T} \big\| [g]_\alpha \big\|_p \frac{\sqrt{T}}{\sqrt{2\alpha+1}} \sqrt{\max\{\log(N_k), p\}} k^\alpha, \qquad (4.1.7)
\end{aligned}
$$

where we have used Hölder continuity of g. Analogously, with $\Phi_s^{(j)} = \sum_{i=0}^{j-1} \mathbf{1}_{[t_i,t_{i+1})}(s)$ $[S(t_j - t_i) - S(t_j - s)]g(t_i)$ for E_3 we obtain

$$
E_3 \leq 2K_{p,D} M \mathrm{e}^{\lambda T} C_Y \frac{\sqrt{T}}{\sqrt{2\alpha+1}} |\!|\!| g |\!|\!|_{p,\infty,Y} \sqrt{\max\{\log(N_k), p\}} k^\alpha \qquad (4.1.8)
$$

using pathwise boundedness of g, i.e., $g(\omega, \cdot) : [0,T] \to \mathcal{L}_2(H,Y)$ being bounded for $\mathbb{P}$-almost every $\omega \in \Omega$, and noting that by (2.3.3), almost surely

$$
\big\| [S(t_j - t_\ell) - S(t_j - s)] g(t_\ell) \big\|_{\gamma(H,X)} \leq 2M \mathrm{e}^{\lambda T} C_Y (s - t_\ell)^\alpha \| g(t_\ell) \|_{\gamma(H,Y)}.
$$

Likewise, with $\Phi_s^{(j)} = \sum_{i=0}^{j-1} \mathbf{1}_{[t_i,t_{i+1})}(s)[S(t_j - t_i) - R_k^{j-i}]g(t_i)$ we obtain

$$
E_4 \leq K_{p,D} C_\alpha \frac{\sqrt{T}}{\sqrt{2\alpha+1}} |\!|\!| g |\!|\!|_{p,\infty,Y} \sqrt{\max\{\log(N_k), p\}} k^\alpha, \qquad (4.1.9)
$$

since R approximates S to order α on Y. The error bound follows from inserting (4.1.6), (4.1.7), (4.1.8), and (4.1.9) into (4.1.5). □

For the *exponential Euler method*, sometimes also referred to as the *splitting scheme*, less regularity of the initial value suffices for the same convergence behaviour. The exponential Euler method is obtained by setting $R_k = S(k)$ in (4.1.3), i.e., we would solve exactly in the absence of noise g.

Corollary 4.2 (Exponential Euler). *Let X and Y be $(2, D)$-smooth Banach spaces such that $Y \hookrightarrow X$. Let $-A$ be the generator of a C_0-semigroup $(S(t))_{t\geq 0}$ on X with $\|S(t)\| \leq M\mathrm{e}^{\lambda t}$ for some $M \geq 1$ and $\lambda \geq 0$. Assume that $g \in L^p_{\mathcal{P}}(\Omega; C([0,T]; \gamma(H,Y)))$ and $g \in L^p_{\mathcal{P}}(\Omega; C^\alpha([0,T]; \gamma(H,X)))$ for some $\alpha \in (0,1]$. Suppose that $Y \hookrightarrow D_A(\alpha,\infty)$ continuously if $\alpha \in (0,1)$ or $Y \hookrightarrow \mathrm{D}(A)$ continuously if $\alpha = 1$. Let $p \in [2,\infty)$ and $u_0 \in L^p_{\mathcal{F}_0}(\Omega; X)$. Denote by U the mild solution of (4.1.1) and by $(U^j)_{j=0,\ldots,N_k}$ the temporal approximations as defined in (4.1.3) obtained with the exponential Euler method $R := S$. Then for $N_k \geq 2$*

$$
\Big\| \max_{0\leq j\leq N_k} \| U(t_j) - U^j \| \Big\|_p \leq C \sqrt{\max\{\log(T/k), p\}} k^\alpha
$$

with constant

$$
C := K_{p,D} M \mathrm{e}^{\lambda T} \frac{\sqrt{T}}{\sqrt{2\alpha+1}} \big(\big\| [g]_\alpha \big\|_p + 2C_Y |\!|\!| g |\!|\!|_{p,\infty,Y} \big),
$$

where $K_{p,D}$ is as in Proposition 2.62 and C_Y denotes the embedding constant of Y into $D_A(\alpha,\infty)$ or $\mathrm{D}(A)$.

In particular, if $Y \hookrightarrow \mathrm{D}(A)$ and g is Lipschitz continuous as a map to $\gamma(H,X)$, the approximations $(U^j)_j$ converge at rate 1 up to a logarithmic correction factor as $k \to 0$.

We emphasize that, in the additive case, convergence rates can be obtained even for rough initial values for the exponential Euler method. This is a distinctive feature of the additive case and does not carry over to the general multiplicative setting (cf. Subsection 4.4.2).

Proof. We split the error as in (4.1.5). For the exponential Euler method, the terms E_1 and E_4 in (4.1.5) vanish due to $S(t_j) - R_k^j = S(jk) - S(k)^j = S(jk) - S(jk) = 0$ and, likewise, $S(t_j - t_i) - R_k^{j-i} = 0$. The error bound follows from inserting the bounds (4.1.7) and (4.1.8) of the remaining terms into (4.1.5). □

4.1.2 Quasi-contractive semigroups

Quasi-contractive semigroups are semigroups for which $\|S(t)\| \leq \mathrm{e}^{\lambda t}$ for some $\lambda \geq 0$ for all $t \geq 0$. Considering such semigroups allows us to eliminate the logarithmic factor for the exponential Euler method. The principle that lies at the heart of our proof is the maximal inequality from Theorem 2.61, which is used to estimate the stochastic convolutions in the error term. Depending on the spatial regularity of the noise g, convergence at rate $\alpha \in (0, 1]$ is attained without a logarithmic correction factor.

Theorem 4.3 (Exponential Euler, quasi-contractive case)**.** *Adopt the notation and assumptions of Corollary 4.2. In addition, assume that $\|S(t)\| \leq \mathrm{e}^{\lambda t}$ for some $\lambda \geq 0$ for all $t \in [0, T]$. Then for $N_k \geq 2$*

$$\Big\| \max_{0 \leq j \leq N_k} \|U(t_j) - U^j\| \Big\|_p \leq C k^\alpha$$

with constant

$$C := \frac{C_{p,D}\sqrt{T}}{\sqrt{2\alpha + 1}} \big(\big\| [g]_\alpha \big\|_p + 2 C_Y \mathrm{e}^{\lambda T} \|g\|_{p,\infty,Y} \big),$$

where $C_{p,D}$ is the constant from Theorem 2.61.

Proof. We bound the error as in (4.1.5), where the first and fourth term vanish as discussed in the proof of Corollary 4.2. We proceed to bound the remaining terms using the maximal inequality from Theorem 2.61 instead of Proposition 2.62 to obtain

$$\begin{aligned} E_2 &\leq \Big\| \sup_{t \in [0,T]} \Big\| \int_0^t S(t-s)[g(s) - g(\lfloor s \rfloor)] \, \mathrm{d}W_H(s) \Big\| \Big\|_p \\ &\leq C_{p,D} \Big\| \Big(\int_0^T \|g(s) - g(\lfloor s \rfloor)\|_{\gamma(H,X)}^2 \, \mathrm{d}s \Big)^{1/2} \Big\|_p \\ &\leq C_{p,D} \Big\| \Big(\sum_{i=0}^{N_k - 1} \int_{t_i}^{t_{i+1}} [g]_\alpha^2 (s - t_i)^{2\alpha} \, \mathrm{d}s \Big)^{1/2} \Big\|_p \\ &\leq \frac{C_{p,D}\sqrt{T}}{\sqrt{2\alpha + 1}} \big\| [g]_\alpha \big\|_p k^\alpha \end{aligned} \tag{4.1.10}$$

by Hölder continuity of g. Analogously, for E_3 we deduce from the semigroup bound (2.3.3) that

$$\begin{aligned}
E_3 &\leq \Big\| \sup_{t\in[0,T]} \Big\| \int_0^t S(t-s)[S(s-\lfloor s\rfloor)-I]g(\lfloor s\rfloor)\,\mathrm{d}W_H(s)\Big\|\Big\|_p \\
&\leq C_{p,D}\Big\|\Big(\int_0^T \|[S(s-\lfloor s\rfloor)-I]g(\lfloor s\rfloor)\|^2_{\gamma(H,X)}\,\mathrm{d}s\Big)^{1/2}\Big\|_p \\
&\leq 2C_{p,D}\mathrm{e}^{\lambda T}C_Y\Big\|\Big(\sum_{i=0}^{N_k-1}\int_{t_i}^{t_{i+1}}(s-t_i)^{2\alpha}\|g(t_i)\|^2_{\gamma(H,Y)}\,\mathrm{d}s\Big)^{1/2}\Big\|_p \\
&\leq 2C_{p,D}\mathrm{e}^{\lambda T}C_Y\frac{\sqrt{T}}{\sqrt{2\alpha+1}}\|\!|g|\!\|_{p,\infty,Y}k^\alpha. \qquad (4.1.11)
\end{aligned}$$

The final error bound follows from adding (4.1.10) and (4.1.11). □

In particular, convergence rate 1 is attained without logarithmic correction factor for spatially sufficiently regular noise g. General, possibly irregular initial values from $L^p_{\mathcal{F}_0}(\Omega;X)$ are still admissible, as the following corollary shows.

Corollary 4.4. *Let X be a $(2,D)$-smooth Banach space and let $-A$ be the generator of a quasi-contractive C_0-semigroup on X with parameter $\lambda > 0$. Assume that $g \in L^p_{\mathcal{P}}(\Omega;C([0,T];\gamma(H,\mathrm{D}(A))))$ is pathwise Lipschitz continuous as a map to $\gamma(H,X)$. Let $p \in [2,\infty)$ and $u_0 \in L^p_{\mathcal{F}_0}(\Omega;X)$. Denote by U the mild solution of (4.1.1) and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in (4.1.3) obtained with the exponential Euler method $R := S$. Then there is a constant $C \geq 0$ depending on $(g,T,p,\alpha,\lambda,X,\mathrm{D}(A))$ such that for $N_k \geq 2$*

$$\Big\|\max_{0\leq j\leq N_k}\|U(t_j)-U^j\|\Big\|_p \leq Ck,$$

i.e., the approximations $(U^j)_j$ converge at rate 1 as $k \to 0$.

4.1.3 Application to the linear Schrödinger equation with additive noise

In this subsection, we study convergence rates of time discretisations of the linear stochastic Schrödinger equation with a potential and additive noise

$$\begin{cases} \mathrm{d}u = -\mathrm{i}(\Delta+V)u\,\mathrm{d}t - \mathrm{i}\,\mathrm{d}W & \text{on } [0,T], \\ u(0) = u_0 \end{cases} \qquad (4.1.12)$$

in $\mathbb{R}^d$ for $d \in \mathbb{N}$, where $(W(t))_{t\geq 0}$ is a square-integrable $\mathbb{K}$-valued Q-Wiener process (cf. Section 2.7), $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$, with respect to a normal filtration $(\mathscr{F}_t)_{t\geq 0}$, V is a $\mathbb{K}$-valued potential and u_0 is an $\mathscr{F}_0$-measurable random variable. Next, we introduce conditions on the dimension and on the regularity of V. With a slight variation of the methods below, one can also consider (4.1.12) on $[0,L]^d$ with periodic boundary conditions. More general domains with Dirichlet or Neumann boundary conditions can be treated as well, but for this, suitable adjustments are needed in the proofs below.

Let $\sigma \geq 0$ and, for this subsection only, write $L^2 = L^2(\mathbb{R}^d)$ and $H^\sigma = H^\sigma(\mathbb{R}^d)$. We will also be using the Bessel potential spaces $H^{\sigma,q}(\mathbb{R}^d)$, which coincide with the classical

Sobolev spaces $W^{\sigma,q}(\mathbb{R}^d)$ if $\sigma \in \mathbb{N}$ and $q \in (1,\infty)$. For details on these spaces, the reader is referred to Section 2.3 and [15, 126].

To ensure the well-posedness of (4.1.12), we assume one of the following mutually exclusive conditions holds.

Assumption 4.5. *Let $\sigma \geq 0$, $d \in \mathbb{N}$, and $V \in L^2$ such that*

(i) $\sigma > \frac{d}{2}$ and $V \in H^\sigma$, or

(ii) $\sigma = 0$ and $V \in H^\beta$ for some $\beta > \frac{d}{2}$, or

(iii) $\sigma \in (0,1)$, $d > 2\sigma$, and $V \in H^\beta$ for some $\beta > \frac{d}{2}$, or

(iv) $\sigma = 1$, $d \geq 2$, and $V \in H^\beta$ for some $\beta > \frac{d}{2}$.

In particular, this assumption implies that $Vu \in H^\sigma$ for any $u \in H^\sigma$ and $\|Vu\|_{H^\sigma} \leq C_V \|u\|_{H^\sigma}$ for some constant $C_V \geq 0$ depending on V. This follows from the algebra property of H^σ in case (i). Note that while (i) is taken verbatim from [2, Prop. 4.1], cases (ii) and (iv) assume less regularity compared to [2] and case (iii) is new. In the second case, Hölder's inequality and the Sobolev embedding $H^\beta \hookrightarrow L^\infty$ for $\beta > \frac{d}{2}$ yield

$$\|Vu\|_{L^2} \leq \|V\|_{L^\infty} \|u\|_{L^2} \lesssim \|V\|_{H^\beta} \|u\|_{L^2}$$

in the case (ii), see [2, Prop. 4.1]. The case (iii) is covered by Lemma 4.6 below. Lastly, $\|Vu\|_{H^1} \lesssim \|u\|_{H^1}$ in the case (iv) follows from Hölder's inequality, once with $p = 2\beta$ and $q = \frac{4\beta}{2\beta-2}$, $\beta > 1$, and once with $p = \infty$ and $q = 2$. Moreover, we make use of the embeddings $H^\beta \hookrightarrow L^\infty$, $H^1 \hookrightarrow L^q$, as well as $H^\beta \hookrightarrow W^{1,2\beta}$, where the latter holds due to $\beta - \frac{d}{2} = \beta(1 - \frac{d}{2\beta}) > \frac{d}{2}(1 - \frac{d}{2\beta}) \geq 1 - \frac{d}{2\beta}$. Altogether, this yields

$$\begin{aligned} \|Vu\|_{H^1}^2 &\lesssim \|Vu\|_{L^2}^2 + \|V \operatorname{grad} u\|_{L^2}^2 + \|(\operatorname{grad} V)u\|_{L^2}^2 \\ &\leq \|V\|_{L^\infty}^2 (\|u\|_{L^2}^2 + \|\operatorname{grad} u\|_{L^2}^2) + \|\operatorname{grad} V\|_{L^{2\beta}}^2 \|u\|_{L^q}^2 \\ &\lesssim (\|V\|_{L^\infty}^2 + \|V\|_{W^{1,2\beta}}^2) \|u\|_{H^1}^2 \lesssim \|V\|_{H^\beta}^2 \|u\|_{H^1}^2. \end{aligned}$$

Hence, multiplication by V is a bounded operator on H^σ if Assumption 4.5 holds.

Lemma 4.6. *Let $\sigma \in (0,1)$, $d \in \mathbb{N}$ such that $d > 2\sigma$, and $V \in H^\beta(\mathbb{R}^d)$ for some $\beta > \frac{d}{2}$. Then $\|Vu\|_{H^\sigma} \leq C_V \|u\|_{H^\sigma}$ for some constant $C_V \geq 0$ for all $u \in H^\sigma(\mathbb{R}^d)$.*

Proof. Let $q_1 = \frac{2d}{d-2\sigma}$ and $q_2 = \frac{d}{\sigma}$. Then $\frac{1}{q_1} + \frac{1}{q_2} = \frac{1}{2}$ and $q_1 < \infty$. By classical Sobolev and Bessel potential space embeddings (cf. Theorem 2.31), $H^{d/2} \hookrightarrow H^{\sigma,q_2}$, $H^\sigma \hookrightarrow L^{q_1}$, and $H^\beta \hookrightarrow C_b(\mathbb{R}^d) \hookrightarrow L^\infty$. Thus, an application of the product estimate from Proposition 2.32 yields

$$\begin{aligned} \|Vu\|_{H^\sigma} &\lesssim \|V\|_{H^{\sigma,q_2}} \|u\|_{L^{q_1}} + \|V\|_{L^\infty} \|u\|_{H^\sigma} \\ &\lesssim (\|V\|_{H^{d/2}} + \|V\|_{H^\beta}) \|u\|_{H^\sigma} \lesssim \|V\|_{H^\beta} \|u\|_{H^\sigma}. \end{aligned}$$ □

Since $-\mathrm{i}\Delta$ generates a C_0-contraction semigroup on H^σ for any $\sigma > 0$ [2, Lem. 2.1], its bounded perturbation $-\mathrm{i}(\Delta+V)$ generates a quasi-contractive semigroup on the same space H^σ [45, Thm. III.1.3]. Thus, we are in the setting of Subsection 4.1.2. Global existence and uniqueness of mild solutions $U \in L^p(\Omega; C([0,T]; H^\sigma))$ to (4.1.12) in H^σ are guaranteed provided that $p \in [2, \infty)$, $u_0 \in L^p_{\mathcal{F}_0}(\Omega; H^\sigma)$, $Q^{1/2} \in \mathcal{L}_2(L^2, H^\sigma)$, and Assumption 4.5 holds. Note that since H^σ is a Hilbert space, $\gamma(L^2, H^\sigma)$ equals the space of Hilbert–Schmidt operators from L^2 to H^σ.

In conclusion, the Schrödinger equation (4.1.12) can be rewritten in the form of (4.1.1) on $X = H^\sigma$ with an H-cylindrical Brownian motion W_H for $H = L^2$.

For exponential Euler, we recover the error bound from [2, Thm. 4.3], showing convergence at rate 1 in the case of sufficiently regular $Q^{1/2}$ under less regularity assumptions on V. Additionally, we obtain an error bound for fractional convergence rates $\alpha \in (0, 1]$ under weaker regularity assumptions on $Q^{1/2}$ and V.

Theorem 4.7. *Let $\sigma \geq 0$, $d \in \mathbb{N}$, and $V \in L^2$ satisfy Assumption 4.5, and let $p \in [2, \infty)$. Assume that $u_0 \in L^p_{\mathcal{F}_0}(\Omega; H^\sigma)$ and $Q^{1/2} \in \mathcal{L}_2(L^2, H^{\sigma+2\alpha})$ for some $\alpha \in (0, 1]$. Denote by U the mild solution of the linear stochastic Schrödinger equation with additive noise* (4.1.12) *and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in* (4.1.3) *obtained with the exponential Euler method $R \coloneqq S$. Then there exists a constant $C \geq 0$ depending on $(V, u_0, T, p, \alpha, \sigma, d)$ such that for $N_k \geq 2$*

$$\left\| \max_{0 \leq j \leq N_k} \|U(t_j) - U^j\|_{H^\sigma} \right\|_p \leq C \|Q^{1/2}\|_{\mathcal{L}_2(L^2, H^{\sigma+2\alpha})} k^\alpha.$$

Proof. As discussed above, $-A = -\mathrm{i}(\Delta + V)$ generates a quasi-contractive semigroup on H^σ. Furthermore, setting $g = -\mathrm{i}Q^{1/2}$ allows us to rewrite (4.1.12) in the form of a stochastic evolution equation (4.1.1). Thus, Theorem 4.3 is applicable with $X = H^\sigma$ and $H = L^2$. It remains to check that $g \in L^p_{\mathcal{P}}(\Omega; C([0,T]; \gamma(H, Y)))$ for some $Y \hookrightarrow D_A(\alpha, \infty)$ and that $g \in L_{\mathcal{P}}(\Omega; C^\alpha([0,T]; \gamma(H, X)))$. The latter holds for any $\alpha \in (0, 1]$ due to g being constant in time. Taking $Y = H^{\sigma+2\alpha} = (H^\sigma, H^{\sigma+2})_{\alpha,2} = (H^\sigma, \mathrm{D}(A))_{\alpha,2} \hookrightarrow (H^\sigma, \mathrm{D}(A))_{\alpha,\infty}$, the first condition is satisfied as well. Corollary 4.2 yields the desired error bound. □

Furthermore, Theorem 4.1 enables us to extend [2, Thm. 4.3] to general discretisation schemes R involving rational approximations, at the price of an additional logarithmic factor. We state it for implicit Euler and Crank–Nicolson.

Theorem 4.8. *Let $\sigma \geq 0$, $d \in \mathbb{N}$, and $V \in L^2$ satisfy Assumption 4.5, and let $p \in [2, \infty)$. Let $(R_k)_{k>0}$ be the implicit Euler method (IE) or the Crank–Nicolson method (CN) and set $\ell = 4$ or $\ell = 3$, respectively. Assume that $u_0 \in L^p_{\mathcal{F}_0}(\Omega; H^{\sigma+\ell\alpha})$ and $Q^{1/2} \in \mathcal{L}_2(L^2, H^{\sigma+\ell\alpha})$ for some $\alpha \in (0, 1]$. Let $(R_k)_{k>0}$ be a time discretisation scheme that is contractive on H^σ and $H^{\sigma+2\alpha}$. Assume R approximates S to order α on $H^{\sigma+2\alpha}$. Denote by U the mild solution of the linear stochastic Schrödinger equation with additive noise* (4.1.12) *and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in* (4.1.3). *Then there exists a constant $C \geq 0$ depending on $(V, u_0, T, p, \alpha, \sigma, d, \ell)$ such that for $N_k \geq 2$*

$$\left\| \max_{0 \leq j \leq N_k} \|U(t_j) - U^j\|_{H^\sigma} \right\|_p \leq C \|Q^{1/2}\|_{\mathcal{L}_2(L^2, H^{\sigma+\ell\alpha})} \sqrt{\max\{\log(T/k), p\}}\, k^\alpha.$$

Proof. This follows from Theorem 4.1 noting that (IE) approximates S to order α on $\mathrm{D}((-A)^{2\alpha})$ and this fractional domain is given by $\mathrm{D}((\mathrm{i}\Delta)^{2\alpha}) = [H^\sigma, H^{\sigma+4}]_\alpha = H^{\sigma+4\alpha}$, which is chosen as the space Y. Here, $[\cdot,\cdot]_\alpha$ denotes the complex interpolation space with parameter α. Likewise, (CN) approximates S to order α on $\mathrm{D}((-A)^{3\alpha/2})$, which is equal to the space $[H^\sigma, H^{\sigma+3}]_\alpha = H^{\sigma+3\alpha}$. □

Comparing this result to Theorem 4.7 for exponential Euler (EE), it becomes apparent that lower-order schemes like (IE) need higher regularity of the noise $Q^{1/2}$ to achieve the same rate of convergence ($\mathcal{L}_2(L^2, H^{\sigma+4\alpha})$ compared to $\mathcal{L}_2(L^2, H^{\sigma+2\alpha})$). For instance, for $Q^{1/2} \in \mathcal{L}_2(L^2, H^{\sigma+2})$, the rates for (EE), (CN), and (IE) are 1, $\frac{2}{3}$, and $\frac{1}{2}$, respectively. If $Q^{1/2} \in \mathcal{L}_2(L^2, H^{\sigma+3})$, (EE) and (CN) have the same convergence rates up to a logarithmic factor, and if $Q^{1/2} \in \mathcal{L}_2(L^2, H^{\sigma+4})$, so does (IE), all provided that V and u_0 are sufficiently smooth.

Note that in the absence of a potential, the same convergence rates are obtained without any limitation on the dimension $d \in \mathbb{N}$ in terms of the parameter σ. An analogue of Theorem 4.8 can be obtained for other implicit Runge–Kutta methods if the space is known on which the scheme approximates the semigroup to a given order.

4.2 Well-posedness

Well-posedness of an equation refers to the existence and uniqueness of solutions to this equation. Moreover, one is interested in how the solution depends on the initial data. The equation we shall study well-posedness of is the stochastic evolution equation with multiplicative noise

$$\begin{cases} \mathrm{d}U + AU\,\mathrm{d}t = F(t,U)\,\mathrm{d}t + G(t,U)\,\mathrm{d}W_H & \text{on } [0,T], \\ U(0) = u_0 \in L^p_{\mathscr{F}_0}(\Omega;X) \end{cases} \tag{4.2.1}$$

for $1 \leq p < \infty$, $T > 0$, and $-A$ generating a C_0-contraction semigroup $(S(t))_{t\geq 0}$ on a 2-smooth Banach space X. In this section, we present progressive measurability, linear growth, and global Lipschitz conditions on u_0, F, and G ensuring the well-posedness of the above equation. Well-posedness shall be understood in the sense of existence and uniqueness of *mild solutions* to (4.2.1) as well as linear growth of the mild solution w.r.t. the initial data u_0. We denote the progressive σ-algebra by $\mathcal{P}$, cf Definition 2.45.

Definition 4.9. A *mild solution* to (4.2.1) is a function $U \in L^0_{\mathcal{P}}(\Omega; C([0,T];X))$ that almost surely solves the fixed point problem

$$U(t) = S(t)u_0 + \int_0^t S(t-s)F(s,U(s))\,\mathrm{d}s + \int_0^t S(t-s)G(s,U(s))\,\mathrm{d}W_H(s) \tag{4.2.2}$$

for all $t \in [0,T]$.

First, we show an auxiliary density result.

Lemma 4.10. *Denote by $\mathcal{E}_1$ and $\mathcal{E}_2$ the sets of all adapted step processes $\phi\colon \Omega\times[0,T] \to X$ and $\phi\colon \Omega \times [0,T] \to \gamma(H,X)$, respectively. Then the following holds.*

(a) $\mathcal{E}_1$ is dense in $L^0_{\mathcal{P}}(\Omega; L^1(0,T;X))$.

(b) $\mathcal{E}_2$ is dense in $L^0_{\mathcal{P}}(\Omega; L^2(0,T;\gamma(H,X)))$.

Proof. Let $n \in \mathbb{N}$, $I_1^n := [0, \frac{T}{2^n}]$, and, for $\ell = 2,\dots,2^n$, let $I_\ell^n := (\frac{\ell-1}{2^n}T, \frac{\ell}{2^n}T]$. Consider the partition $D_n := \sigma(\{I_\ell^n : \ell = 1,\dots,2^n\})$ as the σ-algebra generated by the sets obtained by dividing each interval I_ℓ^{n-1} in half. This defines a filtration $(D_n)_{n\in\mathbb{N}}$. For $t \in (0,T)$, denote by $\lfloor t \rfloor := \max_{0\leq\ell\leq 2^n}\{\frac{\ell}{2^n}T : \frac{\ell}{2^n}T \leq t\}$ and $\lceil t \rceil := \min_{1\leq \ell\leq 2^n}\{\frac{\ell}{2^n}T : \frac{\ell}{2^n}T > t\}$ the left and right interval boundaries of the interval I_ℓ^n containing t.
We start by showing (a). Let $\phi \in L^0_{\mathcal{P}}(\Omega; L^1(0,T;X))$ and define the step process

$$\phi_n := \mathbb{E}(L_n(\phi)|D_n) = \sum_{\ell=1}^{2^n} \mathbf{1}_{I_\ell^n} \frac{1}{|I_\ell^n|}\int_{I_\ell^n} L_n(\phi)(s)\,\mathrm{d}s, \quad L_n(\phi) := \mathbf{1}_{\left[\frac{T}{2^n},T\right]}\phi\Big(\cdot - \frac{T}{2^n}\Big).$$

Fix some $t \in (0,T)$. Since ϕ is progressively measurable, in particular $\phi(s)$ is $\mathscr{F}_s$-measurable for $s \in [0,T]$ and its left translation $L_n(\phi)(s)$ is $(\mathscr{F}_{s-\frac{T}{2^n}})$-measurable for $s \in [\frac{T}{2^n}, T]$. By construction, $\phi_n(t)$ is $(\mathscr{F}_{\lceil t\rceil - \frac{T}{2^n}})$-measurable and thus $\mathscr{F}_t$-measurable, since $\lceil t \rceil - \frac{T}{2^n} = \lfloor t \rfloor \leq t$. Furthermore, $\phi(0) = 0$ is $\mathscr{F}_0$-measurable and $\phi(T)$ is $\mathscr{F}_T$-measurable by an analogous argument. Hence, $\phi_n \in \mathcal{E}_1$.

To show $\phi_n \to \phi$ as $n \to \infty$ in $L^0_{\mathcal{P}}(\Omega; L^1(0,T;X))$, we show convergence in $L^1(0,T;X)$ for almost every $\omega \in \Omega$. For the sake of readability, we fix $\omega \in \Omega$, omit the evaluation in ω, and use the same notation for the functions acting on $L^1(0,T;X)$ as on $L^0_{\mathcal{P}}(L^1(0,T;X))$.

We observe that L_n and $\mathbb{E}(\cdot|D_n)$ are contractive on L^1, since for $\Phi \in L^1(0,T;X)$,

$$\|L_n\Phi\|_1 = \int_{\frac{T}{2^n}}^{T} \Big\|\Phi\Big(t - \frac{T}{2^n}\Big)\Big\| \,\mathrm{d}t = \int_0^{T-\frac{T}{2^n}} \|\Phi(t)\| \,\mathrm{d}t \le \|\Phi\|_1$$

and, by Fubini's theorem,

$$\|\mathbb{E}(\Phi|D_n)\|_1 \le \sum_{\ell=1}^{2^n} \int_{I^n_\ell} \frac{1}{|I^n_\ell|} \int_{I^n_\ell} \|\Phi(s)\| \,\mathrm{d}s \,\mathrm{d}t = \sum_{\ell=1}^{2^n} \int_{I^n_\ell} \|\Phi(s)\| \,\mathrm{d}s = \|\Phi\|_1.$$

Let $\varepsilon > 0$. By density of the continuous functions in $L^1(0,T;X)$, we can choose $\psi, \tilde{\psi} \in C([0,T];X)$ such that $\|L_n(\phi) - \tilde{\psi}\|_1 < \varepsilon$ and $\|\phi - \psi\|_1 < \varepsilon$. Contractivity of L_n and $\mathbb{E}(\cdot|D_n)$ as well as $\phi_n = \mathbb{E}(L_n(\phi)|D_n)$ then imply

$$\begin{aligned}
\|\phi - \phi_n\|_1 &\le \|\phi - \psi\|_1 + \|\psi - L_n(\psi)\|_1 + \|L_n(\psi) - L_n(\phi)\|_1 \\
&\quad + \|L_n(\phi) - \tilde{\psi}\|_1 + \|\tilde{\psi} - \mathbb{E}(\tilde{\psi}|D_n)\|_1 + \|\mathbb{E}(\tilde{\psi}|D_n) - \phi_n\|_1 \\
&\le 2\|\phi - \psi\|_1 + \|\psi - L_n(\psi)\|_1 + 2\|L_n(\phi) - \tilde{\psi}\|_1 + \|\tilde{\psi} - \mathbb{E}(\tilde{\psi}|D_n)\|_1 \\
&< 4\varepsilon + \|\psi - L_n(\psi)\|_1 + \|\tilde{\psi} - \mathbb{E}(\tilde{\psi}|D_n)\|_1
\end{aligned} \tag{4.2.3}$$

for almost every $\omega \in \Omega$. Continuity of ψ on the compact interval $[0,T]$ yields that $\|\psi(t) - \psi(s)\| \le \frac{\varepsilon}{2T}$ uniformly for all $|t-s| \le \frac{T}{2^n}$ with n sufficiently large. Choosing n such that also $\frac{T}{2^n} \le \frac{\varepsilon}{2\|\psi\|_\infty}$, we obtain

$$\begin{aligned}
\|\psi - L_n(\psi)\|_1 &= \int_0^{\frac{T}{2^n}} \|\psi(t)\| \,\mathrm{d}t + \int_{\frac{T}{2^n}}^{T} \Big\|\psi(t) - \psi\Big(t - \frac{T}{2^n}\Big)\Big\| \,\mathrm{d}t \\
&\le \|\psi\|_\infty \cdot \frac{T}{2^n} + \frac{\varepsilon}{2T}\Big(T - \frac{T}{2^n}\Big) \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned} \tag{4.2.4}$$

Continuity of $\tilde{\psi}$ yields that $\|\tilde{\psi}(t) - \tilde{\psi}(s)\| \le \frac{\varepsilon}{T}$ uniformly for all $|t-s| \le \frac{T}{2^n}$ with n sufficiently large, whence

$$\begin{aligned}
\|\tilde{\psi} - \mathbb{E}(\tilde{\psi}|D_n)\|_1 &= \int_0^T \|\tilde{\psi}(t) - \mathbb{E}(\tilde{\psi}|D_n)(t)\| \,\mathrm{d}t \\
&\le \sum_{\ell=1}^{2^n} \int_{I^n_\ell} \frac{1}{|I^n_\ell|} \int_{I^n_\ell} \big\|\tilde{\psi}(t) - \tilde{\psi}(s)\big\| \,\mathrm{d}s \,\mathrm{d}t \\
&\le \sum_{\ell=1}^{2^n} \frac{\varepsilon}{T} |I^n_\ell| \int_{I^n_\ell} \frac{1}{|I^n_\ell|} \,\mathrm{d}t = \frac{\varepsilon}{T} \sum_{\ell=1}^{2^n} |I^n_\ell| = \frac{\varepsilon}{T} \cdot T = \varepsilon.
\end{aligned} \tag{4.2.5}$$

Inserting (4.2.4) and (4.2.5) into (4.2.3) yields $\phi_n(\omega,\cdot) \to \phi(\omega,\cdot)$ in $L^1(0,T;X)$ for almost every $\omega \in \Omega$, that is, $\phi_n \to \phi$ in $L^0_{\mathcal{P}}(\Omega; L^1(0,T;X))$.

We proceed by showing (b). Let $\phi \in L^0_{\mathcal{P}}(\Omega; L^2(0,T;\gamma(H,X)))$. For $m \in \mathbb{N}$, define the stopping time $\tau_m \colon \Omega \to [0,T]$ as

$$\tau_m := \inf\Big\{t \in [0,T] : \int_0^t \|\phi(s)\|^2_{\gamma(H,X)} \,\mathrm{d}s \geq m\Big\} \tag{4.2.6}$$

if the infimum is taken over a non-empty set and as $\tau_m := T$ otherwise. Then $\{\tau_m \leq t\} \in \mathscr{F}_t$ for all $t \in [0,T]$, which implies the progressive measurability of $\mathbf{1}_{[0,\tau_m]}$ due to $\mathbf{1}_{[0,\tau_m]}\big|_{[0,t]\times\Omega} = \mathbf{1}_{[0,\tau_m]\times\{\tau_m \leq t\}}$ for all $t \in [0,T]$. Denote the restriction of ϕ to times until the stopping time by $\phi^m := \phi \mathbf{1}_{[0,\tau_m]}$. It is progressively measurable and

$$\|\phi - \phi^m\|^2_{L^2(0,T;\gamma(H,X))} = \int_{\tau_m}^T \|\phi(t)\|^2_{\gamma(H,X)} \,\mathrm{d}t \to 0, \quad m \to \infty,$$

as for m large enough $\tau_m = T$ by definition of τ_m and square-integrability of ϕ. Thus,

$$\phi^m \to \phi \text{ in } L^0_{\mathcal{P}}(\Omega; L^2(0,T;\gamma(H,X))) \text{ as } m \to \infty. \tag{4.2.7}$$

Define the step function approximations $\phi^m_n := \mathbb{E}(L_n(\phi^m)|D_n)$ of ϕ^m for $n, m \in \mathbb{N}$ analogously to the proof of part (a). By the same reasoning, ϕ^m_n is an adapted step process such that $\phi^m_n \in \mathcal{E}_2$. For $\Phi \in L^2(0,T;\gamma(H,X))$, the contractivity of L_n and $\mathbb{E}(\cdot|D_n)$ on L^2 follows from

$$\|L_n\Phi\|_2^2 = \int_{\frac{T}{2^n}}^T \Big\|\Phi\Big(t - \frac{T}{2^n}\Big)\Big\|^2_{\gamma(H,X)} \,\mathrm{d}t = \int_0^{T-\frac{T}{2^n}} \|\Phi(t)\|^2_{\gamma(H,X)} \,\mathrm{d}t \leq \|\Phi\|_2^2$$

and, using Fubini's theorem and Hölder's inequality,

$$\begin{aligned}
\|\mathbb{E}(\Phi|D_n)\|_2^2 &\leq \sum_{\ell=1}^{2^n} \int_{I^n_\ell} \frac{1}{|I^n_\ell|^2} \Big(\int_{I^n_\ell} \|\Phi(s)\|_{\gamma(H,X)} \,\mathrm{d}s\Big)^2 \,\mathrm{d}t \\
&\leq \sum_{\ell=1}^{2^n} \int_{I^n_\ell} \frac{1}{|I^n_\ell|} \int_{I^n_\ell} \|\Phi(s)\|^2_{\gamma(H,X)} \,\mathrm{d}s \,\mathrm{d}t = \sum_{\ell=1}^{2^n} \int_{I^n_\ell} \|\Phi(s)\|^2_{\gamma(H,X)} \,\mathrm{d}s = \|\Phi\|_2^2.
\end{aligned}$$

Let $\varepsilon > 0$ and $m \in \mathbb{N}$. Analogously to part (a), we can choose $\tilde{\psi}, \psi \in C([0,T];\gamma(H,X))$ such that $\|L_n(\phi^m) - \tilde{\psi}\|_2 \leq \varepsilon$ and $\|\psi - \phi^m\|_2 \leq \varepsilon$. For sufficiently large n such that $\|\psi(t) - \psi(s)\|_{\gamma(H,X)} \leq \frac{\varepsilon}{\sqrt{2T}}$ for $|t-s| \leq \frac{T}{2^n}$ as well as $\frac{T}{2^n} \leq \frac{\varepsilon^2}{2\|\psi\|^2_\infty}$, the translation error of ψ can be bounded by

$$\begin{aligned}
\|\psi - L_n(\psi)\|_2^2 &= \int_0^{\frac{T}{2^n}} \|\psi(t)\|^2_{\gamma(H,X)} \,\mathrm{d}t + \int_{\frac{T}{2^n}}^T \Big\|\psi(t) - \psi\Big(t - \frac{T}{2^n}\Big)\Big\|^2_{\gamma(H,X)} \,\mathrm{d}t \\
&\leq \|\psi\|^2_\infty \cdot \frac{T}{2^n} + \frac{\varepsilon^2}{2T}\Big(T - \frac{T}{2^n}\Big) \leq \frac{\varepsilon^2}{2} + \frac{\varepsilon^2}{2} = \varepsilon^2.
\end{aligned}$$

Choosing n large enough such that also $\|\tilde{\psi}(t) - \tilde{\psi}(s)\|_{\gamma(H,X)} \leq \frac{\varepsilon}{\sqrt{T}}$ for $|t-s| \leq \frac{T}{2^n}$, we obtain

$$\|\tilde{\psi} - \mathbb{E}(\tilde{\psi}|D_n)\|_2^2 \leq \sum_{\ell=1}^{2^n} \int_{I^n_\ell} \frac{1}{|I^n_\ell|^2} \Big(\int_{I^n_\ell} \big\|\tilde{\psi}(t) - \tilde{\psi}(s)\big\|_{\gamma(H,X)} \,\mathrm{d}s\Big)^2 \,\mathrm{d}t$$

$$
\begin{aligned}
&\leq \sum_{\ell=1}^{2^n} \int_{I_\ell^n} \frac{1}{|I_\ell^n|} \int_{I_\ell^n} \|\tilde{\psi}(t) - \tilde{\psi}(s)\|_{\gamma(H,X)}^2 \, \mathrm{d}s \, \mathrm{d}t \\
&\leq \sum_{\ell=1}^{2^n} \frac{\varepsilon^2}{T} |I_\ell^n| \int_{I_\ell^n} \frac{1}{|I_\ell^n|} \, \mathrm{d}t = \frac{\varepsilon^2}{T} \sum_{\ell=1}^{2^n} |I_\ell^n| = \varepsilon^2.
\end{aligned}
$$

Altogether, this yields

$$
\begin{aligned}
\|\phi^m - \phi_n^m\|_2 &\leq \|\phi^m - \psi\|_2 + \|\psi - L_n(\psi)\|_2 + \|L_n(\psi) - L_n(\phi^m)\|_2 \\
&\quad + \|L_n(\phi^m) - \tilde{\psi}\|_2 + \|\tilde{\psi} - \mathbb{E}(\tilde{\psi}|D_n)\|_2 + \|\mathbb{E}(\tilde{\psi}|D_n) - \mathbb{E}(L_n(\phi^m)|D_n)\|_2 \\
&\leq 4\varepsilon + \|\phi^m - \psi\|_2 + \|L_n(\phi^m) - \tilde{\psi}\|_2 \leq 6\varepsilon
\end{aligned}
$$

and thus $\phi_n^m(\omega, \cdot) \to \phi^m(\omega, \cdot)$ in $L^2(0, T; \gamma(H, X))$ for almost every $\omega \in \Omega$. Combining this observation with (4.2.7), we conclude the convergence $\phi_n^m \to \phi$ in $L^0_{\mathcal{P}}(\Omega; L^2(0, T; \gamma(H, X))$ as $m, n \to \infty$. □

The following assumption ensures well-posedness of the stochastic evolution equation (4.2.1), as we will show in Theorem 4.12.

Assumption 4.11. *Let X be a $(2, D)$-smooth Banach space and let $p \in [1, \infty)$. Let $F\colon \Omega \times [0, T] \times X \to X$ and $G\colon \Omega \times [0, T] \times X \to \gamma(H, X)$ be strongly $\mathcal{P} \otimes \mathcal{B}(X)$-measurable, where $F = \tilde{F} + f$ and $G = \tilde{G} + g$ for $f\colon \Omega \times [0, T] \to X$ and $g\colon \Omega \times [0, T] \to \gamma(H, X)$. Suppose that $\tilde{F}(\cdot, \cdot, 0) = 0$, $\tilde{G}(\cdot, \cdot, 0) = 0$, and*

(a) (global Lipschitz continuity on X) *there exist constants $C_{F,X}, C_{G,X} \geq 0$ such that for all $\omega \in \Omega, t \in [0, T]$ and $x \in X$, it holds that*

$$
\|\tilde{F}(\omega, t, x) - \tilde{F}(\omega, t, y)\| \leq C_{F,X}\|x - y\|, \; \|\tilde{G}(\omega, t, x) - \tilde{G}(\omega, t, y)\| \leq C_{G,X}\|x - y\|,
$$

(b) (integrability) *$f \in L^p_{\mathcal{P}}(\Omega; L^1(0, T; X))$, and $g \in L^p_{\mathcal{P}}(\Omega; L^2(0, T; \gamma(H, X)))$.*

We would like to point out that $\tilde{F}$ and f act on different sets: while f depends only on the random input from Ω and time, $\tilde{F}$ also depends on space. This setting allows us to include temporally less smooth nonlinearities. The same applies to $\tilde{G}$ and g. Moreover, strong measurability is to be understood in the stochastic sense, see Remark 2.43.

Combining the progressive measurability and Lipschitz assumptions on $\tilde{F}$ and $\tilde{G}$, we conclude $\tilde{F}(\cdot, \cdot, x) \in L^0_{\mathcal{P}}(\Omega; C([0, T]; X))$ and $\tilde{G}(\cdot, \cdot, x) \in L^0_{\mathcal{P}}(\Omega; C([0, T]; \gamma(H, X)))$ for all $x \in X$. Further note that Assumption 4.11 implies linear growth of F and G on X, i.e., for all $\omega \in \Omega$, $t \in [0, T]$, and $x \in X$,

$$
\|\tilde{F}(\omega, t, x)\| \leq C_{F,X}(1 + \|x\|) \quad \text{and} \quad \|\tilde{G}(\omega, t, x)\|_{\gamma(H,X)} \leq C_{G,X}(1 + \|x\|),
$$

where the constant 1 can be left out, but is included for later use in Theorem 4.14.

The index X of the constants will be omitted whenever the underlying space is clear from the context. Furthermore, for the sake of readability, dependence of u, F, and G on ω will not be explicitly stated.

We now proceed to show well-posedness of the stochastic evolution equation (4.2.1) in the sense of existence and uniqueness of mild solutions (4.2.2). The following well-posedness result is more or less standard [39, Chapters 6,7].

Theorem 4.12. *Suppose that Assumption 4.11 holds for some* $p \in [1,\infty)$ *and assume that* $-A$ *generates a* C_0*-contraction semigroup* $(S(t))_{t\geq 0}$ *on* X*. Let* $u_0 \in L^p_{\mathcal{F}_0}(\Omega; X)$*. Then* (4.2.1) *has a unique mild solution in* $L^p_{\mathcal{P}}(\Omega; C([0,T]; X))$*.*

Proof. Let $\delta \in (0,T]$ to be chosen explicitly later. We show local existence and uniqueness of solutions first and conclude global existence and uniqueness second by concatenating the local solutions. Define the spaces $Z_\delta := L^p(\Omega; C([0,\delta]; X))$, $Z := Z_T$, $Z^{\mathcal{P}}_\delta$ as the subset of all adapted $v \in Z_\delta$ and $Z^{\mathcal{P}} := Z^{\mathcal{P}}_T$. For $v \in Z^{\mathcal{P}}_\delta$, we define the fixed point functional

$$\Gamma v(t) := S(t)u_0 + \int_0^t S(t-r)F(r,v(r))\,\mathrm{d}r + \int_0^t S(t-r)G(r,v(r))\,\mathrm{d}W_H(r). \tag{4.2.8}$$

The problem of finding local mild solutions of (4.2.1) then reduces to finding fixed points $v \in Z^{\mathcal{P}}_\delta$ of Γ. The contraction mapping theorem yields the existence of such unique fixed points, provided that Γ is a contraction that maps $Z^{\mathcal{P}}$ and thus $Z^{\mathcal{P}}_\delta$ into itself. That is, we have to prove *(i)* continuity of paths of Γv via maximal estimates for $v \in Z^{\mathcal{P}}_\delta$ as well as *(ii)* adaptedness of Γv, and that *(iii)* Γ is a (strict) contraction on $Z^{\mathcal{P}}_\delta$. Lastly, we consider the evolution equation on $[\delta, 2\delta]$ with initial value $U(\delta)$ and so forth to extend the solution to larger time intervals in part *(iv)*.

(i) Let $v \in Z^{\mathcal{P}}_\delta$ and fix $\omega \in \Omega$. Consider each term of Γv individually. Strong continuity of S implies continuity of $t \mapsto S(t)u_0$.
For the second term, let w.l.o.g. $s,t \in [0,\delta]$ with $s < t$. Then

$$\begin{aligned}
&\Big\| \int_0^t S(t-r)F(r,v(r))\,\mathrm{d}r - \int_0^s S(s-r)F(r,v(r))\,\mathrm{d}r \Big\| \\
&\leq \int_0^s \|S(s-r)\| \big\| [S(t-s)-I]F(r,v(r)) \big\|\,\mathrm{d}r + \int_s^t \|S(t-r)\| \|F(r,v(r))\|\,\mathrm{d}r \\
&\leq \int_0^s \|[S(t-s)-I]F(r,v(r))\|\,\mathrm{d}r + \int_s^t \|F(r,v(r))\|\,\mathrm{d}r.
\end{aligned} \tag{4.2.9}$$

By strong continuity of S and the assumptions imposed on F, the integrand of the first term vanishes as $s \to t$. Furthermore, the integrands satisfy

$$\begin{aligned}
&\|[S(t-s)-I]F(\cdot,v(\cdot))\|_{L^1(0,\delta;X)} \leq 2\|F(\cdot,v(\cdot))\|_{L^1(0,\delta;X)} \\
&\leq 2(L_{F,X}\delta(1+\|v\|_{\infty,[0,\delta]}) + \|f\|_{L^1(0,\delta;X)}) < \infty
\end{aligned} \tag{4.2.10}$$

uniformly in s. Thus, by dominated convergence, the first summand in (4.2.9) vanishes almost surely as $s \to t$. Moreover, the second summand in (4.2.9) satisfies

$$\begin{aligned}
\int_s^t \|F(r,v(r))\|\,\mathrm{d}r &\leq L_{F,X}\int_s^t (1+\|v(r)\|_X)\,\mathrm{d}r + \int_s^t \|f(r)\|\,\mathrm{d}r \\
&\leq L_{F,X}(t-s)(1+\|v\|_{\infty,[0,\delta]}) + \|f\|_{L^1(s,t;X)},
\end{aligned}$$

which implies that (4.2.9) vanishes as $s \to t$. Hence, the deterministic convolution part in (4.2.8) almost surely has continuous paths.
The existence of a continuous modification of $(\int_0^t S(t-s)G(s,v(s))\,\mathrm{d}W_H(s))_{t\in[0,\delta]}$ is an

immediate consequence of Theorem 2.61 provided that $[\Omega\times[0,\delta] \ni (\omega,s) \mapsto G(\omega,s,v(s)) \in \gamma(H,X)] \in L^0_{\mathcal{P}}(\Omega; L^2(0,\delta;\gamma(H,X)))$. Since Lipschitz continuity of $\tilde{G}$ implies

$$\int_0^\delta \|G(s,v(s))\|^2_{\gamma(H,X)}\,\mathrm{d}s \le C^2_{G,X}\delta\|v\|^2_{\infty,[0,\delta]} + \|g\|^2_{L^2(0,\delta;\gamma(H,X))} < \infty$$

almost surely, this statement follows from Assumption 4.11. Hence, the stochastic convolution term in (4.2.8) and thus Γv have continuous paths almost surely.

(ii) Let $v \in Z^{\mathcal{P}}$ and fix $t \in [0,\delta]$. We show strong $\mathscr{F}_t$-measurability of all three terms in (4.2.8) separately. Since u_0 is $\mathscr{F}_0$-measurable by assumption and thus $\mathscr{F}_t$-measurable, so is $S(t)u_0$.
For $\phi \in L^0_{\mathcal{P}}(\Omega; L^1(0,\delta;X))$, Lemma 4.10(a) implies that there are $\phi_n \in \mathcal{E}_1$ such that $\phi_n \to \phi$ in $L^0_{\mathcal{P}}(\Omega;L^1(0,\delta;X))$. Thus, $\int_0^t \phi_n(s)\,\mathrm{d}s \to \int_0^t \phi(s)\,\mathrm{d}s$ almost surely. Let I^n_k, $\lfloor t\rfloor$, $\lceil t\rceil$ as defined in the proof of Lemma 4.10 with T replaced by δ. Then $\int_0^{\lfloor t\rfloor}\phi_n(s)\,\mathrm{d}s = \sum_{\ell=1}^{2^n}\mathbf{1}_{I^n_\ell}(t)\sum_{k=1}^{\ell-1}\phi_n\big(\frac{k\delta}{2^n}\big)$ is $\mathscr{F}_t$-simple and

$$\Big\|\int_0^t \phi(s)\,\mathrm{d}s - \int_0^{\lfloor t\rfloor}\phi_n(s)\,\mathrm{d}s\Big\|_X \le \int_0^{\lfloor t\rfloor}\|\phi_n(s)-\phi(s)\|_X\,\mathrm{d}s + \int_{\lfloor t\rfloor}^t \|\phi(s)\|_X\,\mathrm{d}s \to 0$$

as $n\to\infty$. Hence, $\int_0^t\phi(s)\,\mathrm{d}s$ is $\mathscr{F}_t$-measurable. Next, we show that $\phi := S(t-\cdot)\psi \in L^0_{\mathcal{P}}(\Omega;L^1(0,t;X))$ for $\psi \in L^0_{\mathcal{P}}(\Omega;L^1(0,t;X))$. By contractivity of S, the assumption on ψ implies $S(t-\cdot)\psi(\omega) \in L^1(0,t;X)$ for almost every $\omega\in\Omega$. For suitable adapted step function approximations $\xi_n = \sum_{\ell=1}^L \mathbf{1}_{(t^n_{\ell-1},t^n_\ell]}\zeta^n_\ell \in \mathcal{E}_1$ of ψ chosen according to Lemma 4.10(a), it holds that $\sum_{\ell=1}^L \mathbf{1}_{(t^n_{\ell-1},t^n_\ell]}S(t-t^n_\ell)\zeta^n_\ell \in \mathcal{E}_1$ and

$$\begin{aligned}&\Big\|\sum_{\ell=1}^L \mathbf{1}_{(t^n_{\ell-1},t^n_\ell]}(s)S(t-t^n_\ell)\zeta^n_\ell - S(t-s)\psi(s)\Big\|_X\\ &\le \sum_{\ell=1}^L \big\|\big(\mathbf{1}_{(t^n_{\ell-1},t^n_\ell]}(s)S(t-t^n_\ell)-S(t-s)\big)\zeta^n_\ell\big\|_X + \|S(t-s)\|\Big\|\sum_{\ell=1}^L\zeta^n_\ell\mathbf{1}_{(t^n_{\ell-1},t^n_\ell]}(s)-\psi(s)\Big\|_X\end{aligned}$$

vanishes as $n\to\infty$ by strong continuity of S and Lemma 4.10(a). Hence, $S(t-\cdot)\psi$ is adapted and thus $S(t-\cdot)\psi \in L^0_{\mathcal{P}}(\Omega;L^1(0,t;X))$.
Lastly, choosing $\psi = [\Omega\times[0,t] \ni (\omega,s)\mapsto F(\omega,s,v(\omega,s)) \in X]$ in the last statement yields adaptedness of the deterministic convolution part of Γv. To do so, it remains to show that $\psi = [\Omega\times[0,\delta]\ni(\omega,s)\mapsto F(\omega,s,v(\omega,s))\in X] \in L^0_{\mathcal{P}}(\Omega;L^1(0,\delta;X))$. Since $f \in L^p_{\mathcal{P}}(\Omega;L^1(0,T;X))$ is progressively measurable by assumption, we prove this for $\tilde{F}(\omega,s,v(\omega,s))$ only. Pathwise absolute integrability was already shown in (4.2.10). Progressive measurability of ψ follows from adaptedness of v, progressive measurability of $\tilde{F}(\cdot,\cdot,x)$ for fixed $x\in X$ and the condition $\tilde{F}(\cdot,\cdot,0)=0$. More precisely, by adaptedness of v we can approximate it by a linear combination of $\mathcal{F}_t$-simple functions $\mathbf{1}_{(t^1,t^2]}\mathbf{1}_A\tilde{x}$ for $t^1<t^2$ in $[0,t]$ for fixed $t\in[0,\delta]$, $A\in\mathscr{F}_t$ and $\tilde{x}\in X$. Since $\tilde{F}(\omega,s,\mathbf{1}_{(t^1,t^2]}(s)\mathbf{1}_A(\omega)\tilde{x}) = \tilde{F}(\omega,s,0)=0$ for $s\notin(t^1,t^2]$ or $\omega\notin A$ by Assumption 4.11, we can reduce to the case $s\in(t^1,t^2]$ and $\omega\in A$. However, in this case the assumed progressive measurability of $\tilde{F}(\cdot,\cdot,x)$ for fixed $x\in X$ implies $\tilde{F}(\omega,s,\mathbf{1}_{(t^1,t^2]}(s)\mathbf{1}_A(\omega)\tilde{x}) = \tilde{F}(\omega,s,\tilde{x}) \in \mathcal{F}_t\times\mathscr{B}([0,t])$.

Progressive measurability of ψ follows by approximation because $\tilde{F}$ is (Lipschitz) continuous in the third argument.
Adaptedness of the stochastic convolution remains to be shown. To this end, let $\phi \in L^0_{\mathcal{P}}(\Omega; L^2(0,\delta;\gamma(H,X)))$. With $\phi^m, \phi^m_n \in L^2_{\mathcal{P}}(\Omega; L^2(0,\delta;\gamma(H,X)))$ and stopping time τ_m as in the proof of Lemma 4.10(b), an application of the maximal inequality from Theorem 2.61 yields

$$\begin{aligned}
&\Big\| \int_0^{(\cdot)} (\phi - \phi^m_n) \,\mathrm{d}W_H \Big\|^2_{L^2(\Omega; L^\infty(0,\delta;\gamma(H,X)))} \\
&\leq \mathbb{E} \sup_{t\in[0,\delta]} \Big\| \int_0^t (\phi - \phi^m) \,\mathrm{d}W_H \Big\|^2_{\gamma(H,X)} + \mathbb{E} \sup_{t\in[0,\delta]} \Big\| \int_0^t (\phi^m - \phi^m_n) \,\mathrm{d}W_H \Big\|^2_{\gamma(H,X)} \\
&\leq \mathbb{E} \sup_{t\in[0,\delta]} \Big\| \int_{\tau_m}^{t\vee\tau_m} \phi \,\mathrm{d}W_H \Big\|^2_{\gamma(H,X)} + C^2_{2,D} \mathbb{E} \|\phi^m - \phi^m_n\|^2_{L^2(0,\delta;\gamma(H,X))} \to 0 \qquad (4.2.11)
\end{aligned}$$

as $m, n \to \infty$. Note that, by definition of τ_m, choosing m sufficiently large yields $\tau_m \geq t$, so that the first term in (4.2.11) vanishes. We conclude $\int_0^{(\cdot)} \phi^m_n \,\mathrm{d}W_H \to \int_0^{(\cdot)} \phi \,\mathrm{d}W_H$ in $L^2(\Omega; L^\infty([0,\delta];\gamma(H,X)))$. In particular, the stochastic integrals almost surely converge pointwise in $[0,\delta]$. Because $\phi^m_n \in \mathcal{E}_2$, we can write $\phi^m_n = \sum_{\ell=1}^L \mathbf{1}_{(t^{nm}_{\ell-1}, t^{nm}_\ell]} \zeta^{nm}_\ell$ for some $L \in \mathbb{N}$, a partition $\{t^{nm}_\ell : \ \ell = 1, \dots, L\}$ of $[0,\delta]$ and $\mathscr{F}_{t^{nm}_{\ell-1}}$-simple $\zeta^{nm}_\ell \colon \Omega \to \gamma(H,X)$. Then

$$t \mapsto \int_0^{\lfloor t \rfloor} \phi^m_n(s) \,\mathrm{d}W_H(s) = \sum_{\ell=1}^L \mathbf{1}_{(t^{nm}_{\ell-1}, t^{nm}_\ell]}(t) \sum_{k=1}^{\ell-1} (W_H(t^{nm}_k) - W_H(t^{nm}_{k-1})) \zeta^{nm}_\ell$$

is an adapted step process and

$$\begin{aligned}
&\Big\| \int_0^t \phi(s) \,\mathrm{d}W_H(s) - \int_0^{\lfloor t \rfloor} \phi^m_n(s) \,\mathrm{d}W_H(s) \Big\|_{\gamma(H,X)} \\
&\leq \Big\| \int_0^t (\phi(s) - \phi^m_n(s)) \,\mathrm{d}W_H(s) \Big\|_{\gamma(H,X)} + \Big\| \int_{\lfloor t \rfloor}^t \phi^m_n(s) \,\mathrm{d}W_H(s) \Big\|_{\gamma(H,X)} \to 0
\end{aligned}$$

as $n \to \infty$. Hence, $\int_0^t \phi(s) \,\mathrm{d}W_H(s)$ is $\mathscr{F}_t$-measurable. Arguing as for the second term, we deduce adaptedness of $\phi = S(t - \cdot)\psi$ for $\psi = [\Omega \times [0,\delta] \ni (\omega, s) \mapsto G(\omega, s, v(\omega, s)) \in \gamma(H,X)] \in L^0_{\mathcal{P}}(\Omega; L^2(0,\delta;\gamma(H,X)))$ from progressive measurability of g and $\tilde{G}(\cdot,\cdot,x)$ for fixed $x \in X$, adaptedness of v, (Lipschitz) continuity of $\tilde{G}$ in the third argument and the condition $\tilde{G}(\cdot,\cdot,0) = 0$ from Assumption 4.11. In conclusion, Γv is adapted and thus $\Gamma v \in Z^{\mathcal{P}}_\delta$ follows from continuity of paths, as shown in part (i).

(iii) Let $v_1, v_2 \in Z^{\mathcal{P}}_\delta$ and $\delta \in (0, \delta_0]$ for $\delta_0 \in (0, T]$ to be fixed later. Lipschitz continuity of $\tilde{F}$ and $\tilde{G}$, contractivity of S, and the maximal inequality from Theorem 2.61 yield

$$\begin{aligned}
\|\Gamma v_1 - \Gamma v_2\|_{Z^{\mathcal{P}}_\delta} \leq &\Big\| \sup_{t\in[0,\delta]} \Big\| \int_0^t S(t-s)[\tilde{F}(s, v_1(s)) - \tilde{F}(s, v_2(s))] \,\mathrm{d}s \Big\| \Big\|_p \\
&+ \Big\| \sup_{t\in[0,\delta]} \Big\| \int_0^t S(t-s)[\tilde{G}(s, v_1(s)) - \tilde{G}(s, v_2(s))] \,\mathrm{d}W_H(s) \Big\| \Big\|_p
\end{aligned}$$

$$\leq C_{F,X}\Big\| \sup_{t\in[0,\delta]} \int_0^t \|v_1(s)-v_2(s)\|\,\mathrm{d}s\Big\|_p + C_{p,D}C_{G,X}\Big\|\Big(\int_0^\delta \|v_1(s)-v_2(s)\|^2\,\mathrm{d}s\Big)^{1/2}\Big\|_p$$
$$\leq (C_{F,X}\delta + C_{p,D}C_{G,X}\delta^{1/2})\Big\| \sup_{t\in[0,\delta]} \|v_1(t)-v_2(t)\|\Big\|_p = C_\delta \|v_1-v_2\|_{Z_\delta^{\mathcal{P}}}$$

for $C_\delta \coloneqq C_{F,X}\delta + C_{p,D}C_{G,X}\delta^{1/2}$. If $\delta \in (0,\delta_0]$ is sufficiently small, it holds that $C_\delta \leq \frac{1}{2}$. Choosing $\delta_0 > 0$ in such a way, the above calculation shows that Γ is a (1/2-)contraction on $Z_\delta^{\mathcal{P}}$ for all $\delta \in (0,\delta_0]$. Hence, the contraction mapping theorem is applicable and yields the desired well-posedness statement on $[0,\delta]$.

(iv) Lastly, we conclude global well-posedness by concatenating local solutions. Denote the local solution to (4.2.1) on $[0,\delta]$ by u_1. Further, consider the stochastic evolution equation

$$\begin{cases} \mathrm{d}u_2 + Au_2\,\mathrm{d}t = F(t,u_2)\,\mathrm{d}t + G(t,u_2)\,\mathrm{d}W_H \quad \text{on } (\delta,2\delta], \\ u_2(\delta) = u_1(\delta) \in L^p(\Omega;X), \end{cases}$$

which by the same argument has a unique mild solution

$$u_2(t) = S(t-\delta)u_1(\delta) + \int_\delta^t S(t-s)F(s,u_2(s))\,\mathrm{d}s + \int_\delta^t S(t-s)G(s,u_2(s))\,\mathrm{d}W_H(s)$$

for $t \in (\delta,2\delta]$. Inserting the formula of the mild solution u_1 at time δ and making use of $S(t-\delta)S(\delta-s) = S(t-s)$, we obtain

$$u_2(t) = S(t)u_0 + \int_0^t S(t-s)F(s,u_2(s))\,\mathrm{d}s + \int_0^t S(t-s)G(s,u_2(s))\,\mathrm{d}W_H(s)$$

for $t \in (\delta,2\delta]$. Setting $u \coloneqq u_1$ on $[0,\delta]$ and $u \coloneqq u_2$ on $(\delta,2\delta]$ yields a mild solution on $[0,2\delta]$. Proceeding analogously, unique mild solutions on $[0,T]$ are obtained. □

In addition to existence and uniqueness, linear growth of the mild solution with respect to the initial data, f, and g can be shown, as the following a priori estimate illustrates. We draw attention to the fact that this linear growth is only valid for higher moments of the mild solution, that is, for $p \in [2,\infty)$. Indeed, the linear growth estimate does not hold in general for $p < 2$ [98, Rem. 4.3]. As a shorthand notation, let

$$\|f\|_{p,q,Z} \coloneqq \|f\|_{L^p(\Omega;L^q(0,T;Z))}, \quad \|\!|g|\!\|_{p,q,Z} \coloneqq \|g\|_{L^p(\Omega;L^q(0,T;\mathcal{L}_2(H,Z)))} \tag{4.2.12}$$

for $p \in [2,\infty)$, $q \in [1,\infty]$, and Banach spaces Z. The notation $\|\!|\cdot|\!\|$ is consistent with its use in Section 4.1 and will be used throughout this work.

Proposition 4.13 (A priori estimate)**.** *Suppose that Assumption 4.11 holds for some $p \in [2,\infty)$ and assume that $-A$ generates a C_0-contraction semigroup $(S(t))_{t\geq 0}$ on X. Let $u_0 \in L^p_{\mathcal{F}_0}(\Omega;X)$. Then*

$$\|U\|_{L^p(\Omega;C([0,T];X))} \leq C_{bdd}^X\big(1 + \|u_0\|_{L^p(\Omega;X)} + \|f\|_{p,1,X} + C_{p,D}\|\!|g|\!\|_{p,2,X}\big),$$

where $C_{bdd}^X \coloneqq (1+C^2T)^{1/2}\mathrm{e}^{(1+C^2T)/2}$ with $C \coloneqq C_{F,X}T^{1/2} + C_{p,D}C_{G,X}$, and $C_{p,D}$ is the constant from Theorem 2.61.

As this will be of interest for linear growth estimates on different spaces, we point out that Lipschitz continuity of $\tilde{F}$ or $\tilde{G}$ is not required for this statement. Linear growth is sufficient, as the following proof illustrates.

Proof. Let $r \in [0,T]$ and $\psi(r) := 1 + \| \sup_{t\in[0,r]} \|U(t)\| \|_p$. From the triangle inequality, Theorem 2.61, linear growth on X, and the Cauchy–Schwarz inequality, we conclude

$$\begin{aligned}\psi(r) &\le 1 + \|u_0\|_{L^p(\Omega;X)} + \Big\| \sup_{t\in[0,r]} \int_0^t \|\tilde{F}(s,U(s)) + f(s)\| \,\mathrm{d}s \Big\|_p \\ &\quad + \Big\| \sup_{t\in[0,r]} \Big\| \int_0^t S(t-s)(\tilde{G}(s,U(s)) + g(s)) \,\mathrm{d}W_H(s) \Big\| \Big\|_p \\ &\le 1 + \|u_0\|_{L^p(\Omega;X)} + C_{F,X} \Big\| \int_0^r 1 + \|U(s)\| \,\mathrm{d}s \Big\|_p + \|f\|_{L^p(\Omega;L^1(0,r;X))} \\ &\quad + C_{p,D}\Big[C_{G,X} \Big\| \Big(\int_0^r (1+\|U(s)\|)^2 \,\mathrm{d}s \Big)^{1/2} \Big\|_p \mathrm{d}s + \|g\|_{L^p(\Omega;L^2(0,r;\gamma(H,X)))} \Big] \\ &\le c_{u_0,f,g} + C_{F,X} \int_0^r \psi(s)\,\mathrm{d}s + C_{p,D} C_{G,X} \Big(\int_0^r \psi(s)^2 \,\mathrm{d}s \Big)^{1/2} \\ &\le c_{u_0,f,g} + C \Big(\int_0^r \psi(s)^2 \,\mathrm{d}s \Big)^{1/2},\end{aligned}$$

where $c_{u_0,f,g} = 1 + \|u_0\|_{L^p(\Omega;X)} + \|f\|_{p,1,X} + C_{p,D} \|\!|g|\!\|_{p,2,X}$ and $C = C_{F,X} T^{1/2} + C_{p,D} C_{G,X}$. Here, we used Minkowski's inequality to pull in the $L^p(\Omega)$ and $L^{p/2}(\Omega)$ norms after rewriting $\|(\cdot)^{1/2}\|_p = \|\cdot\|_{p/2}^{1/2}$. Ultimately, the version of Gronwall's inequality from Lemma 2.67 yields the desired result

$$\psi(T) \le c_{u_0,f,g} (1 + C^2 T)^{1/2} \mathrm{e}^{(1+C^2T)/2}. \qquad \square$$

Lastly, we present a well-posedness result on subspaces $Y \hookrightarrow X$, which does not require Lipschitz continuity of $\tilde{F}$ or $\tilde{G}$ on Y but merely linear growth on this subspace. The reader is referred to Remark 4.15 below for a discussion where we explain why Lipschitz continuity on Y should be avoided.

Theorem 4.14. *Suppose that Assumption 4.11 holds. Let $Y \hookrightarrow X$ be a $(2,D)$-smooth Banach space, $u_0 \in L^p_{\mathcal{F}_0}(\Omega;Y)$, and assume that $-A$ generates a C_0-contraction semigroup $(S(t))_{t\ge 0}$ on both X and Y. Additionally, suppose that $f \in L^p_{\mathcal{P}}(\Omega; L^1(0,T;Y))$, $g \in L^p_{\mathcal{P}}(\Omega; L^2(0,T;\gamma(H,Y)))$ and $F\colon \Omega\times[0,T]\times Y \to Y$, $G\colon \Omega\times[0,T]\times Y \to \gamma(H,Y)$ are strongly $\mathcal{P}\otimes\mathcal{B}(Y)$-measurable, and there are $L_{F,Y}, L_{G,Y} \ge 0$ such that for all $\omega\in\Omega$, $t\in[0,T]$, and $x \in Y$,*

$$\|\tilde{F}(\omega,t,x)\|_Y \le L_{F,Y}(1+\|x\|_Y), \;\; \|\tilde{G}(\omega,t,x)\|_{\gamma(H,Y)} \le L_{F,Y}(1+\|x\|_Y).$$

Under these conditions, the unique mild solution $U \in L^p(\Omega; C([0,T];X))$ to (4.2.1) is in $L^p(\Omega;C([0,T];Y))$ and

$$\|U\|_{L^p(\Omega;C([0,T];Y))} \le C^Y_{bdd}\big(1 + \|u_0\|_{L^p(\Omega;Y)} + \|f\|_{p,1,Y} + C_{p,D}\|\!|g|\!\|_{p,2,Y}\big),$$

where $C^Y_{bdd} := (1+C^2T)^{1/2}\mathrm{e}^{(1+C^2T)/2}$ with $C := L_{F,Y}T^{1/2} + C_{p,D}L_{G,Y}$, and $C_{p,D}$ is the constant from Theorem 2.61.

The constant C appears exponentially in the above. In the special case $p = 2$, $L_{F,Y} = L_{G,Y} = T = 1$, and X Hilbert, this leads to $C^Y_{\mathrm{bdd}} \leq \sqrt{10}\mathrm{e}^5 \leq 470$.

Proof. Recall that by the contraction mapping theorem for $\delta \leq T_0$ and some $T_0 \in (0,1]$ only depending on p, $C_{F,X}$, $C_{G,X}$, and X, one has $U = \lim_{n\to\infty} U_n$ in $L^p(\Omega; C([0,\delta]; X))$, where $U_0 = u_0$ and $U_{n+1} = \Gamma(U_n)$ with Γ as defined in (4.2.8). Since F and G map Y into Y and $\gamma(H,Y)$, respectively, we can also consider Γ as a mapping on $Z^2 := L^p_{\mathcal{P}}(\Omega; L^2(0,\delta;Y))$ to eventually show that U is in $L^p_{\mathcal{P}}(\Omega; C([0,\delta];Y)) \subseteq Z^2$. Note that for $U \in Z^2$, $F(\cdot,U)$ and $G(\cdot,U)$ are progressively measurable as Y- and $\gamma(H,Y)$-valued mappings, respectively [71, Thm. 1.1.6]. Moreover, we claim that for all $v \in Z^2$,

$$\begin{aligned}\|\Gamma(v)\|_{L^p(\Omega;C([0,\delta];Y))} &\leq \|u_0\|_{L^p(\Omega;Y)} + \|f\|_{L^p(\Omega;L^1(0,\delta;Y))} \\ &\quad + C_{p,D}\|g\|_{L^p(\Omega;L^2(0,\delta;\gamma(H,Y)))} + \big(L_{F,Y} + C_{p,D}L_{G,Y}\big)(1 + \|v\|_{Z^2}).\end{aligned} \tag{4.2.13}$$

Indeed, since S is contractive on Y, a combination of the maximal inequality, linear growth of $\tilde{F}$ and $\tilde{G}$ on Y, $\delta \leq 1$, and the Cauchy–Schwarz inequality implies

$$\begin{aligned}&\|\Gamma(v) - S(\cdot)u_0\|_{L^p(\Omega;C([0,\delta];Y))} \leq \|F(\cdot,v)\|_{L^p(\Omega;L^1(0,\delta;Y))} + C_{p,D}\|G(\cdot,v)\|_{L^p(\Omega;L^2(0,\delta;\gamma(H,Y)))} \\ &\quad\leq \|f\|_{L^p(\Omega;L^1(0,\delta;Y))} + L_{F,Y}\big(\delta + \|v\|_{L^p(\Omega;L^1(0,\delta;Y))}\big) \\ &\qquad + C_{p,D}\big(\|g\|_{L^p(\Omega;L^2(0,\delta;\gamma(H,Y)))} + L_{G,Y}\big(\sqrt{\delta} + \|v\|_{L^p(\Omega;L^2(0,\delta;Y))}\big)\big) \\ &\quad\leq \|f\|_{L^p(\Omega;L^1(0,\delta;Y))} + C_{p,D}\|g\|_{L^p(\Omega;L^2(0,\delta;\gamma(H,Y)))} + \big(L_{F,Y} + C_{p,D}L_{G,Y}\big)(1 + \|v\|_{Z^2}).\end{aligned}$$

Therefore, (4.2.13) follows, which in turn implies

$$\begin{aligned}\|\Gamma(v)\|_{Z^2} &\leq \delta^{1/2}\|\Gamma(v)\|_{L^p(\Omega;C([0,\delta];Y))} \\ &\leq \theta(1 + \|u_0\|_{L^p(\Omega;Y)} + \|f\|_{L^p(\Omega;L^1(0,\delta;Y))} + \|g\|_{L^p(\Omega;L^2(0,\delta;\gamma(H,Y)))} + \|v\|_{Z^2})\end{aligned}$$

for $\theta = \delta^{1/2}\max\{1, C_{p,D}, L_{F,Y} + C_{p,D}L_{G,Y}\}$. Choosing $\delta \in (0,T_0]$ such that $\theta \leq \frac{1}{2}$, we iteratively obtain for $n \geq 1$

$$\begin{aligned}&\|U_n\|_{Z^2} \leq \theta(1 + \|u_0\|_{L^p(\Omega;Y)} + \|f\|_{L^p(\Omega;L^1(0,\delta;Y))} + \|g\|_{L^p(\Omega;L^2(0,\delta;\gamma(H,Y)))}\big) + \theta\|U_{n-1}\|_{Z^2} \\ &\quad\leq \theta(1 + \|u_0\|_{L^p(\Omega;Y)} + \|f\|_{L^p(\Omega;L^1(0,\delta;Y))} + \|g\|_{L^p(\Omega;L^2(0,\delta;\gamma(H,Y)))}\big) \\ &\qquad + \theta^2(1 + \|u_0\|_{L^p(\Omega;Y)} + \|f\|_{L^p(\Omega;L^1(0,\delta;Y))} + \|g\|_{L^p(\Omega;L^2(0,\delta;\gamma(H,Y)))} + \|U_{n-2}\|_{Z^2}) \\ &\quad\leq \ldots \leq \sum_{j=1}^n \theta^j(1 + \|u_0\|_{L^p(\Omega;Y)} + \|f\|_{L^p(\Omega;L^1(0,\delta;Y))} + \|g\|_{L^p(\Omega;L^2(0,\delta;\gamma(H,Y)))}\big) + \theta^n\|U_0\|_{Z^2} \\ &\quad\leq 1 + \|f\|_{L^p(\Omega;L^1(0,\delta;Y))} + \|g\|_{L^p(\Omega;L^2(0,\delta;\gamma(H,Y)))} + \|u_0\|_{L^p(\Omega;Y)},\end{aligned}$$

where we have used the finite geometric series formula $\sum_{j=1}^n \theta^j = \theta(1-\theta^n)(1-\theta)^{-1}$.

In conclusion, $(U_n)_{n\in\mathbb{N}}$ is bounded in Z^2. By reflexivity of Y as a 2-smooth Banach space, and thus of Z^2 by Pettis' measurability theorem [71, Cor. 1.3.22], there exist a subsequence $(U_{n_j})_{j\in\mathbb{N}}$ and $V \in Z^2$ such that $U_{n_j} \to V$ weakly in Z^2 and

$$\|V\|_{Z^2} \leq 1 + \|u_0\|_{L^p(\Omega;Y)} + \|f\|_{L^p(\Omega;L^1(0,\delta;Y))} + \|g\|_{L^p(\Omega;L^2(0,\delta;\gamma(H,Y)))}. \tag{4.2.14}$$

Since $U_n \to U$ in $L^p(\Omega; C([0,\delta];X))$, it follows that $V = U$. Since $U = \Gamma(U)$, (4.2.13) and (4.2.14) give that U is in $L^p(\Omega; C([0,\delta];Y))$. The same argument can be applied on $[j\delta, (j+1)\delta]$ using the initial value $U(j\delta) \in L^p(\Omega; Y)$ for $j = 1, 2, \ldots$ to obtain the statement on $[0, T]$.

The final a priori estimate follows as in Proposition 4.13, where we note that the Lipschitz conditions on F and G were not used in the estimate. □

Remark 4.15. In applications, one often takes $X = L^2(\mathcal{O})$ and $Y = H^1(\mathcal{O})$ with $\mathcal{O} \subseteq \mathbb{R}^d$, and F is a Nemytskij operator for a given nonlinearity $\phi : \mathbb{R} \to \mathbb{R}$, i.e. $F(x)(\xi) = \phi(x(\xi))$ for $x \in L^2(\mathcal{O})$ and $\xi \in \mathcal{O}$. Lipschitz continuity of such mappings holds for F seen as a mapping from X to X if ϕ is Lipschitz. Also, linear growth holds for F as a mapping from Y into Y if ϕ is Lipschitz. A less trivial fact is that F is continuous from Y into Y (see [122, Prop. 2.6.4]), but nothing more can be expected. For instance, Lipschitz continuity of $F : Y \to Y$ would require the estimate

$$\|\phi'(x)x' - \phi'(y)y'\|_{L^2(\mathcal{O})} \leq C\|x - y\|_{H^1(\mathcal{O})}.$$

The latter is true if and only if $\|(\phi'(x) - \phi'(y))x'\|_{L^2(\mathcal{O})} \leq \tilde{C}\|x - y\|_{H^1(\mathcal{O})}$. This cannot be expected even if $\phi \in C^\infty(\mathbb{R}^d)$ with bounded derivatives. Indeed, a product of $x - y$ and x' needs to be estimated, but this cannot be done in terms of $\|x - y\|_{H^1(\mathcal{O})}$. Similarly, problems would occur for $Y = H^\alpha(\mathcal{O})$ for other values of $\alpha > 0$. For a detailed exposition which estimates can be expected for $\phi(x) - \phi(y)$, the reader is referred to [122, Section 2.7].

4.3 Stability

Before analysing the convergence of temporal approximations to solutions of the stochastic evolution equation (4.2.1) with multiplicative noise, the question of stability of time discretisation schemes arises. We aim to prove stability of contractive time discretisation schemes under linear growth assumptions on F and G. Stability refers to the fact that the sequence of approximations remains bounded as the number of time steps is increased. This is a necessary but not sufficient condition for convergence of the sequence of approximations. Since we are later interested in *pathwise* uniform error estimates, stability is to be understood in the sense of *pathwise* uniform boundedness of the approximations.

Let $R_k\colon X \to X$ be a contractive time discretisation scheme with time step $k > 0$ on a uniform grid $\{t_j = jk : \ j = 0, \ldots, N_k\} \subseteq [0,T]$ with final time $T = t_{N_k} > 0$ and $N_k = \frac{T}{k} \in \mathbb{N}$ being the number of time steps. We consider the temporal approximations of the mild solution to (4.2.1) given by $U^0 := u_0$ and

$$U^j := R_k U^{j-1} + kR_k F(t_{j-1}, U^{j-1}) + R_k G(t_{j-1}, U^{j-1})\Delta W_j, \tag{4.3.1}$$

$1 \le j \le N_k$, with Wiener increments $\Delta W_j := W_H(t_j) - W_H(t_{j-1})$ (see (2.7.1)) for $1 \le j \le N_k$. The above definition of U^j can be reformulated as the discrete variation-of-constants formula

$$U^j = R_k^j u_0 + k\sum_{i=0}^{j-1} R_k^{j-i} F(t_i, U^i) + \sum_{i=0}^{j-1} R_k^{j-i} G(t_i, U^i)\Delta W_{i+1} \tag{4.3.2}$$

for $j = 0, \ldots, N_k$.

We present two different stability proofs for contractive time discretisation schemes, one for Hilbert spaces X based on a dilation argument and one for 2-smooth Banach spaces X based on a martingale argument. Naturally, the latter is the stronger result in the sense that it implies the former. Nonetheless, the value of including the Hilbert space case resides not only in illustrating a different method of proof but also in significantly lower constant values. In view of a possible implementation, e.g., in adaptive time stepping, this is of considerable importance.

Proposition 4.16 (Stability in Hilbert spaces)**.** *Let X be a Hilbert space, $p \in [2,\infty)$, and $u_0 \in L^p_{\mathcal{F}_0}(\Omega; X)$. Suppose that $F\colon \Omega \times [0,T] \times X \to X$, $G\colon \Omega \times [0,T] \times X \to \mathcal{L}_2(H,X)$ are strongly $\mathcal{P}\otimes\mathcal{B}(X)$-measurable, where $F = \tilde{F} + f$ and $G = \tilde{G} + g$ for $f \in L^p_{\mathcal{P}}(\Omega; C([0,T]; X))$ and $g \in L^p_{\mathcal{P}}(\Omega; C([0,T]; \mathcal{L}_2(H,X)))$. Assume that there are $L_{F,X}, L_{G,X} \ge 0$ such that for all $\omega \in \Omega$, $t \in [0,T]$, and $x \in X$,*

$$\|\tilde{F}(\omega,t,x)\|_X \le L_{F,X}(1 + \|x\|_X),\ \|\tilde{G}(\omega,t,x)\|_{\mathcal{L}_2(H,X)} \le L_{G,X}(1 + \|x\|_X).$$

Let $(R_k)_{k>0}$ be a contractive time discretisation scheme and $N_k \ge 2$. Then the temporal approximations $(U^j)_{j=0,\ldots,N_k}$ obtained via (4.3.1) are stable. That is,

$$1 + \left\| \max_{0\le j\le N_k} \|U^j\| \right\|_p \le C_{stab} c_{u_0,f,g,T}, \tag{4.3.3}$$

where $C_{stab} := (1 + C_{\mathrm{e}}^2 T)^{1/2} \mathrm{e}^{(1+C_{\mathrm{e}}^2 T)/2}$ with $C_{\mathrm{e}} := L_{F,X} T^{1/2} + B_p L_{G,X}$,

$$c_{u_0,f,g,T} := 1 + \|u_0\|_{L^p(\Omega;X)} + T\|f\|_{p,\infty,X} + B_p T^{1/2} |\!|\!| g |\!|\!|_{p,\infty,X},$$

and B_p is the constant from Theorem 2.61.

The exponential dependence in Proposition 4.16 originates from an application of Gronwall's inequality. Therefore, to make the result suitable for numerical applications, some optimisation of the constants was necessary. In the special case that $L_{F,X} = L_{G,X} = T = 1$, and $p = 2$, we have $B_2 = 2$ and one can check that $C_{\text{stab}} = \sqrt{10}\mathrm{e}^5 \le 470$, which seems a reasonable constant for error estimates in applications.

The constant 1 in (4.3.3) can be left out, but is included, since terms of the form $1 + \| \max_{0\le j\le N_k} \|U^j\| \|_p$ will frequently arise from linear growth estimates in Section 4.4. The same holds true for the subsequent estimate (4.3.10) in Proposition 4.19.

Proof. Let $\varphi_N := 1 + \| \max_{0\le j\le N} \|U^j\| \|_p$ and $N \in \{0, \dots, N_k\}$. Then the variation-of-constants formula (4.3.2) and contractivity of R_k allow us to bound

$$\varphi_N \le 1 + \|u_0\|_{L^p(\Omega;X)} + k \sum_{i=0}^{N-1} \Big\| \max_{0\le j\le i} \|F(t_j, U^j)\| \Big\|_p + \Big\| \max_{0\le j\le N} \Big\| \sum_{i=0}^{j-1} R_k^{j-i} G(t_i, U^i)\Delta W_{i+1} \Big\| \Big\|_p. \tag{4.3.4}$$

Invoking linear growth of $\tilde{F}$ and boundedness of f for the second term, we obtain the bound

$$\begin{aligned} k \sum_{i=0}^{N-1} \Big\| \max_{0\le j\le i} \|F(t_j, U^j)\| \Big\|_p &\le k \sum_{i=0}^{N-1} \Big\| \max_{0\le j\le i} (L_{F,X}(1 + \|U^j\|) + \|f(t_j)\|) \Big\|_p \\ &\le k \sum_{i=0}^{N-1} \Big(L_{F,X} \Big(1 + \Big\| \max_{0\le j\le i} \|U^j\| \Big\|_p \Big) + \|f\|_{L^p(\Omega;C([0,T];X))} \Big) \\ &= C_{1,f} t_N + L_{F,X} k \sum_{i=0}^{N-1} \varphi_i \le C_{1,f} t_N + L_{F,X} t_N^{1/2} \Big(k \sum_{i=0}^{N-1} \varphi_i^2 \Big)^{1/2}, \end{aligned} \tag{4.3.5}$$

where we have set $C_{1,f} := \|f\|_{L^p(\Omega;C([0,T];X))}$, and used the Cauchy–Schwarz inequality as well as $Nk = t_N$ in the last line. It remains to bound the last term in (4.3.4).

Since R_k is a contraction, by the Sz.-Nagy dilation theorem as stated in Theorem 2.16, we can find a Hilbert space $\widetilde{X}$, a contractive injection $Q\colon X \to \widetilde{X}$, a contractive projection $P\colon \widetilde{X} \to X$, and a unitary $\widetilde{R}_k$ on $\widetilde{X}$ such that

$$R_k^i = P\widetilde{R}_k^i Q \quad \text{for all } i \ge 0.$$

Let $G^k(s) := G(t_i, U^i)$ and $S^k(s) := \widetilde{R}_k^{-i}$ for $s \in [t_i, t_{i+1})$, $0 \le i \le N_k - 1$. We then deduce from Theorem 2.61 and the triangle inequality in $\ell^2(\{0, \dots, N-1\}; L^p(\Omega; \mathcal{L}_2(H, X)))$ that

$$\begin{aligned} &\Big\| \max_{0\le j\le N} \Big\| \sum_{i=0}^{j-1} R_k^{j-i} G(t_i, U^i)\Delta W_{i+1} \Big\| \Big\|_p \le \Big\| \max_{0\le j\le N} \Big\| \sum_{i=0}^{j-1} \widetilde{R}_k^{j-i} Q G(t_i, U^i)\Delta W_{i+1} \Big\| \Big\|_p \\ &= \Big\| \max_{0\le j\le N} \Big\| \sum_{i=0}^{j-1} \widetilde{R}_k^{-i} Q G(t_i, U^i)\Delta W_{i+1} \Big\| \Big\|_p \le \Big\| \sup_{t\in[0,t_N]} \Big\| \int_0^t S^k(s) Q G^k(s) \,\mathrm{d}W_H(s) \Big\| \Big\|_p \\ &\le B_p \Big\| \Big(\int_0^{t_N} \|S^k(s) Q G^k(s)\|^2_{\mathcal{L}_2(H,X)} \,\mathrm{d}s \Big)^{1/2} \Big\|_p \overset{(*)}{\le} B_p \Big(k \sum_{i=0}^{N-1} \big\| \|G(t_i, U^i)\|_{\mathcal{L}_2(H,X)} \big\|_p^2 \Big)^{1/2} \end{aligned}$$

$$\leq B_p L_{G,X}\bigg(k\sum_{i=0}^{N-1}\varphi_i^2\bigg)^{1/2} + C_{2,g}t_N^{1/2}, \tag{4.3.6}$$

where we have set $C_{2,g} \coloneqq B_p\|g\|_{L^p(\Omega;C([0,T];\mathcal{L}_2(H,X)))}$ and used Minkowski's inequality in $L^{p/2}$ in $(*)$, see Remark 4.17 for further details.

Inserting (4.3.5) and (4.3.6) in (4.3.4) and estimating t_n by T gives the bound

$$\varphi_N \leq 1 + \|u_0\|_{L^p(\Omega;X)} + C_{1,f}T + C_{2,g}T^{1/2} + (L_{F,X}T^{1/2} + B_pL_{G,X})\bigg(k\sum_{i=0}^{N-1}\varphi_i^2\bigg)^{1/2}.$$

Setting $C_\mathrm{e} \coloneqq L_{F,X}T^{1/2} + B_pL_{G,X}$ and $c_{u_0,f,g} \coloneqq 1 + \|u_0\|_{L^p(\Omega;X)} + C_{1,f}T + C_{2,g}T^{1/2}$, we obtain from the discrete version of Gronwall's Lemma 2.68 that

$$\varphi_N \leq c_{u_0,f,g}(1 + C_\mathrm{e}^2kN)^{1/2}\mathrm{e}^{(1+C_\mathrm{e}^2kN)/2}.$$

This implies the desired statement for $N = N_k$ noting that $t_{N_k} = kN_k = T$. □

The "trick" explained in the following remark is used frequently in the following.

Remark 4.17. Working with higher moments $p \in [2,\infty)$ allows us to pull L^p-norms inside the square root of a sum of squares in the following way: Let $\phi \in L^p(\Omega;\ell^2(\{1,\ldots,n\};X))$ for some $n \in \mathbb{N}$. Rewriting L^p-norms of a square root as the square root of a $L^{p/2}$-norm, and vice versa, puts us in a position to apply Minkowski's inequality in $L^{p/2}(\Omega)$, which yields

$$\bigg\|\bigg(\sum_{i=1}^n\|\phi_i\|^2\bigg)^{1/2}\bigg\|_p = \bigg\|\sum_{i=1}^n\|\phi_i\|^2\bigg\|_{p/2}^{1/2} \leq \bigg(\sum_{i=1}^n\big\|\|\phi_i\|^2\big\|_{p/2}\bigg)^{1/2} = \bigg(\sum_{i=1}^n\|\phi_i\|_{L^p(\Omega;X)}^2\bigg)^{1/2}$$

for all $\phi \in \ell^2(\{1,\ldots,n\};L^p(\Omega;X))$. In summary, $\ell^2(\{1,\ldots,n\};L^p(\Omega;X))$ embeds continuously in $L^p(\Omega;\ell^2(\{1,\ldots,n\};X))$.

We now generalise this stability result to 2-smooth Banach spaces. This requires replacing the dilation argument in the proof of Proposition 4.16 by a martingale one based on Pinelis' version of the Burkholder–Rosenthal inequalities from Theorem 2.63, which is the subject of the following lemma. The same martingale argument will later be applied to deduce uniform convergence of some regularisation error in Section 4.6.

Lemma 4.18. *Let X be a $(2,D)$-smooth Banach space, $N \in \mathbb{N}$ with $N \leq N_k$, and $Q\colon \Omega \times [0,T] \to \gamma(H,X)$ be such that $Q_i \coloneqq Q(\cdot,t_i) \in L^p(\Omega;\gamma(H,X))$ is $\mathcal{F}_{t_i}$-measurable for $0 \leq i \leq N-1$. Suppose that $(R_k)_{k>0}$ is contractive. Then*

$$\bigg\|\max_{0\leq j\leq N}\bigg\|\sum_{i=0}^{j-1}R_k^{j-i}Q_i\Delta W_{i+1}\bigg\|\bigg\|_p \leq B_{p,D}\bigg(k\sum_{i=0}^{N-1}\|Q_i\|_{L^p(\Omega;\gamma(H,X))}^2\bigg)^{1/2},$$

where the constant is given by $B_{p,D} = 10p^2(p-1)^{-1}C_{p,D} + 10D\sqrt{p} = 10D\sqrt{p}(10p^2(p-1)^{-1}+1)$ with $C_{p,D}$ denoting the constant from Theorem 2.61.

Proof. The statement follows from an application of Pinelis' version of the Burkholder–Rosenthal inequalities from Theorem 2.63 as illustrated in Step 1. The terms emerging therefrom are simplified in Steps 2 and 3.

Step 1. Define $\widetilde{M}_j \coloneqq \sum_{i=0}^{j-1} R_k^{j-i} Q_i \Delta W_{i+1}$ and $M_j \coloneqq \sum_{i=1}^{j} R_k Q_{i-1} \Delta W_i$ for $0 \leq j \leq N$. Further define $d_j \coloneqq M_j - M_{j-1}$ and $V_j \coloneqq R_k$ for all $1 \leq j \leq N$. Then $\widetilde{M}_0 = M_0 = 0$ by construction and $(\tilde{M}_j)_{j=0}^N$ is adapted because Q_i is $\mathscr{F}_{t_i}$- and thus also $\mathscr{F}_{t_{j-1}}$-measurable and ΔW_{i+1} is $\mathscr{F}_{t_{i+1}}$- and thus also $\mathscr{F}_{t_j}$-measurable for $0 \leq i \leq j-1$. Furthermore, $(M_j)_{j=0}^N$ is an X-valued martingale, since it is adapted and for $0 \leq \ell \leq j \leq N$

$$\mathbb{E}(M_j | \mathscr{F}_{t_\ell}) = \sum_{i=1}^{j} R_k Q_{i-1} \mathbb{E}(\Delta W_i | \mathscr{F}_{t_\ell}) = \sum_{i=1}^{\ell} R_k Q_{i-1} \Delta W_i = M_\ell$$

by independence of ΔW_i of $\mathscr{F}_{t_\ell}$ for all $i \geq \ell + 1$ (see Example 2.53 for details). Moreover, it has conditionally symmetric increments because ΔW_i and $-\Delta W_i$ are conditionally equidistributed, as follows from the definition of H-cylindrical Brownian motions; cf. Definition 2.54. Consequently, $(d_j)_{j=1}^N$ is a martingale difference sequence, and Theorem 2.63 is applicable with better constants. It yields the bound

$$\Big\| \max_{0 \leq j \leq N} \Big\| \sum_{i=0}^{j-1} R_k^{j-i} Q_i \Delta W_{i+1} \Big\| \Big\|_p = \|(\widetilde{M})^*\|_p \leq 5p \|d^*\|_p + 10D\sqrt{p} \|s(M)\|_p, \tag{4.3.7}$$

where $d^* = \max_{1 \leq j \leq N} \|d_j\|$ and $s(M)^2 = \sum_{i=0}^{N-1} \mathbb{E}(\|M_{i+1} - M_i\|^2 \mid \mathscr{F}_{t_i})$.

Step 2. To simplify the first term, we first apply the triangle inequality and Doob's maximal inequality [71, Thm. 3.2.2] before rewriting the Wiener increments as stochastic integrals to apply Itô's isomorphism as in Theorem 2.61. Lastly making use of Minkowski's inequality in $L^{p/2}(\Omega)$, contractivity of R_k and the dominated convergence theorem in $L^p(\Omega)$, it follows that

$$\begin{aligned}
\|d^*\|_p &= \Big\| \max_{1 \leq j \leq N} \|M_j - M_{j-1}\| \Big\|_p \leq 2\|M^*\|_p \leq \frac{2p}{p-1} \|M_N\|_p \\
&= \frac{2p}{p-1} \Big\| \int_0^{t_N} \sum_{i=0}^{N-1} \mathbf{1}_{(t_i, t_{i+1}]}(s) R_k Q_i \, \mathrm{d}W_H(s) \Big\|_{L^p(\Omega;X)} \\
&\leq \frac{2pC_{p,D}}{p-1} \Big\| \Big(\int_0^{t_N} \Big\| \sum_{i=0}^{N-1} \mathbf{1}_{(t_i, t_{i+1}]}(s) R_k Q_i \Big\|_{\gamma(H,X)}^2 \mathrm{d}s \Big)^{1/2} \Big\|_p \\
&= \frac{2pC_{p,D}}{p-1} \Big\| \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \|R_k Q_i\|_{\gamma(H,X)}^2 \, \mathrm{d}s \Big\|_{p/2}^{1/2} \\
&\leq \frac{2pC_{p,D}}{p-1} \Big(k \sum_{i=0}^{N-1} \big\| \|Q_i\|_{\gamma(H,X)}^2 \big\|_{p/2} \Big)^{1/2} = \frac{2pC_{p,D}}{p-1} \Big(k \sum_{i=0}^{N-1} \|Q_i\|_{L^p(\Omega;\gamma(H,X))}^2 \Big)^{1/2}.
\end{aligned} \tag{4.3.8}$$

Step 3. Using that the Wiener increments ΔW_{i+1} are independent of $\mathcal{F}_{t_i}$ and have variance $t_{i+1} - t_i = k$, we can bound the remaining term $\|s(M)\|_p$ in (4.3.7) by

$$\|s(M)\|_p = \Big\| \Big(\sum_{i=0}^{N-1} \mathbb{E}\big(\|R_k Q_i \Delta W_{i+1}\|^2 \mid \mathcal{F}_{t_i} \big) \Big)^{1/2} \Big\|_p$$

$$\le \Big\| \Big(\sum_{i=0}^{N-1} \|R_k Q_i\|^2_{\gamma(H,X)} \mathbb{E}(|\Delta W_{i+1}|^2 \mid \mathscr{F}_{t_i}) \Big)^{1/2} \Big\|_p$$
$$= \Big\| k \sum_{i=0}^{N-1} \|R_k Q_i\|^2_{\gamma(H,X)} \Big\|^{1/2}_{p/2} \le \Big(k \sum_{i=0}^{N-1} \|Q_i\|^2_{L^p(\Omega;\gamma(H,X))} \Big)^{1/2}. \tag{4.3.9}$$

Here, we have used a basic property of conditional expectations from Proposition 2.50. The statement of the theorem is obtained with $B_{p,D} = 10D\sqrt{p}(10p^2(p-1)^{-1}+1)$ from inserting (4.3.8) and (4.3.9) in (4.3.7), noting that $C_{p,D} = 10D\sqrt{p}$. □

This enables us to extend the stability of contractive time discretisation schemes from Hilbert to 2-smooth Banach spaces. Later on, we will also apply Proposition 4.19 in case the space X is replaced by Y in the setting of Section 4.2.

Proposition 4.19 (Stability in 2-smooth Banach spaces)**.** *Let X be a $(2,D)$-smooth Banach space, $p \in [2,\infty)$, and $u_0 \in L^p_{\mathcal{F}_0}(\Omega;X)$. Suppose that $F\colon \Omega \times [0,T] \times X \to X$, $G\colon \Omega \times [0,T] \times X \to \gamma(H,X)$ are strongly $\mathcal{P} \otimes \mathcal{B}(X)$-measurable, where $F = \tilde{F} + f$ and $G = \tilde{G} + g$ for $f \in L^p_{\mathcal{P}}(\Omega; C([0,T];X))$ and $g \in L^p_{\mathcal{P}}(\Omega; C([0,T];\gamma(H,X)))$. Assume that there are $L_{F,X}, L_{G,X} \ge 0$ such that for all $\omega \in \Omega$, $t \in [0,T]$, and $x \in X$,*

$$\|\tilde{F}(\omega,t,x)\|_X \le L_{F,X}(1+\|x\|_X), \; \|\tilde{G}(\omega,t,x)\|_{\gamma(H,X)} \le L_{G,X}(1+\|x\|_X).$$

Let $(R_k)_{k>0}$ be a contractive time discretisation scheme and $N_k \ge 2$. Then the temporal approximations $(U^j)_{j=0,\dots,N_k}$ obtained via (4.3.1) are stable. That is, for all $N_k \in \mathbb{N}$,

$$1 + \Big\| \max_{0\le j\le N_k} \|U^j\| \Big\|_p \le C_{stab} c_{u_0,f,g,T}, \tag{4.3.10}$$

where the constants $C_{stab} := (1+C_{\mathrm{e}}^2 T)^{1/2} \mathrm{e}^{(1+C_{\mathrm{e}}^2 T)/2}$ with $C_{\mathrm{e}} := L_{F,X}T^{1/2} + B_{p,D}L_{G,X}$ and

$$c_{u_0,f,g,T} := 1 + \|u_0\|_{L^p(\Omega;X)} + T\|f\|_{p,\infty,X} + B_{p,D}T^{1/2}\|g\|_{p,\infty,X}$$

are independent of $N_k \in \mathbb{N}$, and $B_{p,D}$ is the constant from Lemma 4.18.

The stability constants C_{stab} and $c_{u_0,f,g,T}$ are of the same form as those in Proposition 4.16 for Hilbert spaces. However, the appearance of $B_{p,D}$ instead of B_p leads to drastically increased constant values. In the special case that $L_{F,X} = L_{G,X} = T = D = 1$, and $p = 2$, $B_{p,D} \approx 580$ and thus $C_{\mathrm{stab}} \approx 8 \cdot 10^{73259}$, which is clearly unfeasible for numerical applications. Even in the Hilbert space case, under the previous assumptions we obtain $B_{p,D} \approx 95$ using that $C_{p,D} = 2$ in this case, which implies $C_{\mathrm{stab}} \approx 7 \cdot 10^{1967}$. Clearly, the constant $C_{\mathrm{stab}} \approx 470$ from Proposition 4.16 is preferable.

Later on, we will also apply Proposition 4.19 in case the space X is replaced by Y in the setting of Section 4.2.

Proof. Let $N \in \{0,\dots,N_k\}$ and $\varphi_N := 1 + \|\max_{0\le j\le N}\|U^j\|\|_p$. Then the variation-of-constants formula (4.3.2) and contractivity of R_k allow us to bound

$$\varphi_N \le 1 + \|u_0\|_{L^p(\Omega;X)} + k\sum_{i=0}^{N-1} \Big\| \max_{0\le j\le i} \|F(t_j,U^j)\| \Big\|_p$$

$$+\Big\|\max_{0\le j\le N}\Big\|\sum_{i=0}^{j-1} R_k^{j-i} G(t_i,U^i)\Delta W_{i+1}\Big\|\Big\|_p. \tag{4.3.11}$$

Invoking linear growth of $\tilde{F}$ and pathwise continuity of f for the third term, we obtain the bound

$$\begin{aligned} k\sum_{i=0}^{N-1}\Big\|\max_{0\le j\le i}\|F(t_j,U^j)\|\Big\|_p &\le k\sum_{i=0}^{N-1}\Big(L_{F,X}\Big(1+\Big\|\max_{0\le j\le i}\|U^j\|\Big\|_p\Big)+\|f\|_{p,\infty,X}\Big)\\ &= t_N\|f\|_{p,\infty,X}+L_{F,X}k\sum_{i=0}^{N-1}\varphi_i \le t_N\|f\|_{p,\infty,X}+L_{F,X}\sqrt{t_N}\Big(k\sum_{i=0}^{N-1}\varphi_i^2\Big)^{1/2}, \end{aligned} \tag{4.3.12}$$

where we have used the Cauchy–Schwarz inequality in the last step.
To the last term in (4.3.11) we apply Lemma 4.18 with $Q_i \coloneqq G(t_i,U^i)$ for $0\le i\le N-1$, which together with the triangle inequality in $\ell^2(\{0,\dots,N-1\};L^p(\Omega;\gamma(H,X)))$ and linear growth of $\tilde{G}$ yields

$$\begin{aligned} &\Big\|\max_{0\le j\le N}\Big\|\sum_{i=0}^{j-1} R_k^{j-i} G(t_i,U^i)\Delta W_{i+1}\Big\|\Big\|_p \le B_{p,D}\Big(k\sum_{i=0}^{N-1}\|G(t_i,U^i)\|^2_{L^p(\Omega;\gamma(H,X))}\Big)^{1/2}\\ &\le B_{p,D}\Big(k\sum_{i=0}^{N-1}\|g(t_i)\|^2_{L^p(\Omega;X)}\Big)^{1/2}+B_{p,D}L_{G,X}\Big(k\sum_{i=0}^{N-1}\|1+U^i\|^2_{L^p(\Omega;X)}\Big)^{1/2}\\ &\le B_{p,D}\Bigg((kN)^{1/2}\Big\|\max_{0\le j\le N-1}\|g(t_j)\|_{\gamma(H,X)}\Big\|_p+L_{G,X}\Bigg(k\sum_{i=0}^{N-1}\Big(1+\Big\|\max_{0\le j\le i}\|U^j\|\Big\|_p\Big)^2\Bigg)^{1/2}\Bigg)\\ &\le B_{p,D}\sqrt{t_N}\|\!|g|\!\|_{p,\infty,X}+B_{p,D}L_{G,X}\Big(k\sum_{i=0}^{N-1}\varphi_i^2\Big)^{1/2}. \end{aligned} \tag{4.3.13}$$

Inserting (4.3.12) and (4.3.13) in (4.3.11) followed by an application of the discrete version of Gronwall's inequality from Lemma 2.68 results in

$$\varphi_N^2 \le \sqrt{C_{\mathrm{e}}}\mathrm{e}^{C_{\mathrm{e}}/2}\big(1+\|u_0\|_{L^p(\Omega;X)}+T\|f\|_{p,\infty,X}+B_{p,D}\sqrt{T}\|\!|g|\!\|_{p,\infty,X}\big)$$

with $C_{\mathrm{e}} \coloneqq 1+L_{F,X}^2 t_N+B_{p,D}^2L_{G,X}^2 t_N$, which implies the desired statement for $N=N_k$ noting that $t_{N_k}=T$. □

Remark 4.20. We point out that Proposition 4.16 and Proposition 4.19 do not require Lipschitz continuity of F or G. Hence, they are also applicable on spaces $Y\hookrightarrow X$, on which $\tilde{F}$ and $\tilde{G}$ are of linear growth. Then (4.3.3) and (4.3.10) hold with the Y-norm of U^j if one replaces $L_{F,X}$ and $L_{G,X}$ by the respective linear growth constants $L_{F,Y}$ and $L_{G,Y}$ on Y. This is also the reason why the additional constant 1 appears in (4.3.3), (4.3.10) and the definition of $c_{u_0,f,g,T}$, it is not required in case $\tilde{F}$ and $\tilde{G}$ are Lipschitz continuous.

4.4 Convergence Rates for Multiplicative Noise

Our aim is to prove rates of convergence of contractive time discretisation schemes for nonlinear stochastic evolution equations of the form

$$\mathrm{d}U + AU\,\mathrm{d}t = F(t,U)\,\mathrm{d}t + G(t,U)\,\mathrm{d}W_H(t), \quad U(0) = u_0 \in L^p(\Omega;X) \tag{4.4.1}$$

with $t \in [0,T]$ on a 2-smooth Banach space X with norm $\|\cdot\|$, where W_H is an H-cylindrical Brownian motion for some Hilbert space H and $p \in [2,\infty)$. The operator $-A$ is assumed to generate a C_0-contraction semigroup $(S(t))_{t\geq 0}$ on X and F and G are assumed to be progressively measurable, of linear growth, and globally Lipschitz as detailed in Assumption 4.11. Hence, the unique mild solution is given by a fixed point of

$$U(t) = S(t)u_0 + \int_0^t S(t-s)F(s,U(s))\,\mathrm{d}s + \int_0^t S(t-s)G(s,U(s))\,\mathrm{d}W_H(s) \tag{4.4.2}$$

for $t \in [0,T]$, see Section 4.2.

To obtain convergence rates for temporal discretisations of the mild solution, we assume additional structure of the nonlinearity F and the noise G. Let Y be another 2-smooth Banach space such that $Y \hookrightarrow X$ and the semigroup $(S(t))_{t\geq 0}$ is also contractive on Y. We will assume that F and G map Y into Y and enjoy the same linear growth conditions on Y as on X. Note that no Lipschitz continuity is assumed on Y in contrast to X. This additional structure resembling the famous Kato setting [84] allows for convergence rates of temporal discretisations of stochastic evolution equations with such nonlinearities for a large class of time discretisation schemes, as detailed in Subsection 4.4.1. The quantitative error estimate in Theorem 4.24 is the main result of this chapter, stating that the additional structure of F and G suffices to obtain the order of the scheme as the convergence rate of the temporal discretisations up to a logarithmic correction factor for sufficiently regular initial data. For the exponential Euler method, the logarithmic correction factor can be omitted, as illustrated in Subsection 4.4.2. The main error estimate of Theorem 4.24 is extended to the full time interval $[0,T]$ for Hilbert spaces X in Subsection 4.4.3. As an application, we revisit Schrödinger's equation in Subsection 4.4.4, now with multiplicative noise, and consider the stochastic Maxwell's equations in Subsection 4.4.5.

4.4.1 General contractive time discretisation schemes

We now detail the assumptions on the structure of F and G on Y. Note that the assumptions also imply that the conditions of Theorems 4.12 and 4.14 hold.

Assumption 4.21. *Let X and Y be 2-smooth Banach spaces such that $Y \hookrightarrow X$ continuously, and let $p \in [2,\infty)$. Let $F\colon \Omega \times [0,T] \times X \to X$ and $G\colon \Omega \times [0,T] \times X \to \gamma(H,X)$ be strongly $\mathcal{P}\otimes\mathcal{B}(X)$-measurable, where $F = \tilde{F} + f$ and $G = \tilde{G} + g$ for $f \in L^p_{\mathcal{P}}(\Omega; C([0,T];X))$ and $g \in L^p_{\mathcal{P}}(\Omega; C([0,T];\gamma(H,X)))$. Suppose that $\tilde{F}(\cdot,\cdot,0) = 0$, $\tilde{G}(\cdot,\cdot,0) = 0$, and*

(a) (global Lipschitz continuity on X) *there exist constants $C_{F,X}, C_{G,X} \geq 0$ such that for all $\omega \in \Omega, t \in [0,T]$, and $x,y \in X$,*

$$\|\tilde{F}(\omega,t,x) - \tilde{F}(\omega,t,y)\| \leq C_{F,X}\|x-y\|, \; \|\tilde{G}(\omega,t,x) - \tilde{G}(\omega,t,y)\|_{\gamma(H,X)} \leq C_{G,X}\|x-y\|,$$

(b) (Hölder continuity with values in X) *for some* $\alpha \in (0,1]$,

$$C_{\alpha,F} := \sup_{\omega\in\Omega, x\in X} [F(\omega,\cdot,x)]_\alpha < \infty, \ C_{\alpha,G} := \sup_{\omega\in\Omega, x\in X} [G(\omega,\cdot,x)]_\alpha < \infty,$$

(c) (Y-invariance) $F\colon \Omega\times[0,T]\times Y \to Y$ *and* $G\colon \Omega\times[0,T]\times Y\to \gamma(H,Y)$ *are strongly* $\mathcal{P}\otimes\mathcal{B}(Y)$*-measurable,* $f \in L^p_{\mathcal{P}}(\Omega; C([0,T];Y))$*, and* $g\in L^p_{\mathcal{P}}(\Omega; C([0,T];\gamma(H,Y)))$*,*

(d) (linear growth on Y) *and there exist constants* $L_{F,Y}, L_{G,Y} \geq 0$ *such that for all* $\omega\in\Omega, t\in[0,T]$*, and* $x\in Y$*,*

$$\|\tilde{F}(\omega,t,x)\|_Y \leq L_{F,Y}(1+\|x\|_Y), \ \|\tilde{G}(\omega,t,x)\|_{\gamma(H,Y)} \leq L_{G,Y}(1+\|x\|_Y).$$

We would like to recall that strong measurability is to be understood in the sense of approximation with simple functions; cf. Definition 2.42 and Remark 2.43, not in the topological sense. Condition (b) can be weakened to the existence of some $\alpha\in(0,1]$ such that

$$\sup_{x\in X}\sup_{0\leq s\leq t\leq T}\frac{F(\cdot,t,x)-F(\cdot,s,x)}{(t-s)^\alpha} \in L^p(\Omega)$$

and likewise for G, i.e., pathwise Hölder continuity uniformly in $x\in X$ is sufficient together with existence of p-th moments of the Hölder seminorms. For the sake of readability, we use Assumption 4.21(b) in the following. Assumption 4.21 implies that (4.4.1) has a unique mild solution.

To bound the error arising from time discretisation of the mild solution, moment bounds of differences of the mild solution at different time points as in the following lemma are required. We recall the shorthand notation $\|\cdot\|_{p,q,Z}$ and $⦀\cdot⦀_{p,q,Z}$ from (4.2.12). Furthermore, we introduce the constants

$$C_{u_0,f,g,Z} := 1 + C^Z_{\mathrm{bdd}}(1+\|u_0\|_{L^p(\Omega;Z)} + \|f\|_{p,1,Z} + C_{p,D}⦀g⦀_{p,2,Z}) \tag{4.4.3}$$

for $Z\in\{X,Y\}$ with C^X_{bdd} and C^Y_{bdd} as in Theorems 4.12 and 4.14, respectively. Then the aforementioned theorems imply the estimate

$$1 + \left\| \sup_{t\in[0,T]} \|U(t)\|_Z \right\|_p \leq C_{u_0,f,g,Z} < \infty \tag{4.4.4}$$

for the mild solution U to (4.4.1) for $Z\in\{X,Y\}$. The constant $C_{p,D} = 10D\sqrt{p}$ used in the following can be replaced by the better constant $B_p = 4\sqrt{p}$ for $p\in(2,\infty)$ or $B_2 = 2$ if X is a Hilbert space; cf. Theorem 2.61.

Lemma 4.22 (α-Hölder continuity of p-th moments of the mild solution)**.** *Suppose that Assumption 4.21 holds for some* $\alpha\in(0,1]$ *and* $p\in[2,\infty)$*. Let* $-A$ *be the generator of a* C_0*-contraction semigroup* $(S(t))_{t\geq 0}$ *on both* X *and* Y*. Suppose that* $Y\hookrightarrow D_A(\alpha,\infty)$ *continuously if* $\alpha\in(0,1)$ *or* $Y\hookrightarrow \mathrm{D}(A)$ *continuously if* $\alpha = 1$*. Let* $u_0\in L^p_{\mathcal{F}_0}(\Omega;Y)$*. Then for all* $0\leq s\leq t\leq T$ *the mild solution* U *of* (4.4.1) *satisfies*

$$(\mathbb{E}\|U(t)-U(s)\|^p)^{1/p} \leq L_1(t-s) + L_2(t-s)^{1/2} + L_3(t-s)^\alpha$$

with constants $L_1 \coloneqq C_{F,X} C_{u_0,f,g,X} + \|f\|_{p,\infty,X}$, $L_2 \coloneqq C_{p,D}(C_{G,X} C_{u_0,f,g,X} + \|g\|_{p,\infty,X})$, *and*

$$L_3 \coloneqq 2C_Y \Big[\|u_0\|_{L^p(\Omega;Y)} + T L_{F,Y} C_{u_0,f,g,Y} + \|f\|_{p,1,Y} + C_{p,D}\big(T^{1/2} L_{G,Y} C_{u_0,f,g,Y} + \|g\|_{p,2,Y}\big)\Big],$$

where $C_{u_0,f,g,X}$ *and* $C_{u_0,f,g,Y}$ *are as defined in* (4.4.3), C_Y *denotes the embedding constant of* Y *into* $D_A(\alpha,\infty)$ *or* $\mathrm{D}(A)$, *and* $C_{p,D}$ *is the constant from Theorem 2.61.*

Proof. Since the conditions of Theorem 4.12 are met, U is pathwise continuous on X. By Theorem 4.14, the pathwise continuity of U follows on Y as well. Moreover, the bound (4.4.4) holds.

Fix $t,s \in [0,T]$ with $s \le t$. From the mild solution formula (4.4.2) we deduce that

$$\begin{aligned}
&(\mathbb{E}\|U(t) - U(s)\|^p)^{1/p} \le \|[S(t) - S(s)]u_0\|_{L^p(\Omega;X)} \\
&+ \Big\| \int_0^s \|[S(t-r) - S(s-r)]F(r,U(r))\| \,\mathrm{d}r \Big\|_p + \Big\| \int_s^t \|S(t-r)F(r,U(r))\| \,\mathrm{d}r \Big\|_p \\
&+ \Big\| \int_0^s [S(t-r) - S(s-r)]G(r,U(r)) \,\mathrm{d}W_H(r) \Big\|_{L^p(\Omega;X)} \\
&+ \Big\| \int_s^t S(t-r)G(r,U(r)) \,\mathrm{d}W_H(r) \Big\|_{L^p(\Omega;X)} \eqqcolon E_1 + E_2 + E_3 + E_4 + E_5,
\end{aligned}$$

where $E_\ell = E_\ell(t,s)$ for $1 \le \ell \le 5$. We proceed to bound these five expressions individually. By the semigroup bound (2.3.3),

$$E_1 \le \|S(t) - S(s)\|_{\mathcal{L}(Y,X)} \|u_0\|_{L^p(\Omega;Y)} \le 2C_Y (t-s)^\alpha \|u_0\|_{L^p(\Omega;Y)}.$$

Using (4.4.4) and (2.3.3) as well as linear growth of $\tilde{F}$ on Y and pathwise integrability of f with values in Y, we obtain

$$\begin{aligned}
E_2 &\le 2C_Y \Big\| \int_0^s [(t-r) - (s-r)]^\alpha \|F(r,U(r))\|_Y \,\mathrm{d}r \Big\|_p \\
&\le 2C_Y (t-s)^\alpha \bigg(s \Big\| \sup_{r\in[0,T]} \|\tilde{F}(r,U(r))\|_Y \Big\|_p + \Big\| \int_0^s \|f(r)\|_Y \,\mathrm{d}r \Big\|_p \bigg) \\
&\le 2C_Y (t-s)^\alpha \bigg(s L_{F,Y} \Big\| \sup_{r\in[0,T]} (1 + \|U(r)\|_Y) \Big\|_p + \|f\|_{p,1,Y} \bigg) \\
&\le 2C_Y \big(L_{F,Y} C_{u_0,f,g,Y} T + \|f\|_{p,1,Y} \big) (t-s)^\alpha.
\end{aligned}$$

Analogously, contractivity of the semigroup, Lipschitz continuity of $\tilde{F}$ on X and boundedness of the mild solution yield

$$E_3 \le (C_{F,X} C_{u_0,f,g,X} + \|f\|_{p,\infty,X})(t-s).$$

For the terms involving a stochastic integral, we apply Theorem 2.61. Additionally using the bound (2.3.3) for semigroup differences, splitting the integral as in E_2, and using linear growth of $\tilde{G}$, (4.4.4), as well as pathwise square-integrability of g results in

$$E_4 \le C_{p,D} \bigg(\mathbb{E} \Big(\int_0^s \|[S(t-r) - S(s-r)]G(r,U(r))\|_{\gamma(H,X)}^2 \,\mathrm{d}r \Big)^{p/2} \bigg)^{1/p}$$

$$\leq 2C_{p,D}C_Y(t-s)^\alpha\Big(s^{1/2}\Big\|\sup_{r\in[0,T]}\|\tilde{G}(r,U(r))\|_{\gamma(H,Y)}\Big\|_p+\Big\|\Big(\int_0^s\|g(r)\|^2_{\gamma(H,Y)}\,\mathrm{d}r\Big)^{1/2}\Big\|_p\Big)$$
$$\leq 2C_{p,D}C_Y\big(T^{1/2}L_{G,Y}C_{u_0,f,g,Y}+\|\!|g|\!\|_{p,2,Y}\big)(t-s)^\alpha.$$

For the last term, the contractivity of the semigroup and Lipschitz continuity of $\tilde{G}$ yield

$$E_5\leq C_{p,D}\Big(\mathbb{E}\Big(\int_s^t\|S(t-r)G(r,U(r))\|^2_{\gamma(H,X)}\,\mathrm{d}r\Big)^{p/2}\Big)^{1/p}$$
$$\leq C_{p,D}(C_{G,X}C_{u_0,f,g,X}+\|\!|g|\!\|_{p,\infty,X})(t-s)^{1/2}.$$

In conclusion, from the five individual bounds we obtain the statement of the lemma

$$\begin{aligned}(\mathbb{E}\|U(t)-U(s)\|^p)^{1/p}\leq{}&(C_{F,X}C_{u_0,f,g,X}+\|f\|_{p,\infty,X})(t-s)\\&+C_{p,D}\big(C_{G,X}C_{u_0,f,g,X}+\|\!|g|\!\|_{p,\infty,X}\big)(t-s)^{1/2}\\&+2C_Y\big[\|u_0\|_{L^p(\Omega;Y)}+(TL_{F,Y}C_{u_0,f,g,Y}+\|f\|_{p,1,Y})\\&+C_{p,D}(T^{1/2}L_{G,Y}C_{u_0,f,g,Y}+\|\!|g|\!\|_{p,2,Y})\big](t-s)^\alpha.\end{aligned}$$ □

Remark 4.23. Suppose that $\alpha\in(0,\frac{1}{2}]$. Lemma 4.22 implies α-Hölder continuity of U in p-th moment. The latter remains true if the pathwise continuity of f and g with values in Y from Assumption 4.21(c) are relaxed to $\|f\|_{p,1,Y}<\infty$ and $\|\!|g|\!\|_{p,2,Y}<\infty$. Performing an additional Hölder argument for E_3 and E_5, the pathwise continuity assumption with values in X can be relaxed to $\|f\|_{p,\frac{1}{1-\alpha},X}<\infty$ and $\|\!|g|\!\|_{p,\frac{2}{1-2\alpha},X}<\infty$, where we use the convention $\frac{1}{0}=\infty$. For our purposes, the above version of the lemma is sufficient, since even pathwise continuity with values in Y is required in Theorem 4.24.

For time discretisation, we employ a contractive time discretisation scheme $R\colon[0,\infty)\to\mathcal{L}(X)$ with time step $k>0$ on a uniform grid $\{t_j=jk:\ j=0,\dots,N_k\}\subseteq[0,T]$ with final time $T=t_{N_k}>0$ and $N_k=\frac{T}{k}\in\mathbb{N}$ being the number of time steps. As in the previous section, the discrete approximation is given by $U^0=u_0$ and

$$U^j=R_kU^{j-1}+kR_kF(t_{j-1},U^{j-1})+R_kG(t_{j-1},U^{j-1})\Delta W_j \tag{4.4.5}$$
$$=R_k^ju_0+k\sum_{i=0}^{j-1}R_k^{j-i}F(t_i,U^i)+\sum_{i=0}^{j-1}R_k^{j-i}G(t_i,U^i)\Delta W_{i+1} \tag{4.4.6}$$

for $j=1,\dots,N_k$ with Wiener increments $\Delta W_j=W_H(t_j)-W_H(t_{j-1})$.

We recall from Definition 2.24 that R *approximates* S *to order* $\alpha>0$ *on* Y or, equivalently, R *converges of order* α *on* Y if there is a constant $C_\alpha\geq 0$ such that for all $u\in Y$

$$\|(S(t_j)-R_k^j)u\|\leq C_\alpha k^\alpha\|u\|_Y.$$

Under the conditions of Assumption 4.21 we conclude from Proposition 4.19 and Remark 4.20 that R is stable not only on X but also on Y provided that $u_0\in L^p_{\mathcal{F}_0}(\Omega;Y)$ and both S and R are contractive on both X and Y. Thus,

$$1+\Big\|\max_{0\leq j\leq N_k}\|U^j\|_Y\Big\|_p\leq K_{u_0,f,g,Y}, \tag{4.4.7}$$

where $K_{u_0,f,g,Y} := C_{\mathrm{stab}} c_{u_0,f,g,T}$ with constants $C_{\mathrm{stab}}, c_{u_0,f,g,T}$ as in Proposition 4.19 applied on Y instead of X. Here and in the following, the better constant values from Proposition 4.16 can be used for C_{stab} and $c_{u_0,f,g,T}$ if the underlying space is a Hilbert space.

We are now in a position to state and prove one of the main results of this thesis.

Theorem 4.24. *Suppose that Assumption 4.21 holds for some* $\alpha \in (0,1]$ *and* $p \in [2,\infty)$. *Let* $-A$ *be the generator of a* C_0*-contraction semigroup* $(S(t))_{t\geq 0}$ *on both* X *and* Y. *Let* $(R_k)_{k>0}$ *be a time discretisation scheme that is contractive on* X *and* Y. *Assume* R *approximates* S *to order* α *on* Y. *Suppose that* $Y \hookrightarrow D_A(\alpha,\infty)$ *continuously if* $\alpha \in (0,1)$ *or* $Y \hookrightarrow \mathrm{D}(A)$ *continuously if* $\alpha = 1$. *Let* $u_0 \in L^p_{\mathcal{F}_0}(\Omega;Y)$. *Denote by* U *the mild solution of* (4.4.1) *and by* $(U^j)_{j=0,\dots,N_k}$ *the temporal approximations as defined in* (4.4.5). *Then for* $N_k \geq 2$

$$\Big\| \max_{0\leq j\leq N_k} \|U(t_j) - U^j\| \Big\|_p \leq C_{\mathrm{e}}\Big(C_1 k + C_2 k^{1/2} + \big(C_3 + C_4 \sqrt{\max\{\log(T/k), p\}}\big) k^\alpha \Big)$$

with constants $C_{\mathrm{e}} := (1 + C^2 T)^{1/2} \exp((1 + C^2 T)/2)$, $C := C_{F,X}\sqrt{T} + C_{p,D} C_{G,X}$, $C_1 := L_1(\frac{1}{2} C_{F,X} T^2 + \frac{1}{\sqrt{3}} C_{p,D} C_{G,X} \sqrt{T})$, $C_2 := L_2(\frac{2}{3} C_{F,X} T + \frac{1}{\sqrt{2}} C_{p,D} C_{G,X}\sqrt{T})$, $C_4 := C_{3,\log}\sqrt{T}$, *and*

$$\begin{aligned}
C_3 &:= C_\alpha \|u_0\|_{L^p(\Omega;Y)} + C_{2,\alpha} T + C_{3,\alpha}\sqrt{T},\\
C_{2,\alpha} &:= \frac{C_{F,X} L_3 + C_{\alpha,F}}{\alpha+1} + \left(\frac{2C_Y}{\alpha+1} + C_\alpha\right)\big(L_{F,Y} K_{u_0,f,g,Y} + \|f\|_{p,\infty,Y}\big),\\
C_{3,\alpha} &:= \frac{C_{p,D}}{\sqrt{2\alpha+1}}\Big(C_{G,X} L_3 + C_{\alpha,G} + 2C_Y\big(L_{G,Y} K_{u_0,f,g,Y} + \|\!|g|\!\|_{p,\infty,Y}\big)\Big),\\
C_{3,\log} &:= K_{p,D} C_\alpha \big(L_{G,Y} K_{u_0,f,g,Y} + \|\!|g|\!\|_{p,\infty,Y}\big),
\end{aligned}$$

where L_1, L_2, L_3 *are as defined in Lemma 4.22,* $K_{u_0,f,g,Y}$ *as in* (4.4.7), $K_{p,D}$ *as in Proposition 2.62,* C_Y *denotes the embedding constant of* Y *into* $D_A(\alpha,\infty)$ *or* $\mathrm{D}(A)$, *and* $C_{p,D}$ *is the constant from Theorem 2.61.*

In particular, the approximations $(U^j)_j$ *converge at rate* $\min\{\alpha, \frac{1}{2}\}$ *up to a logarithmic correction factor as* $k \to 0$.

This convergence result applies to schemes such as exponential Euler, implicit Euler, Crank–Nicolson, and other A-acceptable implicit Runge–Kutta methods such as Radau methods, BDF(2), Lobatto IIA, IIB, and IIC by virtue of Proposition 2.28. If R commutes with the resolvent of A, contractivity of R and S extend to fractional domain spaces and complex interpolation spaces. Hence, contractivity on Y often comes together with contractivity on X.

The constant C_{e} appears exponentially in the above. In the special case that $C_{F,X} = C_{G,X} = T = 1$, $p = 2$, and X is Hilbert, one can check that, similarly to Theorem 4.14, this yields the numerically reasonable value $C_{\mathrm{e}} = \sqrt{10}\mathrm{e}^5 \leq 470$.

If R commutes with the resolvent of A, contractivity of R and S extend to fractional domain spaces and complex interpolation spaces. Hence, contractivity on Y often comes together with contractivity on X. We recall that contractivity of a large class of schemes can be deduced from Proposition 2.28.

Proof. The assumptions of Theorems 4.12 and 4.14 hold, and thus the mild solution U exists and is pathwise continuous on both X and Y, i.e., the bound (4.4.4) holds.

By definition, $U(t_j) = U^j = u_0$ for $j = 0$. Let $N \in \{1, \dots, N_k\}$. Using the discrete variation-of-constants formula (4.4.6), the discretisation error can be split into the three parts

$$\begin{aligned}
E(N) &:= \Big\| \max_{1\le j\le N} \|U(t_j) - U^j\| \Big\|_p \\
&\le \Big\| \max_{1\le j\le N} \|(S(t_j) - R_k^j)u_0\| \Big\|_p \\
&\quad + \Big\| \max_{1\le j\le N} \Big\| \int_0^{t_j} S(t_j - s)F(s, U(s))\,\mathrm{d}s - k\sum_{i=0}^{j-1} R_k^{j-i} F(t_i, U^i) \Big\| \Big\|_p \\
&\quad + \Big\| \max_{1\le j\le N} \Big\| \int_0^{t_j} S(t_j - s)G(s, U(s))\,\mathrm{d}W_H(s) - \sum_{i=0}^{j-1} R_k^{j-i} G(t_i, U^i)\Delta W_{i+1} \Big\| \Big\|_p \\
&=: M_1 + M_2 + M_3
\end{aligned}$$

where $M_\ell = M_\ell(N)$ for $1 \le \ell \le 3$. Since R approximates S to order α on Y, we obtain with the dominated convergence theorem in $L^p(\Omega)$ that

$$M_1 \le C_\alpha k^\alpha \|u_0\|_{L^p(\Omega;Y)}. \tag{4.4.8}$$

To shorten the notation for the discrete terms, we introduce the piecewise constant functions $F^k(s) := F(t_i, U^i)$ and $G^k(s) := G(t_i, U^i)$ for $s \in [t_i, t_{i+1}), 0 \le i \le N_k - 1$ as well as $S^k(s) := R_k^i$ for $s \in (t_{i-1}, t_i], 1 \le i \le N_k$. This allows us to rewrite

$$\begin{aligned}
M_2 &= \Big\| \max_{1\le j\le N} \Big\| \int_0^{t_j} S(t_j - s)F(s, U(s)) - S^k(t_j - s)F^k(s)\,\mathrm{d}s \Big\| \Big\|_p \\
&\le \Big\| \max_{1\le j\le N} \sum_{i=0}^{j-1} \int_{t_i}^{t_{i+1}} \max_{i<\ell\le N} \big\|S(t_\ell - s)F(s, U(s)) - S^k(t_\ell - s)F^k(s)\big\|\,\mathrm{d}s \Big\|_p \\
&\le \Big\| \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \max_{i<\ell\le N} \big\|S(t_\ell - s)[F(s, U(s)) - F(s, U(t_i))]\big\|\,\mathrm{d}s \Big\|_p \\
&\quad + \Big\| \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \max_{i<\ell\le N} \big\|S(t_\ell - s)[F(s, U(t_i)) - F(t_i, U(t_i))]\big\|\,\mathrm{d}s \Big\|_p \\
&\quad + \Big\| \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \max_{i<\ell\le N} \big\|S(t_\ell - s)[F(t_i, U(t_i)) - F(t_i, U^i)]\big\|\,\mathrm{d}s \Big\|_p \\
&\quad + \Big\| \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \max_{i<\ell\le N} \big\|[S(t_\ell - s) - S^k(t_\ell - s)]F^k(s)\big\|\,\mathrm{d}s \Big\|_p \\
&=: M_{2,1} + M_{2,2} + M_{2,3} + M_{2,4}.
\end{aligned}$$

Making use of Minkowski's inequality in $L^p(\Omega)$, contractivity of $(S(t))_{t\ge0}$, and Lipschitz continuity of $\tilde{F}$, we derive the bound

$$M_{2,3} \leq C_{F,X} \sum_{i=0}^{N-1} \Big\| \int_{t_i}^{t_{i+1}} \|U(t_i) - U^i\| \,\mathrm{d}s \Big\|_p \leq C_{F,X} k \sum_{i=0}^{N-1} E(i) \tag{4.4.9}$$

for $M_{2,3}$. Proceeding likewise for $M_{2,1}$, we obtain from Lemma 4.22 that

$$\begin{aligned} M_{2,1} &\leq C_{F,X} \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} (\mathbb{E}\|U(s) - U(t_i)\|^p)^{1/p} \,\mathrm{d}s \\ &\leq C_{F,X} \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} L_1(s-t_i) + L_2(s-t_i)^{1/2} + L_3(s-t_i)^{\alpha} \,\mathrm{d}s \\ &\leq C_{F,X} \sum_{i=0}^{N-1} \Big(\frac{L_1}{2} k^2 + \frac{2L_2}{3} k^{3/2} + \frac{L_3}{\alpha+1} k^{\alpha+1} \Big) \\ &= C_{F,X} t_N \Big(\frac{L_1}{2} k + \frac{2L_2}{3} k^{1/2} + \frac{L_3}{\alpha+1} k^{\alpha} \Big). \end{aligned} \tag{4.4.10}$$

Analogously, uniform Hölder continuity yields

$$\begin{aligned} M_{2,2} &\leq \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \|F(s, U(t_i)) - F(t_i, U(t_i))\|_{L^p(\Omega;X)} \,\mathrm{d}s \\ &\leq \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} (s-t_i)^{\alpha} \,\mathrm{d}s \big\| [F(\cdot, U(t_i))]_{\alpha} \big\|_p \\ &\leq C_{\alpha,F} \sum_{i=0}^{N-1} \frac{k^{\alpha+1}}{\alpha+1} = \frac{C_{\alpha,F}}{\alpha+1} t_N k^{\alpha}. \end{aligned} \tag{4.4.11}$$

Using the semigroup bound (2.3.3) together with the assumed convergence rate α of R on Y, the linear growth assumption and stability as stated in (4.4.7), we obtain

$$\begin{aligned} M_{2,4} &\leq \Big\| \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \max_{i<\ell\leq N} \|[S(t_\ell - s) - S(t_\ell - t_i)] F(t_i, U^i)\| \,\mathrm{d}s \Big\|_p \\ &\quad + \Big\| \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \max_{i<\ell\leq N} \big\| \big[S(t_\ell - t_i) - R_k^{\ell-i}\big] F(t_i, U^i) \big\| \,\mathrm{d}s \Big\|_p \\ &\leq 2C_Y \sum_{i=0}^{N-1} \Big\| \int_{t_i}^{t_{i+1}} (s-t_i)^{\alpha} \|F(t_i, U^i)\|_Y \,\mathrm{d}s \Big\|_p + C_\alpha k^\alpha \sum_{i=0}^{N-1} \Big\| \int_{t_i}^{t_{i+1}} \|F(t_i, U^i)\|_Y \,\mathrm{d}s \Big\|_p \\ &\leq \Big(\frac{2C_Y}{\alpha+1} + C_\alpha \Big) k^{\alpha+1} \sum_{i=0}^{N-1} \big(L_{F,Y} \big(1 + \|U^i\|_{L^p(\Omega;Y)}\big) + \|f(t_i)\|_{L^p(\Omega;Y)} \big) \\ &\leq \Big(\frac{2C_Y}{\alpha+1} + C_\alpha \Big) \big(L_{F,Y} K_{u_0,f,g,Y} + \|f\|_{p,\infty,Y} \big) t_N k^{\alpha}. \end{aligned} \tag{4.4.12}$$

In conclusion from (4.4.9), (4.4.10), (4.4.11), and (4.4.12), M_2 is bounded by

$$M_2 \leq \frac{C_{F,X} L_1}{2} t_N k + \frac{2C_{F,X} L_2}{3} t_N k^{1/2} + C_{2,\alpha} t_N k^{\alpha} + C_{F,X} k \sum_{i=0}^{N-1} E(i)$$

$$\leq \frac{C_{F,X}L_1}{2}t_Nk + \frac{2C_{F,X}L_2}{3}t_Nk^{1/2} + C_{2,\alpha}t_Nk^{\alpha} + C_{F,X}\sqrt{t_N}\Big(k\sum_{i=0}^{N-1}E(i)^2\Big)^{1/2}, \quad (4.4.13)$$

where we have used the Cauchy–Schwarz inequality in the last line.

Let $\lfloor s \rfloor = \max\{t_i : t_i \leq s, 0 \leq i \leq N_k\}$ for $s \in [0,T]$. The remaining term M_3 can be split into

$$\begin{aligned}
M_3 &= \Big\| \max_{1\leq j\leq N} \Big\| \int_0^{t_j} S(t_j - s)G(s,U(s)) - S^k(t_j - s)G^k(s)\,\mathrm{d}W_H(s)\Big\|\Big\|_p \\
&\leq \Big\| \max_{1\leq j\leq N} \Big\| \int_0^{t_j} S(t_j - s)[G(s,U(s)) - G(s,U(\lfloor s \rfloor)]\,\mathrm{d}W_H(s)\Big\|\Big\|_p \\
&\quad + \Big\| \max_{1\leq j\leq N} \Big\| \int_0^{t_j} S(t_j - s)[G(s,U(\lfloor s \rfloor)) - G(\lfloor s \rfloor, U(\lfloor s \rfloor)]\,\mathrm{d}W_H(s)\Big\|\Big\|_p \\
&\quad + \Big\| \max_{1\leq j\leq N} \Big\| \int_0^{t_j} S(t_j - s)[G(\lfloor s \rfloor, U(\lfloor s \rfloor)) - G^k(s)]\,\mathrm{d}W_H(s)\Big\|\Big\|_p \\
&\quad + \Big\| \max_{1\leq j\leq N} \Big\| \int_0^{t_j} [S(t_j - \lfloor s \rfloor) - S(t_j - s)]G^k(s)\,\mathrm{d}W_H(s)\Big\|\Big\|_p \\
&\quad + \Big\| \max_{1\leq j\leq N} \Big\| \int_0^{t_j} [S(t_j - \lfloor s \rfloor) - S^k(t_j - s)]G^k(s)\,\mathrm{d}W_H(s)\Big\|\Big\|_p \\
&=: M_{3,1} + M_{3,2} + M_{3,3} + M_{3,4} + M_{3,5}.
\end{aligned}$$

We bound each term individually. Repeatedly, we use that the norm in $L^p(\Omega; L^2(0,t;Z))$ is bounded by the one in $L^2(0,t;L^p(\Omega;Z))$ for Banach spaces Z and $t \in [0,T]$. This indeed holds true due to

$$\begin{aligned}
\|h\|_{L^p(\Omega;L^2(0,t;Z))} &= \Big\|\Big(\int_0^t \|h(s)\|_Z^2\,\mathrm{d}s\Big)^{1/2}\Big\|_p = \Big\|\int_0^t \|h(s)\|_Z^2\,\mathrm{d}s\Big\|_{p/2}^{1/2} \\
&\leq \Big(\int_0^t \big\|\|h(s)\|_Z^2\big\|_{p/2}\,\mathrm{d}s\Big)^{1/2} = \Big(\int_0^t \|h(s)\|_{L^p(\Omega;Z)}^2\,\mathrm{d}s\Big)^{1/2}, \quad (4.4.14)
\end{aligned}$$

where we have used the triangle inequality in $L^{p/2}(\Omega)$ to estimate the norm of $h \in L^p(\Omega; L^2(0,t;Z))$. This can be seen as the continuous analogue of the estimate discussed in Remark 4.17. Together with an application of the maximal inequality from Theorem 2.61, the Lipschitz continuity of $\tilde{G}$, Lemma 4.22, and the triangle inequality in $L^2(0,t_N)$, this results in

$$\begin{aligned}
M_{3,1} &\leq \Big\| \sup_{t\in[0,t_N]} \Big\| \int_0^t S(t - s)[G(s,U(s)) - G(s,U(\lfloor s \rfloor)]\,\mathrm{d}W_H(s)\Big\|\Big\|_p \\
&\leq C_{p,D}\Big(\mathbb{E}\Big(\int_0^{t_N} \|G(s,U(s)) - G(s,U(\lfloor s \rfloor))\|_{\gamma(H,X)}^2\,\mathrm{d}s\Big)^{p/2}\Big)^{1/p} \\
&\leq C_{p,D}C_{G,X}\Big(\int_0^{t_N} (\mathbb{E}\|U(s) - U(\lfloor s \rfloor)\|^p)^{2/p}\,\mathrm{d}s\Big)^{1/2} \\
&\leq C_{p,D}C_{G,X}\Big(\int_0^{t_N} \big(L_1(s - \lfloor s \rfloor) + L_2(s - \lfloor s \rfloor)^{1/2} + L_3(s - \lfloor s \rfloor)^{\alpha}\big)^2\,\mathrm{d}s\Big)^{1/2}
\end{aligned}$$

$$
\begin{aligned}
&= C_{p,D} C_{G,X} \Big[L_1 \Big(\sum_{i=0}^{N-1} \frac{k^3}{3} \Big)^{1/2} + L_2 \Big(\sum_{i=0}^{N-1} \frac{k^2}{2} \Big)^{1/2} + L_3 \Big(\sum_{i=0}^{N-1} \frac{k^{2\alpha+1}}{2\alpha+1} \Big)^{1/2} \Big] \\
&\le C_{p,D} C_{G,X} \sqrt{t_N} \Big(\frac{L_1}{\sqrt{3}} k + \frac{L_2}{\sqrt{2}} k^{1/2} + \frac{L_3}{\sqrt{2\alpha+1}} k^{\alpha} \Big). \qquad (4.4.15)
\end{aligned}
$$

Again invoking the maximal inequality and (4.4.14), we estimate

$$
\begin{aligned}
M_{3,2} &\le C_{p,D} \Big(\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \big\| \| G(s, U(t_i)) - G(t_i, U(t_i)) \|_{\gamma(H,X)} \big\|_p^2 \, \mathrm{d}s \Big)^{1/2} \\
&\le C_{p,D} \Big(\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} (s - t_i)^{2\alpha} \, \mathrm{d}s \big\| [G(\cdot, U(t_i))]_\alpha \big\|_p^2 \Big)^{1/2} \le \frac{C_{p,D} C_{\alpha,G}}{\sqrt{2\alpha+1}} \sqrt{t_N} k^{\alpha} \qquad (4.4.16)
\end{aligned}
$$

from the uniform Hölder continuity of G. Proceeding analogously for $M_{3,3}$ and then applying Minkowski's inequality in $L^{p/2}(\Omega)$ results in

$$
\begin{aligned}
M_{3,3} &\le \Big\| \sup_{t\in[0,t_N]} \Big\| \int_0^t S(t-s)[G(\lfloor s \rfloor, U(\lfloor s \rfloor) - G^k(s)] \, \mathrm{d}W_H(s) \Big\| \Big\|_p \\
&\le C_{p,D} C_{G,X} \Big(\mathbb{E} \Big(k \sum_{i=0}^{N-1} \| U(t_i) - U^i \|^2 \Big)^{p/2} \Big)^{1/p} \\
&= C_{p,D} C_{G,X} \Big\| k \sum_{\ell=0}^{N-1} \max_{0\le j\le \ell} \| U(t_j) - U^j \|^2 \Big\|_{p/2}^{1/2} \\
&\le C_{p,D} C_{G,X} \Big(k \sum_{\ell=0}^{N-1} \Big\| \max_{0\le j\le \ell} \| U(t_j) - U^j \|^2 \Big\|_{p/2} \Big)^{1/2} = C_{p,D} C_{G,X} \Big(k \sum_{\ell=0}^{N-1} E(\ell)^2 \Big)^{1/2}. \qquad (4.4.17)
\end{aligned}
$$

Since R is contractive on Y by assumption, the conditions of Proposition 4.19 are fulfilled not only on X but also on Y. Thus, we can apply the estimate (4.4.7). Together with the maximal inequality, the semigroup difference bound (2.3.3), Remark 4.17, the ideal property of $\gamma(H, X)$ as in Proposition 2.39, and linear growth of $\tilde{G}$, this yields

$$
\begin{aligned}
M_{3,4} &\le \Big\| \sup_{t\in[0,t_N]} \Big\| \int_0^t S(t-s)[S(s - \lfloor s \rfloor) - I] G^k(s) \, \mathrm{d}W_H(s) \Big\| \Big\|_p \\
&\le C_{p,D} \Big(\mathbb{E} \Big(\int_0^{t_N} \big\| [S(s - \lfloor s \rfloor) - I] G^k(s) \big\|_{\gamma(H,X)}^2 \, \mathrm{d}s \Big)^{p/2} \Big)^{1/p} \\
&\le 2 C_{p,D} C_Y \Big(\mathbb{E} \Big(\sum_{\ell=0}^{N-1} \int_{t_\ell}^{t_{\ell+1}} (s - t_\ell)^{2\alpha} \, \mathrm{d}s \, \| G(t_\ell, U^\ell) \|_{\gamma(H,Y)}^2 \Big)^{p/2} \Big)^{1/p} \\
&\le 2 C_{p,D} C_Y \Big(\sum_{\ell=0}^{N-1} \frac{k^{2\alpha+1}}{2\alpha+1} \Big)^{1/2} \Big\| \max_{0\le j\le N-1} \| \tilde{G}(t_j, U^j) + g(t_j) \|_{\gamma(H,Y)} \Big\|_p \\
&\le \frac{2 C_{p,D} C_Y}{\sqrt{2\alpha+1}} \big(L_{G,Y} K_{u_0,f,g,Y} + \| g \|_{p,\infty,Y} \big) \sqrt{t_N} k^{\alpha}. \qquad (4.4.18)
\end{aligned}
$$

Applying Proposition 2.62 with $\Phi_s^{(j)} = \sum_{i=0}^{j-1} \mathbf{1}_{[t_i,t_{i+1})}(s)[S(t_j-t_i) - R_k^{j-i}]G(t_i,U^i)$ to the remaining term, we conclude that

$$\begin{aligned}
M_{3,5} &= \Big(\mathbb{E} \max_{1\le j\le N} \Big\| \int_0^{t_j} \sum_{i=0}^{j-1} \mathbf{1}_{[t_i,t_{i+1})}(s)[S(t_j-t_i) - R_k^{j-i}]G(t_i,U^i)\,\mathrm{d}W_H(s)\Big\|^p\Big)^{1/p} \\
&\le K_{p,D}\sqrt{\max\{\log(N),p\}} \Big\| \Big(\sum_{\ell=0}^{N-1} \int_{t_\ell}^{t_{\ell+1}} \Big(\max_{1\le j\le N} \| [S(t_j-t_\ell) - R_k^{j-\ell}]G(t_\ell,U^\ell)\|_{\gamma(H,X)}\Big)^2 \,\mathrm{d}s\Big)^{1/2} \Big\|_p \\
&\le K_{p,D}\sqrt{\max\{\log(N),p\}} \Big\| \Big(\sum_{\ell=0}^{N-1} k \big(C_\alpha k^\alpha \|G(t_\ell,U^\ell)\|_{\gamma(H,Y)}\big)^2\Big)^{1/2} \Big\|_p \\
&\le K_{p,D} C_\alpha \sqrt{t_N}\sqrt{\max\{\log(N),p\}} k^\alpha \Big\| \max_{0\le j\le N-1} \|G(t_j,U^j)\|_{\gamma(H,Y)} \Big\|_p \\
&\le K_{p,D} C_\alpha \big(L_{G,Y} K_{u_0,f,g,Y} + \|\!|g|\!\|_{p,\infty,Y}\big)\sqrt{t_N}\sqrt{\max\{\log(N),p\}} k^\alpha \qquad (4.4.19)
\end{aligned}$$

using that R approximates S to order α on Y, the ideal property of $\gamma(H,X)$, linear growth, and stability of R on Y as in (4.4.7). Combining the bounds (4.4.15) to (4.4.19), we deduce

$$\begin{aligned}
M_3 \le{}& \frac{1}{\sqrt{3}} C_{p,D} C_{G,X} L_1 \sqrt{t_N} k + \frac{1}{\sqrt{2}} C_{p,D} C_{G,X} L_2 \sqrt{t_N} k^{1/2} + C_{3,\alpha}\sqrt{t_N} k^\alpha \\
&+ C_{3,\log}\sqrt{t_N}\sqrt{\max\{\log(N),p\}} k^\alpha + C_{p,D} C_{G,X} \Big(k \sum_{\ell=0}^{N-1} E(\ell)^2 \Big)^{1/2}. \qquad (4.4.20)
\end{aligned}$$

Having bounded each term individually in (4.4.8), (4.4.13), and (4.4.20), we conclude

$$E(N) \le C_1 k + C_2 k^{1/2} + C_3 k^\alpha + C_4 \sqrt{\max\{\log(N_k),p\}} k^\alpha + C\Big(k \sum_{\ell=0}^{N-1} E(\ell)^2\Big)^{1/2},$$

noting that $N \le N_k$ and $t_N \le T$. Thus, by the discrete version of Gronwall's Lemma 2.68

$$E(N) \le (1 + C^2 t_N)^{1/2} \mathrm{e}^{(1+C^2 t_N)/2} \Big(C_1 k + C_2 k^{1/2} + C_3 k^\alpha + C_4 \sqrt{\max\{\log(N_k),p\}} k^\alpha \Big)$$

follows. The desired error estimate is obtained for $N = N_k$. As $k \to 0$, the terms with the lowest exponents dominate, i.e.

$$E(N_k) \lesssim k^{1/2} + k + \sqrt{\max\{\log(N_k),p\}} k^\alpha \lesssim \sqrt{\max\{\log(N_k),p\}} k^{\min\{\frac{1}{2},\alpha\}}, \quad (k \to 0). \quad \square$$

Remark 4.25. The result [32, Thm. 1.1] combines Hölder regularity in the p-th moment and bounds on the pointwise strong error to obtain a uniform strong error. Their effective method is based on a sophisticated application of the Kolmogorov–Chentsov continuity theorem, as well as approximation arguments. Let us refer to this method for obtaining uniform strong error estimates as the *Kolmogorov–Chentsov method*. At first sight, one might think that the result can be used to obtain the convergence rate of Theorem 4.24 up to an arbitrary $\varepsilon > 0$. Below, we point out what can precisely be achieved via their method.

Suppose that R approximates S to order $1/2$, a pointwise strong error estimate of rate $1/2$ has already been established, and Assumption 4.21 holds for fixed $p \in [2,\infty)$ and

$\alpha = 1/2$. This means that the fixed data (u_0, f, g) is assumed to have certain $L^p(\Omega)$-integrability. We will check what type of rate the Kolmogorov–Chentsov method yields for

$$\mathrm{E}_k^{p,\infty} := \Big\| \max_{0\le j\le N_k} \|U(t_j) - U^j\| \Big\|_p,$$

and compare it to the rate $\mathrm{E}_k^{p,\infty} \le C_p k^{1/2}\sqrt{\log(T/k)}$ we obtained in Theorem 4.24. We distinguish between three cases.

(i) Integrability of data in $L^2(\Omega)$: In this case, the Kolmogorov–Chentsov method does not apply, so no convergence rate is obtained.

(ii) Integrability of data in $L^p(\Omega)$ for a fixed $p \in (2,\infty)$: the Kolmogorov–Chentsov method gives $\mathrm{E}_k^{p,\infty} \le C_{\gamma,p} k^{\gamma-1/p}$ for any $\gamma \in (1/p, 1/2)$.

(iii) Integrability of data in $L^p(\Omega)$ for all $p \in (2,\infty)$: the Kolmogorov–Chentsov method gives $\mathrm{E}_k^{p,\infty} \le C_{\gamma,p} k^{\gamma}$ for any $\gamma \in (0, 1/2)$.

In the last case, there is an arbitrarily small difference in the error rate. We can obtain this error rate under the assumption that the data is $L^p(\Omega)$-integrable for a fixed $p \in [2,\infty)$. In the case one has this for all $p < \infty$, one needs to choose a very large p in the Kolmogorov–Chentsov method to get close to the desired rate, which in turn produces large constants in the rate estimate.

4.4.2 Exponential Euler method

We analyse the time discretisation error for the special case $R_k := S(k)$ known as the *exponential Euler method* or *splitting scheme*. Trivially, the exponential Euler method is contractive for contraction semigroups. Furthermore, several terms in the error analysis vanish for exponential Euler, since $S(t_j) - R_k^j = S(t_j) - S(k)^j = 0$ for all $0 \le j \le N_k$ by the semigroup property. Consequently, the logarithmic correction factor can be omitted for this scheme.

Corollary 4.26 (Exponential Euler). *Suppose that Assumption 4.21 holds for some $\alpha \in (0,1]$ and $p \in [2,\infty)$. Let $-A$ be the generator of a C_0-contraction semigroup $(S(t))_{t\ge0}$ on both X and Y. Suppose that $Y \hookrightarrow D_A(\alpha,\infty)$ continuously if $\alpha \in (0,1)$ or $Y \hookrightarrow \mathrm{D}(A)$ continuously if $\alpha = 1$. Let $u_0 \in L^p_{\mathcal{F}_0}(\Omega; Y)$. Consider the exponential Euler method $R := S$ for time discretisation. Denote by U the mild solution of (4.4.1) and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in (4.4.5). Then for $N_k \ge 2$*

$$\Big\| \max_{0\le j\le N_k} \|U(t_j) - U^j\| \Big\|_p \le C_{\mathrm{S,e}}\big(C_{\mathrm{S},1} k + C_{\mathrm{S},2} k^{1/2} + C_{\mathrm{S},3} k^{\alpha}\big)$$

with constants $C_{\mathrm{S,e}} := C_{\mathrm{e}}$, $C_{\mathrm{S},1} := C_1$, $C_{\mathrm{S},2} := C_2$ as in Theorem 4.24, $C_{\mathrm{S},3} := C_{\mathrm{S},2,\alpha} T + C_{\mathrm{S},3,\alpha} T^{1/2}$, $C_{\mathrm{S},3,\alpha} := C_{3,\alpha}$, and

$$C_{\mathrm{S},2,\alpha} := \frac{1}{\alpha+1}\big(C_{F,X} L_3 + C_{\alpha,F} + 2C_Y\big(L_{F,Y} K_{u_0,f,g,Y} + \|f\|_{p,\infty,Y}\big)\big),$$

where $C_{\mathrm{e}}, C_1, C_2, C_{3,\alpha}$ are as defined in Theorem 4.24, L_3 as in Lemma 4.22, $K_{u_0,f,g,Y}$ as in (4.4.7), C_Y denotes the embedding constant of Y into $D_A(\alpha,\infty)$ or $D(A)$, and $C_{p,D}$ is the constant from Theorem 2.61.

In particular, the approximations $(U^j)_j$ converge at rate $\min\{\alpha, \frac{1}{2}\}$ as $k \to 0$.

Proof. Adopt the notation from the proof of Theorem 4.24. Contractivity of R on X and Y is immediate from contractivity of S on these spaces. Since $S(t_j) - R_k^j = 0$ for any $j \in \{0, \ldots, N_k\}$, the terms M_1 and $M_{3,5}$ vanish. Moreover, the second term in $M_{2,4}$ vanishes so that

$$M_{2,4} \leq \frac{2C_Y}{\alpha+1}\big(L_{F,Y} K_{u_0,f,g,Y} + \|f\|_{p,\infty,Y}\big) t_N k^{\alpha}.$$

Combining the individual bounds for the remaining terms, the estimate follows from a discrete Gronwall argument, as in the proof of Theorem 4.24. The logarithmic correction factor vanishes due to $M_{3,5} = 0$. □

4.4.3 Error estimates on the full time interval

In this subsection, we will extend the error estimates of Theorem 4.24 and Corollary 4.26 to the full time interval by using suitable Hölder regularity of the paths of the mild solution. Since the proof of the Hölder regularity of the paths relies on a dilation argument, this extension is only valid for Hilbert spaces X.

Example 4.27. *Fix $N \geq 1$. Below, we construct a process $v_N : [0,1] \times \Omega \to \mathbb{R}$ such that $\sup_{t\in[0,1]} \mathbb{E}|v_N(t)|^p \leq 1/N$, but $v_N(t) = 1$ for all t in a neighborhood of $\{i/N : i \in \{1, \ldots, N\}\}$. This shows that information on the pointwise strong error does not provide much insight on the path of v_N in general.*

Indeed, let $\Omega = \{\omega_{m,i} : i \in \{1, \ldots, N\}, m \in \mathbb{N}\}$. For every $i \in \{1, \ldots, N\}$ suppose that $\mathbb{P}(\omega_{m,i}) = \frac{2^{-m}}{N}$. Let $I_N = \bigcup_{m\geq 1} \bigcup_{i=1}^{N} \{\omega_{m,i}\} \times (\frac{i}{N} - \frac{1}{2N}, \frac{i}{N} + \frac{1}{2N})$, and set $v_N(\omega, t) = 1$ if $(\omega, t) \in I_N$. Then one can check that v_N satisfies the required estimates.

The undesired behavior in the above example shows the need for having maximal estimates on the full time interval, i.e. estimates for $\| \sup_{t\in[0,T]} \|U(t) - \tilde{U}(t)\| \|_p$, where $\tilde{U}$ is the process obtained from the discrete approximation using piecewise linear interpolation.

The following simple deterministic result provides a way to connect the uniform error on the interval $[0,T]$ to the error on the grid. It requires generalised Hölder spaces $C^{\Phi}([0,T]; X)$ for some $\Phi\colon [0,T] \to [0,\infty)$, as introduced in Section 2.9. We shall be particularly interested in the function $\Phi(r) = r^{\alpha}(1 + \log(\frac{T}{r}))^{1/2}$ for $r \in (0,T]$ for some $\alpha > 0$ and $\Phi(0) = 0$ in the following.

Lemma 4.28 (Decomposition of the error on the full time interval)**.** *Let $u \in C^{\Phi}([0,T]; X)$ for a non-decreasing function $\Phi\colon [0,T] \to [0,\infty)$ such that $\Phi \neq 0$ on $(0,T]$. Let $\Pi \subseteq [0,T]$ be a finite time grid, and denote by $\tilde{u}\colon \Pi \to X$ an approximation of u, which is extended to $[0,T]$ by setting $\tilde{u}(t) := \tilde{u}(\lfloor t \rfloor_{\Pi})$ for $t \notin \Pi$, where $\lfloor t \rfloor_{\Pi} := \max\{s \in \Pi : s \leq t\}$. Then*

$$\sup_{t\in[0,T]} \|u(t) - \tilde{u}(t)\| \leq \Phi(h) \cdot \|u\|_{C^{\Phi}([0,T];X)} + \sup_{t\in\Pi} \|u(t) - \tilde{u}(t)\|$$

for the maximal time step $h := \sup_{t\in[0,T]}(t - \lfloor t \rfloor_{\Pi}) \leq 2 \sup_{t\in[0,T]} \operatorname{dist}(t, \Pi)$.

Proof. For $t \in [0,T]$, we can write

$$\begin{aligned}\|u(t) - \tilde{u}(t)\| &\leq \|u(t) - u(\lfloor t \rfloor_{\Pi})\| + \|u(\lfloor t \rfloor_{\Pi}) - \tilde{u}(t)\| \\ &\leq \|u\|_{C^{\Phi}([0,T];X)} \cdot \Phi(t - \lfloor t \rfloor_{\Pi}) + \sup_{s\in\Pi} \|u(s) - \tilde{u}(s)\|,\end{aligned}$$

which implies the desired result. □

From the above, we see that in order to estimate the uniform error on $[0,T]$, we need an (optimal) Hölder regularity result for the mild solution U to (4.4.1). To obtain such a result, the main difficulty lies in estimating the stochastic convolution. We point out that $\gamma(H,X) = \mathcal{L}_2(H,X)$, since X is assumed to be a Hilbert space.

Lemma 4.29 (Path regularity of stochastic convolutions)**.** *Let X, Y be separable Hilbert spaces such that $Y \hookrightarrow X$ continuously. Let $-A$ be the generator of a C_0-contraction semigroup $(S(t))_{t\geq 0}$ on both X and Y. Suppose that $Y \hookrightarrow D_A(\alpha,\infty)$ holds for some $\alpha \in (0,1/2]$. Let $q \in (2,\infty]$ be such that $\frac{1}{2} - \frac{1}{q} = \alpha$ and let $2 \leq p < p_0 < \infty$. Suppose that*

$$g \in L^p(\Omega; L^2(0,T;\mathcal{L}_2(H,Y))) \cap L^{p_0}(\Omega; L^q(0,T;\mathcal{L}_2(H,X)))$$

and define $J_g\colon \Omega \times [0,T] \to X$ as the stochastic convolution

$$J_g(t) = \int_0^t S(t-s)g(s)dW_H(s).$$

Then one has $J_g \in L^p(\Omega; C^\Psi([0,T];X))$ for $\Psi\colon (0,T] \to (0,\infty), \Psi(r) := r^\alpha(1+\log(\frac{T}{r}))^{1/2}$ and there exist constants $C_p, C_{\alpha,p,p_0,T} \geq 0$ such that

$$\|J_g\|_{L^p(\Omega;C^\Psi([0,T];X))} \leq C_p \|\!|g|\!\|_{p,2,Y} + C_{\alpha,p,p_0,T}\|\!|g|\!\|_{p_0,q,X}.$$

By a simple rescaling argument, the result extends to quasi-contraction semigroups. Moreover, from the proof below, one can see that a certain Orlicz integrability in Ω is sufficient for g. Note that the above path regularity is optimal for $q=\infty$. Indeed, Lévy's modulus of continuity theorem [114, Thm. 2.7] for a scalar Brownian motion states that a.s.

$$\limsup_{h\downarrow 0} \sup_{t\in[0,1-h]} \frac{B(t+h)-B(t)}{\sqrt{2h\log(1/h)}} = 1,$$

which shows that Ψ cannot be replaced by a "better" function.

Proof of Lemma 4.29. For $0 \leq s < t \leq T$, we can write

$$\begin{aligned}\|J_g(t)-J_g(s)\| &\leq \Big\|(S(t-s)-I)\int_0^s S(s-r)g(r)\,\mathrm{d}W_H(r)\Big\| + \Big\|\int_s^t S(t-r)g(r)\,\mathrm{d}W_H(r)\Big\| \\ &=: T_1(t,s) + T_2(t,s).\end{aligned}$$

For T_1, we can write

$$T_1(t,s) \leq \|S(t-s)-I\|_{\mathcal{L}(Y,X)}\Big\|\int_0^s S(s-r)g(r)\,\mathrm{d}W_H(r)\Big\|_Y \leq 2C_Y(t-s)^\alpha\|J_g(s)\|_Y.$$

Therefore, by Theorem 2.61 (with constant B_p because Y is Hilbert), we obtain

$$\begin{aligned}\Big\|\sup_{0\leq s<t\leq T}\frac{T_1(t,s)}{\Psi(t-s)}\Big\|_p &\leq 2C_Y\Big\|\sup_{0\leq s<t\leq T}\frac{\|J_g(s)\|_Y}{(1+\log(\frac{T}{t-s}))^{1/2}}\Big\|_p \\ &\leq 2C_Y B_p \|\!|g|\!\|_{p,2,Y}.\end{aligned}$$

For T_2, we use the dilation result of [120, Thm. I.7.1] (cf. [63], Theorem 2.16). We can find a Hilbert space $\widetilde{X}$, a contractive injection $Q\colon X \to \widetilde{X}$, a contractive projection

$P\colon \widetilde{X} \to X$, and a unitary C_0-group $(G(t))_{t\in\mathbb{R}}$ on $\widetilde{X}$ such that $S(t) = PG(t)Q$ for $t \geq 0$. Thus, we can write

$$\begin{aligned} T_2(t,s) &= \Big\| \int_s^t PG(t-r)Qg(r)\,\mathrm{d}W_H(r) \Big\|_X \\ &\leq \Big\| \int_s^t G(-r)Qg(r)\,\mathrm{d}W_H(r) \Big\|_{\widetilde{X}} = \|I(t) - I(s)\|_{\widetilde{X}}, \end{aligned}$$

where $I(t) := \int_0^t G(-r)Qg(r)\,\mathrm{d}W_H(r)$. We recall the Besov–Orlicz spaces $B^{\alpha}_{\Phi_2,q}([0,T];X)$ from Section 2.9 and their continuous embedding into $C^{r^{\alpha}|\log r|^{1/2}([0,T];X)}$ (see [109, Formula (2.12)]). From this embedding and Theorem 2.65, we conclude

$$I \in L^p(\Omega; C^{|\cdot|^{\alpha}|\log(\cdot)|^{1/2}}([0,T];\tilde{X}))$$

and thus, by boundedness of $|\log(\cdot)|^{1/2}(1+\log(\frac{T}{\cdot}))^{-1/2}$ on $(0,T]$, $I \in L^p(\Omega; C^{\Psi}([0,T];\tilde{X}))$. Moreover, there are constants $c_{\alpha,T}, C_{\alpha,p,p_0,T} \geq 0$ such that

$$\begin{aligned} \|I\|_{L^p(\Omega;C^{\Psi}([0,T];\tilde{X}))} &\leq c_{\alpha,T}\|I\|_{L^p(\Omega;B^{\alpha}_{\Phi_2,\infty}(0,T;\widetilde{X}))} \\ &\leq C_{\alpha,p,p_0,T}\|G(-\cdot)Qg(\cdot)\|_{L^{p_0}(\Omega;L^q(0,T;\mathcal{L}_2(H,\widetilde{X})))} \\ &\leq C_{\alpha,p,p_0,T}\|g\|_{L^{p_0}(\Omega;L^q(0,T;\mathcal{L}_2(H,X)))}. \end{aligned}$$

It follows that

$$\Big\| \sup_{0\leq s<t\leq T} \frac{T_2(t,s)}{\Psi(t-s)} \Big\|_p \leq \|I\|_{L^p(\Omega;C^{\Psi}([0,T];\tilde{X}))} \leq C_{\alpha,p,p_0,T}|\!|\!|g|\!|\!|_{p_0,q,X}.$$

Now the required estimate follows by combining the estimates for T_1 and T_2 to obtain an estimate for $\|[J_g]_{C^{\Psi}([0,T];X)}\|_p$ and noting that by the maximal inequality from Theorem 2.61 and $Y \hookrightarrow X$,

$$\|J_g\|_{L^p(\Omega;L^{\infty}(0,T;X))} \leq B_p|\!|\!|g|\!|\!|_{p,2,X} \leq B_pC_Y|\!|\!|g|\!|\!|_{p,2,Y}. \qquad \square$$

Remark 4.30. For analytic semigroups on X, the result of Lemma 4.29 even holds if merely $g \in L^{p_0}(\Omega; L^q(0,T;\mathcal{L}_2(H,X)))$, and even $J_g \in L^p(\Omega; B^{\alpha}_{\Phi_2,\infty}(0,T;X))$ (see [109, Thm. 5.1]). In particular, the space Y and contractivity of S are not needed. We do not know if one can take $p_0 = p$ in Lemma 4.29, even in the analytic case. Also, we do not know if the above Besov regularity of J_g holds in the non-analytic case.

Sharp path regularity results such as the one of Lemma 4.29 play an important role in obtaining convergence rates for numerical schemes for SPDEs. In particular, recent other applications of [109] to numerics include [42, 100, 134, 135]. Below, we apply Lemma 4.29 to obtain additional information on the numerical approximation in the Kato setting, and it seems to be the first of its kind for hyperbolic equations.

After these preparations, we can now prove the required path regularity of the mild solution.

Proposition 4.31 (Path regularity of the mild solution)**.** *Let X, Y be separable Hilbert spaces such that $Y \hookrightarrow X$ continuously. Suppose that Assumption 4.21 holds for some*

$\alpha \in (0,1/2]$ and $p \in [2,\infty)$. Let $p_0 \in (p,\infty)$ and $q \in (2,\infty]$ be such that $\frac{1}{2} - \frac{1}{q} = \alpha$, and suppose that f, g, and u_0 additionally satisfy

$$f \in L^{p_0}(\Omega; L^1(0,T;X)), \quad g \in L^{p_0}(\Omega; L^q(0,T;\mathcal{L}_2(H,X))), \quad u_0 \in L^{p_0}_{\mathcal{F}_0}(\Omega;X) \cap L^p_{\mathcal{F}_0}(\Omega;Y).$$

Let $-A$ be the generator of a C_0-contraction semigroup $(S(t))_{t\geq 0}$ on both X and Y. Suppose that $Y \hookrightarrow D_A(\alpha,\infty)$ continuously. Let $\Psi\colon (0,T] \to (0,\infty)$ be given by $\Psi(r) := r^\alpha(1+\log(\frac{T}{r}))^{1/2}$. Then the mild solution to (4.4.1) satisfies $U \in L^p(\Omega; C^\Psi([0,T];X))$ and there exists a constant C depending on $(T,p,p_0,\alpha,\tilde{F},\tilde{G}, X, Y)$ such that

$$\begin{aligned}\|U\|_{L^p(\Omega;C^\Psi([0,T];X))} \leq C\big(1 &+ \|u_0\|_{L^p(\Omega;Y)} + \|f\|_{p,1,Y} + \|g\|_{p,2,Y} \\ &+ \|u_0\|_{L^{p_0}(\Omega;X)} + \|f\|_{p_0,1,X} + \|f\|_{p,\frac{1}{1-\alpha},X} + \|g\|_{p_0,q,X}\big).\end{aligned} \tag{4.4.21}$$

Almost sure pathwise continuity of f and g with values in X and Y as stated in Assumption 4.21 is not required for the statement of the proposition to hold. It is sufficient if f and g are sufficiently regular such that all norms on the left-hand side of (4.4.21) are finite.

Proof. First, we note that

$$\|U\|_{L^p(\Omega;C^\Psi([0,T];X))} \leq \|U\|_{L^p(\Omega;L^\infty(0,T;X))} + \|[U]_{C^\Psi([0,T];X)}\|_p$$

and by the a priori estimate from Proposition 4.13,

$$\|U\|_{L^p(\Omega;L^\infty(0,T;X))} \lesssim 1 + \|u_0\|_{L^p(\Omega;X)} + \|f\|_{p,1,X} + \|g\|_{p,2,X}.$$

Second, we consider the Hölder seminorm term remaining. The mild solution formula (4.4.2) yields an initial value term, a difference of deterministic convolutions, and a stochastic version of the latter. Namely,

$$\begin{aligned}\|[U]_{C^\Psi([0,T];X)}\|_p \leq \big\|[S(\cdot)u_0]_{C^\Psi([0,T];X)}\big\|_p &+ \Big\|\Big[\int_0^\cdot S(\cdot - r)F(r,U(r))\,\mathrm{d}r\Big]_{C^\Psi([0,T];X)}\Big\|_p \\ &+ \big\|[J_{G(\cdot,U(\cdot))}]_{C^\Psi([0,T];X)}\big\|_p.\end{aligned}$$

To the last term, we apply Lemma 4.29 and note that

$$\begin{aligned}\|G(\cdot,U(\cdot))\|_{p,2,Y} &\leq L_{G,Y}\sqrt{T}C_{u_0,f,g,Y} + \|g\|_{p,2,Y} \lesssim 1 + \|u_0\|_{L^p(\Omega;Y)} + \|f\|_{p,1,Y} + \|g\|_{p,2,Y}, \\ \|G(\cdot,U(\cdot))\|_{p_0,q,X} &\leq T^{1/q}\|G(\cdot,U(\cdot))\|_{p_0,\infty,X} + \|g\|_{p_0,q,X} \leq T^{1/q}C_{G,X}\tilde{C}_{u_0,f,g,X} + \|g\|_{p_0,q,X} \\ &\lesssim 1 + \|u_0\|_{L^{p_0}(\Omega;X)} + \|f\|_{p_0,1,X} + \|g\|_{p_0,q,X},\end{aligned}$$

where $\tilde{C}_{u_0,f,g,X}$ is defined as $C_{u_0,f,g,X}$ in (4.4.3) with p replaced by p_0 and we have estimated $\|g\|_{p_0,2,X}$ by $T^\alpha\|g\|_{p_0,q,X}$ via Hölder's inequality. For the initial value term, decay of the semigroup difference of order α follows from (2.3.3), yielding

$$\big\|[S(\cdot)u_0]_{C^\Psi([0,T];X)}\big\|_p \leq 2C_Y\|u_0\|_{L^p(\Omega;Y)}.$$

Analogously, we can bound

$$\Bigg\|\sup_{0\leq s<t\leq T}\frac{\|\int_0^s[S(t-r)-S(s-r)]F(r,U(r))\,\mathrm{d}r\|}{\Psi(t-s)}\Bigg\|_p$$

$$
\begin{aligned}
&\leq 2C_Y \|F(\cdot, U(\cdot))\|_{p,1,Y} \leq 2C_Y (T\|\tilde{F}(\cdot, U(\cdot))\|_{p,\infty,Y} + \|f\|_{p,1,Y}) \\
&\lesssim 1 + \|u_0\|_{L^p(\Omega;Y)} + \|f\|_{p,1,Y} + \vert\!\vert\!\vert g \vert\!\vert\!\vert_{p,2,Y}.
\end{aligned}
$$

For the remaining term, Hölder's inequality with exponent $1/\alpha$ and conjugate exponent $1/(1-\alpha)$ yields

$$
\begin{aligned}
\Big\| \int_s^t S(t-r)F(r, U(r))\,\mathrm{d}r \Big\| &\leq \int_s^t \|\tilde{F}(r, U(r))\|\,\mathrm{d}r + \|f\|_{L^1(s,t;X)} \\
&\leq (t-s)\|\tilde{F}(\cdot, U(\cdot))\|_{L^\infty(s,t;X)} + (t-s)^\alpha \|f\|_{L^{\frac{1}{1-\alpha}}(s,t;X)}
\end{aligned}
$$

almost surely for $0 \leq s < t \leq T$. Consequently,

$$
\begin{aligned}
&\Big\| \sup_{0\leq s<t\leq T} \frac{\| \int_s^t S(t-r)F(r, U(r))\,\mathrm{d}r\|}{\Psi(t-s)} \Big\|_p \\
&\leq \Big\| \sup_{0\leq s<t\leq T} \frac{(t-s)\|\tilde{F}(\cdot, U(\cdot))\|_{p,\infty,X} + (t-s)^\alpha \|f\|_{p,\frac{1}{1-\alpha},X}}{(t-s)^\alpha (1+\log(\frac{T}{t-s}))^{1/2}} \Big\|_p \\
&\lesssim 1 + \|u_0\|_{L^p(\Omega;X)} + \|f\|_{p,\frac{1}{1-\alpha},X} + \vert\!\vert\!\vert g \vert\!\vert\!\vert_{p,2,X},
\end{aligned}
$$

since $\frac{1}{1-\alpha} \in (1, 2]$ and thus $\|f\|_{p,1,X} \lesssim \|f\|_{p,\frac{1}{1-\alpha},X}$. Combining the bounds thus obtained results in the desired path regularity statement. □

Consequently, we can now "upgrade" Theorem 4.24 and Corollary 4.26 to estimates on the full time interval.

Theorem 4.32 (Uniform error on the full interval for general schemes)**.** *Suppose that Assumption 4.21 holds for some $\alpha \in (0, 1/2]$, $p \in [2, \infty)$, and separable Hilbert spaces X and Y. Let $-A$ be the generator of a C_0-contraction semigroup $(S(t))_{t\geq 0}$ on both X and Y. Let $(R_k)_{k>0}$ be a time discretisation scheme that is contractive on X and Y. Assume R approximates S to order α on Y. Suppose that $Y \hookrightarrow D_A(\alpha, \infty)$ continuously. Let $p_0 \in (p, \infty)$ and $q \in (2, \infty]$ be such that $\frac{1}{2} - \frac{1}{q} = \alpha$, and suppose that f, g, and u_0 have additional integrability as X-valued processes*

$$
f \in L^{p_0}(\Omega; L^1(0, T; X)), \quad g \in L^{p_0}(\Omega; L^q(0, T; \mathcal{L}_2(H, X))), \quad u_0 \in L^{p_0}_{\mathcal{F}_0}(\Omega; X) \cap L^p_{\mathcal{F}_0}(\Omega; Y).
$$

Denote by U the mild solution of (4.4.1) and by $(U^j)_{j=0,\ldots,N_k}$ the temporal approximations as defined in (4.4.5). Define the piecewise constant extension $\tilde{U}\colon [0, T] \to L^p(\Omega; X)$ by $\tilde{U}(t) \coloneqq U^j$ for $t \in [t_j, t_{j+1})$, $0 \leq j \leq N_k - 1$, and $\tilde{U}(T) \coloneqq U^{N_k}$. Then for all $N_k \geq 2$ there is a constant $C \geq 0$ depending on $(u_0, T, p, p_0, \alpha, F, G, X, Y)$ such that

$$
\Big\| \sup_{t\in[0,T]} \|U(t) - \tilde{U}(t)\| \Big\|_p \leq C\big(1 + \sqrt{\max\{\log(T/k), p\}}\big)k^\alpha.
$$

Proof. The error bound follows from applying Lemma 4.28 with $\Phi = (\cdot)^\alpha(1+\log(\frac{T}{\cdot}))^{1/2}$ in combination with Theorem 4.24 and Proposition 4.31 to bound the first and second term obtained from the proposition, respectively. □

Thus, we can conclude that Theorem 4.24 and Corollary 4.26 can be improved to a uniform error estimate on $[0,T]$ at the price of a slightly more restrictive integrability condition on f, g, and u_0. Moreover, for the exponential Euler method, an additional square root of a logarithmic factor appears. Recall from [108, Thm. 3] that already for SDEs the error has to grow at least as $\log(T/k)^{1/2}k^{1/2}$ for $k \to 0$. Therefore, for $\alpha = 1/2$, Theorem 4.32 gives the optimal convergence rate for any contractive scheme on a Hilbert space.

In the applications given below, we restrict ourselves to the uniform error estimate on the grid points. By the above result, these statements can be extended to the full interval $[0,T]$ with additionally the square root of a logarithmic factor in the case of exponential Euler by imposing extra integrability conditions on the data.

Remark 4.33. Provided that the semigroup $(S(t))_{t\geq 0}$ has a dilation, analogous results to Theorem 4.32 can be obtained on non-Hilbert 2-smooth Banach spaces. There are two main cases known in which semigroups on non-Hilbert spaces have a dilation: positive semigroups on L^p-spaces for $2 < p < \infty$ [46] and (analytic) semigroups whose generator admits an H^∞-calculus of angle less than $\frac{\pi}{2}$ [48].

4.4.4 Application to the Schrödinger equation

In this subsection, we reconsider the stochastic Schrödinger equation with a potential from Subsection 4.1.3, now with linear multiplicative noise

$$\begin{cases} \mathrm{d}u + \mathrm{i}\Delta u \,\mathrm{d}t = -\mathrm{i}Vu \,\mathrm{d}t - \mathrm{i}u \,\mathrm{d}W \quad \text{on } [0,T], \\ \qquad\quad u(0) = u_0 \end{cases} \tag{4.4.22}$$

and its nonlinear variant with $\phi\colon \mathbb{C} \to \mathbb{C}$ and $\psi\colon \mathbb{C} \to \mathbb{C}$,

$$\begin{cases} \mathrm{d}u + \mathrm{i}\Delta u \,\mathrm{d}t = -\mathrm{i}(Vu + \phi(u)) \,\mathrm{d}t - \mathrm{i}\psi(u) \,\mathrm{d}W \quad \text{on } [0,T], \\ \qquad\quad u(0) = u_0 \end{cases} \tag{4.4.23}$$

in $\mathbb{R}^d$ for $d \in \mathbb{N}$, with a Q-Wiener process $(W(t))_{t\geq 0}$, potential V, and initial value u_0 as introduced in Subsection 4.1.3.

Let $\sigma \geq 0$ and, for this subsection only, write $L^2 = L^2(\mathbb{R}^d;\mathbb{C})$ and $H^\sigma = H^\sigma(\mathbb{R}^d;\mathbb{C})$. We recall that the well-posedness of (4.1.12) required Assumption 4.5 on $\tilde{\sigma} \geq 0$ and $d \in \mathbb{N}$ to hold so that multiplication by V is a bounded operator on the underlying space $H^{\tilde{\sigma}}$. For multiplicative noise, this assumption is also required to hold on $Y = H^{\sigma+\ell\alpha}$, where the choice of ℓ depends on the scheme employed. To facilitate checking the assumptions on Y, we use the following equivalent reformulation of Assumption 4.5:

Assumption 4.34. *Let $\sigma \geq 0$, $d \in \mathbb{N}$ and $V \in L^2$ such that*

(i) $\sigma > \frac{d}{2}$ and $V \in H^\sigma$, or

(ii) $\sigma = 0$ and $V \in H^\beta$ for some $\beta > \frac{d}{2}$, or

(iii) $d = 1$, $\sigma \in (0, \frac{1}{2})$, and $V \in H^\beta$ for some $\beta > \frac{1}{2}$

(iv) $d \geq 2$, $\sigma \in (0,1]$, and $V \in H^\beta$ for some $\beta > \frac{d}{2}$.

Based on the combination of the different cases of Assumption 4.34 for $X = H^\sigma$ and $Y = H^{\sigma+\ell\alpha}$, the following assumption emerges.

Assumption 4.35. *Let $\sigma \geq 0$, $d \in \mathbb{N}$, $\alpha \in \big(0, \frac{1}{2}\big]$, $\ell \in (0,\infty)$, $V \in H^\beta$ for some $\beta > 0$ such that*

(i) $\sigma > \frac{d}{2}$ *and* $\beta = \sigma + \ell\alpha$, *or*

(ii) $\sigma = 0$, $1 \leq d < \ell$, $\alpha > \frac{d}{8}$, *and* $\beta = \ell\alpha$, *or*

(iii) $\sigma = 0$, $d = 1$, $\alpha < \frac{1}{2\ell}$, *and* $\beta > \frac{1}{2}$, *or*

(iv) $\sigma = 0$, $d \geq 2$, $\alpha \leq \frac{1}{\ell}$, *and* $\beta > \frac{d}{2}$, *or*

(v) $d = 1$, $\sigma \in (0, \frac{1}{2})$, $\alpha > \frac{1-2\sigma}{2\ell}$, *and* $V \in H^{\sigma+\ell\alpha}$, *or*

(vi) $d = 1$, $\sigma \in (0, \frac{1}{2})$, $\alpha < \frac{1-2\sigma}{2\ell}$, *and* $\beta > \frac{1}{2}$, *or*

(vii) $2 \leq d < 2\sigma + \ell$, $\sigma \in (0, 1]$, $\alpha > \frac{d-2\sigma}{2\ell}$, *and* $\beta = \sigma + \ell\alpha$, *or*

(viii) $d \geq 2$, $\sigma \in (0, 1]$, $\alpha \leq \frac{1-\sigma}{\ell}$, *and* $\beta > \frac{d}{2}$.

For exponential Euler, we recover the error bound from [2, Thm. 5.5], showing convergence rate $\frac{1}{2}$ for linear noise in the case of sufficiently regular $Q^{1/2}$ and V as well as $\sigma > \frac{d}{2}$. Assuming less regularity of $Q^{1/2}$ and V, we extend their result to fractional convergence rates $\alpha \in (0, \frac{1}{2}]$ as well as the novel cases (ii)-(viii) of Assumption 4.35.

Theorem 4.36. *Suppose that Assumption 4.35 is satisfied for some $\ell \geq 2$ and some σ, d, V, β, and $\alpha \in (0, \frac{1}{2}]$. Further, assume that $Q^{1/2} \in \mathcal{L}_2(L^2, H^\beta)$. Let $p \in [2,\infty)$ and $u_0 \in L^p_{\mathcal{F}_0}(\Omega; H^{\sigma+\ell\alpha})$. Denote by U the mild solution of the linear stochastic Schrödinger equation* (4.4.22) *with multiplicative noise and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in* (4.4.5) *obtained with the exponential Euler method $R := S$. Then there exists a constant $C \geq 0$ depending on $(V, u_0, T, p, \alpha, \sigma, d, \ell)$ such that for $N_k \geq 2$*

$$\Big\| \max_{0 \leq j \leq N_k} \|U(t_j) - U^j\|_{H^\sigma} \Big\|_p \leq C\big(1 + \|Q^{1/2}\|_{\mathcal{L}_2(L^2,H^\beta)}\big) k^\alpha.$$

In particular, the approximations $(U^j)_j$ converge at rate $\frac{1}{2}$ as $k \to 0$ if $Q^{1/2} \in \mathcal{L}_2(L^2, H^{\sigma+1})$, $V \in H^{\sigma+1}$, $\sigma > \frac{d}{2}$, and $u_0 \in L^p_{\mathcal{F}_0}(\Omega; H^{\sigma+1})$.

Proof. By [2, Lem. 2.1], $-A = -\mathrm{i}\Delta$ generates a C_0-contraction semigroup on both Hilbert spaces $X = H^\sigma$ and $Y = H^{\sigma+\ell\alpha}$. Furthermore, setting $F(u) = -\mathrm{i}V \cdot u$ and $G(u) = -\mathrm{i}M_u Q^{1/2}$ for $u \in H^\sigma$ with the multiplication operator M_u allows us to rewrite (4.4.22) in the form of a stochastic evolution equation (4.4.1). It remains to verify the mapping, linear growth, and Lipschitz continuity conditions from Assumption 4.21.

Note that Assumption 4.35 implies that Assumption 4.5 from the additive linear case is satisfied for both σ and $\sigma + \ell\alpha$. In particular, this means that $Vu \in Y = H^{\sigma+\ell\alpha}$ for any $u \in H^{\sigma+\ell\alpha}$ and $\|Vu\|_{H^{\sigma+\ell\alpha}} \leq C_V \|u\|_{H^{\sigma+\ell\alpha}}$ for some constant $C_V \geq 0$. More specifically, it can be shown that $C_V \lesssim \|V\|_{H^\beta}$; cf. Subsection 4.1.3. An analogous statement holds with the same (or a smaller) constant on $X = H^\sigma$. Hence, F maps both X and Y into themselves, and it is of linear growth on Y because of

$$\|F(u)\|_Y = \| -\mathrm{i}V \cdot u\|_{H^{\sigma+\ell\alpha}} \le C_V \|u\|_{H^{\sigma+\ell\alpha}} = C_V\|u\|_Y, \quad u \in Y.$$

Likewise, Lipschitz continuity on X is obtained using the linearity of F.

Set $H = L^2$. Due to

$$\begin{aligned} \|G(u)\|_{\gamma(H,Y)} &= \| -\mathrm{i}M_u \cdot Q^{1/2}\|_{\mathcal{L}_2(L^2,H^{\sigma+\ell\alpha})} \\ &\le \|M_u\|_{\mathcal{L}(H^\beta,H^{\sigma+\ell\alpha})}\|Q^{1/2}\|_{\mathcal{L}_2(L^2,H^\beta)} \\ &\lesssim \|Q^{1/2}\|_{\mathcal{L}_2(L^2,H^\beta)}\|u\|_{H^{\sigma+\ell\alpha}} = \|Q^{1/2}\|_{\mathcal{L}_2(L^2,H^\beta)}\|u\|_Y, \quad u \in Y, \end{aligned} \tag{4.4.24}$$

G is of linear growth on Y. To see this, we estimate the operator norm of M_u from H^β to $H^{\sigma+\ell\alpha}$ using either the Banach algebra property of H^β, a combination of Hölder's inequality and Sobolev embeddings or an argument analogously to Lemma 4.6, as discussed in Subsection 4.1.3. Likewise, we check Lipschitz continuity of G on X with a multiple of $\|Q^{1/2}\|_{\mathcal{L}_2(L^2,H^\beta)}$ as Lipschitz constant using the linearity of G. Measurability and Hölder continuity in time are trivially fulfilled due to F and G depending only on $u \in X$. Thus, Corollary 4.26 is applicable with $X = H^\sigma$, $H = L^2$, and $Y = H^{\sigma+\ell\alpha} \hookrightarrow (H^\sigma, \mathrm{D}(A))_{\alpha,\infty}$, yielding the desired error bound. □

Furthermore, Theorem 4.24 enables us to extend [2, Thm. 5.5] to general discretisation schemes R other than exponential Euler at the price of an additional logarithmic factor. We focus on implicit Euler (IE) and Crank–Nicolson (CN), which approximate the Schrödinger semigroup to rate α on $Y = H^{\sigma+4\alpha}$ and $Y = H^{\sigma+3\alpha}$, respectively (see Theorem 4.8).

Theorem 4.37. *Let $\sigma \ge 0$, $d \in \mathbb{N}$, and $V \in L^2$. Let $(R_k)_{k>0}$ be the implicit Euler method (IE) or the Crank–Nicolson method (CN) and set $\ell_0 := 4$ or $\ell_0 := 3$, respectively. Suppose that Assumption 4.35 is satisfied for some $\ell \ge \ell_0$ and for some $\alpha \in \left(0, \frac{1}{2}\right]$, $\beta > 0$, and $p \in [2,\infty)$. Further, suppose that $u_0 \in L^p_{\mathcal{F}_0}(\Omega; H^{\sigma+\ell\alpha})$ as well as $Q^{1/2} \in \mathcal{L}_2(L^2, H^\beta)$. Denote by U the mild solution of the linear stochastic Schrödinger equation with multiplicative noise (4.4.22) and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in (4.4.5). Then there exists a constant $C \ge 0$ depending on $(V, u_0, T, p, \alpha, \sigma, d, \ell)$ such that for $N_k \ge 2$*

$$\left\| \max_{0\le j\le N_k} \|U(t_j) - U^j\|_{H^\sigma} \right\|_p \le C\big(1 + \|Q^{1/2}\|_{\mathcal{L}_2(L^2,H^\beta)}\big)\sqrt{\max\{\log(T/k), p\}}\,k^\alpha.$$

In particular, (IE) and (CN) converge at rate $\frac{1}{2}$ up to a logarithmic correction as $k \to 0$ if $V \in H^{\sigma+\ell\alpha}$, $Q^{1/2} \in \mathcal{L}_2(L^2, H^{\sigma+\ell\alpha})$, $\sigma > \frac{d}{2}$, and $u_0 \in L^p_{\mathscr{F}_0}(\Omega; H^{\sigma+\ell\alpha})$ with $\ell = 4$ and $\ell = 3$, respectively.

An analogous statement holds for all time discretisation schemes $(R_k)_{k>0}$ which are contractive on H^σ and $H^{\sigma+\ell\alpha}$ and approximate S to order α on $H^{\sigma+\ell\alpha}$. As in the additive case, the conditions on the dimension $d \in \mathbb{N}$ are not required in the absence of a potential. In most cases, choosing $\ell = \ell_0$ is sufficient. However, in the situation of Assumption 4.35(ii) or (vii), choosing a larger ℓ can yield the additional regularity required to solve Schrödinger's equation in higher dimensions.

Proof. We want to apply Theorem 4.24 with $Y = H^{\sigma+\ell\alpha}$ for $\ell \geq \ell_0 \in \{3,4\}$ and X, H, F, G as in Theorem 4.36 for exponential Euler. The proof works analogously, replacing $\ell \geq 2$ by $\ell \geq \ell_0$. It remains to check that (IE) and (CN) are contractive on H^σ and $H^{\sigma+\ell\alpha}$. But since (IE) and (CN) are defined via A and a scaled version of its resolvent, R_k commutes with resolvents of A in both cases. Thus, Proposition 2.28 yields the assertion. □

When passing to the nonlinear situation as in (4.4.23), showing Lipschitz continuity of G requires estimates of the form

$$\|\psi(u) - \psi(v)\|_{H^\sigma} \lesssim \|u - v\|_{H^\sigma}, \quad u, v \in H^\sigma,$$

and similarly for ϕ. However, the best estimate known for $\sigma \in (0,1)$ and $\psi \in C^2$ with bounded first and second derivatives [122, Prop. 2.7.2] is

$$\|\psi(u) - \psi(v)\|_{H^\sigma} \lesssim \|u - v\|_{H^\sigma} + (1 + \|u\|_{H^\sigma} + \|v\|_{H^\sigma})\|u - v\|_{L^\infty}.$$

Since this estimate is nonlinear in u and v, showing Lipschitz continuity of G on H^σ is currently out of reach for $\sigma > 0$. Another reason to restrict our considerations to $\sigma = 0$ in the following is the negative result from Dahlberg [40], see also the survey [17]. It states that for $\sigma + 2\alpha \in (\frac{3}{2}, 1 + \frac{d}{2})$, every mapping ψ such that $\psi \circ u \in H^{\sigma+2\alpha}$ for all $u \in H^{\sigma+2\alpha}$ is affine-linear. Hence, in dimension $d > 1$, the optimal rate $\alpha = \frac{1}{2}$ cannot be expected for all $\sigma > \frac{1}{2}$ for genuinely nonlinear ψ. In particular, Nemytskij maps are not Lipschitz on H^σ for any $\sigma > 0$. For $\sigma = 0$, however, a convergence rate can be obtained.

Theorem 4.38. *Suppose that one of the cases (ii)-(iv) of Assumption 4.35 is satisfied for for $\ell = 2$ and for some d, V, β, and $\alpha \in (0, \frac{1}{2}]$. Further, assume that $Q^{1/2} \in \mathcal{L}_2(L^2, H^\beta)$. Let $p \in [2,\infty)$ and $u_0 \in L^p_{\mathcal{F}_0}(\Omega; H^{2\alpha})$. Let $\phi, \psi\colon \mathbb{C} \to \mathbb{C}$ be Lipschitz continuous and satisfy $\phi(0) = \psi(0) = 0$. Denote by U the mild solution of the nonlinear stochastic Schrödinger equation (4.4.23) with multiplicative noise and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in (4.4.5) obtained with the exponential Euler method $R := S$. Then there exists a constant $C \geq 0$ depending on $(V, u_0, \phi, \psi, T, p, \alpha, d, \ell)$ such that for $N_k \geq 2$*

$$\left\| \max_{0 \leq j \leq N_k} \|U(t_j) - U^j\|_{L^2} \right\|_p \leq C\big(1 + \|Q^{1/2}\|_{\mathcal{L}_2(L^2, H^\beta)}\big) k^\alpha.$$

In particular, the approximations $(U^j)_j$ converge at rate $\frac{1}{2}$ as $k \to 0$ if $Q^{1/2} \in \mathcal{L}_2(L^2, H^1)$, $V \in H^1$, and $u_0 \in L^p_{\mathcal{F}_0}(\Omega; H^1)$ for $d = 1$. In dimension $d \geq 2$, this is attained for $Q^{1/2} \in \mathcal{L}_2(L^2, H^\beta)$, $V \in H^\beta$ for some $\beta > \frac{d}{2}$, and $u_0 \in L^p_{\mathcal{F}_0}(\Omega; H^1)$.

Proof. Let $X = H = L^2$, $Y = H^{2\alpha}$, and $G(u) = -\mathrm{i} M_{\psi \circ u} Q^{1/2}$ for $u \in L^2$. From the linear case, it is already clear that

$$\|G(u) - G(v)\|_{\mathcal{L}_2(L^2, L^2)} = \|M_{\psi\circ u - \psi \circ v} Q^{1/2}\|_{\mathcal{L}_2(L^2, L^2)} \lesssim \|\psi \circ u - \psi \circ v\|_{L^2} \|Q^{1/2}\|_{\mathcal{L}_2(L^2, H^\beta)}.$$

Lipschitz continuity of ψ with Lipschitz constant $C_\psi \geq 0$ implies Lipschitz continuity of G on $X = L^2$ via

$$\|\psi \circ u - \psi \circ v\|_{L^2} \|Q^{1/2}\|_{\mathcal{L}_2(L^2, H^\beta)} \leq C_\psi \|Q^{1/2}\|_{\mathcal{L}_2(L^2, H^\beta)} \|u - v\|_{L^2}.$$

Since from (4.4.24) we can infer

$$\|G(u)\|_{\mathcal{L}_2(L^2,H^{2\alpha})} \lesssim \|M_{\psi\circ u}\|_{\mathcal{L}(H^\beta,H^{2\alpha})}\|Q^{1/2}\|_{\mathcal{L}_2(L^2,H^\beta)} \lesssim \|\psi\circ u\|_{H^{2\alpha}}\|Q^{1/2}\|_{\mathcal{L}_2(L^2,H^\beta)}, \tag{4.4.25}$$

it remains to estimate the norm of the composition $\|\psi \circ u\|_{H^{2\alpha}}$ by a multiple of $\|u\|_{H^{2\alpha}}$ to show linear growth of G on $H^{2\alpha}$. In case $\alpha < \frac{1}{2}$, $2\alpha \in (0,1)$, and thus, by Proposition 2.33, $\|\psi\circ u\|_{H^{2\alpha}} \lesssim \|u\|_{H^{2\alpha}}$. In the remaining cases, $2\alpha = 1$ holds, whence

$$\|\psi\circ u\|_{H^{2\alpha}}^2 = \|\psi\circ u\|_{L^2}^2 + \|\nabla(\psi\circ u)\|_{L^2}^2 \le \|\psi\circ u\|_{L^2}^2 + C_\psi^2\|\nabla u\|_{L^2}^2 \le \max\{1, C_\psi^2\}\|u\|_{H^1}^2,$$

noting that $\operatorname{grad}(\psi\circ u) = \psi'(u)\operatorname{grad} u\mathbf{1}_{\{\psi\neq 0\}}$. In the first inequality, we have invoked Proposition 2.34. Hence, G is of linear growth on $Y = H^{2\alpha}$. In the same way, one can see that $F(u) = -\mathrm{i}(Vu + \phi(u))$ is Lipschitz on X and of linear growth on Y. Finally, the statement of the theorem follows by an application of Corollary 4.26. □

To estimate the composition in (4.4.25), we required $2\alpha \in (0,1]$ to apply the composition estimates. It is an open problem whether such estimates also hold in H^s for $s > 1$. For real-valued functions, results have been obtained for $s < \frac{3}{2}$ in [18, Thm. 18]. These estimates being unknown for $s > 1$ limits us to suboptimal convergence rates for schemes involving rational approximations, at least for nonlinear Schrödinger equations.

Theorem 4.39. *Let $\sigma = 0$, $d \in \mathbb{N}$, and $V \in L^2$. Let $(R_k)_{k>0}$ be the implicit Euler method (IE) or the Crank–Nicolson method (CN) and set $\ell_0 := 4$ or $\ell_0 := 3$, respectively. Suppose that one of the cases (ii)-(iv) of Assumption 4.35 is satisfied for $\ell = \ell_0$ and some $\alpha \in \left(0, \frac{1}{\ell}\right]$, $\beta > 0$, and $p \in [2,\infty)$. Further, suppose that $u_0 \in L^p_{\mathcal{F}_0}(\Omega; H^{\ell\alpha})$ as well as $Q^{1/2} \in \mathcal{L}_2(L^2, H^\beta)$. Let $\phi, \psi : \mathbb{C} \to \mathbb{C}$ be Lipschitz continuous and such that $\phi(0) = \psi(0) = 0$. Denote by U the mild solution of the nonlinear stochastic Schrödinger equation with multiplicative noise (4.4.23) and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in (4.4.5). Then there exists a constant $C \ge 0$ depending on $(V, u_0, \phi, \psi, T, p, \alpha, d, \ell)$ such that for $N_k \ge 2$*

$$\Big\|\max_{0\le j\le N_k}\|U(t_j) - U^j\|_{H^\sigma}\Big\|_p \le C\big(1 + \|Q^{1/2}\|_{\mathcal{L}_2(L^2,H^\beta)}\big)\sqrt{\max\{\log(T/k), p\}}k^\alpha.$$

In particular, in dimension $d = 1$, (IE) converges at rate $\frac{1}{4}$ up to logarithmic correction as $k \to 0$ if $V \in H^1$, $Q^{1/2} \in \mathcal{L}_2(L^2, H^1)$, and $u_0 \in L^p_{\mathscr{F}_0}(\Omega; H^1)$. For the same regularity of V, $Q^{1/2}$, and u_0, (CN) converges at rate $\frac{1}{3}$ up to logarithmic correction as $k \to 0$ in dimension $d = 1$.

This theorem can be generalised to time discretisation schemes $(R_k)_{k>0}$ that are contractive on L^2 and $H^{\ell\alpha}$, and that approximate S to order $\alpha \in (0, \frac{1}{\ell}]$ on $H^{\ell\alpha}$.

4.4.5 Application to Maxwell's equations

As a second example, we consider the stochastic Maxwell's equations

$$\begin{cases} \mathrm{d}U + AU\,\mathrm{d}t = F(U)\,\mathrm{d}t + G(U)\,\mathrm{d}W \quad \text{on } [0,T], \\ \qquad U(0) = (\mathbf{E}_0^\top, \mathbf{H}_0^\top)^\top \end{cases} \tag{4.4.26}$$

with boundary conditions of a perfect conductor as in [28]. It describes the behaviour of the electric and magnetic field $\mathbf{E}$ and $\mathbf{H}$, respectively, on a bounded, simply connected domain $\mathcal{O} \subseteq \mathbb{R}^3$ with smooth boundary, whose unit outward normal vector we denote by $\mathbf{n}$. Here, $A\colon \mathrm{D}(A) \to X := L^2(\mathcal{O})^6$ is the Maxwell operator defined by

$$A\begin{pmatrix}\mathbf{E}\\ \mathbf{H}\end{pmatrix} := \begin{pmatrix}0 & -\varepsilon^{-1}\nabla\times\\ \mu^{-1}\nabla\times & 0\end{pmatrix}\begin{pmatrix}\mathbf{E}\\ \mathbf{H}\end{pmatrix} = \begin{pmatrix}-\varepsilon^{-1}\nabla\times\mathbf{H}\\ \mu^{-1}\nabla\times\mathbf{E}\end{pmatrix}$$

on $\mathrm{D}(A) := H_0(\mathrm{curl},\mathcal{O}) \times H(\mathrm{curl},\mathcal{O})$ with $H(\mathrm{curl},\mathcal{O}) := \{\mathbf{H} \in (L^2(\mathcal{O}))^3 : \nabla\times\mathbf{H} \in L^2(\mathcal{O})^3\}$ and its subspace $H_0(\mathrm{curl},\mathcal{O})$ of those $\mathbf{H}$ with vanishing tangential trace $\mathbf{n}\times\mathbf{H}|_{\partial\mathcal{O}}$. The permittivity and permeability $\varepsilon,\mu \in L^\infty(\mathcal{O})$ are assumed to be uniformly positive, i.e., $\varepsilon,\mu \geq \kappa$ for some constant $\kappa > 0$. We equip the Hilbert space $X = L^2(\mathcal{O})^6 = L^2(\mathcal{O})^3 \times L^2(\mathcal{O})^3$ with the weighted scalar product

$$\left(\begin{pmatrix}\mathbf{E}_1\\ \mathbf{H}_1\end{pmatrix} \,\middle|\, \begin{pmatrix}\mathbf{E}_2\\ \mathbf{H}_2\end{pmatrix}\right)_X := \int_{\mathcal{O}} \big(\varepsilon(x)\,(\mathbf{E}_1(x)\,|\,\mathbf{E}_2(x)) + \mu(x)\,(\mathbf{H}_1(x)\,|\,\mathbf{H}_2(x))\big)\,\mathrm{d}x,$$

where $(\cdot\,|\,\cdot)$ denotes the standard scalar product in $\mathbb{R}^3$, and the corresponding induced norm. Furthermore, W is assumed to be a Q-Wiener process for a symmetric, non-negative operator Q with finite trace. Hence, $Q^{1/2} \in \mathcal{L}_2(H,X)$, where $H = L^2(\mathcal{O})^6$ is equipped with the standard norm.

As $F\colon \Omega\times[0,T]\times X \to X$, we consider the linear drift term given by

$$(\omega,t,V) \mapsto F(\omega,t,V) = \begin{pmatrix}\sigma_1(\cdot,t)\mathbf{E}_V\\ \sigma_2(\cdot,t)\mathbf{H}_V\end{pmatrix}, \quad V = (\mathbf{E}_V^\top, \mathbf{H}_V^\top)^\top \in X, \tag{4.4.27}$$

for sufficiently smooth $\sigma_1,\sigma_2\colon \mathcal{O}\times[0,T] \to \mathbb{R}$. More precisely, we assume boundedness of σ_1,σ_2, and their partial derivatives w.r.t. the spatial variables. In particular, let σ_j be uniformly Lipschitz continuous in time and let $\partial_{x_i}\sigma_j, \sigma_j \in L^\infty(\mathcal{O}\times[0,T])$ for $i = 1,2,3$ and $j = 1,2$. Then F is Lipschitz continuous on X due to linearity of F and

$$\begin{aligned}\|F(t,V)\|_X^2 &= \int_{\mathcal{O}} \big(\varepsilon(x)\|\sigma_1(x,t)\mathbf{E}_V(x)\|^2 + \mu(x)\|\sigma_2(x,t)\mathbf{H}_V(x)\|^2\big)\,\mathrm{d}x\\ &\leq \max\{\|\sigma_1\|_\infty, \|\sigma_2\|_\infty\}^2\|V\|_X^2 =: C_F^2\|V\|_X^2, \quad V = (\mathbf{E}_V^\top, \mathbf{H}_V^\top)^\top,\end{aligned}$$

with $\|\cdot\|$ denoting the standard Euclidean norm in $\mathbb{R}^3$. A straightforward explicit calculation of the curl operator shows that for $\mathbf{E} = (\mathbf{E}_1,\mathbf{E}_2,\mathbf{E}_3) \in L^2(\mathcal{O})^3$,

$$\nabla\times(\sigma_1(\cdot,t)\mathbf{E}) = \sigma_1(\cdot,t)(\nabla\times\mathbf{E}) + \begin{pmatrix}(\partial_{x_2}\sigma_1(\cdot,t))\mathbf{E}_3\\ (\partial_{x_3}\sigma_1(\cdot,t))\mathbf{E}_1\\ (\partial_{x_1}\sigma_1(\cdot,t))\mathbf{E}_2\end{pmatrix} - \begin{pmatrix}(\partial_{x_3}\sigma_1(\cdot,t))\mathbf{E}_2\\ (\partial_{x_1}\sigma_1(\cdot,t))\mathbf{E}_3\\ (\partial_{x_2}\sigma_1(\cdot,t))\mathbf{E}_1\end{pmatrix},$$

and thus

$$\begin{aligned}\|AF(t,V)\|_X^2 &= \left\|\begin{pmatrix}-\varepsilon^{-1}\nabla\times(\sigma_2(\cdot,t)\mathbf{H}_V)\\ \mu^{-1}\nabla\times(\sigma_1(\cdot,t)\mathbf{E}_V)\end{pmatrix}\right\|_X^2\\ &= \int_{\mathcal{O}} \varepsilon(x)\|\varepsilon^{-1}(x)(\nabla\times(\sigma_2(\cdot,t)\mathbf{H}_V))(x)\|^2 + \mu(x)\|\mu^{-1}(x)(\nabla\times(\sigma_1(\cdot,t)\mathbf{E}_V)(x))\|^2\,\mathrm{d}x\\ &\leq 3\max_{j=1,2}\|\sigma_j\|_\infty^2 \int_{\mathcal{O}} \big(\varepsilon(x)\|\varepsilon^{-1}(x)(\nabla\times\mathbf{H}_V)(x)\|^2 + \mu(x)\|\mu^{-1}(x)(\nabla\times\mathbf{E}_V)(x)\|^2\big)\,\mathrm{d}x\end{aligned}$$

$$
\begin{aligned}
&+ 6 \max_{j=1,2} \max_{i=1,2,3} \|\partial_{x_i} \sigma_j\|_\infty^2 \int_{\mathcal{O}} \Big(\frac{\varepsilon(x)}{\mu(x)\varepsilon(x)} \|\mathbf{E}_V(x)\|^2 + \frac{\mu(x)}{\varepsilon(x)\mu(x)} \|\mathbf{H}_V(x)\|^2 \Big) \,\mathrm{d}x \\
&\leq 3C_F^2 \|AV\|_X^2 + \frac{6}{\kappa^2} \Big(\max_{j=1,2} \max_{i=1,2,3} \|\partial_{x_i} \sigma_j\|_\infty^2 \Big) \|V\|_X^2.
\end{aligned}
$$

We conclude linear growth of F on $Y \coloneqq \mathrm{D}(A)$ by

$$
\begin{aligned}
\|F(t,V)\|_{\mathrm{D}(A)}^2 &= \|AF(t,V)\|_X^2 + \|F(t,V)\|_X^2 \\
&\leq \max \Big\{ 3C_F^2, \frac{6}{\kappa^2} \max_{j=1,2} \max_{i=1,2,3} \|\partial_{x_i} \sigma_j\|_\infty^2 + C_F^2 \Big\} \|V\|_{\mathrm{D}(A)}^2.
\end{aligned}
$$

As noise $G(V)$, where $V = (\mathbf{E}_V^\top, \mathbf{H}_V^\top)^\top \in L^2(\mathcal{O})^6$, we consider the Nemytskij map associated to $\operatorname{diag}((-\varepsilon^{-1}\mathbf{E}_V^\top, -\mu^{-1}\mathbf{H}_V^\top))Q^{1/2}$, i.e., for $h \in L^2(\mathcal{O})^6$ and $x \in \mathcal{O}$, we have

$$
(G(V)h)(x) = \begin{pmatrix} -\varepsilon^{-1}(x) \operatorname{diag}(\mathbf{E}_V(x)) & 0 \\ 0 & -\mu^{-1}(x) \operatorname{diag}(\mathbf{H}_V(x)) \end{pmatrix} (Q^{1/2}h)(x) \in \mathbb{R}^6. \tag{4.4.28}
$$

Since for $V_1, V_2 \in L^2(\mathcal{O})^6$,

$$
\|G(V_1 - V_2)\|_{\mathcal{L}_2(H,X)} \leq \kappa^{-1} \|Q^{1/2}\|_{\mathcal{L}_2(H,X)} \|V_1 - V_2\|_X,
$$

$G \colon X \to \mathcal{L}_2(H,X)$ is Lipschitz continuous on X. In order to show linear growth of G on $\mathrm{D}(A)$, higher regularity assumptions have to be imposed on $Q^{1/2}$, as discussed in [28, p. 5]. For $\mathbb{G} \colon X \to \mathcal{L}_2(Q^{1/2}H, X)$ defined via $G = \mathbb{G}Q^{1/2}$ and all $\beta > \frac{3}{2}$, the estimate

$$
\|\mathbb{G}(V)\|_{\mathcal{L}_2(Q^{1/2}H, \mathrm{D}(A))} \lesssim \|Q^{1/2}\|_{\mathcal{L}_2(L^2(\mathcal{O})^6, H^{1+\beta}(\mathcal{O})^6)} (1 + \|V\|_{\mathrm{D}(A)})
$$

is known [28, formula (7)] provided that $Q^{1/2} \in \mathcal{L}_2(L^2(\mathcal{O})^6, H^{1+\beta}(\mathcal{O})^6)$. Taking into account that for an orthonormal basis $(e_\ell)_{\ell \in \mathbb{N}}$ of H, we can rewrite

$$
\|G(V)\|_{\mathcal{L}_2(H,\mathrm{D}(A))} = \sum_{\ell \in \mathbb{N}} \|G(V)e_\ell\|_{\mathrm{D}(A)} = \sum_{\ell \in \mathbb{N}} \|\mathbb{G}(V)Q^{1/2}e_\ell\|_{\mathrm{D}(A)} = \|\mathbb{G}(V)\|_{\mathcal{L}_2(Q^{1/2}H, \mathrm{D}(A))},
$$

this directly implies linear growth of G on $\mathrm{D}(A)$. The choice of the coefficient $\beta > \frac{3}{2}$ stems from the fact that the Sobolev embedding $H^\beta(\mathcal{O}) \hookrightarrow L^\infty(\mathcal{O})$ holds for $\beta > \frac{d}{2} = \frac{3}{2}$ since $\mathcal{O} \subseteq \mathbb{R}^3$ [72, Ex. 9.3.4]. Moreover, we then also have $H^{1+\beta} \hookrightarrow W^{1,\infty}(\mathcal{O})$. Thus, for the embedding into $\mathrm{D}(A)$ to hold, $Q^{1/2}$ is required to map into $H^{1+\beta}(\mathcal{O})^6$.

Theorem 4.40. *Let $p \in [2, \infty)$ and F, G as introduced in* (4.4.27) *and* (4.4.28)*, respectively. Suppose that*

$$
u_0 \in L^p_{\mathcal{F}_0}(\Omega; H_0(\mathrm{curl}, \mathcal{O}) \times H(\mathrm{curl}, \mathcal{O})) \text{ and } Q^{1/2} \in \mathcal{L}_2(L^2(\mathcal{O})^6, H^{1+\beta}(\mathcal{O})^6)
$$

for some $\beta > \frac{3}{2}$. Denote by U the mild solution to the stochastic Maxwell's equations (4.4.26) *with multiplicative noise and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in* (4.4.5) *obtained with the exponential Euler method $R \coloneqq S$. Then there exists a constant $C \geq 0$ depending on $(\sigma_1, \sigma_2, u_0, T, p, \alpha, \varepsilon, \mu, \kappa)$ such that for $N_k \geq 2$*

$$
\Big\| \max_{0 \leq j \leq N_k} \|U(t_j) - U^j\|_{H^\sigma} \Big\|_p \leq C\big(1 + \|Q^{1/2}\|_{\mathcal{L}_2(L^2(\mathcal{O})^6, H^{1+\beta}(\mathcal{O})^6)}\big) k^{1/2},
$$

i.e., the approximations $(U^j)_j$ converge at rate $\frac{1}{2}$ as $k \to 0$.

Proof. This is a consequence of Corollary 4.26 with $\alpha = \frac{1}{2}$ and $Y = \mathrm{D}(A)$. From the above considerations, it follows that the conditions on F and G are met. It remains to verify that Y is Hilbert and $(S(t))_{t\geq 0}$ is a contraction semigroup on both X and Y. Since $Y = \mathrm{D}(A)$ is a Banach space [107, p. 410] and $\lambda - A$ defines an isomorphism between $\mathrm{D}(A)$ and X for $\lambda \in \rho(-A)$, it is also a Hilbert space. Furthermore, $(S(t))_{t\geq 0}$ is a C_0-contraction semigroup on X, via Stone's theorem even a unitary C_0-group [28, Formula (3)]. By definition of the graph norm, this implies contractivity of the semigroup also on $\mathrm{D}(A)$. □

We can extend [28, Thm. 3.3] to schemes involving rational approximations.

Theorem 4.41. *Let $p \in [2, \infty)$ and F, G as introduced in (4.4.27) and (4.4.28), respectively. Suppose that*

$$u_0 \in L^p_{\mathcal{F}_0}(\Omega; H_0(\mathrm{curl}, \mathcal{O}) \times H(\mathrm{curl}, \mathcal{O})) \text{ and } Q^{1/2} \in \mathcal{L}_2(L^2(\mathcal{O})^6, H^{1+\beta}(\mathcal{O})^6)$$

for some $\beta > \frac{3}{2}$. Let $(R_k)_{k>0}$ be a time discretisation scheme that is contractive on $L^2(\mathcal{O})^6$ and $H_0(\mathrm{curl}, \mathcal{O}) \times H(\mathrm{curl}, \mathcal{O})$. Assume R approximates S to order $\frac{1}{2}$ on $H_0(\mathrm{curl}, \mathcal{O}) \times H(\mathrm{curl}, \mathcal{O})$. Denote by U the mild solution to the stochastic Maxwell's equations (4.4.26) with multiplicative noise and by $(U^j)_{j=0,\ldots,N_k}$ the temporal approximations as defined in (4.4.5). Then there exists a constant $C \geq 0$ depending on $(\sigma_1, \sigma_2, u_0, T, p, \alpha, \varepsilon, \mu, \kappa)$ such that for $N_k \geq 2$

$$\Big\| \max_{0\leq j\leq N_k} \|U(t_j) - U^j\|_{H^\sigma} \Big\|_p \leq C\big(1 + \|Q^{1/2}\|_{\mathcal{L}_2(L^2(\mathcal{O})^6, H^{1+\beta}(\mathcal{O})^6)}\big) \sqrt{\max\{\log(T/k), p\}}\, k^{1/2},$$

i.e., the approximations $(U^j)_j$ converge at rate $\frac{1}{2}$ up to a logarithmic correction factor as $k \to 0$. In particular, rate $\frac{1}{2}$ is attained for the implicit Euler method and Crank–Nicolson.

4.5 Convergence Rates for Abstract Wave Equations

In this section, we shall be concerned with rates of convergence for abstract stochastic wave equations of the form

$$\mathrm{d}U + AU\,\mathrm{d}t = \mathbf{F}(t,U)\,\mathrm{d}t + \mathbf{G}(t,U)\,\mathrm{d}W_H(t), \quad U(0) = U_0 = (u_0, v_0) \in L^p(\Omega; X) \tag{4.5.1}$$

on a phase space $X = V \times V_{-1}$ of product structure to be specified later, which takes different regularities of the first and second components of the mild solution into account. We achieve the following convergence rates for the pathwise uniform error E_k^∞ with time step $k > 0$ provided that the noise is sufficiently regular:

- $\mathrm{E}_k^\infty \lesssim k^\alpha \sqrt{\log(T/k)}$ with α close to one (contractive schemes, multiplicative noise);
- $\mathrm{E}_k^\infty \lesssim k$ (exponential Euler, multiplicative noise).

Up to a logarithmic factor, these rates are optimal for the given problem. They provide an alternative proof of [132, Thm. 3.1] for the exponential Euler method under less restrictive regularity assumptions on $\mathbf{F}$ and $\mathbf{G}$ and without making use of the group structure of the semigroup. The latter is crucial in order to extend the convergence result beyond the exponential Euler method. We extend the convergence result to general contractive schemes, which, to the best of our knowledge, is novel.

At the heart of our proof lies the higher Hölder continuity of the first component of the mild solution in V compared to the mild solution vector in X, which emerges from the product structure of the phase space on which the abstract wave equation is considered. This allows for better estimates of those error terms depending on the Hölder continuity of the mild solution. In particular, convergence at rate close to 1 can be achieved even for multiplicative noise, for which the convergence rate was limited to $\frac{1}{2}$ in the setting of Section 4.4. Incorporating the product structure of abstract wave equations into the setting of Section 4.4 leads to the main Theorem 4.48 in Subsection 4.5.1. The exponential Euler method, for which convergence at rate 1 is attained, is covered in Subsection 4.5.2. An extension of the error estimates to the full time interval is presented in Subsection 4.5.3. The results are illustrated for the stochastic wave equation with trace class noise, space-time white noise, and smooth noise in Subsections 4.5.4 to 4.5.6.

4.5.1 General contractive time discretisation schemes

Let V be a separable Hilbert space equipped with the norm $\|\cdot\|_V$. Consider a densely defined, positive, self-adjoint, and invertible operator $\Lambda\colon \mathrm{D}(\Lambda) \subseteq V \to V$. For $\beta \in \mathbb{R}$, define the norm $\|u\|_{V_\beta} := \|\Lambda^{\beta/2}u\|_V$ for $u \in V_\beta$ and, for $\beta \geq 0$, denote the domain of $\Lambda^{\frac{\beta}{2}}$ by V_β and equip it with this norm. For negative β, we denote by V_β the completion of V with respect to $\|\cdot\|_{V_\beta}$. We can thus interpret Λ as an operator mapping from V_1 to V_{-1}, noting that $\|\Lambda u\|_{V_{-1}} = \|\Lambda^{-1/2}\Lambda u\|_V = \|\Lambda^{1/2}u\|_V = \|u\|_{V_1}$. Furthermore, it holds that $V = V_0$. In this section, we consider stochastic evolution equations on the phase space $X := V_0 \times V_{-1} = V \times V_{-1}$. More generally, we introduce the product spaces

$$X_\beta := V_\beta \times V_{\beta-1} = \mathrm{D}(\Lambda^{\frac{\beta}{2}}) \times \mathrm{D}(\Lambda^{\frac{\beta-1}{2}}) \tag{4.5.2}$$

for $\beta \in \mathbb{R}$, equipped with the norm $\|U\|_{X_\beta} := (\|u\|_{V_\beta}^2 + \|v\|_{V_{\beta-1}}^2)^{1/2}$ for $U = (u, v) \in X_\beta$. Clearly, it then holds that $X = X_0$. Also, note that we restrict our considerations to Hilbert spaces in this section, since we are concerned with self-adjoint operators Λ.

The stochastic evolution equation (4.5.1) depends on the nonlinearity $\mathbf{F}\colon \Omega \times [0,T] \times X \to X$ and the multiplicative noise $\mathbf{G}\colon \Omega \times [0,T] \times X \to \gamma(H, X)$ on the phase space X. However, the product structure of X considered in this section motivates an interpretation of (4.5.1) as a system of two evolution equations. We set

$$A = \begin{pmatrix} 0 & -I \\ \Lambda & 0 \end{pmatrix}, \ \mathbf{F}(t, U) = \begin{pmatrix} 0 \\ F(t, u) \end{pmatrix}, \ \mathbf{G}(t, U) = \begin{pmatrix} 0 \\ G(t, u) \end{pmatrix} \text{ for } U = \begin{pmatrix} u \\ v \end{pmatrix} \in X \quad (4.5.3)$$

with domain $\mathrm{D}(A) := \{U \in X : AU \in X\}$. Due to $-\Lambda u \in V_{-1}$ if and only if $u \in V_1$, we find that the domain of A can be rewritten as

$$\mathrm{D}(A) = \{(u, v) \in X : (v, -\Lambda u) \in V_0 \times V_{-1}\} = V_1 \times V_0 = X_1.$$

Likewise, one can show that $\mathrm{D}(A^n) = X_n$ for $n \in \mathbb{N}$. To extend this statement beyond the integer case, we note that for $\delta \in (0, 1)$, real interpolation theory (cf. Section 2.3) yields

$$X_\delta = \mathrm{D}(\Lambda^{\frac{\delta}{2}}) \times \mathrm{D}(\Lambda^{\frac{\delta-1}{2}}) = (X_0, X_1)_{\delta,2} = (X, \mathrm{D}(A))_{\delta,2} = \mathrm{D}(A^\delta).$$

This gives rise to the system of evolution equations

$$\begin{cases} \mathrm{d}u = v \,\mathrm{d}t, \\ \mathrm{d}v = (-\Lambda u + F(t, u)) \,\mathrm{d}t + G(t, u) \,\mathrm{d}W_H(t). \end{cases}$$

This precisely captures the setting of stochastic wave equations when thinking of $v(t)$ as the derivative of $u(t)$, thus yielding a stochastic evolution equation for the derivative with left-hand side $\dot{u}(t)$. The invertibility of Λ is a non-restrictive assumption, because we can always reduce to this case by writing $-\Lambda u + F(t, u) = -(\Lambda + \varepsilon)u + (\varepsilon u + F(t, u))$ without changing the properties of F.

It is not by coincidence that wave equations are commonly considered on the phase space $X = V \times V_{-1}$. Thinking of the physical interpretation of u as the position and v as the velocity, i.e., the derivative of the position, it seems natural to assume regularity of one order lower for the velocity as the second component. But even stepping aside from the physical interpretation, a deeper mathematical result motivates why the phase space is the "correct" space: There simply are no "interesting" generators on $V \times V$ from a PDE perspective, as the following theorem illustrates.

Theorem 4.42 (Corollary 3.14.9 (i) and (iii) in [5])**.** *The following assertions are equivalent.*

(i) The operator $-A$ generates a C_0-semigroup on $V \times V$.

(ii) Λ is bounded.

On the phase space, however, the operator $-A$ from (4.5.3) generates a C_0-semigroup $(S(t))_{t\geq 0}$ given by

$$S(t) = \begin{pmatrix} \cos(t\Lambda^{1/2}) & \Lambda^{-1/2}\sin(t\Lambda^{1/2}) \\ -\Lambda^{1/2}\sin(t\Lambda^{1/2}) & \cos(t\Lambda^{1/2}) \end{pmatrix}, \quad (4.5.4)$$

where we use the spectral theorem for self-adjoint operators to define the matrix entries. Indeed, $S(t) \to I$ strongly as $t \to 0$ due to

$$\lim_{t\to 0} \| \cos(t\Lambda^{1/2})x - x\| = \lim_{t\to 0} \Big\| \int_0^t \sin(s\Lambda^{1/2})\Lambda^{1/2}x \, \mathrm{d}s\Big\| \leq \lim_{t\to 0} t\|\Lambda^{1/2}x\| = 0$$

and, analogously, $\lim_{t\to 0} \| \pm \Lambda^{\mp 1/2} \sin(t\Lambda^{1/2})x\| = 0$ for $x \in \mathrm{D}(\Lambda^{1/2})$. Strong continuity of the semigroup follows by the density of $\mathrm{D}(\Lambda^{1/2})$ and the spectral theorem for self-adjoint operators [5, Theorem 3.14.11]. It is straightforward to see that $(S(t))_{t\geq 0}$ satisfies the semigroup property and that $-A$ is its infinitesimal generator.

Let $\beta \in \mathbb{R}$. Combining the respective one-dimensional statements with the spectral theorem, we obtain that $\sin(t\Lambda^{1/2})$ and $\cos(t\Lambda^{1/2})$ are contractive on V_β, $\sin(0 \cdot \Lambda^{1/2}) = 0$, and that Λ and powers thereof commute with both $\sin(t\Lambda^{1/2})$ and $\cos(t\Lambda^{1/2})$. The trigonometric identity satisfied by $\sin(t\Lambda^{1/2})$ and $\cos(t\Lambda^{1/2})$ implies contractivity of the semigroup, that is,

$$\|S(t)U\|_{X_\beta} \leq \|U\|_{X_\beta}. \tag{4.5.5}$$

Our aim is to derive conditions on F and G rather than $\mathbf{F}$ and $\mathbf{G}$ under which the temporal approximations

$$U^j = R_k^j U_0 + k\sum_{i=0}^{j-1} \mathbf{F}(t_i, U^i) + \sum_{i=0}^{j-1} \Delta W_{i+1} R_k^{j-i}\mathbf{G}(t_i, U^i), \quad 0 \leq j \leq N_k, \tag{4.5.6}$$

converge to the values $U(t_j) = (u(t_j), v(t_j)) \in X$ of the mild solution of (4.5.1) in the time grid points at a certain rate. As will become apparent, rates of convergence exceeding 1/2 can be attained up to a logarithmic correction factor, even for general contractive schemes. The key aspect of this section's main theorem enabling this optimal rate, Theorem 4.24, consists of higher-order Hölder continuity of the first component of the mild solution.

As will be shown, the following assumptions on F and G imply that $\mathbf{F}$ and $\mathbf{G}$ fall within the scope of the setting of Section 4.4.

Assumption 4.43. *Let $\Lambda\colon \mathrm{D}(\Lambda) \subseteq V \to V$ be a densely defined, positive, self-adjoint, and invertible operator on a Hilbert space V. Further, let $p \in [2,\infty)$, $\alpha \in (0,1]$, and $\delta \geq \alpha$. Let $F\colon \Omega \times [0,T] \times V \to V_{-1}$ and $G\colon \Omega \times [0,T] \times V \to \mathcal{L}_2(H, V_{-1})$ be strongly $\mathcal{P} \otimes \mathcal{B}(V)$-measurable, where $F = \tilde{F} + f$ and $G = \tilde{G} + g$ for $f \in L^p_{\mathcal{P}}(\Omega; C([0,T]; V_{-1}))$ and $g \in L^p_{\mathcal{P}}(\Omega; C([0,T]; \mathcal{L}_2(H, V_{-1})))$. Suppose that $\tilde{F}(\cdot,\cdot,0) = 0$, $\tilde{G}(\cdot,\cdot,0) = 0$, and*

(a) (Lipschitz continuity from V to V_{-1}) *there exist constants $C_{F,V}, C_{G,V} \geq 0$ such that for all $\omega \in \Omega, t \in [0,T]$, and $x, y \in V$,*

$$\|\tilde{F}(\omega,t,x) - \tilde{F}(\omega,t,y)\|_{V_{-1}} \leq C_{F,V}\|x-y\|_V,$$
$$\|\tilde{G}(\omega,t,x) - \tilde{G}(\omega,t,y)\|_{\mathcal{L}_2(H,V_{-1})} \leq C_{G,V}\|x-y\|_V,$$

(b) (Hölder continuity with values in V_{-1}) *it holds that*

$$C_{\alpha,F} := \sup_{\omega\in\Omega, x\in V} [\Lambda^{-\frac{1}{2}}F(\omega,\cdot,x)]_\alpha < \infty, \ C_{\alpha,G} := \sup_{\omega\in\Omega, x\in V} [\Lambda^{-\frac{1}{2}}G(\omega,\cdot,x)]_\alpha < \infty,$$

(c) (invariance) $F\colon \Omega\times[0,T]\times V_\delta\to V_{\delta-1}$ *and* $G\colon\Omega\times[0,T]\times V_\delta\to\mathcal{L}_2(H,V_{\delta-1})$ *are strongly* $\mathcal{P}\otimes\mathcal{B}(V_\delta)$*-measurable,* $f\in L^p_{\mathcal{P}}(\Omega;C([0,T];V_{\delta-1}))$*, and* $g\in L^p_{\mathcal{P}}(\Omega;C([0,T];\mathcal{L}_2(H,V_{\delta-1})))$*,*

(d) (linear growth from V_δ to $V_{\delta-1}$) *and there exist constants* $L_{F,\delta},L_{G,\delta}\ge 0$ *such that for all* $\omega\in\Omega$*,* $t\in[0,T]$*, and* $x\in V$*,*

$$\|\tilde F(\omega,t,x)\|_{V_{\delta-1}}\le L_{F,\delta}(1+\|x\|_{V_\delta}),$$
$$\|\tilde G(\omega,t,x)\|_{\mathcal{L}_2(H,V_{\delta-1})}\le L_{G,\delta}(1+\|x\|_{V_\delta}).$$

It is important to note that both $\delta\in[\alpha,1]\subseteq(0,1]$ and $\delta\in(1,2]$ will be considered. Since for $\delta=2$, optimal rates are obtained for the usual schemes, larger values of δ are not considered. Analogously to the definition of $\mathbf{F}$ and $\mathbf{G}$ in (4.5.3), define $\tilde{\mathbf{F}},\mathbf{f},\tilde{\mathbf{G}}$, and $\mathbf{g}$ based on $\tilde F,f,\tilde G$, and g, respectively.

Next, we show that Assumption 4.43 implies the conditions required for well-posedness on both X and X_δ for δ as in Assumption 4.43 and thus (4.5.1) has a unique mild solution on both spaces. Moreover, we verify the conditions for stability of the temporal approximations obtained via a contractive time discretisation scheme. Adopt the notation of the proof of Theorem 4.24, replacing $F,\tilde F,f,G,\tilde G$, and g by $\mathbf{F},\tilde{\mathbf{F}},\mathbf{f},\mathbf{G},\tilde{\mathbf{G}}$, and $\mathbf{g}$, respectively.

Setting $Y:=X_\delta$ for δ as in Assumption 4.43, it is clear from $X=X_0$, invertibility of Λ, $\mathrm{D}(A^n)=X_n$ for $n\in\mathbb{N}$, and the interpolation theory discussed in Section 2.3 that $Y\hookrightarrow X$ and $Y\hookrightarrow D_A(\beta,\infty)$ for any $\beta\in(0,\delta)$. Since V_δ is a separable Hilbert space for $\delta\in\mathbb{R}$, so are X and Y. Contractivity of the semigroup follows from (4.5.5). Note that strong $\mathcal{P}\otimes\mathcal{B}(X)$-measurability of $\mathbf{F}$ and $\mathbf{G}$ immediately follows from the respective assumptions on $\tilde F,\tilde G$ due to the structure (4.5.3), and, likewise, that $\tilde{\mathbf{F}},\tilde{\mathbf{G}}$ vanish in 0. We are left to prove Lipschitz continuity on X, linear growth on Y, and Y-invariance of $\mathbf{F},\mathbf{G}$, and continuity of $\mathbf{f}$ and $\mathbf{g}$ with values in Y. Deducing Y-invariance from Assumption 4.43 is straightforward, noting that

$$\|\mathbf{f}\|_{p,\infty,Y}=\left\|\sup_{t\in[0,T]}\|\mathbf{f}(t)\|_Y\right\|_p=\left\|\sup_{t\in[0,T]}\|f(t)\|_{V_{\delta-1}}\right\|_p=\|f\|_{p,\infty,V_{\delta-1}}\tag{4.5.7}$$

and, likewise, $\|\mathbf{g}\|_{p,\infty,Y}=\|g\|_{p,\infty,V_{\delta-1}}$, where we have used the shorthand notation introduced in 4.2.12. The mapping properties on Y and strong $\mathcal{P}\otimes\mathcal{B}(Y)$-measurability of $\mathbf{F}$ and $\mathbf{G}$ follow from Assumption 4.43(c) because $Y=V_\delta\times V_{\delta-1}$. Linear growth of $\tilde{\mathbf{F}}$ from Y to Y follows from linear growth of $\tilde F$ from V_δ to $V_{\delta-1}$ as stated in Assumption 4.43(d) taking the structure (4.5.3) of $\mathbf{F}$ into account via

$$\|\tilde{\mathbf{F}}(t,U)\|_Y=\|\tilde F(t,u)\|_{V_{\delta-1}}\le L_{F,\delta}(1+\|u\|_{V_\delta})\le L_{F,\delta}(1+\|U\|_Y)$$

for $t\in[0,T]$ and $U=(u,v)\in Y=V_\delta\times V_{\delta-1}$. Analogously, linear growth of $\tilde{\mathbf{G}}$ from Y to $\gamma(H,Y)$ is obtained from

$$\|\tilde{\mathbf{G}}(t,U)\|_Y=\|\tilde G(t,u)\|_{V_{\delta-1}}\le L_{G,\delta}(1+\|u\|_{V_\delta})\le L_{G,\delta}(1+\|U\|_Y).$$

Lipschitz continuity of $\mathbf{F}$ from X to X follows from Assumption 4.43(a) due to

$$\|\mathbf{F}(t,U_1)-\mathbf{F}(t,U_2)\|_X=\|\tilde F(t,u_1)-\tilde F(t,u_2)\|_{V_{-1}}\le C_{F,V}\|u_1-u_2\|_V\le C_{F,V}\|U_1-U_2\|_X$$

for $t \in [0,T]$ and $U_1 = (u_1, v_1), U_2 = (u_2, v_2) \in X$. Analogously,

$$\|\mathbf{G}(t,U_1) - \mathbf{G}(t,U_2)\|_{\gamma(H,X)} = \|\tilde{G}(t,u_1) - \tilde{G}(t,u_2)\|_{\mathcal{L}_2(H,V_{-1})} \leq C_{G,V}\|U_1 - U_2\|_X.$$

Hence, $\mathbf{G}\colon X \to \gamma(H,X)$ is Lipschitz continuous. Via the same argument,

$$[\mathbf{F}(\omega,\cdot,U)]_\alpha = \sup_{0\leq s\leq t\leq T} \frac{\|\mathbf{F}(t,U) - \mathbf{F}(s,U)\|_X}{(t-s)^\alpha} = \sup_{0\leq s\leq t\leq T} \frac{\|\Lambda^{-\frac{1}{2}}[F(t,u) - F(s,u)]\|_V}{(t-s)^\alpha},$$

from which, together with Assumption 4.43(b), we conclude α-Hölder continuity of $\mathbf{F}$ and, analogously, of $\mathbf{G}$. Although not required for neither well-posedness nor stability, this will be of use for the error analysis later.

The above leads to well-posedness of (4.5.1) on X and Y.

Lemma 4.44 (Well-posedness)**.** *Suppose that Assumption 4.43 holds for some $\alpha \in (0,1]$, $\delta \geq \alpha$, and $p \in [2,\infty)$. Let $Y := X_\delta$ as defined in* (4.5.2) *and $U_0 \in L^p_{\mathcal{F}_0}(\Omega;Y)$. Under these conditions there exists a unique mild solution $U \in L^p(\Omega;C([0,T];X))$ to* (4.5.1) *with A, $\mathbf{F}$, and $\mathbf{G}$ as in* (4.5.3)*. Furthermore, it lies in $L^p(\Omega;C([0,T];Y))$ and*

$$\|U\|_{L^p(\Omega;C([0,T];Y))} \leq C^Y_{bdd}\big(1 + \|U_0\|_{L^p(\Omega;Y)} + \|f\|_{p,1,V_{\delta-1}} + B_p \|\!|g|\!\|_{p,2,V_{\delta-1}}\big),$$

where $C^Y_{bdd} := (1+C^2T)^{1/2}\mathrm{e}^{(1+C^2T)/2}$ with $C := L_{F,\delta}T^{1/2} + B_pL_{G,\delta}$, and B_p is the constant from Theorem 2.61.

As established in (4.4.4), the well-posedness on $Z \in \{X,Y\}$ implies

$$1 + \Big\|\sup_{r\in[0,T]} \|U(r)\|_Z\Big\|_p \leq C_{U_0,\mathbf{f},\mathbf{g},Z} < \infty$$

with $C_{U_0,\mathbf{f},\mathbf{g},Z}$ as defined in (4.4.3). In the abstract wave equation setting, the constant simplifies to

$$C_{U_0,\mathbf{f},\mathbf{g},Z} = 1 + C^Z_{\mathrm{bdd}}\big(1 + \|U_0\|_{L^p(\Omega;Z)} + \|f\|_{p,1,Z_2} + \|\!|g|\!\|_{p,2,Z_2}\big), \tag{4.5.8}$$

where C^Z_{bdd} denotes the constant from Lemma 4.44, $Z_2 := V_{-1}$ if $Z = X$, and $Z_2 := V_{\delta-1}$ if $Z = Y = X_\delta$ for $\delta \in (0,2]$.

Lemma 4.45 (Stability)**.** *Suppose that Assumption 4.43 holds for some $\alpha \in (0,1]$, $\delta \geq \alpha$, and $p \in [2,\infty)$. Let $X := X_0$ and $Y := X_\delta$ be as defined in* (4.5.2) *and $U_0 \in L^p_{\mathcal{F}_0}(\Omega;Y)$. Let $(R_k)_{k>0}$ be a time discretisation scheme that is contractive on both X and Y, and let $N_k \geq 2$. Then the temporal approximations $(U^j)_{j=0,\dots,N_k}$ obtained via* (4.5.6) *with $\mathbf{F}$ and $\mathbf{G}$ as in* (4.5.3) *are stable on both X and Y. That is, for $Z \in \{X,Y\}$,*

$$1 + \Big\|\max_{0\leq j\leq N_k} \|U^j\|_Z\Big\|_p \leq C^Z_{stab}c_{U_0,f,g,T,Z},$$

where $C^Z_{stab} := (1+C_Z^2T)^{1/2}\mathrm{e}^{(1+C_Z^2T)/2}$ with $C_X := C_{F,V}T^{1/2} + B_pC_{G,V}$, $C_Y := L_{F,\delta}T^{1/2} + B_pL_{G,\delta}$,

$$c_{U_0,f,g,T,Z} := 1 + \|U_0\|_{L^p(\Omega;Z)} + T\|f\|_{p,\infty,Z_2}T + B_pT^{1/2}\|\!|g|\!\|_{p,\infty,Z_2},$$

where $Z_2 := V_{-1}$ if $Z = X$ and $Z_2 := V_{\delta-1}$ if $Z = Y$, and B_p is the constant from Theorem 2.61.

We denote

$$\begin{aligned} K_{U_0,f,g,Y} &:= C^Y_{\text{stab}} c_{U_0,f,g,T,Y} \\ &= C^Y_{\text{stab}}\big(1 + \|U_0\|_{L^p(\Omega;Y)} + T\|f\|_{p,\infty,V_{\delta-1}} + B_p T^{1/2} \|\!|g|\!\|_{p,\infty,V_{\delta-1}}\big), \end{aligned} \tag{4.5.9}$$

so that $K_{U_0,f,g,Y} = K_{U_0,\mathbf{f},\mathbf{g},Y}$ with $K_{U_0,\mathbf{f},\mathbf{g},Y}$ as defined in (4.4.7).

For future estimates, it is useful to know the decay of differences of the sine and cosine operators $\sin(t\Lambda^{1/2})$ and $\cos(t\Lambda^{1/2})$. We include a short proof for the convenience of the reader.

Lemma 4.46 ([30], page 210)**.** *Let $t \in [0,T]$ and $\Lambda\colon \mathrm{D}(\Lambda) \subseteq V \to V$ be a densely defined, positive, self-adjoint, and invertible operator on a Hilbert space V. Then for all $\alpha \in [0,1]$, we have*

$$\begin{aligned} \|\Lambda^{-\frac{\alpha}{2}}[\sin(t\Lambda^{1/2}) - \sin(s\Lambda^{1/2})]\|_{\mathcal{L}(V)} &\leq 2(t-s)^\alpha, \\ \|\Lambda^{-\frac{\alpha}{2}}[\cos(t\Lambda^{1/2}) - \cos(s\Lambda^{1/2})]\|_{\mathcal{L}(V)} &\leq 2(t-s)^\alpha \end{aligned}$$

for all $0 \leq s \leq t \leq T$.

Proof. The statement is trivially fulfilled for $t = s$. Let $0 \leq s < t \leq T$. We claim that

$$\zeta_\alpha(t,s) := \frac{|\sin(t) - \sin(s)|}{|t-s|^\alpha} \leq 2.$$

Indeed, if $|t-s| \leq 1$, then by the mean value theorem $\zeta_\alpha(t,s) \leq \zeta_1(t,s) \leq 1$. If $|t-s| > 1$, then $\zeta_\alpha(t,s) \leq 2$. Now let $\lambda > 0$. Applying the claim with $t\lambda^{1/2}$ and $s\lambda^{1/2}$ gives

$$\lambda^{-\alpha/2}|\sin(t\lambda^{1/2}) - \sin(s\lambda^{1/2})| \leq 2|t-s|^\alpha.$$

Thus, by the spectral theorem for self-adjoint operators and positivity of Λ, we get the desired statement. The statement for the cosine is proven analogously. □

While the mild solution U has at most 1/2-Hölder continuous paths as follows from Lemma 4.22, the product structure of the stochastic evolution equation results in higher Hölder continuity of the first component u of U, as the following lemma illustrates. In particular, u has Lipschitz continuous paths for sufficiently regular F and G.

Lemma 4.47. *Suppose that Assumption 4.43 holds for some $\alpha \in (0,1]$, $\delta \geq \alpha$, and $p \in [2,\infty)$. Let $X := X_0$ and $Y := X_\delta$ be as defined in* (4.5.2) *and $U_0 \in L^p_{\mathcal{F}_0}(\Omega;Y)$. Then for all $0 \leq s \leq t \leq T$, the first component u of the mild solution U of* (4.5.1) *with A, $\mathbf{F}$, and $\mathbf{G}$ as in* (4.5.3) *satisfies*

$$\|u(t) - u(s)\|_{L^p(\Omega;V)} \leq L(t-s)^\alpha$$

with constant

$$L := 2\Big[\sqrt{2}C_Y\|U_0\|_{L^p(\Omega;Y)} + L_{1,F}T\frac{\alpha+2}{\alpha+1} + B_p L_{2,G} T^{1/2}\Big(1 + \frac{1}{\sqrt{2\alpha+1}}\Big)\Big],$$

where $L_{1,F} := C_Y L_{F,\delta} C_{U_0,\mathbf{f},\mathbf{g},Y} + \|f\|_{p,\infty,V_{\alpha-1}}$, $L_{2,G} := C_Y L_{G,\delta} C_{U_0,\mathbf{f},\mathbf{g},Y} + \|\!|g|\!\|_{p,\infty,V_{\alpha-1}}$, $C_{U_0,\mathbf{f},\mathbf{g},Y}$ is as in (4.5.8)*, C_Y denotes the embedding constant of X_δ into X_α, and B_p is the constant from Theorem 2.61.*

Proof. From the structure (4.5.4) of the semigroup as well as (4.5.3) of **F** and **G**, we deduce the following variation-of-constants formula for the first component of the mild solution.

$$u(t) = \cos(t\Lambda^{1/2})u_0 + \Lambda^{-\frac{1}{2}}\sin(t\Lambda^{1/2})v_0 + \int_0^t \Lambda^{-\frac{1}{2}}\sin((t-r)\Lambda^{1/2})F(r,u(r))\,\mathrm{d}r$$
$$+ \int_0^t \Lambda^{-\frac{1}{2}}\sin((t-r)\Lambda^{1/2})G(r,u(r))\,\mathrm{d}W_H(r).$$

Hence, the difference can be split up as

$$\begin{aligned}
&\|u(t)-u(s)\|_{L^p(\Omega;V)} \\
&\le \big\|[\cos(t\Lambda^{1/2})-\cos(s\Lambda^{1/2})]u_0 + \Lambda^{-\frac{1}{2}}[\sin(t\Lambda^{1/2})-\sin(s\Lambda^{1/2})]v_0\big\|_{L^p(\Omega;V)} \\
&\quad + \Big\|\int_0^s \|\Lambda^{-\frac{1}{2}}[\sin((t-r)\Lambda^{1/2})-\sin((s-r)\Lambda^{1/2})]F(r,u(r))\|_V \,\mathrm{d}r\Big\|_p \\
&\quad + \Big\|\int_s^t \|\Lambda^{-\frac{1}{2}}\sin((t-r)\Lambda^{1/2})F(r,u(r))\|_V \,\mathrm{d}r\Big\|_p \\
&\quad + \Big\|\int_0^s \Lambda^{-\frac{1}{2}}[\sin((t-r)\Lambda^{1/2})-\sin((s-r)\Lambda^{1/2})]G(r,u(r))\,\mathrm{d}W_H(r)\Big\|_{L^p(\Omega;V)} \\
&\quad + \Big\|\int_s^t \Lambda^{-\frac{1}{2}}\sin((t-r)\Lambda^{1/2})G(r,u(r))\,\mathrm{d}W_H(r)\Big\|_{L^p(\Omega;V)} =: E_1+E_2+E_3+E_4+E_5,
\end{aligned}$$

where $E_\ell := E_\ell(t,s)$ for $1 \le \ell \le 5$. We proceed to bound these five expressions individually. Lemma 4.46 yields

$$\begin{aligned}
E_1 &\le \Big\|\big\|[\cos(t\Lambda^{1/2})-\cos(s\Lambda^{1/2})]\Lambda^{-\frac{\alpha}{2}}\big\|_{\mathcal{L}(V)}\|\Lambda^{\frac{\alpha}{2}}u_0\|_V \\
&\quad + \big\|[\sin(t\Lambda^{1/2})-\sin(s\Lambda^{1/2})]\Lambda^{-\frac{\alpha}{2}}\big\|_{\mathcal{L}(V)}\|\Lambda^{\frac{\alpha-1}{2}}v_0\|_V\Big\|_p \\
&\le 2(t-s)^\alpha\big\|\|u_0\|_{V_\alpha}+\|v_0\|_{V_{\alpha-1}}\big\|_p \le 2\sqrt{2}\|U_0\|_{L^p(\Omega;X_\alpha)}\cdot(t-s)^\alpha \\
&\le 2\sqrt{2}C_Y\|U_0\|_{L^p(\Omega;Y)}\cdot(t-s)^\alpha,
\end{aligned}$$

where we have used the embedding $Y = X_\delta \hookrightarrow X_\alpha$ in the last line. Using the same trick of inserting $\Lambda^{-\frac{\alpha}{2}}$, applying Lemma 4.46, and using the embedding $V_{\delta-1} \hookrightarrow V_{\alpha-1}$ as well as linear growth of $\tilde{F}$ from V_δ to $V_{\delta-1}$, we obtain

$$\begin{aligned}
E_2 &\le 2s(t-s)^\alpha\Big\|\sup_{r\in[0,T]}\|\Lambda^{\frac{\alpha-1}{2}}F(r,u(r))\|_V\Big\|_p \\
&\le 2s(t-s)^\alpha\Big(C_Y\Big\|\sup_{r\in[0,T]}\|\tilde{F}(r,u(r))\|_{V_{\delta-1}}\Big\|_p + \Big\|\sup_{r\in[0,T]}\|f(r)\|_{V_{\alpha-1}}\Big\|_p\Big) \\
&\le 2s(t-s)^\alpha\Big(C_Y L_{F,\delta}\Big(1+\Big\|\sup_{r\in[0,T]}\|u(r)\|_{V_\delta}\Big\|_p\Big) + \|f\|_{p,\infty,V_{\alpha-1}}\Big) \le 2L_{1,F}T(t-s)^\alpha.
\end{aligned}$$

Likewise, for the stochastic integral, we conclude

$$E_4 \le 2B_p(C_Y L_{G,\delta}C_{U_0,\mathbf{f},\mathbf{g},Y} + \|g\|_{p,\infty,V_{\alpha-1}})s^{\frac{1}{2}}(t-s)^\alpha \le 2B_p L_{2,G}T^{\frac{1}{2}}(t-s)^\alpha.$$

Recalling that $\sin(0 \cdot \Lambda^{1/2}) = 0$, we can estimate

$$\begin{aligned} E_3 &\leq \Big\| \int_s^t \big\|[\sin((t-r)\Lambda^{1/2}) - \sin(0 \cdot \Lambda^{1/2})]\Lambda^{-\frac{\alpha}{2}}\big\|_{\mathcal{L}(V)} \big\|\Lambda^{\frac{\alpha-1}{2}} F(r, u(r))\big\|_V \, \mathrm{d}r \Big\|_p \\ &\leq 2L_{1,F} \int_s^t (t-r)^\alpha \, \mathrm{d}r \leq \frac{2L_{1,F}}{\alpha+1}(t-s)^{\alpha+1} \leq \frac{2L_{1,F}T}{\alpha+1}(t-s)^\alpha, \end{aligned}$$

and, analogously,

$$E_5 \leq \frac{2B_p L_{2,G}}{\sqrt{2\alpha+1}}(t-s)^{\alpha+\frac{1}{2}} \leq \frac{2B_p L_{2,G} T^{1/2}}{\sqrt{2\alpha+1}}(t-s)^\alpha.$$

Adding the bounds for E_1 to E_5 results in the desired statement. □

Analogous to the considerations in Remark 4.23, the regularity assumptions on f and g can be relaxed in this lemma. Having established Hölder continuity of u of order up to 1, we can derive an error bound attaining the optimal order 1 for sufficiently good schemes and regular nonlinearity, noise and initial values. The following main theorem of this section generalises [132, Thm. 3.1] from exponential Euler to general contractive schemes, as well as more general F and G.

Theorem 4.48. *Suppose that Assumption 4.43 holds for some $\alpha \in (0,1]$, $\delta \geq \alpha$, and $p \in [2,\infty)$. Let $X := X_0$ and $Y := X_\delta$ be as defined in (4.5.2) and $U_0 \in L^p_{\mathcal{F}_0}(\Omega; Y)$. Let A, $\mathbf{F}$, and $\mathbf{G}$ be as in (4.5.3) and let $(R_k)_{k>0}$ be a contractive time discretisation scheme on X which commutes with the resolvent of $-A$. Assume R approximates S to order α on Y. Denote by U the mild solution of (4.5.1) and by $(U^j)_{j=0,\ldots,N_k}$ the temporal approximations as defined in (4.5.6). Then for $N_k \geq 2$*

$$\Big\| \max_{0 \leq j \leq N_k} \|U(t_j) - U^j\| \Big\|_p \leq C_{\mathrm{e}}\big(C_1 + C_2\sqrt{\max\{\log(T/k), p\}}\big)k^\alpha$$

with $C_{\mathrm{e}} := (1 + C^2T)^{1/2}\exp((1+C^2T)/2)$, $C := C_{F,V}\sqrt{T} + B_p C_{G,V}$, $C_2 := KC_\alpha K_G\sqrt{T}$, and

$$\begin{aligned} C_1 := C_\alpha \|U_0\|_{L^p(\Omega;Y)} &+ \frac{T}{\alpha+1}(C_{F,V}L + C_{\alpha,F} + 2C_Y K_F) + C_\alpha K_F T \\ &+ \frac{B_p\sqrt{T}}{\sqrt{2\alpha+1}}(C_{G,V}L + C_{\alpha,G} + 2C_Y K_G), \end{aligned}$$

$K_F := L_{F,\delta}K_{U_0,f,g,Y} + \|f\|_{p,\infty,V_{\delta-1}}$, $K_G := L_{G,\delta}K_{U_0,f,g,Y} + \|g\|_{p,\infty,V_{\delta-1}}$, L as defined in Lemma 4.47, $K_{U_0,f,g,Y}$ as in (4.5.9), $K = 4\exp(1+\frac{1}{2\mathrm{e}})$, C_Y denotes the embedding constant of Y into $D_A(\alpha,\infty)$, and B_p is the constant from Theorem 2.61.

In particular, the approximations $(U^j)_j$ converge at rate $\min\{\alpha, 1\}$ up to a logarithmic correction factor as $k \to 0$.

Possible choices for R in the above include but are not limited to exponential Euler, implicit Euler, Crank–Nicolson, and other A-stable schemes. We recall that the contractivity of a large class of schemes follows from Proposition 2.28.

Proof. As derived in the discussion before Lemma 4.44, the conditions of Theorem 4.24 follow from Assumption 4.43. Now, we make use of Lemma 4.47 to obtain decay at rate α for those terms who limited the rate of convergence in Theorem 4.24 to $\frac{1}{2}$. Adopt the notation of the error terms from the proof of Theorem 4.24. Contractivity of S, Lipschitz continuity of $\tilde{F}$ from V to V_{-1}, and Lemma 4.47 together yield

$$\begin{aligned} M_{2,1} &\leq \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \|\mathbf{F}(s, U(s)) - \mathbf{F}(s, U(t_i))\|_{L^p(\Omega;X)} \,\mathrm{d}s \\ &= \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \|\tilde{F}(s, u(s)) - \tilde{F}(s, u(t_i))\|_{L^p(\Omega;V_{-1})} \,\mathrm{d}s \\ &\leq C_{F,V} \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \|u(s) - u(t_i)\|_{L^p(\Omega;V)} \,\mathrm{d}s \leq C_{F,V} L \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} (s - t_i)^{\alpha} \,\mathrm{d}s \\ &= \frac{C_{F,V} L}{\alpha + 1} t_N k^{\alpha}. \end{aligned}$$

Combining this with the bounds for $M_{2,2}$ to $M_{2,4}$ from Theorem 4.24 leads to

$$M_2 \leq \left(\frac{C_{F,V} L + C_{\alpha,F} + 2C_Y K_F}{\alpha + 1} + C_{\alpha} K_F \right) t_N k^{\alpha} + C_{F,V} \sqrt{t_N} \left(k \sum_{i=0}^{N-1} E(i)^2 \right)^{1/2}.$$

Here, we have used (4.5.7) to pass from the Y-norm of $\mathbf{f}$ to the $V_{\delta-1}$-norm of f appearing in K_F and that $C_{\mathbf{F},X}, L_{\mathbf{F},Y}, C_{\alpha,\mathbf{F}}$ agree with $C_{F,V}, L_{F,\delta}, C_{\alpha,F}$, respectively. For the term $M_{3,1}$, an application of the maximal inequality is required additionally. By the same reasoning as for $M_{2,1}$, we then deduce

$$M_{3,1} \leq B_p C_{G,V} \left(\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \|u(s) - u(t_i)\|^2_{L^p(\Omega;V)} \,\mathrm{d}s \right)^{1/2} \leq \frac{B_p C_{G,V} L}{\sqrt{2\alpha + 1}} \sqrt{t_N} k^{\alpha}.$$

In conclusion from the bounds for $M_{3,1}$ to $M_{3,5}$,

$$\begin{aligned} M_3 \leq{}& \frac{B_p}{\sqrt{2\alpha + 1}} (C_{G,V} L + C_{\alpha,G} + 2C_Y K_G) \sqrt{t_N} k^{\alpha} + K C_{\alpha} K_G \sqrt{t_N} \sqrt{\max\{\log(N), p\}} k^{\alpha} \\ &+ B_p C_{G,V} \left(k \sum_{i=0}^{N-1} E(i)^2 \right)^{1/2}. \end{aligned}$$

The final statement follows by summing the estimates for M_1, M_2 and M_3 and then applying Gronwall's inequality from Lemma 2.68. □

4.5.2 Exponential Euler method

Also in the case of the abstract stochastic wave equation, the logarithmic correction factor vanishes when the exponential Euler method is used. Hence, we obtain convergence at the optimal rate.

Corollary 4.49. *Suppose that Assumption 4.43 holds for some $\alpha \in (0,1]$, $\delta \geq \alpha$, and $p \in [2,\infty)$. Let $X := X_0$ and $Y := X_\delta$ be as defined in* (4.5.2) *and $U_0 \in L^p_{\mathcal{F}_0}(\Omega; Y)$. Consider the exponential Euler method $R := S$ for time discretisation. Denote by U the mild solution of* (4.5.1) *with A, $\mathbf{F}$, and $\mathbf{G}$ as in* (4.5.3)*, and by $(U^j)_{j=0,\ldots,N_k}$ the temporal approximations as defined in* (4.5.6)*. Then for $N_k \geq 2$*

$$\Big\| \max_{j=0,\ldots,N_k} \|U(t_j) - U^j\|_X \Big\|_p \leq C_{\mathrm{S,e}} C_{\mathrm{S}} \cdot k^\alpha$$

with constants $C_{\mathrm{S,e}} := C_{\mathrm{e}}$ as in Theorem 4.48 and

$$C_{\mathrm{S}} := \frac{T}{\alpha+1}(C_{F,V}L + C_{\alpha,F} + 2C_Y K_F) + \frac{B_p\sqrt{T}}{\sqrt{2\alpha+1}}(C_{G,V}L + C_{\alpha,G} + 2C_Y K_G),$$

where L is as defined in Lemma 4.47, K_F and K_G are as in Theorem 4.48, C_Y denotes the embedding constant of Y into $D_A(\alpha,\infty)$, and B_p is the constant from Theorem 2.61.

In particular, the approximations $(U^j)_j$ converge at rate $\min\{\alpha, 1\}$ as $k \to 0$.

4.5.3 Error estimates on the full time interval

In the same way as in the proof of Theorem 4.32, we see that we can extend the error bound in the grid points from Theorem 4.48 to the full time interval.

Corollary 4.50. *Suppose that the conditions of Theorem 4.48 hold for $\alpha \in (0,1/2]$ and $\delta \geq \alpha$. Further suppose that V is separable. Let $p_0 \in (p,\infty)$ and $q \in (2,\infty]$ be such that $\frac{1}{2} - \frac{1}{q} = \alpha$, and suppose that f, g, and U_0 have additional integrability*

$$f \in L^{p_0}(\Omega; L^1(0,T;V)), \quad g \in L^{p_0}(\Omega; L^q(0,T;\mathcal{L}_2(H,V))), \quad U_0 \in L^{p_0}_{\mathcal{F}_0}(\Omega; X) \cap L^p_{\mathcal{F}_0}(\Omega; X_\delta).$$

Denote by U the mild solution of (4.5.1) *with A, $\mathbf{F}$, $\mathbf{G}$ as in* (4.5.3)*, and by $(U^j)_{j=0,\ldots,N_k}$ the temporal approximations as defined in* (4.5.6)*. Define the piecewise constant extension $\tilde{U} \colon [0,T] \to L^p(\Omega; X)$ of $(U^j)_{j=0,\ldots,N_k}$ by $\tilde{U}(t) := U^j$ for $t \in [t_j, t_{j+1})$, $0 \leq j \leq N_k - 1$, and $\tilde{U}(T) := U^{N_k}$. Then for all $N_k \geq 2$ there is a constant $C \geq 0$ depending on $(T, p, p_0, \alpha, u_0, F, G, V, \delta)$ such that*

$$\Big\| \sup_{t\in[0,T]} \|U(t) - \tilde{U}(t)\|_X \Big\|_p \leq C\big(1 + \sqrt{\max\{\log(T/k), p\}}\big)k^\alpha.$$

In case we only estimate the first component u, more can be said about the convergence rate on the full time interval. Under weaker integrability conditions and for general $\alpha \in (0,1]$, we obtain the following.

Corollary 4.51. *Suppose that the conditions of Theorem 4.48 hold. Define the piecewise constant extension $\tilde{U} = (\tilde{u}, \tilde{v}) \colon [0,T] \to L^p(\Omega; X)$ of $(U^j)_{j=0,\ldots,N_k}$ by $\tilde{U}(t) := U^j$ for $t \in [t_j, t_{j+1})$, $0 \leq j \leq N_k - 1$, and $\tilde{U}(T) := U^{N_k}$. Let $\delta_1 := \min\{\delta, 1\}$. Then the following two error estimates hold.*

(a) (general schemes) *It holds that*

$$\Big\| \sup_{t\in[0,T]} \|u(t) - \tilde{u}(t)\|_V \Big\|_p \leq C_{\delta_1} k^{\delta_1} + C_{\mathrm{e}}\big(C_1 + C_2\sqrt{\max\{\log(T/k), p\}}\big)k^\alpha,$$

where $C_{\delta_1} := 2C_{U_0,\mathbf{f},\mathbf{g},X_{\delta_1}} + C_{U_0,\mathbf{f},\mathbf{g},X}$ with $C_{U_0,\mathbf{f},\mathbf{g},X_{\delta_1}}$ and $C_{U_0,\mathbf{f},\mathbf{g},X}$ are as defined in (4.5.8) *and C_{e}, C_1, C_2 are as in Theorem 4.48.*

(b) (exponential Euler) *If* $R = S$ *then*

$$\Big\| \sup_{t\in[0,T]} \|u(t) - \tilde{u}(t)\|_V \Big\|_p \leq (2C_{U_0,\mathbf{f},\mathbf{g},X_{\delta_1}} + C_{U_0,\mathbf{f},\mathbf{g},X})k^{\delta_1} + C_{\mathrm{S,e}}C_{\mathrm{S}} \cdot k^{\alpha},$$

where $C_{U_0,\mathbf{f},\mathbf{g},X_{\delta_1}}$ *and* $C_{U_0,\mathbf{f},\mathbf{g},X}$ *are as defined in* (4.5.8) *and* $C_{\mathrm{S,e}}, C_{\mathrm{S}}$ *are as in Corollary 4.49.*

Proof. Since the mild solution is also a weak solution to (4.5.1), writing $U = (u, v) \in L^p(\Omega; C([0,T]; V \times V_{-1}))$ we see that $(u(t) \,|\, \varphi)_{V_{-1}} - (u_0 \,|\, \varphi)_{V_{-1}} = \int_0^t (v(s) \,|\, \varphi)_{V_{-1}} \,\mathrm{d}s$ for all $\varphi \in V_{-1}$. Therefore, u is continuously differentiable as a (V_{-1})-valued function. Using the interpolation estimate $\|x\|_V \leq \|x\|_{V_{\delta_1-1}}^{\delta_1} \|x\|_{V_{\delta_1}}^{1-\delta_1}$ [64, Formula (1.1), Thm. 2.2], we find that

$$\begin{aligned}\|u(t) - u(s)\|_V &= \|u(t) - u(s)\|_{V_{\delta_1-1}}^{\delta_1} \|u(t) - u(s)\|_{V_{\delta_1}}^{1-\delta_1} \leq 2\|u(t) - u(s)\|_{V_{\delta_1-1}}^{\delta_1} \|u\|_{C([0,T];V_{\delta_1})}^{1-\delta_1} \\ &\leq 2|t-s|^{\delta_1} \|u'\|_{C([0,T];V_{\delta_1-1})}^{\delta_1} \|u\|_{C([0,T];V_{\delta_1})}^{1-\delta_1} \leq 2|t-s|^{\delta_1} \|U\|_{C([0,T];X_{\delta_1})}\end{aligned}$$

almost surely. In the last estimate, we have estimated by the maximum of the respective norms of u and u' prior to using the a priori estimate (4.4.4) w.r.t. the initial values for U. Inserting this estimate in the Hölder seminorm, we find that

$$\begin{aligned}\|u\|_{L^p(\Omega;C^{\delta_1}([0,T];V))} &\leq \|[u]_{C^{\delta_1}([0,T];V)}\|_p + \|u\|_{L^p(\Omega;C([0,T];V))} \\ &\leq 2\|U\|_{L^p(\Omega;C([0,T];X_{\delta_1}))} + \|U\|_{L^p(\Omega;C([0,T];X))} \\ &\leq 2C_{U_0,\mathbf{f},\mathbf{g},X_{\delta_1}} + C_{U_0,\mathbf{f},\mathbf{g},X}.\end{aligned}$$

By Lemma 4.28, we find that for $U^j = (u^j, v^j)$, we can split the error as

$$\sup_{t\in[0,T]} \|u(t) - \tilde{u}(t)\|_V \leq k^{\delta_1} \|u\|_{C^{\delta_1}([0,T];V)} + \max_{j=0,\ldots,N_k} \|u(t_j) - u^j\|_V.$$

Therefore, taking L^p-norms and using the error estimate of Theorem 4.48 we find that

$$\begin{aligned}\Big\| \sup_{t\in[0,T]} \|u(t) - \tilde{u}(t)\|_V \Big\|_p &\leq (2C_{U_0,\mathbf{f},\mathbf{g},X_{\delta_1}} + C_{U_0,\mathbf{f},\mathbf{g},X})k^{\delta_1} + \Big\| \max_{j=0,\ldots,N_k} \|U(t_j) - U^j\|_X \Big\|_p \\ &\leq (2C_{U_0,\mathbf{f},\mathbf{g},X_{\delta_1}} + C_{U_0,\mathbf{f},\mathbf{g},X})k^{\delta_1} + C_{\mathrm{e}}\big(C_1 + C_2\sqrt{\max\{\log(T/k), p\}}\big)k^{\alpha}.\end{aligned}$$

To obtain the second estimate, we use Corollary 4.49 in place of Theorem 4.48. □

4.5.4 Application to the wave equation with trace class noise

As an example, we consider the classical stochastic wave equation on an open and bounded subset $\mathcal{O} \subseteq \mathbb{R}^d$:

$$\begin{cases} \mathrm{d}\dot{u} = (\Delta u + F(u))\,\mathrm{d}t + G(u)\,\mathrm{d}W(t) & \text{on } [0,T], \\ u(0) = u_0, \ \dot{u}(0) = v_0, \end{cases} \tag{4.5.10}$$

with Dirichlet boundary conditions. In the current subsection, we consider trace class noise in L^2 for any $d \in \mathbb{N}$, and in Subsection 4.5.5 space-time white noise in case $d = 1$.

It is well-known that $\Lambda = -\Delta$ is a positive and self-adjoint operator on $L^2(\mathcal{O})$, which is invertible. Let $(W(t))_{t\in[0,T]}$ be a Q-Wiener process with $Q \in \mathcal{L}(L^2(\mathcal{O}))$ so that Q is positive and self-adjoint. Finite-dimensional noise is included, since Q need not be strictly positive. Assume

$$Q^{1/2} \in \mathcal{L}_2(L^2(\mathcal{O}), L^\infty(\mathcal{O})). \tag{4.5.11}$$

In particular, this implies $Q^{1/2} \in \mathcal{L}_2(L^2(\mathcal{O}), L^2(\mathcal{O}))$ and that Q is trace class. The latter follows from the Dunford–Pettis theorem in the version of [117, Cor. A.1.2], which allows us to calculate the trace of Q by integrating over the square-integrable kernel associated with $Q^{1/2}$.

We consider the stochastic wave equation (4.5.10) on $V := L^2(\mathcal{O})$ and set $H := L^2(\mathcal{O})$. For the nonlinearity and the multiplicative noise, we choose Nemytskij operators $F\colon V \to V$ and $G\colon V \to \mathcal{L}_2(H, V) = \mathcal{L}_2(L^2(\mathcal{O}), L^2(\mathcal{O}))$ determined by

$$F(u)(\xi) = \phi(\xi, u(\xi)), \quad (G(u)(h))(\xi) = \psi(\xi, u(\xi))Q^{1/2}h(\xi), \quad \xi \in \mathcal{O}. \tag{4.5.12}$$

Here, the measurable functions $\phi, \psi\colon \mathcal{O} \times \mathbb{R} \to \mathbb{R}$ are Lipschitz and of linear growth in the second coordinate, i.e., there is a constant $L \geq 0$ such that for all $u, u_1, u_2 \in \mathbb{R}$, $\xi \in \mathcal{O}$ it holds that

$$|\phi(\xi, u)| + |\psi(\xi, u)| \leq L(1 + |u|), \quad |\phi(\xi, u_1) - \phi(\xi, u_2)| + |\psi(\xi, u_1) - \psi(\xi, u_2)| \leq L|u_1 - u_2|. \tag{4.5.13}$$

It is clear that F is Lipschitz from V to V. To see that G is of linear growth, note that by (4.5.11)

$$|(G(u)h)(\xi)| = |\psi(\xi, u(\xi))||Q^{1/2}h(\xi)| \leq C_{\psi,Q}(1 + |u(\xi)|)\|h\|_H,$$

where $C_{\psi,Q} := L\|Q^{1/2}\|_{\mathcal{L}(L^2(\mathcal{O}),L^\infty(\mathcal{O}))}$. Therefore, arguing as in [72, Thm. 9.3.6 (3)⇒(4)] by Riesz' theorem we can find $k_u\colon \mathcal{O} \to H$ such that for a.e. $\xi \in \mathcal{O}$ for all $h \in H$, $(k_u(\xi) \,|\, h)_H = (G(u)h)(\xi)$, and $\|k_u(\xi)\|_H \leq C_{\psi,Q}(1+|u(\xi)|)$. Therefore, for an orthonormal basis $(h_n)_{n\in\mathbb{N}}$ of H, we find that

$$\begin{aligned}\|G(u)\|^2_{\mathcal{L}_2(H,V)} &= \sum_{n=1}^{\infty} \|G(u)h_n\|^2_V = \int_{\mathcal{O}} \sum_{n=1}^{\infty} |\,(k_u(\xi) \,|\, h_n)\,|^2 \,\mathrm{d}\xi = \int_{\mathcal{O}} \|k_u(\xi)\|^2_H \,\mathrm{d}\xi \\ &\leq C^2_{\psi,Q}\|1 + |u|\|^2_V \leq C^2_{\psi,Q}(|\mathcal{O}|^{1/2} + \|u\|_V)^2.\end{aligned}$$

with $|\mathcal{O}|$ denoting the Lebesgue measure of the set $\mathcal{O}$. Likewise, we obtain Lipschitz continuity of G noting that

$$|(G(u)h)(\xi) - (G(v)h)(\xi)| = |\psi(\xi, u(\xi)) - \psi(\xi, v(\xi))||Q^{1/2}h(\xi)| \leq C_{\psi,Q}|u(\xi) - v(\xi)|\|h\|_H.$$

In particular, F and G satisfy the required mapping properties of Assumption 4.43 for any $\delta \in (0, 1]$. Moreover, the semigroup associated with (4.5.10) is the wave semigroup $(S(t))_{t\geq 0}$.

As an immediate consequence of Theorem 4.48 and Corollary 4.49, this yields the following convergence estimate generalising [132, Cor. 4.2] to arbitrary contractive schemes and slightly more general Q-Wiener processes W.

Theorem 4.52 (Wave equation with trace class noise in L^2)**.** *Let $\mathcal{O} \subseteq \mathbb{R}^d$, $d \in \mathbb{N}$, be a bounded and open set, $p \in [2, \infty)$, and $0 < \alpha \leq \delta \leq 1$. Let $X := X_0$ and X_δ be as defined in* (4.5.2) *with $V := L^2(\mathcal{O})$, and let $(u_0, v_0) \in L^p_{\mathcal{F}_0}(\Omega; X_\delta)$. Let F and G be the Nemytskij operators as in* (4.5.12) *with ϕ and ψ satisfying* (4.5.13)*. Suppose the covariance operator $Q \in \mathcal{L}(L^2(\mathcal{O}))$ satisfies* (4.5.11)*. Let $(R_k)_{k>0}$ be a time discretisation scheme that is contractive on both X and X_δ. Suppose that R approximates S to order α on X_δ. Denote by U the mild solution of* (4.5.1) *with trace class noise and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in* (4.5.6)*. Then there exists a constant $C \geq 0$ depending on $(u_0, v_0, \phi, \psi, T, p, \alpha, \mathcal{O}, d, V, \delta)$ such that for $N_k \geq 2$*

$$\Big\| \max_{0 \leq j \leq N_k} \|U(t_j) - U^j\|_X \Big\|_p \leq C\big(1 + \|Q^{1/2}\|_{\mathcal{L}(L^2(\mathcal{O}), L^\infty(\mathcal{O}))}\big) \sqrt{\max\{\log(T/k), p\}}\, k^\alpha.$$

In particular, the approximations $(U^j)_j$ converge at rate 1 *if $(u_0, v_0) \in L^p_{\mathcal{F}_0}(\Omega; X_1)$ and the exponential Euler method $R = S$ is used. The logarithmic factor can be omitted in this case.*

In case $\delta = 1$, for implicit Euler and Crank–Nicolson, we can take $\alpha = 1/2$ and $\alpha = 2/3$, respectively. This is due to convergence at rate α on $\mathrm{D}((-A)^{2\alpha})$ and $\mathrm{D}((-A)^{3\alpha/2})$, respectively. Using higher-order schemes, we can come as close to rate 1 as we want. In Theorem 4.54 we show that for smoother noise $\alpha = 1$ can be reached even for implicit Euler.

4.5.5 Application to the wave equation with space-time white noise

We use the same notation as in Subsection 4.5.4, but this time with $\mathcal{O} = (0,1)$ and $Q = I$, so that (4.5.10) is the classical one-dimensional wave equation with space-time white noise. The required mapping properties can be checked as in [132, Cor. 4.3]. For convenience of the reader, we include the details. The functions F and G are defined by (4.5.12), however, in this subsection, we have to consider G as a mapping $G\colon V \to \mathcal{L}_2(H, V_{-1})$.

The eigenvalues of the negative Dirichlet Laplacian $\Lambda = -\Delta$ are $\lambda_j = \pi^2 j^2$, $j \in \mathbb{N}$, with the corresponding orthonormal basis $\{e_j = \sqrt{2}\sin(j\pi\cdot) : j \in \mathbb{N}\}$ of V consisting of eigenfunctions of Λ. Clearly,

$$\sup_{j \in \mathbb{N}} \sup_{\xi \in [0,1]} |e_j(\xi)| \leq \sqrt{2}, \quad \text{and } \|\Lambda^{-\frac{\varepsilon+1}{4}}\|^2_{\mathcal{L}(V)} = \pi^{-(\varepsilon+1)} \sum_{j=1}^\infty j^{-(\varepsilon+1)} =: c_\varepsilon < \infty$$

then hold for every $\varepsilon > 0$. Now let $\varepsilon \in (0,1]$. Using the properties above as well as self-adjointness of Λ, we conclude that

$$\begin{aligned}
\|\Lambda^{-\frac{\varepsilon+1}{4}} G(u)\|^2_{\mathcal{L}_2(H,V)} &= \sum_{i=1}^\infty \sum_{j=1}^\infty |\langle G(u)e_i, \Lambda^{-\frac{\varepsilon+1}{4}} e_j\rangle_V|^2 \\
&= \sum_{i=1}^\infty \sum_{j=1}^\infty \lambda_j^{-\frac{\varepsilon+1}{2}} \Big| \int_{\mathcal{O}} \psi(\xi, u(\xi)) e_i(\xi) e_j(\xi)\, \mathrm{d}\xi \Big|^2 \\
&\leq 2\Big(\sum_{j=1}^\infty \lambda_j^{-\frac{\varepsilon+1}{2}}\Big) \|\psi(\cdot, u(\cdot))\|^2_V \leq 2L^2 c_\varepsilon (|\mathcal{O}|^{1/2} + \|u\|_V)^2.
\end{aligned}$$

Hence, G satisfies the linear growth condition of Assumption 4.43 with $\delta = \frac{1-\varepsilon}{2}$. Repeating the arguments for $\Lambda^{-1/2}[G(u_1) - G(u_2)]$ and using $c_1 = \pi^2/6$ results in

$$\|\Lambda^{-1/2}[G(u_1) - G(u_2)]\|^2_{\mathcal{L}_2(H,V)} \leq 2c_1\|\psi(\cdot, u_1(\cdot)) - \psi(\cdot, u_2(\cdot))\|^2_V \leq \frac{\pi^2 L^2}{3}\|u_1 - u_2\|^2_V.$$

The nonlinearity F has already been considered in Subsection 4.5.4. In conclusion, we obtain the following generalisation of [132, Cor. 4.3] to contractive time discretisation schemes.

Theorem 4.53 (Wave equation with white noise). *Let $\mathcal{O} = (0, 1)$, $p \in [2, \infty)$, and $0 < \alpha \leq \delta < 1/2$. Let $X := X_0$ and X_δ be as defined in (4.5.2) with $V := L^2(\mathcal{O})$, and let $(u_0, v_0) \in L^p_{\mathcal{F}_0}(\Omega; X_\delta)$. Let F and G be Nemytskij operators as above, with ϕ and ψ satisfying (4.5.13). Suppose the covariance operator Q equals I on $L^2(\mathcal{O})$. Let $(R_k)_{k>0}$ be a time discretisation scheme that is contractive on X and X_δ. Assume that R approximates S on X_δ to order α. Denote by U the mild solution of (4.5.1) with space-time white noise and by $(U^j)_{j=0,\ldots,N_k}$ the temporal approximations as defined in (4.5.6). Then there exists a constant $C \geq 0$ depending on $(u_0, v_0, \phi, \psi, T, p, \alpha, \mathcal{O}, d, V, \delta)$ such that for $N_k \geq 2$*

$$\Big\| \max_{0\leq j\leq N_k} \|U(t_j) - U^j\|_X \Big\|_p \leq C\sqrt{\max\{\log(T/k), p\}}\,k^\alpha.$$

In particular, the approximations $(U^j)_j$ converge at rate arbitrarily close to $\frac{1}{2}$ if $(u_0, v_0) \in L^p_{\mathcal{F}_0}(\Omega; X_1)$ and the exponential Euler method $R = S$ is used. The logarithmic factor can be omitted in this case.

For implicit Euler and Crank–Nicolson, we can take $\alpha = \delta/2$ and $\alpha = 2\delta/3$, respectively. Since we can choose δ arbitrarily close to $1/2$, this leads to rates which are almost $1/4$ and $1/3$, respectively.

4.5.6 Application to the wave equation with smooth noise

We have already seen that exponential Euler leads to convergence at any rate $\alpha \in (0, 1]$ depending on the given data. In this section, we show that this can also be attained for other schemes, such as implicit Euler and Crank–Nicolson, under some smoothness conditions on the noise. To avoid problems with boundary conditions, we only consider periodic boundary conditions. Consider

$$\begin{cases} \mathrm{d}\dot{u} = ((\Delta - 1)u + F(u))\,\mathrm{d}t + G(u)\,\mathrm{d}W(t) & \text{on } [0, T], \\ u(0) = u_0,\ \dot{u}(0) = v_0, \end{cases} \tag{4.5.14}$$

with $\Lambda = 1 - \Delta$ and periodic boundary conditions on the d-dimensional torus $\mathbb{T}^d = [0, 1]^d$. For notational convenience, we will write $H^\beta = H^\beta(\mathbb{T}^d) =: V_\beta$ for $\beta \in \mathbb{R}$. Note that $\|\Lambda^{-\beta}\|_{\mathcal{L}(L^2)} \leq 1$ for all $\beta > 0$. The additional $+1$ in the definition of Λ is included in order to ensure invertibility. Of course, F can be suitably redefined so that this is without loss of generality.

Let $\delta \in (1, 2]$ and write $s := \delta - 1$. Let

$$F(u)(\xi) = \phi(u(\xi)), \quad (G(u)(h))(\xi) = \psi(u(\xi))Q^{1/2}h(\xi), \quad \xi \in \mathbb{T}^d. \tag{4.5.15}$$

Here, the measurable functions $\phi, \psi \colon \mathbb{R} \to \mathbb{R}$ are Lipschitz with Lipschitz constants L_ϕ and L_ψ, respectively, and, contrary to the two previous subsections, do not depend on ξ itself. The Lipschitz estimates for F and G follow as in Subsection 4.5.4, since we will assume more restrictive conditions on Q. The growth estimates for F and G as in Assumption 4.43(d) are more complicated. In case $\delta = 2$ the paraproduct constructions from [122] can be avoided, but we will consider the general case.

By the torus version of Proposition 2.33, for $u \in V_\delta$, there is a constant $C_{s,\phi} \geq 0$ such that

$$\|F(u)\|_{V_{\delta-1}} = \|\phi(u)\|_{H^{\delta-1}} \leq C_{s,\phi}(1 + \|u\|_{H^{\delta-1}}) \leq C_{s,\phi}(1 + \|u\|_{H^\delta}) = C_{s,\phi}(1 + \|u\|_{V_\delta}).$$

For G, the estimate is still more complicated. In order to estimate the Hilbert–Schmidt norm of $G(u)$, paraproduct estimates are required, as, for instance, in (4.5.17). These paraproduct estimates involve Bessel potential spaces $H^{s,q}$, which, in general, are not Hilbert spaces. Thus, γ-radonifying operators are required in this application. Let $(\gamma_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence of standard Gaussian random variables taking values in $\mathbb{R}$. Suppose that $\Lambda^{\frac{\delta-1}{2}} Q^{1/2} \colon L^2 \to L^\infty$. Then by [72, Cor. 9.3.3], $Q^{1/2} \in \gamma(H, H^{\beta,q})$ for all $q \in [1,\infty)$ and all $\beta \leq \delta - 1$, and

$$C_{q,\beta} := \|Q^{1/2}\|_{\gamma(H,H^{\beta,q})} \leq \|Q^{1/2}\|_{\gamma(H,H^{\delta-1,q})} \leq c_q \|\Lambda^{\frac{\delta-1}{2}} Q^{1/2}\|_{\mathcal{L}(L^2,L^\infty)}, \tag{4.5.16}$$

where $c_q := \|\gamma_1\|_{L^q(\Omega)}$. Let $(h_n)_{n\in\mathbb{N}}$ be an orthonormal basis of H and fix $N \geq 1$. Let $\eta_N := \sum_{n=1}^N \gamma_n Q^{1/2} h_n \in L^2(\Omega; V_{\delta-1})$. Then $\|\eta_N\|_{L^2(\Omega;V_\beta)} \leq \|Q^{1/2}\|_{\gamma(H,H^{\beta,q})}$ for all $\beta \leq \delta - 1$. It follows that

$$\sum_{n=1}^N \|G(u)h_n\|^2_{V_{\delta-1}} = \|\psi(u)\eta_N\|^2_{L^2(\Omega;V_{\delta-1})}.$$

Next, we estimate $\|\psi(u)\eta_N\|_{V_{\delta-1}}$ pointwise in Ω. By the torus version of Proposition 2.32 (see [1, Prop. 4.1(1)]) and Proposition 2.33, there is a constant $C_{\delta,d,1} \geq 0$ such that

$$\begin{aligned} \|\psi(u)\eta_N\|_{V_{\delta-1}} &= \|\psi(u)\eta_N\|_{H^{\delta-1}} \leq \|\psi(u)\|_{L^{q_1}}\|\eta_N\|_{H^{\delta-1,q_2}} + \|\psi(u)\|_{H^{\delta-1,r_2}}\|\eta_N\|_{L^{r_1}} \\ &\leq L_\psi(\|u\|_{L^{q_1}} + 1)\|\eta_N\|_{H^{\delta-1,q_2}} + L_\psi C_{\delta,d,1}(\|u\|_{H^{\delta-1,r_2}} + 1)\|\eta_N\|_{H^{\delta-1,r_1}}, \end{aligned} \tag{4.5.17}$$

where $\frac{1}{q_1} + \frac{1}{q_2} = \frac{1}{r_1} + \frac{1}{r_2} = \frac{1}{2}$ and $q_1, r_1 \in (2,\infty]$ and $q_2, r_2 \in [2,\infty)$. Taking $r_1 < \infty$ and using (4.5.16), we find that

$$\|\psi(u)\eta_N\|_{L^2(\Omega;V_{\delta-1})} \leq L_\psi C_{q_2,\delta-1}(\|u\|_{L^{q_1}} + 1) + L_\psi C_{\delta,d,1} C_{r_1,\delta-1}(\|u\|_{H^{\delta-1,r_2}} + 1)$$

for suitable constants $C_{q_2,\delta-1}, C_{r_1,\delta-1} \geq 0$. It remains to estimate $\|u\|_{L^{q_1}}$ and $\|u\|_{H^{\delta-1,r_2}}$ by $\|u\|_{H^\delta} = \|u\|_{V_\delta}$ using suitable Sobolev embeddings and choosing $q_1 \in (2,\infty]$ and $r_2 \in (2,\infty)$ suitably. As soon as we have done that, we can let $N \to \infty$ and conclude the required estimate

$$\|G(u)\|_{\mathcal{L}_2(H,V_{\delta-1})} \leq K(1 + \|u\|_{V_\delta}).$$

To obtain $H^\delta \hookrightarrow L^{q_1}$, we consider two cases. If $\delta \leq d/2$ (e.g., $d \in \{1,2\}$) we can take $q_1 < \infty$ arbitrary. If $\delta > d/2$, then we take $q_1 = \frac{2d}{d-2\delta}$, and thus $q_2 = \frac{d}{\delta}$.

To obtain $H^\delta \hookrightarrow H^{\delta-1,r_2}$ we consider two cases. If $d \in \{1,2\}$, then we can take $r_2 \in (2,\infty)$ arbitrary. If $d \geq 3$, then we set $r_2 = \frac{2d}{d-2}$, and thus $r_1 = d$.

Theorem 4.54 (Wave equation with smooth noise)**.** *Let $X := X_0$ and X_δ be as defined in (4.5.2) with $V := L^2(\mathcal{O})$, $p \in [2, \infty)$, and $0 < \alpha \le 1 < \delta \le 2$. Suppose that $(u_0, v_0) \in L^p_{\mathcal{F}_0}(\Omega; X_\delta)$. Let F and G be Nemytskij operators as in (4.5.15) with Lipschitz functions ϕ and ψ. Suppose the covariance operator Q on $L^2(\mathcal{O})$ satisfies $\Lambda^{\frac{\delta-1}{2}} Q^{1/2} \in \mathcal{L}(L^2(\mathbb{T}^d), L^\infty(\mathbb{T}^d))$. Let $(R_k)_{k>0}$ be a time discretisation scheme that is contractive on both X and X_δ. Assume that R approximates S to order α on X_δ. Denote by U the mild solution of (4.5.14) driven by a Q-Wiener process W and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in (4.5.6). Then there exists a constant $C \ge 0$ depending on $(u_0, v_0, \phi, \psi, T, p, \alpha, d, V, \delta)$ such that for $N_k \ge 2$*

$$\Big\| \max_{0 \le j \le N_k} \|U(t_j) - U^j\|_X \Big\|_p \le C\big(1 + \|\Lambda^{(\delta-1)/2} Q^{1/2}\|_{\mathcal{L}(L^2(\mathbb{T}^d), L^\infty(\mathbb{T}^d))}\big) \sqrt{\max\{\log(T/k), p\}}\, k^\alpha.$$

The above result is not useful for exponential Euler, since Theorem 4.52 is better in that case. However, if we specialize to implicit Euler and Crank–Nicolson, then we obtain rates $\alpha = \frac{\delta}{2}$ and $\alpha = \min\{\frac{2}{3}\delta, 1\}$, respectively. In particular, this leads to convergence of order one if $\delta = 2$ for many numerical schemes. Note that $\delta = 2$ approximately corresponds to a noise W that belongs to $H^{1,q}(\mathbb{T}^d)$ for all $q < \infty$.

Remark 4.55. Theorem 4.54 gives an explanation for the convergence rates obtained in [132, Fig. 6.1, right figure] from numerical experiments. There, trace class noise determined by $\psi(u) = u$ and Q with eigenvalues $q_j = j^{-\beta}$, $j \in \mathbb{N}$, $\beta = 1.1$ has been investigated. Denote by $(e_j)_{j\in\mathbb{N}}$ the orthonormal basis of V and by $\lambda_j = Cj^2$ the eigenvalues of Λ as in Subsection 4.5.5 for some constant $C > 0$. We calculate that

$$\Lambda^{\frac{\delta-1}{2}} Q^{\frac{1}{2}} e_j = q_j^{\frac{1}{2}} \Lambda^{\frac{\delta-1}{2}} e_j = j^{-\frac{\beta}{2}} \lambda_j^{\frac{\delta-1}{2}} e_j = C^{\frac{\delta-1}{2}} j^{\delta-1-\frac{\beta}{2}} e_j$$

for $j \in \mathbb{N}$. Thus, $\Lambda^{\frac{\delta-1}{2}} Q^{\frac{1}{2}}$ maps L^2 into L^∞ if $\delta \le 1 + \frac{\beta}{2}$. Setting $\delta := \min\{1 + \frac{\beta}{2}, 2\} = 1 + \frac{1.1}{2} = 1.55$, we derive convergence at rate $\frac{\delta}{2} = 0.775$ for implicit Euler and $\min\{\frac{2}{3}\delta, 1\} = 1$ for Crank–Nicolson. Taking numerical errors into account, this corresponds to the numerical convergence rates obtained in [132, Fig. 6.1, right figure].

4.6 Pathwise Uniform Convergence for Irregular Nonlinearities

As in the preceding sections, our aim is to prove pathwise uniform convergence of discretisations

$$U^0 \coloneqq u_0, \quad U^j \coloneqq R_k U^{j-1} + k R_k F(t_{j-1}, U^{j-1}) + R_k G(t_{j-1}, U^{j-1}) \Delta W_j, \tag{4.6.1}$$

$j = 1, \ldots, N_k$, obtained from contractive time discretisation schemes $(R_k)_{k>0}$ to the mild solution of a nonlinear stochastic evolution equation of the form

$$\mathrm{d}U + AU \,\mathrm{d}t = F(t, U) \,\mathrm{d}t + G(t, U) \,\mathrm{d}W_H(t), \quad U(0) = u_0 \in L^p(\Omega; X). \tag{4.6.2}$$

In contrast to Section 4.4, we shall not assume any further regularity in the structure of the nonlinearity $F\colon \Omega \times [0,T] \times X \to X$ or the noise $G\colon \Omega \times [0,T] \times X \to \gamma(H,X)$. Instead, we merely assume global Lipschitz continuity and progressive measurability on X, and impose no further conditions on the images $F(\Omega \times [0,T] \times Y)$ for some $Y \hookrightarrow X$ or even on $F(\Omega \times [0,T] \times X)$ being proper and more regular subspaces of X, or $\gamma(H,X)$ in the case of G. Moreover, we allow rough initial data $u_0 \in L^p(\Omega; X)$.

Naturally, allowing such irregular nonlinearities and initial data, we cannot expect convergence of a certain rate. Even in the linear, deterministic case, convergence can be arbitrarily slow for irregular initial data. Hence, the main result of this section in Theorem 4.57 is a qualitative pathwise uniform convergence statement in the irregular setting described above.

Subsection 4.6.1 is devoted to the proof of the main theorem, which relies on martingale techniques. An example of an irregular stochastic Schrödinger equation, for which none of the quantified convergence results are applicable, is considered in Subsection 4.6.2.

4.6.1 The convergence result

The following assumption summarises the conditions imposed on F and G. We will later see that they are sufficient for pathwise uniform convergence. Since no space Y with additional regularity is required, the space index X is omitted in the constants.

Assumption 4.56. *Let X be a $(2,D)$-smooth Banach space and $p \in [2,\infty)$. Let $F\colon \Omega \times [0,T] \times X \to X$ and $G\colon \Omega \times [0,T] \times X \to \gamma(H,X)$ be strongly $\mathcal{P} \otimes \mathcal{B}(X)$-measurable. Suppose that*

(a) (global Lipschitz continuity) *there exist constants $C_F, C_G \geq 0$ such that for all $\omega \in \Omega$, $t \in [0,T]$, and $x, y \in X$,*

$$\|F(\omega,t,x) - F(\omega,t,y)\| \leq C_F \|x-y\|, \ \|G(\omega,t,x) - G(\omega,t,y)\|_{\gamma(H,X)} \leq C_G \|x-y\|,$$

(b) (linear growth) *there exist constants $L_F, L_G \geq 0$ such that for all $\omega \in \Omega$, $t \in [0,T]$, and $x \in X$,*

$$\|F(\omega,t,x)\| \leq L_F(1 + \|x\|), \quad \|G(\omega,t,x)\|_{\gamma(H,X)} \leq L_G(1 + \|x\|),$$

(c) (Hölder continuity) *for some* $\alpha \in (0,1]$,

$$C_{\alpha,F} := \sup_{\omega\in\Omega, x\in X} [F(\omega,\cdot,x)]_\alpha < \infty, \; C_{\alpha,G} := \sup_{\omega\in\Omega, x\in X} [G(\omega,\cdot,x)]_\alpha < \infty.$$

Assumption 4.56 implies Assumption 4.11 with $\tilde{F} := F - F(\cdot,\cdot,0)$, $f := F(\cdot,\cdot,0)$, and likewise for G, whence (4.6.2) has a unique mild solution. Contrary to Section 4.4, we do not decompose $F = \tilde{F} + f$ or $G = \tilde{G} + g$ in this section, as the main purpose of this decomposition was to relax the integrability assumptions with values in Y. Since the main objective of this section is to not impose any conditions on Y, we do not perform this decomposition. As a consequence, we can no longer impose conditions on the value at time 0 without loss of generality. Hence, linear growth of F is assumed directly instead of deducing it from Lipschitz continuity. Compared to Assumption 4.11, only condition (c) is added. Under these assumptions, we obtain the main result of this section on pathwise uniform convergence of the temporal approximations.

Theorem 4.57. *Suppose that Assumption 4.56 holds for some* $p \in [2,\infty)$ *and* $\alpha \in (0,1]$. *Further suppose that* $-A$ *generates a* C_0*-contraction semigroup on both* X *and* $\mathrm{D}(A)$, *and let* $u_0 \in L^p_{\mathcal{F}_0}(\Omega;X)$. *Let* $(R_k)_{k>0}$ *be a time discretisation scheme that is contractive on* X *and* $\mathrm{D}(A)$. *Assume* R *approximates* S *to order* α *on* $\mathrm{D}(A)$. *Denote by* U *the mild solution of* (4.6.2) *and by* $(U^j)_{j=0,\dots,N_k}$ *the temporal approximations as defined in* (4.6.1). *Define the piecewise constant extension* $\tilde{U}\colon [0,T] \to L^p(\Omega;X)$ *by* $\tilde{U}(t) := U^j$ *for* $t \in [t_j, t_{j+1})$, $0 \le j \le N_k - 1$, *and* $\tilde{U}(T) := U^{N_k}$. *Then*

$$\lim_{k\to 0} \Big\| \sup_{t\in[0,T]} \|U(t) - \tilde{U}(t)\| \Big\|_p = 0. \tag{4.6.3}$$

The main ingredient of the proof of this theorem consists of regularising the nonlinearity, the noise, and the initial values by

$${}_mF := mR(m,-A)F, \quad {}_mG := mR(m,-A)G, \quad {}_mu_0 := mR(m,-A)u_0 \tag{4.6.4}$$

for $m \in \mathbb{N}$. By construction, ${}_mF$ maps to $\mathrm{D}(A)$, ${}_mG$ maps to $\gamma(H,\mathrm{D}(A))$, and ${}_mu_0 \in L^p(\Omega;\mathrm{D}(A))$, giving the desired additional regularity in structure. Assumption 4.56 also implies existence and uniqueness of the mild solution ${}_mU$ of the regularised problem

$${}_mU + A{}_mU \,\mathrm{d}t = {}_mF({}_mU)\,\mathrm{d}t + {}_mG({}_mU)\,\mathrm{d}W_H(t), \quad {}_mU(0) = {}_mu_0 \in X \tag{4.6.5}$$

for $m \in \mathbb{N}$. It is given by a fixed point of

$${}_mU(t) = S(t){}_mu_0 + \int_0^t S(t-s){}_mF({}_mU(s))\,\mathrm{d}s + \int_0^t S(t-s){}_mG({}_mU(s))\,\mathrm{d}W_H(s).$$

The following proposition lists useful properties of the regularised quantities.

Proposition 4.58. *Let Assumption 4.56 hold and let* ${}_mF, {}_mG, {}_mu_0$ *be as defined in* (4.6.4) *for* $m \in \mathbb{N}$. *Suppose that* $-A$ *generates a* C_0*-contraction semigroup* $(S(t))_{t\ge 0}$. *Let* $u_0 \in L^p_{\mathcal{F}_0}(\Omega;X)$. *Then the following statements hold.*

(a) ($\mathrm{D}(A)$-invariance) ${}_mF\colon \Omega \times [0,T] \times \mathrm{D}(A) \to \mathrm{D}(A)$ *and* ${}_mG\colon \Omega \times [0,T] \times \mathrm{D}(A) \to \gamma(H,\mathrm{D}(A))$ *are strongly* $\mathcal{P} \otimes \mathcal{B}(\mathrm{D}(A))$*-measurable and* ${}_mu_0 \in L^p_{\mathcal{F}_0}(\Omega;\mathrm{D}(A))$.

(b) (uniform Lipschitz continuity) ${}_mF$ *and* ${}_mG$ *are uniformly Lipschitz continuous with the same Lipschitz constants as* F *and* G*, i.e., for all* $m \in \mathbb{N}$*,* $\omega \in \Omega$*,* $t \in [0,T]$*, and* $x, y \in X$*,*

$$\|{}_mF(\omega,t,x) - {}_mF(\omega,t,y)\| \le C_F\|x-y\|,$$
$$\|{}_mG(\omega,t,x) - {}_mG(\omega,t,y)\|_{\gamma(H,X)} \le C_G\|x-y\|.$$

(c) (linear growth on $\mathrm{D}(A)$) *For all* $m \in \mathbb{N}$*, there are constants* $L_{F,m}, L_{G,m} \ge 0$ *such that for all* $\omega \in \Omega$*,* $t \in [0,T]$*, and* $x \in Y$*,*

$$\|{}_mF(\omega,t,x)\|_{\mathrm{D}(A)} \le L_{F,m}(1 + \|x\|_{\mathrm{D}(A)}),$$
$$\|{}_mG(\omega,t,x)\|_{\gamma(H,\mathrm{D}(A))} \le L_{G,m}(1 + \|x\|_{\mathrm{D}(A)}).$$

(d) (pointwise convergence) *As* $m \to \infty$*,* ${}_mF$ *and* ${}_mG$ *converge pointwise to* F *and* G*, respectively. Moreover,* ${}_mu_0 \to u_0$ *in* $L^p(\Omega; X)$ *as* $m \to \infty$*.*

Proof. (a) Continuity of $F(\omega,t,\cdot)\colon X \to X$ for all $\omega \in \Omega$ and $t \in [0,T]$ follows from Assumption 4.11(a), and thus also continuity as a mapping $F(\omega,t,\cdot)\colon \mathrm{D}(A) \to X$. From the identity $AR(m,-A) = I - mR(m,-A)$ and continuity of the resolvent on X, we obtain continuity of $R(m,-A)\colon X \to \mathrm{D}(A)$. Consequently, ${}_mF(\omega,t,\cdot)\colon \mathrm{D}(A) \to \mathrm{D}(A)$ is continuous. Hence, strong $\mathcal{P} \otimes \mathcal{B}(\mathrm{D}(A))$-measurability of ${}_mF\colon \Omega \times [0,T] \times \mathrm{D}(A) \to \mathrm{D}(A)$ follows from strong $\mathcal{P}\otimes\mathcal{B}(X)$-measurability as stated in Assumption 4.11. Likewise, strong $\mathcal{P} \otimes \mathcal{B}(\mathrm{D}(A))$-measurability of ${}_mG$ can be derived. Lastly, since $R(m,-A)$ maps to $\mathrm{D}(A)$ and $u_0 \in L^p_{\mathcal{F}_0}(\Omega; X)$, it holds that ${}_mu_0 \in L^p_{\mathcal{F}_0}(\Omega; \mathrm{D}(A))$.

(b) First, we recall a folklore result from semigroup theory, which is an immediate consequence of Theorem 2.8(c): For contraction semigroups, the norm of the resolvent $R(\lambda,-A)$ is bounded by $\|R(\lambda,-A)\|_{\mathcal{L}(X)} \le (\operatorname{Re}\lambda)^{-1}$ for all $\operatorname{Re}\lambda > 0$. Hence, $\|mR(m,-A)\|_{\mathcal{L}(X)} \le m \cdot \frac{1}{m} = 1$ is contractive. Together with this observation, Lipschitz continuity of F and G implies uniform Lipschitz continuity of ${}_mF$ and ${}_mG$ with the same Lipschitz constant, respectively.

(c) By assumption, F is of linear growth on X. Linear growth of ${}_mF$ on $\mathrm{D}(A)$ with $L_{F,m} = (2m+1)L_F$ follows from the identity $AR(m,-A) = I - mR(m,-A)$ via

$$\begin{aligned}
\|{}_mF(\omega,t,x)\|_{\mathrm{D}(A)} &= \|mAR(m,-A)F(\omega,t,x)\| + \|mR(m,-A)F(\omega,t,x)\| \\
&\le \|m[I - mR(m,-A)]F(\omega,t,x)\| + \|F(\omega,t,x)\| \\
&\le (2m+1)\|F(\omega,t,x)\| \le (2m+1)L_F(1 + \|x\|).
\end{aligned}$$

Linear growth of ${}_mG$ on $\mathrm{D}(A)$ with $L_{G,m} = (2m+1)L_G$ follows analogously.

(d) It suffices to prove that $mR(m,-A) \to I$ in the strong operator topology as $m \to \infty$. Since A is densely defined and closed as the negative generator of a C_0-semigroup, this follows from Lemma 2.9. □

For a comprehensive discretisation error analysis, we start by investigating the continuous regularisation error. The following lemma will prove helpful in doing so for $Z \in \{X, \gamma(H,X)\}$ and with ψ chosen based on the nonlinearity F or noise G.

Lemma 4.59. *Let Z be a Banach space, $\psi\colon \Omega \times [0,T] \to Z$ have continuous paths almost surely and assume that*

$$\sup_{t\in[0,T]} \|\psi(\cdot,t)\|_Z \in L^p(\Omega).$$

Let $R_n, R \in \mathcal{L}(Z)$, $n \in \mathbb{N}$, be such that $R_n \to R$ strongly as $n \to \infty$. Then

$$\lim_{n\to\infty} \Big\| \sup_{t\in[0,T]} \|(R_n - R)\psi(\cdot,t)\|_Z \Big\|_p = 0.$$

Proof. By continuity of paths of ψ, the set $\psi(\omega,[0,T]) \subseteq Z$ is compact for a.e. $\omega \in \Omega$. Since by assumption R_n converges to R in the strong operator topology, Proposition 2.17 yields uniform convergence of R_n to R on compact sets in Z as $n \to \infty$ for a.e. $\omega \in \Omega$. Hence,

$$\sup_{t\in[0,T]} \|(R_n - R)\psi(\omega,t)\|_Z \xrightarrow[n\to\infty]{} 0 \qquad \text{for a.e. } \omega \in \Omega.$$

Due to the assumed integrability of the supremum of ψ in time, the desired statement follows from dominated convergence in $L^p(\Omega)$. □

Lemma 4.60 (Convergence of the continuous regularisation)**.** *Suppose that Assumption 4.56 holds for some $p \in [2,\infty)$. Suppose that $-A$ generates a C_0-contraction semigroup $(S(t))_{t\geq 0}$ on X and let $u_0 \in L^p_{\mathcal{F}_0}(\Omega;X)$. Denote by U the mild solution of* (4.6.2) *and by ${}_mU$ the mild solution of* (4.6.5) *with ${}_mF$, ${}_mG$, and ${}_mu_0$ as defined in* (4.6.4) *for $m \in \mathbb{N}$. Then*

$$\lim_{m\to\infty} \Big\| \sup_{t\in[0,T]} \|U(t) - {}_mU(t)\| \Big\|_p = 0. \tag{4.6.6}$$

Proof. Let ${}_mV \coloneqq U - {}_mU$ and $\tau \in [0,T]$. Then ${}_mV$ is given by

$$\begin{aligned} {}_mV(t) = S(t)[u_0 - {}_mu_0] &+ \int_0^t S(t-s)[F(s,U(s)) - {}_mF(s,{}_mU(s))]\,\mathrm{d}s \\ &+ \int_0^t S(t-s)[G(s,U(s)) - {}_mG(s,{}_mU(s))]\,\mathrm{d}W_H(s), \end{aligned}$$

which implies

$$\begin{aligned} {}_mE_1(\tau) &\coloneqq \Big\| \sup_{t\in[0,\tau]} \|{}_mV(t)\| \Big\|_p \\ &\leq \Big\| \sup_{t\in[0,\tau]} \|S(t)[u_0 - {}_mu_0]\| \Big\|_p + \Big\| \sup_{t\in[0,\tau]} \int_0^t \|S(t-s)[F(s,U(s)) - {}_mF(s,{}_mU(s))]\|\,\mathrm{d}s \Big\|_p \\ &\quad + \Big\| \sup_{t\in[0,\tau]} \Big\| \int_0^t S(t-s)[G(s,U(s)) - {}_mG(s,{}_mU(s))]\,\mathrm{d}W_H(s) \Big\| \Big\|_p \\ &=: {}_mE_{1,1}(\tau) + {}_mE_{1,2}(\tau) + {}_mE_{1,3}(\tau). \end{aligned}$$

We proceed to bound the terms individually. For the initial value term, contractivity of S, strong convergence of $mR(m,-A)$ to I on X as obtained from Lemma 2.9, and dominated convergence in $L^p(\Omega)$ yield the existence of $m_0 \in \mathbb{N}$ such that for all $m \geq m_0$,

$${}_mE_{1,1}(\tau) \leq \|u_0 - {}_mu_0\|_{L^p(\Omega;X)} = \|[I - mR(m,-A)]u_0\|_{L^p(\Omega;X)} < \frac{\varepsilon}{3}. \tag{4.6.7}$$

Next, we estimate

$$\begin{aligned}
{}_mE_{1,2}(\tau) &\le \Big\| \int_0^\tau \|F(s,U(s)) - {}_mF(s,U(s))\| \,\mathrm{d}s \Big\|_p + C_F \Big\| \int_0^\tau \|U(s) - {}_mU(s)\| \,\mathrm{d}s \Big\|_p \\
&\le \tau \Big\| \sup_{s\in[0,\tau]} \|F(s,U(s)) - {}_mF(s,U(s))\| \Big\|_p + C_F \int_0^\tau \Big\| \sup_{r\in[0,s]} \|{}_mV(r)\| \Big\|_p \,\mathrm{d}s
\end{aligned}$$

using contractivity of S and uniform Lipschitz continuity of ${}_mF$. The second term is a multiple of the integral of the full error ${}_mE_1$ from 0 to τ, allowing for a Gronwall argument after bounding the first term. By Theorem 4.12, U has continuous paths almost surely. Combined with continuity of F in time and space as derived from Assumption 4.56, this implies that $\psi\colon \Omega\times[0,T] \to X$, $\psi(\omega,t) := F(\omega,t,U(\omega,t))$ also has continuous paths almost surely. Furthermore, linear growth and Theorem 4.12 imply

$$\Big\| \sup_{t\in[0,\tau]} \|F(t,U(t))\| \Big\|_p \le L_F\big(1 + \|U\|_{L^p(\Omega;C([0,T];X))}\big) < \infty.$$

Hence, Lemma 4.59 applied to ψ on $Z = X$ with $R_n = nR(n,-A)$ and $R = I$ yields the existence of $m_1 \in \mathbb{N}$ such that for all $m \ge m_1$,

$$\Big\| \sup_{s\in[0,\tau]} \|F(s,U(s)) - {}_mF(s,U(s))\| \Big\|_p < \frac{\varepsilon}{3T}.$$

In conclusion, for $m \ge m_1$ the Cauchy–Schwarz inequality yields

$${}_mE_{1,2}(\tau) < \frac{\varepsilon}{3} + C_F \int_0^\tau {}_mE_1(s)\,\mathrm{d}s \le \frac{\varepsilon}{3} + C_F T^{1/2}\Big(\int_0^\tau {}_mE_1(s)^2\,\mathrm{d}s\Big)^{1/2}. \tag{4.6.8}$$

It remains to estimate the part of the error induced by noise. Via the maximal inequality from Theorem 2.61, the triangle inequality in $L^p(\Omega; L^2(0,\tau;\gamma(H,X)))$, uniform Lipschitz continuity of ${}_mG$, and Fubini's theorem, we obtain

$$\begin{aligned}
{}_mE_{1,3}(\tau) &\le C_{p,D} \Big\| \Big(\int_0^\tau \|G(s,U(s)) - {}_mG(s,{}_mU(s))\|^2_{\gamma(H,X)} \,\mathrm{d}s \Big)^{1/2} \Big\|_p \\
&\le C_{p,D} \Big\| \Big(\int_0^\tau \|G(s,U(s)) - {}_mG(s,U(s))\|^2_{\gamma(H,X)} \,\mathrm{d}s \Big)^{1/2} \Big\|_p \\
&\quad + C_{p,D} C_G \Big\| \int_0^\tau \|U(s) - {}_mU(s)\|^2 \,\mathrm{d}s \Big\|_{p/2}^{1/2} \\
&\le C_{p,D} \tau^{1/2} \Big\| \sup_{s\in[0,\tau]} \big\|G(s,U(s)) - {}_mG(s,U(s))\big\|_{\gamma(H,X)} \Big\|_p \\
&\quad + C_{p,D} C_G \Big(\int_0^\tau \Big\| \sup_{r\in[0,s]} \|U(r) - {}_mU(r)\| \Big\|_p^2 \,\mathrm{d}s \Big)^{1/2}.
\end{aligned}$$

By the left ideal property of $\gamma(H,X)$, see Proposition 2.39, the resolvent $R(m,-A)$ extends to a linear and bounded operator on $\gamma(H,X)$ for $m \in \rho(-A)$. Hence, arguing as for the nonlinear terms, Lemma 4.59 with $Z = \gamma(H,X)$ and $\psi(\omega,t) = G(\omega,t,U(\omega,t))$ yields the existence of $m_2 \in \mathbb{N}$ such that for all $m \ge m_2$,

$$\Big\| \sup_{s\in[0,\tau]} \|G(s,U(s)) - {}_mG(s,U(s))\|_{\gamma(H,X)} \Big\|_p < \frac{\varepsilon}{3C_{p,D}T^{1/2}}.$$

Hence, for $m \geq m_2$

$$ {}_mE_{1,3}(\tau) \leq \frac{\varepsilon}{3} + C_{p,D}C_G\Big(\int_0^\tau {}_mE_1(s)^2\,\mathrm{d}s\Big)^{1/2}. \tag{4.6.9} $$

Altogether, we deduce from (4.6.7), (4.6.8), and (4.6.9) that

$$ {}_mE_1(\tau) \leq \varepsilon + \beta_{p,D,T}\Big(\int_0^\tau {}_mE_1(s)^2\,\mathrm{d}s\Big)^{1/2} $$

with $\beta_{p,D,T} := C_F T^{1/2} + C_{p,D}C_G$. An application of the continuous version of Gronwall's inequality from Lemma 2.67 yields

$$ {}_mE_1(\tau) \leq \varepsilon \cdot (1 + \beta_{p,D,T}^2\tau)^{1/2}\exp\Big(\frac{1}{2} + \frac{1}{2}\beta_{p,D,T}^2\tau\Big). $$

The required statement follows by setting $\tau = T$. □

The regularised discrete approximation obtained by applying the contractive time discretisation scheme as in (4.6.1) to the regularised evolution equation is given by the variation-of-constants formula

$$ {}_mU^j := R_k^j({}_mu_0) + k\sum_{i=0}^{j-1} R_k^{j-i}({}_mF(t_i, {}_mU^i)) + \sum_{i=0}^{j-1} R_k^{j-i}({}_mG(t_i, {}_mU^i)\Delta W_{i+1}) \tag{4.6.10} $$

for $j = 0, \ldots, N_k$.

The next error investigated is the numerical discretisation error for the regularised problem, where we make use of the fact that ${}_mF$ and ${}_mG$ are mapping into spaces with additional regularity. By means of the regularisation, we are now in the position to apply the results from Section 4.4.

Corollary 4.61 (Convergence of the regularised discretisation). *Suppose that Assumption 4.56 holds for some $p \in [2, \infty)$ and $\alpha \in (0, 1]$. Further suppose that $-A$ generates a C_0-contraction semigroup on both X and $\mathrm{D}(A)$, and let $u_0 \in L^p_{\mathcal{F}_0}(\Omega; X)$. Let $(R_k)_{k>0}$ be a time discretisation scheme that is contractive on both X and $\mathrm{D}(A)$. Assume R approximates S to order α on $\mathrm{D}(A)$. Let $m \in \mathbb{N}$. Denote by ${}_mU$ the mild solution of (4.6.5) with ${}_mF, {}_mG, {}_mu_0$ as defined in (4.6.4) and by $({}_mU^j)_{j=0,\ldots,N_k}$ the temporal approximations as defined in (4.6.10). Then*

$$ \lim_{k\to 0}\Big\|\max_{0\leq j\leq N_k}\|{}_mU(t_j) - {}_mU^j\|\Big\|_p = 0. \tag{4.6.11} $$

Proof. First, we note that $\mathrm{D}(A)$ is a 2-smooth Banach space like X, see Section 2.4. Global Lipschitz continuity of ${}_mF$ and ${}_mG$ on X, $\mathrm{D}(A)$-invariance, and linear growth on $\mathrm{D}(A)$ as stated in Assumption 4.21(a), (c), and (d) were already proven in Proposition 4.58. Hölder continuity of ${}_mF$ and ${}_mG$ as in Assumption 4.21(b) follows immediately from the respective Assumption 4.56(c) on F and G. Lastly, ${}_mu_0 = mR(m, -A)u_0 \in L^p_{\mathcal{F}_0}(\Omega; \mathrm{D}(A))$ due to the regularising property of the resolvent.

Hence, Theorem 4.24 is applicable to ${}_mU$ and its discretisation $({}_mU^j)_{j=0,\ldots,N_k}$ with $Y = \mathrm{D}(A)$, nonlinearity ${}_mF$, noise ${}_mG$, and initial values ${}_mu_0$. It yields the desired convergence, even with a rate depending on the Hölder continuity in time of F and G. □

Note that the convergence of the regularised discretisation is not uniform in the regularisation parameter $m \in \mathbb{N}$. This leads to additional challenges in proving the main result, which we now move on to.

Proof of Theorem 4.57. Let $\lfloor t \rfloor := t_j$ for $t \in [t_j, t_{j+1})$, $0 \le j \le N_k - 1$, and $\lfloor T \rfloor := T$. Then

$$\Big\| \sup_{t\in[0,T]} \|U(t) - \tilde{U}(t)\| \Big\|_p \le \Big\| \sup_{t\in[0,T]} \|U(t) - U(\lfloor t \rfloor)\| \Big\|_p + \Big\| \max_{0\le j\le N_k} \|U(t_j) - U^j\| \Big\|_p.$$

Theorem 4.12 implies pathwise continuity of the mild solution U. Clearly, U is also uniformly continuous on $[0,T]$, which together with dominated convergence in $L^p(\Omega)$ yields convergence of the first term to 0 as $k \to 0$. It remains to show convergence of the discretisation error. To this end, let $N \in \{0,\dots,N_k\}$ and fix some $m \in \mathbb{N}$ to be determined later. We further decompose the discretisation error at the first $N+1$ grid points into the three parts

$$\begin{aligned} E(N) &:= \Big\| \max_{0\le j\le N} \|U(t_j) - U^j\| \Big\|_p \\ &\le \Big\| \max_{0\le j\le N} \|U(t_j) - {}_mU(t_j)\| \Big\|_p + \Big\| \max_{0\le j\le N} \|{}_mU(t_j) - {}_mU^j\| \Big\|_p + \Big\| \max_{0\le j\le N} \|{}_mU^j - U^j\| \Big\|_p \\ &=: {}_mE_1(N) + {}_mE_2(N) + {}_mE_3(N). \end{aligned} \tag{4.6.12}$$

Note that ${}_mE_1(N) \to 0$ uniformly in N as $m \to \infty$ as a consequence of Lemma 4.60. Moreover, ${}_mE_2(N_k) \to 0$ as $k \to 0$ follows from Corollary 4.61. It remains to bound the remaining term ${}_mE_3(N)$. This will be done in terms of ${}_mE_1(N_k)$ and ${}_mE_2(N_k)$, which converge in the desired manner, and $E(i)$, $0 \le i \le N-1$, which is dealt with via a Gronwall argument, as illustrated in Step 1. The bound of ${}_mE_3(N)$ from the claim below is obtained in Step 2 of the proof.

Claim. Let $\varepsilon > 0$. We claim that there exist $m_0 = m_0(\varepsilon) \in \mathbb{N}$ and a constant $C = C(p,D,F,G,T) \ge 0$ such that for $m \ge m_0$,

$${}_mE_3(N) \le \frac{\varepsilon}{2} + C[{}_mE_1(N_k) + {}_mE_2(N_k)] + C\Big(k \sum_{i=0}^{N-1} E(i)^2\Big)^{1/2}. \tag{4.6.13}$$

Step 1. Suppose that the claim holds true. We show that it implies the convergence of $E(N_k)$ to 0 as $k \to 0$. Indeed, noting that ${}_mE_i(N) \le {}_mE_i(N_k)$ for $i = 1,2$, we conclude from (4.6.12) and (4.6.13) that

$$E(N) \le \frac{\varepsilon}{2} + (C+1)[{}_mE_1(N_k) + {}_mE_2(N_k)] + C\Big(k \sum_{i=0}^{N-1} E(i)^2\Big)^{1/2}.$$

An application of Gronwall's inequality from Lemma 2.68 results in

$$E(N) \le \Big(\frac{\varepsilon}{2} + (C+1)[{}_mE_1(N_k) + {}_mE_2(N_k)]\Big)(1 + C^2 t_N)^{1/2} \exp\Big(\frac{1 + C^2 t_N}{2}\Big)$$

for all $m \geq m_0$. By Lemma 4.60, there exists $m_1 \in \mathbb{N}$ such that

$$ {}_mE_1(N_k) \leq \Big\| \sup_{t\in[0,T]} \|U(t) - {}_mU(t)\| \Big\|_p \leq \frac{\varepsilon}{2(C+1)} $$

for all $m \geq m_1$ and $N_k \in \mathbb{N}$. Hence, for $m \geq \max\{m_0, m_1\}$ fixed, we obtain

$$ E(N) \leq \big(\varepsilon + (C+1)\,{}_mE_2(N_k)\big)(1 + C^2T)^{1/2} \exp\left(\frac{1+C^2T}{2}\right). $$

Corollary 4.61 gives ${}_mE_2(N_k) \to 0$ as $N_k \to \infty$ or, equivalently, $k \to 0$. Since $\varepsilon > 0$ was chosen arbitrarily, we conclude $E(N_k) \to 0$ as $k \to 0$, which proves the desired convergence statement.

Step 2. We proceed to prove the claim (4.6.13). The error can be divided into an initial value part, a nonlinear part, and a noise part according to

$$ \begin{aligned} {}_mE_3(N) &= \Big\| \max_{0\leq j\leq N} \|U^j - {}_mU^j\| \Big\|_p \\ &\leq \Big\| \max_{0\leq j\leq N} \|R_k^j[u_0 - {}_mu_0]\| \Big\|_p + \Big\| \max_{0\leq j\leq N} \Big\| k \sum_{i=0}^{j-1} R_k^{j-i}[F(t_i, U^i) - {}_mF(t_i, {}_mU^i)] \Big\| \Big\|_p \\ &\quad + \Big\| \max_{0\leq j\leq N} \Big\| \sum_{i=0}^{j-1} R_k^{j-i}[G(t_i, U^i) - {}_mG(t_i, {}_mU^i)]\Delta W_{i+1} \Big\| \Big\|_p \\ &=: E_{3,1} + E_{3,2} + E_{3,3}, \end{aligned} \tag{4.6.14} $$

where the dependence $E_{3,\ell} = E_{3,\ell}(N, m, k)$ for $\ell = 1, 2, 3$ is omitted in the notation. We bound all three terms individually. First, we observe that by contractivity of R_k and pointwise convergence of $mR(m, -A) \to I$, there exists $m_2 \in \mathbb{N}$ such that for all $m \geq m_2$

$$ E_{3,1} \leq \|u_0 - {}_mu_0\|_{L^p(\Omega;X)} = \|(I - mR(m, -A))u_0\|_{L^p(\Omega;X)} < \frac{\varepsilon}{6}. \tag{4.6.15} $$

Second, we consider the nonlinear part of the error. For $0 \leq j \leq N$ and $0 \leq i \leq j-1$, we estimate

$$ \begin{aligned} &\|R_k^{j-i}[F(t_i, U^i) - {}_mF(t_i, {}_mU^i)]\| \\ &\qquad \leq \|F(t_i, U^i) - F(t_i, U(t_i))\| + \|F(t_i, U(t_i)) - {}_mF(t_i, U(t_i))\| \\ &\qquad\quad + \|{}_mF(t_i, U(t_i)) - {}_mF(t_i, {}_mU(t_i))\| + \|{}_mF(t_i, {}_mU(t_i)) - {}_mF(t_i, {}_mU^i)\| \\ &\qquad \leq C_F\|U(t_i) - U^i\| + \|F(t_i, U(t_i)) - {}_mF(t_i, U(t_i))\| \\ &\qquad\quad + C_F\|U(t_i) - {}_mU(t_i)\| + C_F\|{}_mU(t_i) - {}_mU^i\|, \end{aligned} \tag{4.6.16} $$

where in the last step we have used uniform Lipschitz continuity of ${}_mF$ and F. The motivation for splitting the error in this manner is that the difference between F and its regularised counterpart ${}_mF$ is then evaluated in the values $U(t_i)$ of the mild solution at the time grid points. Since the mild solution has continuous paths, this enables us to apply Lemma 4.59, as seen in the proof of Lemma 4.60. This yields uniform convergence, in particular uniformly in the number of time steps. Summing over i, multiplying by k, taking the maximum over all j and taking norms in $L^p(\Omega)$, we conclude from Minkowski's inequality in $L^p(\Omega)$ that

$$\begin{aligned}
E_{3,2} &= \Big\| \max_{0\le j\le N} \Big\| k \sum_{i=0}^{j-1} R_k^{j-i}[F(t_i, U^i) - {}_mF(t_i, {}_mU^i)] \Big\| \Big\|_p \\
&\le C_F \Big\| k \sum_{i=0}^{N-1} \|U(t_i) - U^i\| \Big\|_p + C_F \Big\| k \sum_{i=0}^{N-1} \|U(t_i) - {}_mU(t_i)\| \Big\|_p \\
&\quad + C_F \Big\| k \sum_{i=0}^{N-1} \|{}_mU(t_i) - {}_mU^i\| \Big\|_p + \Big\| k \sum_{i=0}^{N-1} \|F(t_i, U(t_i)) - {}_mF(t_i, U(t_i))\| \Big\|_p \\
&\le C_F k \sum_{i=0}^{N-1} E(i) + C_F T[{}_mE_1(N_k) + {}_mE_2(N_k)] + T \Big\| \sup_{t\in[0,T]} \|F(t, U(t)) - {}_mF(t, U(t))\| \Big\|_p .
\end{aligned}$$

Analogously to the proof of Lemma 4.60, we derive from Lemma 4.59 that there exists $m_3 \in \mathbb{N}$ such that for all $m \ge m_3$,

$$\Big\| \sup_{t\in[0,T]} \|F(t, U(t)) - {}_mF(t, U(t))\| \Big\|_p < \frac{\varepsilon}{6T}.$$

Consequently, from the Cauchy–Schwarz inequality, it follows that for all $m \ge m_3$,

$$E_{3,2} \le C_F \sqrt{T} \Big(k \sum_{i=0}^{N-1} E(i)^2 \Big)^{1/2} + C_F T[{}_mE_1(N_k) + {}_mE_2(N_k)] + \frac{\varepsilon}{6}. \tag{4.6.17}$$

To bound the last term $E_{3,3}$ in (4.6.14), we make use of the martingale argument from Lemma 4.18, which we apply with $Q_i := G(t_i, U^i) - {}_mG(t_i, {}_mU^i)$. This yields

$$E_{3,3} \le B_{p,D} \Big(k \sum_{i=0}^{N-1} \big\| \|[G(t_i, U^i) - {}_mG(t_i, {}_mU^i)]\|_{\gamma(H,X)} \big\|_p^2 \Big)^{1/2}$$

with $B_{p,D} = 10D\sqrt{p}(10p^2(p-1)^{-1} + 1)$. As for the nonlinear terms in (4.6.16), we split the term $\|G(t_i, U^i) - {}_mG(t_i, {}_mU^i)\|_{\gamma(H,X)}$ in such a way that the difference of G and ${}_mG$ is evaluated at $U(t_i)$ rather than the discrete approximations U^i or ${}_mU^i$. After an application of the triangle inequality in $\ell^2(\{0, \dots, N-1\}; L^p(\Omega; \gamma(H, X)))$, this results in

$$\begin{aligned}
E_{3,3} &\le B_{p,D} \Big[C_G \Big(k \sum_{i=0}^{N-1} \|U(t_i) - U^i\|_{L^p(\Omega;X)}^2 \Big)^{1/2} \\
&\quad + C_G \Big(k \sum_{i=0}^{N-1} \|U(t_i) - {}_mU(t_i)\|_{L^p(\Omega;X)}^2 \Big)^{1/2} + C_G \Big(k \sum_{i=0}^{N-1} \|{}_mU(t_i) - {}_mU^i\|_{L^p(\Omega;X)}^2 \Big)^{1/2} \\
&\quad + \Big(k \sum_{i=0}^{N-1} \big\| \|G(t_i, U(t_i)) - {}_mG(t_i, U(t_i))\|_{\gamma(H,X)} \big\|_p^2 \Big)^{1/2} \Big] \\
&\le B_{p,D} \Big[C_G \Big(k \sum_{i=0}^{N-1} E(i)^2 \Big)^{1/2} + C_G \sqrt{T}[{}_mE_1(N_k) + {}_mE_2(N_k)]
\end{aligned}$$

$$+ \sqrt{T} \Big\| \sup_{t\in[0,T]} \|G(t,U(t)) - {}_mG(t,U(t))\|_{\gamma(H,X)} \Big\|_p \Big].$$

We recall from the proof of Lemma 4.60 that the left ideal property of $\gamma(H,X)$ allows us to apply Lemma 4.59 on $Z = \gamma(H,X)$. We infer that there is $m_4 \in \mathbb{N}$ such that for all $m \geq m_4$, the bound

$$\Big\| \sup_{t\in[0,T]} \|G(t,U(t)) - {}_mG(t,U(t))\| \Big\|_p < \frac{\varepsilon}{6B_{p,D}\sqrt{T}}$$

holds. Thus, for $m \geq m_4$,

$$E_{3,3} \leq B_{p,D}C_G\Big(k\sum_{i=0}^{N-1} E(i)^2\Big)^{1/2} + B_{p,D}C_G\sqrt{T}[{}_mE_1(N_k) + {}_mE_2(N_k)] + \frac{\varepsilon}{6}. \tag{4.6.18}$$

Inserting the bounds (4.6.15), (4.6.17), and (4.6.18) into (4.6.14) proves the claim (4.6.13) with $C := \max\{\sqrt{T}, 1\} \cdot (C_F\sqrt{T} + B_{p,D}C_G)$ and $m_0 := \max\{m_2, m_3, m_4\}$. □

4.6.2 Application to the irregular Schrödinger equation

To illustrate the convergence results from Subsection 4.6.1, we reconsider the stochastic Schrödinger equation with a potential and linear multiplicative noise

$$\begin{cases} \mathrm{d}u + \mathrm{i}\Delta u\,\mathrm{d}t = -\mathrm{i}Vu\,\mathrm{d}t - \mathrm{i}u\,\mathrm{d}W \quad \text{on } [0,T], \\ \qquad u(0) = u_0 \end{cases} \tag{4.6.19}$$

and its nonlinear variant with $\phi\colon \mathbb{C} \to \mathbb{C}$ and $\psi\colon \mathbb{C} \to \mathbb{C}$,

$$\begin{cases} \mathrm{d}u + \mathrm{i}\Delta u\,\mathrm{d}t = -\mathrm{i}(Vu + \phi(u))\,\mathrm{d}t - \mathrm{i}\psi(u)\,\mathrm{d}W \quad \text{on } [0,T], \\ \qquad u(0) = u_0, \end{cases} \tag{4.6.20}$$

which we already encountered in Subsection 4.4.4. In Theorems 4.37 and 4.39, we have obtained pathwise uniform convergence rates for both variants under regularity conditions on the potential, the covariance of the noise, and the initial data. We aim at relaxing the regularity conditions imposed on the potential V as well as the covariance operator Q and allowing for rough initial data u_0 while maintaining pathwise uniform convergence.

Let $\sigma \geq 0$ and write $L^2 = L^2(\mathbb{R}^d)$, $L^\infty = L^\infty(\mathbb{R}^d)$, and $H^\sigma = H^\sigma(\mathbb{R}^d)$. We keep the notation from Subsection 4.4.4 and will also be using the Bessel potential spaces $H^{\sigma,q} = H^{\sigma,q}(\mathbb{R}^d)$, which coincide with the classical Sobolev spaces $W^{\sigma,q}(\mathbb{R}^d)$ if $\sigma \in \mathbb{N}$ and $q \in (1,\infty)$. For details on these spaces, we refer the interested reader to Section 2.3 and [15, 126].

We are concerned with covariance operators $Q \in \mathcal{L}(L^2)$ of trace class. More precisely, we assume that for an orthonormal basis $(h_n)_{n\in\mathbb{N}}$ of L^2, the covariance operator admits the decomposition

$$Q = \sum_{n\in\mathbb{N}} \lambda_n (h_n \otimes h_n) \text{ with } \sum_{n\in\mathbb{N}} \lambda_n = C_\lambda < \infty, \quad \sup_{n\in\mathbb{N}} \big(\|h_n\|_{L^\infty} + \|h_n\|_{H^{\sigma,d/\sigma}}\big) < \infty \tag{4.6.21}$$

for some constant $C_\lambda \geq 0$, where we interpret $H^{\sigma,d/\sigma}$ as L^∞ for $\sigma = 0$. The conditions (4.6.21) are equivalent to $Q^{1/2} \in \mathcal{L}(L^2, L^\infty \cap H^{\sigma,d/\sigma})$. While the last condition constitutes an additional regularity assumption on Q, a wide range of operators is still covered due to the Sobolev index of $H^{\sigma,d/\sigma}$ being 0. In particular, $H^{\sigma,d/\sigma}$-regularity does not result in any Hölder regularity, not even continuity.

The following theorem on the linear Schrödinger equation covers general $\sigma \geq 0$. More general nonlinearities can be treated when restricting considerations to only the case $\sigma = 0$ (see Theorem 4.63).

Theorem 4.62. *Let $d \in \mathbb{N}$, $\sigma \in [0, \frac{d}{2})$, and $p \in [2, \infty)$. Assume that $V \in L^\infty \cap H^{\sigma,\frac{d}{\sigma}}$ and $u_0 \in L^p_{\mathcal{F}_0}(\Omega; H^\sigma)$. Suppose that the covariance operator $Q \in \mathcal{L}(L^2)$ satisfies (4.6.21). Let $(R_k)_{k>0}$ be a time discretisation scheme that is contractive on H^σ and $H^{\sigma+2}$. Assume R approximates S to some order $\alpha \in (0,1]$ on $H^{\sigma+2}$. Denote by U the mild solution of the linear stochastic Schrödinger equation with multiplicative noise (4.6.19) and by $(U^j)_{j=0,\ldots,N_k}$ the temporal approximations as defined in (4.6.1). Define the piecewise constant extension $\tilde{U}\colon [0,T] \to L^p(\Omega; X)$ by $\tilde{U}(t) := U^j$ for $t \in [t_j, t_{j+1})$, $0 \leq j \leq N_k - 1$, and $\tilde{U}(T) := U^{N_k}$. Then*

$$\lim_{k\to 0} \left\| \sup_{t\in[0,T]} \|U(t) - \tilde{U}(t)\|_{H^\sigma} \right\|_p = 0. \tag{4.6.22}$$

Proof. Let $X := H^\sigma$. The semigroup generated by $-A = -\mathrm{i}\Delta$ is contractive on both X and $\mathrm{D}(A) = H^{\sigma+2}$ [2, Lem. 2.1]. We claim that Assumption 4.56 is satisfied for $F\colon \Omega \times [0,T] \times X \to X$, $F(\omega,t,u) := -\mathrm{i}Vu$ and $G\colon \Omega \times [0,T] \times X \to \gamma(H,X)$, $G(\omega,t,u) := -\mathrm{i}M_uQ^{1/2}$ with M_u denoting the multiplication operator associated with u. At first, we show Lipschitz continuity of F on X. Let $q_1 = \frac{2d}{d-2\sigma}$ and $q_2 = \frac{d}{\sigma}$. Then $\frac{1}{q_1} + \frac{1}{q_2} = \frac{1}{2}$ and $q_1 < \infty$ because $d > 2\sigma$. By classical Sobolev and Bessel potential space embeddings, H^σ embeds into L^{q_1} (see Theorem 2.31). Hence, an application of the product estimate from Proposition 2.32 yields

$$\begin{aligned} \|F(u)\|_{H^\sigma} = \|V \cdot u\|_{H^\sigma} &\lesssim \|V\|_{H^{\sigma,q_2}} \|u\|_{L^{q_1}} + \|V\|_{L^\infty} \|u\|_{H^\sigma} \\ &\lesssim (\|V\|_{H^{\sigma,d/\sigma}} + \|V\|_{L^\infty}) \|u\|_{H^\sigma} \end{aligned} \tag{4.6.23}$$

for $u \in H^\sigma$, i.e., linear growth of F. Lipschitz continuity of F follows from the above considerations, noting that F is linear.

Next, Lipschitz continuity and linear growth of G are to be derived from the trace class condition of Q. Set $H := L^2$ and write $Q = \sum_{n\in\mathbb{N}} \lambda_n (h_n \otimes h_n)$ with $(h_n)_{n\in\mathbb{N}}$ and λ_n as in (4.6.21). Since H^σ is a Hilbert space, it suffices to consider Hilbert–Schmidt norms. Using the product estimate from (4.6.23) in the inequality marked with $(*)$, we calculate

$$\begin{aligned} \|G(u)\|^2_{\mathcal{L}_2(L^2,H^\sigma)} = \|M_u Q^{1/2}\|^2_{\mathcal{L}_2(L^2,H^\sigma)} &= \sum_{n\in\mathbb{N}} \|uQ^{1/2}h_n\|^2_{H^\sigma} = \sum_{n\in\mathbb{N}} \lambda_n \|u \cdot h_n\|^2_{H^\sigma} \\ &\overset{(*)}{\lesssim} \sum_{n\in\mathbb{N}} \lambda_n (\|h_n\|_{H^{\sigma,d/\sigma}} + \|h_n\|_{L^\infty})^2 \|u\|^2_{H^\sigma} \\ &\leq C_\lambda \sup_{n\in\mathbb{N}} \big(\|h_n\|_{H^{\sigma,d/\sigma}} + \|h_n\|_{L^\infty}\big)^2 \|u\|^2_{H^\sigma}. \end{aligned}$$

Linearity of G yields Lipschitz continuity of G. In conclusion, Assumption 4.56 is satisfied. Hence, Theorem 4.57 is applicable and yields the desired convergence statement. □

Note that the convergence can be arbitrarily slow. More precisely, in the general case, there is no $\alpha > 0$ such that

$$\Big\| \sup_{t\in[0,T]} \|U(t) - \tilde{U}(t)\|_{H^\sigma} \Big\|_p \lesssim k^\alpha.$$

Theorem 4.24 is not applicable in the setting of Theorem 4.62. Clearly, rough initial data prohibits the application of quantified results. However, even for smooth initial data, a convergence rate is out of reach due to the lack of regularity of the potential V and the covariance Q. Since the embedding $H^\sigma \hookrightarrow L^{\frac{2d}{d-2\sigma}}$ and the product estimate of Proposition 2.32 are sharp, there is no $\tilde{\sigma} > \sigma$ such that F or G are mappings of linear growth on $H^{\tilde{\sigma}}$. Consequently, the smoother space Y required for a convergence rate in Theorem 4.24 cannot be found in the setting of this subsection.

To cover nonlinearities as in (4.6.20), we restrict our considerations to $\sigma = 0$ for the same reasons as in Subsection 4.4.4.

Theorem 4.63. *Let $d \in \mathbb{N}$ and $p \in [2,\infty)$. Assume that $V \in L^\infty$ and $u_0 \in L^p_{\mathcal{F}_0}(\Omega; L^2)$. Suppose that the covariance operator $Q \in \mathcal{L}(L^2)$ satisfies (4.6.21). Let $(R_k)_{k>0}$ be a time discretisation scheme that is contractive on L^2 and H^2. Assume R approximates S to some order $\alpha \in (0,1]$ on H^2. Let $\phi, \psi\colon \mathbb{C} \to \mathbb{C}$ be Lipschitz continuous and such that $\phi(0) = \psi(0) = 0$. Denote by U the mild solution of the nonlinear stochastic Schrödinger equation with multiplicative noise (4.6.20) and by $(U^j)_{j=0,\dots,N_k}$ the temporal approximations as defined in (4.6.1). Define the piecewise constant extension $\tilde{U}\colon [0,T] \to L^p(\Omega; X)$ by $\tilde{U}(t) := U^j$ for $t \in [t_j, t_{j+1})$, $0 \le j \le N_k - 1$, and $\tilde{U}(T) := U^{N_k}$. Then*

$$\lim_{k\to 0} \Big\| \sup_{t\in[0,T]} \|U(t) - \tilde{U}(t)\|_{L^2} \Big\|_p = 0. \tag{4.6.24}$$

Naturally, the result extends to non-vanishing σ in specific cases where Lipschitz continuity of F and G on H^σ is known.

Proof. We show that Theorem 4.57 is applicable with $F\colon \Omega\times[0,T]\times X \to X$, $F(\omega,t,u) := -\mathrm{i}(Vu + \phi(u))$ and $G\colon \Omega \times [0,T] \times X \to \gamma(H,X)$, $G(\omega,t,u) := -\mathrm{i}M_{\psi(u)}Q^{1/2}$ on $X = L^2$. Analogously to the proof of Theorem 4.62, we obtain the bound

$$\begin{aligned}\|G(u) - G(w)\|_{\mathcal{L}_2(L^2,L^2)} &\lesssim \sqrt{2C_\lambda}\Big(\sup_{n\in\mathbb{N}} \|h_n\|_{L^\infty}\Big) \|\psi(u) - \psi(w)\|_{L^2} \\ &\le \sqrt{2C_\lambda}C_\psi \Big(\sup_{n\in\mathbb{N}} \|h_n\|_{L^\infty}\Big)\|u - w\|_{L^2}\end{aligned}$$

for $u, w \in L^2$, from which we can deduce Lipschitz continuity of G. Linear growth of G follows from $G(0) = 0$. In the same way, one can see that $F(u) = -\mathrm{i}(Vu+\phi(u))$ is Lipschitz and of linear growth on L^2. The statement directly follows from Theorem 4.57. □

Bibliography

[1] A. Agresti and M. C. Veraar. Stochastic maximal $L^p(L^q)$-regularity for second order systems with periodic boundary conditions. *To appear in Ann. Inst. Henri Poincaré Probab. Stat., arXiv preprint arXiv:2106.01274*, 2021.

[2] R. Anton and D. Cohen. Exponential integrators for stochastic Schrödinger equations driven by Itô noise. *J. Comput. Math.*, 36(2):276–309, 2018.

[3] R. Anton, D. Cohen, S. Larsson, and X. Wang. Full discretization of semilinear stochastic wave equations driven by multiplicative noise. *SIAM J. Numer. Anal.*, 54(2):1093–1119, 2016.

[4] R. Anton, D. Cohen, and L. Quer-Sardanyons. A fully discrete approximation of the one-dimensional stochastic heat equation. *IMA J. Numer. Anal.*, 40(1):247–284, 2020.

[5] W. Arendt, C. J. K. Batty, M. Hieber, and F. Neubrander. *Vector-valued Laplace transforms and Cauchy problems*, volume 96 of *Monographs in Mathematics*. Birkhäuser/Springer Basel AG, Basel, second edition, 2011.

[6] W. Arendt, I. Chalendar, and R. Eymard. Galerkin approximation of linear problems in Banach and Hilbert spaces. *IMA J. Numer. Anal.*, 42(1):165–198, 2022.

[7] P. Auscher and P. Tchamitchian. Square roots of elliptic second order divergence operators on strongly Lipschitz domains: L^2 theory. *Journal d'Analyse Mathématique*, 90(1):1–12, 2003.

[8] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Numer. Anal.*, 45(3):1005–1034, 2007.

[9] I. Babuška, R. Tempone, and G. E. Zouraris. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal.*, 42(2):800–825, 2004.

[10] M. Bachmayr, A. Cohen, D. Dũng, and C. Schwab. Fully discrete approximation of parametric and stochastic elliptic PDEs. *SIAM J. Numer. Anal.*, 55(5):2151–2186, 2017.

[11] L. Banjai, G. J. Lord, and J. Molla. Strong convergence of a Verlet integrator for the semilinear stochastic wave equation. *SIAM J. Numer. Anal.*, 59(4):1976–2003, 2021.

[12] A. Barth and A. Lang. L^p and almost sure convergence of a Milstein scheme for stochastic partial differential equations. *Stochastic Process. Appl.*, 123(5):1563–1587, 2013.

[13] S. Becker and A. Jentzen. Strong convergence rates for nonlinearity-truncated Euler-type approximations of stochastic Ginzburg-Landau equations. *Stochastic Process. Appl.*, 129(1):28–69, 2019.

[14] A. Berg, D. Cohen, and G. Dujardin. Approximated exponential integrators for the stochastic Manakov equation. *J. Comput. Dyn.*, 10(2):323–344, 2023.

[15] J. Bergh and J. Löfström. *Interpolation spaces. An introduction.* Grundlehren der Mathematischen Wissenschaften, No. 223. Springer-Verlag, Berlin-New York, 1976.

[16] H. Bessaih, E. Hausenblas, T. A. Randrianasolo, and P. A. Razafimandimby. Numerical approximation of stochastic evolution equations: Convergence in scale of Hilbert spaces. *J. Comput. Appl. Math.*, 343:250–274, 2018.

[17] G. Bourdaud. An introduction to composition operators in Sobolev spaces. *Eurasian Math. J.*, 14(1):39–54, 2023.

[18] G. Bourdaud and W. Sickel. Composition operators on function spaces with fractional order of smoothness. In *Harmonic analysis and nonlinear partial differential equations*, volume B26 of *RIMS Kôkyûroku Bessatsu*, pages 93–132. Res. Inst. Math. Sci. (RIMS), Kyoto, 2011.

[19] D. Braess. *Finite Elemente : Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie.* Springer-Verlag, Berlin, fifth edition, 2013.

[20] C.-E. Bréhier and D. Cohen. Analysis of a splitting scheme for a class of nonlinear stochastic Schrödinger equations. *Appl. Numer. Math.*, 186:57–83, 2023.

[21] P. Brenner and V. Thomée. On rational approximations of semigroups. *SIAM J. Numer. Anal.*, 16(4):683–694, 1979.

[22] I. G. Bubnov. Report on the works of Professor Timoshenko which were awarded the Zhuranskyi Prize. Symposium of the Institute of Communication Engineers, 81, All Union Special Planning Office (SPB) (in Russian), 1913.

[23] M. Cai, D. Cohen, and X. Wang. Strong convergence rates for a full discretization of stochastic wave equation with nonlinear damping. *arXiv preprint arXiv:2307.01975*, 2023.

[24] E. Carlen and P. Krée. L^p estimates on iterated stochastic integrals. *Ann. Probab.*, 19(1):354–368, 1991.

[25] R. Chill and A. Elst. Weak and strong approximation of semigroups on Hilbert spaces. *Integral Equations and Operator Theory*, 90(1):1–22, 2018.

[26] A. Chkifa, A. Cohen, and C. Schwab. Breaking the curse of dimensionality in sparse polynomial approximation of parametric PDEs. *J. Math. Pures Appl. (9)*, 103(2):400–428, 2015.

[27] A. Cohen, R. DeVore, and C. Schwab. Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE's. *Anal. Appl. (Singap.)*, 9(1):11–47, 2011.

[28] D. Cohen, J. Cui, J. Hong, and L. Sun. Exponential integrators for stochastic Maxwell's equations driven by Itô noise. *J. Comput. Phys.*, 410:109382, 21, 2020.

[29] D. Cohen and A. Lang. Numerical approximation and simulation of the stochastic wave equation on the sphere. *Calcolo*, 59(3):Paper No. 32, 32, 2022.

[30] D. Cohen, S. Larsson, and M. Sigg. A trigonometric method for the linear stochastic wave equation. *SIAM J. Numer. Anal.*, 51(1):204–222, 2013.

[31] D. Cohen and L. Quer-Sardanyons. A fully discrete approximation of the one-dimensional stochastic wave equation. *IMA J. Numer. Anal.*, 36(1):400–420, 2016.

[32] S. Cox, M. Hutzenthaler, A. Jentzen, J. M. A. M. van Neerven, and T. Welti. Convergence in Hölder norms with applications to Monte Carlo methods in infinite dimensions. *IMA J. Numer. Anal.*, 41(1):493–548, 2021.

[33] S. Cox, A. Jentzen, and F. Lindner. Weak convergence rates for temporal numerical approximations of stochastic wave equations with multiplicative noise. *arXiv preprint arXiv:1901.05535*, 2019.

[34] S. Cox and J. M. A. M. van Neerven. Pathwise Hölder convergence of the implicit-linear Euler scheme for semi-linear SPDEs with multiplicative noise. *Numer. Math.*, 125(2):259–345, 2013.

[35] M. Crouzeix, S. Larsson, S. Piskarëv, and V. Thomée. The stability of rational approximations of analytic semigroups. *BIT*, 33(1):74–84, 1993.

[36] J. Cui. Explicit numerical methods for high dimensional stochastic nonlinear Schrödinger equation: Divergence, regularity and convergence. *arXiv preprint arXiv:2112.10177*, 2021.

[37] J. Cui and J. Hong. Analysis of a splitting scheme for damped stochastic nonlinear Schrödinger equation with multiplicative noise. *SIAM J. Numer. Anal.*, 56(4):2045–2069, 2018.

[38] J. Cui, J. Hong, Z. Liu, and W. Zhou. Strong convergence rate of splitting schemes for stochastic nonlinear Schrödinger equations. *J. Differential Equations*, 266(9):5625–5663, 2019.

[39] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.

[40] B. Dahlberg. A note on Sobolev spaces. In *Proc. Sympos. Pure Math*, volume 35, pages 183–185, 1979.

[41] G. G. Dahlquist. A special stability problem for linear multistep methods. *Nordisk Tidskr. Informationsbehandling (BIT)*, 3:27–43, 1963.

[42] L. Diening, M. Hofmanová, and J. Wichmann. An averaged space-time discretization of the stochastic p-Laplace system. *Numer. Math.*, 153(2-3):557–609, 2023.

[43] B. Dörich and M. Hochbruck. Exponential integrators for quasilinear wave-type equations. *SIAM J. Numer. Anal.*, 60(3):1472–1493, 2022.

[44] M. Eigel, C. J. Gittelson, C. Schwab, and E. Zander. A convergent adaptive stochastic Galerkin finite element method with quasi-optimal spatial meshes. *ESAIM Math. Model. Numer. Anal.*, 49(5):1367–1398, 2015.

[45] K.-J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000.

[46] G. Fendler. Dilations of one parameter semigroups of positive contractions on L^p spaces. *Canadian Journal of Mathematics*, 49(4):736–748, 1997.

[47] P. Frauenfelder, C. Schwab, and R. A. Todor. Finite elements for elliptic problems with stochastic coefficients. *Comput. Methods Appl. Mech. Engrg.*, 194(2-5):205–228, 2005.

[48] A. M. Fröhlich and L. Weis. H^∞ calculus and dilations. *Bulletin de la Société Mathématique de France*, 134(4):487–508, 2006.

[49] H. Fujita, N. Saito, and T. Suzuki. *Operator theory and numerical methods*, volume 30 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 2001.

[50] D. Funaro. *Polynomial approximation of differential equations*, volume 8 of *Lecture Notes in Physics*. Springer-Verlag, Berlin, 1992.

[51] B. G. Galerkin. Sterzhni i plastiny. ryady v nekotorykh voprosakh uprugogo ravnovesiya sterzhnei i plastin (rods and plates series occuring in some problems of elastic equilibrium of rods and plates) (in russian). *Vestnik Inzhenerov i Tekhnikov, Petrograd*, 19:897–908, 1915.

[52] S. Geiss and M. Scheutzow. Sharpness of Lenglart's domination inequality and a sharp monotone version. *Electron. Commun. Probab.*, 26:8, 2021. Id/No 44.

[53] R. G. Ghanem and P. D. Spanos. *Stochastic finite elements: a spectral approach.* Springer-Verlag, New York, 1991.

[54] A. Gomilko, S. Kosowicz, and Y. Tomilov. A general approach to approximation theory of operator semigroups. *J. Math. Pures Appl. (9)*, 127:216–267, 2019.

[55] A. Gomilko and Y. Tomilov. On convergence rates in approximation theory for operator semigroups. *J. Funct. Anal.*, 266(5):3040–3082, 2014.

[56] D. Gottlieb and D. Xiu. Galerkin method for wave equations with uncertain coefficients. *Commun. Comput. Phys.*, 3(2):505–518, 2008.

[57] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 69 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.

[58] T. H. Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Annals of Mathematics*, pages 292–296, 1919.

[59] I. Gyöngy and A. Millet. Rate of convergence of space time approximations for stochastic evolution equations. *Potential Anal.*, 30(1):29–64, 2009.

[60] M. Haase. *The functional calculus for sectorial operators*, volume 169 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 2006.

[61] K. W. Hajduk and J. C. Robinson. Explicit characterisation of the fractional power spaces of the dirichlet laplacian and stokes operators, 2021. arXiv-preprint 2108.03139.

[62] P. Harms and M. S. Müller. Weak convergence rates for stochastic evolution equations and applications to nonlinear stochastic wave, HJMM, stochastic Schrödinger and linearized stochastic Korteweg–de Vries equations. *Z. Angew. Math. Phys.*, 70(1):Paper No. 16, 28, 2019.

[63] E. Hausenblas and J. Seidler. Stochastic convolutions driven by martingales: Maximal inequalities and exponential integrability. *Stoch. Anal. Appl.*, 26(1):98–119, 2008.

[64] M. Hegland. Variable Hilbert scales and their interpolation inequalities with applications to Tikhonov regularization. *Appl. Anal.*, 59(1-4):207–223, 1995.

[65] R. Hersh and T. Kato. High-accuracy stable difference schemes for well-posed initial value problems. *SIAM J. Numer. Anal.*, 16(4):670–682, 1979.

[66] E. Hille and R. S. Phillips. *Functional analysis and semi-groups*, volume Vol. 31 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1957. rev. ed.

[67] M. Hochbruck and T. Pažur. Error analysis of implicit Euler methods for quasilinear hyperbolic evolution equations. *Numer. Math.*, 135(2):547–569, 2017.

[68] M. Hochbruck, T. Pažur, and R. Schnaubelt. Error analysis of implicit Runge-Kutta methods for quasilinear hyperbolic evolution equations. *Numer. Math.*, 138(3):557–579, 2018.

[69] J. M. Holte. Discrete Gronwall lemma and applications. In *MAA-NCS meeting at the University of North Dakota*, volume 24, pages 1–7, 2009.

[70] J. Hong, B. Hou, and L. Sun. Energy-preserving fully-discrete schemes for nonlinear stochastic wave equations with multiplicative noise. *J. Comput. Phys.*, 451:Paper No. 110829, 20, 2022.

[71] T. Hytönen, J. M. A. M. van Neerven, M. C. Veraar, and L. Weis. *Analysis in Banach spaces. Vol. I. Martingales and Littlewood-Paley theory*, volume 63 of *Ergebnisse der Mathematik und ihrer Grenzgebiete.* Springer, Cham, 2016.

[72] T. Hytönen, J. M. A. M. van Neerven, M. C. Veraar, and L. Weis. *Analysis in Banach Spaces. Volume II. Probabilistic Methods and Operator Theory*, volume 67 of *Ergebnisse der Mathematik und ihrer Grenzgebiete.* Springer, Cham, 2017.

[73] K. Itô. Stochastic integration. In *Vector and operator valued measures and applications (Proc. Sympos., Alta, Utah, 1972)*, pages 141–148. Academic Press, New York-London, 1973.

[74] K. Ito and F. Kappel. The Trotter-Kato theorem and approximation of PDEs. *Math. Comp.*, 67(221):21–44, 1998.

[75] L. Jacobe de Naurois, A. Jentzen, and T. Welti. Weak convergence rates for spatial spectral Galerkin approximations of semilinear stochastic wave equations with multiplicative noise. *Appl. Math. Optim.*, 84(suppl. 2):S1187–S1217, 2021.

[76] A. Jentzen and P. E. Kloeden. The numerical approximation of stochastic partial differential equations. *Milan J. Math.*, 77:205–244, 2009.

[77] A. Jentzen and P. E. Kloeden. *Taylor approximations for stochastic partial differential equations*, volume 83 of *CBMS-NSF Regional Conference Series in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.

[78] A. Jentzen and M. Röckner. A Milstein scheme for SPDEs. *Found. Comput. Math.*, 15(2):313–362, 2015.

[79] M. Kac and A. J. F. Siegert. An explicit representation of a stationary Gaussian process. *Ann. Math. Statistics*, 18:438–442, 1947.

[80] O. Kallenberg and R. Sztencel. Some dimension-free features of vector-valued martingales. *Probab. Theory Related Fields*, 88(2):215–247, 1991.

[81] M. Kamrani and D. Blömker. Pathwise convergence of a numerical method for stochastic partial differential equations with correlated noise and local Lipschitz condition. *J. Comput. Appl. Math.*, 323:123–135, 2017.

[82] K. Karhunen. Über lineare Methoden in der Wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae Ser. A. I. Math.-Phys.*, 1947(37):79, 1947.

[83] T. Kato. A generalization of the Heinz inequality. *Proc. Japan Acad.*, 37:305–308, 1961.

[84] T. Kato. Quasi-linear equations of evolution, with applications to partial differential equations. In *Spectral theory and differential equations (Proc. Sympos.)*, Lecture Notes in Math., Vol. 448, pages 25–70. Springer, Berlin, 1975.

[85] K. Klioba and C. Seifert. Approximation of random evolution equations. *arXiv preprint arXiv:2404.07660*, 2024.

[86] K. Klioba and M. Veraar. Pathwise uniform convergence of time discretization schemes for SPDEs. *IMA J. Numer. Anal.*, page drae055, 2024.

[87] K. Klioba and M. Veraar. Temporal approximation of stochastic evolution equations with irregular nonlinearities. *J. Evol. Equ.*, 24(2):Paper No. 43, 2024.

[88] H. Komatsu. Fractional powers of operators. *Pacific J. Math.*, 19:285–346, 1966.

[89] B. Kovács and C. Lubich. Stability and convergence of time discretizations of quasilinear evolution equations of Kato type. *Numer. Math.*, 138(2):365–388, 2018.

[90] M. Kovács. On the convergence of rational approximations of semigroups on intermediate spaces. *Math. Comp.*, 76(257):273–286, 2007.

[91] M. Kovács, A. Lang, and A. Petersson. Weak convergence of fully discrete finite element approximations of semilinear hyperbolic SPDE with additive noise. *ESAIM Math. Model. Numer. Anal.*, 54(6):2199–2227, 2020.

[92] M. Kovács, S. Larsson, and F. Lindgren. Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise. *BIT*, 52(1):85–108, 2012.

[93] M. Kovács, S. Larsson, and F. Lindgren. Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise II. Fully discrete schemes. *BIT*, 53(2):497–525, 2013.

[94] M. Kovács, S. Larsson, and F. Lindgren. On the backward Euler approximation of the stochastic Allen-Cahn equation. *J. Appl. Probab.*, 52(2):323–338, 2015.

[95] M. Kovács, S. Larsson, and F. Lindgren. On the discretisation in time of the stochastic Allen-Cahn equation. *Math. Nachr.*, 291(5-6):966–995, 2018.

[96] M. Kovács, S. Larsson, and F. Saedpanah. Finite element approximation of the linear stochastic wave equation with additive noise. *SIAM J. Numer. Anal.*, 48(2):408–427, 2010.

[97] R. Kruse. Optimal error estimates of Galerkin finite element methods for stochastic partial differential equations with multiplicative noise. *IMA J. Numer. Anal.*, 34(1):217–251, 2014.

[98] N. Krylov. An analytic approach to SPDEs. *Stochastic partial differential equations: six perspectives*, 64, 1999.

[99] S. Larsson, V. Thomée, and L. B. Wahlbin. Finite-element methods for a strongly damped wave equation. *IMA J. Numer. Anal.*, 11(1):115–142, 1991.

[100] K.-N. Le and J. Wichmann. A class of space-time discretizations for the stochastic p-Stokes system. *Stochastic Process. Appl.*, 177:Paper No. 104443, 36, 2024.

[101] O. P. Le Maître and O. M. Knio. *Spectral methods for uncertainty quantification.* Scientific Computation. Springer, New York, 2010.

[102] P. Lesky. Die Charakterisierung der klassischen orthogonalen Polynome durch Sturm-Liouvillesche Differentialgleichungen. *Archive for Rational Mechanics and Analysis*, 10(1):341–351, Jan 1962.

[103] M. Loève. Fonctions aléatoires de second ordre. *Revue Sci.*, 84:195–206, 1946.

[104] G. J. Lord, C. E. Powell, and T. Shardlow. *An introduction to computational stochastic PDEs.* Cambridge Texts in Applied Mathematics. Cambridge University Press, New York, 2014.

[105] A. Lunardi. *Analytic semigroups and optimal regularity in parabolic problems.* Progress in Nonlinear Differential Equations and their Applications, 16. Birkhäuser Verlag, Basel, 1995.

[106] A. Lunardi. On the Ornstein-Uhlenbeck operator in L^2 spaces with respect to invariant measures. *Trans. Amer. Math. Soc.*, 349(1):155–169, 1997.

[107] P. Monk. *Finite element methods for Maxwell's equations.* Oxford University Press, 2003.

[108] T. Müller-Gronbach. The optimal uniform approximation of systems of stochastic differential equations. *Ann. Appl. Probab.*, 12(2):664–690, 2002.

[109] M. Ondreját and M. Veraar. On temporal regularity of stochastic convolutions in 2-smooth Banach spaces. *Ann. Inst. Henri Poincaré Probab. Stat.*, 56(3):1792–1808, 2020.

[110] E.-M. Ouhabaz. *Analysis of Heat Equations on Domains.* London Mathematical Society Monographs. Princeton University Press, Princeton, N.J., 2005.

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

[112] B. J. Pettis. On integration in vector spaces. *Transactions of the American Mathematical Society*, 44(2):277, 1938.

[113] I. Pinelis. Optimum bounds for the distributions of martingales in Banach spaces. *The Annals of Probability*, 22(4):1679, 1994.

[114] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, Berlin, third edition, 1999.

[115] J. J. Schäffer. On unitary dilations of contractions. *Proc. Amer. Math. Soc.*, 6:322, 1955.

[116] R. Schnaubelt. Error analysis of the implicit Euler scheme for the Maxwell–Kerr system. Technical Report 3, Karlsruher Institut für Technologie (KIT), 2023.

[117] B. Simon. Schrödinger semigroups. *Bull. Amer. Math. Soc. (N.S.)*, 7(3):447–526, 1982.

[118] O. G. Smolyanov, H. v. Weizsäcker, and O. Wittich. Chernoff's theorem and discrete time approximations of Brownian motion on manifolds. *Potential analysis*, 26(1), 2006.

[119] T. J. Sullivan. *Introduction to uncertainty quantification*, volume 63 of *Texts in Applied Mathematics*. Springer, Cham, 2015.

[120] B. Sz.-Nagy, C. Foias, H. Bercovici, and L. Kérchy. *Harmonic analysis of operators on Hilbert space*. Universitext. Springer, New York, second enlarged edition, 2010.

[121] M. Talagrand. *Upper and lower bounds for stochastic processes—decomposition theorems*, volume 60 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics*. Springer, Cham, second edition, 2021.

[122] M. E. Taylor. *Tools for PDE : pseudodifferential operators, paradifferential operators, and layer potentials*. Mathematical surveys and monographs. American Mathematical Society, Providence, R.I., 2007.

[123] S. Thomaschewski. *Form Methods for autonomous and non-autonomous Cauchy Problems*. PhD thesis, Universität Ulm, https://oparu.uni-ulm.de/xmlui/handle/123456789/68, 2003. Dissertation.

[124] V. Thomée. *Galerkin finite element methods for parabolic problems*, volume 25 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006.

[125] R. A. Todor and C. Schwab. Convergence rates for sparse chaos approximations of elliptic problems with stochastic coefficients. *IMA J. Numer. Anal.*, 27(2):232–261, 2007.

[126] H. Triebel. *Interpolation theory, function spaces, differential operators*. Johann Ambrosius Barth, Heidelberg, second edition, 1995.

[127] J. M. A. M. van Neerven and M. C. Veraar. Maximal inequalities for stochastic convolutions in 2-smooth Banach spaces and applications to stochastic evolution equations. *Philos. Trans. Roy. Soc. A*, 378(2185):20190622, 21, 2020.

[128] J. M. A. M. van Neerven and M. C. Veraar. Maximal inequalities for stochastic convolutions and pathwise uniform convergence of time discretisation schemes. *Stoch. Partial Differ. Equ. Anal. Comput.*, 10(2):516–581, 2022.

[129] J. M. A. M. van Neerven, M. C. Veraar, and L. Weis. Stochastic integration in UMD Banach spaces. *Ann. Probab.*, 35(4):1438–1478, 2007.

[130] J. M. A. M. van Neerven, M. C. Veraar, and L. Weis. Stochastic integration in Banach spaces—a survey. In *Stochastic analysis: a series of lectures*, volume 68 of *Progr. Probab.*, pages 297–332. Birkhäuser/Springer, Basel, 2015.

[131] C. von Hallern and A. Rössler. An analysis of the Milstein scheme for SPDEs without a commutative noise condition. In *Monte Carlo and quasi-Monte Carlo methods*, volume 324 of *Springer Proc. Math. Stat.*, pages 503–521. Springer, Cham, 2020.

[132] X. Wang. An exponential integrator scheme for time discretization of nonlinear stochastic wave equation. *J. Sci. Comput.*, 64(1):234–263, 2015.

[133] X. Wang, S. Gan, and J. Tang. Higher order strong approximations of semilinear stochastic wave equation with additive space-time white noise. *SIAM J. Sci. Comput.*, 36(6):A2611–A2632, 2014.

[134] J. Wichmann. Temporal regularity of symmetric stochastic p-Stokes systems. *arXiv preprint arXiv:2209.02796*, 2022.

[135] J. Wichmann. On temporal regularity of strong solutions to stochastic p-Laplace systems. *SIAM J. Math. Anal.*, 55(4):3713–3730, 2023.

[136] N. Wiener. The Homogeneous Chaos. *Amer. J. Math.*, 60(4):897–936, 1938.

[137] D. Williams. *Probability with martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

[138] D. Xiu. *Numerical methods for stochastic computations.* Princeton University Press, Princeton, NJ, 2010. A spectral method approach.

[139] D. Xiu and J. S. Hesthaven. High-order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.*, 27(3):1118–1139, 2005.

[140] D. Xiu and G. E. Karniadakis. Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. *Comput. Methods Appl. Mech. Engrg.*, 191(43):4927–4948, 2002.

[141] D. Xiu and G. E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.

[142] D. Xiu and G. E. Karniadakis. A new stochastic approach to transient heat conduction modeling with uncertainty. *Int. J. Heat Mass Transf.*, 46(24):4681–4693, 2003.

[143] D. Xiu and J. Shen. Efficient stochastic Galerkin methods for random diffusion equations. *J. Comput. Phys.*, 228(2):266–281, 2009.

[144] J. Xu and L. Zikatanov. Some observations on Babuška and Brezzi theories. *Numer. Math.*, 94(1):195–202, 2003.

[145] I. Yaroslavtsev. Why stochastics? And why in Banach spaces?—Epidemic example. *Nieuw Arch. Wiskd. (5)*, 21(4):245–250, 2020.

[146] P. Zorin-Kranich. A_1 Fefferman-Stein inequality for maximal functions of martingales in uniformly smooth spaces. *Electron. J. Probab.*, 26:Paper No. 120, 18, 2021.