TECHNISCHE UNIVERSITÄT WIEN
VIENNA UNIVERSITY OF TECHNOLOGY

Diese Dissertation haben begutachtet:

DISSERTATION

Numerical Methods For Stochastic Partial Differential Equations

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors der technischen Wissenschaften unter der Leitung von

Univ.Prof. Jens Markus Melenk, PhD
E101 Institut für Analysis und Scientific Computing
Technische Universität Wien

Prof. Dr. Josef Teichmann
Departement Mathematik
Eidgenössische Technische Hochschule Zürich

eingereicht an der Technischen Universität Wien
Fakultät für Mathematik und Geoinformation

von

Dipl.-Ing. Philipp Dörsek, MSc
0326026
Döblinger Hauptstraße 7A/64
A-1190 Wien

Wien, am 26. August 2011

Kurzfassung

Wir betrachten stochastische partielle Differentialgleichungen sowohl von einer analytischen als auch von einer numerischen Perspektive. Wir führen gewichtete Räume von Funktionen auf den Zustandsräumen unendlichdimensionaler stochastischer Gleichungen ein. Mittels einer Erweiterung der klassischen Feller-Eigenschaft positiver Halbgruppen auf dem Raum der stetigen Funktionen auf einem lokalkompakten Raum, die im Unendlichen abklingen, leiten wir hinreichende Bedingungen für die starke Stetigkeit der von einem Markovprozess in endlicher oder unendlicher Dimension induzierten Halbgruppe her. Unter Verwendung der starken Stetigkeit und neuer invarianter Teilräume erhalten wir Taylor-Entwicklungen der Markov-Halbgruppen der Lösungsprozesse von stochastischen partiellen Differentialgleichungen.

Diese Resultate werden auf die numerische Analysis von Splitting- und Kubatur-Approximationen von stochastischen partiellen Differentialgleichungen vom Da Prato-Zabczyk-Typ angewendet. Wir erhalten dieselben optimalen Konvergenzraten wie im endlichdimensionalen Rahmen. Als numerisches Beispiel simulieren wir die Heath-Jarrow-Morton-Gleichung der Zinstheorie.

Abschließend leiten wir Fehlerabschätzungen für die stochastischen Navier-Stokes-Gleichungen auf dem zweidimensionalen Torus her. Die Abschätzungen sind optimal in der Zeit, aber die Konstanten hängen von der Ordnung einer Ortsdiskretisierung durch eine Spektralmethode ab. Numerische Rechnungen bestätigen die Anwendbarkeit der vorgeschlagenen Methoden.

Abstract

We consider stochastic partial differential equations, both from an analytical and a numerical point of view. We introduce weighted spaces of functions on state spaces of infinite-dimensional stochastic equations. Through an extension of the classical Feller property of positive semigroups on the space of functions decaying at infinity on a locally compact space, we derive sufficient conditions for the strong continuity of the semigroup induced by a Markov process in finite and infinite dimension. Using the strong continuity and novel invariant subspaces, we obtain Taylor expansions of Markov semigroups of solution processes of stochastic partial differential equations.

These results are applied to the numerical analysis of splitting and cubature approximations for stochastic partial differential equations of Da Prato-Zabczyk type. We recover the same optimal rates of convergence as in the finite-dimensional setting. As a numerical example, we simulate the Heath-Jarrow-Morton equation of interest rate theory.

Finally, we derive error estimates for discretisations of the stochastic Navier-Stokes equations on the two-dimensional torus. The estimates are optimal in time, but the constants depend on the order of the spatial discretisation by a spectral method. Numerical calculations confirm the applicability of the suggested schemes.

Acknowledgements

While there are many people who have helped me in many different ways to be able to write this thesis, I want to start by thanking my advisors, Professors Markus Melenk and Josef Teichmann, for their support and guidance. They allowed me to participate in their working groups and gave me in this way the chance to work on a very interesting research topic.

At Vienna University of Technology, I enjoyed very much the atmosphere of the two research groups I was part of, on the one hand the numerical analysis group, on the other the financial and actuarial mathematics group. Here I was also part of the doctoral school on Partial Differential Equations in Technical Systems of the Vienna University of Technology, and of the doctoral school on Differential Equation Models in Science and Engineering sponsored by the Austrian Science Foundation FWF. I recall fondly the very interesting summer schools at the beautiful Weissensee in Carinthia.

At ETH Zürich, I participated in the research group on mathematical finance and actuarial mathematics. I worked together very closely with Dejan Veluščček. Together, we implemented a framework for the weak approximation of stochastic differential equations, which became the founding pillar for the simulation code used to produce the numerical results for the Heath-Jarrow-Morton equation in Chapter 4 and the stochastic Navier-Stokes equations in Chapter 6.

Many of the results were greatly improved through inspiring discussions. Of course, my advisors gave many suggestions for corrections. Regarding Chapter 2, fruitful discussions with Michael Kaltenbäck and Georg Grafendorfer, who commented on early drafts, allowed me to simplify the theorems significantly.

Finally, I want to thank my girlfriend, Melina, for her understanding, and my parents for their constant support through their years. Without it, I could never have finished this work.

Contents

Introduction

The aim of this PhD thesis is to develop a novel mathematical framework for the numerical analysis of stochastic partial differential equations and to demonstrate its applicability. While there are many different ways in which randomness can enter into a mathematical model given by a partial differential equation, we consider here the case of *random forcing*. This means that there is a driving process, typically a finite or infinite dimensional Brownian motion or, more generally, a Lévy process or even a semimartingale, which puts randomness into a partial differential equation. As it is our belief that a finite dimensional driving process captures the most important phenomena, we will restrict ourselves to this case. This belief is justified by results such as [47]. There, it is proved that random input on finitely many, correctly chosen Fourier modes is sufficient to make the solution process of the Navier-Stokes equations on the two-dimensional torus ergodic.

In the application of numerical methods to stochastic differential equations, we distinguish different types of approximations. While pathwise and strong methods aim to obtain convergence in the original probability space of the problem, weak methods are built such that expected values of functionals of the solution are accurately obtained. As this is the approach that is needed in the numerical evaluation of pricing problems in mathematical finance and also in the simulation of ergodic processes such as the stochastic Navier-Stokes equations in the setting of Hairer and Mattingly, see [47], we shall focus on this type of method. Note that pathwise and strong approximations have been derived also for stochastic partial differential equations, see, e.g., [55] for an overview of recent results.

The kind of numerical schemes we shall be working with is the class of *cubature* and *splitting schemes*. Cubature on Wiener space, introduced in [65, 70], is a method of approximating the expected value of a solution of a stochastic differential equation by the solution of certain deterministic differential equations. These result from replacing Brownian motion in the original equation by certain well-chosen paths of bounded variation, making these problems well-defined in a pathwise manner. Chapter 1 contains a summary of these methods. We interpret them from a semigroup perspective, defined on the space $\mathrm{BUC}(\mathbb{R}^N)$ of bounded and uniformly continuous functions. This allows us to obtain slightly sharper error estimates than those given in [70] by the use of spaces of functions with uniformly continuous derivatives.

Splitting methods rely on the fact that autonomous stochastic differential equations are Markovian and thus define an operator semigroup on a suitable space of functions defined on the state space of the stochastic differential equation with values in $\mathbb{R}$. Using Itô's formula, it is possible to define the generator of the Markov semigroup and calculate Taylor expansions. Under the assumption that the generator can be split into several generators of Markov semigroups, each of which is easier to simulate than the original problem, we can concatenate the corresponding split Markov semigroups to obtain a simulation method for the original problem. There is a straightforward way to perform such a splitting for semigroups induced by stochastic differential equations, and its use in applications in computational finance was pioneered in [80]. It has since become an important tool in the approximation of expected values of functions of stochastic processes. The reason for this is the simplicity of the approach: by an adequate splitting, we are able to reduce a complicated stochastic differential problem to several simple ordinary differential equation problems, one for each split Markov semigroup. Thus, we can reuse tested solvers for these problems and obtain efficient numerical codes. The disadvantage of splitting schemes is that they have an inherent order barrier of 2, see [15], if no assumptions on commutators of the generators are made. A way around this is provided by extrapolation schemes, see [82], and also Section 4.2 of the present work.

The most fundamental problem of the above approach to this problem is, however, that the assumptions of the method are far too strict for a practical application. It is required that both the coefficients of the stochastic differential equations as well as the payoff are bounded and C^∞-bounded, an assumption that is essentially never fulfilled in real-world problems. Steps around this were taken on the one hand in [6], where the presence of the unbounded operator in a stochastic partial differential equation was dealt with by strong assumptions on the vector fields. On the other hand, in [2, 105], the restrictions for stochastic ordinary differential equations were relaxed, allowing the use of linearly bounded and Lipschitz continuous coefficients and polynomially growing payoffs.

Let us remark here shortly why the space of bounded and C^∞-bounded functions is an inadequate choice for the payoffs if the coefficients of the problem are no longer assumed to be bounded. To derive error estimates, we are forced to apply, in one way or another, the vector fields defining the problem to the payoff. Now, even if the payoff might initially have compact support, after arbitrarily short time, we expect the evolved payoff to be nonvanishing on an unbounded set unless restrictive assumptions are made on the volatilities such as in [6]. Thus, applying the vector field to the evolved payoff is expected to yield an unbounded function, showing that we leave the setting of bounded and C^∞-bounded functions, even for stochastic differential equations with smooth and Lipschitz-continuous coefficients. The situation is clearly even worse for examples such as the Heston model, where the coefficients are not even smooth any longer. We are therefore led to consider larger classes of functions for the payoffs.

Our approach to this problem is, as explained in the beginning, to provide a new mathematical framework. We take the route of strongly continuous semigroups. Strong continuity is in many senses a "via regia" towards approximation schemes via splitting schemes (e.g., Trotter-type formulae, Chernov's theorem, etc), and therefore a very desirable feature. Moreover, it allows us to derive estimates of the rate of convergence in a rather standard manner by using results such as [54, 43, 49, 44].

It is well-known that the world of stochastic Markov processes on general state spaces is tied to strongly continuous semigroups in two ways: either through the Feller property, or through invariant measures. In both cases we can construct an appropriate Banach space, $C_0(X)$ and $L^p(X, \mu)$, respectively, where the Markov semigroups act in a strongly continuous way. However, neither the existence of invariant measures nor the Feller property are generic properties of Markov processes – this holds true in particular in infinite dimension.

The situation is even worse for the Feller property, where we have a strong connection to locally compact state spaces and continuous functions vanishing at infinity. It therefore seems natural to ask for a framework extending the Feller property towards unbounded payoffs and non-locally compact spaces. Moreover, the framework should be as generic as possible to remain applicable to general stochastic partial differential equations. From the viewpoint of applications, the new concept is useful if we are able to prove rates of convergence for substantially larger classes of payoffs and equations with the presented method.

It turns out that the Feller property can be extended to a larger class of state spaces by replacing the space $C_0(X)$ of functions vanishing at infinity by a space $\mathcal{B}^\psi(X)$ of functions which have their growth controlled by ψ, and this theory is introduced in Chapter 2. Instead of the notion of a point at infinity, we assume that the sets $K_R := \{x \in X : \psi(x) \leqslant R\}$ are compact. Such an assumption is viable in infinite dimension if we endow the dual space of a normed space with the weak-$*$ topology. In particular, separable Hilbert spaces, usually used as state space in stochastic partial differential equations, are contained in this approach. It requires us, however, to work with the weak-$*$ topology. As continuous dependence on initial data for stochastic partial differential equations is usually only obtained for the norm topology, we relate usual spaces of strongly continuous functions to the newly defined spaces.

Chapter 3 applies the setting of weighted spaces to Markov semigroups. Sufficient conditions for strong continuity are provided. Furthermore, under the assumption that the stochastic process is the solution of a stochastic partial differential equation, Taylor expansions of the Markov semigroup are derived.

Chapters 4 and 5 are devoted to the derivation of rates of convergence for splitting and cubature approximations to stochastic partial differential equations. Optimal rates are obtained for sufficiently smooth functions.

Finally, Chapter 6 deals with the problem that initiated this research, the numerical approximation of the stochastic Navier-Stokes equations on the two-dimensional torus. While we are unable to reproduce the results obtained in Chapters 4 and 5 in this case, we are still able to construct splitting and cubature approximations and prove their convergence under a restriction on the time step size by the use of a spectral Galerkin approximation.

The articles [31, 30, 32] have resulted from the research performed for this thesis.

Chapter 1

Cubature On Wiener Space: A Semigroup Perspective

Introduced by Kusuoka, Lyons and Victoir [64, 70], the method of cubature on Wiener space quickly became an important numerical tool for applications and a topic of major research. Its fundamental idea is to use a combination of stochastic Taylor expansion and an innovative replacement of iterated Stratonovich integrals to construct weak approximation schemes for stochastic differential equations, reducing them to ordinary differential equation problems. In particular, it is in principle possible to find cubature paths of arbitrarily high order, a consequence of Tchakaloff's theorem (see [5] and [70, Theorem A.1]), even though this general existence result is nonconstructive and suboptimal in the number of paths needed to obtain a given order of convergence. Explicit paths up to order 11 for a single Brownian motion have been constructed, see [45].

In this chapter, we illustrate the fundamentals of the numerical methods that are at the basis of the approach to stochastic partial differential equations used in this work. Section 1.1 provides a short overview of the basics of the method of cubature on Wiener space. A convergence analysis from a semigroup perspective is given. Under the assumption that the coefficients of the stochastic differential equation are bounded and C^{∞}-bounded, we consider the space $\mathrm{BUC}(\mathbb{R}^N)$ of bounded and uniformly continuous functions to be the correct setting for cubature methods. While rates of convergence can be expected only for sufficiently smooth functions, we prove that strong convergence is retained on the entire space $\mathrm{BUC}(\mathbb{R}^N)$ (but see [64, 70, 66] for the use of smoothing effects to obtain optimal rates of convergence fo nonsmooth functions, and Section 5.2.3 for an extension of these results to unbounded payoffs).

In Section 1.2, we focus on splitting methods. A splitting-up approach to stochastic ordinary and partial differential equations based on the Lie-Trotter theorem was used by many authors, e.g., [8, 9, 38, 7, 10, 91, 67, 104, 53, 40, 42, 41]. This, however, can only yield weak order one methods, making high accuracy unattainable, see [87]. In [83, 86], a weak second-order method based on a splitting of the drift is proposed, but this method requires the explicit use of derivatives of the volatilities and is thus difficult to implement for general classes of equations. Ninomiya and Victoir [80] introduced a weak second-order method, which is a variant of the well-known Strang splitting. Their approach has immediate advantages in the simplicity of its implementation: we only have to solve for one single vector field at any given step in the algorithm. Hence, such a splitting is very attractive for use in the simulation of stochastic differential equations: well-tested, robust and efficient solvers for the corresponding deterministic problems can be used.

1.1 Cubature on Wiener space

In the following, we shall use standard notions from stochastic analysis freely. Please refer to Appendix B for an introduction to the fundamentals, and for further references.

For m, $n \in \mathbb{N}$, let $C_b^\infty(\mathbb{R}^m; \mathbb{R}^n)$ denote the space of functions $f\colon \mathbb{R}^m \to \mathbb{R}^n$ that are bounded, infinitely often differentiable, and have all partial derivatives bounded. Such functions are also called bounded and C^∞-bounded, the latter alone only signifying that all partial derivatives are bounded, but not necessarily the function itself.

Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geqslant 0})$ denote a filtered probability space satisfying the usual conditions, and suppose that $(B_t^j)_{j=1,\ldots,d}$ is a Brownian motion defined on it. The method of cubature on Wiener space introduced by Kusuoka, Lyons and Victoir in [65, 70] is a weak approximation scheme for a (for simplicity autonomous) stochastic differential equation of the form

$$\text{(1.1)} \qquad \mathrm{d}x(t, x_0) = \mu(x(t, x_0))\mathrm{d}t + \sum_{j=1}^{d} \sigma_j(x(t, x_0))\mathrm{d}B_t^j, \quad x(0, x_0) = x_0,$$

with state space $\mathbb{R}^N$, where μ, $\sigma_j \in C_b^\infty(\mathbb{R}^N; \mathbb{R}^N)$, or its equivalent Stratonovich formulation

$$\text{(1.2)} \qquad \mathrm{d}x(t, x_0) = \mu_0(x(t, x_0))\mathrm{d}t + \sum_{j=1}^{d} \sigma_j(x(t, x_0)) \circ \mathrm{d}B_t^j, \quad x(0, x_0) = x_0,$$

where $\mu_0 := \mu - \frac{1}{2} D\sigma_j \cdot \sigma_j$ is the Stratonovich corrected drift. This means that we define a family $(Q_{(t)})_{t\geqslant 0}$ of operators such that $Q_{(t/n)}^n$ converges, in some sense specified below, to P_t for $n \to \infty$. Here, $P_t f(x_0) := \mathbb{E}[f(x(t, x_0))]$ denotes the Markov transition operator defined by the process $(x(t, x_0))_{t \geqslant 0}$, and $f : \mathbb{R}^N \to \mathbb{R}$ is measurable.

To define spaces of functions where we can expect convergence of discretisation schemes, let us proceed as follows. $\mathrm{C}_b(\mathbb{R}^N)$ denotes the space of bounded, continuous functions, endowed with the supremum norm,

$$\|f\|_{\mathrm{C}_b(\mathbb{R}^N)} := \sup_{x \in \mathbb{R}^N} |f(x)|. \tag{1.3}$$

$\mathrm{BUC}(\mathbb{R}^N) \subset \mathrm{C}_b(\mathbb{R}^N)$ is its subspace constisting of bounded and uniformly continuous functions. When deriving Taylor expansions for Markov semigroups, we need to use smooth functions. We define

$$\begin{aligned} \mathrm{C}_b^k(\mathbb{R}^N) := \Big\{ f \in \mathrm{C}^k(\mathbb{R}^N) \colon & \frac{\partial^j}{\partial x_{i_1} \ldots \partial x_{i_j}} f \in \mathrm{C}_b(\mathbb{R}^N) \\ & \text{for all multiindices } (i_1, \ldots, i_j) \in \{1, \ldots, N\}^j, \\ & j = 0, \ldots, k \Big\}, \end{aligned} \tag{1.4}$$

endowed with the norm

$$\|f\|_{\mathrm{C}_b^k(\mathbb{R}^N)} := \sum_{j=0}^{k} \sum_{(i_1, \ldots, i_j) \in \{1, \ldots, N\}^j} \Big\| \frac{\partial^j}{\partial x_{i_1} \ldots \partial x_{i_j}} f \Big\|_{\mathrm{C}_b(\mathbb{R}^N)}. \tag{1.5}$$

The closed subspace $\mathrm{BUC}^k(\mathbb{R}^N) \subset \mathrm{C}_b^k(\mathbb{R}^N)$ is given by

$$\begin{aligned} \mathrm{BUC}^k(\mathbb{R}^N) := \Big\{ f \in \mathrm{C}^k(\mathbb{R}^N) \colon & \frac{\partial^j}{\partial x_{i_1} \ldots \partial x_{i_j}} f \in \mathrm{BUC}(\mathbb{R}^N) \\ & \text{for all multiindices } (i_1, \ldots, i_j) \in \{1, \ldots, N\}^j, \\ & j = 0, \ldots, k \Big\}. \end{aligned} \tag{1.6}$$

Proposition 1.1. $\mathrm{C}_b^\infty(\mathbb{R}^N)$ *is dense in* $\mathrm{BUC}(\mathbb{R}^N)$.

Proof. For $f \in \mathrm{BUC}(\mathbb{R}^N)$, set $f_\varepsilon(x) := \mathbb{E}[f(x + B_\varepsilon)]$, where $(B_t)_{t \geqslant 0}$ is a standard n-dimensional Brownian motion. It is easy to see that $f_\varepsilon \in \mathrm{C}_b^\infty(\mathbb{R}^N)$ for all $\varepsilon > 0$, and by [108, p. 399, Proposition], we see that

$$\lim_{\varepsilon \to 0+} \|f - f_\varepsilon\|_{\mathrm{C}_b(\mathbb{R}^N)} = 0. \tag{1.7}$$ □

The fundamental idea is to introduce a finite set of *cubature paths* which replicate the expectation of iterated Stratonovich integrals of Brownian motion. Using the stochastic Taylor expansion, this lets us replace Brownian motion in (1.1) with the cubature paths to obtain an estimate for the local error. Global error bounds are then derived from stability estimates of the exact and approximate solution operators.

1.1.1 The stochastic Taylor expansion

We shall only give a short overview of the stochastic Taylor expansion. More details can be found in [59, 70].

Let $f \in C_b^\infty(\mathbb{R}^N)$. We are interested in the behaviour of the stochastic process $(f(x(t, x_0)))_{t \geqslant 0}$. By the chain rule of Stratonovich calculus, Proposition B.16, we obtain

$$\begin{aligned} f(x(t,x_0)) - f(x_0) = &\int_0^t Df(x(s,x_0))\mu_0(x(s,x_0))\mathrm{d}s \\ &+ \sum_{j=1}^d \int_0^t Df(x(s,x_0))\sigma_j(x(s,x_0)) \circ \mathrm{d}B_s^j, \end{aligned} \tag{1.8}$$

which is the pillar of the stochastic Taylor expansion, similarly as the deterministic chain rule is the pillar of the deterministic Taylor expansion. As $\mathbb{E}[(B_s^j)^2] = s$, we expect the first term on the right hand side above to behave like t in the limit $t \to 0$, the second one like $t^{1/2}$. Thus, if we want to obtain a certain rate of convergence in t, we should expect that we need an expansion of higher order in B_s than in s. This leads to the consideration of a *weighting* of multiindices done in the following manner: Let $\alpha := (j_1, \dots, j_k)$ be a multiindex. Its degree $\deg(\alpha)$ is defined by

$$\deg(\alpha) := k + |\{\ell \colon j_\ell = 0\}|, \tag{1.9}$$

that is, every component with $j_\ell = 0$ is counted twice. The empty multiindex is denoted by $\varnothing$ and satisfies $\deg(\varnothing) = 0$. We set

$$\mathcal{A} := \{\alpha \text{ multiindex}\}, \qquad \mathcal{A}_m := \{\alpha \in \mathcal{A} \colon \deg(\alpha) \leqslant m\}, \tag{1.10}$$

$$\mathcal{A}^* := \{\alpha \in \mathcal{A} \colon \alpha \notin \{\varnothing, (0)\}\}, \quad \mathcal{A}_m^* := \mathcal{A}^* \cap \mathcal{A}_m. \tag{1.11}$$

For compactness of notation, we denote $B_s^0 := s$, and interpret $Vf(x) := Df(x) \cdot V(x)$ for a vector field $V \colon \mathbb{R}^N \to \mathbb{R}^N$. Furthermore, we set $V_j := \sigma_j$, $j = 1, \dots, d$

and $V_0 := \mu_0$, and denote the iterated Stratonovich integrals by

$$I_t^{(j_1,\dots j_k)}(g)(x_0) := \int\cdots\int_{0<t_1<\cdots<t_k<t} g(x(t_1,x_0)) \circ \mathrm{d}B_{t_1}^{j_1} \cdots \circ \mathrm{d}B_{t_k}^{j_k} \tag{1.12}$$

and $I_t^{(j_1,\dots j_k)} := I_t^{(j_1,\dots j_k)}(1)$.

Proposition 1.2. *For every $m \geqslant 0$, we have the* stochastic Taylor expansion

$$\begin{aligned} f(x(\Delta t, x_0)) - f(x_0) = &\sum_{(j_1,\dots j_k)\in\mathcal{A}_m} V_{j_1}\dots V_{j_k} f(x_0) I_{\Delta t}^{(j_1,\dots j_k)} \\ &+ R_m(\Delta t, x_0, f). \end{aligned} \tag{1.13}$$

The remainder term $R_m(\Delta t, x_0, f)$ is given by

$$R_m(\Delta t, x_0, f) = \sum_{\substack{(j_1,\dots j_k)\in\mathcal{A}_m \\ (j_0,\dots j_k)\notin\mathcal{A}_m}} I_{\Delta t}^{(j_0,\dots j_k)}(V_{j_0}\dots V_{j_k} f)(x_0), \tag{1.14}$$

and satisfies for every $T > 0$ the estimate

$$\begin{aligned} &\sup_{x_0\in\mathbb{R}^N} \sqrt{\mathbb{E}[R_m(\Delta t, x_0, f)^2]} \\ &\leqslant C(\Delta t)^{(m+1)/2} \sup_{(j_1,\dots j_k)\in\mathcal{A}_{m+2}\setminus\mathcal{A}_m} \|V_{j_1}\dots V_{j_k} f\|_{C_b(\mathbb{R}^N)} \end{aligned} \tag{1.15}$$

with a constant $C = C_T > 0$ independent of $\Delta t \in [0, T]$ and f.

Its proof is essentially a straightforward application of the Stratonovich chain rule, together with an estimate of the Stratonovich integrals by transforming to Itô form and using the Itô isometry. Note that the condition $f \in C_b^\infty(\mathbb{R}^N)$ is clearly too strong. The above error estimate shows that $f \in C_b^{m+2}(\mathbb{R}^N)$ suffices.

1.1.2 Formulation of the method

Cubature on Wiener space replaces the paths B_t^j of Brownian motion by deterministic paths of bounded variation in such a manner that the expected values $\mathbb{E}[I_{(j_1,\dots,j_k)}]$ of iterated Stratonovich integrals remain unchanged up to degree m, $\deg(j_1,\dots,j_k) \leqslant m$. These values can actually be determined explicitly by algebraic methods, see [4, Proposition 1.3]. The order m then determines the rate of convergence the algorithm exhibits for sufficiently smooth functions f.

Consider functions $\omega_i^j\colon [0,1] \to \mathbb{R}$ of bounded variation with $\omega_i^j(0) = 0$ and weights $\lambda_i \geqslant 0$, $j = 1, \ldots, d$, $\omega_i^0(t) = t$, $i = 1, \ldots, M$. We assume that for any multiindex $\alpha = (j_1, \ldots, j_k)$ with $\deg(\alpha) \leqslant m$,

$$\mathbb{E}[I_1^{(j_1,\ldots,j_k)}] = \sum_{i=1}^{M} \lambda_i \int \cdots \int_{0<s_1<\cdots<s_k<1} \mathrm{d}\omega_i^{j_1}(s_1) \ldots \mathrm{d}\omega_i^{j_k}(s_k). \tag{1.16}$$

Such a collection $(\omega_i, \lambda_i)_{i=1,\ldots,M}$ of paths and weights is called a *cubature formula of order m*. A straightforward application of the chain rule shows that the transformed paths $\omega_i^{(\Delta t),j}(s) := \sqrt{\Delta t}\omega_i^j(\frac{s}{\Delta t})$, $s \in [0, \Delta t]$, $j = 1, \ldots, d$, $\omega_i^{(\Delta t),0}(s) := s$, satisfy

$$\mathbb{E}[I_{\Delta t}^{(j_1,\ldots,j_k)}] = \sum_{i=1}^{M} \lambda_i \int \cdots \int_{0<s_1<\cdots<s_k<\Delta t} \mathrm{d}\omega_i^{(\Delta t),j_1}(s_1) \ldots \mathrm{d}\omega_i^{(\Delta t),j_k}(s_k). \tag{1.17}$$

Define the approximation operator $Q_{(\Delta t)}$ by

$$Q_{(\Delta t)} f(x_0) := \sum_{i=1}^{M} \lambda_i f(x(t, x_0; \omega_i^{(\Delta t)})), \tag{1.18}$$

where $x(s, x_0; \omega_i^{(\Delta t)})$ solves the ordinary differential equation

$$\mathrm{d}x(s, x_0; \omega_i^{(\Delta t)}) = \sum_{j=0}^{d} V_j(x(s, x_0; \omega_i^{(\Delta t)}))\mathrm{d}\omega_i^{(\Delta t)}(s). \tag{1.19}$$

Proposition 1.3. *For $f \in \mathrm{BUC}^{m+1}(\mathbb{R}^N)$,*

$$Q_{(\Delta t)} f(x_0) = \sum_{(j_1,\ldots,j_k)\in\mathcal{A}_m} V_{j_1} \ldots V_{j_k} f(x_0)\mathbb{E}[I_{\Delta t}^{(j_1,\ldots,j_k)}] + \tilde{R}_m(\Delta t, x_0, f). \tag{1.20}$$

For every $T > 0$, the residual $\tilde{R}_m(\Delta t, x_0, f) \in \mathrm{BUC}^{m+1}(\mathbb{R}^N)$ satisfies

$$|\tilde{R}_m(\Delta t, x_0, f)| \leqslant C(\Delta t)^{\frac{m+1}{2}} \sup_{\substack{(j_1,\ldots,j_k)\in\mathcal{A}_m \\ (j_0,j_1,\ldots,j_k)\notin\mathcal{A}_m}} \|V_{j_1} \ldots V_{j_k} f\|_{C_b(\mathbb{R}^N)} \tag{1.21}$$

with a constant $C = C_T > 0$ independent of $\Delta t \in [0, T]$ and f.

Proof. Setting

$$I_t^{(j_1,\dots j_k)}(\omega_i^{(\Delta t)}, g) \tag{1.22}$$
$$:= \int \cdots \int_{0<s_1<\cdots<s_k<t} g(x(s_1, x_0; \omega_i^{(\Delta t)}))d\omega_i^{(\Delta t),j}(s_1) \dots d\omega_i^{(\Delta t),j}(s_k)$$

and $I_t^{(j_1,\dots j_k)}(\omega_i^{(\Delta t)}) := I_t^{(j_1,\dots j_k)}(\omega_i^{(\Delta t)}, 1)$, the usual, deterministic Taylor expansion yields

$$f(x(t, x_0; \omega_i^{(\Delta t)})) - f(x_0)$$
$$= \sum_{(j_1,\dots,j_k)\in\mathcal{A}_m} V_{j_1} \dots V_{j_k} f(x_0) I_t^{(j_1,\dots j_k)}(\omega_i^{(\Delta t)}) + r_{m,i}(t, x_0, f), \tag{1.23}$$

where the residual $r_{m,i}$ is given explicitly by

$$r_{m,i}(t, x_0, f) = \sum_{\substack{(j_1,\dots j_k)\in\mathcal{A}_m \\ (j_0,\dots j_k)\notin\mathcal{A}_m}} I_t^{(j_0,\dots j_k)}(\omega_i^{(\Delta t)}, V_{j_0} \dots V_{j_k} f)(x_0). \tag{1.24}$$

As $f \in \mathrm{BUC}^{m+1}(\mathbb{R}^N)$, we see that $V_{j_0} \dots V_{j_k} f \in \mathrm{BUC}(\mathbb{R}^N)$. Hence,

$$|r_{m,i}(t, x_0, f)| \leqslant C t^{(m+1)/2} \sup_{\substack{(j_1,\dots,j_k)\in\mathcal{A}_m \\ (j_0,j_1\dots,j_k)\notin\mathcal{A}_m}} \|V_{j_0} \dots V_{j_k} f\|_{C_b(\mathbb{R}^N)}. \tag{1.25}$$

Multiplying with λ_i and summing up over $i = 1, \dots, M$ allows us to conclude. □

We now obtain from Proposition 1.2 and 1.3 that

$$\|P_{\Delta t} f - Q_{(\Delta t t)} f\|_{C_b(\mathbb{R}^N)}$$
$$\leqslant C(\Delta t)^{\frac{m+1}{2}} \sup_{(j_1,\dots,j_k)\in\mathcal{A}_{m+2}\setminus\mathcal{A}_m} \|V_{j_1} \dots V_{j_k} f\|_{C_b(\mathbb{R}^N)}. \tag{1.26}$$

As the coefficients are bounded and C^∞-bounded, we see that

$$\|V_{j_1} \dots V_{j_k} f\|_{C_b(\mathbb{R}^N)} \leqslant \|f\|_{C_b^k(\mathbb{R}^N)}. \tag{1.27}$$

Furthermore, for smooth f,

$$\|P_t f\|_{C_b^k(\mathbb{R}^N)} \leqslant C\|f\|_{C_b^k(\mathbb{R}^N)} \quad \text{for } t \in [0, T], \tag{1.28}$$

which altogether proves a global rate of convergence of $(m-1)/2$ for smooth functions f:

Proposition 1.4. *Given $T > 0$, there exists a constant $C = C_T > 0$ such that for all $f \in C_b^{m+2}(\mathbb{R}^N)$, $t \in [0, T]$ and $n \in \mathbb{N}$.*

$$\|P_t f - Q_{(t/n)}^n f\|_{C_b(\mathbb{R}^N)} \leqslant C n^{-(m-1)/2} \|f\|_{C_b^{m+2}(\mathbb{R}^N)}. \tag{1.29}$$

Remark 1.5. Note that the smoothness assumptions on f are slightly worse than those which are obtained in splitting schemes, compare, e.g., with Proposition 1.13. This results from the use of Stratonovich integrals in the stochastic Taylor expansion. An alternative approach to the Taylor expansion for P_t, performed below by the use of semigroup methods in spaces $\mathrm{BUC}^k(\mathbb{R}^N)$, will allow us to recover the same kind of estimates as those for splitting schemes for odd m.

1.1.3 A semigroup interpretation

We want to consider stochastic ordinary differential equations on the state space $\mathbb{R}^N$ from the perspective of strongly continuous semigroups. This interpretation makes clear which kinds of function spaces should be considered for estimation of rates of convergence. Furthermore, it will be at the basis of the analysis of the method in the more general settings of subsequent chapters.

Let $P_t f(x_0) := \mathbb{E}[f(x(t, x_0))]$ denote the Markov semigroup defined by (1.1). Then, $P_t \in L(\mathrm{BUC}(\mathbb{R}^N))$. In fact, $(P_t)_{t \geqslant 0}$ even defines a strongly continuous semigroup of contractions on $\mathrm{BUC}(\mathbb{R}^N)$. See Section A.1 for an overview of strongly continuous semigroups.

Proposition 1.6. *For every $t \in [0, \infty)$, the operator $P_t \colon \mathrm{BUC}(\mathbb{R}^N) \to \mathrm{BUC}(\mathbb{R}^N)$ is well-defined and a contraction, that is, $P_t f \in \mathrm{BUC}(\mathbb{R}^N)$ and $\|P_t f\|_{C_b(\mathbb{R}^N)} \leqslant \|f\|_{C_b(\mathbb{R}^N)}$ for all $f \in \mathrm{BUC}(\mathbb{R}^N)$. Furthermore, $\lim_{t \to 0+} \|P_t f - f\|_{C_b(\mathbb{R}^N)} = 0$.*

It follows that $(P_t)_{t \geqslant 0}$ is a strongly continuous semigroup of contractions on $\mathrm{BUC}(\mathbb{R}^N)$.

Proof. The proof is done similarly as in [108, Proposition, p. 399]. Fix $f \in \mathrm{BUC}(\mathbb{R}^N)$.

We first prove that $P_t f \in \mathrm{BUC}(\mathbb{R}^N)$. Let $\varepsilon > 0$ be given. As there exists $\delta > 0$ such that for every $x \in \mathbb{R}^N$, we have that $|f(y) - f(x)| \leqslant \varepsilon$ whenever

$|y-x| \leqslant \delta$,

$$\begin{aligned}
|P_t f(x_1) - P_t f(x_2)| &\leqslant \mathbb{E}[|f(x(t,x_1)) - f(x(t,x_2))|] \\
&= \mathbb{E}[|f(x(t,x_1)) - f(x(t,x_2))|\chi_{[|x(t,x_1)-x(t,x_2)|\leqslant\delta]}] \\
&\quad + \mathbb{E}[|f(x(t,x_1)) - f(x(t,x_2))|\chi_{[|x(t,x_1)-x(t,x_2)|>\delta]}] \\
&\leqslant \varepsilon + 2\|f\|_{C_b(\mathbb{R}^N)}\mathbb{P}([|x(t,x_1) - x(t,x_2)| > \delta]). \qquad (1.30)
\end{aligned}$$

Here, $\chi_A(x) := 1$ if $x \in A$, 0 otherwise, denotes the indicator function of the set A. By Chebyshev's inequality,

$$\mathbb{P}([|x(t,x_1) - x(t,x_2)| > \delta]) \leqslant \delta^{-2}\mathbb{E}\Big[\big|x(t,x_1) - x(t,x_2)\big|^2\Big]. \qquad (1.31)$$

Due to

$$\begin{aligned}
x(t,x_1) - x(t,x_2) &= x_1 - x_2 + \int_0^t (\mu(x(s,x_1)) - \mu(x(s,x_2)))\mathrm{d}s \\
&\quad + \sum_{j=1}^d \int_0^t (\sigma_j(x(s,x_1)) - \sigma_j(x(s,x_2)))\mathrm{d}B_s^j, \qquad (1.32)
\end{aligned}$$

we can apply the Itô isometry and the Lipschitz continuity of the coefficients to obtain

$$\begin{aligned}
\mathbb{E}\Big[\big|\int_0^t (\sigma_j(x(s,x_1)) - \sigma_j(x(s,x_2)))\mathrm{d}s\big|^2\Big] &= \int_0^t \mathbb{E}[|\sigma_j(x(s,x_1)) - \sigma_j(x(s,x_2))|^2]\mathrm{d}s \\
&\leqslant \int_0^t C\mathbb{E}[|x(s,x_1) - x(s,x_2)|^2]\mathrm{d}s, \qquad (1.33)
\end{aligned}$$

and similarly, from the Jensen inequality,

$$\begin{aligned}
&\mathbb{E}\Big[\big|\int_0^t (\mu(x(s,x_1)) - \mu(x(s,x_2)))\mathrm{d}s\big|^2\Big] \\
&\qquad \leqslant \int_0^t C\mathbb{E}[|x(s,x_1) - x(s,x_2)|^2]\mathrm{d}s. \qquad (1.34)
\end{aligned}$$

Hence, the Gronwall inequality yields

$$\mathbb{E}[|x(t,x_1) - x(t,x_2)|^2] \leqslant \exp(Ct)|x_1 - x_2|^2, \qquad (1.35)$$

and we deduce

$$\mathbb{P}([|x(t,x_1) - x(t,x_2)| > \delta]) \leqslant \delta^{-2}C|x_1 - x_2|^2. \qquad (1.36)$$

Plugging this into (1.30), we obtain $P_t f \in \mathrm{BUC}(\mathbb{R}^N)$, and the contraction property $\|P_t f\|_{C_b(\mathbb{R}^N)} \leqslant \|f\|_{C_b(\mathbb{R}^N)}$ is obvious.

We now prove that $P_t f \to f$ uniformly as $t \to 0+$. Denote the transition probability of the process $(x(t,x_0))_{t\geqslant 0}$ by $\mu_t(x_0,\cdot) := \mathbb{P}^{x(t,x_0)}$, i.e., $\mu_t(x_0,A) := \mathbb{P}[x(t,x_0) \in A]$ for all Borel sets $A \subset \mathbb{R}^N$. Choosing $\delta > 0$ as above, it follows that

$$
\begin{aligned}
|P_t f(x_0) - f(x_0)| &\leqslant \int_{\mathbb{R}^N} |f(x) - f(x_0)| \mu_t(x_0, \mathrm{d}x) \\
&= \int_{|x_0 - x| \leqslant \delta} |f(x) - f(x_0)| \mu_t(x_0, \mathrm{d}x) \\
&\quad + \int_{|x_0 - x| > \delta} |f(x) - f(x_0)| \mu_t(x_0, \mathrm{d}x) \\
&\leqslant \varepsilon + 2\|f\|_{C_b(\mathbb{R}^N)} \int_{|x - x_0| > \delta} \mu_t(x_0, \mathrm{d}x). \qquad (1.37)
\end{aligned}
$$

As

$$
\int_{|x - x_0| > \delta} \mu_t(x_0, \mathrm{d}x) = \mathbb{P}([|x(t,x_0) - x_0| > \delta]) \leqslant \delta^{-2} \mathbb{E}[(x(t,x_0) - x_0)^2], \qquad (1.38)
$$

the result follows from

$$
\begin{aligned}
&\mathbb{E}[(x(t,x_0) - x_0)^2] \\
&\quad = \mathbb{E}\Big[\Big(\int_0^t \mu(x(s,x_0))\mathrm{d}s + \sum_{j=1}^d \int_0^t \sigma_j(x(s,x_0))\mathrm{d}B_s^j\Big)^2\Big], \qquad (1.39)
\end{aligned}
$$

where we use that

$$
\mathbb{E}\Big[\Big(\int_0^t \mu(x(s,x_0))\mathrm{d}s\Big)^2\Big] \leqslant Ct^2, \qquad (1.40)
$$

$$
\mathbb{E}\Big[\Big(\int_0^t \sigma_j(x(s,x_0))\mathrm{d}B_s^j\Big)^2\Big] = \mathbb{E}\Big[\int_0^t \sigma_j(x(s,x_0))^2 \mathrm{d}s\Big] \leqslant Ct \qquad (1.41)
$$

by the Itô isometry, and

$$
\mathbb{E}\Big[\int_0^t \mu(x(s,x_0))\mathrm{d}s \int_0^t \sigma_j(x(s,x_0))\mathrm{d}B_s^j\Big] \leqslant Ct^{3/2} \qquad (1.42)
$$

by the Cauchy-Schwarz inequality. □

A similar argument as in the first part of the above proof shows that $Q_{(\Delta t)} \in L(\mathrm{BUC}(\mathbb{R}^N))$. Therefore, the following result on the strong convergence of $Q^n_{(t/n)}$ to P_t follows from Proposition 1.4 by a density argument.

Proposition 1.7. *For every $f \in \mathrm{BUC}(\mathbb{R}^N)$ and $t \geqslant 0$,*

$$\lim_{n\to\infty} \|P_t f - Q^n_{(t/n)} f\|_{\mathrm{C}_b(\mathbb{R}^N)} = 0. \tag{1.43}$$

Another important consequence of the semigroup property of $(P_t)_{t\geqslant 0}$ is the existence of the infinitesimal generator. The following result collects some of its properties.

Proposition 1.8. *Denote the infinitesimal generator of $(P_t)_{t\geqslant 0}$ by $\mathcal{G}$ with $\operatorname{dom}\mathcal{G}$. The space $\mathrm{C}^\infty_b(\mathbb{R}^N)$ is a core of $\mathcal{G}$, i.e., for all $f \in \operatorname{dom}\mathcal{G}$, there exists a sequence $f_n \in \mathrm{C}^\infty_b(\mathbb{R}^N)$ such that*

$$\lim_{n\to\infty} \|f - f_n\|_{\mathrm{C}_b(\mathbb{R}^N)} + \lim_{n\to\infty} \|\mathcal{G}f - \mathcal{G}f_n\|_{\mathrm{C}_b(\mathbb{R}^N)} = 0. \tag{1.44}$$

Furthermore, $\mathrm{BUC}^2(\mathbb{R}^N) \subset \operatorname{dom}\mathcal{G}$,

$$\mathcal{G}f = \mu_0 f + \sum_{j=1}^{d} \sigma_j^2 f, \quad \text{and} \tag{1.45}$$

$$\|\mathcal{G}f\|_{\mathrm{C}_b(\mathbb{R}^N)} \leqslant C\|f\|_{\mathrm{C}^2_b(\mathbb{R}^N)} \quad \text{for } f \in \mathrm{C}^2_b(\mathbb{R}^N). \tag{1.46}$$

Thus, $\mathcal{G}$ is a differential operator of second order.

Proof. That $\mathrm{C}^\infty_b(\mathbb{R}^N) \subset \operatorname{dom}\mathcal{G}$ and the representation (1.45) of $\mathcal{G}$ on $\mathrm{C}^\infty_b(\mathbb{R}^N)$ follow from Itô's formula. Applying that $\mathcal{G}$ is a closed operator and the right hand side of (1.45) is continuous as operator $\tilde{\mathcal{G}}\colon \mathrm{BUC}^2(\mathbb{R}^N) \to \mathrm{BUC}(\mathbb{R}^N)$ by the smoothness and boundedness of the vector fields, this formula extends to $\mathrm{BUC}^2(\mathbb{R}^N)$ in the following way: for $f \in \mathrm{BUC}^2(\mathbb{R}^N)$, choose a sequence $(f_n)_{n\in\mathbb{N}}$ in $\mathrm{C}^\infty_b(\mathbb{R}^N)$ with $\lim_{n\to\infty}\|f - f_n\|_{C^2_b(\mathbb{R}^N)} = 0$ (existence of such a sequence is proved as in Proposition 1.1). Then, $\tilde{\mathcal{G}}f_n$ converges to $\tilde{\mathcal{G}}f$ in $\mathrm{BUC}(\mathbb{R}^N)$. The closedness of $\mathcal{G}$ yields $f \in \operatorname{dom}\mathcal{G}$ and $\mathcal{G}f = \tilde{\mathcal{G}}f$, whence the claim.

As $\mathrm{C}^\infty_b(\mathbb{R}^N)$ is dense in $\mathrm{BUC}(\mathbb{R}^N)$ by Proposition 1.1 and as it is invariant with respect to the semigroup $(P_t)_{t\geqslant 0}$ by Proposition B.21, Proposition A.5(vi) proves that $\mathrm{C}^\infty_b(\mathbb{R}^N)$ is a core of $\mathcal{G}$. □

Proposition 1.8 allows us to obtain a finer smoothness assumption in the Taylor expansion than Proposition 1.2.

Corollary 1.9. *Assume that $f \in \mathrm{BUC}^{2(k+1)}(\mathbb{R}^N)$. Then,*

$$P_t f = \sum_{j=0}^{k} \frac{t^j}{j!}\mathcal{G}^j f + R_{2k+1}(t,f), \tag{1.47}$$

where the residual $R_{2k+1}(t,f) \in \mathrm{BUC}(\mathbb{R}^N)$ satisfies

$$\|R_{2k+1}(t,f)\|_{\mathrm{C}_b(\mathbb{R}^N)} \leqslant C t^{k+1} \|f\|_{\mathrm{C}_b^{2(k+1)}(\mathbb{R}^N)}. \tag{1.48}$$

Proof. Induction on Proposition 1.8 yields that $\mathcal{G}^j : \mathrm{BUC}^{2j}(\mathbb{R}^N) \to \mathrm{BUC}(\mathbb{R}^N)$ is bounded. Hence, we can apply Proposition A.5(v) to obtain the claim. □

Remark 1.10. Corollary 1.9 is stronger than Proposition 1.2, as $\mathrm{C}_b^{2k+3}(\mathbb{R}^N) \subsetneq \mathrm{BUC}^{2(k+1)}(\mathbb{R}^N)$.

Corollary 1.11. *Assume that m is odd. For $T > 0$, there exists a constant $C = C_T > 0$ such that for $t \in [0,T]$, $f \in \mathrm{BUC}^{m+1}(\mathbb{R}^N)$ and $n \in \mathbb{N}$,*

$$\|P_t f - Q_{(t/n)}^n f\|_{\mathrm{C}_b(\mathbb{R}^N)} \leqslant C n^{-(m-1)/2} \|f\|_{\mathrm{C}_b^{m+1}(\mathbb{R}^N)}. \tag{1.49}$$

Proof. Let $f \in \mathrm{C}_b^\infty(\mathbb{R}^N)$. By Proposition 1.4,

$$\|P_{\Delta t} f - Q_{(\Delta t)} f\|_{\mathrm{C}_b^{m+1}(\mathbb{R}^N)} \leqslant C(\Delta t)^{(m+1)/2} \|f\|_{\mathrm{C}_b^{m+2}(\mathbb{R}^N)}. \tag{1.50}$$

Proposition 1.3 and Corollary 1.9 show that both $P_{\Delta t} f$ and $Q_{(\Delta t)} f$ have Taylor expansions with residual of order $(m+1)/2$ controlled by $\|f\|_{\mathrm{C}_b^{m+1}(\mathbb{R}^N)}$. Hence, the error estimate extends to this space, and a telescoping argument proves the claim. □

1.2 Ninomiya-Victoir splitting

In [80], Ninomiya and Victoir introduced splitting methods and interpreted them as a variant of cubature on Wiener space. More specifically, they considered the problem of approximating the Markov semigroup corresponding to the solution of (1.2). To this end, they defined auxiliary problems

$$\frac{\mathrm{d}}{\mathrm{d}t} x^0(t,x_0) = \mu_0(x^0(t,x_0)), \qquad x^0(0,x_0) = x_0 \quad \text{and} \tag{1.51a}$$

$$\mathrm{d}x^j(t,x_0) = \sigma_j(x^j(t,x_0)) \circ \mathrm{d}B_t^j, \quad x^j(0,x_0) = x_0 \quad \text{for } j = 1,\dots,d. \tag{1.51b}$$

These problems are easy to solve numerically, as they only involve the solution of an ordinary differential equation each, where we evaluate at the random time B_t^j for $j = 1, \dots, d$. Defining the split Markov semigroups $P_t^j f(x_0) := \mathbb{E}[f(x^j(t, x_0))]$, they use the discretisation

$$Q_t f(x_0) := \frac{1}{2} P_{t/2}^0 \left(P_t^1 \dots P_t^d + P_t^d \dots P_t^1 \right) P_{t/2}^0. \tag{1.52}$$

The expected values appearing in the definition of the semigroups P_t^j, $j = 1, \dots, d$, can be discretised further by Gauss-Hermite quadrature with three points to preserve the optimal rate of convergence in a fully discrete scheme.

In light of Proposition 1.6, we see that not only $(P_t)_{t\geqslant 0}$, but also $(P_t^j)_{t\geqslant 0}$, $j = 0, \dots, d$, are strongly continuous semigroups of contractions on the space $\mathrm{BUC}(\mathbb{R}^N)$. This means that (1.52) can be seen as an *exponential splitting* for $(P_t)_{t\geqslant 0}$.

Proposition 1.8 allows us now to relate the generator $\mathcal{G}$ of $(P_t)_{t\geqslant 0}$ with the generators $\mathcal{G}_j$ of $(P_t^j)_{t\geqslant 0}$, $j = 0, \dots, d$, on a space that is a core for all the generators simultaneously.

Proposition 1.12. *We have the equality*

$$\mathcal{G} f = \sum_{j=0}^{d} \mathcal{G}_j f \quad \text{for all } f \in C_b^\infty(\mathbb{R}^N). \tag{1.53}$$

Proposition 1.8 shows that (1.53) extends to the intersection of the domains of $\mathcal{G}$ and $\mathcal{G}_j$, $j = 0, \dots, d$. The theory of [49] now applies to yield the following result.

Proposition 1.13. *Let $f \in \mathrm{BUC}(\mathbb{R}^N)$ be such that*

$$\|\mathcal{G}_{j_1}\mathcal{G}_{j_2}\mathcal{G}_{j_3} P_t f\|_{C_b(\mathbb{R}^N)} \leqslant C \quad \text{for } t \in [0, T]. \tag{1.54}$$

Then,

$$\|(Q_{(T/n)}^n - P_T) f\|_{C_b(\mathbb{R}^N)} \leqslant C_{f,T}\, n^{-2}, \tag{1.55}$$

where $C_{f,T} > 0$ is independent of $n \in \mathbb{N}$.

In particular, the Ninomiya-Victoir splitting converges of second order for $f \in \mathrm{BUC}^6(\mathbb{R}^N)$. More precisely,

$$\|(Q_{(T/n)}^n - P_T) f\|_{C_b(\mathbb{R}^N)} \leqslant C_T\, n^{-2} \|f\|_{C_b^6(\mathbb{R}^N)}. \tag{1.56}$$

Chapter 2

Weighted Spaces And A Generalised Feller Condition

The aim of this chapter is the construction of a novel theoretical framework for the application of splitting and cubature methods to stochastic partial differential equations under realistic conditions, improving significantly on the results from [6]. To this end, we define in Section 2.1 Banach spaces of real-valued functions with controlled growth on possibly infinite-dimensional state spaces for which a Riesz respresentation theorem holds true, i.e., every continuous linear functional can be represented by an integral with respect to a certain finite measure. Elements of these spaces are analysed, obtaining similar properties as for the space $C_0(X)$ of functions decaying at infinity for X a locally compact space.

These results allow us to prove in Section 2.2 that on these spaces, semigroups of positive, bounded operators $(P_t)_{t\geqslant 0}$ with $\lim_{t\to 0+} P_t f(x) = f(x)$ are in fact strongly continuous. This relaxes the assumptions of the classical Feller condition (see, e.g., [57, p. 315]) in two ways: first, the functions can be unbounded, and second, the state space can be infinite-dimensional (not locally compact).

Subsequently, we consider in Section 2.3 the case that X is the dual space of a separable Banach space. The correct topology is in this case the weak-$*$ topology, and we prove that under certain assumptions on the weight function typically satisfied in applications, the elements of weighted spaces are sequentially weak-$*$ continuous. We define a corresponding notion of differentiability, and relate it to the usual setting of strongly differentiable functions by the use of compact embeddings. As a stepping stone for Taylor expansions of Markov semigroups induced by stochastic differential equations on these spaces, we analyse directional derivatives along vector fields and prove norm bounds for such derivations.

During the final stages of this thesis, the author found out about work by

M. Röckner and Z. Sobol [94, 95, 96]. In [95], they introduce spaces C_V corresponding exactly to our spaces $\mathcal{B}^{\psi}(X)$. In particular, in [95, Theorem 5.1], they prove a Riesz representation theorem for this function space over general completely regular topological spaces, similarly as our Theorem 2.5. Their focus is different than ours: they construct solutions of martingale problems in the sense of Stroock and Varadhan, we perform an analysis of numerical methods. They do not construct a hierarchy of spaces of differentiable functions in the setting of weighted spaces, hence our results can be seen as extending [95]. Furthermore, they restrict themselves to additive noise (which can, however, be infinite-dimensional), whereas we allow nonlinear coefficients.

2.1 Riesz representation for weighted spaces

In this section we show that we can actually obtain a variant of the Riesz representation theorem even on spaces that are not locally compact.

Definition 2.1. Let X be a topological space, and $\varphi\colon X \to (0,\infty)$ be bounded from below by some $\delta > 0$. For a Banach space $(Y, \|\cdot\|_Y)$, we set

$$\mathrm{B}^{\varphi}(X;Y) := \left\{ f\colon X \to Y\colon \sup_{x\in X} \varphi(x)^{-1}\|f(x)\|_Y < \infty \right\}, \tag{2.1}$$

endowed with the φ-norm

$$\|f\|_{\varphi} := \sup_{x\in X} \varphi(x)^{-1}\|f(x)\|_Y. \tag{2.2}$$

If $Y = \mathbb{R}$, we define $\mathrm{B}^{\varphi}(X) := \mathrm{B}^{\varphi}(X;\mathbb{R})$.

It is easy to see that $\mathrm{B}^{\varphi}(X;Y)$ is a Banach space. Furthermore, it is clear that $\mathrm{C}_b(X;Y) \subset \mathrm{B}^{\varphi}(X;Y)$, where $\mathrm{C}_b(X;Y)$ denotes the space of continuous, bounded functions $f\colon X \to Y$, endowed with the norm $\|f\|_{\mathrm{C}_b(X;Y)} := \sup_{x\in X}\|f(x)\|_Y$.

Definition 2.2. Consider a completely regular Hausdorff topological space X (i.e. $\mathrm{T}_{3.5}$; see [18, Chapitre IX § 1 Définition 1]). A function $\psi\colon X \to (0,\infty)$ is called *admissible weight function* if the sets $K_R := \{x \in X\colon \psi(x) \leqslant R\}$ are compact for all $R > 0$. We call the pair (X,ψ) a *weighted space*.

Such a function ψ is lower semicontinuous and bounded from below, and any such space X is σ-compact due to $\bigcup_{n\in\mathbb{N}} K_n = X$.

Definition 2.3. Let ψ be an admissible weight function on the completely regular Hausdorff space X. We define $\mathcal{B}^\psi(X;Y)$ as the closure of $C_b(X;Y)$ in $B^\psi(X;Y)$. If $Y = \mathbb{R}$, we set $\mathcal{B}^\psi(X) := \mathcal{B}^\psi(X;\mathbb{R})$.

By definition, the normed space $\mathcal{B}^\psi(X;Y)$ is a Banach space.

Remark 2.4. Suppose X compact. Then the choice $\psi(x) := 1$ for all $x \in X$ is admissible. On general spaces ψ will necessarily grow due to the compactness of K_R, which means that $f \in \mathcal{B}^\psi(X;Y)$ typically is unbounded, but its growth is restricted by the growth of ψ. Therefore, we call elements of $\mathcal{B}^\psi(X;Y)$ *functions with growth controlled by* ψ.

Theorem 2.5 (Riesz representation for $\mathcal{B}^\psi(X)$)**.** *Given a weighted space (X,ψ), let $\ell\colon \mathcal{B}^\psi(X) \to \mathbb{R}$ be a continuous linear functional, $\ell \in \mathcal{B}^\psi(X)^*$. There exists a finite signed Radon measure μ on X such that*

$$\ell(f) = \int_X f(x)\mu(\mathrm{d}x) \quad \text{for all } f \in \mathcal{B}^\psi(X). \tag{2.3}$$

Furthermore,

$$\int_X \psi(x)|\mu|(\mathrm{d}x) = \|\ell\|_{\mathcal{B}^\psi(X)^*}, \tag{2.4}$$

where $|\mu|$ denotes the total variation measure of μ.

As every such measure defines a continuous linear functional on $\mathcal{B}^\psi(X)$, this completely characterises the dual space of $\mathcal{B}^\psi(X)$.

Proof. Clearly, $\ell|_{C_b(X)}$ is a continuous linear functional on $C_b(X)$, as

$$\|f\|_\psi \leqslant \left(\inf_{x\in X}\psi(x)\right)^{-1}\|f\|_{C_b(X)} \quad \text{for } f \in C_b(X). \tag{2.5}$$

We thus have to ensure condition *(M)* of [16, § 5 Proposition 5]. Defining $K := K_{\varepsilon^{-1}\|\ell\|_{\mathcal{B}^\psi(X)^*}}$, we see that for $g \in C_b(X)$ with $|g| \leqslant 1$ and $g|_K = 0$,

$$\|g\|_\psi = \sup_{x\in X\setminus K}\psi(x)^{-1}|g(x)| \leqslant \varepsilon\|\ell\|^{-1}_{\mathcal{B}^\psi(X)^*}\|g\|_{C_b(X)} \leqslant \varepsilon\|\ell\|^{-1}_{\mathcal{B}^\psi(X)^*}, \tag{2.6}$$

and thus $|\ell(g)| \leqslant \varepsilon$. Hence we obtain existence of a finite, uniquely determined signed Radon measure μ with $\ell(f) = \int_X f(x)\mu(\mathrm{d}x)$ for all $f \in C_b(X)$ (see also [13, Chapter 2 Theorem 2.2]).

To determine $\int_X \psi(x)|\mu|(\mathrm{d}x)$, we apply [16, § 5 Proposition 1b)]: ψ is lower semicontinuous and every positive $g \in C_b(X)$ with $g \leqslant \psi$ satisfies $\|g\|_\psi \leqslant 1$. Therefore,

$$\int_X \psi(x)|\mu|(\mathrm{d}x) = \sup_{\substack{g \in C_b(X) \\ |g| \leqslant \psi}} |\ell(g)| \leqslant \|\ell\|_{\mathcal{B}^\psi(X)^*}. \tag{2.7}$$

The density of $C_b(X)$ in $\mathcal{B}^\psi(X)$ yields

$$\begin{aligned} \|\ell\|_{\mathcal{B}^\psi(X)^*} &= \sup_{g \in C_b(X)} \|g\|_\psi^{-1} |\ell(g)| = \sup_{g \in C_b(X)} \|g\|_\psi^{-1} \Big| \int_X g(x)\mu(\mathrm{d}x) \Big| \\ &\leqslant \int_X \psi(x)|\mu|(\mathrm{d}x). \end{aligned} \tag{2.8}$$

Hence, $\int_X \psi(x)|\mu|(\mathrm{d}x) = \|\ell\|_{\mathcal{B}^\psi(X)^*}$.

For the proof of $\ell(f) = \int_X f(x)\mu(\mathrm{d}x)$ for all $f \in \mathcal{B}^\psi(X)$, note that $f \mapsto \int_X f(x)\mu(\mathrm{d}x)$ defines a continuous linear functional on $\mathcal{B}^\psi(X)$ due to the integrability of ψ with respect to $|\mu|$. As both expressions agree on a dense subset, we obtain the desired equality. □

Remark 2.6. While the result in [13, Chapter 2 Theorem 2.2] is applicable even for spaces which are not completely regular, in contrast to [16, § 5 Proposition 5], we do not see how to prove $\int_X \psi(x)|\mu|(\mathrm{d}x) < \infty$ in that situation. However, this bound is important in our further results, see the proof of Theorem 2.11.

Corollary 2.7. *Let $\ell\colon \mathcal{B}^\psi(X) \to \mathbb{R}$ be a positive linear functional, that is, $\ell(f) \geqslant 0$ whenever $f(x) \geqslant 0$ for all $x \in X$. Then, there exists a (positive) measure μ with $\ell(f) = \int_X f(x)\mu(\mathrm{d}x)$ for every $f \in \mathcal{B}^\psi(X)$.*

Proof. We only have to prove that ℓ is continuous. Assume otherwise. Then, there exists a sequence $(f_n)_{n\in\mathbb{N}}$, $f_n \in \mathcal{B}^\psi(X)$, such that $\|f_n\|_\psi = 1$, but $|\ell(f_n)| \geqslant n^3$. As $|\ell(f)| \leqslant \ell(|f|)$ for any $f \in \mathcal{B}^\psi(X)$, we can assume without loss of generality that $f_n \geqslant 0$ for all $n \in \mathbb{N}$. As $\sum_{n\in\mathbb{N}} n^{-2}\|f_n\|_\psi < \infty$, the limit $f := \sum_{n\in\mathbb{N}} n^{-2} f_n \in \mathcal{B}^\psi(X)$ is well-defined and $f \geqslant 0$. We obtain a contradiction due to $n \leqslant \ell(n^{-2} f_n) \leqslant \ell(f)$. □

The following results emphasise the analogy in structure of $\mathcal{B}^\psi(X)$ and the space of functions vanishing at infinity on a locally compact space.

Theorem 2.8. *Let $f : X \to \mathbb{R}$. Then, $f \in \mathcal{B}^{\psi}(X)$ if and only if $f|_{K_R} \in \mathrm{C}(K_R)$ for all $R > 0$ and*

$$\lim_{R\to\infty} \sup_{x\in X\setminus K_R} \psi(x)^{-1}|f(x)| = 0. \tag{2.9}$$

In particular, $f \in \mathcal{B}^{\psi}(X)$ for every $f \in \mathrm{C}(X)$ satisfying (2.9).

Proof. Assume that $f \in \mathcal{B}^{\psi}(X)$. For $g \in \mathrm{C}_b(X)$ with $\|f - g\|_{\psi} < \frac{\varepsilon}{2}$,

$$\psi(x)^{-1}|f(x)| \leqslant \frac{\varepsilon}{2} + \psi(x)^{-1}|g(x)| \quad \text{for } x \in X, \tag{2.10}$$

the last term being bounded by $\frac{\varepsilon}{2}$ for $x \in X\setminus K_R$ with $R := 2\varepsilon^{-1}\|g\|_{\mathrm{C}_b(X)}$. Thus,

$$\sup_{x\in X\setminus K_R} \psi(x)^{-1}|f(x)| \leqslant \varepsilon, \tag{2.11}$$

which proves (2.9).

Next, we prove that for any $R > 0$, $f|_{K_R}$ is continuous. With g as above,

$$\sup_{x\in K_R} |f(x) - g(x)| \leqslant R \sup_{x\in K_R} \psi(x)^{-1}|f(x) - g(x)| \leqslant \frac{\varepsilon}{2}R, \tag{2.12}$$

which means that $f|_{K_R}$ is a uniform limit of continuous functions and hence continuous.

For the other direction, set $f_n := \min(\max(f(\cdot), -n), n) = (f_n \vee n) \wedge n$. We prove first that $f_n \in \mathcal{B}^{\psi}(X)$. As $f|_{K_R} \in \mathrm{C}(K_R)$, we see that $f_n|_{K_R} \in \mathrm{C}(K_R)$. K_R is compact in a completely regular space. We can embed X into a compact space Y by [18, Chapitre IX § 1 Proposition 3, Proposition 4]. Applying the Tietze extension theorem [18, Chapitre IX § 4 Théorème 2] to the set K_R, which is also compact and therefore closed in Y, we obtain existence of $g_{n,R} \in \mathrm{C}_b(X)$ with $g_{n,R}|_{K_R} = f_n|_{K_R}$ and $\sup_{x\in X}|g_{n,R}(x)| \leqslant n$ for all $x \in X$. (2.9) yields

$$\|f_n - g_{n,R}\|_{\psi} \leqslant \sup_{x\in X\setminus K_R} \psi(x)^{-1}|f_n(x) - g_{n,R}(x)| \leqslant 2nR^{-1}, \tag{2.13}$$

hence $f_n \in \mathcal{B}^{\psi}(X)$. Next, choose $R > 0$ such that $\sup_{x\in X\setminus K_R} \psi(x)^{-1}|f(x)| < \varepsilon$. With $n > \sup_{x\in K_R}|f(x)|$, $f(x) = f_n(x)$ on K_R. Therefore,

$$\|f - f_n\|_{\psi} \leqslant \sup_{x\in X\setminus K_R} \psi(x)^{-1}|f(x) - f_n(x)| \leqslant 2\varepsilon, \tag{2.14}$$

which shows that $f \in \mathcal{B}^{\psi}(X)$. $\square$

The next result shows that not only the residual of f outside of K_R grows more slowly than ψ, but also the supremum of f on the set K_R (which is attained due to compactness of K_R and Theorem 2.8) grows more slowly than R.

Corollary 2.9. *If $f \in \mathcal{B}^{\psi}(X)$, then*

$$\lim_{R\to\infty} R^{-1} \sup_{x\in K_R} |f(x)| = 0. \tag{2.15}$$

Proof. Define the functions F and G by

$$F(R) := R^{-1} \sup_{\psi(x)=R} |f(x)| \quad \text{and} \quad G(R) := R^{-1} \sup_{x\in K_R} |f(x)|. \tag{2.16}$$

We claim $\lim_{R\to\infty} G(R) = 0$. We see that

$$\sup_{x\in X\setminus K_R} \psi(x)^{-1}|f(x)| = \sup_{R'>0} F(R') \tag{2.17}$$

and it follows by Theorem 2.8 that $\lim_{R\to\infty} F(R) = 0$. Now,

$$G(R) = R^{-1} \sup_{R'\leqslant R} R'F(R'). \tag{2.18}$$

Given $\varepsilon > 0$, choose R_0 large enough such that $F(R) < \frac{\varepsilon}{2}$ for all $R > R_0$. It follows that

$$\begin{aligned} G(R) &= \max(R^{-1} \sup_{R'\leqslant R_0} R'F(R'), R^{-1} \sup_{R_0<R'\leqslant R} R'F(R')) \\ &\leqslant R^{-1} \sup_{R'\leqslant R_0} R'F(R') + \frac{\varepsilon}{2}. \end{aligned} \tag{2.19}$$

Choosing $R > \frac{2\sup_{R'\leqslant R_0} R'F(R')}{\varepsilon}$, we obtain $G(R) < \varepsilon$. The proof is thus complete. □

Theorem 2.10. *For every $f \in \mathcal{B}^{\psi}(X)$ with $\sup_{x\in X} f(x) > 0$, there exists $z \in X$ such that*

$$\psi(x)^{-1}f(x) \leqslant \psi(z)^{-1}f(z) \quad \textit{for all } x \in X. \tag{2.20}$$

Proof. Let $\alpha := \sup_{x\in X} \psi(x)^{-1}f(x) > 0$. By Theorem 2.8, there exists $R > 0$ such that $\sup_{\psi(x)>R} \psi(x)^{-1}f(x) \leqslant \frac{\alpha}{2}$, whence

$$\alpha = \sup_{x\in K_R} \psi(x)^{-1}f(x). \tag{2.21}$$

Define $h := \psi^{-1}\max(f,0)$. Then, $\alpha = \sup_{x\in K_R} h(x)$. Furthermore, ψ^{-1} is upper semicontinuous, $\max(f,0)$ is continuous on K_R by Theorem 2.8 and both are nonnegative. Thus, h is upper semicontinuous (see [17, Chap. IV § 6 Proposition 2]) and by [17, Chapitre IV § 6 Théorème 3] attains its maximum at some point $z \in K_R$, i.e., $\alpha = \psi(z)^{-1}f(z)$ □

2.2 A generalised Feller condition

The generalised Feller property will allow us to speak about strongly continuous semigroups on spaces of functions with growth controlled by ψ. We consider a weighted supremum norm instead of the supremum norm. Hence, from the point of view of applications, we will still be able to control the pointwise error of numerical approximations.

Let $(P_t)_{t\geqslant 0}$ be a family of bounded linear operators $P_t\colon \mathcal{B}^\psi(X) \to \mathcal{B}^\psi(X)$ with the following properties:

F1. $P_0 = I$, the identity on $\mathcal{B}^\psi(X)$,

F2. $P_{t+s} = P_t P_s$ for all t, $s \geqslant 0$,

F3. for all $f \in \mathcal{B}^\psi(X)$ and $x \in X$, $\lim_{t\to 0+} P_t f(x) = f(x)$,

F4. there exist a constant $C \in \mathbb{R}$ and $\varepsilon > 0$ such that for all $t \in [0,\varepsilon]$, $\|P_t\|_{L(\mathcal{B}^\psi(X))} \leqslant C$,

F5. P_t is positive for all $t \geqslant 0$, that is, for $f \in \mathcal{B}^\psi(X)$, $f \geqslant 0$, we have $P_t f \geqslant 0$.

Alluding to [57, Chapter 17], such a family of operators will be called a *generalised Feller semigroup*. Here, for $(B, \|\cdot\|_B)$ a Banach space, $L(B)$ denotes the space of bounded linear operators $T\colon B \to B$ with the norm

$$\|T\|_{L(B)} = \sup_{\|x\|_B \leqslant 1} \|Tx\|_B. \tag{2.22}$$

We shall now prove that semigroups satisfying F1 to F4 are actually strongly continuous, a direct consequence of Lebesgue's dominated convergence theorem with respect to measures existing due to Riesz representation.

Theorem 2.11. *Let $(P_t)_{t\geqslant 0}$ satisfy F1 to F4. Then, $(P_t)_{t\geqslant 0}$ is strongly continuous on $\mathcal{B}^\psi(X)$, i.e.,*

$$\lim_{t\to 0+} \|P_t f - f\|_\psi = 0 \quad \textit{for all } f \in \mathcal{B}^\psi(X). \tag{2.23}$$

Proof. By Proposition A.3, we only have to prove that $t \mapsto \ell(P_t f)$ is right continuous at zero for every $f \in \mathcal{B}^{\psi}(X)$ and every continuous linear functional $\ell \colon \mathcal{B}^{\psi}(X) \to \mathbb{R}$. Due to Theorem 2.5, we know that there exists a signed measure ν on X such that $\ell(g) = \int_X g \mathrm{d}\nu$ for every $g \in \mathcal{B}^{\psi}(X)$. By F4, we see that for every $t \in [0, \varepsilon]$,

$$|P_t f(x)| \leqslant C\psi(x). \tag{2.24}$$

Using (2.4), the dominated convergence theorem yields

$$\lim_{t \to 0+} \int_X P_t f(x)\nu(\mathrm{d}x) = \int_X f(x)\nu(\mathrm{d}x), \tag{2.25}$$

and the claim follows. Here, the integrability of ψ with respect to the total variation measure $|\nu|$ enters in an essential way. □

Remark 2.12. As Chris Rogers remarked, state space transformation of the type $x \mapsto \phi(x) := \frac{x}{\sqrt{1+\|x\|^2}}$ transform unbounded state spaces into bounded ones. The weight function ψ is then used to rescale real valued functions $f : X \to \mathbb{R}$ via $\tilde{f} := f/\psi$ in order to investigate $\tilde{f} \circ \phi^{-1}$ on $\phi(X)$. This function will often have a continuous extension to the closure of $\phi(X)$, which – in the appropriate topology – will be often compact. This relates the generalised Feller property to the classical Feller property. Note that in our situation, however, ψ is typically not continuous for infinite dimensional X.

We can establish a positive maximum principle in case that the semigroup P_t grows like $\exp(\alpha t)$ with respect to the operator norm on $\mathcal{B}^{\psi}(X)$.

Theorem 2.13. *Let $\mathcal{G}$ be an operator on $\mathcal{B}^{\psi}(X)$ with domain D, and $\omega \in \mathbb{R}$. $\mathcal{G}$ is closable with its closure $\overline{\mathcal{G}}$ generating a generalised Feller semigroup $(P_t)_{t \geqslant 0}$ with $\|P_t\|_{L(\mathcal{B}^{\psi}(X))} \leqslant \exp(\omega t)$ for all $t \geqslant 0$ if and only if*

(i) D is dense,

(ii) $\mathcal{G} - \lambda_0$ has dense image for some $\lambda_0 > \omega$, and

(iii) $\mathcal{G}$ satisfies the generalised positive maximum principle, *that is, for $f \in D$ with $(\psi^{-1} f) \vee 0 \leqslant \psi(z)^{-1} f(z)$ for some $z \in X$, $\mathcal{G} f(z) \leqslant \omega f(z)$.*

Here, $a \vee b := \max(a, b)$. Note that $(\psi^{-1} f) \vee 0 = \psi^{-1}(f \vee 0)$ as $\psi > 0$. Therefore, $(\psi^{-1} f) \vee 0 \leqslant \psi^{-1}(z) f(z)$ is equivalent to

$$\|f \vee 0\|_{\psi} \leqslant \psi^{-1}(z) f(z). \tag{2.26}$$

Proof. We mimic the proof of [57, Theorem 17.11]. Assume first that $(P_t)_{t\geqslant 0}$ is a generalised Feller semigroup satisfying

$$\|P_t\|_{L(\mathcal{B}^\psi(X))} \leqslant \exp(\omega t), \tag{2.27}$$

and $\mathcal{G}$ with domain D is its generator. For $f \in D$ with $\|f \vee 0\|_\psi \leqslant \psi^{-1}(z) f(z)$,

$$\begin{aligned} P_t f(z) &\leqslant P_t(f \vee 0)(z) \leqslant \psi(z)\|P_t(f \vee 0)\|_\psi \leqslant \psi(z)\exp(\omega t)\|f \vee 0\|_\psi \\ &\leqslant \exp(\omega t) f(z), \end{aligned} \tag{2.28}$$

and due to the continuity of point evaluation, we obtain the inequality $\mathcal{G}f(z) \leqslant \omega f(z)$ in the limit $t \to 0+$. Thus, $\mathcal{G}$ satisfies the generalised positive maximum principle. The density of D and $(\mathcal{G} - \lambda_0)D$ follows from the Lumer-Phillips theorem, Proposition A.6, as $(\exp(-\omega t)P_t)_{t\geqslant 0}$ is a strongly continuous semigroup of contractions.

For the other direction, let $f \in D$ be arbitrary, and define $g := (\operatorname{sgn} f(z))f$, where z is chosen such that $\psi(z)^{-1}|f(z)| = \|f\|_\psi$ (this is possible due to Theorem 2.10). Clearly, $g \in D$ and $\psi(x)^{-1}g(x) \leqslant \psi(z)^{-1}g(z)$, so the generalised positive maximum principle yields $\mathcal{G}g(z) \leqslant \omega g(z)$. Thus, for $\lambda > 0$,

$$\begin{aligned} \|(\lambda - (\mathcal{G} - \omega))f\|_\psi &\geqslant \psi(z)^{-1}\left(\lambda g(z) - (\mathcal{G} - \omega)g(z)\right) \geqslant \psi(z)^{-1}\lambda g(z) \\ &= \lambda\|f\|_\psi. \end{aligned} \tag{2.29}$$

From this, closability of $\mathcal{G}$ follows: if $(f_n)_{n\in\mathbb{N}}$ in D are given such that both $\lim_{n\to\infty}\|f_n\|_\psi = 0$ and $\lim_{n\to\infty}\|\mathcal{G}f_n - g\|_\psi = 0$, there exist $(g_m)_{m\in\mathbb{N}}$ in D with $\lim_{m\to\infty}\|g_m - g\|_\psi = 0$. Thus, for any $\lambda > 0$ and m, $n \in \mathbb{N}$,

$$\|(\lambda - (\mathcal{G} - \omega))(g_m + \lambda f_n)\|_\psi \geqslant \lambda\|g_m + \lambda f_n\|. \tag{2.30}$$

Taking the limit $n \to \infty$, dividing by λ and taking the limit $\lambda \to \infty$, we obtain $\|g_m - g\|_\psi \geqslant \|g_m\|_\psi$, and the limit $m \to \infty$ yields $g = 0$. This proves the closability of $\mathcal{G}$, and the closure $\overline{\mathcal{G}}$ of $\mathcal{G}$ with domain $\mathcal{D}$ satisfies

$$\|(\lambda - (\overline{\mathcal{G}} - \omega))f\|_\psi \geqslant \lambda\|f\|_\psi \quad \text{for all } \lambda > 0 \text{ and } f \in \mathcal{D}. \tag{2.31}$$

Thus, $\overline{\mathcal{G}} - \omega$ is dissipative. The Lumer-Phillips theorem, Proposition A.6, yields that $\overline{\mathcal{G}}$ generates a semigroup with $\|P_t\|_{L(\mathcal{B}^\psi(X))} \leqslant \exp(\omega t)$ for all $t \geqslant 0$.

We now prove positivity of $R_\lambda := (\lambda - \overline{\mathcal{G}})^{-1}$ for every $\lambda > \omega$, which yields that P_t is positive for every $t \geqslant 0$ (by an application of [57, Corollary V.5.5]). To this end, we show that given $g \in \mathcal{B}^\psi(X)$ such that the solution $f \in \mathcal{D}$ of

$(\lambda - \overline{\mathcal{G}})f = g$ is not positive, g cannot be positive, either. By assumption, $\alpha := \inf_{x\in X} \psi(x)^{-1} f(x) < 0$. Given a sequence of functions $(f_n)_{n\in\mathbb{N}}$ in D converging to f such that $\mathcal{G} f_n$ converges to $\overline{\mathcal{G}} f$, we see that we can assume without loss of generality that for every $n \in \mathbb{N}$, $\alpha_n := \inf_{x\in X} \psi(x)^{-1} f_n(x) < 0$, and we have that $\lim_{n\to\infty} \alpha_n = \alpha$. Theorem 2.10 yields the existence of $z_n \in X$ with $\psi(z_n)^{-1} f_n(z_n) = \alpha_n$. By the positive maximum principle, $\mathcal{G} f_n(z_n) \geqslant \omega f_n(z_n)$. Thus,

$$
\begin{aligned}
\inf_{x\in X} \psi(x)^{-1} g(x) &= \lim_{n\to\infty} \inf_{x\in X} \psi(x)^{-1} (\lambda - \mathcal{G}) f_n(x) \\
&\leqslant \lim_{n\to\infty} \psi(z_n)^{-1} (\lambda - \mathcal{G}) f_n(z_n) \\
&\leqslant \lim_{n\to\infty} \psi(z_n)^{-1} (\lambda - \omega) f_n(z_n) \\
&= (\lambda - \omega) \lim_{n\to\infty} \inf_{x\in X} \psi(x)^{-1} f_n(x) \\
&= (\lambda - \omega) \inf_{x\in X} \psi(x)^{-1} f(x) = (\lambda - \omega)\alpha < 0, \qquad (2.32)
\end{aligned}
$$

that is, g is not positive. □

2.3 Results on dual spaces

In this section we consider a special class of state spaces that will be crucial for our applications to stochastic partial differential equations: dual spaces of Banach spaces equipped with the weak-$*$ topology. We remark that the weak topology on Hilbert spaces and sequential weak continuity was also used by Maslowski and Seidler [72] to prove ergodicity of stochastic partial differential equations.

Assume that $(X, \|\cdot\|_X)$ is the dual space of some Banach space $(W, \|\cdot\|_W)$. We will use the weak-$*$ topology on X, and denote this space by X_{w*}. Such a space is clearly endowed with a uniform structure in the sense of [17, Chapitre II § 1 Définition 1], and thus completely regular Hausdorff [18, Chapitre IX § 1 Théorème 2]. Consider a lower semicontinuous function $\psi\colon X \to (0, \infty)$. Due to the Banach-Alaoglu theorem [99, Theorem 3.15], compactness of K_R follows from boundedness, which gives us a simple way to prove the admissibility of ψ.

Assume from now on that (X_{w*}, ψ) is a weighted space. We shall always understand $K_R := \{x \in X\colon \psi(x) \leqslant R\}$ to be endowed with the weak-$*$ topology.

Afterwards, we shall also consider the issue of differentiability of functions in $\mathcal{B}^{\psi}_{k}(X_{w*})$. This motivates the next definition.

Definition 2.14. Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space. A function ψ is called *D-admissible weight function* if and only if it is an admissible weight function and for every $x \in X$, there exists some $R > 0$ such that $B_\varepsilon(x) \subset K_R$ for some $\varepsilon > 0$, where $B_\varepsilon(x) := \{y \in X : \|y - x\|_X < \varepsilon\}$ is the open ε-ball around x.

It is called *C-admissible weight function* if and only if φ is bounded from below, weak-$*$ lower semicontinuous, and if for every $x \in X$, there exists some $\varepsilon > 0$ such that φ is bounded on the closed ε-ball $C_\varepsilon(x) := \{z \in X : \|z - x\|_X \leqslant \varepsilon\}$.

Remark 2.15. We do not require C-admissible weight functions to be admissible. However, ψ is D-admissible if and only if it is admissible and C-admissible.

Example 2.16. Typical examples for weight functions are of the form $\psi(x) = \rho(\|x\|)$, where $\rho : [0, \infty) \to (0, \infty)$ is increasing, left-continuous, and satisfies $\lim_{\xi \to \infty} \rho(\xi) = +\infty$. We will call such weight functions to be *of type ρ*. In this case,

$$K_R = C_r(0) := \{x \in X : \|x\|_X \leqslant r\}, \tag{2.33}$$

where $r := \max\{p \in \mathbb{R} : \rho(p) \leqslant R\}$, and $C_r(0)$ is weak-$*$ compact by the Banach-Alaoglu theorem. Note that $\rho(r) \leqslant R$ by left continuity. Clearly, any such weight function is D-admissible. Below, we will consider choices such as $\rho(t) = (1 + t^2)^{s/2}$, $s \geqslant 2$, $\rho(t) = \cosh(\beta t)$, $\beta > 0$, and $\rho(t) = \exp(\eta t^2)$, $\eta > 0$.

2.3.1 Approximation by smooth functions

We want to give an approximation result for functions in $\mathcal{B}^\psi(X_{w*})$ by cylindrical functions.

Definition 2.17. Let $(X, \|\cdot\|_X)$ be the dual space of a Banach space $(W, \|\cdot\|_W)$, and let $(Y, \|\cdot\|_Y)$ be a Banach space. For $N \in \mathbb{N}$, set

$$\begin{aligned} \mathcal{A}_N(X; Y) := \big\{ g(\langle \cdot, w_1 \rangle, \ldots, \langle \cdot, w_N \rangle) : \ & g \in C_b^\infty(\mathbb{R}^N; Y) \\ & \text{and } w_j \in W, \ j = 1, \ldots, N \big\}. \end{aligned} \tag{2.34}$$

$\mathcal{A}(X; Y) := \bigcup_{N \in \mathbb{N}} \mathcal{A}_N(X; Y)$ is called the space of *bounded smooth cylindrical functions* on X with values in Y. For $Y = \mathbb{R}$, we set $\mathcal{A}_N(X) := \mathcal{A}_N(X; \mathbb{R})$ and $\mathcal{A}(X) := \mathcal{A}(X; \mathbb{R})$.

Clearly, $\mathcal{A}(X; Y) \subset \mathcal{B}^\psi(X_{w*}; Y)$.

Theorem 2.18. *If $(X, \|\cdot\|_X)$ is the dual space of a Banach space $(W, \|\cdot\|_W)$, the closure of $\mathcal{A}(X)$ in $\mathrm{B}^\psi(X_{w*})$ coincides with $\mathcal{B}^\psi(X_{w*})$.*

Proof. We prove first by the Stone-Weierstrass theorem [99] that $\mathcal{A}$ is dense in $\mathrm{C}_b(K_R)$ for any $R > 0$. First, it is obvious that $\mathcal{A}(X)$ is an algebra, as $\mathcal{A}_N(X) \cdot \mathcal{A}_M(X) \subset \mathcal{A}_{N+M}(X)$ for all N and M with obvious notation, and $\mathcal{A}_N(X) \subset \mathcal{A}_{N+1}(X)$ for all $N \in \mathbb{N}$. Moreover, for any $x_1 \neq x_2$, x_1, $x_2 \in K_R$, there exists some $w \in W$ with $\langle x_1, w\rangle \neq \langle x_2, w\rangle$, which yields that $\mathcal{A}_1(X)$ separates points. As the constant functions are in $\mathcal{A}(X)$, we obtain density in $\mathrm{C}_b(K_R)$.

Let $f \in \mathrm{C}_b(X_{w*})$. For every $R > 0$ and $\varepsilon > 0$, there exists some $N \in \mathbb{N}$ and $\tilde{f}_{R,\varepsilon} \in \mathcal{A}_N(X) \subset \mathcal{B}^\psi(X)$ with

$$\sup_{x \in K_R} |f(x) - \tilde{f}_{R,\varepsilon}(x)| < \varepsilon. \tag{2.35}$$

By definition, $\tilde{f}_{R,\varepsilon} = \tilde{g} \circ h$ with $h(x) = (\langle x, w_j\rangle)_{j=1,\dots,N}$ for some $w_j \in W$, $j = 1, \dots, N$, and $\tilde{g} \in \mathrm{C}_b^\infty(\mathbb{R}^N)$. As K_R is compact, $h(K_R) \subset \mathbb{R}^N$ is compact. By the Tietze extension theorem [18, Chapitre IX § 4 Théorème 2], we can extend $\tilde{g}|_{h(K_R)}$ to a continuous function $\hat{g}$ on $\mathbb{R}^N$ with $\sup_{y \in \mathbb{R}^N} |\hat{g}(y)| \leqslant \sup_{x \in K_R} |\tilde{f}_{R,\varepsilon}(x)|$. Applying [19, Proposition IV.21, Proposition IV.20], we see that convolution of $\hat{g}$ with a mollifier yields a function $g \in \mathrm{C}_b^\infty(\mathbb{R}^N)$ with $\sup_{y \in \mathbb{R}^N} |g(y)| \leqslant \sup_{x \in K_R} |\tilde{f}_{R,\varepsilon}(x)|$ and $\sup_{y \in h(K_R)} |g(y) - \tilde{g}(y)| < \varepsilon$. Assuming without loss of generality that

$$\sup_{x \in K_R} |\tilde{f}_{R,\varepsilon}(x)| \leqslant 2 \sup_{x \in K_R} |f(x)|, \tag{2.36}$$

we see that $f_{R,\varepsilon} := g \circ h$ satisfies

$$\sup_{x \in K_R} |f(x) - f_{R,\varepsilon}(x)| < 2\varepsilon \quad \text{and} \quad \sup_{x \in X} |f_{R,\varepsilon}(x)| \leqslant 2 \sup_{x \in X} |f(x)|, \tag{2.37}$$

independently of R and ε. Therefore, as $\psi(x) \geqslant \delta$ for all $x \in X$,

$$\begin{aligned} \|f - f_{R,\varepsilon}\|_\psi &\leqslant \sup_{x \in K_R} \psi(x)^{-1} |f(x) - f_{R,\varepsilon}(x)| + \sup_{\psi(x) > R} \psi(x)^{-1} |f(x) - f_{R,\varepsilon}(x)| \\ &\leqslant \delta^{-1} \sup_{x \in K_R} |f(x) - f_{R,\varepsilon}(x)| + 3R^{-1} \sup_{x \in X} |f(x)|. \end{aligned} \tag{2.38}$$

The result follows. □

The definition of $\mathcal{A}(X)$ is not “optimal” in the sense that it will contain too many functions. The following result is significantly better in this respect.

Theorem 2.19. *Let $(X, \|\cdot\|_X)$ be the dual space of the Banach space $(W, \|\cdot\|_W)$. Assume that W is separable, and let $\{w_j : j \in \mathbb{N}\} \subset W$ be a countable set which separates the points of $X = W^*$. Define*

$$\widetilde{\mathcal{A}}_N(X) := \left\{g(\langle\cdot, w_1\rangle, \dots, \langle\cdot, w_N\rangle) : g \in \mathrm{C}_b^\infty(\mathbb{R}^N)\right\}, \tag{2.39}$$

and $\widetilde{\mathcal{A}}(X) := \bigcup_{N\in\mathbb{N}} \widetilde{\mathcal{A}}_N(X) \subset \mathcal{B}^\psi(X_{w*})$. Then, $\widetilde{\mathcal{A}}(X)$ is dense in $\mathcal{B}^\psi(X_{w*})$.

Proof. The proof is done in the same way as for Theorem 2.18, using that for any $x_1, x_2 \in X$ with $x_1 \neq x_2$, there exists some $j \in \mathbb{N}$ with $\langle x_1, w_j\rangle \neq \langle x_2, w_j\rangle$. □

Remark 2.20. A possible choice for $\{w_j : j \in \mathbb{N}\}$ is given by any countable dense set in Y. For X a separable Hilbert space, we can use any orthonormal basis $(e_n)_{n\in\mathbb{N}}$. We note that the specific choice of the w_j does not make any difference, which was also observed in [48, Remark 5.9].

2.3.2 Connections with weak-$*$ continuity

As we define the spaces $\mathcal{B}^\psi(X_{w*})$ with respect to the weak-$*$ topology, it is not surprising that there is a characterisation of its functions by weak-$*$ continuity. The next result makes this precise.

Lemma 2.21. *Assume that $(X, \|\cdot\|_X)$ is the dual space of a separable Banach space $(W, \|\cdot\|_W)$.*

(i) $f \in \mathcal{B}^\psi(X_{w*})$ if and only if f satisfies (2.9) *and $f|_{K_R}$ is sequentially weak-$*$ continuous for any $R > 0$.*

(ii) Assume that for every $r > 0$, there exists some $R > 0$ with $C_r(0) \subset K_R$. Then, every $f \in \mathcal{B}^\psi(X_{w*})$ is sequentially weak-$*$ continuous. In particular, in this case, $\mathcal{B}^\psi(X_{w*}) \subset \mathrm{C}(X)$.

Here, $\mathrm{C}(X)$ denotes the set of functions $f : X \to \mathbb{R}$ continuous in the norm topology.

Proof. By Theorem 2.8, we only have to equate sequential weak-$*$ and weak-$*$ continuity of $f|_{K_R}$ for any $R > 0$. By compactness, K_R is bounded by the Banach-Steinhaus theorem [19, Théorème II.1], as for any $y \in Y$,

$$\sup_{x\in K_R} |\langle x, y\rangle| < \infty. \tag{2.40}$$

Thus, [19, Théorème III.25] shows that the topology of K_R is metrisable, which means that weak-$*$ continuity and sequential weak-$*$ continuity coincide. Therefore, any function f is sequentially weak-$*$ continuous if and only if it is weak-$*$ continuous on K_R, and the first claim follows.

For the second claim, note that any weak-$*$ converging sequence $(x_n)_{n\in\mathbb{N}}$ is bounded by the Banach-Steinhaus theorem. Thus, by assumption, $(x_n)_{n\in\mathbb{N}}$ stays in K_R for some $R > 0$, and the weak-$*$ continuity of $f|_{K_R}$ yields the result. Finally, every such f is continuous with respect to the norm topology, as every norm convergent sequence converges weak-$*$, as well. □

Remark 2.22. Unfortunately, the condition in the second part of the above Lemma is stronger than D-admissibility. This is shown by the example of X a separable Hilbert space, and

$$\psi(x) := 1 + \|x\| + \sum_{n\in\mathbb{N}} n\chi_{M_n}(x). \tag{2.41}$$

Here,

$$M_n := \left\{x \in X : |\langle c, e_n - x\rangle| < 2^{-n-2}\right\} \tag{2.42}$$

and $c = \sum_{n\in\mathbb{N}} 2^{-n} e_n$ with $(e_n)_{n\in\mathbb{N}}$ an orthonormal basis of X. As the M_n are pairwise disjoint, at most one term in the sum is nonzero in the definition of ψ. Therefore, ψ is locally finite, whence D-admissible. However, we see that the supremum of ψ on $C_1(0)$ is infinite. Even worse: consider a function $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = 0$ for $x \leqslant 0$, and $f|_{(0,\infty)}$ is continuous with $f(2^{-n}) = 1$ and $f|_{[5\cdot 2^{-n-2}, 3\cdot 2^{-n-1}]} = 0$ for $n \in \mathbb{N}$. Then, for $g(x) := f(\langle c, x\rangle)$, $g \in \mathcal{B}^{\psi}(X_{w*})$, as it can be approximated in the ψ-norm by $g_n := g\chi_{\{x\in X :\, |\langle c,x\rangle| > 3\cdot 2^{-n-1}\}}$.

But g is not weakly continuous at zero ($g(e_n) = 1$ for $n \in \mathbb{N}$, but $g(0) = 0$). Thus, not every space $\mathcal{B}^{\psi}(X_{w*})$ with D-admissible ψ allows a characterisation by sequential weak-$*$ continuity.

2.3.3 Differentiable functions and $\mathcal{B}^{\psi}$ spaces

To consider differentiable functions with controlled growth, we need to refine our definitions from before.

Definition 2.23. Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be Banach spaces. $L_j(X; Y)$ denotes the linear space of bounded, j-linear maps $a : X^j \to Y$, which is a Banach space with respect to the norm

$$\|a\|_{L_j(X;Y)} := \sup_{\|h_i\|_X \leqslant 1, i=1,\dots,j} \|a(h_1,\dots,h_j)\|_Y. \tag{2.43}$$

For $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ Banach spaces, $\mathrm{C}^k(X;Y)$ denotes the space of k times continuously Fréchet differentiable functions $f\colon X \to Y$.

Definition 2.24. Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space, $k \geqslant 0$, and $\varphi = (\varphi_j)_{j=0,\dots,k}$, $\varphi_j\colon X \to (0,\infty)$ bounded from below by some $\delta > 0$, $j = 0,\dots,k$, and $(Y, \|\cdot\|_Y)$ a Banach space. We set

$$\begin{aligned}\mathrm{B}^\varphi_k(X;Y) := \Big\{f \in \mathrm{C}^k(X;Y)\colon & \sup_{x\in X} \varphi_j(x)^{-1}\|D^j f(x)\|_{L_j(X;Y)} < \infty \\ & \text{for } j = 0,\dots,k\Big\}.\end{aligned} \tag{2.44}$$

$\mathrm{B}^\varphi_k(X;Y)$ is called the *enveloping space* and is endowed with the norm

$$\|f\|_{\varphi,k} := \|f\|_{\varphi_0} + \sum_{j=1}^{k} |f|_{\varphi_j,j}, \tag{2.45}$$

where the seminorms $|\cdot|_{\varphi_j,j}$ are given by

$$|f|_{\varphi_j,j} := \sup_{x\in X} \varphi_j(x)^{-1}\|D^j f(x)\|_{L_j(X;Y)}. \tag{2.46}$$

If $Y = \mathbb{R}$, we define $\mathrm{B}^\varphi_k(X) := \mathrm{B}^\varphi_k(X;\mathbb{R})$.

Theorem 2.25. *Let $k \in \mathbb{N}$, and assume that $\varphi = (\varphi_j)_{j=0,\dots,k}$ is a vector of C-admissible weight functions. Then, $\mathrm{B}^\varphi_k(X_{w*};Y)$ is a Banach space.*

Proof. Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in this space. It is clear that f_n admits a pointwise limit f. Moreover, it follows that for every $x \in X$, there exists $\varepsilon > 0$ such that $(f_n|_{C_\varepsilon(x)})_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathrm{C}^k(C_\varepsilon(x);Y)$. Here, $C_\varepsilon(x) := \{z \in X\colon \|z - x\|_X \leqslant \varepsilon\}$ denotes the closed ε-ball around x. But this entails that $f|_{C_\varepsilon(x)} \in \mathrm{C}^k(C_\varepsilon(x);Y)$. As differentiability is a local property, we see that $f \in \mathrm{C}^k(X;Y)$. The necessary estimates for f and its derivatives are now easy to see. □

Remark 2.26. Note that not every admissible weight function is D-admissible, as already the counterexample $X = \mathbb{R}$, $\psi(x) := 1 + x^2 + x^{-1}\chi_{(0,\infty)}$ with $\chi_A(x) := 1$ for $x \in A$ and 0 for $x \notin A$ the indicator of A shows. However, such an assumption is necessary to be able to transfer differentiability properties to limits when using weighted supremum norms.

Let us consider a concrete example. Choose the admissible weight function $\psi(x) := 1 + x^2 + x^{-2}\chi_{(0,\infty)}$ on $X = \mathbb{R}$. Let $f_n \in C_b^\infty(\mathbb{R})$, $n \in \mathbb{N}$, be such that

$$f_n(x) := \begin{cases} 0, & x \leqslant 0, \\ 1, & x \geqslant n^{-1}, \end{cases} \tag{2.47}$$

and monotone on $(0, n^{-1})$ such that $0 \leqslant f_n(x) \leqslant 1$ for all $x \in \mathbb{R}$. This can be done in such a way that $|f_n'(x)| \leqslant Cn$ on $(0, n^{-1})$ for some $C > 0$ independent of $n \in \mathbb{N}$, for example by choosing f_1 as required and setting $f_n(x) := f_1(nx)$. Then, for $n \in \mathbb{N}$ and $m \geqslant n$,

$$\|f_n - f_m\|_\psi \leqslant 2 \sup_{x \in (0,n^{-1})} \psi(x)^{-1} = 2(1+n^2)^{-1} \quad \text{and} \tag{2.48}$$

$$|f_n|_{\psi,1} \leqslant 2 \sup_{x \in (0,n^{-1})} \psi(x)^{-1} Cn = 2Cn(1+n^2)^{-1}, \tag{2.49}$$

from which

$$|f_n - f_m|_{\psi,1} \leqslant 2C \left(\frac{n}{(1+n)^2} + \frac{m}{(1+m)^2} \right). \tag{2.50}$$

It follows that $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $B_1^\psi(\mathbb{R})$. As evaluation functionals are continuous, we see that the only candidate for the limit is $f = \chi_{(0,\infty)}$. This function, however, is not in $B_1^\psi(\mathbb{R})$, and is not even continuous.

Note that this is not a contradiction to the characterisation of $\mathcal{B}^\psi(\mathbb{R})$ by continuity, as no set $K_R := \{x \in \mathbb{R} \colon \psi(x) \leqslant R\}$ contains a neighbourhood of $x = 0$.

The aim is now to consider differentiability in the setting of $\mathcal{B}^\psi$-spaces in such a way that we can analyse vector fields and determine their effects on differentiability.

Definition 2.27. Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space and $(Y, \|\cdot\|_Y)$ be a Banach space. Let $\psi = (\psi_j)_{j=0,\dots,k}$ with ψ_j D-admissible weight functions for $j = 0, \dots, k$, $k \in \mathbb{N}$. The space $\mathcal{B}_k^\psi(X_{w*}; Y)$ is the closure of $\mathcal{A}(X; Y)$ in $B_k^\psi(X; Y)$. For $Y = \mathbb{R}$, we set $\mathcal{B}_k^\psi(X_{w*}) := \mathcal{B}_k^\psi(X_{w*}; \mathbb{R})$.

$\mathcal{B}_k^\psi(X_{w*})$ is a Banach space by Theorem 2.25. This definition coincides with the earlier one given for completely regular spaces X by Theorem 2.18.

If X is a Hilbert space and ψ is a function of the norm, the approximating functions can be chosen in a simple manner.

Theorem 2.28. *Let $(X, \|\cdot\|_X)$ be a separable Hilbert space and $\psi = (\psi_j)_{j=0,\dots,k}$, ψ_j of type ρ, $j = 0, \dots, k$ (see Example 2.16). Then, for every $\varepsilon > 0$ and $f \in \mathcal{B}_k^\psi(X_w; Y)$, there exists an orthogonal projection $\pi\colon X \to X$ of finite rank such that $\|f - f \circ \pi\|_{\psi,k} < \varepsilon$.*

Proof. Given ε and f as in the statement of the theorem, we know that there exists $f_\varepsilon \in \mathcal{A}(X; Y)$ such that $\|f - f_\varepsilon\|_{\psi,k} < \varepsilon/2$. This implies existence of $N \in \mathbb{N}$, $g_\varepsilon \in C_b^\infty(\mathbb{R}^N; Y)$ and $e_j \in X$, $j = 1, \dots, N$ with $f_\varepsilon(x) = g_\varepsilon(\langle x, e_1\rangle, \dots, \langle x, e_N\rangle)$. With π the orthogonal projection onto $\operatorname{span}\{e_j : j = 1, \dots, N\}$, $\|\pi x\|_X \leqslant \|x\|_X$ for $x \in X$ yields

$$|f_\varepsilon - f \circ \pi|_{\psi_j,j} \leqslant \sup_{x\in X} \psi_j(x)^{-1} \|D^j f_\varepsilon(\pi x) - D^j f(\pi x)\|_{L_j(X;Y)} < \varepsilon/2, \tag{2.51}$$

whence the triangle inequality yields the claim. □

Theorem 2.29. *Assume that $(X, \|\cdot\|_X)$ is the dual space of the separable Banach space $(W, \|\cdot\|_W)$. Let $(w_n)_{n\in\mathbb{N}}$ in W be such that* $\operatorname{span}\{w_j : j \in \mathbb{N}\}$ *is dense in W. Then, for any Banach space $(Y, \|\cdot\|_Y)$, vector $\psi = (\psi_j)_{j=0,\dots,k}$ of D-admissible weight functions, and $k \geqslant 0$, the space $\tilde{\mathcal{A}}(X; Y) := \bigcup_{N\in\mathbb{N}} \tilde{\mathcal{A}}_N(X; Y)$ is dense in $\mathcal{B}_k^\psi(X; Y)$, where*

$$\tilde{\mathcal{A}}_N(X; Y) := \left\{ g(\langle \cdot, w_1\rangle, \dots, \langle \cdot, w_N\rangle) : g \in C_b^\infty(\mathbb{R}^N; Y) \right\}. \tag{2.52}$$

Proof. We only need to show that for every $f \in \mathcal{A}_N(X; Y)$ and every $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbb{N}$ and $f_\varepsilon \in \tilde{\mathcal{A}}_{N_\varepsilon}(X; Y)$ with $\|f - f_\varepsilon\|_{\psi,k} < \varepsilon$. Similarly as in the proof of Theorem 2.18, we can restrict ourselves to K_R for some appropriate $R > 0$, as the error on $X \backslash K_R$ is less than ε for R large enough. As K_R is bounded, there exists $r > 0$ with $\|x\|_X \leqslant r$ for all $x \in K_R$.

Let $f = g \circ h$, where $g \in C_b^\infty(\mathbb{R}^N; Y)$ and $h\colon X \to \mathbb{R}^N$ is given by $h(x) = (\langle x, \omega_j\rangle)_{j=1}^N$ for some $\omega_j \in W$, $j = 1, \dots, N$. The smoothness of g yields existence of $C > 0$ such that

$$\begin{aligned} &\|D^j g(\xi) - D^j g(\eta)\|_{L_j(\mathbb{R}^N;Y)} < C \max_{r=1,\dots,N} |\xi_r - \eta_r| \\ &\text{for all } \xi = (\xi_r)_{r=1,\dots,N},\ \eta = (\eta_r)_{r=1,\dots,N} \in \mathbb{R}^N \text{ and } j = 0, \dots, k. \end{aligned} \tag{2.53}$$

By assumption, we can find $N_\varepsilon \in \mathbb{N}$ and $A = (\alpha_{j\ell}) \in \mathbb{R}^{N\times N_\varepsilon}$ such that for every $j = 1, \ldots, N$, $\|\omega_j - \sum_{\ell=1}^{N_\varepsilon} \alpha_{j\ell} w_\ell\|_W < r^{-1}C^{-1}\varepsilon$. Define now

$$g_\varepsilon \colon \mathbb{R}^{N_\varepsilon} \to Y, \quad x \mapsto g_\varepsilon(x) := g(Ax). \tag{2.54}$$

It follows that for $f_\varepsilon := g_\varepsilon(\langle\cdot, w_1\rangle, \ldots, \langle\cdot, w_{N_\varepsilon}\rangle)$ and $x \in K_R$,

$$\begin{aligned} |D^j f(x) - D^j f_\varepsilon(x)| &= \Big|D^j g(h(x)) - D^j g\Big(\big(\textstyle\sum_{\ell=1}^{N_\varepsilon} \alpha_{j\ell}\langle x, w_\ell\rangle\big)_{j=1}^N\Big)\Big| \\ &\leqslant C \max_{r=1,\ldots,N} |\langle x, \omega_r - \textstyle\sum_{\ell=1}^{N_\varepsilon} \alpha_{j\ell} w_\ell\rangle| \leqslant \varepsilon. \end{aligned} \tag{2.55}$$

This proves the claim. □

We provide a particularly interesting class of functions that is dense in $\mathcal{B}_k^\psi(X_{w*})$. Recall that a Banach space $(Z, \|\cdot\|_Z)$ has the *approximation property* if and only if for every $K \subset X$ compact and $\varepsilon > 0$ there exists an operator T of finite rank such that $\|Tx - x\|_Z < \varepsilon$ for all $x \in K$, see [68, Definition 1.e.1]. Note that every separable Hilbert space has the approximation property as it has a Schauder basis, see the discussion after [68, Definition 1.e.1].

Definition 2.30. Let $(Z, \|\cdot\|_Z)$, $(Y, \|\cdot\|_Y)$ be Banach spaces. We define

$$\mathrm{C}_b^k(Z;Y) := \left\{f \in \mathrm{C}^k(Z;Y) \colon \|f\|_{\mathrm{C}_b^k(Z;Y)} < \infty\right\}, \tag{2.56}$$

equipped with the norm

$$\|f\|_{\mathrm{C}_b^k(Z;Y)} := \sum_{j=0}^k \|D^j f\|_{C_b(Z;L_j(Z;Y))}. \tag{2.57}$$

Here, we have set $\|f\|_{C_b(U;V)} := \sup_{u\in U}\|f(u)\|_V$ for two normed linear spaces $(U, \|\cdot\|_U)$ and $(V, \|\cdot\|_V)$.

Theorem 2.31. *Let $(X, \|\cdot\|_X)$ be a separable, reflexive Banach space, endowed with a vector $\psi = (\psi_j)_{j=0,\ldots,k}$ of D-admissible weight functions, $k \geqslant 0$. Assume that X is compactly embedded into another Banach space $(Z, \|\cdot\|_Z)$ with the approximation property.*

Then, for every Banach space $(Y, \|\cdot\|_Y)$, we have that

$$\mathrm{C}_b^k(Z;Y) \subset \mathcal{B}_k^\psi(X_w;Y). \tag{2.58}$$

Furthermore, $\mathrm{C}_b^k(Z;Y)$ is dense in $\mathcal{B}_k^\psi(X_w;Y)$.

Proof. Note first that due to the reflexivity of X, we have that $X_w = X_{w*}$.

Denote the compact embedding of X into Z by $\iota\colon X \to Z$. [68, Theorem 1.e.4] proves that there exists a sequence $(\iota_n)_{n\in\mathbb{N}}$ of finite rank operators, $\iota_n\colon X \to Z$, such that for all $r > 0$, $\lim_{n\to\infty} \sup_{\|x\|_X \leqslant r} \|\iota_n x - \iota x\|_Z = 0$.

Let $f \in C_b^k(Z; Y)$ be given. Define $f_n := f \circ \iota_n$; we need to prove that $\lim_{n\to\infty} \|f_n - f\|_{\psi,k} = 0$. Given $\varepsilon > 0$, choose $R := \varepsilon^{-1} \|f\|_{C_b^k(Z;Y)}$ such that $\sum_{j=0}^k \sup_{x\in X\setminus K_R} \psi(x)^{-1} \|D^j f\|_{L_j(X;Y)} \leqslant \varepsilon$. As K_R is bounded, it follows that $\iota(K_R)$ is compact in the norm topology of Z. Hence, for every $\varepsilon > 0$ there exists some $\delta > 0$ such that if $x \in K_R$ and $z \in Z$ with $\|\iota x - z\|_Z < \delta$, then $\|D^j f(\iota x) - D^j f(z)\|_{L_j(Z;Y)} < \varepsilon$ for $j = 0, \dots, k$. Choose now $n_0 \in \mathbb{N}$ large enough such that $\|\iota x - \iota_n x\|_Z < \delta$ for all $x \in K_R$ and $n \geqslant n_0$. Hence, $\|f_n - f\|_{\psi,k} < \varepsilon$, which proves $f \in \mathcal{B}_k^{\psi}(X_w; Y)$.

To prove the density, we apply Theorem 2.29. As $\iota\colon X \to Z$ is injective and X is reflexive, we see that $\iota^*\colon Z^* \to X^* = W$ has dense range. Hence, we can choose a sequence $(\zeta_n)_{n\in\mathbb{N}}$ in Z^* such that $\operatorname{span}\{w_j : j \in \mathbb{N}\}$ is dense in W, where $w_j = \iota^* \zeta_j \in W$, $j \in \mathbb{N}$. Defining $\tilde{\mathcal{A}}(X; Y)$ with this sequence, we see that every $f = g(\langle\cdot, w_1\rangle, \dots, \langle\cdot, w_N\rangle) \in \tilde{\mathcal{A}}_N(X; Y)$ can be extended to a function $\tilde{f}\colon Z \to Y$ with $\tilde{f}(x) = f(x)$ for $x \in X$ by virtue of

$$\tilde{f} := g(\langle\cdot, \zeta_1\rangle, \dots, \langle\cdot, \zeta_n\rangle). \tag{2.59}$$

Clearly, $\tilde{f} \in C_b^k(Z; Y)$. The proof is thus complete. □

For $f \in \mathcal{A}(X; Y)$, $D^j f(x)$ is not a general multilinear form for $x \in X$ and $j = 0, \dots, k$. It is actually *completely continuous*. Let us recall first the definition of this property.

Definition 2.32. For Banach spaces $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$, we define $V_j(X; Y) \subset L_j(X; Y)$ to be the linear space of *completely continuous multilinear forms*, i.e., $a \in V_j(X; Y)$ if and only if for all sequences $x_{i,n}$ converging weakly to x_i, $i = 1, \dots, j$, $a(x_{1,n}, \dots, x_{j,n})$ converges strongly to $a(x_1, \dots, x_j)$.

Proposition 2.33. *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be Banach spaces. $V_j(X; Y)$ is closed in $L_j(X; Y)$.*

In particular, $(V_j(X; Y), \|\cdot\|_{L_j(X;Y)})$ is a Banach space.

Proof. A more general result (showing that the completely continuous multilinear maps even form an *ideal*) can be found in [98, Theorem 2.3]. We reproduce the easy proof of the statement given above.

Let $a \in L_j(X;Y)$ be in the closure of $V_j(X;Y)$. Then, given $\varepsilon > 0$, there exists $a' \in V_j(X;Y)$ with $\|a - a'\|_{L_j(X;Y)} < \varepsilon$. If $x_{i,n}$ converges weakly to x_i, $i = 1,\dots,j$, there exists $n_0 \in \mathbb{N}$ such that for $n \geqslant n_0$, $\|a'(x_1,\dots,x_j) - a'(x_{1,n},\dots,x_{j,n})\|_Y < \varepsilon$. As we can assume without loss of generality that $\|x_{i,n}\| \leqslant 1$ for $i = 1,\dots,j$ and $n \in \mathbb{N}$,

$$\begin{aligned}\|a(x_1,\dots,x_j) - a(x_{1,n},\dots,x_{j,n})\|_Y &\leqslant \|a(x_1,\dots,x_j) - a'(x_1,\dots,x_j)\|_Y \\ &\quad + \|a'(x_1,\dots,x_j) - a'(x_{1,n},\dots,x_{j,n})\|_Y \\ &\quad + \|a'(x_{1,n},\dots,x_{j,n}) - a(x_{1,n},\dots,x_{j,n})\|_Y \\ &< 3\varepsilon, \end{aligned} \tag{2.60}$$

and the result follows. □

Corollary 2.34. *Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space endowed with a vector ψ of D-admissible weight functions, and $(Y, \|\cdot\|_Y)$ a Banach space. If $f \in \mathcal{B}^{\psi}_k(X_{w*};Y)$, then $D^j f \in \mathcal{B}^{\psi_j}(X_w; V_j(X;Y))$ for $j = 0,\dots,k$.*

Proof. For $f \in \mathcal{A}(X;Y)$, $f = g \circ \pi$ with $\pi\colon X \to \mathbb{R}^N$ bounded and linear and $g \in C^\infty_b(\mathbb{R}^N;Y)$, whence

$$D^j f(x)(x_1,\dots,x_j) = D^j g(\pi x)(\pi x_1,\dots,\pi x_j), \tag{2.61}$$

and as π is a compact operator, we obtain $D^i f(x) \in V_j(X;Y)$. Furthermore, we see clearly that $D^j f$ is again bounded and cylindrical. Hence, this result extends to $\mathcal{B}^{\psi}_k(X_{w*};Y)$ by a density argument due to Proposition 2.33. □

If X is a Hilbert space, we have the following converse of Corollary 2.34. This generalises the result of Lemma 2.21 to the current setting.

Theorem 2.35. *Let $(X, \|\cdot\|_X)$ be a separable Hilbert space and $\psi = (\psi_j)_{j=0,\dots,k}$, ψ_j of type ρ, $j = 0,\dots,k$ (see Example 2.16), and $(Y, \|\cdot\|_Y)$ a Banach space.*

$f \in \mathcal{B}^{\psi}_k(X_w;Y)$ if and only if

(i) $f \in B^{\psi}_k(X;Y)$,

(ii) $D^j f\colon X \to V_j(X;Y)$ is sequentially completely continuous for $j = 0,\dots,k$ in the sense that if $(x_n)_{n\in\mathbb{N}}$ converges weakly to x, then

$$\lim_{n\to\infty} \|D^j f(x) - D^j f(x_n)\|_{L_j(X;Y)} = 0, \quad \textit{and} \tag{2.62}$$

(iii) $\lim_{R\to\infty} \sup_{\psi_j(x)>R} \psi_j(x)^{-1} \|D^j f(x)\|_{L_j(X;Y)} = 0$ *for* $j = 0, \dots, k$.

Proof. Corollary 2.34, together with a similar argument as in Theorem 2.8 and Lemma 2.21, proves the first part of the equivalence.

For the converse, let $f \in \mathrm{B}_k^{\psi}(X;Y)$ be given with $D^j f \in \mathcal{B}^{\psi}(X_w; V_j(X;Y))$ for $j = 0, \dots, k$. Given $\varepsilon > 0$, there exists by assumption $r > 0$ such that

$$\psi_j(x)^{-1} \|D^j f(x)\|_{L_j(X;Y)} < \varepsilon \quad \text{for } j = 0, \dots, k \text{ and } \|x\|_X > r. \tag{2.63}$$

For every $j = 0, \dots, k$, the mapping $g_j(x, x_1, \dots, x_j) := D^j f(x)(x_1, \dots, x_j)$ is continuous $C_r(0) \times C_1(0)^j \to Y$, where $C_r(0)$ and $C_1(0)$ denote the closed balls of radius r and 1 in X, respectively, and are endowed with the weak topology. Indeed: under the given assumptions, $C_r(0) \times C_1(0)^j$ is metrisable by [19, Théorème III.25], hence we only have to prove sequential continuity. If $(x_n^0, x_n^1, \dots, x_n^j)$ converges weakly to $(x^0, x^1, \dots, x^j)$ in $C_r(0) \times C_1(0)^j$, then $\lim_{n\to\infty} \|D^j f(x^0) - D^j f(x_n^0)\|_{L_j(X;Y)} = 0$. As $D^j f(x) \in V_j(X;Y)$ for all $x \in X$, we can choose $n_0 \in \mathbb{N}$ such that $\|D^j f(x^0)(x^1, \dots, x^j) - D^j f(x^0)(x_n^1, \dots, x_n^j)\|_Y < \varepsilon$ and $\|D^j f(x^0) - D^j f(x_n^0)\|_{L_j(X;Y)} < \varepsilon$ for $n \geqslant n_0$. Hence,

$$\begin{aligned} &\|D^j f(x^0)(x^1, \dots, x^j) - D^j f(x_n^0)(x_n^1, \dots, x_n^j)\|_Y \\ &\qquad \leqslant \|D^j f(x^0)(x^1, \dots, x^j) - D^j f(x^0)(x_n^1, \dots, x_n^j)\|_Y \\ &\qquad\quad + \|D^j f(x^0)(x_n^1, \dots, x_n^j) - D^j f(x_n^0)(x_n^1, \dots, x_n^j)\|_Y \leqslant 2\varepsilon, \end{aligned} \tag{2.64}$$

whence the stated continuity of g_j. As the set $C_r(0) \times C_1(0)^j$ is weakly compact, g_j is uniformly continuous, i.e., there exists a weak neighbourhood U of 0 in X such that for all $j = 0, \dots, k$,

$$\begin{aligned} &\|g_j(x_0)(x_1, \dots, x_j) - g_j(y_0)(y_1, \dots, y_j)\|_Y < \varepsilon \\ &\quad \text{for } (x_i)_{i=0}^j, (y_i)_{i=0}^j \in C_r(0) \times C_1(0)^j \text{ with } x_i - y_i \in U,\ i = 0, \dots, j. \end{aligned} \tag{2.65}$$

By definition of the weak topology, there exist $\xi_\ell \in X$ with $\|\xi_\ell\|_X = 1$, $\ell = 1, \dots, M$, and $\delta > 0$ such that $x \in U$ if $|\langle x, \xi_\ell \rangle_X| < \delta$, $\ell = 1, \dots, M$. Let π denote the orthogonal projection onto $\operatorname{span}\{\xi_\ell \colon \ell = 1, \dots, M\}$. Then, $x - \pi x \in U$ for all $x \in X$. It follows that

$$\begin{aligned} &\|g_j(x_0)(x_1, \dots, x_j) - g_j(\pi x_0)(\pi x_1, \dots, \pi x_j)\|_Y < \varepsilon \\ &\qquad \text{for } (x_i)_{i=0,\dots,j} \in C_r(0) \times C_1(0)^j. \end{aligned} \tag{2.66}$$

Defining $\tilde{f} := f \circ \pi$, it is easy to see that

$$D^j \tilde{f}(x_0)(x_1, \dots, x_j) = D^j f(\pi x_0)(\pi x_1, \dots, \pi x_j). \tag{2.67}$$

As $\tilde{f}$ is only defined on a finite-dimensional space, we can cut it off and smoothen it in a straightforward manner, see also the proof of Theorem 2.18, ensuring uniform convergence on the image of $C_r(0) \times C_1(0)^j$ under π. Hence, we can construct $\hat{f} \in \mathcal{A}(X;Y)$ arbitrarily close to f in the norm of $\mathrm{B}_k^\psi(X;Y)$, and it follows that $f \in \mathcal{B}_k^\psi(X_w;Y)$. □

2.3.4 $\mathcal{B}^\psi$ multipliers

Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space and $(Y\|\cdot\|_Y)$ be a Banach space. Consider the space $\mathcal{B}_k^\psi(X_{w*};Y)$. Assume that for some other Banach spaces $(Z, \|\cdot\|_Z)$ and $(W, \|\cdot\|_W)$, there exists a bilinear mapping $\mathcal{M}\colon Y \times Z \to W$. We want to consider an operator mapping functions $f \in \mathcal{B}_k^\psi(X_{w*};Y)$ to functions $x \mapsto \mathcal{M}(f(x), g(x))$, where $g\colon X \to Z$. This raises the question what assumptions we have to take on g such that this mapping is bounded into another $\mathcal{B}^\psi$ space. First, we settle the boundedness issue in the enveloping space.

Theorem 2.36. *Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space and $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$ and $(W, \|\cdot\|_W)$ be Banach spaces. Given a bounded bilinear mapping $\mathcal{M}\colon Y \times Z \to W$, i.e., for some constant $M > 0$,*

$$\|\mathcal{M}(y,z)\|_W \leqslant M\|y\|_Y\|z\|_Z \quad \text{for all } y \in Y \text{ and } z \in Z. \tag{2.68}$$

For some $k \geqslant 0$, let $\varphi^{(1)} := (\varphi_j^{(1)})_{j=0,\dots,k}$, $\varphi^{(2)} := (\varphi_j^{(2)})_{j=0,\dots,k}$ be vectors of C-admissible weight functions. Set

$$\varphi := (\varphi_j)_{j=0,\dots,k}, \text{ where } \varphi_j(x) := \sum_{i=0}^{j} \binom{j}{i} \varphi_i^{(1)}(x)\varphi_{j-i}^{(2)}(x),\ j = 0, \dots, k. \tag{2.69}$$

Then, φ is a vector of C-admissible weight functions on X, and the mapping

$$\mathcal{M}\colon \mathrm{B}_k^{\varphi^{(1)}}(X;Y) \times \mathrm{B}_k^{\varphi^{(2)}}(X;Z) \to \mathrm{B}_k^{\varphi}(X;W), \tag{2.70a}$$

$$\mathcal{M}(f,g)(x) := \mathcal{M}(f(x), g(x)) \quad \text{for all } x \in X, \tag{2.70b}$$

is well-defined, bilinear, and satisfies

$$\|\mathcal{M}(f,g)\|_{\tilde{\psi},k} \leqslant M\|f\|_{\psi,k}\|g\|_{\varphi,k} \tag{2.71}$$

If $\varphi^{(1)}$ consists of D-admissible weight functions, φ consists of D-admissible weight functions, as well.

Proof. It is clear that φ consists of C-admissible weight functions. To see that φ consists of D-admissible weight functions if $\varphi^{(1)}$ does, we proceed as follows. Set $K^{(1)}_{j,R} := \left\{x \in X : \varphi^{(1)}_j(x) \leqslant R\right\}$ and $K_{j,R} := \{x \in X : \varphi_j(x) \leqslant R\}$. First of all, note that

$$K_{j,R} \subset K^{(1)}_{j,R/\delta}, \tag{2.72}$$

where $\delta := \inf_{x\in X} \varphi^{(2)}_0(x)$. Thus, $K_{j,R}$ is precompact. As the sum and product of lower semicontinuous functions is also lower semicontinuous, it is also closed, thus compact. It follows that φ_j is admissible, whence D-admissible.

For $j = 1, \dots, k$ and $x \in X$, the operator D^j_x defines a symmetric, bounded j-linear map on $\mathrm{B}^{\varphi^{(1)}}_k(X;Y)$ and on $\mathrm{B}^{\varphi^{(2)}}_k(X;Y)$. Hence, the Leibniz rule yields

$$\begin{aligned} D^j\mathcal{M}(f(x),g(x))(h_1,\dots,h_j) &= \sum_{i=0}^{j} \frac{1}{i!(j-i)!}\times \\ &\times \sum_{\sigma\in\mathcal{S}_j} \mathcal{M}(D^i f(x)(h_{\sigma_1},\dots,h_{\sigma_i}), D^{j-i}g(x)(h_{\sigma_{i+1}},\dots,h_{\sigma_j})). \end{aligned} \tag{2.73}$$

It follows that

$$\begin{aligned} \|D^j\mathcal{M}(f(x),g(x))\|_{L_j(X;W)} &\leqslant M\sum_{i=0}^{j}\binom{j}{i}\varphi^{(1)}_i(x)\varphi^{(2)}_{j-i}(x)\|f\|_{\varphi^{(1)},j}\|g\|_{\varphi^{(2)},j} \\ &= M\varphi_j(x)\|f\|_{\varphi^{(1)},j}\|g\|_{\varphi^{(2)},j}, \end{aligned} \tag{2.74}$$

which shows the claimed norm estimate for $\mathcal{M}(f,g)$. □

Definition 2.37. Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space and $(Z, \|\cdot\|_Z)$ be a Banach space. Let φ be a vector of C-admissible weight functions. We say that $g \in \mathcal{C}^{\varphi}_k(X_{w*}; Z)$ if and only if

(i) $g \in \mathrm{B}^{\varphi}_k(X;Z)$,

(ii) for $x \in X$ and $j = 0, \dots, k$, the mapping $(h_1, \dots, h_j) \mapsto D^j g(x)(h_1, \dots, h_j)$ is continuous from the weak-$*$ topology on $C_1(0)^j$ to the weak topology on Z, and

(iii) for $j = 0, \dots, k$, $x \mapsto D^j g(x)(h_1, \dots, h_j)$ is continuous from the weak-$*$ topology on $C_r(0)$ to the weak topology on Z uniformly in $h_1, \dots, h_j \in C_1(0)$ for all $r > 0$, i.e., given $r > 0$, $x \in C_r(0)$ and a weak neighbourhood V of 0 in Z, there exists a weak-$*$ neighbourhood U of 0 in X such that for all $h_i \in C_1(0)$, $i = 1, \dots, j$, and $y \in C_r(0)$ with $x - y \in U$,

$$D^j g(x)(h_1, \dots, h_j) - D^j g(y)(h_1, \dots, h_j) \in V. \tag{2.75}$$

Recall that $C_r(0) := \{x \in X : \|x\|_X \leqslant r\}$ is the closed ball of radius r in X.

It is easy to see that $\mathcal{A}(X; Z) \subset \mathcal{C}_k^{\varphi}(X_{w*}; Z)$. However, the closure of $\mathcal{A}(X; Z)$ in $\mathrm{B}_k^{\varphi}(X; Z)$ does not even contain all bounded linear operators. As the weak-$*$ topology is metrisable on bounded sets by [19, Théorème III.25], we see that the continuity requirements in (ii) and (iii) can also be formulated using sequences as follows:

(ii') for $x \in X$, $j = 0, \dots, k$ and sequences $(h_n^i)_{n \in \mathbb{N}}$ converging weak-$*$ to h^i, $i = 1, \dots, j$, $D^j g(x)(h_n^1, \dots, h_n^j)$ converges weakly to $D^j g(x)(h^1, \dots, h^j)$, and

(iii') for every $x \in X$, sequence $(x_n)_{n \in \mathbb{N}}$ converging weak-$*$ to $x \in X$ and weak neighbourhood V of 0 in Z, there exists $n_0 \in \mathbb{N}$ such that for all $n \geqslant n_0$ and $h_i \in X$ with $\|h_i\|_X \leqslant 1$,

$$D^j g(x)(h_1, \dots, h_j) - D^j g(x_n)(h_1, \dots, h_j) \in V. \tag{2.76}$$

Similarly as Theorem 2.31, the following result yields that classically differentiable functions on a larger space are included in $\mathcal{C}_k^{\varphi}(X_w; Z)$.

Theorem 2.38. *Let $(X, \|\cdot\|_X)$ be a separable Hilbert space endowed with a vector $\varphi = (\varphi_j)_{j=0,\dots,k}$ of C-admissible weight functions, $k \geqslant 0$, $(Z, \|\cdot\|_Z)$ a reflexive Banach space. Assume that there exist Banach spaces $(\tilde{X}, \|\cdot\|_{\tilde{X}})$, $(\tilde{Z}, \|\cdot\|_{\tilde{Z}})$ such that X is compactly embedded in $\tilde{X}$, and Z continuously embedded in $\tilde{Z}$.*

Suppose $g \colon \tilde{X} \to \tilde{Z}$ satisfies $g \in \mathrm{C}^k(\tilde{X}; \tilde{Z})$, $g(X) \subset Z$ and $g \in \mathrm{B}_k^{\varphi}(X; Z)$. Then $g \in \mathcal{C}_k^{\varphi}(X_w; Z)$.

Proof. Both X and Z are reflexive. Hence, the respective weak and weak-$*$ topologies coincide.

We first prove (ii') above. Let, therefore, $(h^i_n)_{n\in\mathbb{N}}$ be sequences converging weakly to h^i in X, $i = 1, \dots, j$, $j = 0, \dots, k$, and fix $x \in X$. It follows that $(h^i_n)_{n\in\mathbb{N}}$ converges strongly to h^i in $\tilde{X}$. As $g \in \mathrm{C}^k(\tilde{X}; \tilde{Z})$, we see that $D^j g(x)(h^1_n, \dots, h^j_n)$ converges strongly to $D^j g(x)(h^1, \dots, h^j)$ in $\tilde{Z}$.

On the other hand, $g \in \mathrm{B}^{\varphi}_k(X; Z)$ yields that $(D^j g(x)(h^1_n, \dots, h^j_n))_{n\in\mathbb{N}}$ is a bounded sequence in Z. Thus, by [19, Théorème III.27], every subsequence admits a subsequence converging weakly in Z, and by the continuous embedding $Z \to \tilde{Z}$, it follows that all these limits have to agree with $D^j g(x)(h^1, \dots, h^j)$. We obtain weak convergence of $(D^j g(x)(h^1_n, \dots, h^j_n)_{n\in\mathbb{N}}$ to $D^j g(x)(h^1, \dots, h^j)$, proving (ii').

The proof of (iii') is similar. Given a sequence $(x_n)_{n\in\mathbb{N}}$ converging weakly to x in X, $(x_n)_{n\in\mathbb{N}}$ converges strongly in $\tilde{X}$. Hence, for all $\zeta \in \tilde{Z}^*$ and $\varepsilon > 0$, we can choose $n_0 \in \mathbb{N}$ such that for all $n \geqslant n_0$ and $h_i \in \tilde{X}$ with $\|h_i\|_{\tilde{X}} \leqslant 1$, $i = 1, \dots, j$,

$$|\zeta\big(D^j g(x)(h_1, \dots, h_j) - D^j g(x_n)(h_1, \dots, h_j)\big)| < \varepsilon. \tag{2.77}$$

As the injection $\iota\colon Z \to \tilde{Z}$ is injective and continuous, $\tilde{Z}^*$ is dense in Z^*, and (2.77) extends to $\zeta \in Z^*$ and $h_i \in X$ with $\|h_i\|_X \leqslant 1$, $i = 1, \dots, j$, by possibly adjusting n_0 to account for the operator norm of ι. This proves (iii'). As $g \in \mathrm{B}^{\varphi}_k(X; Z)$ by assumption, we obtain that $g \in \mathcal{C}^{\varphi}_k(X_w; Z)$, as claimed. □

Theorem 2.39. *Given a separable Hilbert space $(X, \|\cdot\|_X)$ endowed with D-admissible weight functions $\psi = (\psi_j)_{j=0,\dots,k}$ and C-admissible weight functions $\varphi = (\varphi_j)_{j=0,\dots,k}$, $k \geqslant 0$, and Banach spaces $(Y, \|\cdot\|_Y)$, $(Z, \|\cdot\|_Z)$ and $(W, \|\cdot\|_W)$. Assume that the bounded bilinear mapping $\mathcal{M}\colon Y \times Z \to W$ satisfies $\lim_{n\to\infty}\|\mathcal{M}(y, z - z_n)\|_W = 0$ for each $y \in Y$ and every sequence $(z_n)_{n\in\mathbb{N}}$ converging weakly to $z \in Z$. Define $\tilde{\psi} := (\tilde{\psi}_j)_{j=0,\dots,k}$ according to (2.69), i.e.,*

$$\tilde{\psi}_j(x) := \sum_{i=0}^{j} \binom{j}{i} \psi_i(x)\varphi_{j-i}(x) \quad \text{for } j = 0, \dots, k \text{ and } x \in X, \tag{2.78}$$

and suppose that $\tilde{\psi}_j$ is of type ρ, $j = 0, \dots, k$ (see Example 2.16). Then, $\mathcal{M}(f, g) \in \mathcal{B}^{\tilde{\psi}}_k(X_w; W)$ for all $f \in \mathcal{B}^{\psi}_k(X_w; Y)$ and $g \in \mathcal{C}^{\varphi}_k(X_w; Z)$.

If $\psi_i(x) = \rho^{\psi}_i(\|x\|_X)$ and $\varphi_i(x) = \rho^{\varphi}_i(\|x\|_X)$, then $\tilde{\psi}_i$ also has the required representation. The assumption on $\mathcal{M}$ is satisfied if, e.g., $\mathcal{M}(y, z) = \mathcal{M}_0(y, \kappa z)$ with $\mathcal{M}_0$ a bounded bilinear form and κ a compact operator, or if $Y = Z^*$ and $\mathcal{M}$ is the dual pairing of Z with its dual space.

Proof. By Theorem 2.36, we have to satisfy the conditions of Theorem 2.35. Due to (2.73), it is sufficient to show that $g_{i,j}\colon X \to V_j(X;Y)$ is sequentially weakly continuous and $\lim_{R\to\infty}\sup_{\tilde\psi_j(x)>R}\tilde\psi_j(x)^{-1}\|g_{i,j}(x)\|_{L_j(X;Y)} = 0$, where

$$g_{i,j}(x)(h_1,\dots,h_j) := \mathcal{M}(f_i(x)(h_1,\dots,h_i), g_{j-i}(x)(h_{i+1},\dots,h_j)) \quad \text{with} \tag{2.79}$$

$$f_i(x)(h_1,\dots,h_i) := D^i f(x)(h_1,\dots,h_i) \quad \text{and} \tag{2.80}$$

$$g_i(x)(h_1,\dots,h_i) := D^i g(x)(h_1,\dots,h_i). \tag{2.81}$$

From the assumptions, it follows similarly as in the proof of Theorem 2.36 that $g_{i,j}(x) \in L_j(X;Y)$. To see that $g_{i,j}(x) \in V_j(X;Y)$, assume that $(h_n^p)_{n\in\mathbb{N}}$ converges weakly to h^p, $p = 1,\dots,j$. Then, as $D^i f(x) \in V_i(X;Y)$,

$$\lim_{n\to\infty}\|D^i f(x)(h^1,\dots,h^i) - D^i f(x)(h_n^1,\dots,h_n^i)\|_Y = 0. \tag{2.82}$$

Furthermore, as $g \in \mathcal{C}_k^\varphi(X_w;Z)$, $D^{j-i}g(x)(h_n^{i+1},\dots,h_n^j)$ converges weakly to $D^{j-i}g(x)(h^{i+1},\dots,h^j)$. Hence, by the boundedness of $\mathcal{M}$,

$$\begin{aligned}
&\|g_{i,j}(x)(h^1,\dots,h^j) - g_{i,j}(x)(h_n^1,\dots,h_n^j)\|_W\\
&\leqslant M\|f_i(x)(h^1,\dots,h^i) - f_i(x)(h_n^1,\dots,h_n^i)\|_Y\|g_{j-i}(x)(h_n^{i+1},\dots,h_n^j)\|_Z\\
&\quad + \|\mathcal{M}(f_i(x)(h^1,\dots,h^i), g_{j-i}(x)(h^{i+1},\dots,h^j))\\
&\qquad - \mathcal{M}(f_i(x)(h^1,\dots,h^i), g_{j-i}(x)(h_n^{i+1},\dots,h_n^j))\|_W.
\end{aligned} \tag{2.83}$$

By the assumptions on g and $\mathcal{M}$, we see that the above expression converges to 0, and it follows that $g_{i,j}(x)$ is completely continuous.

Next, we prove that $x \mapsto g_{i,j}(x)$ is sequentially continuous from the weak topology on X to the norm topology on $L_j(X;W)$. If $(x_n)_{n\in\mathbb{N}}$ converges weakly to $x \in X$ and $h_p \in C_1(0)$, $p = 1,\dots,j$ is fixed,

$$\begin{aligned}
&\|g_{i,j}(x)(h_1,\dots,h_j) - g_{i,j}(x_n)(h_1,\dots,h_j)\|_W\\
&\leqslant M\|f_i(x)(h_1,\dots,h_i) - f_i(x_n)(h_1,\dots,h_i)\|_Y\|g_{j-i}(h_{i+1},\dots,h_j)\|_Z\\
&\quad + \|\mathcal{M}(f_i(x)(h_1,\dots,h_i), g_{j-i}(x)(h_{i+1},\dots,h_j))\\
&\qquad - \mathcal{M}(f_i(x)(h_1,\dots,h_i), g_{j-i}(x_n)(h_{i+1},\dots,h_j))\|_W.
\end{aligned} \tag{2.84}$$

Again, the assumptions on g and $\mathcal{M}$ yield that the above expression converges to 0, and sequential weak continuity follows.

Finally, we have to ensure that

$$\lim_{R\to\infty}\sup_{\tilde\psi_j(x)>R}\tilde\psi_j(x)^{-1}\|g_{i,j}(x)\|_{L_j(X;Y)} = 0. \tag{2.85}$$

But

$$
\begin{aligned}
\tilde{\psi}_j(x)^{-1}\|g_{i,j}(x)\|_{L_j(X;W)} &\leqslant M\tilde{\psi}_j(x)^{-1}\varphi_{j-i}(x)\|f_i(x)\|_{L_i(X;Y)}\|g_j\|_{\varphi_{j-i}} \\
&\leqslant M\psi_i(x)^{-1}\|f_i(x)\|_{L_i(X;Y)}\|g_j\|_{\varphi_{j-i}}.
\end{aligned}
\tag{2.86}
$$

Corollary 2.34 implies $f_i \in \mathcal{B}^{\psi_i}(X_w; V_i(X;Y))$, and the result follows. □

Remark 2.40. Given the setup of Theorem 2.39, assume that $g \in \mathrm{B}_k^{\varphi}(X;Z)$ is such that the conclusion holds true. Choose $Y := Z^*$ and $\mathcal{M}(y,z) := \langle y, z\rangle_{Z^*,Z}$, the dual pairing of Z^* and Z. We set $f \equiv y \in Z^*$; clearly, $f \in \mathcal{A}(X;Z^*)$. Hence, Theorem 2.35 implies that $X \to V_j(X;\mathbb{R})$, $x \mapsto g_j(x)$, is sequentially completely continuous, where

$$
g_j(x)(h_1,\dots,h_j) := \langle y, D^j g(x)(h_1,\dots,h_j)\rangle_{Z^*,Z}. \tag{2.87}
$$

It follows that $(h_1,\dots,h_j) \mapsto D^j g(x)(h_1,\dots,h_j)$ is continuous from the weak-$*$ topology on $C_1(0)^j$ to the weak topology of Z. Fix $\varepsilon > 0$, $r > 0$ and $x \in C_r(0)$. Equating again sequential weak continuity and weak continuity on $C_r(0)$, we obtain existence of a weak neighbourhood U of 0 in X such that for $y \in C_r(0)$ with $x - y \in U$,

$$
\|g_j(x) - g_j(y)\|_{L_j(X;\mathbb{R})} < \varepsilon. \tag{2.88}
$$

This implies that for all $h_i \in C_1(0)$, $i = 1,\dots,j$,

$$
|\langle y, D^j g(x)(h_1,\dots,h_j) - D^j g(y)(h_1,\dots,h_j)\rangle_{Z^*,Z}| < \varepsilon. \tag{2.89}
$$

Altogether, we see that $g \in \mathcal{C}_k^{\varphi}(X_w;Z)$. Hence, $g \in \mathcal{C}_k^{\varphi}(X_w;Z)$ is necessary and sufficient for g to be a general multiplier in $\mathcal{B}^{\psi}$ spaces.

2.3.5 Vector fields and $\mathcal{B}^{\psi}$ spaces

We want to construct a Lie derivative in the $\mathcal{B}^{\psi}$ setting, i.e., a directional derivative $\mathcal{L}_V f$ of a function $f\colon X \to Y$ along a vector field $V\colon X \to X$. As we also need to consider vector fields that are only defined on proper subspaces of the given space, e.g., the derivative along an unbounded operator, we shall state our results for this more general setting.

The following result is clear from our definitions.

Theorem 2.41. *Let $(X, \|\cdot\|_X)$ be the dual space of a separable Banach space, and $(Y, \|\cdot\|_Y)$ a Banach space. Furthermore, let $\psi = (\psi_j)_{j=0,\dots,k}$ be a vector of D-admissible weight functions, and set $\hat{\psi} := (\psi_{j+1})_{j=0,\dots,k-1}$. Then, the linear mapping*

$$D\colon \mathcal{B}_k^{\psi}(X_{w*}; Y) \to \mathcal{B}_{k-1}^{\hat{\psi}}(X_{w*}; L(X;Y)), \quad f \mapsto Df, \tag{2.90}$$

is a bounded operator.

Corollary 2.42. *Suppose $(X, \|\cdot\|_X)$, $(Z, \|\cdot\|_Z)$ are separable Hilbert spaces with $Z \subset X$ continuously embedded. Let $\psi = (\psi_j)_{j=0,\dots,k}$ be a vector of D-admissible weight functions on X, $\hat{\psi} = (\hat{\psi}_j)_{j=0,\dots,k-1}$ a vector of D-admissible weight functions on Z with $\psi_{j+1}(x) \leqslant \hat{\psi}_j(x)$ for $j = 0, \dots, k-1$, and $\varphi = (\varphi_j)_{j=0,\dots,k-1}$ a vector of C-admissible weight functions on Z. Define $\tilde{\psi} := (\tilde{\psi}_j)_{j=0,\dots,k}$ by*

$$\tilde{\psi}_j(x) := \sum_{i=0}^{j} \binom{j}{i} \hat{\psi}_i(x)\varphi_{j-i}(x) \quad \text{for } j = 0, \dots, k-1 \text{ and } z \in Z, \tag{2.91}$$

and assume $\tilde{\psi}_j$ is of type ρ, $j = 0, \dots, k$ (see Example 2.16).

Then, the mapping

$$\mathcal{L}\colon \mathcal{B}_k^{\psi}(X_{w*}) \times \mathcal{C}_{k-1}^{\varphi}(Z_{w*}; X) \to \mathcal{B}_{k-1}^{\tilde{\psi}}(Z_{w*}), \quad (f, V) \mapsto \mathcal{L}_V f, \tag{2.92}$$

given by

$$\mathcal{L}_V f(x) := Df(x)V(x), \tag{2.93}$$

is bounded and bilinear.

Proof. Theorem 2.41 together with the assumed relation between ψ and $\hat{\psi}$ prove that the mapping $D\colon \mathcal{B}_k^{\psi}(X_{w*}) \to \mathcal{B}_{k-1}^{\hat{\psi}}(Z_{w*}; L(X;\mathbb{R}))$ is continuous. Consider $\mathcal{M}\colon L(X;\mathbb{R}) \times X \to \mathbb{R}$, given by $\mathcal{M}(x^*, x) = \langle x^*, x\rangle_{L(X;\mathbb{R}),X}$, the dual pairing. It follows that

$$\mathcal{L}_V f = \mathcal{M}(Df, V), \tag{2.94}$$

and Theorem 2.39 proves the claim. □

Let us consider two special cases.

Corollary 2.43. *Let $(H, \|\cdot\|_H)$ be a separable Hilbert space, $(Z, \|\cdot\|_Z)$ a continuously embedded Hilbert space. Define the D-admissible weight functions $\psi_j(x) := \cosh(\|x\|_H)$ on H and $\hat{\psi}_j(x) := \cosh(\|x\|_Z)$ on Z and the C-admissible weight functions $\varphi_j(x) := 1$ on Z, $j \geqslant 0$. Then, for every $k \geqslant 0$, the mapping*

$$\mathcal{L}\colon \mathcal{B}^{\psi}_k(X_{w*}) \times \mathcal{C}^{\varphi}_{k-1}(Z_{w*}; X) \to \mathcal{B}^{\hat{\psi}}_{k-1}(Z_{w*}), \quad (f, V) \mapsto \mathcal{L}_V f, \tag{2.95}$$

given by $\mathcal{L}_V f(x) := Df(x)V(x)$, is bounded and bilinear.

Remark 2.44. If $Z = H$, this has the simple interpretation that bounded vector fields map cosh-weighted spaces into themselves.

Proof. This is straightforward from Corollary 2.42, as the $\tilde{\psi}_j$ defined there is only a multiple of $\hat{\psi}_j$ in this case. □

The following special case is very useful in the analysis of stochastic partial differential equations of Da Prato-Zabczyk type.

Corollary 2.45. *Let $(H, \|\cdot\|_H)$ be a separable Hilbert space, $(Z, \|\cdot\|_Z)$ a continuously embedded Hilbert space. Fix $n \in \mathbb{N}$. Define the D-admissible weight functions $\psi_j(x) := (1 + \|x\|_H^2)^{(n-j)/2}$ on H and $\hat{\psi}_j(x) := (1 + \|x\|_Z^2)^{(n-j)/2}$ on Z, $j = 0, \dots, n-1$, and the C-admissible weight functions $\varphi_0(x) := (1 + \|x\|_Z^2)^{1/2}$ and $\varphi_j(x) := 1$ on Z, $j \in \mathbb{N}$. Then, for $k \leqslant n-1$, the mapping*

$$\mathcal{L}\colon \mathcal{B}^{\psi}_k(X_{w*}) \times \mathcal{C}^{\varphi}_{k-1}(Z_{w*}; X) \to \mathcal{B}^{\hat{\psi}}_{k-1}(Z_{w*}), \quad (f, V) \mapsto \mathcal{L}_V f, \tag{2.96}$$

given by $\mathcal{L}_V f(x) := Df(x)V(x)$, is bounded and bilinear.

Remark 2.46. This has the interpretation that linearly bounded vector fields $Z \to X$ with bounded derivatives (hence also Lipschitz continuous) map polynomially bounded functions to polynomially bounded functions, with the same weights.

Proof. Calculating

$$\begin{aligned}\tilde{\psi}_j(x) &:= (1 + \|x\|_Z^2)^{(n-1)/2}(1 + \|x\|_Z^2)^{1/2} + \sum_{i=0}^{j} \binom{j}{i} (1 + \|x\|_Z^2)^{(n-i-1)/2} \\ &\leqslant C\hat{\psi}_j(x), \qquad (2.97)\end{aligned}$$

the claim again follows from an application of Corollary 2.42. □

Let us consider some concrete examples.

Example 2.47. Let $(X, \|\cdot\|_H)$, $(Y, \|\cdot\|_Z)$ be separable Hilbert spaces, and $A\colon X \to Y$ a continuous linear operator. It is easy to see that $V_A \in \mathcal{C}^{\varphi}_k(X_w; Y)$, where $V_A(x) := Ax$ and φ is chosen as in Corollary 2.45. We check the definition: $V_A \in \mathrm{B}^{\varphi}_k(X; Y)$ is obvious, as $DV_A(x)(h) = Ah$ and $D^j V_A = 0$ for $j \geqslant 2$. Furthermore, if $(h_n)_{n\in\mathbb{N}}$ converges weakly to $h \in X$, then Ah_n converges weakly to $Ah \in Y$. Finally, $DV_A(x)$ is independent of $x \in X$, and the claim follows.

Note that the assumptions are satisfied for densely defined and closed operators $A\colon \operatorname{dom} A \subset H \to H$ on separable Hilbert spaces, in particular for infinitesimal generators of strongly continuous semigroups (see Proposition A.5(i)).

Example 2.48. Given a separable Hilbert space $(H, \|\cdot\|_H)$ of functions defined on a bounded set $D \subset \mathbb{R}^d$ with smooth boundary. Let $G\colon H \to H$ be a *Nemytskii* or *superposition operator*, i.e., with some $g\colon \mathbb{R} \to \mathbb{R}$,

$$\Phi(f)(x) = \varphi(f(x)) \quad \text{for } x \in D \text{ and } f \in H. \tag{2.98}$$

For an analysis of the mapping properties of such operators on diverse state spaces, see [3, 100].

Assume that $H = \mathrm{H}^s(D)$, the usual Sobolev space of s times weakly differentiable functions on D with weak derivatives in $\mathrm{L}^2(D)$, where $s > d/2$. Then, [100, p. 381, Theorem 2], together with the Sobolev embedding theorem (e.g., [100, p. 32, Theorem 1]), proves that $G\colon \mathrm{H}^s(D) \to \mathrm{H}^s(D)$ is infinitely often Fréchet differentiable if $g' \in \mathrm{C}^\infty(\mathbb{R})$ and $g(0) = 0$. Similarly, [77, Lemma 1.3.3] proves that $G\colon \mathrm{H}^s_0(D) \to \mathrm{H}^s_0(D)$ is C^∞ if $s > d/2$ and $g \in \mathrm{C}^\infty(\mathbb{R})$, where $\mathrm{H}^s_0(D)$ denotes the subspace of $\mathrm{H}^s(D)$ consisting of functions vanishing on ∂D together with all derivatives up to order $s - 1$.

Let us consider the case $H = \mathrm{H}^s(D)$ with $s - 1 > d/2$, and $g' \in \mathrm{C}^\infty_b(\mathbb{R})$ with $g(0) = 0$. We want to prove that $G \in \mathcal{C}^{\varphi}_s(H_w; H)$, where φ is chosen as in Corollary 2.45. First, [100, p. 381, Theorem 2] cited above yields $G \in \mathrm{B}^{\varphi}_s(H; H)$, as we have the exact representation

$$D^j G(u)(h_1, \ldots, h_j) = g^{(j)}(u) h_1 \cdots h_j \quad \text{for } u,\ h_1, \ldots, h_j \in \mathrm{H}^k(D). \tag{2.99}$$

Furthermore, $G\colon \mathrm{H}^{s-1}(D) \to \mathrm{H}^{s-1}(D)$ is infinitely often differentiable. The Rellich-Kondrachev theorem [100, p. 82, Theorems 1, 2] proves that the inclusion $\mathrm{H}^s(D) \to \mathrm{H}^{s-1}(D)$ is compact. Hence, Theorem 2.38 yields the claim.

Such operators are of interest in the context of analysis and numerics for stochastic partial differential equations, consider, e.g., [107, 62]. Note that G will be neither weakly continuous nor Fréchet differentiable if $H = \mathrm{L}^2(D)$ unless it is affine; see [3, Section 3.6] and [97, Section 1.3].

Chapter 3

Stochastic Processes And Weighted Spaces

In this chapter, we consider semigroups induced by Markov processes, in particular those solving stochastic partial differential equations, on weighted spaces. Section 3.1 establishes sufficient conditions such that the semigroup generated by a Markov process is strongly continuous on an appropriate weighted space. In Section 3.2, we analyse these conditions for solutions of stochastic partial differential equations. Moreover, we provide Taylor expansions of the Markov semigroup through an explicit representation of its infinitesimal generator using vector fields on $\mathcal{B}^{\psi}(X_{w*})$. Finally, Section 3.3 presents results on the smoothing effects of stochastic partial differential equations with sectorial generator, i.e., analytic semigroup.

Note that for simplicity and ease of representation, we restrict ourselves to equations driven by Brownian motions. It is possible to deal with more general Lévy driving processes in a similar manner, see [105] in this regard.

3.1 Strong continuity and Markov semigroups

Assume that $(X, \|\cdot\|_X)$ is the dual space of a separable Banach space $(W, \|\cdot\|_W)$. Again, we write X_{w*} for X endowed with the weak-$*$ topology.

Assumption 3.1. $(x(t, x_0))_{t\geqslant 0}$ is a time homogeneous Markov process with values in X on some stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geqslant 0})$ satisfying the usual conditions, started at $x_0 \in X$. It has right continuous trajectories with respect to the weak-$*$ topology on X.

We want to derive conditions on $(x(t, x_0))_{t\geqslant 0}$ such that the Markov semigroup $(P_t)_{t\geqslant 0}$ of $(x(t, x_0))_{t\geqslant 0}$, given by $P_t f(x_0) := \mathbb{E}\left[f(x(t, x_0))\right]$, is strongly continuous on the space $\mathcal{B}^{\psi}(X_{w*})$ for an adequate weight function ψ.

Assumption 3.2. ψ is an admissible weight function on X. There exist constants $C > 0$ and $\varepsilon > 0$ with

$$\mathbb{E}[\psi(x(t,x_0))] \leqslant C\psi(x_0) \quad \text{for all } x_0 \in X \text{ and } t \in [0,\varepsilon]. \tag{3.1}$$

Inequality (3.1) is related to boundedness of the transition operator on $\mathcal{B}^{\psi}(X_{w*})$, and to some supermartingale property. This is formulated in the following lemma.

Lemma 3.3. *Suppose Assumptions 3.1 and 3.2. Then* $|\mathbb{E}[f(x(t,x_0))]| \leqslant C\psi(x_0)$ *for all* $f \in \mathcal{B}^{\psi}(X_{w*})$*,* $x_0 \in X$ *and* $t \in [0,\varepsilon]$*.*

Furthermore, the condition

$$\mathbb{E}[\psi(x(t,x_0))] \leqslant \exp(\omega t)\psi(x_0) \quad \text{for all } x_0 \in X \text{ and } t \in [0,\varepsilon]. \tag{3.2}$$

is equivalent to the process $\exp(-\omega t)\psi(x(t,x_0))$ *being a supermartingale in its own filtration. This implies*

$$|\mathbb{E}[f(x(t,x_0))]| \leqslant \exp(\omega t)\psi(x_0) \quad \text{for } x_0 \in X \text{ and } t \geqslant 0 \tag{3.3}$$

for all $f \in \mathcal{B}^{\psi}(X_{w*})$*.*

Proof. This is clear from the definitions. □

Lemma 3.4. *Suppose Assumptions 3.1 and 3.2. Then*

$$\lim_{t\to 0+} \mathbb{E}[f(x(t,x_0))] = f(x_0) \quad \text{for all } f \in \mathcal{B}^{\psi}(X_{w*}) \text{ and } x_0 \in X. \tag{3.4}$$

Proof. Denoting by χ_A the indicator function of the set A, we choose $R > \psi(x_0)$ and consider

$$\begin{aligned}|\mathbb{E}[f(x(t,x_0))] - f(x_0)| \leqslant &\mathbb{E}\left[|f(x(t,x_0)) - f(x_0)|\chi_{[\psi(x(t,x_0))\leqslant R]}\right]\\ &+ \mathbb{E}\left[|f(x(t,x_0))|\chi_{[\psi(x(t,x_0))>R]}\right]\\ &+ f(x_0)\mathbb{P}[\psi(x(t,x_0)) > R]\,.\end{aligned} \tag{3.5}$$

By the Markov inequality,

$$\mathbb{P}[\psi(x(t,x_0)) > R] \leqslant R^{-1}\mathbb{E}[\psi(x(t,x_0))] \leqslant CR^{-1}\psi(x_0). \tag{3.6}$$

Given $\varepsilon > 0$, Theorem 2.8 shows that for some $R > 0$, $|f(x)| \leqslant \varepsilon\psi(x)$ if $\psi(x) > R$. Therefore,

$$\mathbb{E}\left[|f(x(t,x_0))|\chi_{[\psi(x(t,x_0))>R]}\right] \leqslant C\varepsilon\psi(x_0). \tag{3.7}$$

Finally, given $R > 0$, $\sup_{\psi(x)\leqslant R}|f(x)| < \infty$ by weak continuity. By dominated convergence, $\lim_{t\to 0+}\mathbb{E}\left[|f(x(t,x_0)) - f(x_0)|\chi_{[\psi(x(t,x_0))\leqslant R]}\right] = 0$. □

Theorem 3.5. *Suppose Assumptions 3.1 and 3.2. Let $\{w_j : j \in \mathbb{N}\} \subset W$ be a countable set which separates the points of X. Assume that for any $t > 0$, $j \in \mathbb{N}$ and sequence $(x_n)_{n\in\mathbb{N}}$ converging weak-$*$ to some $x_0 \in X$,*

$$\lim_{n\to\infty} \langle x(t, x_n), w_j\rangle = \langle x(t, x_0), w_j\rangle \quad \textit{almost surely.} \tag{3.8}$$

Then, $P_t f(x_0) := \mathbb{E}[f(x(t, x_0))]$ satisfies the generalised Feller property and is therefore a strongly continuous semigroup on $\mathcal{B}^\psi(X_{w*})$.

The condition given here is weaker than assuming that the map $x_0 \mapsto x(t, x_0)$ is almost surely weak-$*$ continuous, as the nullset can depend on t, x_0, the sequence $(x_n)_{n\in\mathbb{N}}$, and $j \in \mathbb{N}$ (even though the dependence on j can be removed, as a countable union of nullsets is again a nullset). If X is a separable Hilbert space, $\{w_j : j \in \mathbb{N}\}$ can be chosen to be an orthonormal basis.

Proof. Let $f = g \circ h$ with $g \in C_b^\infty(\mathbb{R}^n)$ and $h(x) = (\langle x, y_j\rangle)_{j=1,\ldots,n}$. Such functions are dense in $\mathcal{B}^\psi(X_{w*})$ by Theorem 2.19. By Lemma 2.21, we only have to prove sequential weak-$*$ continuity of $P_t f$ for $f \in \mathcal{B}^\psi(X_{w*})$. By assumption, for any weak-$*$ converging sequence $(x_n)_{n\in\mathbb{N}}$ with limit x_0 , $\lim_{n\to\infty} h(x(t, x_n)) = h(x(t, x_0))$ almost surely. The dominated convergence theorem yields $P_t f \in \mathcal{B}^\psi(X_{w*})$. The result now follows from Lemma 3.4 and Theorem 2.11. □

Example 3.6. Suppose $x(t, x_0) = x_0 + L_t$, where L_t is a càdlàg Lévy process with jumps bounded by some constant $c > 0$ in X. Then, by Fernique's theorem [85, Theorem 4.4], it follows that $\mathbb{E}[\exp(\beta\|L_t\|)] < \infty$ for all $\beta > 0$. Choosing $\psi(x) := \cosh(\beta\|x\|)$, we see that $\psi(x + y) \leqslant 2\psi(x)\psi(y)$. Hence,

$$\mathbb{E}[\psi(x(t, x_0))] \leqslant 2\mathbb{E}[\psi(L_t)]\psi(x_0). \tag{3.9}$$

We obtain from Theorem 3.5 that every càdlàg Lévy process on a Hilbert space with bounded jumps induces a strongly continuous semigroup on a cosh-weighted space $\mathcal{B}^\psi(X_{w*})$.

The continuity assumptions of Theorem 3.5 are typically not easy to verify directly in the weak-$*$ topology. The following theorem yields a simpler approach by using a compact embedding in a reflexive setting.

Theorem 3.7. *Suppose Assumptions 3.1 and 3.2, and that X is reflexive. Let $(Z, \|\cdot\|_Z)$ be another Banach space such that X is compactly embedded in Z.*

Furthermore, suppose that the Markov process $(x(t,x_0))_{t\geqslant 0}$ on X can be extended to Z, and that for any $f \in C_b(Z)$, the mapping $z_0 \mapsto \mathbb{E}[f(x(t,z_0))]$ is continuous with respect to the norm topology of Z.

Then, $P_t f(x_0) := \mathbb{E}[f(x(t,x_0))]$ satisfies the generalised Feller property and is therefore a strongly continuous semigroup on $\mathcal{B}^{\psi}(X_{w*})$.

Remark 3.8. Note that for concrete examples, we often work the other way round: first, we prove existence of the process on Z, then we prove the invariance and continuity properties for $x(t,x_0)$ on X and Z. It is actually a result on preservation of regularity, when showing that $x(t,x_0) \in X$ almost surely if $x_0 \in X$.

Proof. We only need to prove that there exists a dense subset of $\mathcal{B}^{\psi}(X_{w*})$ that is mapped into $\mathcal{B}^{\psi}(X_{w*})$. But Theorem 2.31 and the assumptions of the theorem show that this is satisfied for $C_b(Z)$, as $\sup_{z_0\in Z}|\mathbb{E}[f(x(t,z_0))]| \leqslant \sup_{z_0\in Z}|f(z_0)|$. □

Example 3.9. Continuity in norm topologies, as required in Theorem 3.7, is often satisfied in applications for stochastic partial differential equations, see Proposition B.20. The classical Rellich-Kondrachov type embedding theorems, see [19, Théorème IX.16], yield compact embeddings for problems on bounded domains.

Theorem 3.10. *Suppose Assumptions 3.1 and 3.2, and that X is a separable Hilbert space with scalar product $\langle \cdot, \cdot \rangle$ and countable orthonormal basis $(e_j)_{j\in\mathbb{N}}$. Denoting by π_M the orthogonal projection onto the span of the first M basis vectors, assume that for $j \in \mathbb{N}$,*

$$\lim_{M\to\infty} \sup_{x_0\in X} \psi(x_0)^{-1}\mathbb{E}\left[|\langle x(t,x_0), e_j\rangle - \langle x(t,\pi_M x_0), e_j\rangle|\right] = 0. \tag{3.10}$$

Then, the semigroup $(P_t)_{t\geqslant 0}$ defined by $P_t f(x_0) := \mathbb{E}[f(x(t,x_0))]$ satisfies the generalised Feller property and is therefore strongly continuous on $\mathcal{B}^{\psi}(X_{w*})$.

Proof. For f a bounded and smooth cylinder function with $f = f \circ \pi_N$, consider $g_M := P_t(f \circ \pi_N) \circ \pi_M$. We prove that g_M converges to $P_t(f \circ \pi_N)$. For any $x_0 \in X$, the smoothness of f yields

$$\begin{aligned} |P_t(f\circ\pi_N)(x_0) - g_M(x_0)| &\leqslant \mathbb{E}\left[|f(\pi_N x(t,x_0)) - f(\pi_N x(t,\pi_M x_0))|\right] \\ &\leqslant C_f \mathbb{E}\left[\|\pi_N(x(t,x_0) - x(t,\pi_M x_0))\|\right] \\ &\leqslant C_f \sum_{j=1}^{N} \mathbb{E}\left[|\langle x(t,x_0), e_j\rangle - \langle x(t,\pi_M x_0), e_j\rangle|\right], \end{aligned} \tag{3.11}$$

whence $P_t\mathcal{B}^\psi(X_{w*}) \subset \mathcal{B}^\psi(X_{w*})$, see Remark 2.20. By Lemma 3.3, $P_t \in L(\mathcal{B}^\psi(X_{w*}))$. Again, the result follows from Lemma 3.4 and Theorem 2.11. □

Example 3.11. The assumptions of Theorem 3.10 are satisfied for the stochastic Navier-Stokes equation on the two-dimensional torus with additive noise (a similar result is found in [48, Theorem 5.10]). The first estimate in [48, Theorem A.3] proves the condition of Theorem 3.10, where the weight function is $\psi(x) = \exp(\eta\|x\|^2)$ with $\eta > 0$ chosen in such a way that $\mathbb{E}[\psi(x(t, x_0))] \leqslant K\psi(x_0)$ for small t.

3.2 Application to stochastic partial differential equations

In finite dimensions, the stochastic Taylor expansion (see [59, Chapter 5]) is an important tool in the derivation of both strong and weak approximation schemes for stochastic differential equations. In infinite dimensions, the situation is more complicated. A fundamental issue is that if $(x(t))_{t\geqslant 0}$ is the solution of a stochastic partial differential equation on the Hilbert space H with unbounded infinitesimal generator A, then $Ax(t)$ is not well-defined in general. This can be dealt with by considering initial conditions that lie in the domain of a power of A, and requiring that the vector fields leave these domains of powers of A invariant, see [6].

We suggest an alternative approach, making use of the infinitesimal generator of the strongly continuous Markov semigroup.

Assumption 3.12. Let $(H, \|\cdot\|_H)$ be a Hilbert space and $(A, \operatorname{dom} A)$ the infinitesimal generator of a strongly continuous semigroup $(S_t)_{t\geqslant 0}$ of pseudo-contractions on H. Fix $\ell_0 \in \mathbb{N}$. For $\ell = 0, \dots, \ell_0$, let H_ℓ be subspaces of H endowed with Hilbert norms $\|\cdot\|_{H_\ell}$. Assume that $H_0 = H$, and that for $\ell = 0, \dots, \ell_0 - 1$, $H_{\ell+1} \subset H_\ell$ with continuous and dense embedding, and $A\colon H_{\ell+1} \to H_\ell$ is a bounded linear operator. Furthermore, assume that for $\ell = 0, \dots, \ell_0$, $S_t(H_\ell) \subset H_\ell$ for $t \geqslant 0$, and that $(S_t)_{t\geqslant 0}$ is strongly continuous on H_ℓ.

On H_ℓ, we define D-admissible weight functions

$$\psi_\ell^s(x) := \left(1 + \|x\|_{H_\ell}^2\right)^{s/2}, \quad s \geqslant 1, \quad \ell = 0, \dots, \ell_0, \quad \psi^s := \psi_0^s, \tag{3.12}$$

and the C-admissible weight functions

$$\varphi_{\ell,0}(x) := \left(1 + \|x\|_{H_\ell}^2\right)^{1/2}, \quad \varphi_{\ell,j}(x) := 1, \quad j \geqslant 1. \tag{3.13}$$

Define the vectors of weight functions $\psi_\ell^{(n)} := (\psi_\ell^{n-j})_{j=0,\ldots,k}$, $k < n$, and $\varphi_\ell := (\varphi_{\ell,j})_{j=0,\ldots,k}$.

Assumption 3.13. For some $k_0 \in \mathbb{N}$ and $\ell = 0, \ldots, \ell_0$,

$$V_0 \in \mathcal{C}_{k_0}^{\varphi_\ell}((H_\ell)_w, (H_\ell)_w) \quad \text{and} \tag{3.14}$$

$$V_j \in \mathcal{C}_{k_0}^{\varphi_\ell}((H_\ell)_w, (H_\ell)_w) \quad \text{for } j = 1, \ldots, d. \tag{3.15}$$

Remark 3.14. In the following results, the sharp smoothness requirements on the vector fields vary. The ones given above are sufficient everywhere, and are the most general ones under which a result as in Lemma 4.10 can be expected.

For $x \in H_\ell$, $\ell = 0, \ldots, \ell_0$, we can then consider the Da Prato-Zabczyk equation

$$\mathrm{d}x(t, x_0) = Ax(t, x_0)\mathrm{d}t + \sum_{j=0}^{d} V_j(x(t, x_0)) \circ \mathrm{d}B_t^j, \quad x(t, x_0) = x_0, \tag{3.16}$$

on H_ℓ, where $B_t^0 = t$ and $(B_t^j)_{j=1,\ldots,d}$ is a d-dimensional Brownian motion. As the assumptions on the vector fields V_j essentially mean that they are Lipschitz continuous with bounded derivatives, whence linearly bounded, all these equations have unique solutions in H_ℓ if $x \in H_\ell$, agreeing with each other for sufficiently smooth initial conditions if we vary ℓ.

Assumption 3.15. The Markov semigroup $(P_t)_{t \geqslant 0}$, $P_t f(x_0) := \mathbb{E}[f(x(t, x_0))]$, is strongly continuous on $\mathcal{B}^{\psi_\ell^n}((H_\ell)_{w*})$ for all $n \in \mathbb{N}$ and $\ell = 0, \ldots, \ell_0$.

Recall that Section 3.1 collects several conditions ensuring Assumption 3.15.

An interesting fact is that we can prove pseudocontractivity of $(P_t)_{t \geqslant 0}$ under the assumption of pseudocontractivity of the semigroup generated by A.

Theorem 3.16. *Consider a solution of* (3.16), *where A generates a pseudo-contractive semigroup and the vector fields V_j are Lipschitz continuous. Then,*

$$\|P_t\|_{L(\mathcal{B}^{\psi_\ell^n}((H_\ell)_w))} \leqslant \exp(\omega t) \quad \textit{for some } \omega > 0. \tag{3.17}$$

Remark 3.17. The proof is somehow twisted in infinite dimension and does not follow the usual finite dimensional lines of proving that the local martingale part of $\psi_\ell^n(x(t, x_0))$ is in fact a martingale, and therefore Ito's formula yields the result: we use the Szőkefalvi-Nagy theorem [93, p. 452, Théorème IV] to move to a larger Hilbert space $\mathcal{H}_\ell \supset H_\ell$ containing H_ℓ as a closed subspace and where we can write the solution process $x(t, x_0) = \pi \mathcal{U}_t Y(t, x_0)$ as orthogonal projection.

Proof. We proceed similarly as in [106]. Take $\ell = 0$ without any restriction and set $\psi = \psi_0^n$. Additionally we assume that A generates a contractive semigroup on H by adding the growth to V_0. [93, p. 452, Théorème IV] yields existence of a larger Hilbert space $\mathcal{H} \supset H$, where the semigroup generated by A lifts to a unitary group $\mathcal{U}$ with generator $\mathcal{A}$. The projection onto H is denoted by π. We consider the stochastic partial differential equation prolonged to $\mathcal{H}$,

$$\mathrm{d}X(t, x_0) = \mathcal{A}X(t, x_0)\mathrm{d}t + \sum_{j=0}^{d} V_j(\pi(X(t, x_0))) \circ \mathrm{d}B_t^j. \tag{3.18}$$

Rewriting the above equation using Itô integrals and switching to a "coordinate system" which moves with velocity $x \mapsto \mathcal{A}x$, we obtain a new stochastic differential equation

$$\mathrm{d}Y(t, x_0) = \sum_{j=0}^{d} \mathcal{V}_j(t, Y(t, x_0))\mathrm{d}B_t^j \tag{3.19}$$

with Lipschitz continuous vector fields

$$\mathcal{V}_0(t, y) = \mathcal{U}_{-t}\tilde{V}_0(\pi\mathcal{U}_t y) \quad \text{and} \tag{3.20}$$

$$\mathcal{V}_j(t, y) = \mathcal{U}_{-t}V_j(\pi\mathcal{U}_t y) \quad \text{for } t \geqslant 0,\ y \in \mathcal{H} \text{ and } j = 0, \ldots, d, \tag{3.21}$$

where $\tilde{V}_0(x) := V_0(x) - \frac{1}{2}\sum_{j=1}^{d} DV_j(x)V_j(x)$ is the Itô drift. It follows that $x(t, x_0) = \pi\mathcal{U}_t Y(t, x_0)$ for $t \geqslant 0$ and $x_0 \in H$.

Proposition B.20 yields $\sup_{t\in[0,\varepsilon]} \mathbb{E}[\|Y(t, x_0)\|^p] < \infty$ for $p \geqslant 2$. Ito's formula applied to

$$\psi_{\mathcal{H}}(Y(t, x_0)) := (1 + \|Y(t, x_0)\|^2)^{n/2} \tag{3.22}$$

together with linear growth and Gronwall's inequality then yields the result; more precisely, defining

$$\mathcal{L}_t f(x) := Df(x) \cdot \mathcal{V}_0(t, x) + \frac{1}{2}\sum_{j=1}^{d} D^2 f(x)(\mathcal{V}_j(t, x), \mathcal{V}_j(t, x)), \tag{3.23}$$

we see that

$$\begin{aligned}\mathbb{E}[\psi_{\mathcal{H}}(Y(t,x_0))] =&\psi_{\mathcal{H}}(x_0) + \int_0^t \mathbb{E}[\mathcal{L}_t(\psi_{\mathcal{H}})(Y(s,x_0)]ds\\ \leqslant&\psi(x_0) + \omega\int_0^t \mathbb{E}[\psi_{\mathcal{H}}(Y(s,x_0))]ds,\end{aligned} \tag{3.24}$$

where the constant ω depends on the Lipschitz bounds of the vector fields V_j, $j = 0,\dots,d$. Noting that $\psi(x_0) = \psi_{\mathcal{H}}(x_0)$, we consider $x(t,x_0) = \pi\mathcal{U}_t Y(t,x_0)$ and realise that, due to $\|\pi\mathcal{U}_t y\|_H \leqslant \|y\|_{\mathcal{H}}$ for $y \in \mathcal{H}$,

$$\mathbb{E}[\psi(x(t,x_0)] \leqslant \mathbb{E}[\psi_{\mathcal{H}}(Y(t,x_0))]] \leqslant \exp(\omega t)\psi_{\mathcal{H}}(x_0) = \exp(\omega t)\psi(x_0), \tag{3.25}$$

which is the desired result. □

3.2.1 Weak continuity

One approach to satisfy Assumption 3.15 is the following.

Assumption 3.18. Given Assumption 3.12, we suppose the existence of a Hilbert space H_{-1} such that H_ℓ is compactly and densely embedded in $H_{\ell-1}$ for $\ell = 0,\dots,\ell_0$, and that $V_j\colon H_\ell \to H_\ell$ is Lipschitz continuous for $j = 0,\dots,d$ and $\ell = -1,\dots,\ell_0$.

Theorem 3.19. *Suppose Assumption 3.18, and that for all $\ell = -1,\dots,\ell_0$, $V_j\colon H_\ell \to H_\ell$ is Lipschitz continuous for $j = 0,\dots,d$. Then, Assumption 3.15 is satisfied.*

Proof. We apply Proposition B.20 to prove that for all $\ell = -1,\dots,\ell_0$, $n \geqslant 2$ and $T > 0$, there exists some constant $K_T > 0$ such that $\mathbb{E}[\psi_\ell^n(x(t,x_0))] \leqslant K_T\psi_\ell^n(x_0)$ for all $x_0 \in H_\ell$ and $t \in [0,T]$. Thus, the result follows from Theorem 3.7. □

We now give some examples such that Assumption 3.18 is satisfied.

Lemma 3.20. *Let $(\tilde{H}, \|\cdot\|_{\tilde{H}})$ be a separable Hilbert space. Assume that the operator $A\colon \operatorname{dom} A \subset \tilde{H} \to \tilde{H}$ generates a strongly continuous semigroup and admits a compact resolvent. Then,* $\operatorname{dom} A^{\ell+1}$ *is compactly embedded in* $\operatorname{dom} A^\ell$, *$\ell \geqslant 0$. Hence, the choice $H_\ell := \operatorname{dom} A^{\ell+1}$, $\ell = -1,\dots,\ell_0$, and $H := H_0$ satisfies Assumption 3.18.*

Here, $\operatorname{dom} A^\ell$ *is endowed with the norm* $\|x\|_{\operatorname{dom} A^\ell} := \left(\sum_{j=0}^{\ell}\|A^j\ell\|_{\tilde{H}}^2\right)^{1/2}$.

Proof. As A has a compact resolvent and generates a strongly continuous semigroup, there exists some $\lambda_0 \in \mathbb{R}$ such that $\lambda_0 - A$ is continuously invertible and $(\lambda_0 - A)^{-1}\colon \tilde{H} \to \tilde{H}$ is compact. Clearly, $(\lambda_0 - A)^{\ell}\colon \operatorname{dom} A^{\ell} \to \tilde{H}$ is continuously invertible.

If a sequence $(x_n)_{n\in\mathbb{N}}$ converges weakly in $\operatorname{dom} A^{\ell+1}$ to some $x \in \operatorname{dom} A^{\ell+1}$, then $(\lambda_0 - A)^{\ell+1}x_n$ converges weakly to $(\lambda_0 - A)^{\ell+1}x$. It follows by the compactness of $(\lambda_0 - A)^{-1}$ that $(\lambda_0 - A)^{\ell}x_n$ converges strongly to $(\lambda_0 - A)^{\ell}x$. This proves that $\operatorname{dom} A^{\ell+1}$ is compactly embedded in $\operatorname{dom} A^{\ell}$. □

Remark 3.21. Under the assumption that the semigroup generated by A consists of compact operators, a condition that is stronger than the existence of a compact resolvent of A (see [84, Theorem 2.3.3]), an argument as in [72, Theorem 2.2] shows directly that $(P_t)_{t\geqslant 0}$ is strongly continuous on $\mathcal{B}^{\psi_0^n}(H_w)$, $n \geqslant 2$. The unboundedness of A still requires us to consider directional derivatives along A only on subspaces of H.

In many situations, in particular for stochastic partial differential equations on unbounded sets, the generator A does not admit a compact resolvent. We give an exemplary construction of Hilbert spaces H_ℓ of functions $(0,\infty) \to \mathbb{R}$, compactly embedded in each other, such that the differential operator $\frac{\mathrm{d}}{\mathrm{d}x}$ satisfies $\frac{\mathrm{d}}{\mathrm{d}x}H_\ell \subset H_{\ell-1}$, $\ell = 0,\dots,\ell_0$. These spaces can be used to embed the Heath-Jarrow-Morton equation of interest rate theory into our setting. This will be performed in Section 4.3.

With $\alpha \in \mathbb{R}$ and $w_\alpha(x) := \exp(\alpha x)$, $x \in \mathbb{R}_+$, we set $\mathrm{L}^2_\alpha(\mathbb{R}_+) := \mathrm{L}^2(\mathbb{R}_+, w_\alpha)$ and $\mathrm{H}^k_\alpha(\mathbb{R}_+) := \mathrm{H}^k(\mathbb{R}_+, w_\alpha)$. Here and in the following, $\mathbb{R}_+ := (0,\infty)$, and for $D \subset \mathbb{R}^N$, the weighted Lebesgue and Sobolev spaces are

$$\mathrm{L}^2(D,w) := \Big\{ f\colon D \to \mathbb{R}\colon \|f\|_{\mathrm{L}^2(D,w)} < \infty \Big\} \quad \text{and} \tag{3.26}$$

$$\mathrm{H}^k(D,w) := \Big\{ f\colon D \to \mathbb{R}\colon \|f\|_{\mathrm{H}^k(D,w)} < \infty \Big\} \tag{3.27}$$

with norms

$$\|f\|_{\mathrm{L}^2(D,w)} := \Big(\int_D f(x)^2 w(x)\mathrm{d}x\Big)^{1/2} \quad \text{and} \tag{3.28}$$

$$\|f\|_{\mathrm{H}^k(D,w)} := \Big(\sum_{j=0}^{k} \|f^{(j)}\|^2_{\mathrm{L}^2(D,w)}\Big)^{1/2}. \tag{3.29}$$

Proposition 3.22. *For every $\alpha > 0$, the space $\mathrm{H}^1(\mathbb{R}_+) \cap \mathrm{L}^2_\alpha(\mathbb{R}_+)$ with norm*

$$\|f\| := \left(\|f\|^2_{\mathrm{H}^1(\mathbb{R}_+)} + \|f\|^2_{\mathrm{L}^2_\alpha(\mathbb{R}_+)} \right)^{1/2} \tag{3.30}$$

is compactly embedded in $\mathrm{L}^2(\mathbb{R}_+)$.

Note that the proof shows that an analogous result holds true for any weight function w with $\lim_{x\to+\infty} w(x) = +\infty$.

Proof. We apply [19, Théorème IV.26]. For any $\tau > 0$,

$$\begin{aligned} \int_{\mathbb{R}_+} |f(x+\tau) - f(x)|^2 \mathrm{d}x &\leqslant \int_{\mathbb{R}_+} \int_0^\tau |f'(x+s)|^2 \mathrm{d}s\mathrm{d}x \\ &= \int_0^\tau \int_{\mathbb{R}_+} |f'(x+s)|^2 \mathrm{d}x\mathrm{d}s \\ &\leqslant \tau \|f\|_{\mathrm{H}^1(\mathbb{R}_+)}, \end{aligned} \tag{3.31}$$

and for any $R > 0$,

$$\begin{aligned} \int_R^\infty |f(x)|^2 \mathrm{d}x &\leqslant \exp(-\alpha R) \int_R^\infty |f(x)|^2 \exp(\alpha x) \mathrm{d}x \\ &\leqslant \exp(-\alpha R) \|f\|_{\mathrm{L}^2_\alpha(\mathbb{R}_+)}. \end{aligned} \tag{3.32}$$

These estimates prove the claim. □

Corollary 3.23. *For any $\alpha, \beta \in \mathbb{R}$ with $\beta > \alpha$ and integer $k \geqslant 0$, $\mathrm{H}^{k+1}_\beta(\mathbb{R}_+)$ is compactly embedded in $\mathrm{H}^k_\alpha(\mathbb{R}_+)$.*

Proof. Assume first $k = 0$. Then, Proposition 3.22 shows that $\mathrm{H}^1_{\beta-\alpha}(\mathbb{R}_+)$ is compactly embedded in $\mathrm{L}^2(\mathbb{R}_+)$.

The mapping $T\colon \mathrm{L}^2(\mathbb{R}_+) \to \mathrm{L}^2_\alpha(\mathbb{R}_+)$, $f \mapsto \exp(-\frac{\alpha}{2}x)f$, is an isometric isomorphism, and $T(\mathrm{H}^1_{\beta-\alpha}(\mathbb{R}_+)) = \mathrm{H}^1_\beta(\mathbb{R}_+)$, where the norms $\|T^{-1}f\|_{\mathrm{H}^1_{\beta-\alpha}(\mathbb{R}_+)}$ and $\|f\|_{\mathrm{H}^1_\beta(\mathbb{R}_+)}$ are equivalent. It follows that $\mathrm{H}^1_\beta(\mathbb{R}_+)$ is compactly embedded in $\mathrm{L}^2_\alpha(\mathbb{R}_+)$. The full result follows by induction. □

Given a strictly increasing sequence $\alpha = \alpha_{-1} < \alpha_0 < \cdots < \alpha_{\ell_0} < \infty$ of real numbers, we define the spaces

$$H_\ell := \left\{ h \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}_+) \colon h' \in \mathrm{H}^{\ell+1}_{\alpha_\ell}(\mathbb{R}_+) \right\}, \quad \ell = -1, \ldots, \ell_0, \tag{3.33}$$

endowed with the norm

$$\|h\|_{H_\ell} := \left(|h(0)|^2 + \|h'\|^2_{\mathrm{H}^{\ell+1}_{\alpha_\ell}(\mathbb{R}_+)} \right)^{1/2}. \tag{3.34}$$

Clearly, the spaces H_ℓ are Hilbert spaces for $\ell = -1, \ldots, \ell_0$. Furthermore, $H_\ell \subset H_{\ell-1}$ and $A := \frac{\mathrm{d}}{\mathrm{d}x} \colon H_\ell \to H_{\ell-1}$ is continuous for $\ell = 0, \ldots, \ell_0$, and $H_{-1} = \mathcal{H}_\alpha$.

Theorem 3.24. *H_ℓ is compactly embedded in $H_{\ell-1}$ for $\ell = 0, \ldots, \ell_0$.*

Proof. We have to prove that if a sequence $(h_n)_{n \in \mathbb{N}}$ in H_ℓ converges weakly to some $h \in H_\ell$, it converges strongly in $H_{\ell-1}$. As evaluation functionals are continuous on H_ℓ, we see that $\lim_{n\to\infty} h_n(0) = h(0)$ follows from weak convergence in H_ℓ. By Corollary 3.23, we see that h'_n converges strongly to h' in $H^\ell_{\alpha_{\ell-1}}(\mathbb{R}_+)$. This proves the result. □

This means that we have constructed spaces H_ℓ, $\ell = -1, \ldots, \ell_0$, such that the Heath-Jarrow-Morton equation of interest rate theory satisfies Assumption 3.18. Note that requiring Assumption 3.13 is actually not untypical in this context and is even weaker than [37, (A1), p. 135]. Thus, all results on covergence of numerical approximations for stochastic partial differential equations we will derive below can be applied to the Heath-Jarrow-Morton equation, and this will be described in detail in Section 4.3.

3.2.2 Taylor expansions

We are now in the situation to derive Taylor expansions for Markov semigroups of stochastic partial differential equations.

Theorem 3.25. *Given Assumptions 3.12, 3.13 and 3.15. Consider the strongly continuous semigroup $(P_t)_{t\geqslant 0}$ on the space $\mathcal{B}^{\psi^n_\ell}((H_\ell)_w)$ with $n \geqslant 4$. Denote its generator by $(\mathcal{G}, \operatorname{dom}\mathcal{G})$.*

Then, $\mathcal{B}_2^{\psi^{(n)}_{\ell-1}}((H_{\ell-1})_w) \subset \operatorname{dom}\mathcal{G}$, and

$$\mathcal{G}f(x) = Df(x)(Ax) + \mathcal{L}_{V_0}f(x) + \frac{1}{2}\sum_{j=1}^{d} \mathcal{L}^2_{V_j} f(x) \tag{3.35}$$

for $f \in \mathcal{B}_2^{\psi^{(n)}_{\ell-1}}((H_{\ell-1})_w)$ and $x \in H_\ell$.

Here, the Lie derivative $\mathcal{L}$ is defined as in Corollary 2.45.

Proof. By the Itô formula given in Proposition B.19, it follows that for $f \in \mathcal{A}(H_{\ell-1})$, we have $f \in \operatorname{dom}\mathcal{G}$, and (3.35) is satisfied.

We extend this representation as follows. Given $f \in \mathcal{B}_2^{\psi_{\ell-1}^{(n)}}((H_{\ell-1})_w)$, choose a sequence $(f_n)_{n\in\mathbb{N}}$ in $\mathcal{A}(H_{\ell-1})$ such that $\lim_{n\to\infty}\|f - f_n\|_{\psi_{\ell-1}^{(n)},2} = 0$. Corollary 2.45 shows that the right hand side of (3.35) defines a continuous linear operator $\tilde{\mathcal{G}}\colon \mathcal{B}_2^{\psi_{\ell-1}^{(n)}}((H_{\ell-1})_w) \to \mathcal{B}^{\psi_\ell^n}((H_\ell)_w)$. Hence, $\lim_{n\to\infty}\|\tilde{\mathcal{G}}f - \mathcal{G}f_n\|_{\psi_\ell^{(n)}} = 0$. The closedness of $\mathcal{G}$ (see Proposition A.5(i)) thus yields that $f \in \operatorname{dom}\mathcal{G}$ and $\mathcal{G}f = \tilde{\mathcal{G}}f$. The claim follows. □

The next result follows directly from Corollary 2.45 and the explicit representation in (3.35).

Corollary 3.26. *Let* $0 \leqslant k \leqslant k_0 - 1$. *Given Assumptions 3.12, 3.13 and 3.15, the infinitesimal generator* $\mathcal{G}$ *satisfies the mapping property*

$$\mathcal{G}\colon \mathcal{B}_{k+2}^{\psi_{\ell-1}^{(n)}}((H_{\ell-1})_w) \to \mathcal{B}_k^{\psi_\ell^{(n)}}((H_\ell)_w), \quad \ell = 1, \dots, \ell_0. \tag{3.36}$$

Induction now yields:

Corollary 3.27. *Let* $j \leqslant \ell \leqslant \ell_0$ *and* $0 \leqslant k \leqslant k_0 - 2j + 1$. *Given Assumptions 3.12, 3.13 and 3.15, the powers of the infinitesimal generator* $\mathcal{G}$ *satisfy*

$$\mathcal{G}^j\colon \mathcal{B}_{k+2j}^{\psi_{\ell-j}^{(n)}}((H_{\ell-j})_w) \to \mathcal{B}_k^{\psi_\ell^{(n)}}((H_\ell)_w). \tag{3.37}$$

They are given explicitly by taking the powers of (3.35).

Proposition A.5(v) yields a Taylor expansion of $P_t f$ for smooth f.

Corollary 3.28. *Let* $f \in \mathcal{B}_{2(k+1)}^{\psi_{\ell-(k+1)}^{(n)}}((H_{\ell-(k+1)})_w)$, $k+1 \leqslant \ell \leqslant \ell_0$, *and* $2(k+1) \leqslant k_0 - 1$. *Given Assumptions 3.12, 3.13 and 3.15,*

$$P_t f = \sum_{j=0}^{k} \frac{t^j}{j!}\mathcal{G}^j f + t^{k+1} R_{t,k} f, \tag{3.38}$$

where the linear operator $R_{t,k}\colon \mathcal{B}_{2(k+1)}^{\psi_{\ell-(k+1)}^{(n)}}((H_{\ell-(k+1)})_w) \to \mathcal{B}^{\psi_\ell^n}((H_\ell)_w)$ *satisfies*

$$\|R_{t,k}f\|_{\psi_\ell^n} \leqslant C_T \|f\|_{\psi_{\ell-(k+1)}^{(n)},2(k+1)} \quad \text{for } t \in [0,T] \tag{3.39}$$

for a constant $C_T > 0$ *independent of* f.

3.3 Smoothing effects with analytic semigroups

Suppose that Assumptions 3.12, 3.13 and 3.15 are satisfied, where $H_\ell = \operatorname{dom} A^\ell$, $\ell = 0, \dots, \ell_0$, and that $\operatorname{dom} A$ is compactly embedded in H. Furthermore, suppose that the operator A generates an analytic semigroup, and, without loss of generality, that 0 is in the resolvent set of A. Consult Section A.2 for a short overview of analytic semigroups and fractional powers of operators. By Proposition A.8(iv), there exists $\delta > 0$ such that for all $\gamma > 0$, we can find some $M_\gamma > 0$ with

$$\|(-A)^\gamma S_t x\|_H \leqslant M_\gamma t^{-\gamma} \exp(-\delta t)\|x\|_H \quad \text{for all } x \in H \text{ and } t > 0, \tag{3.40}$$

where $(-A)^\gamma$ denotes the γ fractional power of $-A$ and $S_t := \exp(tA)$ denotes the semigroup generated by A. We want to use this property to derive smoothing effects of the mapping $x_0 \mapsto x(t, x_0)$. More precisely, under appropriate assumptions on the coefficients V_j, we want to prove that $x(t, x_0) \in \operatorname{dom} A^k$ for $x_0 \in X$.

To prove a regularising effect, let us consider the mild formulation of the stochastic partial differential equation. We have that

$$x(t, x_0) = S_t x_0 + \int_0^t S_{t-s}\tilde{V}_0(x(s, x_0))\mathrm{d}s + \sum_{j=1}^d \int_0^t S_{t-s}V_j(x(s, x_0))\mathrm{d}B_s^j, \tag{3.41}$$

$\tilde{V}_0(x) = V_0(x) + \frac{1}{2}\sum_{j=1}^d DV_j(x)V_j(x)$ being the Itô drift.

As stated above, $S_t x_0 \in \operatorname{dom} A$, and for every $\gamma > 0$, $\|A^\gamma S_t x_0\|_H \leqslant M_\gamma t^{-\gamma}\|x_0\|_H$ for $t \in [0, T]$. To estimate the other terms, we use the following auxiliary results. Denote by $\mathrm{L}^p([0, T]; H)$ the space of measurable functions $f\colon [0, T] \to H$ with $\int_0^T \|f(s)\|_H^p \mathrm{d}s < \infty$, endowed with the norm

$$\|f\|_{\mathrm{L}^p([0,T];H)} := \Big(\int_0^T \|f(s)\|_H^p \mathrm{d}s\Big)^{1/p}. \tag{3.42}$$

Proposition 3.29. *Suppose that $V\colon H \to H$ is Lipschitz continuous. Let $f \in \mathrm{L}^p([0, T]; H)$ with $p > 5/4$, and set $F(t) := \int_0^t S_{t-s}V(f(s))\mathrm{d}s$, $t \in [0, T]$. Then, there exists $C > 0$ such that*

$$\sup_{t\in[0,T]} \|(-A)^{1/5}F(t)\|_H \leqslant C(1 + \|f\|_{\mathrm{L}^p([0,T];H)}). \tag{3.43}$$

Proof. We calculate

$$\|(-A)^{1/5}F(t)\|_H \leqslant C\int_0^t (t-s)^{-1/5}\|V(f(s))\|_H \mathrm{d}s. \tag{3.44}$$

Linear boundedness of V and the Hölder inequality yield the claim. □

Proposition 3.30. *Suppose $V\colon H \to H$ is Lipschitz continuous. Given a one-dimensional Brownian motion $(B_t)_{t\geqslant 0}$ on $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geqslant 0})$. Let $f\colon \Omega\times[0,T] \to H$ be progressively measurable with $\int_0^T \mathbb{E}[\|f(s)\|_H^p]\mathrm{d}s < \infty$, $p \geqslant 4$, and set $F(t) := \int_0^t S_{t-s}V(f(s))\mathrm{d}B_s$, $t \in [0,T]$. Then, there exists $C > 0$ such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|(-A)^{1/5}F(t)\|_H^p\right] \leqslant C\left(1 + \int_0^t \mathbb{E}[\|f(s)\|_H^p]\mathrm{d}s\right). \tag{3.45}$$

Proof. By the Burkholder-Davis-Gundy inequality [28, Theorem 5.2.4],

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]} \|(-A)^{1/5}F(s)\|_H^p\right] &\leqslant C\mathbb{E}\left[\Big(\int_0^t \|(-A)^{1/5}S_{t-s}V(f(s))\|_H^2\mathrm{d}s\Big)^{p/2}\right]\\ &\leqslant C\mathbb{E}\left[\Big(\int_0^t (t-s)^{-2/5}\|V(f(s))\|_H^2\mathrm{d}s\Big)^{p/2}\right].\end{aligned} \tag{3.46}$$

The Cauchy-Schwarz inequality yields

$$\int_0^t (t-s)^{-2/5}\|V(f(s))\|_H^2\mathrm{d}s \leqslant C\Big(\int_0^t \|V(f(s))\|^4\mathrm{d}s\Big)^{1/2}. \tag{3.47}$$

Linear boundedness of V and the Jensen inequality prove the claim. □

Remark 3.31. The restriction to $p \geqslant 4$ is arbitary. Instead, every $p > 2$ is possible. Note, however, that the case $p = 2$ is not admissible.

Corollary 3.32. *Suppose the assumptions made at the beginning of this section. Then, for every $p \geqslant 4$, there exists $C > 0$ such that for $x_0 \in H$ and $t \in (0,T]$,*

$$\mathbb{E}[\|(-A)^{1/5}x(t,x_0)\|_H^p] \leqslant Ct^{-1/5}(1 + \|x_0\|_H^p). \tag{3.48}$$

In particular, $\mathbb{E}[\psi_{1/5}^n(x(t,x_0))] \leqslant Ct^{-1/5}\psi_0^n(x_0)$ for $n \geqslant 4$, $t \in (0,T]$ and $x_0 \in H$. Here, we have set $\psi_{1/5}^s(x) := (1 + \|x\|_{\mathrm{dom}(-A)^{1/5}}^2)^{s/2}$ for $x \in \mathrm{dom}(-A)^{1/5}$, where $\|x\|_{\mathrm{dom}(-A)^{1/5}} := (\|x\|_H^2 + \|(-A)^{1/5}x\|_H^2)^{1/2}$.

Proof. By the above results, we see that

$$\mathbb{E}[\|(-A)^{1/5}x(t,x_0)\|_H^p] \leqslant Ct^{-1/5}\|x_0\|_H^p + C\int_0^t \mathbb{E}[\|x(s,x_0)\|_H^p]\mathrm{d}s. \tag{3.49}$$

As $\mathbb{E}[\|x(s,x_0)\|_H^p] \leqslant C(1+\|x_0\|_H^p)$ by Proposition B.20, the result follows. □

Theorem 3.33. *The operator* $P_t\colon \mathcal{B}^{\psi_{1/5}^n}((\mathrm{dom}(-A)^{1/5})_w) \to \mathcal{B}^{\psi_0^n}(H_w)$ *is well-defined for* $n \geqslant 4$ *and* $t \in (0,T]$*, and satisfies*

$$\|P_t f\|_{\mathcal{B}^{\psi_0^n}(H_w)} \leqslant Ct^{-1/5}\|f\|_{\mathcal{B}^{\psi_{1/5}^n}((\mathrm{dom}(-A)^{1/5})_w)} \tag{3.50}$$

for $f \in \mathcal{B}^{\psi_{1/5}^n}((\mathrm{dom}(-A)^{1/5})_w)$ *and* $t \in (0,T]$.

Proof. The norm bound follows immediately from Corollary 3.32. To see that P_t satisfies the given mapping property, note that P_t leaves $\mathcal{B}^{\psi_0^n}(H_w)$ invariant. □

The following result shows that we are actually in the same situation on $\mathrm{dom}(-A)^{1/5}$ as we were on H.

Lemma 3.34. *Under the assumptions given at the beginning of this section,* $\tilde{V}_0$, $V_j\colon \mathrm{dom}(-A)^\gamma \to \mathrm{dom}(-A)^\gamma$, $j=1,\dots,d$, $\gamma \in (0,1)$, *are well-defined and Lipschitz continuous mappings.*

Proof. By [12, p. 170, Theorem 6.1], the domain of the fractional power $(-A)^\alpha$ agrees with a certain interpolation space,

$$\mathrm{dom}(-A)^\gamma = (H, \mathrm{dom}\, A)_{\gamma,2}, \quad \gamma \in (0,1). \tag{3.51}$$

Hence, Assumption 3.13 together with [21, Proposition 3] yields the claim. □

Corollary 3.35. *The operator* $P_t\colon \mathcal{B}^{\psi_\ell^n}((\mathrm{dom}\, A^\ell)_w) \to \mathcal{B}^{\psi_0^n}(H_w)$ *is well-defined for* $n \geqslant 4$ *and* $t \in (0,T]$*, and satisfies*

$$\|P_t f\|_{\mathcal{B}^{\psi_0^n}(H_w)} \leqslant Ct^{-\ell}\|f\|_{\mathcal{B}^{\psi_\ell^n}((\mathrm{dom}\, A^\ell)_w)} \tag{3.52}$$

for $f \in \mathcal{B}^{\psi_\ell^n}((\mathrm{dom}\, A^\ell)_w)$ *and* $t \in (0,T]$.

Proof. Lemma 3.34 shows that we can consider the equation (3.41) also on the space $\mathrm{dom}(-A)^{1/5}$, and that we are in the setting of Theorem 3.33. An induction yields the claim. □

Note that we can actually prove a similar result for spaces of differentiable functions due to Theorem 2.31, as well. Hence, we can obtain optimal convergence estimates for functions defined on $\operatorname{dom} A^\ell$ and initial values in $\operatorname{dom} A^\ell$ by the use of graded time steps, see, e.g., [39, 101, 102].

Remark 3.36. Assuming that the vector fields $\tilde{V}_0$, V_j, $j = 1, \dots, d$, themselves have a smoothing effect, such a result can be obtained in a less technical manner. To be precise, suppose that $\tilde{V}_0, V_j \colon H \to \operatorname{dom} A^k$, $j = 1, \dots, d$, with

$$\|\tilde{V}_0(x)\|_{\operatorname{dom} A^k} + \sum_{j=1}^{d} \|V_j(x)\|_{\operatorname{dom} A^k} \leqslant C\,(1 + \|x\|_H) \quad \text{for all } x \in H. \tag{3.53}$$

That is, we only require linear growth of the mappings $H \to \operatorname{dom} A^k$, but not Lipschitz continuity. Such an assumption is often satisfied in applications in mathematical finance, see [35]. Then, we easily see that $P_t \colon \mathcal{B}^{\psi^n_{\ell+k}}((\operatorname{dom} A^{\ell+k})_w) \to \mathcal{B}^{\psi^n_\ell}((\operatorname{dom} A^\ell)_w)$ is well-defined and the bound

$$\|P_t\|_{L(\mathcal{B}^{\psi^n_{\ell+k}}((\operatorname{dom} A^{\ell+k})_w), \mathcal{B}^{\psi^n_\ell}((\operatorname{dom} A^\ell)_w))} \leqslant C t^{-k} \tag{3.54}$$

holds true, without having to resort to interpolation theory.

Chapter 4

Splitting Schemes For Stochastic (Partial) Differential Equations

The aim of this chapter is to show how splitting methods can be applied to stochastic partial differential equations under realistic conditions, improving significantly on the results from [6]. The results of Chapter 3 allow us to use the theory of exponential splittings to prove optimal rates of convergence of splitting schemes for stochastic (partial) differential equations with linearly growing characteristics and for sets of functions with controlled growth.

In Section 4.1, we apply the results of Chapter 3 to the derivation of estimates of the rate of convergence of splitting schemes. Under the assumptions made in Section 3.2, we obtain optimal estimates for sufficiently smooth functions. Section 4.2 contains an analysis of extrapolation schemes based on the symmetrically weighted sequential splitting. For smooth functions, an asymptotic expansion in n^{-2} is obtained, allowing fast error reduction in this case. Section 4.3 is concerned with an application of the theory to the Heath-Jarrow-Morton equation of interest rate theory. We are able to calibrate the model to given data, and price a swaption using the calibrated model.

4.1 Error estimates

Assume the setting of Section 3.2, i.e., we suppose that Assumptions 3.12, 3.13 and 3.15 are satisfied. We consider two splitting schemes for (3.16), the Lie-Trotter scheme (i.e., the Euler scheme in a geometric integrator version), and the Ninomiya-Victoir scheme. We derive convergence estimates for both splittings.

We define $z^0(t,x)$, $z^j(t,x)$, $j=1,\ldots,d$ as the solutions of

$$\frac{\mathrm{d}}{\mathrm{d}t}z^0(t,x_0) = Az^0(t,x_0) + V_0(z^0(t,x_0)), \tag{4.1}$$

$$\mathrm{d}z^j(t,x_0) = V_j(z^j(t,x_0)) \circ \mathrm{d}B^j_t, \quad j=1,\ldots,d. \tag{4.2}$$

This exact manner of splitting up the stochastic partial differential equation is not mandatory in our setting as it is in approaches guided by Lyons-Victoir cubature [70, 80, 6], but it is very helpful – for $j=1,\ldots,d$, the processes $z^j(t,x)$ are given through evaluation of the flow of the vector field V_j at random times given by W^j_t: $z^j(t,x) = \mathrm{Fl}^{V_j}_{B^j_t}(x)$, where $\mathrm{Fl}^{V_j}_s$ denotes the flow defined by V_j. Note that only the equation for $z^0(t,x_0)$ contains the unbounded operator A, but that this equation is a deterministic evolution equation on H.

By Theorem 3.16, $\|P^j_t\|_{L(\mathcal{B}^{\psi^n_\ell}((H_\ell)_w))} \leqslant \exp(\omega_j t)$ with some constants $\omega_j \in \mathbb{R}$, $j=0,\ldots,d$.

Remark 4.1. For the split semigroups, we can also prove pseudocontractivity directly without invoking the Szőkefalvi-Nagy theorem. Indeed, for P^j_t, $j=1,\ldots,d$, we can apply Itô's formula. For P^0_t, we use the mild formulation

$$z^0(t,x_0) = \exp(tA)x_0 + \int_0^t \exp((t-s)A)V_0(z^0(s,x_0))\mathrm{d}s, \tag{4.3}$$

where $\exp(tA)$ denotes the semigroup generated by A at time t. As A is pseudocontractive, we can assume without loss of generality that A is contractive by adding the growth to V_0. Denoting the linear growth bound of V_0 in H_ℓ by L, $\|V_0(x)\|_{H_\ell} \leqslant L(1+\|x\|_{H_\ell})$, this yields

$$\begin{aligned}\|z^0(t,x_0)\|_{H_\ell} &\leqslant \|x_0\|_{H_\ell} + \int_0^t \|V_0(z^0(s,x_0))\|_{H_\ell}\mathrm{d}s \\ &\leqslant \|x_0\|_{H_\ell} + \int_0^t L(1+\|z^0(s,x_0)\|_{H_\ell})\mathrm{d}s.\end{aligned} \tag{4.4}$$

From the Gronwall inequality,

$$\|z^0(t,x_0)\|_{H_\ell} \leqslant (\|x_0\|_{H_\ell} + Lt)\exp(Lt). \tag{4.5}$$

Thus,

$$\begin{aligned}1+\|z^0(t,x_0)\|^2_{H_\ell} &\leqslant 1 + (\|x_0\|_{H_\ell} + Lt)^2\exp(2Lt) \\ &\leqslant (1+\|x_0\|_{H_\ell})^2(1+L^2t^2)\exp(2Lt),\end{aligned} \tag{4.6}$$

which proves the bound

$$\psi_\ell^n(z^0(t,x_0)) \leqslant \exp(\omega t)\psi_\ell^n(x_0) \tag{4.7}$$

for $t \geqslant 0$ with $\omega = 4Ln$.

This, together with the fact that the split semigroups approximate P_t strongly on $\mathcal{B}^\psi((H_\ell)_w)$ (see Corollary 4.13), yields an alternative proof of Theorem 3.16.

We now define two well-known splitting schemes and prove optimal rates of convergence on spaces of sufficiently smooth functions in our general setting.

Definition 4.2 (Lie-Trotter splitting)**.** One step of the Lie-Trotter splitting reads

$$Q^{\mathrm{LT}}_{(\Delta t)} := P^0_{\Delta t} P^1_{\Delta t} \cdots P^d_{\Delta t}, \tag{4.8}$$

which is a geometric integrator version of the well-known Euler scheme.

Definition 4.3 (Ninomiya-Victoir splitting)**.** One step of the Ninomiya-Victoir splitting reads

$$Q^{\mathrm{NV}}_{(\Delta t)} := \frac{1}{2} P^0_{\Delta t/2} \left(P^1_{\Delta t} \cdots P^d_{\Delta t} + P^d_{\Delta t} \cdots P^1_{\Delta t} \right) P^0_{\Delta t/2}, \tag{4.9}$$

which should in theory improve the weak rate of convergence of the Lie-Trotter scheme by one order.

The Ninomiya-Victoir splitting can be seen as a variant of the classical Strang splitting, generalised to a sum of more than 2 generators.

Let $\mathcal{G}_j$ with domain $\operatorname{dom}\mathcal{G}_j$ be the infinitesimal generator of $(P^j_t)_{t\geqslant 0}$, where $(P^j_t)_{t\geqslant 0}$ is considered on $\mathcal{B}^{\psi^n_\ell}((H_\ell)_w)$ with some fixed $0 \leqslant \ell \leqslant \ell_0$. The function spaces defined below will be fundamental for proving convergence estimates.

Definition 4.4. Let $p \geqslant 1$ be given. We say that $f \in \mathcal{M}^p_T$ if and only if $f \in \mathcal{B}^{\psi^n_{\ell_0}}((H_{\ell_0})_w)$, $P_t f \in \operatorname{dom}\mathcal{G}^p \cap \bigcap_{j_1,\dots,j_p=0}^d \operatorname{dom}\mathcal{G}_{j_1}\dots\mathcal{G}_{j_p}$ for $t \in [0,T]$,

$$C_f := \sup_{\substack{t\in[0,T]\\ j_1,\dots,j_p=0,\dots,d}} \|\mathcal{G}_{j_1}\cdots\mathcal{G}_{j_p} P_t f\|_{\psi^n_{\ell_0}} < \infty \quad \text{and} \tag{4.10}$$

$$\mathcal{G}^i P_t f = \left(\sum_{j=0}^d \mathcal{G}_j\right)^i P_t f, \quad i = 1,\dots,p. \tag{4.11}$$

Proposition 4.5. *Let $Q_{\Delta t}$ be a splitting for $P_{\Delta t}$ of classical order p. For $f \in \mathcal{M}_T^{p+1}$, the splitting converges of optimal order, that is, with a constant C_f independent of $n \in \mathbb{N}$ and $\Delta t > 0$, we have that for $n\Delta t \leqslant T$,*

$$\|P_{n\Delta t} f - Q^n_{(\Delta t)} f\|_{\psi^n_{\ell_0}} \leqslant C_f (\Delta t)^p. \tag{4.12}$$

Proof. Set $g := P_t f \in \operatorname{dom} \mathcal{G} \cap \bigcap_{j=0}^d \mathcal{G}_j$. The results in [49, Proof of Theorem 3.4, Section 4.1, Section 4.4] prove existence of a family of linear operators $T_{(\Delta t)} \colon \mathcal{B}^{\psi^n_{\ell_0}}((H_{\ell_0})_w) \to \mathcal{B}^{\psi^n_{\ell_0}}((H_{\ell_0})_w)$ that are uniformly bounded, i.e.,

$$\sup_{\Delta t \in [0,\varepsilon]} \|T_{(\Delta t)}\|_{L(\mathcal{B}^{\psi^n_{\ell_0}}((H_{\ell_0})_w))} \leqslant C_\varepsilon < \infty \quad \text{for some } \varepsilon > 0, \tag{4.13}$$

such that the difference of the Taylor expansions of $P_{\Delta t} g$ and $Q_{(\Delta t)} g$ of order p is given by

$$P_{\Delta t} g - Q_{(\Delta t)} g = (\Delta t)^{p+1} T_{(\Delta t)} \mathcal{E}_{p+1} g, \tag{4.14}$$

where $\mathcal{E}_{p+1}$ is a linear combination of the operators $\mathcal{G}_{j_1} \cdots \mathcal{G}_{j_{p+1}}$, $j_1, \ldots, j_{p+1} = 0, \ldots, d$. Here, we apply that by assumption, $\mathcal{G}^{p+1}$ is itself a linear combination of these operators when applied to g. Thus,

$$\|P_{\Delta t} g - Q_{(\Delta t)} g\|_{\psi^n_{\ell_0}} \leqslant C_f (\Delta t)^{p+1} \|T_{(\Delta t)}\|_{L(\mathcal{B}^{\psi^n_{\ell_0}}((H_{\ell_0})_w))} \leqslant C_f (\Delta t)^{p+1}. \tag{4.15}$$

It follows that

$$\begin{aligned} \|P_{n\Delta t} f - Q^n_{(\Delta t)} f\|_{\psi^n_{\ell_0}} &\leqslant C_f (\Delta t)^{p+1} \sum_{i=1}^n \|Q^j_{(\Delta t)}\|_{L(\mathcal{B}^{\psi^n_{\ell_0}}((H_{\ell_0})_w))} \\ &\leqslant C_f (\Delta t)^p, \end{aligned} \tag{4.16}$$

which proves the result. □

For the Lie-Trotter scheme, we set $\mathcal{M}_T^{\mathrm{LT}} := \mathcal{M}_T^2$, and for the Ninomiya-Victoir scheme, $\mathcal{M}_T^{\mathrm{NV}} := \mathcal{M}_T^3$. The following results are now an easy consequence of Proposition 4.5.

Corollary 4.6. *For $f \in \mathcal{M}_T^{\mathrm{LT}}$ there exists a constant $C_f > 0$ such that for all $t \in [0, T]$ and $m \in \mathbb{N}$,*

$$\|P_t f - (Q^{\mathrm{LT}}_{(t/m)})^m f\|_{\psi^n_{\ell_0}} \leqslant C_f m^{-1}. \tag{4.17}$$

Hence, for $f \in \mathcal{M}_T^{\mathrm{LT}}$, the Euler splitting scheme converges of optimal order.

Corollary 4.7. *For $f \in \mathcal{M}_T^{\mathrm{NV}}$ there exists a constant $C_f > 0$ such that for all $t \in [0, T]$ and $m \in \mathbb{N}$,*

$$\|P_t f - (Q^{\mathrm{NV}}_{(t/m)})^m f\|_{\psi^n_{\ell_0}} \leqslant C_f m^{-2}. \tag{4.18}$$

Hence, for $f \in \mathcal{M}_T^{\mathrm{NV}}$, the Ninomiya-Victoir splitting scheme converges of optimal order.

Remark 4.8. It is possible to consider other splittings than the Lie-Trotter or the Ninomiya-Victoir schemes. It is, however, not possible to obtain higher rates of convergence due to inherent limits of generic splitting schemes with positive coefficients (see [15]), and positivity of coefficients is mandatory in the probabilistic setting under concern. To obtain higher order methods, we can either resort to extrapolation, see Section 4.2, or to cubature methods, see Chapter 5.

We derive easy conditions guaranteeing $f \in \mathcal{M}_T^{\mathrm{NV}}$.

Lemma 4.9. *Let $0 \leqslant k \leqslant k_0 - 1$ and $1 \leqslant \ell \leqslant \ell_0$. Then,*

$$\sum_{j=0}^{d} \mathcal{G}_j f = \mathcal{G} f \quad \text{for all } f \in \mathcal{B}^{\psi^n_\ell}_{k+2}((H_\ell)_w). \tag{4.19}$$

Proof. This follows directly from Corollary 3.26 applied to $\mathcal{G}_j$ and $\mathcal{G}$. □

Next, we prove that P_t leaves smooth functions invariant.

Lemma 4.10. *Let $1 \leqslant \ell \leqslant \ell_0$, $n \geqslant 2$ and $0 \leqslant k \leqslant k_0$. Then, $P_t \mathcal{B}^{\psi^{(n)}_\ell}_k((H_\ell)_w) \subset \mathcal{B}^{\psi^{(n)}_\ell}_k((H_\ell)_w)$, and $\sup_{t \in [0,T]} \|P_t f\|_{\psi^{(n)}_\ell, k} \leqslant K_T \|f\|_{\psi^{(n)}_\ell, k}$ with some constant K_T independent of f.*

Proof. Proposition B.21 yields: for all $T > 0$ and $p \in [1, \infty)$, there exists $C_{p,T} > 0$ such that

$$\begin{aligned} &\mathbb{E}[\|D^j_{x_0} x(t, x_0)\|^p_{L_j(H_\ell; H_\ell)}] \leqslant C_T \\ &\qquad \text{for all } x \in H_\ell,\ \ell = 0, \dots, \ell_0,\ j = 1, \cdots, k \text{ and } t \in [0, T]. \end{aligned} \tag{4.20}$$

Moreover, the mappings $x_0 \mapsto D^j_{x_0} x(t, x_0)$ are almost surely norm continuous. It follows that $P_t(\mathrm{C}^k_b(H_\ell)) \subset \mathrm{C}^k_b(H_\ell)$ for all $t \in [0, T]$ and $\ell = 0, \dots, \ell_0$ (note that we can apply [109, Proposition 4.8c)] instead of [29, Proposition 7.4.1] to obtain this sharper result). Furthermore, for $f \in C_b^k(H_\ell)$, the Cauchy-Schwarz inequality yields that for x_0, $x_1 \in H_\ell$ such that $\|x_1\|_{H_\ell} \leqslant 1$,

$$\begin{aligned} |DP_t f(x_0)(x_1)| &\leqslant \mathbb{E}[\|Dx(t,x_0)\|_{L(H_\ell;H_\ell)}\|Df(x(t,x_0))\|_{L(H_\ell;\mathbb{R})}] \\ &\leqslant |f|_{\psi_\ell^{n-1},1}\mathbb{E}[\|Dx(t,x_0)\|_{L(H_\ell;H_\ell)}\psi_\ell^{n-1}(x(t,x_0))] \\ &\leqslant |f|_{\psi_\ell^{n-1},1}\mathbb{E}[\|Dx(t,x_0)\|^2_{L(H_\ell;H_\ell)}]^{1/2}\mathbb{E}[\psi_\ell^{n-1}(x(t,x_0))^2]^{1/2} \\ &\leqslant C_T|f|_{\psi_\ell^{n-1},1}\psi_\ell^{n-1}(x), \end{aligned} \tag{4.21}$$

where we apply that $\psi_\ell^n(x)^2 = \psi_\ell^{2n}(x)$ and that by Assumption 3.15,

$$\mathbb{E}[\psi_\ell^{2(n-1)}(x(t,x_0))] \leqslant C_T\psi_\ell^{2(n-1)}(x_0). \tag{4.22}$$

Thus,

$$\|P_t f\|_{\psi_\ell^{(n)},1} \leqslant C_T\|f\|_{\psi_\ell^{(n)},1} \quad \text{for all } f \in C_b^1(H_{\ell-1}). \tag{4.23}$$

A similar argument applies for higher derivatives. Theorem 2.31 now shows $P_t(\mathcal{B}_k^{\psi_\ell^{(n)}}((H_\ell)_w)) \subset \mathcal{B}_k^{\psi_\ell^{(n)}}((H_\ell)_w)$. This proves the claim. □

Theorem 4.11. *Suppose $\ell_0 \geqslant 4$ and $k_0 \geqslant 6$. Choose $1 \leqslant \ell \leqslant \ell_0 - 4$. Then, $\mathcal{B}_6^{\psi_\ell^{(n)}}((H_\ell)_w) \subset \mathcal{M}_T^{\mathrm{NV}}$. In particular, $C_b^6(H) \subset \mathcal{M}_T^{\mathrm{NV}}$.*

Proof. By Lemma 4.10, $\|P_t f\|_{\psi_\ell^{(n)},6} \leqslant K_T\|f\|_{\psi_\ell^{(n)},6} < \infty$ for all $t \in [0,T]$. The first claim follows by Corollary 3.27 together with iterating Lemma 4.9.

The second claim follows from Theorem 2.31. □

The following theorem follows analogously.

Theorem 4.12. *Suppose $\ell_0 \geqslant 3$ and $k_0 \geqslant 4$. Choose $1 \leqslant \ell \leqslant \ell_0 - 3$. Then, $\mathcal{B}_4^{\psi_\ell^{(n)}}((H_\ell)_w) \subset \mathcal{M}_T^{\mathrm{LT}}$. In particular, $C_b^4(H) \subset \mathcal{M}_T^{\mathrm{LT}}$.*

Corollary 4.13. *Let $f \in \mathcal{B}^{\psi_{\ell_0}^n}((H_{\ell_0})_w)$. Then, for any $t > 0$,*

$$\lim_{n\to\infty}\|P_t f - (Q_{(t/n)}^{\mathrm{LT}})^n f\|_{\psi_{\ell_0}^n} = \lim_{n\to\infty}\|P_t f - (Q_{(t/n)}^{\mathrm{NV}})^n f\|_{\psi_{\ell_0}^n} = 0, \tag{4.24}$$

that is, the Lie-Trotter and Ninomiya-Victoir splittings converge strongly on the space $\mathcal{B}^{\psi_{\ell_0}^n}((H_{\ell_0})_w)$.

Proof. This follows from the density of bounded, smooth, cylindrical functions in $\mathcal{B}^{\psi^n_{\ell_0}}((H_{\ell_0})_w)$, see Theorem 2.18. □

Example 4.14. Assume that $V_0 \equiv 0$ and that the V_j are constant, $j = 1, \dots, d$. This includes, in particular, stochastic heat and wave equations on bounded domains with additive noise. It is easy to see that if $A\colon \operatorname{dom} A \to X$ admits a compact resolvent, we are in the situation described above, and the Ninomiya-Victoir splitting converges of optimal order.

Example 4.15. Finite-dimensional problems with Lipschitz-continuous coefficients are also included in this setting. Here, A can be chosen to be zero, and the embedding is trivially compact due to the local compactness of finite-dimensional spaces.

4.2 Extrapolation

As noted above, the order that can be attained by splitting schemes is limited to two in our setting. Therefore, it is interesting to ask whether there is an alternative approach to constructing methods of higher order. While using extrapolation is well known, Gyöngy and Krylov provide in [43, 44] an approach to this problem which is well adapted to our setting. We shall consider their approach from the perspective of strongly continuous semigroups and obtain error estimates for the extrapolated symmetrically weighted sequential splitting.

To this end, we shall use the following setup.

Assumption 4.16. $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ is a separable Banach space. For $j = 0, \dots, d$, $\mathcal{G}_j\colon \operatorname{dom} \mathcal{G}_j \subset \mathcal{V} \to \mathcal{V}$ are infinitesimal generators of strongly continuous semigroups of pseudocontractions $P^j_t := \exp(t\mathcal{G}_j)$. There exists a sequence of Banach spaces $(\mathcal{V}_\ell, \|\cdot\|_{\mathcal{V}_\ell})$ with the following properties:

(i) $\mathcal{V}_0 = \mathcal{V}$ with coinciding norms.

(ii) $\mathcal{V}_{\ell+1}$ is a dense subset of $\mathcal{V}_\ell$, $\ell \geqslant 0$.

(iii) $\mathcal{V}_1 \subset \operatorname{dom} \mathcal{G}_j$, $j = 0, \dots, d$, and $\mathcal{G}_j(\mathcal{V}_{\ell+1}) \subset \mathcal{V}_\ell$, $\ell \geqslant 0$, $j = 0, \dots, d$.

(iv) $P^j_t(\mathcal{V}_\ell) \subset \mathcal{V}_\ell$ and $P^j_t\colon \mathcal{V}_\ell \to \mathcal{V}_\ell$ is continuous, $\ell \geqslant 0$, $t \geqslant 0$, $j = 0, \dots, d$.

Finally, the operator $\tilde{\mathcal{G}} := \sum_{j=0}^d \mathcal{G}_j$, defined on $\mathcal{V}_1$, is closable, and its closure generates a strongly continuous semigroup $(P_t)_{t\geqslant 0}$ of pseudocontractions.

Assumption 4.16 implies that $\mathcal{V}_\ell$ is a core of $\mathcal{G}_j$, $\ell \geqslant 0$, $j = 0, \dots, d$ (see Proposition A.5(vi)), and $\mathcal{V}_1$ is a core for $\mathcal{G}$ by assumption. Setting

$$\overrightarrow{Q}_{(\Delta t)} := P^0_{\Delta t} \cdots P^d_{\Delta t} \quad \text{and} \quad \overleftarrow{Q}_{(\Delta t)} := P^d_{\Delta t} \cdots P^0_{\Delta t}, \tag{4.25}$$

the Chernoff product formula [34, Theorem III.5.2] shows that

$$\lim_{n\to\infty} \| P_t f - \overrightarrow{Q}^{\,n}_{(t/n)} f \|_{\mathcal{V}} = 0 \quad \text{for all } f \in \mathcal{V}, \tag{4.26}$$

the limit being uniform for $t \in [0, T]$, $T > 0$ arbitrary, and similarly for $\overleftarrow{Q}_{(\Delta t)}$. Assume finally that $\mathcal{G}(\mathcal{V}_{\ell+1}) \subset B_\ell$, $P_t(\mathcal{V}_\ell) \subset \mathcal{V}_\ell$ and $P_t \colon \mathcal{V}_\ell \to \mathcal{V}_\ell$ is continuous, $\ell \geqslant 0$, $t \geqslant 0$.

Proposition 4.17. *Given Assumption 4.16, there exist elements $\overrightarrow{f}_k, \overleftarrow{f}_k \in \mathcal{V}_0$, $k = 1, \dots, m$, such that for every $f \in \mathcal{V}_{2(m+1)}$,*

$$\overrightarrow{Q}^{\,n}_{(T/n)} f - P_T f = \sum_{k=1}^{m} n^{-k} \overrightarrow{f}_k + n^{-m-1} \overrightarrow{r}_{m,n}, \quad \textit{and} \tag{4.27}$$

$$\overleftarrow{Q}^{\,n}_{(T/n)} f - P_T f = \sum_{k=1}^{m} n^{-k} \overleftarrow{f}_k + n^{-m-1} \overleftarrow{r}_{m,n}, \tag{4.28}$$

where $(\overrightarrow{r}_{m,n})_{n\in\mathbb{N}}$, $(\overleftarrow{r}_{m,n})_{n\in\mathbb{N}}$ are families of elements of $\mathcal{V}$ such that $\|\overrightarrow{r}_{m,n}\|_{\mathcal{V}}$, $\|\overleftarrow{r}_{m,n}\|_{\mathcal{V}} \leqslant C_m$ with some constant $C_m \geqslant 0$ independent of n.

Proof. This follows in a straightforward manner from the results in [44, Section 5], as the assumptions there are clearly satisfied for generators of strongly continuous semigroups if the spaces $\mathcal{V}_\ell$ are invariant with respect to these semigroups. □

In [82], it is used that so-called Fujiwara splittings have the advantage that we not only obtain an asymptotic expansion in n^{-1}, but actually in n^{-2}. This means that we gain two orders of convergence per extrapolation step, not only one. We now prove that this holds true for the symmetrically weighted sequential splitting, as well.

Theorem 4.18. *Given Assumption 4.16, we have that*

$$\overrightarrow{f}_{2\kappa+1} + \overleftarrow{f}_{2\kappa+1} = 0 \quad \textit{for } \kappa = 0, \dots, \left\lfloor \frac{m-1}{2} \right\rfloor. \tag{4.29}$$

Thus, setting $g_\kappa := \frac{1}{2}(\overrightarrow{f}_{2\kappa} + \overleftarrow{f}_{2\kappa})$*,* $\kappa \in \mathbb{N}$*, we obtain*

$$\frac{1}{2}\left(\overrightarrow{Q}^n_{(T/n)} f + \overleftarrow{Q}^n_{(T/n)} f\right) - P_T f = \sum_{\kappa=1}^{\lfloor \frac{m-1}{2} \rfloor} n^{-2\kappa} g_\kappa + n^{-m-1} r_{m,n}, \tag{4.30}$$

where $r_{m,n} = \frac{1}{2}(\overrightarrow{r}_{m,n} + \overleftarrow{r}_{m,n})$*, and is thus bounded in the norm of* $\mathcal{V}$ *independently of* $n \in \mathbb{N}$*.*

Proof. Consider the exact representation of $\overrightarrow{f}_k$ and $\overleftarrow{f}_k$ resulting from [44, Theorem 22], that is,

$$\overrightarrow{f}_k = \sum_{(\sigma,\gamma)\in A(2k)} \overrightarrow{c}_k(\sigma,\gamma) S_\sigma u_\gamma. \tag{4.31}$$

Here, for a sequence $\sigma = (\beta_1, \dots, \beta_j) \in \mathcal{I}$ of multinumbers, $\beta_i \in \mathcal{M}$, $i = 1, \dots, j$, $S_\sigma = \mathcal{R} L_{\beta_1} \cdots \mathcal{R} L_{\beta_j}$, where

$$\mathcal{M} = \{\alpha_1 \dots \alpha_j \colon \alpha_i \in \{0, \dots, d\},\, j \in \mathbb{N}\} \tag{4.32}$$

is the set of multinumbers and

$$\mathcal{I} = \{(\beta_1, \dots, \beta_j) \colon \beta_i \in \mathcal{M},\, j \in \mathbb{N}\}\,. \tag{4.33}$$

For $g \in W_0 := \{f \colon [0,T] \to \mathcal{V}_0 \colon f \text{ is weakly right continuous}\}$, we denote the solution operator of the integral equation

$$u(t) = \int_0^t \mathcal{G} u(s) \mathrm{d}s + \int_0^t g(s) \mathrm{d}s \quad \text{for } t \in [0,T] \tag{4.34}$$

by $\mathcal{R} \colon W_0 \to W_0$, $\mathcal{R} g := u$. $\mathcal{G}_\beta$ is defined recursively by $\mathcal{G}_{\alpha r} = -\mathcal{G}_\alpha \mathcal{G}_r$, $v_\gamma = \mathcal{G}_\gamma u$ with $u(t) = P_t f$, and

$$A(i) = \{(\sigma,\gamma) \colon \sigma \in \mathcal{I},\, \gamma \in \mathcal{M},\, |\sigma| + |\gamma| \leqslant i\}\,, \tag{4.35}$$

and similarly for $\overleftarrow{f}_k$. Note that the coefficients $\overrightarrow{c}_k(\sigma,\gamma)$, $\overleftarrow{c}_k(\sigma,\gamma)$ are independent of $\mathcal{G}_r$. We see that if we can prove in an algebraic manner that the above claim holds true, then we obtain it for arbitrary strongly continuous semigroups.

Assume therefore that $\mathcal{V}$ is finite-dimensional. Then, P^j_t is invertible for all $t \geqslant 0$ and $(P^j_t)_{t\in\mathbb{R}}$ is a strongly continuous group if we set $P^j_{-t} := (P^j_t)^{-1}$ for $t > 0$. Setting

$$E(n) := \frac{1}{2}\left(\overrightarrow{Q}^n_{(T/n)} f + \overleftarrow{Q}^n_{(T/n)} f\right) - P_T, \tag{4.36}$$

we see $E(-n) = E(n)$ for all $n \in \mathbb{N}$ by noting that

$$\left(\overrightarrow{Q}_{(-t)}\right)^{-1} = \overleftarrow{Q}_{(t)}, \tag{4.37}$$

whence $n \mapsto E(n)$ is an even function. This entails that all odd terms in the asymptotic expansion of $E(n)$ have to vanish. Thus, the claim holds true if we choose the $\mathcal{G}_j$ to be arbitrary matrices in a finite-dimensional setting, showing that the coefficients $\overrightarrow{c}_k(\sigma, \gamma)$ and $\overleftarrow{c}_k(\sigma, \gamma)$ have to be such that

$$\overrightarrow{f}_{2\kappa+1} + \overleftarrow{f}_{2\kappa+1} = 0 \quad \text{for } \kappa = 1, \ldots, \left\lfloor \frac{m-1}{2} \right\rfloor. \tag{4.38}$$

The result follows. □

Remark 4.19. The above result is closely connected to the theory of geometric integrators: in that nomenclature, (4.37) states that $\overleftarrow{Q}_{(t)}$ and $\overrightarrow{Q}_{(t)}$ are adjoint to each other, see [46, p. 42]. In this setting, it is well-known that combining a method with its adjoint increases the order of convergence.

The above results are clearly applicable to the setting of Section 4.1. Consider the approximation

$$P_t \approx Q_{t,n} := \frac{1}{2}\left((P^0_{t/n} \ldots P^d_{t/n})^n + (P^d_{t/n} \ldots P^0_{t/n})^n\right). \tag{4.39}$$

We obtain the following result.

Theorem 4.20. *Let $(\delta_j)_{j=1,\ldots,m}$, $\delta_j \in \mathbb{N}$, be pairwise distinct and let $(\theta_j)_{j=1,\ldots,m}$ be such that $\sum_{j=1}^m \theta_j = 1$ and $\sum_{j=1}^m \theta_j \delta^{-2\kappa} = 0$, $\kappa = 1, \ldots, m$. Assume furthermore that $f \in \mathcal{B}^{\psi_\ell^{(n)}}_{8m}((H_\ell)_w)$ with $0 \leqslant \ell \leqslant \ell_0 - 4m - 1$. Then,*

$$\|P_T f - \sum_{j=1}^m \theta_j Q_{T,n\delta_j} f\|_{\mathcal{B}^{\psi^n_{\ell_0}}((H_{\ell_0})_w)} \leqslant C_f n^{-2m}, \tag{4.40}$$

where C_f depends on the choice of the points $(\delta_j)_{j=1,\ldots,m}$ and on f, but not on n.

Remark 4.21. For the case $m = 1$, this result is worse than Theorem 4.11, as we have to assume more smoothness of f. Its importance is however, of course, its applicability to obtain methods of even higher order through extrapolation. Nevertheless, we want to remark that the degree of smoothness required in Proposition 4.17 appears to be suboptimal, and we expect that the conclusion of Theorem 4.20 holds true for $f \in \mathcal{B}^{\psi_\ell^{(n)}}_{2(2m+1)}((H_\ell)_w)$ with $0 \leqslant \ell \leqslant \ell_0 - 2(2m+1) - 1$.

4.3 Numerical example: the HJM equation of interest rate theory

We consider the approximation of the Heath-Jarrow-Morton equation of interest rate theory. For a background on interest rate modelling, see [20, 78, 25, 36]. The state space consists of forward yield curves $r\colon [0,\infty) \to \mathbb{R}$, the parameter being time to maturity. In its more natural Itô formulation, the equation reads

$$(4.41)\qquad \mathrm{d}r(t,r_0) = (Ar(t,r_0) + \alpha_{\mathrm{HJM}}(r(t,r_0)))\mathrm{d}t + \sum_{j=1}^{d} \sigma_j(r(t,r_0))\mathrm{d}B_t^j,$$
$$r(0,r_0) = r_0.$$

Here, $A = \frac{\mathrm{d}}{\mathrm{d}x}$. Hence, this problem can be interpreted as a stochastically perturbed transport equation.

If we want to price financial derivatives by taking expectations, it is necessary to use the *risk-neutral measure* as underlying probability measure. Under this measure, loosely speaking, traded financial derivatives become (local) martingales, see [36, Section 4.3]. For our purposes, this means that α_{HJM} and σ_j are coupled by the *Heath-Jarrow-Morton drift condition*, see, e.g., [35, Lemma 4.3.3], [25, p. 61], [36, Theorem 6.1]. Thus,

$$(4.42)\qquad \alpha_{HJM}(r)(x) = \sum_{j=1}^{d} \sigma_j(r)(x) \int_0^x \sigma_j(r)(\xi)\mathrm{d}\xi.$$

Recall from Section 3.2.1 the choice of spaces

$$(4.43)\qquad H_\ell := \left\{ h \in \mathrm{L}^1_{\mathrm{loc}}(\mathbb{R}_+)\colon h' \in \mathrm{H}^{\ell+1}_{\alpha_\ell}(\mathbb{R}_+) \right\},$$

endowed with the norm

$$(4.44)\qquad \|h\|_{H_\ell} := \left(|h(0)|^2 + \|h'\|^2_{\mathrm{H}^{\ell+1}_{\alpha_\ell}(\mathbb{R}_+)} \right)^{1/2}.$$

We remarked there that $A(H_\ell) \subset H_{\ell-1}$, $\ell = 0, \ldots, \ell_0$. On every H_ℓ, the infinitesimal generator of the shift semigroup $(S_t)_{t\geqslant 0}$, $S_t f(x) = f(t+x)$, equals A on the dense set of infinitely often differentiable functions with compact support.

As we want to apply a second order splitting, we fix $\ell_0 = 6$; for the first order splitting, the choice $\ell_0 = 4$ is adequate. Defining ψ^s_ℓ and $\varphi_{\ell,j}$ according to (3.12) and (3.13), we see that we need to satisfy Assumptions 3.12, 3.13 and 3.15. Assumption 3.12 is clear by definition. Theorem 3.24 proves Assumption 3.18, hence Theorem 3.19 yields Assumption 3.15.

It remains to choose vector fields σ_j, $j = 0, \ldots, d$, in such a way that Assumption 3.13 is fulfilled. We do this in the following way: for $j = 1, \ldots, d$, let $\sigma_j(r) = g_j(r)\lambda_j$, where $\lambda_j \in H_{\ell_0}$ satisfies $\lim_{x\to\infty} \lambda_j^{(\ell)}(x) = 0$ for $\ell = 0, \ldots, \ell_0$, and $g \in \mathcal{A}(H_{-1})$. As $\sigma_j \in \mathcal{A}(H_{-1}; H_{\ell_0})$, we see that $\sigma_j \in \mathcal{C}_k^{\varphi}(H_\ell; H_\ell)$, $\ell = 0, \ldots, \ell_0$, $j = 1, \ldots, d$. This choice is inspired by the results from [37].

Finally, consider the Stratonovich drift $V_0(r) = \alpha_{\mathrm{HJM}}(r) - \frac{1}{2}\sum_{j=1}^d D\sigma_j(r)\sigma_j(r)$. As argued above, the Itô drift has to have the form

$$\alpha_{\mathrm{HJM}}(r)(x) = \sum_{j=1}^{d} \sigma_j(r)(x) \int_0^x \sigma_j(r)(\xi)\mathrm{d}\xi \quad \text{for } x \in \mathbb{R}_+ \text{ and } r \in H_{-1}. \tag{4.45}$$

In our case, this expression simplifies to

$$\alpha_{\mathrm{HJM}}(r)(x) = \sum_{j=1}^{d} g_j(r)^2 \lambda_j(x) \int_0^x \lambda_j(\xi)\mathrm{d}\xi, \tag{4.46}$$

and we see that $\alpha_{HJM} \in \mathcal{A}(H_{-1}; H_{\ell_0})$: we have that

$$\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}\left(\lambda_j(x)\int_0^x \lambda_j(\xi)\mathrm{d}\xi\right) = \lambda_j^{(\ell)}(x)\int_0^x \lambda_j(\xi)\mathrm{d}\xi + \sum_{i=0}^{\ell-1} \lambda_j^{(i)}\lambda_j^{(\ell-1-i)}(x). \tag{4.47}$$

The $\mathrm{L}^2_{\alpha_{\ell_0}}(\mathbb{R}_+)$ norm of the first term is bounded due to [35, equation (5.3)]. Estimating the norms of the terms in the sum similarly as on [35, p. 79] using [35, equations (5.7), (5.8)], it follows that α_{HJM} takes its values in H_{ℓ_0}. Thus, as $g_j \in \mathcal{A}(H_{-1})$, $j = 1, \ldots, d$, the claim follows.

There is one final term in V_0, the Stratonovich correction. It equals

$$-\frac{1}{2}\sum_{j=1}^{d} D\sigma_j(r)\sigma_j(r) = -\frac{1}{2}\sum_{j=1}^{d} g_j(r)Dg_j(r)(\lambda_j)\lambda_j. \tag{4.48}$$

This is again in $\mathcal{A}(H_{-1}; H_{\ell_0})$. Hence, we obtain that $V_0 \in \mathcal{A}(H_{-1}; H_{\ell_0})$, and Assumption 3.13 is satisfied.

As a concrete choice, let us assume that for $j = 1, \ldots, d$ and some mesh $(t_i)_{i=1,\ldots,M}$ of time points,

$$g_j(r) = \gamma_j(r(t_1), \ldots, r(t_M)), \tag{4.49}$$

where $\gamma_j \in C_b^\infty(\mathbb{R}^M)$. Such a choice is admissible as point evaluations define continuous linear functionals on H_{-1}, and has the economic interpretation of using benchmark forward rates to drive the process [37, p. 135]. Furthermore, for some $N \in \mathbb{N}$,

$$\lambda_j(x) := \sum_{i=0}^{N} \alpha_{ji} x^i \exp(-\beta_j x), \tag{4.50}$$

where the $\beta_j \in (0, \infty)$ are chosen in such a way that $\lambda_j \in H_{\ell_0}$, and $\alpha_{ji} \in \mathbb{R}$.

To determine the parameters of our model, we calibrate to the caplet prices given in [60, Section 2.6]. *Caplets* are financial derivatives on rates, paying off a certain amount if the rate is larger than a fixed strike, providing insurance against rising rates. More precisely, with the price of a zero coupon bond at time t with maturity T given by

$$P(t, T) := \exp\left(-\int_t^T r(t)(\tau)\mathrm{d}\tau\right), \tag{4.51}$$

we define the LIBOR rate with maturity T for $\delta > 0$ by

$$L(t, T) := \frac{1}{\delta}\left(\frac{P(t, T)}{P(t, T + \delta)} - 1\right) \tag{4.52}$$

and the payoff of the caplet on the LIBOR at T, which is settled at $T + \delta$, with strike K by

$$C_{T+\delta}(T, K) := (L(T, T) - K)_+ . \tag{4.53}$$

Here, $x_+ := \max(x, 0)$ denotes the positive part of $x \in \mathbb{R}$. Note that while the value of $C_{T+\delta}(T, K)$ is determined at time T, the cash flow happens only at $T + \delta$. The LIBOR rate is defined in such a way that

$$1 + \delta L(t, T) = \frac{P(t, T)}{P(t, T + \delta)}, \tag{4.54}$$

i.e., discrete time interest over the time interval $[T, T + \delta]$ with rate $L(t, T)$ corresponds to the bond structure. By standard no arbitrage arguments, we obtain that the fair value of the caplet at time $t < T$ is given by

$$C_t(T, K) = \mathbb{E}[B_{T+\delta}^{-1} C_{T+\delta}(T, K)], \tag{4.55}$$

where $B_t := \exp\left(\int_0^t r(s)(0)\mathrm{d}s\right)$ denotes the *money market account*, where money is continuously compounded by the short rate $r(t)(0)$

To fully discretise the stochastic partial differential equation, we approximate r by a piecewise affine and continuous function. Choosing $\Delta x = \Delta t$, we see that $\frac{\mathrm{d}}{\mathrm{d}t} r(t, r_0) = \frac{\mathrm{d}}{\mathrm{d}x} r(t, r_0)$ is solved exactly by the shift, whence we do not incur any additional error from the space discretisation. We apply the symmetrically weighted sequential splitting analysed in Section 4.2. In order to solve the remaining deterministic problem, $\frac{\mathrm{d}}{\mathrm{d}t} r(t, r_0) = Ar(t, r_0) + V_0(r(t, r_0))$, we again perform a splitting into the equations

$$\frac{\mathrm{d}}{\mathrm{d}t} r(t, r_0) = \frac{\mathrm{d}}{\mathrm{d}x} r(t, r_0), \tag{4.56}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} r(t, r_0) = g_j(r(t, r_0))^2 \lambda_j \int_0^{\cdot} \lambda_j(\xi)\mathrm{d}\xi, \tag{4.57}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} r(t, r_0) = -\frac{1}{2} g_j(r(t, r_0)) D g_j(r(t, r_0))(\lambda_j)\lambda_j. \tag{4.58}$$

Embedding this in the symmetrically weighted sequential splitting, we see that we preserve the rate of convergence of 2 also for this scheme.

In the calibration, we set $d = 3$, i.e., we use a three factor model. The functions γ_j are of the form

$$\gamma(r) = \frac{(\sum_{i=1}^m a_i r(t_i))(1 + \sum_{i=1}^m b_i(r_i)}{1 + \sum_{i=1}^m c_i r(t_i)^2}, \tag{4.59}$$

with the parameters a_i, b_i, c_i and t_i. We assume that $\gamma_1 = \gamma_2$, with $m = 2$, and γ_3 with $m = 1$. The λ_j are all chosen independently, with $N = 3$ each. Hence, in total, we have 24 parameters. The final error obtained in the calibration after 500 Levenberg-Marquardt steps using the code by Lourakis [69] was 292 basis points, where the error is measured in the fit to the implied volatility surface given in [60, Section 2.6]. The fit is shown in Figure 4.3. The calibration time was 7.5 minutes, in which 1557 evaluations of the entire implied volatility surface were performed, using 2048 quasi-Monte Carlo paths each.

As an application, we price an at the money payer swaption. A *payer swaption* is the right, but not the obligation, to enter a payer swap at a certain future date T with a certain, a priori determined fixed rate K. A *payer swap* in turn is an exchange of a fixed rate versus a floating rate, i.e., at certain pre-determined points in time, the owner of the payer swap receives the floating rate, and pays the fixed rate. Hence, a payer swap protects against changing rates. The price

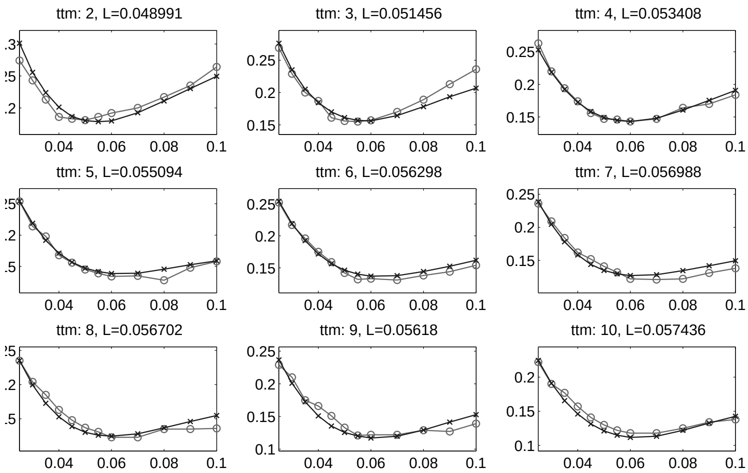

Figure 4.1: Fit of caplet volatilites

Circles correspond to market prices, crosses to the prices given by the model. ttm denotes the time to maturity of the time slice, L the current LIBOR rate for the corresponding period.

of a payer swap with fixed rate K at time 0 is given by

$$PS_0 := 1 - P(0, T_N) - \sum_{i=1}^{N} \delta K P(0, T_i), \tag{4.60}$$

by a no arbitrage argument, where $T_i = i\delta$, i.e., we assume equidistant payment dates. The *at the money* price of a payer swap is the amount for K such that $PS_0 = 0$. A no arbitrage argument shows that the time 0 price of a payer swaption is given by

$$\mathbb{E}[B_T^{-1}(1 - P(T, T_N) - \sum_{i=1}^{N} \delta K P(T, T_i))_+], \tag{4.61}$$

where now $T_i = T + i\delta$, and $x_+ := \max(x, 0)$ again denotes the positive part.

In our numerical example, we let $\delta = .25$ and $N = 12$. At time 0, the at the money value of the swap is $K_{\text{ATM}} = 0.0442608$. Using this as strike in the swaption, we obtain the reference value 0.01192380 by solving the problem with

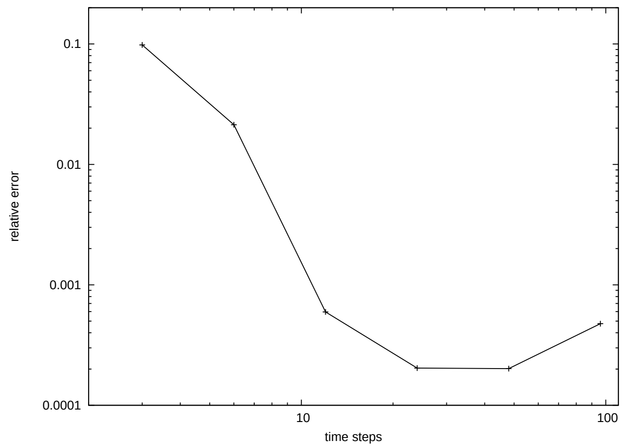

Figure 4.2: Error, swaption prices

2^{20} quasi-Monte Carlo paths and 120 time steps. The numbers in Figure 4.2 were obtained using 2^{16} quasi-Monte Carlo paths. The rate of convergence is approximately 1.7. We note that with 12 time steps, we obtain a relative error of less than $1e-3$, proving the viability of the new method.

Chapter 5

Cubature Schemes For Stochastic (Partial) Differential Equations

In Chapter 4, we saw a way of extending Ninomiya-Victoir-type splitting schemes to stochastic partial differential equations. However, as noted there, it is not possible to obtain splitting schemes of orders higher than 2 for generic equations. On the other hand, cubature formulas are available of high degree, see [45] for paths resulting in rates of convergence up to 5 for single factor problems. This means that a proof of rates of convergence for cubature methods would allow us to obtain high order methods for the numerical simulation of stochastic partial differential equations without having to resort to extrapolation.

Similarly as for splitting methods, we need to obtain results ensuring that

(i) the approximation operators $(Q_{(t)})_{t\geqslant 0}$ defined by the cubature method are power bounded in an appropriate weighted ψ-norm, and that

(ii) a local error estimate of the correct order holds true.

Together, this will yield convergence of high order for sufficiently smooth functions, similarly as in Theorem 4.11.

In this chapter, we will mainly focus on the proof of stability of cubature schemes with respect to weighted norms. In Section 5.1, we show that cubature is always stable in finite dimensions. In infinite dimensions, we first consider stochastic ordinary differential equations where the vector fields have a nonsmooth time dependence. Afterwards, the method of the moving frame yields stability for cubature approximations of Da Prato-Zabczyk equations where the generator is pseudo-contractive. In every case, convergence is easily obtained by an application of the local expansions obtained in Section 3.2.2, and in Section 5.2, we formulate corresponding results. Finally, in Section 5.2.3, we show that in finite dimensions, we have smoothing effects in $\mathcal{B}^{\psi}$ spaces under the UFG condition, allowing us to obtain optimal estimates of the rate of convergence for nonsmooth payoffs.

5.1 Stability of cubature schemes

We prove stability of cubature on Wiener space in the setting of weighted spaces. See Section 1.1.2 for the definition of cubature paths.

5.1.1 Finite dimensional state space

Given a Stratonovich SDE on $\mathbb{R}^N$,

$$\mathrm{d}x(t,x_0) = \sum_{j=0}^{d} V_j(x(t,x_0)) \circ \mathrm{d}B_t^j, \quad x(0,x_0) = x_0, \tag{5.1}$$

with vector fields $V_j\colon \mathbb{R}^N \to \mathbb{R}^N$, the cubature discretisation of the Markov semigroup $P_t f(x_0) := \mathbb{E}[f(x(t,x_0))]$ reads

$$Q_{(\Delta t)} f(x_0) := \sum_{i=1}^{N} \lambda_i f(x(\Delta t, x_0; \omega_i^{(\Delta t)})), \tag{5.2}$$

where $x(s,x_0;\omega_i^{(\Delta t)})$ is the solution of the problem

$$\mathrm{d}x(s,x_0;\omega_i^{(\Delta t)}) = \sum_{j=0}^{d} V_j(x(s,x_0;\omega_i^{(\Delta t)}))\mathrm{d}\omega_i^{(\Delta t),j}(s), \tag{5.3a}$$

$$x(0,x_0;\omega_i^{(\Delta t)}) = x_0. \tag{5.3b}$$

Theorem 5.1. *Let the cubature formula $(\omega_i^{(\Delta t)}, \lambda_i)_{i=1}^M$ be of order $m \geqslant 1$. Suppose that $\psi\colon \mathbb{R}^N \to \mathbb{R}$ is an admissible weight function, and assume that*

$$|V_i V_j \psi(x)| + |V_i \psi(x)| \leqslant C\psi(x) \quad \text{for } i = 0, \cdots, d \text{ and } j = 1, \cdots, d, \tag{5.4}$$

where we require that all the necessary derivatives are well-defined.

Then, there exists a constant $C > 0$ independent of $\Delta t > 0$ such that

$$Q_{(\Delta t)}\psi(x_0) \leqslant \exp(C\Delta t)\psi(x_0) \quad \text{for all } x_0 \in \mathbb{R}^N. \tag{5.5}$$

Proof. We define the intermediate operator

$$Q_{(\Delta t,s)}f(x_0) := \sum_{i=1}^{N} \lambda_i f(x(s,x_0;\omega_i^{(\Delta t)})) \quad \text{for } s \in [0,t] \tag{5.6}$$

and note that $Q_{(\Delta t)} = Q_{(\Delta t,\Delta t)}$. We see that

$$\begin{aligned}\psi(x(s,x_0;\omega_i^{(\Delta t)})) &= \psi(x_0) + \sum_{j=0}^{d}\int_0^s V_j\psi(x(r,x_0;\omega_i^{(\Delta t)}))\mathrm{d}\omega_i^{(\Delta t),j}(r)\\ &= \psi(x_0) + \int_0^s V_0\psi(x(r,x_0;\omega_i^{(\Delta t)}))\mathrm{d}r + \sum_{j=1}^{d} V_j\psi(x_0)\omega_i^{(\Delta t),j}(s)\\ &\quad + \sum_{j=1}^{d}\sum_{k=0}^{d}\int_0^s\int_0^r V_kV_j\psi(x(q,x_0;\omega_i^{(\Delta t)}))\mathrm{d}\omega_i^{(\Delta t),k}(q)\mathrm{d}\omega_i^{(\Delta t),j}(r). \qquad (5.7)\end{aligned}$$

Note that

$$\int_0^s V_0\psi(x(r,x_0;\omega_i^{(\Delta t)}))\mathrm{d}r \leqslant C\int_0^s \psi(x(r,x_0;\omega_i^{(\Delta t)}))\mathrm{d}r. \tag{5.8}$$

Furthermore, as $|\omega_i^{(\Delta t),j}(s)| \leqslant C(\Delta t)^{1/2}$ and $|\frac{\partial}{\partial s}\omega_i^{(\Delta t),j}(s)| \leqslant C(\Delta t)^{-1/2}$, Fubini's theorem yields

$$\begin{aligned}&\int_0^s\int_0^r V_kV_j\psi(x(q,x_0;\omega_i^{(\Delta t)}))\mathrm{d}\omega_i^{(\Delta t),k}(q)\mathrm{d}\omega_i^{(\Delta t),j}(r)\\ &\qquad\leqslant C\int_0^s |\omega_i^{(\Delta t),j}(s)-\omega_i^{(\Delta t),j}(q)|\psi(x(q,x_0;\omega_i^{(\Delta t)}))\left|\frac{\partial}{\partial q}\omega_i^{(\Delta t),j}(q)\right|\mathrm{d}q\\ &\qquad\leqslant C\int_0^s \psi(x(q,x_0;\omega_i^{(\Delta t)}))\mathrm{d}q. \qquad (5.9)\end{aligned}$$

Thus, we see that

$$\begin{aligned}Q_{(\Delta t,s)}\psi(x_0) &= \sum_{i=1}^{N}\lambda_i\psi(x(s,x_0;\omega_i^{(\Delta t)}))\\ &\leqslant \psi(x_0) + \sum_{j=1}^{d} V_j\psi(x_0)\sum_{i=1}^{N}\lambda_i\omega_i^{(\Delta t),j}(s) + C\int_0^s Q_{(\Delta t,r)}\psi(x_0)\mathrm{d}r. \qquad (5.10)\end{aligned}$$

Defining $\alpha_{\Delta t,s}(x) := \sum_{j=1}^{d} V_j\psi(x) \sum_{i=1}^{N} \lambda_i \omega_i^{(\Delta t),j}(s)$, Gronwall's inequality yields that

$$\begin{aligned} Q_{(\Delta t,s)}\psi(x_0) &\leqslant \psi(x_0) + \alpha_{\Delta t,s}(x_0) \\ &\quad + \int_0^s (\psi(x_0) + \alpha_{\Delta t,r}(x_0))\, C \exp(C(s-r))\mathrm{d}r. \end{aligned} \tag{5.11}$$

Note that $\alpha_{\Delta t,\Delta t}(x) = 0$ by the equality $\sum_{i=1}^{N} \lambda_i \omega_i^{(\Delta t),j}(\Delta t) = 0$. Furthermore,

$$\alpha_{\Delta t,s}(x) \leqslant C\sqrt{\Delta t}\psi(x) \leqslant \frac{C}{2}(1+\Delta t)\psi(x) \leqslant \frac{C}{2}\exp(\Delta t)\psi(x). \tag{5.12}$$

This proves

$$\begin{aligned} Q_{(\Delta t)}\psi(x_0) = Q_{(\Delta t,\Delta t)}\psi(x_0) &\leqslant \psi(x_0)\left(1 + \frac{C}{2}\exp(\Delta t)(\exp(C\Delta t) - 1)\right) \\ &\leqslant \exp(\tilde{C}\Delta t)\psi(x_0), \end{aligned} \tag{5.13}$$

where $\tilde{C} = \max(C^2/2, C+1)$, that is, the required estimate. □

5.1.2 Time-dependent stochastic ordinary differential equations on Hilbert space

Let H be a Hilbert space, and consider the nonautonomous stochastic ordinary differential equation

$$\mathrm{d}x(t,x_0) = \sum_{j=0}^{d} V_j(t, x(t,x_0)) \circ \mathrm{d}B_t^j, \quad x(0,x_0) = x, \tag{5.14}$$

on H. The cubature approximations of (5.14) read

$$\mathrm{d}x(s,x_0;t,\omega_i^{(\Delta t)}) = \sum_{j=0}^{d} V_j(t+s, x(s,x_0;t,\omega_i^{(\Delta t)}))\mathrm{d}\omega_i^{(\Delta t),j}(s), \tag{5.15a}$$

$$x(0,x_0;t,\omega_i^{(\Delta t)}) = x_0, \tag{5.15b}$$

The approximation operator is given by

$$Q_{(\Delta t)}^t f(x_0) := \sum_{i=1}^{N} \lambda_i f(x(\Delta t, x_0; t, \omega_i^{(\Delta t)})). \tag{5.16}$$

Definition 5.2. A cubature formula $(\omega_i^{(\Delta t)}, \lambda_i)_{i=1,\dots,N}$ is called *symmetric* if for every $i \in \{1, \dots, N\}$, there exists some $i' \in \{1, \dots, N\}$ such that $\lambda_i = \lambda_{i'}$ and

$$\omega_i^{(\Delta t),j}(s) = -\omega_{i'}^{(\Delta t),j}(s) \quad \text{for all } s \in [0, \Delta t] \text{ and } j = 1, \dots, d. \tag{5.17}$$

It is called *weakly symmetric* if for $j = 1, \dots, d$,

$$\sum_{i=1}^{N} \lambda_i \omega_i^{(\Delta t),j}(s) = 0 \quad \text{for } s \in [0, \Delta t]. \tag{5.18}$$

Remark 5.3. Clearly, all symmetric cubature formulas are also weakly symmetric. Note that many known cubature formulas are actually symmetric. Moreover, a non-symmetric cubature formula can be made symmetric by adding the negatives of the paths with the same weights to it and finally halving all the weights. This will at most double the number of paths. Thus, if we use a cubature formula with a small number of paths in high dimensions, we can also find a symmetric cubature formula with this property.

Theorem 5.4. *Suppose that the cubature formula used in the definition of* $Q^t_{(\Delta t)}$ *is weakly symmetric. Let* ψ *be an admissible weight function on* H *and suppose*

$$\|D\psi(x)\| \leqslant C(1 + \|x\|^2)^{-1/2}\psi(x) \quad \text{and} \tag{5.19}$$

$$\|D^2\psi(x)\| \leqslant C(1 + \|x\|^2)^{-1}\psi(x) \tag{5.20}$$

with some constant $C > 0$*, Furthermore, assume that for some constant* $C > 0$ *independent of* t*,*

$$\|V_j(t, x)\| \leqslant C(1 + \|x\|^2)^{1/2} \quad \text{for } j = 0, \dots, d,\ x \in X \text{ and } t \in [0, T], \tag{5.21}$$

and that $x \mapsto V_j(\Delta t, x)$ *is continuously differentiable with derivative bounded uniformly in* $t \in [0, T]$ *for* $j = 1, \dots, d$*.*

Then, there exists a constant $C > 0$ *such that for all* $t \in [0, T]$ *and* $\Delta t \in [0, T - t]$*,*

$$Q^t_{(\Delta t)}\psi(x_0) \leqslant \exp(Ct)\psi(x_0) \quad \text{for all } x_0 \in H. \tag{5.22}$$

Proof. Define the intermediate approximation for $s \in [0, \Delta t]$ by

$$Q^t_{(\Delta t, s)} f(x_0) := \sum_{i=1}^{N} \lambda_i f(x(s, x_0; t, \omega_i^{(\Delta t)})). \tag{5.23}$$

As above, we note that $Q^t_{(\Delta t,\Delta t)} = Q^t_{(\Delta t)}$. For $0 \leqslant s \leqslant \Delta t$,

$$\begin{aligned}
&\psi(x(s,x_0;t,\omega_i^{(\Delta t)})) = \psi(x_0)+ \\
&\sum_{j=0}^{d}\int_0^s D\psi(x(r,x_0;t,\omega_i^{(\Delta t)}))V_j(t+r,x(r,x_0;t,\omega_i^{(\Delta t)}))\mathrm{d}\omega_i^{(\Delta t),j}(r).
\end{aligned}\tag{5.24}$$

Consider $g_j(r,x) := D\psi(x)V_j(t+r,x)$. Then,

$$\begin{aligned}
g_j(\rho,x(r,x_0;t,\omega_i^{(\Delta t)})) =&g_j(\rho,x_0) + \sum_{k=0}^{d}\int_0^r D_x g_j(\rho,x(q,x_0;t,\omega_i^{(\Delta t)}))\times \\
&\times V_k(t+q,x(q,x_0;t,\omega_i^{(\Delta t)}))\mathrm{d}\omega_i^{(\Delta t),k}(q).
\end{aligned}\tag{5.25}$$

From (5.19), (5.20) and (5.21), we obtain that for $0 \leqslant s \leqslant \Delta t \leqslant T$,

$$\begin{aligned}
|g_0(r,x)| = |D\psi(x)V_0(t+r,x)| &\leqslant C\|D\psi(x)\|\cdot\|V_0(t+r,x)\| \\
&\leqslant C\psi(x).
\end{aligned}\tag{5.26}$$

We argue in a similar manner for $D_x g_j(r,x)V_k(t+q,x)$, $j = 1,\dots,d$, $k = 0,\dots,d$, to obtain that for $0 \leqslant q \leqslant r \leqslant \Delta t$,

$$\begin{aligned}
|D_x g_j(r,x)V_k(t+q,x)| &= |D^2\psi(x)(V_j(t+r,x),V_k(t+q,x)) \\
&\qquad + D\psi(x)D_xV_j(t+r,x)V_k(t+q,x)| \\
&\leqslant C\psi(x).
\end{aligned}\tag{5.27}$$

By an application of Fubini's theorem, similarly as in the proof of Theorem 5.1,

$$\begin{aligned}
\psi(x(s,x_0;t,\omega_i^{(\Delta t)})) &= \psi(x_0) + \int_0^s g_0(r,x(r,x_0;t,\omega_i^{(\Delta t)}))\mathrm{d}r \\
&\quad + \sum_{j=1}^{d}\int_0^s g_j(r,x_0)\mathrm{d}\omega_i^{(\Delta t),j}(r) \\
&\quad + \sum_{j=1}^{d}\sum_{k=0}^{d}\int_0^s\int_0^r D_x g_j(r,x(q,x_0;t,\omega_i^{(\Delta t)}))\times \\
&\quad \times V_k(t+q,x(q,x_0;t,\omega_i^{(\Delta t)}))\mathrm{d}\omega_i^{(\Delta t),k}(q)\mathrm{d}\omega_i^{(\Delta t),j}(r) \\
&\leqslant \psi(x_0) + C\int_0^s \psi(x(r,x_0;t,\omega_i^{(\Delta t)}))\mathrm{d}r \\
&\quad + \sum_{j=1}^{d}\int_0^s g_j(r,x_0)\mathrm{d}\omega_i^{(\Delta t),j}(r),
\end{aligned}\tag{5.28}$$

where we apply that $\Delta t \leqslant T$. As from the weak symmetry of the cubature paths,

$$\begin{aligned}\sum_{i=1}^{N}\lambda_i\sum_{j=1}^{d}\int_0^s g_j(r,x)\mathrm{d}\omega_i^{(\Delta t),j}(r) &= \sum_{j=1}^{d}\int_0^s g_j(r,x)\mathrm{d}\Big(\sum_{i=1}^{N}\lambda_i\omega_i^{(\Delta t),j}(r)\Big)\\ &= 0,\end{aligned} \tag{5.29}$$

we obtain

$$Q_{(\Delta t,s)}\psi(x_0) \leqslant \psi(x_0) + C\int_0^s Q_{(\Delta t,r)}\psi(x_0)\mathrm{d}r. \tag{5.30}$$

An application of Gronwall's lemma yields $Q_{(\Delta t)}\psi(x_0) \leqslant \exp(C\Delta t)\psi(x_0)$, which proves the result. □

Remark 5.5. It is clear that the given assumptions on the vector fields and the weight function are not the only ones possible. Instead, we could also require the vector fields to be bounded uniformly in $t \in [0,T]$, and allow the weight function to satisfy $\|D\psi(x)\| + \|D^2\psi(x)\| \leqslant C\psi(x)$. While the situation of Theorem 5.4 corresponds to polynomially growing weight functions and linearly bounded vector fields, this variant corresponds to exponentially growing weight functions and bounded vector fields, cf. Corollaries 2.45 and 2.43.

Such an approach might be more appropriate when dealing with exponentials of stochastic processes such as Lévy processes. These are ubiquitous in applications in mathematical finance as they ensure nonnegativity of the price process in a simple manner.

5.1.3 Da Prato-Zabczyk equations

Suppose now that

$$\mathrm{d}x(t,x_0) = Ax(t,x_0)\mathrm{d}t + \sum_{j=0}^{d} V_j(x(t,x_0))\circ \mathrm{d}B_t^j, \quad x(0,x_0) = x_0, \tag{5.31}$$

is a stochastic partial differential equation of Da Prato-Zabczyk type on some Hilbert space H. Refer to Section B.3 for an overview of the theory of such equations. Here, solutions are understood in the mild sense,

$$x(t,x_0) = \exp(tA)x + \sum_{j=0}^{d}\int_0^t \exp((t-s)A)V_j(x(t,x_0))\circ \mathrm{d}B_s^j, \tag{5.32}$$

and we also define the cubature discretisations in the mild sense,

$$x(t,x_0)(\omega_i^{(\Delta t)}) = \exp(tA)x + \sum_{j=0}^{d}\int_0^t \exp((t-s)A)V_j(x(t,x_0)(\omega_i^{(\Delta t)}))\mathrm{d}\omega_i^{(\Delta t),j}(s). \tag{5.33}$$

Again, the approximation of the Markov semigroup $P_t f(x_0) := \mathbb{E}[f(x(t,x_0))]$ is given by

$$Q_{(\Delta t)}f(x_0) := \sum_{i=1}^{N}\lambda_i f(x(t,x_0;\omega_i^{(\Delta t)})). \tag{5.34}$$

Theorem 5.6. *Suppose that A is the generator of a group $S_t = \exp(tA)$, $t \in \mathbb{R}$, and that the cubature formula used in the definition of $Q_{(\Delta t)}$ is weakly symmetric. Let ψ be an admissible weight function on H. With some constant $C > 0$, let $\psi(S_t x) \leqslant \exp(Ct)\psi(x)$ for all $x \in H$ and $t > 0$, and*

$$\|D\psi(x)\| \leqslant C(1+\|x\|^2)^{-1/2}\psi(x) \quad \text{and} \tag{5.35a}$$

$$\|D^2\psi(x)\| \leqslant C(1+\|x\|^2)^{-1}\psi(x). \tag{5.35b}$$

Furthermore, assume that

$$\|V_j(x)\| \leqslant C(1+\|x\|^2)^{1/2} \quad \text{for } j = 0,\dots,d, \tag{5.36}$$

and that V_j is continuously differentiable with bounded derivative for $j = 1,\dots,d$.

Then, for any $T > 0$, there exists a constant $C > 0$ such that for every $\Delta t \in [0,T]$,

$$Q_{(\Delta t)}\psi(x_0) \leqslant \exp(C\Delta t)\psi(x_0) \quad \text{for all } x_0 \in H. \tag{5.37}$$

Proof. We apply the method of the moving frame from [106]. This yields that $x(t,x_0) = S_t y(t,x_0)$, where $(y(t,y_0))_{t\geqslant 0}$ satisfies the Hilbert space ordinary stochastic differential equation

$$\mathrm{d}y(t,y_0) = \sum_{j=0}^{d}\tilde{V}_j(t,y(t,y_0)) \circ \mathrm{d}B_t^j, \quad y(0,y_0) = y_0, \tag{5.38}$$

with $\tilde{V}_j(t,y) = S_{-t}V_j(S_t y)$. Thus, rewriting the cubature discretisations of $(x(t,x_0))_{t\geqslant 0}$ using $(y(t,x_0))_{t\geqslant 0}$,

$$\mathrm{d}y(s,x_0;\omega_i^{(\Delta t)}) = \sum_{j=0}^{d}\tilde{V}_j(s,y(s,x_0;\omega_i^{(\Delta t)}))\mathrm{d}\omega_i^{(\Delta t),j}(s), \tag{5.39}$$

we see that, if we define

$$\tilde{Q}_{(\Delta t)} f(y_0) := \sum_{i=1}^{N} \lambda_i f(y(\Delta t, y_0; \omega_i^{(\Delta t)})) \tag{5.40}$$

for $f\colon H \to \mathbb{R}$, then $Q_{(\Delta t)} h(x_0) = \tilde{Q}_{(\Delta t)} g(x_0)$, where $g(y) := h(S_{\Delta t} y)$. In particular,

$$Q_{(\Delta t)} \psi(x_0) = \tilde{Q}_{(\Delta t)} (\psi \circ S_{\Delta t})(x_0) \leqslant \exp(C\Delta t) \tilde{Q}_{(\Delta t)} \psi(x_0), \tag{5.41}$$

where we apply the assumptions on ψ and the positivity of $\tilde{Q}_{(\Delta t)}$.

Hence, we are in the situation of Theorem 5.4: the estimates for ψ are clear by assumption, and for $\tilde{V}_j(s, y)$, we note that, as $s \in [0, T]$,

$$\|\tilde{V}_j(s, y)\| = \|S_{-s} V_j(S_s y)\| \leqslant C(1 + \|x\|^2)^{1/2} \quad \text{for } j = 0, \dots, d \tag{5.42}$$

and

$$\|D_y \tilde{V}_j(s, y)\| = \|S_{-s} D_y V_j(S_s y) S_s\| \leqslant C \quad \text{for } j = 1, \dots, d. \tag{5.43}$$

An appeal to Theorem 5.4 yields

$$\tilde{Q}_{(\Delta t)} \psi(x_0) \leqslant \exp(C\Delta t) \psi(x_0), \tag{5.44}$$

and the result follows. □

The Szőkefalvi-Nagy theorem allows us to obtain a corresponding result for pseudocontractive semigroups.

Corollary 5.7. *Suppose that A is the generator of a semigroup of pseudocontractions $S_t = \exp(tA)$, $t \geqslant 0$. Let $\psi(x) = \rho(\|x\|^2)$ with some increasing and left continuous function $\rho\colon [0, \infty) \to (0, \infty)$ with $\lim_{\xi \to \infty} \rho(\xi) = +\infty$ (see also Example 2.16) that satisfies $\rho(Cu) \leqslant C\rho(u)$ for all $u \geqslant 0$ and $C > 0$, is twice differentiable and satisfies*

$$\rho'(u) \leqslant C(1 + u)^{-1} \rho(u) \quad \text{and} \quad \rho''(u) \leqslant C(1 + u)^{-2} \rho(u). \tag{5.45}$$

Furthermore, assume that $\|V_j(x)\| \leqslant C(1 + \|x\|^2)^{1/2}$ for $j = 0, \dots, d$, and that V_j is continuously differentiable with bounded derivative for $j = 1, \dots, d$.

Then, for any $T > 0$, there exists a constant $C > 0$ such that for every $\Delta t \in [0, T]$, the operator $Q_{(\Delta t)}$ satisfies

$$Q_{(\Delta t)} \psi(x_0) \leqslant \exp(C\Delta t) \psi(x_0) \quad \text{for all } x_0 \in H. \tag{5.46}$$

Proof. Assume without loss of generality that $(S_t)_{t\geqslant 0}$ is a semigroup of contractions (otherwise, add any growth of $(S_t)_{t\geqslant 0}$ to V_0). By the Szőkefalvi-Nagy theorem [93, p. 452, Théorème IV], we see that we can find a Hilbert space $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ containing H as a closed subspace and a strongly continuous group $(\mathcal{S}_t)_{t\in\mathbb{R}}$ of unitary mappings such that $S_t = \pi\mathcal{S}_t$, where $\pi\colon \mathcal{H} \to H$ is the orthogonal projection.

Defining $\psi_{\mathcal{H}}(y) := \rho(\|y\|_{\mathcal{H}}^2)$ and $V_j^{\mathcal{H}}(y) := V_j(\pi y)$, it is easy to see that the assumptions of Theorem 5.6 are satisfied. The results of [106] prove that $x(t,x_0) = \pi x^{\mathcal{H}}(t,x_0)$, where

$$x^{\mathcal{H}}(t,x_0) = \mathcal{S}_t x_0 + \sum_{j=0}^{d} \int_0^t \mathcal{S}_{t-s} V_j^{\mathcal{H}}(x^{\mathcal{H}}(t,x_0)) \circ \mathrm{d}B_t^j, \tag{5.47}$$

and similarly for the cubature approximations. Setting

$$Q_{(\Delta t)}^{\mathcal{H}} f(x_0) := \sum_{i=1}^{N} \lambda_i f(x^{\mathcal{H}}(\Delta t, x_0; \omega_i^{(\Delta t)})), \tag{5.48}$$

Theorem 5.6 yields that $Q_{(\Delta t)}^{\mathcal{H}}\psi_{\mathcal{H}}(y) \leqslant \exp(C\Delta t)\psi_{\mathcal{H}}(y)$, and from $\psi_{\mathcal{H}}(x) = \psi(x)$ for $x \in H$ we obtain that for $x \in H$,

$$Q_{(\Delta t)}\psi(x_0) = \sum_{i=1}^{N} \lambda_i \rho(\|\pi x^{\mathcal{H}}(\Delta t, x_0; \omega_i^{(\Delta t)})\|^2) \leqslant \sum_{i=1}^{N} \lambda_i \rho(\|x^{\mathcal{H}}(\Delta t, x_0; \omega_i^{(\Delta t)})\|_{\mathcal{H}}^2)$$

$$= Q_{(\Delta t)}^{\mathcal{H}}\psi_{\mathcal{H}}(x_0) \leqslant \exp(C\Delta t)\psi_{\mathcal{H}}(x_0) = \exp(C\Delta t)\psi(x_0). \tag{5.49}$$

The result is thus proved. □

5.2 Convergence estimates of cubature schemes

We are now ready to prove rates of convergence for cubature on Wiener space on weighted spaces. We shall only prove these results in the infinite-dimensional setting; corresponding results in finite dimensions are obtained in a similar manner. Consider therefore the setting of Section 3.2, i.e., let Assumptions 3.12, 3.13 and 3.15 be satisfied.

5.2.1 Taylor expansion of cubature approximations

We prove a local expansion of cubature approximations in weighted spaces. For the definition of the terms used, see also Sections 1.1.1 and 1.1.2. This can be seen as generalising related results in [6, proof of Theorem 4.4] to weighted spaces.

Theorem 5.8. *Assume that the cubature formula* $(\omega_i^{(\Delta t)}, \lambda_i)_{i=1,\dots,N}$ *is of odd order* $m = 2k+1$. *For* $f \in \mathcal{B}_{2(k+1)}^{\psi_{\ell-(k+1)}^{(n)}}((H_{\ell-(k+1)})_w)$, $k+1 \leqslant \ell \leqslant \ell_0$, $s \geqslant 2(k+2)$,

$$Q_{(\Delta t)} f = \sum_{j=0}^{k} \frac{(\Delta t)^j}{j!} \mathcal{G}^j f + (\Delta t)^{k+1} \hat{R}_{\Delta t,k} f, \tag{5.50}$$

where the linear operator $\hat{R}_{\Delta t,k} : \mathcal{B}_{2(k+1)}^{\psi_{\ell-(k+1)}^{(n)}}((H_{\ell-(k+1)})_w) \to \mathcal{B}^{\psi_\ell^n}((H_\ell)_w)$ *satisfies*

$$\|\hat{R}_{\Delta t,k} f\|_{\psi_\ell^n} \leqslant C_T \|f\|_{\psi_{\ell-(k+1)}^{(n)}, 2(k+1)} \quad \text{for } \Delta t \in [0, T] \tag{5.51}$$

for a constant $C_T > 0$ *independent of* f.

Proof. Under the assumptions on the vector fields, we have for every $f \in \mathcal{A}(H_{\ell-(k+1)})$ the Taylor expansion

$$\begin{aligned} f(x(\Delta t, x_0; (\omega_i^{(\Delta t)})) \\ = \sum_{(i_1,\dots,i_k)\in\mathcal{A}_m} V_{i_1} \dots V_{i_k} f(x_0) I_{\Delta t}^{(i_1,\dots,i_k)}(\omega_i^{(\Delta t)}) + \hat{R}_{\Delta t,k}^i f(x_0), \end{aligned} \tag{5.52}$$

where we define the iterated integrals by

$$\begin{aligned} I_{\Delta t}^{(i_1,\dots,i_k)}(\omega_i^{(\Delta t)}, g) \\ := \int_{0<t_1<\dots<t_k<\Delta t} g(x(t_1, x_0; \omega_i^{(\Delta t)})) \mathrm{d}\omega_i^{(t),i_1}(t_1) \dots \mathrm{d}\omega_i^{(t),i_k}(t_k), \end{aligned} \tag{5.53}$$

$I_{\Delta t}^{(i_1,\dots,i_k)}(\omega_i^{(\Delta t)}) := I_{\Delta t}^{(i_1,\dots,i_k)}(\omega_i^{(\Delta t)}, 1)$, the remainder term $\hat{R}_{\Delta t,k}^i f$ satisfies

$$\hat{R}_{\Delta t,k}^i f(x) = \sum_{\substack{(i_1,\dots,i_k)\in\mathcal{A}_m \\ (i_0,i_1,\dots,i_k)\notin\mathcal{A}_m}} I_{\Delta t}^{(i_0,\dots,i_k)}(\omega_i^{(\Delta t)}, f_{(i_0,\dots,i_k)}), \tag{5.54}$$

and we set $\beta_0(x) := Ax + V_0(x)$, $\beta_j(x) := V_j(x)$, $j = 1,\dots,d$, and $f_{(i_0,\dots,i_k)} := \beta_{i_0} \dots \beta_{i_k} f$, $(i_0,\dots,i_k) \in \{0,\dots,d\}^{k+1}$. Summing up, the scaling of the cubature paths proves that the remainder term is as claimed. To see that the initial terms have the given form, we use the order $2k+1$ of the cubature and the explicit formula of $\mathcal{G}$ from Theorem 3.25. A density argument proves the result. □

5.2.2 The rate of convergence

Corollary 5.9. *For $f \in \mathcal{B}_{2(k+1)}^{\psi_{\ell-(k+1)}^{(n)}}((H_{\ell-(k+1)})_w)$, $k+1 \leqslant \ell \leqslant \ell_0$, $n > 2(k+1)$, $2(k+1) \leqslant k_0$,*

$$\|P_T f - Q^n_{(T/n)} f\|_{\psi^n_\ell} \leqslant C_T n^{-k} \|f\|_{\psi^{(n)}_{\ell-(k+1)}, 2(k+1)} \tag{5.55}$$

with a constant C_T independent of f.

Proof. The local estimate follows from a combination of Corollary 3.28 and Theorem 5.8. The stability of $Q_{(T/n)}$ from Corollary 5.7 and Lemma 4.10 prove the claim. □

5.2.3 Smoothing effects under the UFG condition

Under the UFG condition, it is proved in [65, 70] that even for nonsmooth payoffs f, we can obtain the optimal rate of convergence by using non-equidistant grids due to the smoothing effects of $P_t f$ in the direction of the vector fields V_j. The aim of this section is to show how a corresponding result can also be obtained even for growing payoffs. In particular, we will focus on exponentially growing payoffs through the choice of the weight function $\cosh(\alpha|x|)$. This has important applications in mathematical finance, where one frequently models the log price as the solution of a stochastic differential equation, and thus, all payoffs will be a function of the exponential of the stochastic process. Other weight functions are equally possible.

Consider the finite dimensional situation, $H = \mathbb{R}^N$ for some $N \in \mathbb{N}$, and $A = 0$. Suppose that all vector fields $V_j \colon \mathbb{R}^N \to \mathbb{R}^N$ are bounded and C^∞-bounded. We choose the D-admissible weight function $\psi(x) := \cosh(\alpha|x|)$ for some $\alpha > 0$. Here, $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^N$.

Proposition 5.10. *For any $\alpha > 0$, there exists $C > 0$ such that*

$$\mathbb{E}[\cosh(\alpha|x(t, x_0)|)] \leqslant \exp(Ct)\cosh(\alpha|x|). \tag{5.56}$$

Proof. For any $k \in \mathbb{N}$,

$$D^k \cosh(\alpha|x|)(h_1, \ldots, h_k) \leqslant C_k \cosh(\alpha|x|) \prod_{j=1}^{k} |h_j|. \tag{5.57}$$

With $\mathcal{G}$ the generator of P_t, we thus obtain from the boundedness of the vector fields that $\mathcal{G}\cosh(\alpha|x|) \leqslant C\cosh(\alpha|x|)$. It follows that

$$\begin{aligned} \mathbb{E}[\cosh(\alpha|x(t,x_0)|)] &= \cosh(\alpha|x_0|) + \int_0^t \mathbb{E}[\mathcal{G}\cosh(\alpha|x(s,x_0|)]\mathrm{d}s \\ &\leqslant \cosh(\alpha|x_0|) + \int_0^t C\mathbb{E}[\cosh(\alpha|x(s,x_0)|)]\mathrm{d}s. \end{aligned} \tag{5.58}$$

The Gronwall inequality proves the claim. □

Corollary 5.11. *For any $p \in [1,\infty)$ and $T > 0$, there exists $C_T > 0$ such that*

$$\mathbb{E}[\cosh(\alpha|x(t,x_0)|)^p]^{1/p} \leqslant C_T \cosh(\alpha|x_0|) \quad \textit{for all } t \in [0,T]. \tag{5.59}$$

Proof. We only need to note that for any $p \in [1,\infty)$, there exists some constant $C > 0$ with $C^{-1}\cosh(pu) \leqslant \cosh(u)^p \leqslant C\cosh(pu)$ for all $u \in [0,\infty)$, and apply Proposition 5.10. □

We formulate now the ellipticity assumptions that are necessary to obtain smoothing effects. We follow [26].

The UFG condition. There exists $\ell \in \mathbb{N}$ such that for every $\alpha \in \mathcal{A}^*$, there exist $\varphi_{\alpha,\beta} \in C_b^\infty(\mathbb{R}^N)$, $\beta \in \mathcal{A}_\ell^*$, such that

$$V_{[\alpha]} = \sum_{\beta\in\mathcal{A}_\ell^*} \varphi_{\alpha,\beta} V_{[\beta]}. \tag{5.60}$$

The V0 condition. For some $\varphi_\beta \in C_b^\infty(\mathbb{R}^N)$, $\beta \in \mathcal{A}_2^*$,

$$V_0 = \sum_{\beta\in\mathcal{A}_2^*} \varphi_\beta V_{[\beta]}. \tag{5.61}$$

Theorem 5.12. *Assume that the UFG and V0 conditions are satisfied. Then, for any $f \in C_b^\infty(\mathbb{R}^N)$, any $k, m \geqslant 0$ and any $i_1, \dots, i_{k+m} = 0, 1, \dots, d$,*

$$\|V_{i_1}\dots V_{i_k} P_t V_{i_{k+1}} \dots V_{i_{k+m}} f\|_\psi \leqslant C t^{-\deg(i_1,\dots,i_{k+m})/2} \|f\|_\psi. \tag{5.62}$$

Proof. We apply [63, Corollary 2.17] to obtain that for each $x_0 \in \mathbb{R}^N$, there exists a real-valued random variable $\pi_t^{x_0}$, depending on k and $i_1, \dots, i_{k+m}$, with

$$V_{i_1}\dots V_{i_k} P_t V_{i_{k+1}} \dots V_{i_{k+m}} f(x_0) = \mathbb{E}[f(x(t,x_0))\pi_t^{x_0}]. \tag{5.63}$$

Furthermore, for each $p \in [1, \infty)$, there exists a constant $C > 0$ independent of t with

$$\sup_{x_0 \in \mathbb{R}^N} \mathbb{E}[|\pi_t^{x_0}|^p] \leqslant C t^{-\deg(i_1, \dots, i_{k+m})/2}. \tag{5.64}$$

It follows that for $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$,

$$\begin{aligned} \|V_{i_1} \dots V_{i_k} P_t V_{i_{k+1}} \dots V_{i_{k+m}} f\|_\psi &\leqslant \sup_{x_0 \in \mathbb{R}^N} \psi(x_0)^{-1} \mathbb{E}[|f(x(t, x_0))| \cdot |\pi_t^{x_0}|] \\ &\leqslant \|f\|_\psi \sup_{x_0 \in \mathbb{R}^N} \psi(x_0)^{-1} \times \\ &\quad \times \mathbb{E}[\psi(x(t, x_0))^p]^{1/p} \cdot \mathbb{E}[|\pi_t^{x_0}|^q]^{1/q} \\ &\leqslant C t^{-\deg(i_1, \dots, i_{k+m})/2} \|f\|_\psi, \end{aligned} \tag{5.65}$$

where we apply Corollary 5.11. □

Corollary 5.13. *Assume that the UFG and V0 conditions are satisfied. Then, for any mesh* $0 = t_0 < \cdots < t_n = T$ *and* $f \in C_b^\infty(\mathbb{R}^N)$,

$$\begin{aligned} &\|P_T f - Q_{t_1 - t_0} \dots Q_{t_n - t_{n-1}} f\|_\psi \\ &\qquad \leqslant C |f|_{\psi,1} \Big((t_n - t_{n-1})^{1/2} + \sum_{i=1}^{n-1} \frac{(t_i - t_{i-1})^{(m+1)/2}}{(T - t_i)^{m/2}} \Big). \end{aligned} \tag{5.66}$$

Here, we use the vector of weight functions (ψ, ψ).

Proof. We proceed as in the proofs of [70, Proposition 3.6], [26, Lemma 3.5]. First, note that

$$\begin{aligned} \|P_{\Delta t} f - f\|_\psi &\leqslant \sup_{x_0 \in \mathbb{R}^N} \psi(x_0)^{-1} \mathbb{E}[|f(x(\Delta t, x_0)) - f(x_0)|] \\ &\leqslant \sup_{x_0 \in \mathbb{R}^N} \psi(x_0)^{-1} \times \\ &\qquad \mathbb{E}[\sup_{s \in [0,1]} |\nabla f(s x(\Delta t, x_0) + (1 - s) x_0)| \cdot |x(\Delta t, x_0) - x_0|] \\ &\leqslant |f|_{\psi,1} \sup_{x_0 \in \mathbb{R}^N} \psi(x_0)^{-1} \mathbb{E}[\sup_{s \in [0,1]} \psi(s x(\Delta t, x_0) + (1 - s) x_0)^2]^{1/2} \times \\ &\qquad \times \mathbb{E}[|x(\Delta t, x_0) - x_0|^2]^{1/2}. \end{aligned} \tag{5.67}$$

As $|sx(\Delta t, x_0) + (1-s)x_0| \leqslant \max(|x(\Delta t, x_0)|, |x_0|)$ for all $s \in [0,1]$ and cosh is monotonic on $[0,\infty)$, we see that Corollary 5.11 yields

$$|P_{\Delta t} f - f|_\psi \leqslant C(\Delta t)^{1/2} |f|_{\psi,1}. \tag{5.68}$$

By Theorem 5.12, we obtain

$$\begin{aligned} \|(P_{\Delta t} - Q_{\Delta t}) P_{T-t} f\|_\psi &\leqslant (\Delta t)^{(m+1)/2} \sum_{\substack{(i_1,\dots,i_k)\in\mathcal{A}_m \\ (i_0,i_1,\dots,i_k)\notin\mathcal{A}_m}} \|V_{i_0} V_{i_1} \dots V_{i_k} P_{T-t} f\|_\psi \\ &\leqslant C(\Delta t)^{(m+1)/2} (T-t)^{-m/2} |f|_{\psi,1}. \end{aligned} \tag{5.69}$$

Summing up in the usual manner, the claim follows. □

Corollary 5.14. *Under the UFG and V0 assumptions, the cubature method converges of optimal order for* $f \in \mathcal{B}_1^\psi(\mathbb{R}^N)$ *on graded meshes such as the ones suggested in [70, Example 3.7].*

Proof. This follows directly from Corollary 5.13 together with the density of $C_b^\infty(\mathbb{R}^N)$ in $\mathcal{B}^\psi(\mathbb{R}^N)$. □

Chapter 6

Splitting And Cubature For The Stochastic Navier-Stokes Equations

The issue of turbulence in fluid flows is an essentially unsolved problem. From the perspective of numerical analysis, its main difficulty is that a direct numerical simulation (DNS), resolving all relevant temporal and spatial scales, is unavailable for many practically relevant geometries. Hence, we can only use results from underresolved simulations, which are often useless due to their severely reduced accuracy.

This has led to models dealing with the closure problem, see, e.g., [88, 14]. These models deal with underresolution by introducing an approximation of the effects taking place on scales smaller than those that are resolved.

We are concerned with a different approach to turbulence modelling. In the last years, the introduction of noise into the equations of fluid dynamics has become the focus of research (see, e.g., [11, 61, 28, 76, 1]). In particular, Hairer and Mattingly proved in [47, 48] that the stochastic Navier-Stokes equations on the two-dimensional torus with finite-dimensional, additive noise have ergodic dynamics, and estimated the rate of convergence to the invariant measure.

We consider the problem of weak approximation of the solution of the stochastic Navier-Stokes equations. In contrast to [51], we propose a simulation scheme, based either on splitting or cubature approximations. The advantage of such an approach is that it is trivial to parallelise, as every path can be simulated independently. In the case of splitting schemes, we can furthermore reuse well-tested, robust and fast solvers for the deterministic Navier-Stokes or Euler equations to obtain solvers for the stochastic Navier-Stokes equations with minimal effort.

To derive rates of convergence, we employ the theory of Chapters 2 and

3. While we are unable to prove rates of convergence on the continuous level, a discretisation by a spectral Galerkin scheme allows us to obtain an optimal convergence estimate in time.

This chapter is organised as follows. In Section 6.1, we recall the definition of the stochastic Navier-Stokes equations in the setting of Hairer and Mattingly and consider them from the perspective of the results of Chapter 2. Our analysis profits greatly from the fundamental results shown by Hairer and Mattingly in [73, 47, 48]. Section 6.2 is devoted to the derivation of estimates for the error done by a spectral Galerkin approximation. Section 6.3 presents the main results of this paper, estimates for full discretisations of the stochastic Navier-Stokes equations by splitting and cubature schemes. In Section 6.4, we present the results of numerical calculations for a model problem with ergodic dynamics, and in Section 6.5, we sum up our results.

6.1 The stochastic Navier-Stokes equations and weighted spaces

Consider, as in [47, 48], the vorticity formulation of the stochastic Navier-Stokes equations on the two-dimensional torus $\mathbb{T}^2$,

$$\begin{aligned} \text{(6.1)}\quad \mathrm{d}w(t,w_0) &= \nu\Delta w(t,w_0)\mathrm{d}t + B(\mathcal{K}w(t,w_0), w(t,w_0))\mathrm{d}t + \sum_{j=1}^{d} q_j f_{k_j} \mathrm{d}W_t^j, \\ w(0,w_0) &= w_0. \end{aligned}$$

The state space is $\mathbb{L}^2$, the space of mean zero square integrable functions, with norm $\|\cdot\|$ and scalar product $\langle\cdot,\cdot\rangle$. Furthermore, Δ is the Laplacian, $\mathcal{K}$ the inverse of the rotation $\nabla \wedge u = \partial_2 u_1 - \partial_1 u_2$ in the space of divergence free vector fields, $\nabla \wedge (\mathcal{K}w) = w$ and $\nabla \cdot \mathcal{K}w = 0$, $B(u,w) = -(u\cdot\nabla)w$ the Navier-Stokes nonlinearity, and $(W_t^j)_{j=1,\cdots,d}$ a d-dimensional Brownian motion. The q_j are nonvanishing real numbers, $q_j \in \mathbb{R}\setminus\{0\}$, and f_k are the orthonormal eigenfunctions of Δ on $\mathbb{T}^2$,

$$\text{(6.2)}\qquad f_k(x) = \begin{cases} (2\pi^2)^{-1/2}\sin(k\cdot x), & k\in\mathbb{Z}_+^2, \\ (2\pi^2)^{-1/2}\cos(k\cdot x), & \text{else,} \end{cases}$$

where

$$\text{(6.3)}\qquad \mathbb{Z}_+^2 := \left\{k=(k_1,k_2)\in\mathbb{Z}^2\colon \text{either } k_2>0, \text{ or } k_2=0 \text{ and } k_1>0\right\}.$$

Solvability of this equation is settled in [73].

We also define the Sobolev spaces of divergence-free, mean zero functions $\mathbb{H}^s$, $s \in \mathbb{R}$, with norm $\|\sum_{k\in\mathbb{Z}^2} w_k f_k\|_s := \sqrt{\sum_{k\in\mathbb{Z}^2}(k_1^2 + k_2^2)^s |w_k|^2}$, which is non-degenerate due to the mean zero condition (the term for $k = (0,0)$ vanishes). We note, in particular, that

$$-\langle \Delta w, w\rangle = \|w\|_1^2. \tag{6.4}$$

Similarly as in [48, Section 5.3], we introduce the weight function $\psi_\eta(w) := \exp(\eta\|w\|^2)$ with some $\eta > 0$ and consider the weighted space $\mathcal{B}^{\psi_\eta}(\mathbb{L}^2_w)$.

Proposition 6.1. *The Markov semigroup $(P_t)_{t\geqslant 0}$ defined through $P_t f(w_0) := \mathbb{E}[f(w(t,w_0))]$ is strongly continuous on $\mathcal{B}^{\psi_\eta}(\mathbb{L}^2_w)$ for $\eta > 0$ small enough.*

Proof. This follows from Theorem 3.10 and [48, Theorem A.3]. A very similar result is proved in [48, Theorem 5.10]. □

Contrary to the approach used in [80, 105], we are not able to split this problem into a part corresponding fully to the drift and another for the diffusion: the process $y(t,w_0)_t := w_0 + \sum_{j=1}^d q_j W_t^j f_{k_j}$ corresponding to the diffusion does not satisfy $\mathbb{E}[\psi_\eta(y(t,w_0))] \leqslant K\psi_\eta(w_0)$ with $K > 0$ constant for t small enough, which means that we cannot use standard Ninomiya-Victoir splittings.

Thus, we split up the equation differently. For a given $\varepsilon \in (0,1)$, we introduce the deterministic vorticity equation,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} w^1(t,w_0) &= (1-\varepsilon)\nu\Delta w^1(t,w_0) + B(\mathcal{K}w^1(t,w_0), w^1(t,w_0)), \\ w^1(0,w_0) &= w_0, \end{aligned} \tag{6.5}$$

and a stochastic heat equation defining an Ornstein-Uhlenbeck process on $\mathbb{L}^2$,

$$\mathrm{d}w^2(t,w_0) = \varepsilon\nu\Delta w^2(t,w_0)\mathrm{d}t + \sum_{j=1}^d q_j f_{k_j}\mathrm{d}W_t^j, \quad w^2(0,w_0) = w_0. \tag{6.6}$$

Define by $P_t^1 f(w_0) := \mathbb{E}[f(w^1(t,w_0))]$ and $P_t^2 f(w_0) := \mathbb{E}[f(w^2(t,w_0))]$ the Markov semigroups corresponding to w^1 and w^2.

Lemma 6.2. *For $\eta > 0$, $(P_t^1)_{t\geqslant 0}$ defines a strongly continuous semigroup on $\mathcal{B}^{\psi_\eta}(\mathbb{L}^2_w)$ with $\|P_t^1\|_{L(\mathcal{B}^{\psi_\eta}(\mathbb{L}^2_w))} \leqslant 1$.*

Proof. The strong continuity is obtained using Theorem 3.10. The necessary bounds are proved by applying [48, Theorem A.3]; see also [48, Theorem 5.10].

The deterministic vorticity equations have $\mathbb{L}^2$-contractive dynamics, as

$$\begin{aligned}\|w^1(t,w_0)\|^2 &= \|w_0\|^2 \\ &\quad + \int_0^t \langle \varepsilon\nu\Delta w^1(s,w_0) + B(\mathcal{K}w^1(s,w_0), w^1(s,w_0)), w^1(s,w_0)\rangle \mathrm{d}s \\ &\leqslant \|w_0\|^2, \end{aligned} \tag{6.7}$$

which yields the norm bound. The proof is thus complete. □

The cumbersome proof of the following proposition is postponed to Section 6.6.

Proposition 6.3. *If $\eta > 0$ is small enough, there exists $\omega > 0$ such that the process $t \mapsto \exp(-\omega t)\psi_\eta(w^2(t,w_0))$ is a positive supermartingale, i.e.*

$$\mathbb{E}[\psi_\eta(w^2(t,w_0)] \leqslant \exp(\omega t)\psi_\eta(w^2(t,w_0)). \tag{6.8}$$

Lemma 6.4. *For $\eta > 0$ small enough, $(P_t^2)_{t\geqslant 0}$ is strongly continuous on $\mathcal{B}^{\psi_\eta}(\mathbb{L}^2_w)$ with bound $\|P_t^2\|_{L(\mathcal{B}^{\psi_\eta}(\mathbb{L}^2_w))} \leqslant \exp(\omega t)$.*

Proof. Clear from Proposition 6.3 (see also Example 3.6). □

6.2 Spectral Galerkin approximations

For the stochastic Navier-Stokes equations, we cannot argue directly as in Chapters 4 or 5: there do not appear to be useful weight functions on spaces of more regular functions (such spaces are nevertheless invariant with respect to the dynamics of (6.1); see [73, Section 3.4] in this regard). We will therefore settle with a weaker result: we shall prove that spectral Galerkin approximations using Fourier modes up to degree N yield a convergent scheme, which can then be approximated by a splitting or a cubature scheme with N-dependent error bound. As the N-dependence of the estimate is given explicitly, we can derive convergent schemes by choosing the time step size small enough in relation to N.

Consider therefore the spectral Galerkin approximation of (6.1),

$$\begin{aligned} \mathrm{d}w_N(t, w_0) &= \nu \Delta w_N(t, w_0)\mathrm{d}t \\ &\quad + \pi_N B(\mathcal{K} w_N(t, w_0), w_N(t, w_0))\mathrm{d}t + \sum_{j=1}^{d} q_j f_{k_j} \mathrm{d}W_t^j, \end{aligned} \tag{6.9a}$$

$$w_N(0, w_0) = \pi_N w_0, \tag{6.9b}$$

see also [33], where $\pi_N : \mathbb{L}^2 \to \mathbb{L}^2$ is the projection onto the space $\mathcal{H}_N$ of tensor products of trigonometric polynomials of degree N,

$$\mathcal{H}_N := \operatorname{span}\left\{ f_k : \max_{i=1,2} |k_i| \leqslant N \right\}, \tag{6.10}$$

and N is assumed to be large enough so that $f_{k_j} \in \mathcal{H}_N$ for $j = 1, \dots, d$. Its split semigroups are given by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} w_N^1(t, w_0) &= \pi_N B(\mathcal{K} w_N^1(t, w_0), w_N^1(t, w_0)), \\ w_N^1(0, w_0) &= w_0, \qquad \text{and} \end{aligned} \tag{6.11}$$

$$\begin{aligned} \mathrm{d}w_N^2(t, w_0) &= \nu \Delta w_N^2(t, w_0)\mathrm{d}t + \sum_{j=1}^{d} q_j f_{k_j} \mathrm{d}W_t^j, \\ w_N^2(0, w_0) &= w_0. \end{aligned} \tag{6.12}$$

The choice $\varepsilon = 1$ made here is not admissible above: in the space continuous setting, the results from [48] do not allow us to apply Theorem 3.10 to conclude that P_t^1 is strongly continuous for this choice. (Note, however, that Theorem 2.31 might be applicable, as the velocity formulation admits solutions in L^2, which, by [48, equation (38)], implies solvability for the initial vorticity in $\mathbb{H}^{-1}$). As $\mathcal{H}_N$ is finite-dimensional, however, we do not have to distinguish between different topologies, and it follows that the Markov semigroups P_t^N, $P_t^{N,1}$ and $P_t^{N,2}$ of w_N, w_N^1 and w_N^2 are strongly continuous on $\mathcal{B}^{\psi_\eta}(\mathcal{H}_N)$ if $\eta > 0$ is small enough. In case that a solver for deterministic Navier-Stokes equations is available, it is also possible to use $\varepsilon < 1$ here (the case $\varepsilon = 1$ corresponds to splitting up into a deterministic Euler equation).

We now estimate the error of the spectral Galerkin approximation.

Proposition 6.5. *For any $\alpha > 0$, $w_0 \in \mathbb{L}^2$ and $t > 0$,*

$$\|w(t,w_0) - w_N(t,w_0)\|^2 \leqslant CN^{-1}\|w(t,w_0)\|_1^2 + C_\alpha N^{-1} \exp\left(C_\alpha t + \frac{\alpha}{2}\int_0^t \|w(\sigma,w_0)\|_1^2 \mathrm{d}\sigma\right) \int_0^t \|w(s,w_0)\|_1^4 \mathrm{d}s. \tag{6.13}$$

Proof. Let $e_N(t) := \pi_N w(t,w_0) - w_N(t,w_0) \in \mathcal{H}_N$ and $\eta_N(t) := w(t,w_0) - \pi_N w(t,w_0)$. Then,

$$\begin{aligned} \mathrm{d}e_N(t) = \nu\Delta e_N(t) &+ \pi_N\left(B(\mathcal{K}w_N(t,w_0), e_N(t)) + B(\mathcal{K}e_N(t), \pi_N w(t,w_0))\right)\mathrm{d}t \\ &+ \pi_N\left(B(\mathcal{K}\pi_N w(t,w_0), \eta_N(t)) + B(\mathcal{K}\eta_N(t), w(t,w_0))\right)\mathrm{d}t. \end{aligned} \tag{6.14}$$

It results that

$$\begin{aligned} \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|e_N(t)\|^2 = -\nu\|e_N(t)\|_1^2\mathrm{d}t &+ \langle B(\mathcal{K}e_N(t), \pi_N w(t,w_0)), e_N(t)\rangle \\ &+ \langle B(\mathcal{K}\pi_N w(t,w_0), \eta_N(t)) + B(\mathcal{K}\eta_N(t), w(t,w_0)), e_N(t)\rangle. \end{aligned} \tag{6.15}$$

We now proceed similarly as in [47, Proof of Lemma 4.10, point 3]. For any $\delta > 0$, we estimate

$$|\langle B(\mathcal{K}h, w), \zeta\rangle| \leqslant \delta\|\zeta\|_1^2 + \frac{C}{4\alpha^2\delta}\|\zeta\|^2 + \frac{\alpha}{4}\|w\|_1^2\|h\|^2. \tag{6.16}$$

This yields

$$\begin{aligned} |\langle B(\mathcal{K}e_N(t), \pi_N w(t,w_0)), e_N(t)\rangle| &\leqslant \delta\|e_N(t)\|_1^2 + \frac{C}{4\alpha^2\delta}\|e_N(t)\|^2 \\ &+ \frac{\alpha}{4}\|\pi_N w(t,w_0)\|_1^2\|e_N(t)\|^2 \quad \text{and} \end{aligned} \tag{6.17}$$

$$\begin{aligned} |\langle B(\mathcal{K}\eta_N(t), w(t,w_0)), e_N(t)\rangle| &\leqslant \delta\|e_N(t)\|_1^2 + \frac{C}{4\alpha^2\delta}\|e_N(t)\|^2 \\ &+ \frac{\alpha}{4}\|w(t,w_0)\|_1^2\|\eta_N(t)\|^2. \end{aligned} \tag{6.18}$$

For the final term, we apply

$$|\langle B(\mathcal{K}h, w), \zeta\rangle| \leqslant \delta\|\zeta\|_1^2 + \frac{C}{4\delta}\|h\|_1^2\|w\|^2, \tag{6.19}$$

which shows

$$\begin{aligned} |\langle B(\mathcal{K}\pi_N w(t,w_0), \eta_N(t)), e_N(t)\rangle| &\leqslant \delta\|e_N(t)\|_1^2 \\ &+ \frac{C}{4\delta}\|\pi_N w(t,w_0)\|_1^2\|\eta_N(t)\|^2. \end{aligned} \tag{6.20}$$

Choosing $\delta = \frac{\nu}{6}$ and combining the above estimates yields

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|e_N(t)\|^2 &\leqslant -\frac{\nu}{2}\|e_N(t)\|_1^2 + \frac{3C}{\alpha^2\nu}\|e_N(t)\|^2 + \frac{\alpha}{4}\|\pi_N w(t,w_0)\|_1^2\|e_N(t)\|^2 \\ &\quad + \left(\frac{\alpha}{4}\|w(t,w_0)\|_1^2 + \frac{3C}{\nu}\|\pi_N w(t,w_0)\|_1^2\right)\|\eta_N(t)\|^2. \end{aligned} \tag{6.21}$$

Using $\|\pi_N w\|_1 \leqslant \|w\|_1$, we obtain

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|e_N(t)\|^2 \leqslant \left(C_\alpha + \frac{\alpha}{2}\|w(t,w_0)\|_1^2\right)\frac{1}{2}\|e_N\|^2 \\ + C_\alpha\|w(t,w_0)\|_1^2\|\eta_N(t)\|^2.\end{aligned} \tag{6.22}$$

An application of Gronwall's inequality yields, as $e_N(0) = 0$,

$$\begin{aligned}\frac{1}{2}\|e_N(t)\|^2 \leqslant \int_0^t C_\alpha\|w(s,w_0)\|_1^2\|\eta_N(s)\|^2 \times \\ \times \exp\left(C_\alpha(t-s) + \frac{\alpha}{2}\int_s^t\|w(\sigma,w_0)\|_1^2\mathrm{d}\sigma\right)\mathrm{d}s.\end{aligned} \tag{6.23}$$

As $\|w-\pi_N w\| \leqslant CN^{-1}\|w\|_1$, we see that $\|\eta_N(t)\| \leqslant CN^{-1}\|w(t,w_0)\|_1$, whence

$$\begin{aligned}\frac{1}{2}\|e_N(t)\|^2 &\leqslant C_\alpha N^{-1}\int_0^t\|w(s,w_0)\|_1^4 \times \\ &\qquad \times \exp\left(C_\alpha(t-s) + \frac{\alpha}{2}\int_s^t\|w(\sigma,w_0)\|_1^2\mathrm{d}\sigma\right)\mathrm{d}s \\ &\leqslant C_\alpha N^{-1}\exp\left(C_\alpha t + \frac{\alpha}{2}\int_0^t\|w(\sigma,w_0)\|_1^2\mathrm{d}\sigma\right)\int_0^t\|w(s,w_0)\|_1^4\mathrm{d}s.\end{aligned} \tag{6.24}$$

The result follows due to

$$\|w(t,w_0) - w_N(t,w_0)\| \leqslant \|e_N(t)\| + CN^{-1}\|w(t,w_0)\|_1. \tag{6.25}$$ □

Corollary 6.6. *For any $w_0 \in \mathbb{H}^1$ and $T \geqslant 0$, there exists a constant $C = C_{w_0,T} > 0$ such that for any $t \in [0,T]$,*

$$\mathbb{E}\left[\|w(t,w_0) - w_N(t,w_0)\|^2\right] \leqslant CN^{-1}. \tag{6.26}$$

Proof. From Proposition 6.5 and an application of the Cauchy-Schwarz inequality, we see that we need to prove

$$\mathbb{E}[\|w(t,w_0)\|_1^2] + \mathbb{E}\left[\exp\left(\alpha \int_0^t \|w(\sigma,w_0)\|_1^2 \mathrm{d}\sigma\right)\right] + \mathbb{E}\left[\left(\int_0^t \|w(s,w_0)\|_1^4 \mathrm{d}s\right)^2\right] \leqslant K \tag{6.27}$$

for all $t \in [0,T]$ with some $K = K_{t,w_0} > 0$. For the first and third term, this follows from [73, Theorem 3.7], and for the second, from [47, Lemma 4.10]. □

Remark 6.7. Actually, it seems quite plausible here that the assumption $w_0 \in \mathbb{H}^1$ is too strong. Indeed, the results in [74] show that if $w_0 \in \mathbb{L}^2$, then $w(t,w_0) \in \mathbb{H}^s$ for all $s > 0$ for subsequent times, and [75, Lemma A.3] gives some quantitative estimates. It remains unclear to us however how this can be used to prove an estimate for $\mathbb{E}\left[\left(\int_0^t \|w(s,w_0)\|_1^4 \mathrm{d}s\right)^2\right]$.

The estimate from Corollary 6.6 allows us to estimate the pointwise approximation error of the weak approximation of the stochastic Navier-Stokes equation by the spectral Galerkin scheme.

Theorem 6.8. *Assume* $\varphi \in \mathcal{B}^{\psi_\eta}(\mathbb{L}^2_w) \cap \mathrm{C}^1(\mathbb{L}^2)$ *with*

$$C_\varphi := \sup_{w \in \mathbb{L}^2} \psi_{\tilde{\eta}}(w)^{-1} \|D\varphi(w)\| < \infty \tag{6.28}$$

for some $\tilde{\eta} \in [0, \eta/2]$. *Then, for* $w \in \mathbb{H}^1$ *and* $T \geqslant 0$, *there exists a constant* $C = C_{w,T,\varphi}$ *such that for all* $t \in [0,T]$,

$$|P_t\varphi(w) - P_t^N(\varphi|_{\mathcal{H}_N})(w)| \leqslant CN^{-1}. \tag{6.29}$$

Proof. By the fundamental theorem of calculus,

$$\begin{aligned} &|\varphi(w(t,w_0)) - \varphi(w_N(t,w_0))| \\ &\quad\leqslant \int_0^1 \|D\varphi(\theta w(t,w_0) + (1-\theta)w_N(t,w_0))\| \times \\ &\qquad\times \|w(t,w_0) - w_N(t,w_0))\|\mathrm{d}\theta. \end{aligned} \tag{6.30}$$

The assumption on φ together with the convexity of $w \mapsto \exp(\tilde{\eta}\|w\|^2)$ yields

$$\begin{aligned} &\|D\varphi(\theta w(t,w_0) + (1-\theta)w_N(t,w_0))\| \\ &\quad\leqslant C_\varphi \left(\exp(\tilde{\eta}\|w(t,w_0)\|^2) + \exp(\tilde{\eta}\|w_N(t,w_0)\|^2)\right). \end{aligned} \tag{6.31}$$

Therefore, the Cauchy-Schwarz inequality implies

$$|P_t\varphi(w) - P_t^N(\varphi|_{\mathcal{H}_N}(w)| \leqslant C_\varphi \mathbb{E}\left[\|w(t,w_0) - w_N(t,w_0)\|^2\right]^{1/2} \times$$
$$\times \left(\mathbb{E}[\exp(2\tilde{\eta}\|w(t,w_0)\|^2)]^{1/2} + \mathbb{E}[\exp(2\tilde{\eta}\|w_N(t,w_0)\|^2)]^{1/2}\right). \tag{6.32}$$

Note that the estimate in [47, Lemma 4.10, 1.] also holds true for $w_N(t,w_0)$ instead of $w(t,w_0)$. Therefore, Corollary 6.6 proves the claimed estimate. □

In the discrete setting, it is easy to analyse the differential operators corresponding to the split semigroups. For $k \geqslant 0$, we consider the vector of weight functions $(\psi_\eta)_{j=0,\dots,k}$, which we shall also denote by ψ_η. We denote by $\mathcal{G}_j^N$ with domain $\operatorname{dom}\mathcal{G}_j^N$ the infinitesimal generator of $(P_t^{N,j})_{t\geqslant 0}$, $j = 1, 2$, and by $\mathcal{G}^N$ with domain $\operatorname{dom}\mathcal{G}^N$ the infinitesimal generator of $(P_t^N)_{t\geqslant 0}$.

Lemma 6.9. *For any $\varepsilon > 0$,*

$$\mathcal{B}_2^{\psi_{\tilde{\eta}}}(\mathcal{H}_N) \subset \operatorname{dom}\mathcal{G}^N \cap \operatorname{dom}\mathcal{G}_1^N \cap \operatorname{dom}\mathcal{G}_2^N. \tag{6.33}$$

For $k \geqslant 0$, $\mathcal{G}^N$, $\mathcal{G}_j^N\colon \mathcal{B}_{k+2}^{\psi_{\tilde{\eta}}}(\mathcal{H}_N) \to \mathcal{B}_k^{\psi_{\tilde{\eta}+\varepsilon}}(\mathcal{H}_N)$, $j = 1, 2$, are continuous operators, and

$$\|\mathcal{G}^N\|_{L(\mathcal{B}_{k+2}^{\psi_{\tilde{\eta}}}(\mathcal{H}_N);\mathcal{B}_k^{\psi_{\tilde{\eta}+\varepsilon}}(\mathcal{H}_N))} + \|\mathcal{G}_j^N\|_{L(\mathcal{B}_{k+2}^{\psi_{\tilde{\eta}}}(\mathcal{H}_N);\mathcal{B}_k^{\psi_{\tilde{\eta}+\varepsilon}}(\mathcal{H}_N))}$$
$$\leqslant CN^2, \quad j = 1, 2. \tag{6.34}$$

Furthermore,

$$\mathcal{G}^N\varphi = \mathcal{G}_1^N\varphi + \mathcal{G}_2^N\varphi \quad \textit{for all } \varphi \in \mathcal{B}_2^{\psi_{\tilde{\eta}}}(\mathcal{H}^N). \tag{6.35}$$

Proof. For $\varphi \in \mathcal{B}_{k+2}^{\psi_{\tilde{\eta}}}(\mathcal{H}_N)$, we see by the fundamental theorem of calculus and the estimates in [48, Appendix] that with $\alpha > 0$,

$$|\mathcal{G}_1^N\varphi(w)| = |D\varphi(w)\,(\pi_N B(\mathcal{K}w, w))| \leqslant \|D\varphi(w)\| \cdot \left(N^{1+\alpha}\|w\|^2\right)$$
$$\leqslant CN^2 \exp(\varepsilon\|w\|^2)\|D\varphi(w)\|, \tag{6.36}$$

and similarly, by Itô's formula,

$$|\mathcal{G}_2^N\varphi(w)| = |D\varphi(w)\nu\Delta w + \frac{1}{2}\sum_{j=1}^{d} D^2\varphi(w)(q_j f_{k_j}, q_j f_{k_j})|$$
$$\leqslant \|D\varphi(w)\| \cdot \nu N^2\|w\| + C\|D^2\varphi(w)\|$$
$$\leqslant CN^2 \exp(\varepsilon\|w\|^2)\left(\|D\varphi(w)\| + \|D^2\varphi(w)\|\right). \tag{6.37}$$

The result for $\mathcal{G}^N$ is proved in a similar manner. The equality (6.35) is a consequence of Itô's formula if $\varphi \in C_b^\infty(\mathcal{H}_N)$, and a density argument proves it for the general case. □

6.3 Rates of convergence

We are now in the situation to prove estimates for the convergence of both splitting schemes and cubature methods.

6.3.1 Splitting methods

Lemma 6.10. *For all $k \geqslant 0$, $P_t^N \mathcal{B}_k^{\psi_{\tilde{\eta}}}(\mathcal{H}_N) \subset \mathcal{B}_k^{\psi_{\tilde{\eta}}}(\mathcal{H}_N)$, and we have the estimate $\sup_{t\in[0,T]} \|P_t^N \varphi\|_{\psi_{\tilde{\eta}},k} \leqslant K_T \|\varphi\|_{\psi_{\tilde{\eta}},k}$ with some constant K_T independent of φ.*

Proof. This is proved using similar estimates as those given in [47, Lemma 4.10, 1. and 3.]. □

Using Lemma 6.9, the method of [49] yields the following convergence estimate.

Theorem 6.11. *Let $Q_{(\Delta t)}^N := P_{\Delta t/2}^{N,1} P_{\Delta t}^{N,2} P_{\Delta t/2}^{N,1}$ denote the Strang splitting approximation of $P_{\Delta t}^N$ using $P_{\Delta t}^{N,1}$ and $P_{\Delta t}^{N,2}$. For any $\tilde{\eta} < \eta/2$, there exists $C = C_{T,\tilde{\eta}} > 0$ such that for all $\varphi \in \mathcal{B}_6^{\psi_{\tilde{\eta}}}(\mathcal{H}_N)$ and $n \in \mathbb{N}$,*

$$\|P_T^N \varphi - (Q_{(T/n)}^N)^n \varphi\|_{\psi_\eta} \leqslant C_T N^6 n^{-2} \|\varphi\|_{\psi_{\tilde{\eta}},6}. \tag{6.38}$$

Note that if $\varphi \in C^6(\mathbb{L}^2)$ is such that for some $\tilde{\eta} < \eta$,

$$\sup_{w\in\mathbb{L}^2} \psi_{\tilde{\eta}}(w)^{-1} \|D^j \varphi(w)\|_{L_j(\mathbb{L}^2;\mathbb{R})} < \infty \quad \text{for } j = 0, \ldots, 6, \tag{6.39}$$

then $\varphi|_{\mathcal{H}_N} \in \mathcal{B}_6^{\psi_\eta}(\mathcal{H}_N)$ for all $N \in \mathbb{N}$ with uniformly bounded norms. Furthermore, (6.39) with $\tilde{\eta} < \eta/2$ implies (6.28). Thus, we obtain the following result.

Corollary 6.12. *Assume that φ satisfies (6.39) with $\tilde{\eta} < \eta/2$. For any $T > 0$ and $w_0 \in \mathbb{H}^1$, there exists $C = C_{w_0,T,\varphi} > 0$ such that for all $n \in \mathbb{N}$*

$$|P_T \varphi(w_0) - (Q_{(T/n)}^N)^n \varphi|_{\mathcal{H}_N}(w_0)| \leqslant C \left(N^{-1} + N^6 n^{-2}\right). \tag{6.40}$$

Proof. The combination of Theorem 6.8 and Theorem 6.11 allows us to conclude the desired estimate. □

Remark 6.13. We see here an important advantage of the second order splitting in comparison to a possible first order splitting. There, in the second term, the instability would be of the order N^4, but the convergence would only be of first order, n^{-1}. Therefore, we can choose n significantly smaller here while still obtaining a stable method. Nevertheless, we have to stress that the given error estimate is far from what we would expect to obtain, see also the numerical results in Section 6.4.

6.3.2 Cubature methods

We define cubature approximations for the spectral Galerkin discretisation of the stochastic Navier-Stokes equations. See Section 1.1.2 for the definition of cubature paths. The approximations are given by

$$\begin{aligned} \mathrm{d}w_N(s, w_0; \omega_i^{(\Delta t)}) &= \Big(\nu\Delta w_N(s, w_0; \omega_i^{(\Delta t)}) + \pi_N B(\mathcal{K} w_N(s, w_0; \omega_i^{(\Delta t)}))\Big)\,\mathrm{d}s \\ &\quad + \sum_{j=1}^{d} q_j f_{k_j} \mathrm{d}\omega_i^{(\Delta t),j}(s). \end{aligned} \tag{6.41}$$

Here, we apply that the noise is purely additive, entailing that the Itô and Stratonovich integrals of the noise terms coincide. The cubature approximation of the Markov semigroup $P^N_{\Delta t}$ reads

$$Q^N_{(\Delta t)} f(w_0) := \sum_{i=1}^{M} \lambda_i f(w_N(\Delta t, w_0; \omega_i^{(\Delta t)})). \tag{6.42}$$

To prove stability of the cubature approximation, we require that the quadrature formula induced by the cubature scheme is symmetric, i.e., for all $i = 1, \dots, M$, there exists a unique $i' \in \{1, \dots, M\}$ such that $\lambda_i = \lambda_{i'}$ and $\omega_i^j(\Delta t) = -\omega_{i'}^j(\Delta t)$ for $j = 1, \dots, d$. This induces a corresponding symmetry for $\omega^{(\Delta t)}$. Many known cubature formulas satisfy such a property, consider, e.g., the paths given in [70]. Moreover, given an arbitrary cubature formula, it is easy to construct a symmetric one from it by adding the reflected paths.

Our use of this assumption is to prove an estimate for the moment generating function of the cubature paths at Δt.

Lemma 6.14. *Assume that the quadrature formula induced by the cubature scheme is symmetric. Then, for all continuous* $f : \mathbb{R}^d \to \mathbb{R}$,

$$\begin{aligned}\sum_{i=1}^{M} \lambda_i f(\omega_i^{(\Delta t),1}(\Delta t), \dots, \omega_i^{(\Delta t),d}(\Delta t)) \\ = \frac{1}{2}\sum_{i=1}^{M} \lambda_i \Big(f(\omega_i^{(\Delta t),1}(\Delta t), \dots, \omega_i^{(\Delta t),d}(\Delta t)) \\ + f(-\omega_i^{(\Delta t),1}(\Delta t), \dots, -\omega_i^{(\Delta t),d}(\Delta t))\Big).\end{aligned} \tag{6.43}$$

In particular, $\sum_{i=1}^{M} \lambda_i f(\omega_i^{(\Delta t),1}(\Delta t), \dots, \omega_i^{(\Delta t),d}(\Delta t)) = 0$ *if* f *is odd.*

This implies

$$\sum_{i=1}^{M} \lambda_i \exp\left(\sum_{j=1}^{d} u_j \omega_i^{(\Delta t),j}(\Delta t)\right) \leqslant \exp\left(\frac{C}{2}\Delta t \sum_{j=1}^{d} u_j^2\right). \tag{6.44}$$

Proof. The first two claims are clear. For the estimate of the moment generating function, note that, as $|\omega_i^{(\Delta t),j}(\Delta t)| \leqslant C\sqrt{\Delta t}$ and $(2\ell)! \leqslant 2^\ell \ell!$,

$$\begin{aligned}\sum_{i=1}^{M} \lambda_i \exp\Big(\sum_{j=1}^{d} u_j \omega_i^{(\Delta t),j}(\Delta t)\Big) &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i=1}^{M} \lambda_i \Big(\sum_{j=1}^{d} u_j \omega_i^{(\Delta t),j}(\Delta t)\Big)^k \\ &= \sum_{\ell=0}^{\infty} \frac{1}{(2\ell)!} \sum_{i=1}^{M} \lambda_i \Big(\sum_{j=1}^{d} u_j \omega_i^{(\Delta t),j}(\Delta t)\Big)^{2\ell} \leqslant \exp\Big(\frac{C}{2}\Delta t \sum_{j=1}^{d} u_j^2\Big),\end{aligned} \tag{6.45}$$

which proves the given estimate. □

Theorem 6.15. *Assume that the quadrature formula induced by the cubature scheme is symmetric. Then, there exist* $\eta_0 > 0$ *and* $\varepsilon > 0$*, depending only on the given problem data, but not on the discretisation parameter* N*, such that with a constant* $C > 0$ *independent of* Δt *and* N,

$$\|Q_{(\Delta t)}^N f\|_{\psi_\eta} \leqslant \exp(C\Delta t)\|f\|_{\psi_\eta} \quad \text{for } \Delta t \in (0, \varepsilon],\ \eta \in (0, \eta_0],\ \text{and } f \in \mathcal{B}^{\psi_\eta}(\mathcal{H}_N). \tag{6.46}$$

Proof. Set $w_N(s) := w_N(s, w_0; \omega_i^{(\Delta t)})$ and $V^N(w_N) := \nu\Delta w_N + \pi_N B(\mathcal{K}w_N, w_N)$. For every $\alpha \geqslant 0$,

$$\begin{aligned}\exp(\alpha s)\|w_N(s)\|^2 - \|w_N(0)\|^2 \\ &= \int_0^s \exp(\alpha r)(\alpha\|w_N(r)\|^2 + 2\langle V^N(w_N(r)), w_N(r)\rangle)\mathrm{d}r \\ &\quad + 2\sum_{j=1}^d \int_0^s \exp(\alpha r)\langle q_j f_{k_j}, w_N(r)\rangle \mathrm{d}\omega_i^{(\Delta t),j}(r).\end{aligned} \tag{6.47}$$

Applying Fubini's theorem and integration by parts to

$$\begin{aligned}\int_\sigma^\tau \exp(\alpha r)\mathrm{d}\omega_i^{(\Delta t),j}(r) &= \exp(\alpha\tau)\omega_i^{(\Delta t),j}(\tau) - \exp(\alpha\sigma)\omega_i^{(\Delta t),j}(\sigma) \\ &\quad - \alpha\int_\sigma^\tau \omega_i^{(\Delta t),j}(r)\exp(\alpha r)\mathrm{d}r,\end{aligned} \tag{6.48}$$

we obtain that

$$\begin{aligned}\int_0^s \exp(\alpha r)\langle q_j f_{k_j}, w_N(r)\rangle \mathrm{d}\omega_i^{(\Delta t),j}(r) &= \langle q_j f_{k_j}, w_N(0)\rangle \int_0^s \exp(\alpha r)\mathrm{d}\omega_i^{(\Delta t),j}(r) \\ &\quad + \int_0^s \exp(\alpha r)\int_0^r \langle q_j f_{k_j}, V^N(w_N(q))\rangle \mathrm{d}q\,\mathrm{d}\omega_i^{(\Delta t),j}(r) \\ &\quad + \sum_{i=1}^d \int_0^s \exp(\alpha r)\int_0^r \langle q_j f_{k_j}, q_i f_{k_i}\rangle \mathrm{d}\omega_i^{(\Delta t),i}(q)\mathrm{d}\omega_i^{(\Delta t),j)}(r).\end{aligned} \tag{6.49}$$

An application of Young's inequality yields

$$\begin{aligned}\int_0^s \exp(\alpha r)\langle q_j f_{k_j}, w_N(r)\rangle \mathrm{d}\omega_i^{(\Delta t),j}(r) &\leqslant \langle q_j f_{k_j}, w_N(0)\rangle \exp(\alpha s)\omega_i^{(\Delta t),j}(s) \\ &\quad + C\exp(\alpha s)\|w_N(0)\|^2\Delta t + C\exp(\alpha s)s^2 \\ &\quad + C\sqrt{\Delta t}\int_0^s \exp(\alpha q)\|V^N(w_N(q))\|_{-3}\mathrm{d}q + C\exp(\alpha s)s.\end{aligned} \tag{6.50}$$

Hence, as $\langle w_N, V^N(w_N)\rangle = -\nu\|w_N\|_1^2$ and $\|V^N(w_N)\|_{-3} \leqslant \|w_N\| + C\|w_N\|^2$,

$$\begin{aligned}\|w_N(\Delta t)\|^2 &\leqslant \big(\exp(-\alpha\Delta t) + C\Delta t\big)\|w_N(0)\|^2 \\ &+ 2\sum_{j=1}^{d}\langle q_j f_{k_j}, w_N(0)\rangle\omega_i^{(\Delta t),j}(\Delta t) + C\Delta t + C(\Delta t)^2 \\ &+ \int_0^{\Delta t}\exp(\alpha(q-\Delta t))\Big((\alpha + C\sqrt{\Delta t})\|w_N(q)\|^2 \\ &\quad + C\sqrt{\Delta t}\|w_N(q)\| - 2\nu\|w_N(q)\|_1^2\Big)\mathrm{d}q.\end{aligned} \tag{6.51}$$

Fix $\alpha = \nu$. As $\|w_N\|_1 \geqslant \|w_N\|$, we can choose $\varepsilon > 0$ such that for $\Delta t \in (0, \varepsilon]$,

$$\nu\|w_N\|^2 + C\sqrt{\Delta t}(\|w_N\| + \|w_N\|^2) - 2\nu\|w_N\|_1^2 \leqslant C\Delta t. \tag{6.52}$$

By Lemma 6.14,

$$\begin{aligned}\sum_{i=1}^{M}\lambda_i \exp\Big(2\eta\sum_{j=1}^{d}\langle q_j f_{k_j}, w_N(0)\rangle\omega_i^{(\Delta t),j}(\Delta t)\Big) & \\ \leqslant \exp\Big(\eta^2 C\Delta t\sum_{j=1}^{d}\langle q_j f_{k_j}, w_N(0)\rangle^2\Big) & \\ \leqslant \exp(\eta^2 C\Delta t\|w_N(0)\|^2). &\end{aligned} \tag{6.53}$$

Hence, for $\Delta t \in (0, \varepsilon]$,

$$\begin{aligned}&\sum_{i=1}^{M}\lambda_i \exp(\eta\|w_N(\Delta t, w_0; \omega_i^{(\Delta t),j})\|^2) \\ &\quad\leqslant \exp\Big(C\Delta t + \eta\|w_N(0)\|^2\big(\exp(-\nu\Delta t) + \eta C\Delta t\big)\Big).\end{aligned} \tag{6.54}$$

Choosing $\eta_0 > 0$ small enough, we see that

$$\exp(-\nu\Delta t) + \eta C\Delta t \leqslant 1 \quad \text{for } \Delta t \in (0, \varepsilon] \text{ and } \eta \in (0, \eta_0]. \tag{6.55}$$

The claim is thus proved. □

Remark 6.16. It is clear from the proof that a corresponding result can also be shown in the space continuous case. As remarked before in the context of the splitting scheme, however, we are not able to derive rates of convergence in this setting, which is why we focus on the space discrete case.

As it is straightforward to obtain a Taylor expansion of $Q^N_{(\Delta t)}$ by the fundamental theorem of calculus (see [70, 26, 6] and Proposition 1.3), we have the following result.

Theorem 6.17. *Fix $\eta > 0$ small enough. Given $T > 0$ and $\tilde{\eta} < \eta/2$ and assuming that m is odd, there exist constants $\varepsilon > 0$ and $C = C_{T,\tilde{\eta}} > 0$ such that for all $\varphi \in \mathcal{B}_6^{\psi_{\tilde{\eta}}}(\mathcal{H}_N)$ and $n \in \mathbb{N}$ with $T/n < \varepsilon$,*

$$\|P_T\varphi - (Q^N_{(T/n)})^n\varphi\|_{\psi_\eta} \leqslant CN^{m+1}n^{-\frac{m-1}{2}}\|\varphi\|_{\psi_{\tilde{\eta}},6}. \tag{6.56}$$

The following result is a version of Corollary 6.12 for cubature approximations.

Corollary 6.18. *Suppose m odd, and fix $\eta > 0$ small enough. Assume that φ satisfies (6.39) with $\tilde{\eta} < \eta/2$. For any $T > 0$ and $w_0 \in \mathbb{H}^1$, there exists $\varepsilon > 0$ and $C = C_{w_0,T,\varphi} > 0$ such that for all $n \in \mathbb{N}$ with $T/n < \varepsilon$,*

$$|P_T\varphi(w_0) - (Q^N_{(T/n)})^n\varphi|_{\mathcal{H}_N}(w_0)| \leqslant C\left(N^{-1} + N^{m+1}n^{-\frac{m-1}{2}}\right). \tag{6.57}$$

6.4 Numerical examples

We consider the problem of approximating (6.1) with $\nu = 10^{-2}$, $w_0 = 0$, $d = 4$, $q_j = 1$, $j = 1, \dots, 4$, and $k_1 = (1,0)$, $k_2 = (-1,0)$, $k_3 = (1,1)$ and $k_4 = (-1,-1)$. [47, Example 2.5] shows that the dynamics generated by this process are ergodic. We aim to find estimates for $\mathbb{E}[\|w(1,0)\|]$, $\mathbb{E}[\|w(1,0)\|_{-1}]$ and $\mathbb{E}[\|w(1,0)\|_{+1}]$. We remark that the first and second values are related to the mean enstrophy and energy, respectively. Furthermore, control of the $\mathbb{H}^1$ norm of $w(1,0)$ means control of the $\mathbb{H}^2$ norm of $\mathcal{K}w(1,0)$, which in turn implies that we can take point evaluations of $\mathcal{K}w(1,0)$ due to the Sobolev embedding theorems in two dimensions. This is important in the evaluation of cross correlations.

Our numerical simulations are performed using a splitting scheme, the symmetrically weighted sequential splitting

$$Q^N_{T,n} := \frac{1}{2}\left((P^{N,1}_{T/n}P^{N,2}_{T/n})^n + (P^{N,2}_{T/n}P^{N,1}_{T/n})^n\right), \tag{6.58}$$

going back at least to [103, equation (25)] and being of second order for problems that are smooth enough.

We apply a Monte Carlo method. For a single realisation, we have to solve, alternatingly, a time-dependent Euler equation and an Ornstein-Uhlenbeck equation. Note that the solution of the Ornstein-Uhlenbeck equation follows a Gaussian process, and its distribution is therefore explicitly known. To discretise the Euler equation, we apply the standard RK4 scheme. While the Heun method, i.e., an RK2 scheme, provides the correct order such that the entire approximation is of second order, see [79], it has suboptimal stability properties, leading to strong step size restrictions, see [23, Section D.2.5]. In this regard, see also [52] for issues of stability of the Euler-Maruyama scheme for equations with non-globally Lipschitz coefficients. As we apply the FFT to determine the value of $(\mathcal{K}w_N \cdot \nabla)w_N$ efficiently, we observe aliasing effects, which are reduced by the use of the 2/3 dealiasing, see [24, Section 3.3.2].

To find the expected values in the definition of $P^{N,2}_{T/n}$, we use quasi-Monte Carlo integration, applying the Sobol′ sequences of Joe and Kuo [56]. Also, instead of simulating both terms in the definition of $Q^N_{T,n}$, we use a Bernoulli random variable to generate either of them, retaining the order of the approximation.

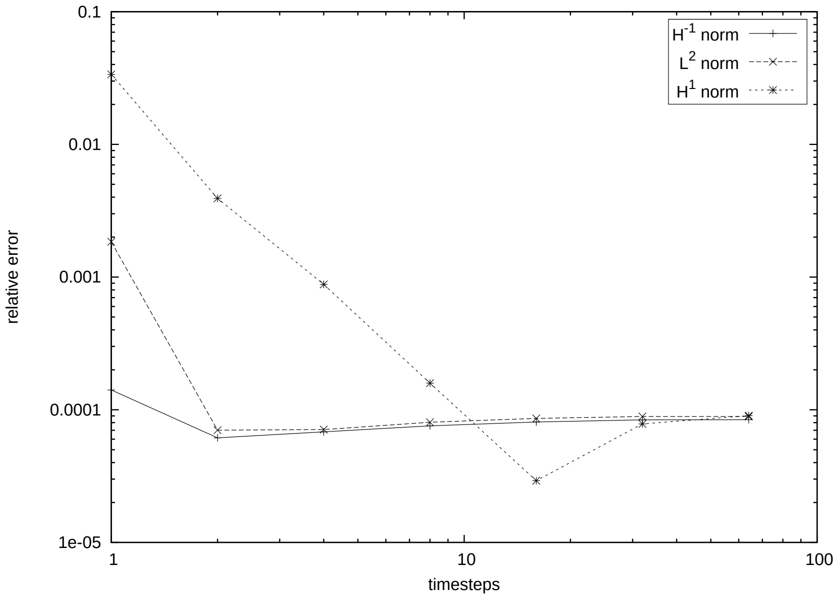

Figure 6.1: Error plot, increasing number of timesteps

Figures 6.4, 6.4 and 6.4 present the results of numerical calculations with increasing number of timesteps, Fourier modes, and quasi-Monte Carlo paths. All errors are relative, and were calculated through comparison with a reference

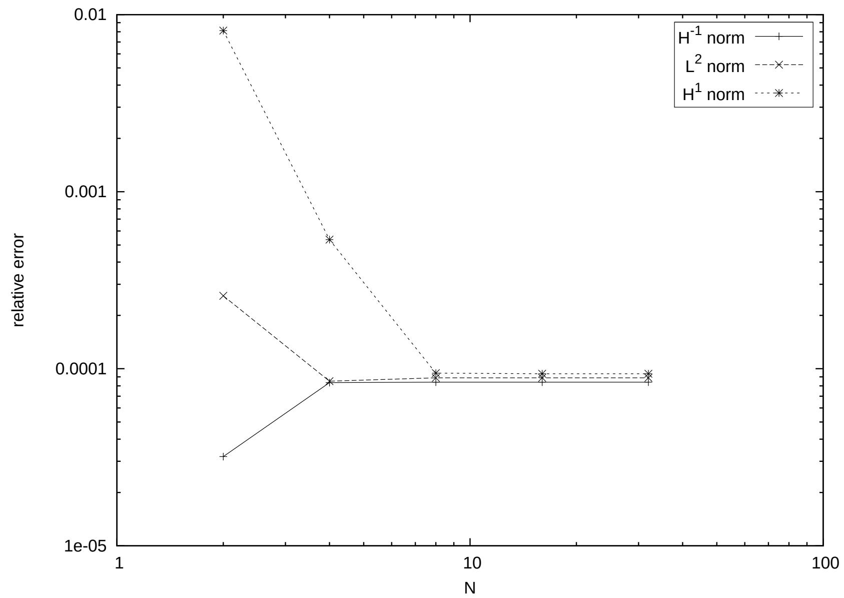

Figure 6.2: Error plot, increasing number of Fourier modes

solution found using $K = 2^{20}$ quasi-Monte Carlo paths, $N = 32$ and $n = 128$ timesteps. There, we obtained the approximate values

$$\mathbb{E}[\|w(1,0)\|_{-1}] \sim 1.138449630686444, \tag{6.59}$$

$$\mathbb{E}[\|w(1,0)\|] \sim 1.319968848291092, \quad \text{and} \tag{6.60}$$

$$\mathbb{E}[\|w(1,0)\|_{+1}] \sim 1.620419847035606. \tag{6.61}$$

In Figure 6.4, we chose the other parameters to be $K = 2^{16}$ and $N = 32$; in Figure 6.4, $K = 2^{16}$ and $n = 128$; and in Figure 6.4, $N = 32$ and $n = 64$.

We clearly see that mainly the number of quasi-Monte Carlo paths limits the attainable accuracy. Nevertheless, with $2^{12} = 4096$ paths, we obtain a relative error of less than 10^{-3}, and that calculation took approximately 60 seconds running on 16 cores of a Primergy RX200 S6 spotting 4 Intel Xeon CPU X5650 processor, each of which provides 6 cores. In Figure 6.4, we observe that we obtain a rate of convergence of about 2.5 for the $\mathbb{H}^1$ norm with respect to the

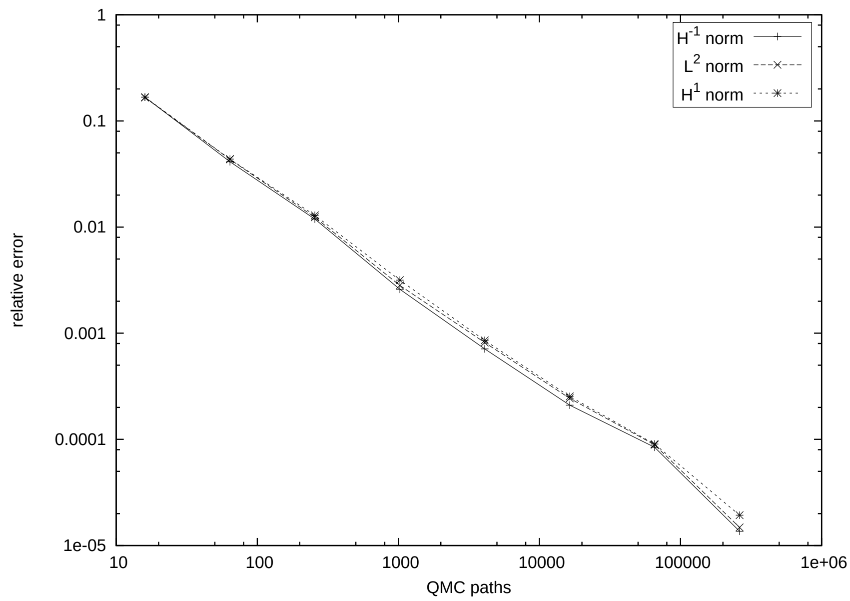

Figure 6.3: Error plot, increasing number of quasi-Monte Carlo paths

number of time steps, which is even more than the theoretically predicted rate of 2 and seems to result from the fact that we compare with numerical estimates instead of the exact value. The solution of the model problem is smooth (see also [74] in this regard), and indeed, Figure 6.4 exhibits spectral convergence in the number of Fourier modes.

6.5 Conclusion

We have introduced and analysed novel high order approximation schemes for the stochastic Navier-Stokes equations on the 2D torus. We prove high order accuracy in time and give precise estimates for the dependence on the order of the spectral Galerkin discretisation. Using high order cubature paths, it is possible to attain convergence of arbitrary order in time.

From a practical point of view, the splitting schemes presented in this work have the important advantage that well-tested and robust solvers for the deterministic Navier-Stokes and Euler equations can be reused. Furthermore, the algorithm makes increasing the dimension of the driving Brownian motion easy.

Numerical examples establish the applicability of the method to some simple, but relevant functionals.

6.6 Proof of Proposition 6.3

Lemma 6.19. *For $N \sim \mathcal{N}(0,1)$, $j = 1, \dots, d$, and S, A, $B \in \mathbb{R}$ with $C \in \mathbb{R}$ small enough,*

$$\begin{aligned}\mathbb{E}[\exp(C(S^2 &+ 2SABN + (BN)^2))] \\ &= \frac{1}{(1-2CB^2)^{1/2}} \exp\left(\left(1 + \frac{2CA^2B^2}{1-2CB^2}\right) CS^2\right). \end{aligned} \tag{6.62}$$

Proof. A direct calculation yields

$$\begin{aligned}\mathbb{E}[\exp(C(S^2 &+ 2SABN + (BN)^2))] \\ &= \int_{\mathbb{R}} \exp(C(S^2 + 2SABy + (By)^2)) \frac{1}{(2\pi)^{1/2}} \exp\left(-\frac{1}{2}y^2\right) \mathrm{d}y \\ &= \frac{1}{(2\pi)^{1/2}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}(1-2CB^2)\left(y - \frac{2CSAB}{1-2CB^2}\right)^2\right) \mathrm{d}y \times \\ &\quad \times \exp\left(\left(1 + \frac{2CA^2B^2}{1-2CB^2}\right) CS^2\right) \\ &= \frac{1}{(1-2CB^2)^{1/2}} \exp\left(\left(1 + \frac{2CA^2B^2}{1-2CB^2}\right) CS^2\right), \end{aligned}$$

which proves the result. □

Corollary 6.20. *For independent $N_j \sim \mathcal{N}(0,1)$, $j = 1, \dots, d$, and S, A_j, $B_j \in \mathbb{R}$ with $C \in \mathbb{R}$ small enough,*

$$\begin{aligned}\mathbb{E}[\exp(C(S^2 &+ \sum_{j=1}^d 2SA_jB_jN_j + \sum_{j=1}^d (B_jN_j)^2))] \\ &= \frac{1}{\prod_{j=1}^d (1-2CB_j^2)^{1/2}} \exp\left(\left(1 + \sum_{j=1}^d \frac{2CA_j^2B_j^2}{1-2CB_j^2}\right) CS^2\right). \end{aligned} \tag{6.63}$$

Proof of Proposition 6.3. Note that

$$w^2(t, w_0) = \exp(t\varepsilon\nu\Delta)w_0 + \int_0^t \exp((t-s)\varepsilon\nu\Delta) \sum_{j=1}^d q_j f_{k_j} \mathrm{d}W_s^j. \tag{6.64}$$

Denoting by λ_j the eigenvalue of f_{k_j} with respect to the operator $\varepsilon\nu\Delta$, $\varepsilon\nu\Delta f_{k_j} = \lambda_j f_{k_j}$, we see that

$$\int_0^t \exp((t-s)\varepsilon\nu\Delta)Q\mathrm{d}W_s = \sum_{j=1}^d \int_0^t \exp((t-s)\lambda_j)q_j f_{k_j}\mathrm{d}W_s^j. \tag{6.65}$$

The coefficient $Z_t^j := \int_0^t \exp((t-s)\tilde{\lambda}_{k_j})\mathrm{d}W_s^j$ is normally distributed, more precisely, $Z_t^j \sim \mathcal{N}\left(0, \frac{1-\exp(2t\tilde{\lambda}_{k_j})}{-2\tilde{\lambda}_{k_j}}\right)$. In particular, with $S(t) := \exp(t\varepsilon\nu\Delta)$,

$$P_t^2\psi(w) = \mathbb{E}\left[\exp\left(\eta\|S(t)w + \sum_{j=1}^d q_j Z_t^j f_{k_j}\|_0^2\right)\right]. \tag{6.66}$$

Note

$$\begin{aligned}\|S(t)w + \sum_{j=1}^d q_j Z_t^j f_{k_j}\|^2 &= \|S(t)w\|^2 \\ &+ 2\sum_{j=1}^d \frac{\langle S(t)w, q_j f_{k_j}\rangle}{\|S(t)w\|\cdot\|q_j f_{k_j}\|}\|S(t)w\|\cdot(\|q_j f_{k_j}\|Z_t^j) + \sum_{j=1}^d (\|q_j f_{k_j}\|Z_t^j)^2,\end{aligned} \tag{6.67}$$

and apply Corollary 6.20 with $C = \eta$, $S = \|S(t)w\|$, $A_j = \frac{\langle S(t)w, q_j f_{k_j}\rangle}{\|S(t)w\|_0\cdot\|q_j f_{k_j}\|_0}$ and $B_j = \|q_j f_{k_r}\|\left(\frac{1-\exp(2t\tilde{\lambda}_{k_j})}{-2\tilde{\lambda}_{k_j}}\right)^{1/2}$. As $A_j^2 \leqslant 1$ and

$$1 - 2CB_j^2 = 1 - 2\eta\|q_j f_{k_j}\|^2 \frac{1-\exp(2t\tilde{\lambda}_{k_j})}{-2\tilde{\lambda}_{k_j}} \geqslant \exp(2\omega t) \tag{6.68}$$

for $0 > \omega \geqslant \tilde{\lambda}_{k_j}$ and $0 < \eta \leqslant \frac{-\omega}{\|q_j f_{k_j}\|^2}$ and, similarly,

$$1 + \sum_{j=1}^d \frac{2CA_j^2 B_j^2}{1-2CB_j^2} \leqslant \exp(2\alpha t) \tag{6.69}$$

for $\alpha > 0$ and $\eta \leqslant \min_{j=1,\dots,d} \frac{2\alpha}{(d-1)\|q_j f_{k_j}\|^2}$, we obtain

$$P_t^2\psi_\eta(w) \leqslant \exp(-dt\omega)\exp(\eta\|w\|^2) = \exp(-dt\omega)\psi_\eta(w), \tag{6.70}$$

the required result. □

Appendix A

Strongly Continuous Semigroups

We give a short overview on the theory of strongly continuous semigroups on Banach spaces. Some standard references are [50, 22, 84, 34].

A.1 Basic definitions and results

Definition A.1. Let $(B, \|\cdot\|_B)$ be a Banach space. A family $(S_t)_{t\geqslant 0}$ of bounded linear operators on B is called a *semigroup of operators* if and only if

(i) $S_0 = I$, the identity operator on B, and

(ii) $S_{t+s} = S_t S_s$ for all $t, s \geqslant 0$.

It is called *strongly continuous* if, moreover,

(iii) for all $x \in B$, $\lim_{t\to 0+} \|S_t x - x\|_B = 0$.

We collect several important properties of semigroups.

Proposition A.2. *For every strongly continuous semigroup $(S_t)_{t\geqslant 0}$ on $(B, \|\cdot\|_B)$, there exist constants $M \geqslant 1$, $\omega \in \mathbb{R}$ such that*

$$\|S_t x\|_B \leqslant M \exp(t\omega)\|x\|_B \quad \textit{for all } t \geqslant 0 \textit{ and } x \in B. \tag{A.1}$$

The following result is well-known and given in [34, Theorem I.5.8].

Proposition A.3. *A semigroup $(S_t)_{t\geqslant 0}$ on $(B, \|\cdot\|_B)$ is strongly continuous if and only if it is* weakly continuous, *i.e., for all $\varphi \in B^*$ and $x \in B$,*

$$\lim_{t\to 0+} \varphi(S_t x) = \varphi(x). \tag{A.2}$$

Definition A.4. Given a strongly continuous semigroup $(S_t)_{t\geqslant 0}$ on $(B, \|\cdot\|_B)$, we define its *infinitesimal generator* by

$$A\colon \operatorname{dom} A \subset B \to B, \quad Ax := \lim_{t\to 0+} t^{-1}(S_t x - x). \tag{A.3}$$

Its domain is

$$\operatorname{dom} A := \left\{ x \in B\colon \text{the limit } \lim_{t\to 0+} t^{-1}(S_t x - x) \text{ exists in } B \right\}. \tag{A.4}$$

If A is the infinitesimal generator of $(S_t)_{t\geqslant 0}$, we also write $\exp(tA) := S_t$.

Recall that a linear operator $A\colon \operatorname{dom} A \subset B \to B$ on $(B, \|\cdot\|_B)$ is called *closed* if and only if for all sequences $(x_n)_{n\in\mathbb{N}}$ in B with $x_n \to x$ and $Ax_n \to y$ in the norm topology of B, we have that $x \in \operatorname{dom} A$ and $Ax = y$. It is called *densely defined* if and only if $\operatorname{dom} A$ is dense in B.

Proposition A.5. *Let $(S_t)_{t\geqslant 0}$ be a strongly continuous semigroup on $(B, \|\cdot\|_B)$ with infinitesimal generator $A\colon \operatorname{dom} A \subset B \to B$.*

(i) A is a closed and densely defined operator.

(ii) For all $t \geqslant 0$, $S_t(\operatorname{dom} A) \subset \operatorname{dom} A$, and $AS_t x = S_t Ax$ for all $x \in \operatorname{dom} A$.

(iii) For $x \in \operatorname{dom} A$, the mapping $t \mapsto S_t x$ is continuously differentiable, and $\frac{\mathrm{d}}{\mathrm{d}t} S_t x = AS_t x$.

(iv) There exists $\omega \in \mathbb{R}$ such that $\lambda - A$ is invertible for $\lambda > \omega$, and the inverse is given by the integral

$$(\lambda - A)^{-1} x = \int_0^\infty \exp(-\lambda s) S_s x \mathrm{d}s. \tag{A.5}$$

The integral is an improper Riemann integral in the norm topology of B.

(v) Let $k \in \mathbb{N}$. If $x \in \operatorname{dom} A^{k+1}$, then $S_t x \in \operatorname{dom} A^{k+1}$ for all $t \geqslant 0$, $t \mapsto S_t x$ is $k+1$ times continuously differentiable, and

$$S_t x = \sum_{j=0}^{k} \frac{t^j}{j!} A^j x + t^{k+1} r_t x, \tag{A.6}$$

where the remainder is explicitly given by

$$r_t x = t^{-(k+1)} \int_0^t \frac{(t-s)^k}{k!} S_s A^{k+1} x \mathrm{d}s \tag{A.7}$$

and satisfies $\|r_t x\|_B \leqslant C \|A^{k+1} x\|_B$.

(vi) *Let $D \subset \operatorname{dom} A$ be dense in B. Assume furthermore that $S_t(D) \subset D$ for all $t > 0$. Then, for every $x \in \operatorname{dom} A$, there exists a sequence $(x_n)_{n\in\mathbb{N}}$ in D such that $\lim_{n\to\infty}(\|x_n - x\|_B + \|Ax_n - Ax\|_B) = 0$. In this case, we say that D is a* core *for A.*

The last point is given in [34, Proposition II.1.7].

The Hille-Yosida theorem gives necessary and sufficient conditions for an unbounded operator on a Banach space to be the infinitesimal generator of a strongly continuous semigroup. In this work, we only need the following result, which is stated, e.g., in [34, Theorem II.3.15].

Proposition A.6 (Lumer-Phillips)**.** *A densely defined operator $A\colon \operatorname{dom} A \subset B \to B$ on $(B, \|\cdot\|_B)$ is the infinitesimal generator of a strongly continuous semigroup of contractions (i.e., $\|S_t x\|_B \leqslant \|x\|_B$ for all $t \geqslant 0$ and $x \in X$) if and only if*

(i) *A is* dissipative, *i.e., $\|(\lambda - A)x\|_B \geqslant \lambda\|x\|_B$, and*

(ii) *$(\lambda - A)B$ is dense in B for some $\lambda > 0$.*

In this case, $(\lambda - A)B$ is dense in B for all $\lambda > 0$.

A.2 Analytic semigroups and fractional powers

We shall need fractional powers of the infinitesimal generator of a semigroup. We only consider these for $(S_t)_{t\geqslant 0}$ analytic. Here, a strongly continuous semigroup $(S_t)_{t\geqslant 0}$ is called *analytic* if and only if $S_t(B) \subset \operatorname{dom} A$ for all $t > 0$, and there exists $C > 0$ such that $\|AS_t x\|_B \leqslant Ct^{-1}\|x\|_B$ for $t > 0$ and $x \in B$. See [84, pp. 60] for equivalent definitions and more background.

Definition A.7. Let $\alpha \in (0, 1)$, and assume that $A\colon \operatorname{dom} A \subset B \to B$ is boundedly invertible and generates an analytic semigroup on $(B, \|\cdot\|_B)$. The fractional power $(-A)^\alpha$ is the inverse of the bounded operator $(-A)^{-\alpha}$ given by

$$(-A)^{-\alpha} = \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty t^{-\alpha}(t - A)^{-1} \mathrm{d}t, \tag{A.8}$$

the integral being taken in the uniform operator topology. Its domain is given by $\operatorname{dom}(-A)^\alpha := (-A)^{-\alpha}(B)$. For $\alpha = n + s$, $n \in \mathbb{N}$, $s \in (0, 1)$, we set $(-A)^\alpha := (-A)^n(-A)^s$, with domain $\operatorname{dom}(-A)^\alpha := (-A)^{-s}(\operatorname{dom}(-A)^n)$.

Proposition A.8. *Let $A\colon \operatorname{dom} A \subset B \to B$ be boundedly invertible and generate an analytic semigroup $(S_t)_{t\geqslant 0}$ on $(B, \|\cdot\|_B)$.*

(i) $\operatorname{dom}(-A)^\beta \subset \operatorname{dom}(-A)^\alpha$ *for* $0 < \alpha < \beta$.

(ii) For $\alpha > 0$, $(-A)^\alpha\colon \operatorname{dom}(-A)^\alpha \subset B \to B$ is a densely defined and closed operator.

(iii) For α, $\beta > 0$, $(-A)^{\alpha+\beta} = (-A)^\alpha(-A)^\beta$ on $\operatorname{dom}(-A)^{\alpha+\beta}$. In particular, integer powers agree with their usual definitions.

(iv) There exists $\delta > 0$ such that for $\alpha > 0$, $S_t(B) \subset \operatorname{dom}(-A)^\alpha$ and $\|(-A)^\alpha S_t x\|_B \leqslant M_\alpha t^{-\alpha}\exp(-\delta t)$ for all $t \geqslant 0$ with some constant M_α independent of t and x.

(v) For $x \in \operatorname{dom}(-A)^\alpha$, $(-A)^\alpha S_t x = S_t(-A)^\alpha x$. In particular, $(S_t)_{t\geqslant 0}$ defines an analytic semigroup on $(\operatorname{dom}(-A)^\alpha, \|\cdot\|_{\operatorname{dom}(-A)^\alpha})$, where $\|x\|_{\operatorname{dom}(-A)^\alpha} := \|x\|_B + \|(-A)^\alpha x\|_B$ is the graph norm.

Appendix B

Stochastic Ordinary And Partial Differential Equations

We give a short overview of the tools of stochastic analysis used regularly in this thesis. Standard textbooks are [58, 59, 57, 92, 81, 90, 71]. Books on stochastic partial differential equations are [28, 29, 27, 89, 85, 62].

B.1 The Itô integral

Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geqslant 0})$ be a filtered, complete probability space satisfying the usual conditions, i.e.,

(i) $\mathcal{F}_0$ contains all $\mathbb{P}$-nullsets, and

(ii) the filtration $(\mathcal{F}_t)_{t\geqslant 0}$ is right continuous, i.e., $\mathcal{F}_s = \bigcap_{t>s} \mathcal{F}_t$ for $s \geqslant 0$.

Here, a filtration is an increasing family of σ-fields on Ω, all contained in $\mathcal{F}$. We denote the expected value with respect to $\mathbb{P}$ by $\mathbb{E}$, i.e., for a random variable $X\colon \Omega \to \mathbb{R}$, we set $\mathbb{E}[X] := \int_\Omega X(\omega)\mathbb{P}(\mathrm{d}\omega)$. For X a real-valued, integrable random variable, conditional expectation with respect to a σ-field $\tilde{\mathcal{F}}$ is denoted by $\mathbb{E}[X|\tilde{\mathcal{F}}]$. See [57, Chapter 5] for background on conditional expectations.

Definition B.1 (Stochastic process; adaptedness)**.** Let $(M, \mathcal{G})$ be a measurable space. A *stochastic process* is a family $(X_t)_{t\in I}$ of random variables on Ω with values in $(M, \mathcal{G})$, indexed by some set I, in our case usually $[0, \infty)$ or $[0, T]$ for some $T > 0$. A stochastic process $(X_t)_{t\geqslant 0}$ is called *adapted* if and only if X_t is $\mathcal{F}_t$-measurable for all $t \geqslant 0$.

Definition B.2 (Brownian motion)**.** A *d-dimensional Brownian motion relative to the filtration* $(\mathcal{F}_t)_{t\geqslant 0}$ is an adapted process $(B_t)_{t\geqslant 0}$ of $\mathbb{R}^d$-valued random variables such that

(i) $B_0 = 0$ almost surely,

(ii) $B_{t+h} - B_t$ is independent of $\mathcal{F}_t$ for $t \geqslant 0$ and $h \geqslant 0$,

(iii) $(B_t)_{t\geqslant 0}$ has almost surely continuous paths, i.e., the mapping $t \mapsto B_t$ is almost surely continuous,

(iv) $B_{t+h} - B_t$ is distributed according to a d-dimensional normal distribution with mean 0 and covariance matrix hI_d, where I_d is the d-dimensional identity matrix.

We shall always assume that Brownian motions are given relative to the filtration used in the definition of the underlying filtered probability space.

To construct the Itô integral, one typically proceeds as follows: First, the definition is given for certain simple integrands where the approach is natural. Then, one proves the Itô isometry. Finally, by determining the closure of the space of simple integrands, a density argument yields the Itô integral on its natural domain of definition.

Definition B.3. A real-valued stochastic process $(X_t)_{t\geqslant 0}$ is called *elementary* if and only if there exist real numbers $(t_j)_{j=0}^{\infty}$ with $0 = t_0 < t_1 < \dots$ and $\mathcal{F}_{t_j}$-measurable random variables ξ_j such that

$$X_t = \sum_{j=0}^{\infty} \xi_j \chi_{[t_j, t_{j+1})}(t). \tag{B.1}$$

Definition B.4 (Itô integral, simple integrands)**.** For an elementary stochastic process $X_t = \sum_{j=0}^{\infty} \xi_j \chi_{[t_j,t_{j+1})}(t)$, we define the Itô integral with respect to a one-dimensional Brownian motion $(B_t)_{t\geqslant 0}$ by

$$\int_0^t X_s \mathrm{d}B_s := \sum_{j=0}^{\infty} \xi_j (B_{t_{j+1}\wedge t} - B_{t_j \wedge t}). \tag{B.2}$$

Clearly, the Itô integral is a linear operator.

Proposition B.5 (Itô isometry)**.** *For elementary integrands* $(X_t)_{t\geqslant 0}$,

$$\mathbb{E}\Big[\Big(\int_0^t X_s \mathrm{d}B_s\Big)^2\Big] = \int_0^t \mathbb{E}[X_s^2] \mathrm{d}s. \tag{B.3}$$

The space of adapted processes is too large to be able to define the Itô integral. The correct space of integrands is defined as follows.

Definition B.6. Let $(M, \mathcal{G})$ be a measurable space. A stochastic process $(X_t)_{t \geqslant 0}$ is called *progressively measurable* if and only if the mapping $[0, T] \times \Omega \to M$, $(t, \omega) \mapsto X_t(\omega)$, is $\mathcal{B}_{[0,T]} \otimes \mathcal{F}_T$-$\mathcal{G}$-measurable for all $T > 0$, where $\mathcal{B}_{[0,T]}$ denotes the Borel sets in $[0, T]$.

Note that progressively measurable processes are always adapted. Conversely, for every adapted process $(\tilde{X}_t)_{t \geqslant 0}$, there exists a progressively measurable process $(X_t)_{t \geqslant 0}$ such that $\mathbb{P}[X_t = \tilde{X}_t] = 1$ for all $t \geqslant 0$, i.e., there exists a progressively measurable *modification*.

Proposition B.7 (Itô integral, progressively measurable integrands)**.** *The closure of the space of elementary integrands with respect to the topology induced by the norm* $[X]_T := \left(\int_0^T \mathbb{E}[X_t^2] \mathrm{d}t\right)^{1/2}$ *equals the space of real-valued progressively measurable processes. In particular, for all such processes, the Itô integral is well-defined and satisfies the Itô isometry, which now reads*

$$\mathbb{E}\Big[\big(\int_0^t X_s \mathrm{d}B_s\big)^2\Big] = [X]_t^2 \quad \text{for } t \in [0, T]. \tag{B.4}$$

While Proposition B.7 allows us to define the Itô integral $Y_t := \int_0^t X_s \mathrm{d}B_s$ for a progressively measurable processes, it does not allow us to speak about path properties of the process $(Y_t)_{t \geqslant 0}$, the reason being that this integral is only defined up to modification. This leads to the following approach.

Proposition B.8. *Let* $(M_t)_{t \geqslant 0}$ *be a* martingale, *i.e.,* $(M_t)_{t \geqslant 0}$ *is a real-valued, adapted stochastic process consisting of integrable random variables with*

$$\mathbb{E}[M_t | \mathcal{F}_s] = M_s \quad \text{for all } 0 \leqslant s \leqslant t. \tag{B.5}$$

Then, $(M_t)_{t \geqslant 0}$ *has a modification that is almost surely càdlàg, i.e., for almost all* $\omega \in \Omega$*, the mapping* $t \mapsto M_t(\omega)$ *is right continuous and has left limits.*

This result also holds true if $(M_t)_{t \geqslant 0}$ is a *supermartingale*, i.e., a real-valued adapted stochastic process of integrable random variables with

$$\mathbb{E}[M_t | \mathcal{F}_s] \leqslant M_s \quad \text{for all } 0 \leqslant s \leqslant t, \tag{B.6}$$

or a *submartingale*, i.e., $(-M_t)_{t \geqslant 0}$ is a supermartingale.

Proposition B.9. *Let $(X_t)_{t\geqslant 0}$ be a progressively measurable process satisfying $[X]_T < \infty$. Then, the Itô integral $(Y_t)_{t\geqslant 0}$, $Y_t = \int_0^t X_s \mathrm{d}B_s$ for $t \geqslant 0$, is a martingale.*

Hence, there exists a modification of $(Y_t)_{t\geqslant 0}$ that is càdlàg. In the future, we shall always choose this modification.

Choosing $X_t = 1$ for all $t \geqslant 0$, we see that this contains the martingale property of Brownian motion itself. We collect some features of the Itô integral.

Proposition B.10. *Let $(X_t)_{t\geqslant 0}$, $(X_t^1)_{t\geqslant 0}$, $(X_t^2)_{t\geqslant 0}$ be progressively measurable with $[X]_T$, $[X^1]_T$, $[X^2]_T < \infty$.*

(i) $\int_0^0 X_s \mathrm{d}B_s = 0$.

(ii) The mapping $t \mapsto \int_0^t X_s \mathrm{d}B_s$ *is almost surely continuous.*

(iii) For $\alpha_1, \alpha_2 \in \mathbb{R}$,

$$\int_0^t (\alpha_1 X_s^1 + \alpha_2 X_s^2) \mathrm{d}B_s = \alpha_1 \int_0^t X_s^1 \mathrm{d}B_s + \alpha_2 \int_0^t X_s^2 \mathrm{d}B_s, \tag{B.7}$$

i.e., the Itô integral is a linear operator.

Remark B.11. By a localisation argument, it is possible to extend the Itô integral to all progressively measurable processes $(X_t)_{t\geqslant 0}$ that only satisfy the property

$$\mathbb{P}\Big[\int_0^T X_t^2 \mathrm{d}t < \infty\Big] = 1 \quad \text{for all } T \geqslant 0. \tag{B.8}$$

In this case, the stochastic integral still defines a linear operator and $\int_0^t X_s \mathrm{d}B_s$ still has almost surely continuous paths, but we no longer obtain martingales, but only *local martingales*, i.e., processes that become martingales when stopped at appropriate stopping times.

Proposition B.12 (Itô formula)**.** *Let $f\colon [0,\infty) \times \mathbb{R} \to \mathbb{R}$, $(t,x) \mapsto f(t,x)$, be a function once differentiable with respect to t and twice with respect to x. Assume that $(X_t)_{t\geqslant 0}$ is an* Itô process*, i.e., can be written in the form $X_t = X_0 + \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}B_s$, where X_0 is a constant, $(\mu_t)_{t\geqslant 0}$ is adapted with* *almost every path Lebesgue integrable on* $[0, T]$*, and* $(\sigma_t)_{t\geqslant 0}$ *is progressively measurable and satisfies* (B.8)*. Then,*

$$
\begin{aligned}
f(t, X_t) = f(0, X_0) &+ \int_0^t f_{,t}(s, X_s)\mathrm{d}s + \int_0^t f_{,x}(s, X_s)(\mu_s \mathrm{d}s + \sigma_s \mathrm{d}B_s) \\
&+ \frac{1}{2}\int_0^t f_{,xx}(s, X_s)\sigma_s^2 \mathrm{d}s.
\end{aligned}
\tag{B.9}
$$

This is also written in the form

$$
\mathrm{d}f(t, X_t) = f_{,t}(t, X_t)\mathrm{d}t + f_{,x}(t, X_t)(\mu_t \mathrm{d}t + \sigma_t \mathrm{d}B_t) + \frac{1}{2} f_{,xx}(t, X_t)\sigma_t^2 \mathrm{d}t. \tag{B.10}
$$

The last term in the above expressions shows the deviation of this chain rule for the Itô calculus from the usual rules of deterministic calculus. The differential notation given in (B.10) cannot be directly defined, as Brownian motion is almost surely nowhere differentiable. It is only to be seen as shorthand notation for the corresponding integral expression.

Remark B.13. If $(X_t)_{t\geqslant 0}$ is $\mathbb{R}^N$-valued such that every component is an Itô process with respect to a d-dimensional Brownian motion $(B_t)_{t\geqslant 0}$,

$$
X_t^k = X_0^k + \int_0^t \mu_s^k \mathrm{d}s + \sum_{j=1}^d \sigma_s^{k,j} \mathrm{d}B_s^j, \quad k = 1, \dots, N, \tag{B.11}
$$

a corresponding formula holds true; see, e.g., [58, Theorem 3.3.6]. (B.11) is customarily written as

$$
\mathrm{d}X_t^k = \mu_t^k \mathrm{d}t + \sum_{j=1}^d \sigma_t^{k,j} \mathrm{d}B_t^j, \quad k = 1, \dots, N. \tag{B.12}
$$

Definition B.14. Let $(X_t)_{t\geqslant 0}$ be an Itô process with values in $\mathbb{R}^N$ satisfying (B.11), and let $(Y_t)_{t\geqslant 0}$ be a progressively measurable process such that $(Y_t^k \mu_t^k)_{t\geqslant 0}$ has almost surely Lebesgue integrable paths for $k = 1, \dots, N$ and $(Y_t^k \sigma_t^{k,j})_{t\geqslant 0}$ satisfies (B.8) for all $k = 1, \dots, N$ and $j = 1, \dots, d$. Then, the Itô integral $\int_0^t Y_s \mathrm{d}X_s$ is defined by

$$
\int_0^t Y_s \mathrm{d}X_s := \sum_{k=1}^N \int_0^t Y_s^k \mu_s^k \mathrm{d}s + \sum_{k=1}^N \sum_{j=1}^d \int_0^t Y_s^k \sigma_s^{k,j} \mathrm{d}B_s^j. \tag{B.13}
$$

While the Itô integral does not reproduce the chain rule from deterministic calculus, there is another stochastic integral that does.

Definition B.15 (Stratonovich integral)**.** Let $(X_t^1)_{t\geqslant 0}$, $(X_t^2)_{t\geqslant 0}$ be real-valued Itô diffusions,

$$\mathrm{d}X_t^k = \mu_t^k \mathrm{d}t + \sum_{j=1}^{d} \sigma_t^{k,j} \mathrm{d}B_t^j, \quad k = 1, 2. \tag{B.14}$$

Then, the *Stratonovich integral* $\int_0^t X_s^1 \circ \mathrm{d}X_s^2$ is defined by

$$\int_0^t X_s^1 \circ \mathrm{d}X_s^2 := \int_0^t X_s^1 \mathrm{d}X_s^2 + \sum_{j=1}^{d} \frac{1}{2} \int_0^t \sigma_s^{1,j} \sigma_s^{2,j} \mathrm{d}s. \tag{B.15}$$

Proposition B.16. *Let $(X_t)_{t\geqslant 0}$ be an Itô process with values in $\mathbb{R}^N$, and assume that $f\colon \mathbb{R}^N \to \mathbb{R}$ is three times continuously differentiable. Then,*

$$f(X_t) = f(X_0) + \sum_{k=1}^{N} \int_0^t f_{,x^k}(X_s) \circ \mathrm{d}X_s^k. \tag{B.16}$$

B.2 Stochastic ordinary differential equations

In this thesis, we analyse numerical methods for stochastic differential equations of the form

$$\mathrm{d}x(t, x_0) = \alpha(x(t, x_0))\mathrm{d}t + \sum_{j=1}^{d} \sigma_j(x(t, x_0))\mathrm{d}B_t^j, \quad x(0, x_0) = x_0 \tag{B.17}$$

on $\mathbb{R}^N$. Here, $(B_t)_{t\geqslant 0}$ is a d-dimensional Brownian motion, α, $\sigma_j\colon \mathbb{R}^N \to \mathbb{R}^N$ are vector fields, and $(x(t, x_0))_{t\geqslant 0}$ is a stochastic process with values in $\mathbb{R}^N$ satisfying the equations above, i.e., as we again need to interpret the differentials as integrals,

$$x(t, x_0) = x_0 + \int_0^t \alpha(x(s, x_0))\mathrm{d}s + \sum_{j=1}^{d} \int_0^t \sigma_j(x(s, x_0))\mathrm{d}B_s^j. \tag{B.18}$$

For us, x_0 will typically be a constant, but in general, it can be any $\mathcal{F}_0$-measurable random variable. If the vector fields are regular enough, any Itô equation can be rewritten into an equivalent Stratonovich form,

$$\mathrm{d}x(t,x_0) = \alpha_0(x(t,x_0))\mathrm{d}t + \sum_{j=1}^{d} \sigma_j(x(t,x_0)) \circ \mathrm{d}B_t^j, \quad x(0,x_0) = x_0, \tag{B.19}$$

where $\alpha_0(x) := \alpha(x) - \frac{1}{2}\sum_{j=1}^{d} D\sigma_j(x)\sigma_j(x)$ denotes the *Stratonovich corrected drift*. When dealing with Stratonovich equations, we also write

$$\mathrm{d}x(t,x_0) = \sum_{j=0}^{d} V_j(x(t,x_0)) \circ \mathrm{d}B_t^j, \quad x(0,x_0) = x_0, \tag{B.20}$$

with vector fields $V_j\colon \mathbb{R}^N \to \mathbb{R}^N$, where we set $B_t^0 = t$ to shorten the notation.

The fundamental result on solvability of such equations is the following, which is essentially a copy of the corresponding theorem for ordinary differential equations.

Proposition B.17. *Assume that α, σ_j are Lipschitz continuous vector fields. Then, there exists a unique solution $(x(t,x_0))_{t\geqslant 0}$ of* (B.17) *with almost surely continuous paths that is adapted to the filtration generated by $(B_t)_{t\geqslant 0}$, i.e., the smallest filtration making $(B_t)_{t\geqslant 0}$ adapted (and, hence, also to $(\mathcal{F}_t)_{t\geqslant 0}$). In particular, all integrals appearing in* (B.17) *are well-defined. Furthermore, $\sup_{t\in[0,T]}\mathbb{E}[|x(t,x_0)|^2] < \infty$ for $T \geqslant 0$, the mapping $x_0 \mapsto x(t,x_0)$ is almost surely Lipschitz continuous, and $\mathbb{E}[|x(t,x_1) - x(t,x_2)|^2] \leqslant C|x_1 - x_2|^2$ for all x_1, $x_2 \in \mathbb{R}^N$.*

Here, $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^N$. The proof of the theorem is done by Picard iterations.

We remark that extensions are possible in many directions, in particular, α can be allowed to only satisfy a one-sided Lipschitz condition, see [71]. Solutions as obtained in Proposition B.17 are also called *strong solutions*. If the coefficients are less regular, it can still be possible to solve (B.17) on a larger probability space (*weak solutions*). As we shall not need this more general notion in this work, we refer the reader to [58, 81] for more details on existence and properties of weak solutions.

One of the properties of solutions of stochastic differential equations extensively used in this work is the *Markov property*.

Proposition B.18. *Assume that the vector fields α, σ_j, $j = 1, \ldots, d$ are Lipschitz continuous. Then, the solution $(x(t, x_0))_{t \geqslant 0}$ of* (B.17) *satisfies*

$$\mathbb{E}[f(x(t+h, x_0))|\mathcal{F}_t^x] = \mathbb{E}[f(x(h, \xi))]|_{\xi = x(t, x_0)}, \tag{B.21}$$

where $f\colon \mathbb{R}^N \to \mathbb{R}$ is a bounded and measurable function. Here, $(\mathcal{F}_t^x)$ denotes the filtration generated by $(x(t, x_0))_{t \geqslant 0}$.

Define the Markov semigroup $(P_t)_{t \geqslant 0}$ of $x(t, x_0)$ by $P_t f(x_0) := \mathbb{E}[f(t, x_0)]$. Proposition B.17 shows that $P_t f$ is bounded and continuous if f is. Furthermore, $(P_t)_{t \geqslant 0}$ actually is a semigroup on the bounded and continuous functions, endowed with the supremum norm $\|f\| := \sup_{x \in \mathbb{R}^N} |f(x)|$, as boundedness of P_t, $t \geqslant 0$, follows from the monotony of the integral, $P_0 = I$ is obvious, and Propostion B.18 yields $P_{t+s} = P_t P_s$ for t, $s \geqslant 0$.

B.3 Stochastic partial differential equations

Let $(H, \|\cdot\|_H)$ be a separable Hilbert space. For vector fields α, $\sigma_j\colon H \to H$, and $A\colon \operatorname{dom} A \subset H \to H$ the infinitesimal generator of a strongly continuous semigroup on H, see Appendix A, we want to consider the equation

$$\mathrm{d}x(t, x_0) = (Ax(t, x_0) + \alpha(x(t, x_0)))\mathrm{d}t + \sum_{j=1}^{d} \sigma_j(x(t, x_0))\mathrm{d}B_t^j, \tag{B.22}$$

where $(B_t)_{t \geqslant 0}$ is a d-dimensional Brownian motion. Here, we need to take stochastic integrals with values in H. As we restrict ourselves to finite-dimensional driving noise, these can be constructed as in Section B.1, as the Itô isometry holds true for Hilbert space-valued elementary integrands; it reads

$$\mathbb{E}[\|\int_0^t X_s \mathrm{d}B_s^j\|_H^2] = \int_0^t \mathbb{E}[\|X_s\|_H^2]\mathrm{d}s \quad \text{for } j = 1, \ldots, d. \tag{B.23}$$

We shall only state the following generalisation of Proposition B.12, which can be found in [29, Theorem 7.2.1].

Proposition B.19. *Assume that $(X_t)_{t \geqslant 0}$ is an Itô process with values in H driven by a d-dimensional Brownian motion, i.e.,*

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sum_{j=1}^{d} \sigma_t^j \mathrm{d}B_t^j. \tag{B.24}$$

Then, for every $f\colon [0,\infty)\times H \to \mathbb{R}$, $(t,x)\mapsto f(t,x)$, *once differentiable with respect to* t *and twice with respect to* x, *uniformly continuous on bounded subsets of* $[0,\infty)\times H$ *together with its derivatives,*

$$\begin{aligned} \mathrm{d}f(t,X_t) &= D_t f(t,X_t)\mathrm{d}t + D_x f(t,X_t)\mathrm{d}X_t \\ &\quad + \frac{1}{2}\sum_{j=1}^{d} D_x^2 f(t,X_t)(\sigma_t^j,\sigma_t^j)\mathrm{d}t. \end{aligned} \tag{B.25}$$

For the general case of infinite-dimensional driving noise, see, e.g., [27], or any other book on stochastic partial differential equations cited at the beginning of this chapter.

As in the finite-dimensional case, assuming that σ_j is Fréchet differentiable, it is possible to transform to Stratonovich form, the Stratonovich corrected drift being

$$\alpha_0(x) := \alpha(x) - \frac{1}{2}\sum_{j=1}^{d} D\sigma_j(x)\sigma_j(x). \tag{B.26}$$

Moreover, the finite-dimensional case is contained in the infinite-dimensional case by setting $H=\mathbb{R}^N$ and $A=0$.

Similarly to deterministic partial differential equations, it is usually not possible to solve (B.22) in the strong sense, i.e., taking classical derivatives (the differentiability requirements correspond to $x(t,x_0)\in \operatorname{dom} A$). Instead, using the semigroup $S_t := \exp(tA)$ generated by A, we consider the *mild formulation*

$$x(t,x_0) = S_t x_0 + \int_0^t S_{t-s}\alpha(x(s,x_0))\mathrm{d}s + \sum_{j=1}^{d}\int_0^t S_{t-s}\sigma_j(x(s,x_0))\mathrm{d}B_s^j. \tag{B.27}$$

Again, we shall restrict ourselves to deterministic initial conditions, but remark that an extension to $\mathcal{F}_0$-measurable random variables is possible.

Proposition B.20. *Assume that* α, σ_j, $j=1,\dots,d$, *are Lipschitz continuous. Then, there exists a unique solution* $(x(t,x_0))_{t\geqslant 0}$ *of* (B.27) *with almost surely continuous paths, and the mapping* $x_0\mapsto x(t,x_0)$ *is almost surely Lipschitz continuous. Furthermore,* $\sup_{t\in[0,T]}\mathbb{E}[\|x(t,x_0)\|_H^p] \leqslant C_T(1+\|x_0\|_H^p)$ *with some* $C_T>0$ *for all* $T>0$.

Define $P_t f(x_0) := \mathbb{E}[f(x(t,x_0))]$. As in the finite-dimensional case, $(P_t)_{t\geqslant 0}$ defines a semigroup (the proof of Proposition B.18 given in [81, p. 115] clearly generalises, being applicable to all cases where strong solutions exist for arbitrary square integrable initial values; see also [92, p.371]). The next result proves differentiability with respect to the initial value.

Proposition B.21. *Assume that α, σ_j, $j = 1, \dots, d$ are k times continuously Fréchet differentiable with bounded derivatives (α and σ_j, $j = 1, \dots, d$, do not have to be bounded themselves). Then, $x_0 \mapsto x(t, x_0)$ is almost surely k times continuously Fréchet differentiable, and the derivatives are given by taking formal derivatives in* (B.27)*. Furthermore, for all $T \geqslant 0$ there exists some constant $C_T > 0$ independent of x_0 with $\sup_{t\in[0,T]} \mathbb{E}[\|D^j_{x_0} x(t, x_0)\|^p_{L_j(H;H)}] \leqslant C_T$.*

Here, $L_j(H; H)$ denotes the Banach space of bounded j-linear maps $H^j \to H$, endowed with the norm

$$\|a\|_{L_j(H;H)} := \sup_{\substack{\|h_i\|\leqslant 1 \\ i=1,\dots,j}} \|a(h_1, \dots, h_j)\|_H; \tag{B.28}$$

see also Definition 2.23. Hence, P_t preserves differentiability if the coefficients are smooth enough.

Bibliography

[1] Sergio Albeverio, Franco Flandoli, and Yakov G. Sinai. *SPDE in hydrodynamic: recent progress and prospects*, volume 1942 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2008. Lectures given at the C.I.M.E. Summer School held in Cetraro, August 29–September 3, 2005, Edited by Giuseppe Da Prato and Michael Röckner.

[2] Aurélien Alfonsi. High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Math. Comp.*, 79(269):209–237, 2010.

[3] Jürgen Appell and Petr P. Zabrejko. *Nonlinear superposition operators*, volume 95 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1990.

[4] Fabrice Baudoin. *An introduction to the geometry of stochastic flows*. Imperial College Press, London, 2004.

[5] Christian Bayer and Josef Teichmann. The proof of Tchakaloff's theorem. *Proc. Amer. Math. Soc.*, 134(10):3035–3040 (electronic), 2006.

[6] Christian Bayer and Josef Teichmann. Cubature on Wiener space in infinite dimension. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 464(2097):2493–2516, 2008.

[7] A. Bensoussan. Splitting up method in the context of stochastic PDE. In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume 176 of *Lecture Notes in Control and Inform. Sci.*, pages 22–31. Springer, Berlin, 1992.

[8] A. Bensoussan and R. Glowinski. Approximation of Zakai equation by the splitting up method. In *Stochastic systems and optimization (Warsaw, 1988)*, volume 136 of *Lecture Notes in Control and Inform. Sci.*, pages 257–265. Springer, Berlin, 1989.

[9] A. Bensoussan, R. Glowinski, and A. Răşcanu. Approximation of the Zakai equation by the splitting up method. *SIAM J. Control Optim.*, 28(6):1420–1431, 1990.

[10] A. Bensoussan, R. Glowinski, and A. Răşcanu. Approximation of some stochastic differential equations by the splitting up method. *Appl. Math. Optim.*, 25(1):81–106, 1992.

[11] A. Bensoussan and R. Temam. Équations stochastiques du type Navier-Stokes. *J. Functional Analysis*, 13:195–222, 1973.

[12] Alain Bensoussan, Giuseppe Da Prato, Michel C. Delfour, and Sanjoy K. Mitter. *Representation and control of infinite dimensional systems*. Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, second edition, 2007.

[13] Christian Berg, Jens Peter Reus Christensen, and Paul Ressel. *Harmonic analysis on semigroups*, volume 100 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1984. Theory of positive definite and related functions.

[14] L. C. Berselli, T. Iliescu, and W. J. Layton. *Mathematics of large eddy simulation of turbulent flows*. Scientific Computation. Springer-Verlag, Berlin, 2006.

[15] Sergio Blanes and Fernando Casas. On the necessity of negative coefficients for operator splitting schemes of order higher than two. *Appl. Numer. Math.*, 54(1):23–37, 2005.

[16] Nicolas Bourbaki. *Éléments de mathématique. Fasc. XXXV. Livre VI: Intégration. Chapitre IX: Intégration sur les espaces topologiques séparés*. Actualités Scientifiques et Industrielles, No. 1343. Hermann, Paris, 1969.

[17] Nicolas Bourbaki. *Éléments de mathématique. Topologie générale. Chapitres 1 à 4*. Hermann, Paris, 1971.

[18] Nicolas Bourbaki. *Elements of mathematics. General topology. Chapters 5–10. (Éléments de mathématique. Topologie générale. Chapitres 5 à 10.) Reprint of the 1974 original.* Berlin: Springer, 1974.

[19] Haïm Brézis. *Functional analysis. Theory and applications. (Analyse fonctionnelle. Théorie et applications.)*. Collection Mathématiques Appliquées pour la Maîtrise. Paris: Masson. 248 p. , 1994.

[20] Damiano Brigo and Fabio Mercurio. *Interest rate models—theory and practice*. Springer Finance. Springer-Verlag, Berlin, 2001.

[21] Félix E. Browder. Remarks on nonlinear interpolation in Banach spaces. *J. Functional Analysis*, 4:390–403, 1969.

[22] Paul L. Butzer and Hubert Berens. *Semi-groups of operators and approximation*. Die Grundlehren der mathematischen Wissenschaften, Band 145. Springer-Verlag New York Inc., New York, 1967.

[23] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral methods*. Scientific Computation. Springer-Verlag, Berlin, 2006. Fundamentals in single domains.

[24] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral methods*. Scientific Computation. Springer, Berlin, 2007. Evolution to complex geometries and applications to fluid dynamics.

[25] René A. Carmona and Michael R. Tehranchi. *Interest rate models: an infinite dimensional stochastic analysis perspective*. Springer Finance. Springer-Verlag, Berlin, 2006.

[26] Dan Crisan and Saadia Ghazali. On the convergence rates of a general class of weak approximations of SDEs. In *Stochastic differential equations: theory and applications*, volume 2 of *Interdiscip. Math. Sci.*, pages 221–248. World Sci. Publ., Hackensack, NJ, 2007.

[27] Giuseppe Da Prato. *An introduction to infinite-dimensional analysis*. Universitext. Springer-Verlag, Berlin, 2006. Revised and extended from the 2001 original by Da Prato.

[28] Giuseppe Da Prato and Jerzy Zabczyk. *Ergodicity for infinite-dimensional systems*, volume 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.

[29] Giuseppe Da Prato and Jerzy Zabczyk. *Second order partial differential equations in Hilbert spaces*, volume 293 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2002.

[30] Philipp Dörsek. Semigroup Splitting And Cubature Approximations For The Stochastic Navier-Stokes Equations. *ArXiv e-prints*, May 2011.

[31] Philipp Dörsek and Josef Teichmann. A Semigroup Point Of View On Splitting Schemes For Stochastic (Partial) Differential Equations. *ArXiv e-prints*, November 2010.

[32] Philipp Dörsek, Josef Teichmann, and Dejan Velušček. Cubature methods for stochastic (partial) differential equations in weighted spaces. in preparation, 2011.

[33] Weinan E and Jonathan C. Mattingly. Ergodicity for the Navier-Stokes equation with degenerate random forcing: finite-dimensional approximation. *Comm. Pure Appl. Math.*, 54(11):1386–1402, 2001.

[34] Klaus-Jochen Engel and Rainer Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.

[35] Damir Filipović. *Consistency problems for Heath-Jarrow-Morton interest rate models*, volume 1760 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001.

[36] Damir Filipović. *Term-structure models*. Springer Finance. Springer-Verlag, Berlin, 2009. A graduate course.

[37] Damir Filipović and Josef Teichmann. On the geometry of the term structure of interest rates. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041):129–167, 2004. Stochastic analysis with applications to mathematical finance.

[38] Patrick Florchinger and François Le Gland. Time-discretization of the Zakai equation for diffusion processes observed in correlated noise. *Stochastics Stochastics Rep.*, 35(4):233–256, 1991.

[39] W. Gui and I. Babuška. The *h*, *p* and *h-p* versions of the finite element method in 1 dimension. II. The error analysis of the *h*- and *h-p* versions. *Numer. Math.*, 49(6):613–657, 1986.

[40] István Gyöngy. Approximations of stochastic partial differential equations. In *Stochastic partial differential equations and applications (Trento, 2002)*, volume 227 of *Lecture Notes in Pure and Appl. Math.*, pages 287–307. Dekker, New York, 2002.

[41] István Gyöngy and Nicolai Krylov. On the rate of convergence of splitting-up approximations for SPDEs. In *Stochastic inequalities and applications*, volume 56 of *Progr. Probab.*, pages 301–321. Birkhäuser, Basel, 2003.

[42] István Gyöngy and Nicolai Krylov. On the splitting-up method and stochastic partial differential equations. *Ann. Probab.*, 31(2):564–591, 2003.

[43] István Gyöngy and Nicolai Krylov. Expansion of solutions of parameterized equations and acceleration of numerical methods. *Illinois J. Math.*, 50(1-4):473–514 (electronic), 2006.

[44] István Gyöngy and Nicolai Krylov. Accelerated numerical schemes for pdes and spdes. In Dan Crisan, editor, *Stochastic Analysis 2010*, pages 131–168. Springer Berlin Heidelberg, 2011.

[45] Lajos Gergely Gyurkó and Terry J. Lyons. Efficient and practical implementations of cubature on wiener space. In Dan Crisan, editor, *Stochastic Analysis 2010*, pages 73–111. Springer Berlin Heidelberg, 2011.

[46] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, second edition, 2006. Structure-preserving algorithms for ordinary differential equations.

[47] Martin Hairer and Jonathan C. Mattingly. Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing. *Ann. of Math. (2)*, 164(3):993–1032, 2006.

[48] Martin Hairer and Jonathan C. Mattingly. Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations. *Ann. Probab.*, 36(6):2050–2091, 2008.

[49] Eskil Hansen and Alexander Ostermann. Exponential splitting for unbounded operators. *Math. Comp.*, 78(267):1485–1496, 2009.

[50] Einar Hille and Ralph S. Phillips. *Functional analysis and semi-groups*. American Mathematical Society Colloquium Publications, vol. 31. American Mathematical Society, Providence, R. I., 1957. rev. ed.

[51] Thomas Y. Hou, Wuan Luo, Boris Rozovskii, and Hao-Min Zhou. Wiener chaos expansions and numerical solutions of randomly forced equations of fluid mechanics. *J. Comput. Phys.*, 216(2):687–706, 2006.

[52] Martin Hutzenthaler, Arnulf Jentzen, and Peter E. Kloeden. Strong and weak divergence in finite time of euler's method for stochastic differential equations with non-globally lipschitz continuous coefficients. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 467(2130):1563–1576, 2011.

[53] Kazufumi Ito and Boris Rozovskii. Approximation of the Kushner equation for nonlinear filtering. *SIAM J. Control Optim.*, 38(3):893–915 (electronic), 2000.

[54] Tobias Jahnke and Christian Lubich. Error bounds for exponential operator splittings. *BIT*, 40(4):735–744, 2000.

[55] A. Jentzen and P. E. Kloeden. The numerical approximation of stochastic partial differential equations. *Milan J. Math.*, 77:205–244, 2009.

[56] Stephen Joe and Frances Y. Kuo. Constructing Sobol′ sequences with better two-dimensional projections. *SIAM J. Sci. Comput.*, 30(5):2635–2654, 2008.

[57] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, 1997.

[58] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[59] Peter E. Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.

[60] Wolfgang Kluge. *Time-inhomogeneous Lévy processes in interest rate and credit risk models*. PhD thesis, University of Freiburg, 2005.

[61] Peter Kotelenez. A stochastic Navier-Stokes equation for the vorticity of a two-dimensional fluid. *Ann. Appl. Probab.*, 5(4):1126–1160, 1995.

[62] Peter Kotelenez. *Stochastic ordinary and stochastic partial differential equations*, volume 58 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2008. Transition from microscopic to macroscopic equations.

[63] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. III. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 34(2):391–442, 1987.

[64] Shigeo Kusuoka. Approximation of expectation of diffusion process and mathematical finance. In *Taniguchi Conference on Mathematics Nara '98*, volume 31 of *Adv. Stud. Pure Math.*, pages 147–165. Math. Soc. Japan, Tokyo, 2001.

[65] Shigeo Kusuoka. Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus. In *Advances in mathematical economics. Vol. 6*, volume 6 of *Adv. Math. Econ.*, pages 69–83. Springer, Tokyo, 2004.

[66] Shigeo Kusuoka. Gaussian K-Scheme. *Journal of Mathematical Sciences of the University of Tokyo*, (15), 2009.

[67] François Le Gland. Splitting-up approximation for SPDEs and SDEs with application to nonlinear filtering. In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume 176 of *Lecture Notes in Control and Inform. Sci.*, pages 177–187. Springer, Berlin, 1992.

[68] Joram Lindenstrauss and Lior Tzafriri. *Classical Banach spaces. I.* Springer-Verlag, Berlin, 1977. Sequence spaces, Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 92.

[69] M.I.A. Lourakis. levmar: Levenberg-marquardt nonlinear least squares algorithms in C/C++. [web page] `http://www.ics.forth.gr/~lourakis/levmar/`, Jul. 2004. [Accessed on 14 Jun. 2011.].

[70] Terry Lyons and Nicolas Victoir. Cubature on Wiener space. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041):169–198, 2004. Stochastic analysis with applications to mathematical finance.

[71] Xuerong Mao. *Stochastic differential equations and applications*. Horwood Publishing Limited, Chichester, second edition, 2008.

[72] Bohdan Maslowski and Jan Seidler. On sequentially weakly Feller solutions to SPDE's. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.*, 10(2):69–78, 1999.

[73] Jonathan C. Mattingly. *The Stochastic Navier-Stokes Equation – Energy Estimates and Phase Space Contraction*. PhD thesis, Princeton University, 1998.

[74] Jonathan C. Mattingly. The dissipative scale of the stochastics Navier-Stokes equation: regularization and analyticity. *J. Statist. Phys.*, 108(5-6):1157–1179, 2002. Dedicated to David Ruelle and Yasha Sinai on the occasion of their 65th birthdays.

[75] Jonathan C. Mattingly and Étienne Pardoux. Malliavin calculus for the stochastic 2D Navier-Stokes equation. *Comm. Pure Appl. Math.*, 59(12):1742–1790, 2006.

[76] R. Mikulevicius and B. L. Rozovskii. Stochastic Navier-Stokes equations for turbulent flows. *SIAM J. Math. Anal.*, 35(5):1250–1310 (electronic), 2004.

[77] Salah-Eldin A. Mohammed, Tusheng Zhang, and Huaizhong Zhao. The stable manifold theorem for semilinear stochastic evolution equations and stochastic partial differential equations. *Mem. Amer. Math. Soc.*, 196(917):vi+105, 2008.

[78] Marek Musiela and Marek Rutkowski. *Martingale methods in financial modelling*, volume 36 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, second edition, 2005.

[79] Mariko Ninomiya and Syoiti Ninomiya. A new higher-order weak approximation scheme for stochastic differential equations and the Runge-Kutta method. *Finance Stoch.*, 13(3):415–443, 2009.

[80] Syoiti Ninomiya and Nicolas Victoir. Weak approximation of stochastic differential equations and application to derivative pricing. *Appl. Math. Finance*, 15(1-2):107–121, 2008.

[81] Bernt Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.

[82] K. Oshima, J. Teichmann, and D. Velušček. A new extrapolation method for weak approximation schemes with applications. *Ann. Appl. Probab., to appear*, 2011.

[83] Hans Christian Öttinger. *Stochastic processes in polymeric fluids*. Springer-Verlag, Berlin, 1996. Tools and examples for developing simulation algorithms.

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

[85] S. Peszat and J. Zabczyk. *Stochastic partial differential equations with Lévy noise*, volume 113 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2007. An evolution equation approach.

[86] W. P. Petersen. A general implicit splitting for stabilizing numerical simulations of Itô stochastic differential equations. *SIAM J. Numer. Anal.*, 35(4):1439–1451 (electronic), 1998.

[87] Eckhard Platen. An introduction to numerical methods for stochastic differential equations. In *Acta numerica, 1999*, volume 8 of *Acta Numer.*, pages 197–246. Cambridge Univ. Press, Cambridge, 1999.

[88] Stephen B. Pope. *Turbulent flows*. Cambridge University Press, Cambridge, 2000.

[89] Claudia Prévôt and Michael Röckner. *A concise course on stochastic partial differential equations*, volume 1905 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.

[90] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2004. Stochastic Modelling and Applied Probability.

[91] A. Răşcanu and C. Tudor. Approximation of stochastic equations by the splitting up method. In *Qualitative problems for differential equations and control theory*, pages 277–287. World Sci. Publ., River Edge, NJ, 1995.

[92] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.

[93] Frédéric Riesz and Béla Sz.-Nagy. *Leçons d'analyse fonctionnelle*. Gauthier-Villars, Paris, 1955. 3ème éd.

[94] Michael Röckner and Zeev Sobol. A new approach to Kolmogorov equations in infinite dimensions and applications to stochastic generalized Burgers equations. *C. R. Math. Acad. Sci. Paris*, 338(12):945–949, 2004.

[95] Michael Röckner and Zeev Sobol. Kolmogorov equations in infinite dimensions: well-posedness and regularity of solutions, with applications to stochastic generalized Burgers equations. *Ann. Probab.*, 34(2):663–727, 2006.

[96] Michael Röckner and Zeev Sobol. A new approach to Kolmogorov equations in infinite dimensions and applications to the stochastic 2D Navier-Stokes equation. *C. R. Math. Acad. Sci. Paris*, 345(5):289–292, 2007.

[97] Tomáš Roubíček. *Nonlinear partial differential equations with applications*, volume 153 of *International Series of Numerical Mathematics*. Birkhäuser Verlag, Basel, 2005.

[98] David K. Ruch. Completely continuous and related multilinear operators. *Rend. Circ. Mat. Palermo (2)*, 45(3):377–396, 1996.

[99] Walter Rudin. *Functional analysis*. McGraw-Hill Book Co., New York, 1973. McGraw-Hill Series in Higher Mathematics.

[100] Thomas Runst and Winfried Sickel. *Sobolev spaces of fractional order, Nemytskij operators, and nonlinear partial differential equations*, volume 3 of *de Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin, 1996.

[101] D. Schötzau. *hp-DGFEM for parabolic evolution problems*. PhD thesis, ETH Zürich, 1999.

[102] Dominik Schötzau and Christoph Schwab. Time discretization of parabolic problems by the *hp*-version of the discontinuous Galerkin finite element method. *SIAM J. Numer. Anal.*, 38(3):837–875, 2000.

[103] Gilbert Strang. Accurate partial difference methods. I. Linear Cauchy problems. *Arch. Rational Mech. Anal.*, 12:392–402, 1963.

[104] M. Sun and R. Glowinski. Pathwise approximation and simulation for the Zakai filtering equation through operator splitting. *Calcolo*, 30(3):219–239 (1994), 1993.

[105] Hideyuki Tanaka and Arturo Kohatsu-Higa. An operator approach for Markov chain weak approximations with an application to infinite activity Lévy driven SDEs. *Ann. Appl. Probab.*, 19(3):1026–1062, 2009.

[106] Josef Teichmann. Another approach to some rough and stochastic partial differential equations. *Stoch. Dynam.*, 11(2–3):535–550, 2011.

[107] John B. Walsh. An introduction to stochastic partial differential equations. In *École d'été de probabilités de Saint-Flour, XIV—1984*, volume 1180 of *Lecture Notes in Math.*, pages 265–439. Springer, Berlin, 1986.

[108] Kôsaku Yosida. *Functional analysis*, volume 123 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, sixth edition, 1980.

[109] Eberhard Zeidler. *Nonlinear functional analysis and its applications. I.* Springer-Verlag, New York, 1986. Fixed-point theorems, Translated from the German by Peter R. Wadsack.

Dipl.-Ing. Philipp Dörsek MSc
Döblinger Hauptstraße 7A/64, A-1190 Wien
philipp.doersek@math.ethz.ch
+43 676 96 235 69

Geboren am 29. April 1985
Österreichische Staatsbürgerschaft

Berufliche Tätigkeiten

Jan. 2011-heute — **ETH Zürich**
Assistent am Departement für Mathematik

Jan. 2009-Dez. 2010 — **TU Wien und Fonds zur Förderung wissenschaftlicher Forschung**
Forschungsassistent in den Kollegs *Differential Equation Models in Science and Engineering* und *Partial Differential Equations*

Okt. 2008-Dez. 2008 — **Wiener Wissenschafts- und Technologiefonds**
Forschungsassistent im Projekt *Mathematik und Kreditrisiken*

Okt. 2007-Sept. 2008 — **TU Wien**
Übungsgruppenleiter am Institut für Analysis und Scientific Computing und am Institut für Statistik

2002-2006 — **Avalon Computer Ges.m.b.H., Maurer Lange G. 57/3/1, A-1230 Wien**
Ferialpraktikant Systemadministration und Programmierung

Juli 2001 — **IAF-Service Ges.m.b.H.**
Ferialpraktikant Administration

Ausbildung

Okt. 2008-Okt. 2011 — **ETH Zürich und TU Wien**
Doktoratsstudium unter J. Markus Melenk (TU Wien) und Josef Teichmann (ETH Zürich); Thema der Dissertation: Entwicklung effizienter numerischer Methoden zur Simulation von stochastischen partiellen Differentialgleichungen mit Anwendungen in Technik und Finanzmathematik

Okt. 2003-Okt. 2008 — **TU Wien**
Diplom-Ingenieur in *Technischer Mathematik* mit Schwerpunkt *Mathematik in den Naturwissenschaften* mit Auszeichnung

Sept. 2006-Sept. 2007 — **Brunel University West London**
Master of Science with Distinction in *Computational Mathematics with Modelling*

Sept. 1995-Juni 2003 — **Öffentliches Schottengymnasium der Benediktiner in Wien**
Matura mit ausgezeichnetem Erfolg

Sonstige Tätigkeiten

Juli 2009-Juni 2011 **Fachschaft Doktorat TU Wien**
Studienvertreter und Mitglied der Studienkommission

Aug. 2008 **Solidarités Jeunesses**
Mitarbeit am Workcamp Le Créneau in Auvergne, Frankreich

Sept. 2006-Sept. 2007 **Student Representationship Brunel University West London**
Studienvertreter

Okt. 2004-Juni 2006 **Fachschaft Technische Mathematik TU Wien**
Erstsemestrigentutor

Sprachen

Deutsch	Muttersprache
Englisch	Verhandlungssicher
Französisch	Fließend

IT

Anwendungskenntnisse	MS Office, Windows, Linux
Programmiersprachen	Matlab, C, C++, Perl

Wissenschaftliche Veröffentlichungen

- *Semigroup Splitting And Cubature Approximations For The Stochastic Navier-Stokes Equations*, 2011.
- *A Semigroup Point Of View On Splitting Schemes For Stochastic (Partial) Differential Equations*, 2010. (mit J. Teichmann)
- *Adaptive hp-FEM for the contact problem with Tresca friction in linear elasticity: the primal-dual formulation and a posteriori error estimation*, Applied Numerical Mathematics 60, S. 689-704, Elsevier Verlag, 2010. (mit J. M. Melenk)
- *hp-FEM for the contact problem with Tresca friction in linear elasticity: the primal formulation*, in J.S. Hesthaven, E.M. Rønquist, Spectral and High Order Methods for Partial Differential Equations, Lecture Notes in Computational Science and Engineering, Vol. 76 of LNCSE, S. 1-17, Springer Verlag, 2011. (mit J. M. Melenk)
- *hp-Finite Element Methods For Variational Inequalities*, Diplomarbeit, TU Wien, 2008.
- *h- and p-version Finite Element Methods for Elasto-Plasticity*, MSc-Arbeit, Brunel University West London, 2007.