Thesis for the Degree of Licentiate of Philosophy

Variational Methods for Moments of Solutions to Stochastic Differential Equations

Kristin Kirchner

CHALMERS | GÖTEBORGS UNIVERSITET

Division of Applied Mathematics and Statistics
Department of Mathematical Sciences
Chalmers University of Technology
and University of Gothenburg
Gothenburg, Sweden, 2016

Variational Methods for Moments of Solutions to Stochastic Differential Equations
Kristin Kirchner

Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg, Sweden
Phone: +46 (0)31 772 53 57

Author e-mail: `kristin.kirchner@chalmers.se`

Printed in Gothenburg, Sweden 2016

Variational Methods for Moments of Solutions to Stochastic Differential Equations

Kristin Kirchner

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

Abstract

Numerical methods for stochastic differential equations typically estimate moments of the solution from sampled paths. Instead, we pursue the approach proposed by A. Lang, S. Larsson, and Ch. Schwab[1], who derived well-posed deterministic, tensorized evolution equations for the second moment and the covariance of the solution to a parabolic stochastic partial differential equation driven by additive Wiener noise.

In Paper I we consider parabolic stochastic partial differential equations with multiplicative Lévy noise of affine type. For the second moment of the mild solution, a deterministic space-time variational problem is derived. It is posed on projective and injective tensor product spaces as trial and test spaces. Well-posedness is proven under appropriate assumptions on the noise term. From these results, a deterministic equation for the covariance function is deduced.

These deterministic equations in variational form are used in Paper II to derive numerical methods for approximating the first and second moment of the solution to a stochastic ordinary differential equation driven by additive or multiplicative Wiener noise. For the canonical examples with additive noise (Ornstein–Uhlenbeck process) and multiplicative noise (geometric Brownian motion) we first recall the variational problems satisfied by the first and the second moments of the solution processes and discuss their well-posedness in detail. For the considered examples, well-posedness beyond the assumptions on the multiplicative noise term made in Paper I are proven. We propose Petrov–Galerkin discretizations based on tensor product piecewise polynomials and analyze their stability and convergence in the natural norms.

Keywords: Stochastic ordinary and partial differential equations, Additive and multiplicative noise, Space-time variational problem, Hilbert tensor product space, Projective and injective tensor product space, Petrov–Galerkin discretization.

---

[1] A. Lang, S. Larsson, and Ch. Schwab, Covariance structure of parabolic stochastic partial differential equations, Stoch. PDE: Anal. Comp., 1(2013), pp. 351–364.

List of included papers

The following papers are included in this thesis:

I Kristin Kirchner, Annika Lang, and Stig Larsson. *Covariance structure of parabolic stochastic partial differential equations with multiplicative Lévy noise.* Preprint, arXiv:1506.00624.

II Roman Andreev and Kristin Kirchner. *Numerical methods for the 2nd moment of stochastic ODEs.* Preprint, arXiv:1611.02164.

Publication not included in the thesis:

Kristin Kirchner, Karsten Urban, and Oliver Zeeb. *Maxwell's Equations for Conductors with Impedance Boundary Conditions: Discontinuous Galerkin and Reduced Basis Methods.* M2AN, Math. Model. Numer. Anal., 50(6):1763–1787, 2016.

Acknowledgments

For the constant support and his unique way of motivating me, whenever it is necessary, I would like to express my gratitude to Stig Larsson. I would like to thank Annika Lang for the careful reading of this thesis and her valuable comments. I am very grateful to Sonja Cox for agreeing on review the thesis.

During the elaboration of this work many people inside and outside of academia have supported me. I would like to thank Roman Andreev for the collaboration on a common research project, for productive four weeks, while he was visiting Chalmers, and for sharpening my way of expressing mathematics in the sense of Karl R. Popper

Wer's nicht einfach und klar sagen kann,
der soll schweigen und weiterarbeiten, bis er's klar sagen kann.

Thanks are also due to my fellow Ph.D. students, colleagues, and friends Adam A., Alexey L., Anna P., Claes A., Dmitry Z., Jonas W., and Tobias Ö., who have made me feel welcome at the Department of Mathematical Sciences from the beginning. In particular, would like to thank Adam Andersson for fruitful discussions and the organization of interesting reading courses.

I am grateful to my family; to my mother for her practical advice in my daily life; to my father and his wife for sharing the passion for music with me.

Finally, I thank David B. and Henrike H. for being my everyday incentive *to be a rock and not to roll.*

Kristin Kirchner
Gothenburg, November 10, 2016

Contents

Part A

Introduction

Introduction

Many models in, e.g., finance, biology, physics, and social sciences are based on ordinary or partial differential equations. In order to improve their applicability to the reality, one has to take uncertainties into account. These uncertainties can be related to the parameters, to the geometry of the physical domain, to boundary or initial conditions, or to the source terms. In this thesis we focus on the latter scenario. More precisely, we consider ordinary and partial differential evolution equations driven by an additive or multiplicative noise term. Under appropriate assumptions, existence and uniqueness of a solution to such an equation is ensured. This solution is then a square-integrable stochastic process with values in a certain state space. In particular, the finiteness of the first and second moment is guaranteed. In applications, often not the solution process itself, but only its moments are of interest. In the case, when the solution process is Gaussian, its distribution is even completely characterized by the first two moments.

The numerical approximation of moments of the solution process to a stochastic differential equation is typically based on sampling methods such as Monte Carlo. However, Monte Carlo methods are, in general, computationally expensive due to the convergence order $1/2$ of the Monte Carlo estimation and the high cost for computing sample paths of solutions to stochastic differential equations.

An alternative approach has been suggested in [11], where the first and second moment of the solution process to a parabolic stochastic partial differential equation driven by additive Wiener noise have been described as solutions to deterministic evolution equations which can be formulated as well-posed linear space-time variational problems. Hence, instead of estimating moments from computationally expensive sample paths, one can apply numerical methods to the deterministic variational problems satisfied by the first and second moment. The main promise of this approach is in potential savings in computing time and memory through space-time compressive schemes, e.g., using adaptive wavelet methods or low-rank tensor approximations.

The first aim of this thesis is to extend the result of [11] to parabolic stochastic partial differential equation driven by multiplicative Lévy noise in Paper I.

Afterwards, in Paper II, we focus on deriving numerical methods for solving these variational problems in the case of stochastic ordinary differential equations driven by additive or multiplicative Wiener noise.

Throughout the following sections, let H and U be separable Hilbert spaces over $\mathbb{R}$ with respect to the inner products $(\cdot,\cdot)_H$ and $(\cdot,\cdot)_U$, respectively. On H we denote the induced norm by $\|\cdot\|_H$ and similarly for U. If $(E, \|\cdot\|_E)$ is a normed vector space, then $S(E) := \{x \in E : \|x\|_E = 1\}$ denotes its unit sphere and E' its dual, i.e, the space of all linear continuous mappings from E to $\mathbb{R}$.

1. Operator theory and tensor product spaces

In the following we provide the definitions and results from the theory of operators and tensor product spaces which are needed for the derivation of the deterministic equations in Paper I and II.

For this purpose, we first introduce the relevant operator classes in Subsection 1.1, where we focus particularly on the notion of Schatten class operators. In Subsection 1.2 we present different notions of tensor product spaces and some of their properties. The deterministic variational problems derived in Paper I and II are posed on tensor products of vector-valued function spaces as trial and test spaces. The vector spaces needed to define these trial and test spaces are introduced in Subsections 1.3 and 1.4. Finally, in Subsection 1.5 we establish a connection between the notion of Schatten class operators on the one hand and tensor products of Bochner spaces on the other hand.

1.1. Special classes of bounded linear operators. In this subsection we present different classes of linear operators which are of relevance for our analysis. For a detailed overview of operator classes we refer to [4, 5].

1.1.1. *Bounded operators.* A linear operator $T\colon U \to H$ is called bounded or continuous if it has a finite operator norm:

$$\|T\|_{\mathcal{L}(U;H)} := \sup_{x \in S(U)} \|Tx\|_H < +\infty.$$

With the above norm, the space of all continuous linear operators from U to H is a Banach space denoted by $\mathcal{L}(U;H)$. We write $\mathcal{L}(U)$ whenever $U = H$.

1.1.2. *Compact operators.* An operator $T\colon U \to H$ is compact if the image of any bounded set in U (or equivalently the closed unit ball in U) under T is relatively compact in H, meaning that its closure is compact. We denote the set of all compact operators mapping from U to H by $\mathcal{K}(U;H)$ and use the abbreviation $\mathcal{K}(U)$ if $U = H$.

Equivalently, cf. [21, §X.2], one can define the subspace $\mathcal{K}(U;H) \subset \mathcal{L}(U;H)$ as the closure of all finite-rank operators mapping from U to H, i.e., $T\colon U \to H$ is compact if and only if there exists a sequence of finite-rank operators $T_n \in \mathcal{L}(U;H)$ converging to T in the norm topology on $\mathcal{L}(U;H)$:

$$\dim \operatorname{Im}(T_n) < +\infty \quad \forall n \in \mathbb{N}, \qquad \lim_{n\to\infty} \|T - T_n\|_{\mathcal{L}(U;H)} = 0.$$

In our analysis we use the latter characterization of compact operators.

We introduce the adjoint $T^*\colon H \to U$ of a linear operator $T\colon U \to H$, i.e.,

$$(Tx, \phi)_H = (x, T^*\phi)_U \quad \forall x \in U,\ \forall \phi \in H.$$

If $U = H$ and $T^* = T$ then T is called self-adjoint.

Self-adjoint compact operators have real-valued spectra and they generate orthonormal bases consisting of eigenvectors, see [5, Cor. X.3.5]. Since we refer to this property several times, we summarize it in the following theorem.

THEOREM 1.1 (Spectral theorem for self-adjoint compact operators). *Let $T \in \mathcal{K}(U)$ be self-adjoint. Then there exists an orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$ of U and a real-valued sequence $\{\gamma_n\}_{n\in\mathbb{N}}$, which has 0 as its only accumulation point, such that $Te_n = \gamma_n e_n$ for all $n \in \mathbb{N}$.*

1.1.3. *Schatten class operators.* A continuous linear operator $T \in \mathcal{L}(U; H)$ is called a Schatten class operator of p-th order or a p-Schatten class operator for $p \in [1, \infty)$ if T has a finite p-Schatten norm:

$$\|T\|_{\mathcal{L}_p(U;H)} := \left(\sum_{n\in\mathbb{N}} s_n(T)^p \right)^{1/p} < +\infty,$$

where

$$s_1(T) \ge s_2(T) \ge \ldots \ge s_n(T) \ge \ldots \ge 0$$

are the singular values of T, i.e., the eigenvalues of the operator $|T| := (T^*T)^{1/2}$. The space of all Schatten class operators of p-th order mapping from U to H denoted by $\mathcal{L}_p(U; H)$ is a Banach space with respect to $\| \cdot \|_{\mathcal{L}_p(U;H)}$. Again, we use an abbreviation when $U = H$ and write $\mathcal{L}_p(U)$ in this case. The Schatten norm is monotone in p, i.e.,

$$\|T\|_{\mathcal{L}_1(U;H)} \ge \|T\|_{\mathcal{L}_p(U;H)} \ge \|T\|_{\mathcal{L}_{p'}(U;H)} \ge \|T\|_{\mathcal{L}(U;H)}$$

for $1 \le p \le p' < +\infty$ and, moreover, every Schatten class operator is compact. Therefore, the introduced operator spaces satisfy the following relation:

$$\mathcal{L}_1(U; H) \subset \mathcal{L}_p(U; H) \subset \mathcal{L}_{p'}(U; H) \subset \mathcal{K}(U; H) \subset \mathcal{L}(U; H).$$

1.1.4. *Trace class and Hilbert–Schmidt operators.* Schatten class operators of first order mapping from U into U are also called trace class operators. Their name originates from the following fact: For $T \in \mathcal{L}_1(U)$ the trace defined by

$$\operatorname{tr}(T) := \sum_{n\in\mathbb{N}} (Te_n, e_n)_U$$

is finite and independent of the choice of the orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$ of U. Moreover, it holds $|\operatorname{tr}(T)| \le \operatorname{tr}(|T|) = \|T\|_{\mathcal{L}_1(U)}$, cf. [2, Prop. C.1]. For self-adjoint, nonnegative trace class operators the trace coincides with the 1-Schatten norm, i.e., $\operatorname{tr}(T) = \|T\|_{\mathcal{L}_1(U)}$ for all $T \in \mathcal{L}_1^+(U)$, where

$$\mathcal{L}_1^+(U) := \{T \in \mathcal{L}_1(U) : T^* = T,\ (Tx, x)_U \ge 0 \quad \forall x \in U\}.$$

This is due to the equality of $|T|$ and T for $T \in \mathcal{L}_1^+(U)$.

The 2-Schatten norm of $T : U \to H$ satisfies

$$\|T\|^2_{\mathcal{L}_2(U;H)} = \operatorname{tr}(T^*T) = \sum_{n\in\mathbb{N}} \|Te_n\|^2_H$$

for any orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$ of U. In contrast to all other Schatten norms when $p \neq 2$, this norm originates from the inner product given by

$$(S,T)_{\mathcal{L}_2(U;H)} := \sum_{n\in\mathbb{N}} (Se_n, Te_n)_H, \quad S,T \in \mathcal{L}_2(U;H),$$

which is often referred to as the Hilbert–Schmidt inner product between S and T. The 2-Schatten norm is known as Hilbert–Schmidt norm and 2-Schatten class operators are also called Hilbert–Schmidt operators.

1.2. Tensor product spaces. In addition to H and U, let also $(\widetilde{H}, (\cdot,\cdot)_{\widetilde{H}})$ and $(\widetilde{U}, (\cdot,\cdot)_{\widetilde{U}})$ denote separable Hilbert spaces over $\mathbb{R}$. We first introduce the algebraic tensor product $H \otimes \widetilde{H}$ as the vector space consisting of all finite sums of the form

$$\sum_{k=1}^{N} \phi_k \otimes \widetilde{\phi}_k, \quad \phi_k \in H, \, \widetilde{\phi}_k \in \widetilde{H}, \, k = 1, \ldots, N,$$

equipped with the obvious algebraic operations. There are several ways to define a norm on this vector space, and taking the closure with respect to the different norms yields different Banach spaces. For our purposes, the three notions of tensor products introduced below are important. We refer to [8, 17] for a general introduction into the theory of tensor product spaces.

1.2.1. *Hilbert tensor product.* The Hilbert tensor product space $H \hat{\otimes}_2 \widetilde{H}$ is defined as the completion of the algebraic tensor product space $H \otimes \widetilde{H}$ with respect to the norm induced by the inner product

$$(\Phi, \Psi)_{H\hat{\otimes}_2\widetilde{H}} := \sum_{k=1}^{N}\sum_{\ell=1}^{M} (\phi_k, \psi_\ell)_H (\widetilde{\phi}_k, \widetilde{\psi}_\ell)_{\widetilde{H}},$$

which is independent of the choice of the representations $\Phi = \sum_{k=1}^{N} \phi_k \otimes \widetilde{\phi}_k$ and $\Psi = \sum_{\ell=1}^{M} \psi_\ell \otimes \widetilde{\psi}_\ell$ of Φ, $\Psi \in H \otimes \widetilde{H}$. For $H = \widetilde{H}$ we abbreviate the notation for this space by $H_2 := H \hat{\otimes}_2 H$ and for the inner product by $(\cdot,\cdot)_2$ as well as for the corresponding norm by $\|\cdot\|_2$.

1.2.2. *Projective tensor product.* The projective tensor product space denoted by $H \hat{\otimes}_\pi \widetilde{H}$ is obtained by taking the closure of the algebraic tensor product space $H \otimes \widetilde{H}$ with respect to the projective norm defined for $\Phi \in H \otimes \widetilde{H}$ by

$$\|\Phi\|_{H\hat{\otimes}_\pi\widetilde{H}} := \inf\left\{ \sum_{k=1}^{N} \|\phi_k\|_H \|\widetilde{\phi}_k\|_{\widetilde{H}} : \Phi = \sum_{k=1}^{N} \phi_k \otimes \widetilde{\phi}_k \right\}.$$

We write $H_\pi := H \hat{\otimes}_\pi H$ and $\|\cdot\|_\pi := \|\cdot\|_{H\hat{\otimes}_\pi H}$, whenever $H = \widetilde{H}$.

1.2.3. *Injective tensor product.* The injective norm of an element Φ in the algebraic tensor product space $H \otimes \widetilde{H}$ is defined as

$$\|\Phi\|_{H\hat{\otimes}_\varepsilon\widetilde{H}} := \sup\left\{ \left| \sum_{k=1}^{N} f(\phi_k)\, g(\widetilde{\phi}_k) \right| : f \in S(H'), \, g \in S(\widetilde{H}') \right\},$$

where $\sum_{k=1}^N \phi_k \otimes \widetilde{\phi}_k$ is any representation of $\Phi \in H \otimes \widetilde{H}$. Note that the value of the supremum is independent of the choice of the representation of Φ, cf. [17, p. 45]. The completion of $H \otimes \widetilde{H}$ with respect to this norm is called injective tensor product space and denoted by $H \mathbin{\hat{\otimes}_\varepsilon} \widetilde{H}$. If $H = \widetilde{H}$, the abbreviations $H_\epsilon := H \mathbin{\hat{\otimes}_\varepsilon} H$ as well as $\|\cdot\|_\epsilon := \|\cdot\|_{H \hat{\otimes}_\varepsilon H}$ are used.

1.2.4. *Some remarks.* We note that the projective and injective tensor product spaces in 1.2.2 and 1.2.3 are Banach spaces, which are in general not reflexive, cf. [17, Thm. 4.21].

An immediate consequence of the definition of the above norms is the following chain of continuous embeddings, see [17, Prop. 6.1(a)]:

$$H \mathbin{\hat{\otimes}_\pi} \widetilde{H} \hookrightarrow H \mathbin{\hat{\otimes}_2} \widetilde{H} \hookrightarrow H \mathbin{\hat{\otimes}_\varepsilon} \widetilde{H}.$$

Here, the embedding constants are all equal to 1.

Another important fact when dealing with linear operators on tensor product spaces is the following: For $T \in \mathcal{L}(U; H)$ and $S \in \mathcal{L}(\widetilde{U}; \widetilde{H})$ setting

$$(T \otimes S)(x \otimes \widetilde{x}) = (Tx) \otimes (S\widetilde{x}), \quad x \in U, \ \widetilde{x} \in \widetilde{U},$$

and extending this definition by linearity to elements in $U \otimes \widetilde{U}$ yields a well-defined linear operator $T \otimes S$ mapping between the algebraic tensor product spaces $U \otimes \widetilde{U}$ and $H \otimes \widetilde{H}$. This operator admits a unique extension to a continuous linear operator $T \hat{\otimes}_\iota S \colon U \hat{\otimes}_\iota \widetilde{U} \to H \hat{\otimes}_\iota \widetilde{H}$ and it holds

$$\|T \hat{\otimes}_\iota S\|_{\mathcal{L}(U \hat{\otimes}_\iota \widetilde{U}; H \hat{\otimes}_\iota \widetilde{H})} = \|T\|_{\mathcal{L}(U;H)} \|S\|_{\mathcal{L}(\widetilde{U};\widetilde{H})}$$

for all types of tensor spaces $\iota \in \{2, \pi, \epsilon\}$ considered above, see [17, Propositions 2.3 & 3.2] and Lemma 3.1 (ii) in Paper I.

1.3. Self-adjoint unbounded operators and semigroups. In the following we recall the notions of C_0-semigroups on Hilbert spaces, their generators, as well as fractional powers of generators and the Hilbert spaces they induce. For a general introduction to semigroup theory we refer to [6, 15].

Let $A \colon \mathcal{D}(A) \to H$ be a linear operator defined on a dense subspace $\mathcal{D}(A)$ of H. Furthermore we assume that A is self-adjoint and positive definite, i.e.,

$$(A\phi, \psi)_H = (\phi, A\psi)_H \quad \text{and} \quad (A\vartheta, \vartheta)_H > 0$$

for all $\phi, \psi, \vartheta \in \mathcal{D}(A)$, $\vartheta \neq 0$, and that A has a compact inverse $A^{-1} \in \mathcal{K}(H)$. The application of Theorem 1.1 to A^{-1} shows that there exists an orthonormal basis $\{e_n\}_{n\in\mathbb{N}}$ of H consisting of eigenvectors of A and an increasing sequence of corresponding eigenvalues $\{\lambda_n\}_{n\in\mathbb{N}} \subset \mathbb{R}_{>0}$, i.e., $Ae_n = \lambda_n e_n$ for all $n \geq 1$.

In particular, $-A$ is a closed operator, densely defined on $\mathcal{D}(-A) = \mathcal{D}(A)$ and its spectrum consists only of negative real numbers. By the Hille–Yosida theorem [6, Thm. 3.5], it is thus the generator of a C_0-semigroup $(S(t), t \geq 0)$. Here, $(S(t), t \geq 0)$ is called a C_0-semigroup or strongly continuous semigroup if the following hold for all $\phi \in H$ and $t_1, t_2 \geq 0$:

$$\begin{aligned} S(0)\phi &= \phi, \\ S(t_1 + t_2) &= S(t_1)S(t_2), \\ \lim_{t\to 0^+} S(t)\phi &= \phi, \end{aligned}$$

and $-A$ is its generator if

$$\lim_{t\to 0^+} \frac{S(t)\phi - \phi}{t} = -A\phi \quad \forall \phi \in \mathcal{D}(A),$$

where $\lim_{t\to 0^+}$ denotes the one-sided limit from above 0 in H. Moreover, its elements are contractions (i.e., $\|S(t)\|_{\mathcal{L}(H)} \leq 1$ for all $t \geq 0$), the semigroup is analytic [6, Cor. 4.7], and for $r \geq 0$ the fractional power operator

$$A^{r/2} \colon \mathcal{D}(A^{r/2}) \to H, \quad A^{r/2}\phi := \sum_{n\in\mathbb{N}} \lambda_n^{r/2} (\phi, e_n)_H \, e_n$$

is well-defined on a domain $\mathcal{D}(A^{r/2}) \subset H$, cf. [15, Ch. 2], which can be characterized as

$$\mathcal{D}(A^{r/2}) = \left\{ \phi \in H : \sum_{n\in\mathbb{N}} \lambda_n^r (\phi, e_n)_H^2 < +\infty \right\}.$$

We define the Hilbert space $\dot{H}^r$ as the closure of $\mathcal{D}(A^{r/2})$ with respect to the norm induced by the inner product

$$(\phi, \psi)_{\dot{H}^r} := (A^{r/2}\phi, A^{r/2}\psi)_H,$$

and introduce the space $\dot{H}^{-r}$ as the identification of the dual space of $\dot{H}^r$ via the inner product on H. In this way we obtain the following scale of Hilbert spaces:

$$\dot{H}^s \hookrightarrow \dot{H}^r \hookrightarrow \dot{H}^0 \cong H \cong \dot{H}^{-0} \hookrightarrow \dot{H}^{-r} \hookrightarrow \dot{H}^{-s}$$

for $0 \leq r \leq s$. We denote $H^* := \dot{H}^{-0}$ as well as $V := \dot{H}^1$ and $V^* := \dot{H}^{-1}$ for the case $r = 1$ and emphasize the following Gelfand triple of densely embedded Hilbert spaces

$$V \hookrightarrow H \cong H^* \hookrightarrow V^*.$$

The dual pairing between V^* and V is denoted by ${}_{V^*}\langle \cdot, \cdot \rangle_V$ or by ${}_V\langle \cdot, \cdot \rangle_{V^*}$. We stress the difference between the spaces V' and V^*: V' denotes the dual of V in its original sense, i.e., all continuous linear mappings from V to $\mathbb{R}$, while V^* is the identification of the dual via the inner product on H. Therefore, we have the relation: $\phi \in V^*$ if and only if ${}_{V^*}\langle \phi, \cdot \rangle_V \in V'$.

1.4. Bochner spaces. In order to define the trial and test spaces for the variational formulations of the evolution equations, we introduce a special class of function spaces: Bochner or Lebesgue–Bochner spaces. For the definition of the Bochner integral we refer to [21, §V.5].

For finite $T > 0$ and $p \in [1, \infty)$ we consider the time interval $J := (0, T)$ and the Bochner space $L_p(J; H)$ of Bochner-measurable, p-integrable functions mapping from J to the Hilbert space H, which is itself a Banach space with respect to the norm

$$\|u\|_{L_p(J;H)} := \left(\int_J \|u(t)\|_H^p \, \mathrm{d}t \right)^{1/p}.$$

For future reference, we introduce the abbreviation $\mathcal{W} := L_2(J;H)$ for the case $p = 2$, and note that $\mathcal{W}$ is a Hilbert space with respect to the obvious inner product inducing the norm $\|\cdot\|_{L_2(J;H)}$.

Let $u \in L_1(J;H)$ be an H-valued Bochner-integrable function. Following [3, Ch. XVIII, §1, Def. 3] we define the distributional derivative $\partial_t u$ of u as the H-valued distribution satisfying

$$((\partial_t u)(v), \phi)_H = -\int_J \frac{\mathrm{d}v}{\mathrm{d}t}(t)(u(t),\phi)_H \, \mathrm{d}t \quad \forall (v,\phi) \in C_0^\infty(J;\mathbb{R}) \times H.$$

Recall the Gelfand triple $V \hookrightarrow H \hookrightarrow V^*$ from the previous Subsection 1.3. The definition of the distributional derivative of an H-valued function above implies that the for a V^*-valued function u the distributional derivative $\partial_t u$ is defined as the V^*-valued distribution with

$${}_{V^*}\langle (\partial_t u)(v), \phi\rangle_V = -\int_J \frac{\mathrm{d}v}{\mathrm{d}t}(t) \, {}_{V^*}\langle u(t), \phi\rangle_V \, \mathrm{d}t \quad \forall (v,\phi) \in C_0^\infty(J;\mathbb{R}) \times V.$$

After having defined the Bochner space $L_2(J;V^*)$ and the distributional derivative we can now introduce the space

$$H^1(J;V^*) := \{u \in L_2(J;V^*) : \partial_t u \in L_2(J;V^*)\}$$

and equip it with the norm

$$\|u\|_{H^1(J;V^*)} := \left(\|u\|_{L_2(J;V^*)}^2 + \|\partial_t u\|_{L_2(J;V^*)}^2 \right)^{1/2}.$$

With respect to the corresponding inner product, $H^1(J;V^*)$ is a Hilbert space.

In the following the formulation of the variational problems is based on trial and test spaces originating from the spaces

$$\mathcal{X} := L_2(J;V), \qquad \widehat{\mathcal{Y}} := L_2(J;V) \cap H^1(J;V^*).$$

These spaces are Hilbert spaces: $\mathcal{X}$ with respect to the Bochner inner product $(\cdot,\cdot)_{\mathcal{X}} = (\cdot,\cdot)_{L_2(J;V)}$ and $\widehat{\mathcal{Y}}$ being equipped with the graph norm

$$\|v\|_{\mathcal{Y}} := \left(\|v\|_{L_2(J;V)}^2 + \|\partial_t v\|_{L_2(J;V^*)}^2 \right)^{1/2}$$

and the obvious corresponding inner product.

It is a well-known result [3, Ch. XVIII, §1, Thm. 1] that $\widehat{\mathcal{Y}} \hookrightarrow C(\bar{J};H)$, where $C(\bar{J};H)$ denotes the space of continuous H-valued functions on the closure $\bar{J} := [0,T]$ of J. Therefore, the values $v(0)$ and $v(T)$ are are well-defined in H for $v \in \widehat{\mathcal{Y}}$ and the following are closed subspaces of $\widehat{\mathcal{Y}}$:

$$\mathcal{Y}_0 := \{v \in \widehat{\mathcal{Y}} : v(0) = 0 \text{ in } H\}, \quad \mathcal{Y} := \{v \in \widehat{\mathcal{Y}} : v(T) = 0 \text{ in } H\},$$

which are equipped with the same norm $\|\cdot\|_{\mathcal{Y}}$ as $\widehat{\mathcal{Y}}$. We note that

$$|\!|\!|v|\!|\!|_{\mathcal{Y}} := \Big(\|v\|_{L_2(J;V)}^2 + \|\partial_t v\|_{L_2(J;V^*)}^2 + \|v(0)\|_H^2\Big)^{1/2}$$

defines an equivalent norm on $\mathcal{Y}$, which has the following advantages:

(i) The embedding constant in $\mathcal{Y} \hookrightarrow C(\bar{J}; H)$ is smaller: Fix $t \in J$. The integration of $\frac{\mathrm{d}}{\mathrm{d}t}\|v(t)\|_H^2 = 2\,{}_{V^*}\langle \partial_t v(t), v(t)\rangle_V$ for $v \in \mathcal{Y}$ over the interval $[t, T]$ yields (recall that $v(T) = 0$ for $v \in \mathcal{Y}$):

$$\begin{aligned}\|v(t)\|_H^2 &\le 2\,\|\partial_t v\|_{L_2((t,T);V^*)}\|v\|_{L_2((t,T);V)} \\ &\le \|v\|_{L_2((t,T);V)}^2 + \|\partial_t v\|_{L_2((t,T);V^*)}^2.\end{aligned}$$

This already shows that the embedding constant with respect to $\|\cdot\|_{\mathcal{Y}}$ is bounded by 1. By integrating from 0 to t instead we obtain:

$$\|v(t)\|_H^2 \le \|v(0)\|_H^2 + \|v\|_{L_2((0,t);V)}^2 + \|\partial_t v\|_{L_2((0,t);V^*)}^2.$$

Adding the two inequalities shows that $\|v(t)\|_H \le \frac{1}{\sqrt{2}}|\!|\!|v|\!|\!|_{\mathcal{Y}}$. For sharpness of these bounds we refer to the example in §2.2 in Paper II.

(ii) If we define the evolution operator $b\colon \mathcal{X} \to \mathcal{Y}'$ by

$$(bu)(v) := \int_J {}_V\langle u(t), (-\partial_t + A)v(t)\rangle_{V^*}\,\mathrm{d}t,$$

then $b \in \mathcal{L}(\mathcal{X}; \mathcal{Y}')$ is an isometry, i.e., $|\!|\!|bu|\!|\!|_{\mathcal{Y}'} = \|u\|_{\mathcal{X}}$, where $|\!|\!|f|\!|\!|_{\mathcal{Y}'} := \sup_{v\in\mathcal{Y}\setminus\{0\}} \frac{|f(v)|}{|\!|\!|v|\!|\!|_{\mathcal{Y}}}$, see Subsection 2.2.

In particular, since the latter property is useful for the error analysis of numerical methods, the results in Paper II are formulated with respect to the norm $|\!|\!|\cdot|\!|\!|_{\mathcal{Y}}$, while in, e.g., [11, 19, 20] and also in Paper I the norm $\|\cdot\|_{\mathcal{Y}}$ is used.

1.5. Relating tensor product spaces and Schatten class operators. In the following we establish a connection between the tensor products $\mathcal{W}_\pi$ and $\mathcal{W}_2$ of the Bochner space $\mathcal{W} = L_2(J; H)$ on the one hand, and Schatten class operators of order 1 and 2 on the other hand. In addition, we show that compact operators in $\mathcal{K}(\mathcal{W})$ are related to the elements in the injective tensor product space $\mathcal{W}_\epsilon$.

For this purpose we first define an $H \otimes H$-valued kernel k on $J \times J$ as an element in the algebraic tensor product space $\mathcal{W} \otimes \mathcal{W}$, which is, in addition, symmetric, i.e.,

$$k(s,t) = k(t,s) \quad \text{for a.e. } (s,t) \in J \times J.$$

Due to the Riesz representation theorem, we can define the action of the linear integral operator $T_k\colon \mathcal{W} \to \mathcal{W}$ associated with the kernel k on $w \in \mathcal{W}$ as the unique element $T_k w \in \mathcal{W}$ satisfying

$$(T_k w, v)_{\mathcal{W}} = \int_J\int_J (k(s,t), w(s) \otimes v(t))_{H_2}\,\mathrm{d}s\,\mathrm{d}t = (k, w \otimes v)_{\mathcal{W}_2} \quad \forall v \in \mathcal{W}.$$

The next proposition illustrates the relation between the introduced tensor product spaces and compact / 1-Schatten class / 2-Schatten class operators.

Proposition 1.2. *Let k be an $H\otimes H$-valued kernel k on $J\times J$. The integral operator $T_k\colon \mathcal{W}\to\mathcal{W}$ associated with the kernel k is self-adjoint and satisfies*

$$\|T_k\|_{\mathcal{L}(\mathcal{W})}=\|k\|_{\epsilon},\quad \|T_k\|_{\mathcal{L}_1(\mathcal{W})}=\|k\|_{\pi},\quad \text{and}\quad \|T_k\|_{\mathcal{L}_2(\mathcal{W})}=\|k\|_2.$$

In other words, the mapping $\mathcal{J}\colon k\mapsto T_k$ extends to an isometric isomorphism between the spaces

$$\mathcal{W}_{\epsilon}^{\mathrm{sym}}\overset{\mathcal{J}}{\cong}\mathcal{K}^{\mathrm{sym}}(\mathcal{W}),\quad \mathcal{W}_{\pi}^{\mathrm{sym}}\overset{\mathcal{J}}{\cong}\mathcal{L}_1^{\mathrm{sym}}(\mathcal{W}),\quad \text{and}\quad \mathcal{W}_{2}^{\mathrm{sym}}\overset{\mathcal{J}}{\cong}\mathcal{L}_2^{\mathrm{sym}}(\mathcal{W}),$$

where the superscript $^{\mathrm{sym}}$ indicates for the tensor product spaces to take the closure of only the symmetric functions in the algebraic tensor product space $\mathcal{W}\otimes\mathcal{W}$ with respect to the tensor norms and for the operator spaces the closed subspaces of self-adjoint operators.

Proof. Self-adjointness of T_k follows from the symmetry of the kernel k. In order to prove the identities of the norms, let $k=\sum_{n=1}^{N}k_n^1\otimes k_n^2$ be any representation of $k\in\mathcal{W}\otimes\mathcal{W}$. Then we obtain for the operator norm of the induced operator T_k:

$$\begin{aligned}\|T_k\|_{\mathcal{L}(\mathcal{W})}&=\sup_{w\in S(\mathcal{W})}\|T_k w\|_{\mathcal{W}}=\sup_{w,v\in S(\mathcal{W})}(T_k w,v)_{\mathcal{W}}\\&=\sup_{w,v\in S(\mathcal{W})}\sum_{n=1}^{N}(k_n^1,w)_{\mathcal{W}}(k_n^2,v)_{\mathcal{W}}=\sup_{f,g\in S(\mathcal{W}')}\sum_{n=1}^{N}f(k_n^1)g(k_n^2)=\|k\|_{\epsilon}.\end{aligned}$$

In this calculation the Riesz representation theorem justifies taking the supremum over $f,g\in S(\mathcal{W}')$ instead of over $w,v\in S(\mathcal{W})$. Therefore, the integral operator $T_k\colon\mathcal{W}\to\mathcal{W}$ is continuous if and only if its kernel k is an element of the tensor product space $\mathcal{W}_{\epsilon}^{\mathrm{sym}}$. The identity

$$(T_k w,v)_{\mathcal{W}}=\sum_{n=1}^{N}(k_n^1,w)_{\mathcal{W}}(k_n^2,v)_{\mathcal{W}}=\Big(\sum_{n=1}^{N}(k_n^1,w)_{\mathcal{W}}\,k_n^2,v\Big)_{\mathcal{W}}\quad\forall v\in\mathcal{W}$$

shows that $T_k w=\sum_{n=1}^{N}(k_n^1,w)_{\mathcal{W}}\,k_n^2$ for all $w\in\mathcal{W}$, i.e., T_k is a finite-rank operator and thus compact if $k\in\mathcal{W}\otimes\mathcal{W}$. In the more general case when $k\in\mathcal{W}_{\epsilon}^{\mathrm{sym}}$, we can find a sequence of kernels in $\mathcal{W}\otimes\mathcal{W}$ converging to k with respect to the injective norm $\|\cdot\|_{\epsilon}$. Due to the isometry property derived above, also T_k can be approximated by self-adjoint finite-rank operators in $\mathcal{L}(\mathcal{W})$ and, hence, $T_k\in\mathcal{K}^{\mathrm{sym}}(\mathcal{W})$, see Subsection 1.1.2.

The application of Theorem 1.1 to $T_k\in\mathcal{K}^{\mathrm{sym}}(\mathcal{W})$ yields the existence of an orthonormal basis $\{e_i\}_{i\in\mathbb{N}}$ of $\mathcal{W}$ consisting of eigenvectors of T_k with corresponding eigenvalues $\{\gamma_i\}_{i\in\mathbb{N}}\subset\mathbb{R}$. The observation $(k,e_i\otimes e_j)_{\mathcal{W}_2}=(T_k e_i,e_j)_{\mathcal{W}}=\delta_{ij}\gamma_i$, where δ_{ij} denotes the Kronecker delta, shows that k can be expanded in $\mathcal{W}_2$ as $k=\sum_{i\in\mathbb{N}}\gamma_i(e_i\otimes e_i)$ and we obtain the estimate

$$\|k\|_{\pi}\le\sum_{i\in\mathbb{N}}|\gamma_i|=\mathrm{tr}(|T_k|)=\|T_k\|_{\mathcal{L}_1(\mathcal{W})}.$$

The reverse inequality follows from the Cauchy–Schwarz inequality for sums and Parseval's identity:

$$\begin{aligned}\|T_k\|_{\mathcal{L}_1(\mathcal{W})} &= \sum_{i\in\mathbb{N}} |\gamma_i| = \sum_{i\in\mathbb{N}} |(k, e_i \otimes e_i)_{\mathcal{W}_2}| = \sum_{i\in\mathbb{N}} \left| \sum_{n=1}^{N} (k_n^1, e_i)_{\mathcal{W}} (k_n^2, e_i)_{\mathcal{W}} \right| \\ &\leq \sum_{n=1}^{N} \left(\sum_{i\in\mathbb{N}} (k_n^1, e_i)_{\mathcal{W}}^2 \right)^{1/2} \left(\sum_{i\in\mathbb{N}} (k_n^2, e_i)_{\mathcal{W}}^2 \right)^{1/2} = \sum_{n=1}^{N} \|k_n^1\|_{\mathcal{W}} \|k_n^2\|_{\mathcal{W}}.\end{aligned}$$

Since the representation of k is arbitrary, we may take the infimum over all representations of k in $\mathcal{W} \otimes \mathcal{W}$ and obtain $\|T_k\|_{\mathcal{L}_1(\mathcal{W})} \leq \|k\|_\pi$. Thus, $\mathcal{J}$ extends to an isometric isomorphism between the spaces $\mathcal{W}_\pi^{\mathrm{sym}}$ and $\mathcal{L}_1^{\mathrm{sym}}(\mathcal{W})$.

For the Hilbert–Schmidt norm of T_k we calculate

$$\begin{aligned}\|T_k\|_{\mathcal{L}_2(\mathcal{W})}^2 &= \sum_{i\in\mathbb{N}} \|T_k e_i\|_{\mathcal{W}}^2 = \sum_{i\in\mathbb{N}} \sum_{j\in\mathbb{N}} (T_k e_i, e_j)_{\mathcal{W}}^2 \\ &= \sum_{i\in\mathbb{N}} \sum_{j\in\mathbb{N}} \sum_{m=1}^{N} \sum_{n=1}^{N} (k_m^1, e_i)_{\mathcal{W}} (k_n^1, e_i)_{\mathcal{W}} (k_m^2, e_j)_{\mathcal{W}} (k_n^2, e_j)_{\mathcal{W}} \\ &= \sum_{m=1}^{N} \sum_{n=1}^{N} (k_m^1, k_n^1)_{\mathcal{W}} (k_m^2, k_n^2)_{\mathcal{W}} = \|k\|_2^2.\end{aligned}$$

The space of self-adjoint Hilbert–Schmidt operators $\mathcal{L}_2^{\mathrm{sym}}(\mathcal{W})$ is thus isometrically isomorphic to the Hilbert tensor product space $\mathcal{W}_2^{\mathrm{sym}}$. □

We note that Proposition 1.2 can be seen as a generalization of Mercer's theorem [13], which considers continuous real-valued kernels $k \in C(\bar{J} \times \bar{J}; \mathbb{R})$.

In the literature, kernels in $\mathcal{W}_\pi$ and $\mathcal{W}_2 \cong L^2(J \times J; H_2)$ are often called Fredholm kernels [7] and Hilbert–Schmidt kernels [21, §VII.3, Example 1], respectively.

As a by-product of Proposition 1.2 we obtain an explicit way to calculate the projective norm of a kernel $k \in \mathcal{W}_\pi^{\mathrm{sym}}$ which is positive semi-definite, i.e.,

$$\int_J \int_J (k(s,t), v(s) \otimes v(t))_{H_2} \, \mathrm{d}s \, \mathrm{d}t \geq 0 \quad \forall v \in \mathcal{W}.$$

For this purpose, we first introduce the real-valued operator δ on the algebraic tensor product space $\mathcal{W} \otimes \mathcal{W}$ as follows: if $k = \sum_{n=1}^{N} k_n^1 \otimes k_n^2$ is any representation of $k \in \mathcal{W} \otimes \mathcal{W}$, then

$$\delta(k) := \sum_{n=1}^{N} \int_J (k_n^1(t), k_n^2(t))_H \, \mathrm{d}t.$$

The operator δ is bounded with respect to the projective norm:

$$|\delta(k)| \leq \sum_{n=1}^{N} \int_J \|k_n^1(t)\|_H \|k_n^2(t)\|_H \, \mathrm{d}t \leq \sum_{n=1}^{N} \|k_n^1\|_{\mathcal{W}} \|k_n^2\|_{\mathcal{W}},$$

where this estimate holds for any representation of $k \in \mathcal{W} \otimes \mathcal{W}$. For this reason and due to the density of $\mathcal{W} \otimes \mathcal{W}$ in $\mathcal{W}_\pi$, there exists a unique linear continuous extension of δ to a functional on the projective tensor product space $\mathcal{W}_\pi$.

If the kernel $k \in \mathcal{W}_\pi^{\mathrm{sym}}$ is positive semi-definite, then the associated integral operator $T_k \colon \mathcal{W} \to \mathcal{W}$ is an element of $\mathcal{L}_1^+(\mathcal{W})$ and, as in the proof of Proposition 1.2, its kernel admits an expansion $\sum_{i \in \mathbb{N}} \gamma_i (e_i \otimes e_i)$ in $\mathcal{W}_\pi$, where the coefficients γ_i are the nonnegative eigenvalues of $T_k \in \mathcal{L}_1^+(\mathcal{W})$. Therefore, we obtain the identity

$$\|T_k\|_{\mathcal{L}_1(\mathcal{W})} = \sum_{i \in \mathbb{N}} \gamma_i = \sum_{i \in \mathbb{N}} \gamma_i (e_i, e_i)_{\mathcal{W}} = \sum_{i \in \mathbb{N}} \gamma_i \, \delta(e_i \otimes e_i) = \delta(k),$$

and the projective norm of a positive semi-definite kernel $k \in \mathcal{W}_\pi^{\mathrm{sym}}$ is given by

$$\|k\|_\pi = \|T_k\|_{\mathcal{L}_1(\mathcal{W})} = \delta(k).$$

2. Deterministic initial value problems

We consider the abstract inhomogeneous initial value problem

$$\text{(IVP)} \qquad u'(t) + Au(t) = f(t), \quad t \in J = (0, T), \qquad u(0) = u_0,$$

for a right-hand side $f \in L_1(J; H)$ and an initial data $u_0 \in H$. Here, $u'(t) \in H$ denotes the strong derivative of the H-valued function u, i.e.,

$$u'(t) = \lim_{h \to 0} \frac{u(t+h) - u(t)}{h} \quad \forall t \in J.$$

We assume that $-A \colon \mathcal{D}(A) \to H$ is the generator of an analytic C_0-semigroup of contractions $(S(t), t \geq 0)$, see Subsection 1.3. In the homogeneous case when $f = 0$, Problem (IVP) is often referred to as the Cauchy problem (relative to the operator $-A$) in the literature.

A classical solution u to (IVP) on $[0, T)$ is an H-valued function which is continuous on $[0, T)$, continuously differentiable on J, takes values in $\mathcal{D}(A)$ on J, i.e.,

$$u \in C([0, T); H) \cap C((0, T); \mathcal{D}(A)), \quad u' \in C((0, T); H),$$

and satisfies (IVP). One knows [15, Ch. 4, Cor. 3.3] that the homogeneous initial value problem has a unique classical solution for every initial data $u_0 \in H$, since $-A$ is the generator of an analytic C_0-semigroup. It is given by $u(t) = S(t)u_0$ for all $t \in [0, T)$. From this result, it is evident that the inhomogeneous problem in (IVP) has at most one solution. However, for a general right-hand side $f \in L_1(J; H)$ the definition of a classical solution is often too restrictive in order to ensure existence.

Consider, e.g., the simple example when $A = 0$. Then the initial value problem (IVP) does not have a classical solution unless f is continuous. But also continuity of f on $\bar{J} = [0, T]$ is not sufficient to guarantee existence of a classical solution to (IVP) when $-A$ is the generator of a C_0-semigroup – for a counterexample see [12, Example 4.1.7].

For this reason, generalized solution concepts have been introduced. Note that for $f \in L_1(J; H)$ and $A = 0$, the initial value problem (IVP) always has a solution, which is differentiable almost everywhere and satisfies $u'(t) = f(t)$ for almost every $t \in J$. Namely, the function given by $u(t) = u_0 + \int_0^t f(s)\,\mathrm{d}s$ satisfies these properties. This motivates the following notion of a strong solution to (IVP) for a general generator $-A$, see [15, Ch. 4, Def. 2.8].

A function $u\colon \bar{J} \to H$, which is differentiable almost everywhere on $\bar{J}$ such that $u' \in L_1(J;H)$ is called a strong solution of (IVP) if $u(0) = u_0$ and

$$u'(t) + Au(t) = f(t) \quad \text{for a.e. } t \in \bar{J}.$$

We note that this definition implies integrability of the strong solution u itself and of Au on J, i.e.,

$$\int_J \|u(t)\|_H + \|Au(t)\|_H \,\mathrm{d}t < +\infty,$$

and that the following hold for almost every $t \in \bar{J}$:

$$u(t) \in \mathcal{D}(A), \quad u(t) = u_0 - \int_0^t Au(s)\,\mathrm{d}s + \int_0^t f(s)\,\mathrm{d}s.$$

Since $-A$ is the generator of an analytic C_0-semigroup one knows that the initial value problem (IVP) has a unique strong solution for every $u_0 \in H$ if f is locally α-Hölder continuous on $(0,T]$ with exponent $\alpha > 0$, see [15, Ch. 4, Cor. 3.3]. However, a solution concept for (IVP) which admits a unique solution for every $f \in L_1(J;H)$ would be preferable.

In the following, we present two different approaches to widen the notion of a solution to the problem (IVP):

(i) Motivated by the fact that the unique classical solution to the homogeneous problem for $u_0 \in H$ is given by $u(t) = S(t)u_0$, one can define generalized solutions in terms of the analytic semigroup. These functions are then continuous on $\bar{J}$, but in general not differentiable and they do not necessarily take values in $\mathcal{D}(A)$ on J, see [12, 15].

(ii) In [19, 20] it has been proposed to treat the initial value problem (IVP) with variational formulations posed on Bochner and vector-valued Sobolev spaces as trial and test spaces as, e.g., the spaces introduced in Subsection 1.4. If the pair of trial–test spaces is chosen appropriately, one can show that the resulting variational problem satisfies an inf-sup condition and that existence and uniqueness of a solution in this variational sense is guaranteed.

2.1. The semigroup approach. For every $f \in L_1(J;H)$ the function $u\colon \bar{J} \to H$ defined by

$$u(t) = S(t)u_0 + \int_0^t S(t-s)f(s)\,\mathrm{d}s, \quad t \in \bar{J},$$

is continuous, i.e., $u \in C(\bar{J};H)$. Furthermore, if the initial value problem (IVP) has a classical solution, then it is given by this function. In this way, it may be considered as a generalized solution called the mild solution of (IVP).

As already mentioned, every classical solution is a mild solution. Since we assume that the semigroup $(S(t), t \geq 0)$ is analytic, it is furthermore ensured that for every $u_0 \in H$, the mild solution is also a classical solution if f is locally α-Hölder continuous on $(0, T]$ with exponent $\alpha > 0$, see [15, Ch. 4, Cor. 3.3].

2.2. The variational approach. Every variational formulation of the initial value problem (IVP) is posed on a trial and a test space. These spaces have to be balanced in such a way that the resulting solution operator is a bijection between the dual of the test space and the trial space, since then existence and uniqueness of a variational solution in the trial space is ensured.

In order to introduce the trial and test spaces of the variational problems below, we assume that A is a self-adjoint, positive definite operator with a compact inverse as in Subsections 1.3–1.4. In addition, we recall the Gelfand triple $V \hookrightarrow H \cong H^* \hookrightarrow V^*$ as well as the Hilbert spaces

$$\mathcal{X} = L_2(J; V), \quad \widehat{\mathcal{Y}} = L_2(J; V) \cap H^1(J; V^*),$$

and the closed linear subspaces $\mathcal{Y}_0, \mathcal{Y} \subset \widehat{\mathcal{Y}}$ of functions vanishing at time $t = 0$ and $t = T$, respectively.

There are different ways to derive a well-posed variational formulation of the initial value problem (IVP), but all of them have in common that the resulting solution concept is less restrictive than the ones of the classical and strong solution presented above.

For the first approach, let $f \in L_2(J; V^*)$. Furthermore, we assume that the solution u to (IVP) has a square-integrable V^*-valued distributional derivative $\partial_t u$, that u takes values in $V = \mathcal{D}(A^{1/2})$ almost everywhere in $\bar{J}$, and that u is square-integrable on J with respect to V. In other words, u is a well-defined element of the space $\widehat{\mathcal{Y}}$. If the initial value u_0 equals 0, a variational problem corresponding to (IVP) is given by [20]:

$$\text{(VP1)} \qquad \text{Find} \quad u \in \mathcal{Y}_0 \quad \text{s.t.} \quad b_*(u, v) = \ell_*(v) \quad \forall v \in \mathcal{X},$$

where the bilinear form $b_* \colon \mathcal{Y}_0 \times \mathcal{X} \to \mathbb{R}$ is defined by

$$b_*(w, v) := \int_J {}_{V^*}\langle \partial_t w(t) + A w(t), v(t) \rangle_V \, \mathrm{d}t \quad \forall (w, v) \in \mathcal{Y}_0 \times \mathcal{X},$$

and $\ell_*(v) := \int_J {}_{V^*}\langle f(t), v(t) \rangle_V \, \mathrm{d}t$ for $v \in \mathcal{X}$.

In order to cope with non-vanishing initial data $u(0) = u_0 \neq 0$, one can, e.g., choose a function $u_p \in \widehat{\mathcal{Y}}$ with $u_p(0) = u_0$ and consider the problem

$$\text{Find} \quad \widetilde{u} \in \mathcal{Y}_0 \quad \text{s.t.} \quad b_*(\widetilde{u}, v) = \widetilde{\ell}_*(v) \quad \forall v \in \mathcal{X},$$

where $\widetilde{\ell}_*(v) := \ell_*(v) - b_*(u_p, v)$ for $v \in \mathcal{X}$. The function $u := u_p + \widetilde{u} \in \widehat{\mathcal{Y}}$ solves (IVP) then in the variational sense.

A problem arising with this approach is that the function $u_p \in \mathcal{X}$ has to be chosen and, in particular, this function has to have V-regularity almost everywhere. Thus, e.g., the choice $u_p(t) := e^{-t} u_0$ is only admissible for $u_0 \in V$.

If $u_0 \in H \setminus V$ one can, e.g., enforce the initial value with a multiplier in the variational problem, see [19]. For this approach, we consider the the vector space

$$\widehat{\mathcal{X}} := \{(v, \phi) : v \in \mathcal{X},\ \phi \in H\}$$

equipped with the following algebraic operations:

$$\begin{aligned} (v,\phi) + (w,\psi) &:= (v + w, \phi + \psi) && \forall (v,\phi), (w,\psi) \in \widehat{\mathcal{X}}, \\ \lambda(v,\phi) &:= (\lambda v, \lambda\phi) && \forall (v,\phi) \in \widehat{\mathcal{X}},\ \lambda \in \mathbb{R}, \end{aligned}$$

and the graph norm:

$$\|(v,\phi)\|_{\widehat{\mathcal{X}}} := \left(\|v\|_{\mathcal{X}}^2 + \|\phi\|_H^2\right)^{1/2}.$$

We define the bilinear form $\widehat{b}\colon \widehat{\mathcal{Y}} \times \widehat{\mathcal{X}} \to \mathbb{R}$ for $w \in \widehat{\mathcal{Y}}$ and $(v,\phi) \in \widehat{\mathcal{X}}$ by

$$\widehat{b}(w,(v,\phi)) := \int_J {}_{V^*}\langle \partial_t w(t) + Aw(t), v(t)\rangle_V \,\mathrm{d}t + (w(0), \phi)_H,$$

where again $\partial_t w$ denotes the distributional derivative of w. A variational solution to (IVP) is then given by the function u satisfying

$$\text{(VP2)} \qquad \text{Find} \quad u \in \widehat{\mathcal{Y}} \quad \text{s.t.} \quad \widehat{b}(u,(v,\phi)) = \widehat{\ell}(v,\phi) \quad \forall (v,\phi) \in \widehat{\mathcal{X}},$$

where $\widehat{\ell}(v,\phi) := \int_J {}_{V^*}\langle f(t), v(t)\rangle_V \,\mathrm{d}t + (u_0,\phi)_H$ for $(v,\phi) \in \widehat{\mathcal{X}}$.

An alternative approach [20] is to choose the trial and test spaces in such a way that the variational problem incorporates the initial condition as a "natural boundary condition". For this purpose, we first note that the following integration by parts formula holds for functions $w, v \in \widehat{\mathcal{Y}}$ with V^*-valued distributional derivatives $\partial_t w$ and $\partial_t v$:

$$\begin{aligned} \int_J {}_{V^*}\langle \partial_t w(t), v(t)\rangle_V \,\mathrm{d}t = &- \int_J {}_V\langle w(t), \partial_t v(t)\rangle_{V^*} \,\mathrm{d}t \\ &+ (w(T), v(T))_H - (w(0), v(0))_H. \end{aligned}$$

After multiplying the initial value problem (IVP) with a test function $v \in \mathcal{Y}$ (i.e., $v(T) = 0$) and integrating over J, the application of the above identity yields the following variational problem:

$$\text{(VP3)} \qquad \text{Find} \quad u \in \mathcal{X} \quad \text{s.t.} \quad b(u,v) = \ell(v) \quad \forall v \in \mathcal{Y},$$

with the bilinear form $b\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ defined by

$$b(w,v) := \int_J {}_V\langle w(t), -\partial_t v(t) + Av(t)\rangle_{V^*} \,\mathrm{d}t, \quad \forall (w,v) \in \mathcal{X} \times \mathcal{Y}$$

and the right-hand side $\ell(v) := \int_J {}_{V^*}\langle f(t), v(t)\rangle_V \,\mathrm{d}t + (u_0, v(0))_H$. Note that we have included the term arising from the initial value in the functional ℓ.

In the following theorem we address well-posedness of the three presented variational problems.

Theorem 2.1. *The bilinear forms* $b_*: \mathcal{Y}_0 \times \mathcal{X} \to \mathbb{R}$, $\widehat{b}: \widehat{\mathcal{Y}} \times \widehat{\mathcal{X}} \to \mathbb{R}$, *and* $b: \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ *in* (VP1), (VP2), *and* (VP3) *are continuous and there exist constants* $\gamma_*, \widehat{\gamma}, \gamma > 0$ *such that the following inf-sup and surjectivity conditions are satisfied:*

$$\inf_{w \in S(\mathcal{Y}_0)} \sup_{v \in S(\mathcal{X})} b_*(w,v) \geq \gamma^*, \qquad \forall v \in \mathcal{X} \setminus \{0\} : \quad \sup_{w \in S(\mathcal{Y}_0)} b_*(w,v) > 0,$$

$$\inf_{w \in S(\widehat{\mathcal{Y}})} \sup_{v \in S(\widehat{\mathcal{X}})} \widehat{b}(w,v) \geq \widehat{\gamma}, \qquad \forall v \in \widehat{\mathcal{X}} \setminus \{0\} : \quad \sup_{w \in S(\widehat{\mathcal{Y}})} \widehat{b}(w,v) > 0,$$

$$\inf_{w \in S(\mathcal{X})} \sup_{v \in S(\mathcal{Y})} b(w,v) \geq \gamma, \qquad \forall v \in \mathcal{Y} \setminus \{0\} : \quad \sup_{w \in S(\mathcal{X})} b(w,v) > 0.$$

Furthermore, for any $f \in L_2(J; V^*)$, *the functionals* ℓ_* *and* $\widehat{\ell}$ *in* (VP1) *and* (VP2) *are linear and continuous on* $\mathcal{X}$ *and* $\widehat{\mathcal{X}}$, *respectively. More precisely, the following estimates hold:*

$$\|\ell_*\|_{\mathcal{X}'} \leq \|f\|_{L_2(J;V^*)}, \quad \|\widehat{\ell}\|_{\widehat{\mathcal{X}}'} \leq \|f\|_{L_2(J;V^*)} + \|u_0\|_H.$$

The functional ℓ *in* (VP3) *is continuous on* $\mathcal{Y}$ *for every* $f \in \mathcal{Y}^*$ *and* $u_0 \in H$ *with* $\|\ell\|_{\mathcal{Y}'} \leq \|f\|_{\mathcal{Y}^*} + \|u_0\|_H$, *where* $\mathcal{Y}^*$ *denotes the identification of the dual of* $\mathcal{Y}$ *via the inner product on* $L_2(J;H)$.

Proof. For the proof of the inf-sup and surjectivity conditions, see [19, Thm. 5.1] and [20, Thm. 2.2].

The bounds for $\|\ell_*\|_{\mathcal{X}'}$ and $\|\widehat{\ell}\|_{\widehat{\mathcal{X}}'}$ are readily seen. In order to derive the bound for $\|\ell\|_{\mathcal{Y}'}$, we recall that the embedding constant in $\mathcal{Y} \hookrightarrow C(\bar{J};H)$ equals 1, see Subsection 1.4. Thus, we obtain for $v \in \mathcal{Y}$:

$$|\ell(v)| \leq \|f\|_{\mathcal{Y}^*}\|v\|_{\mathcal{Y}} + \|u_0\|_H\|v(0)\|_H \leq (\|f\|_{\mathcal{Y}^*} + \|u_0\|_H)\|v\|_{\mathcal{Y}}. \qquad \square$$

We close this section by drawing some conclusions from Theorem 2.1.

- The bilinear forms b_*, $\widehat{b}$, and b induce boundedly invertible continuous linear operators $b_* \in \mathcal{L}(\mathcal{Y}_0; \mathcal{X}')$, $\widehat{b} \in \mathcal{L}(\widehat{\mathcal{Y}}; \widehat{\mathcal{X}}')$, and $b \in \mathcal{L}(\mathcal{X}; \mathcal{Y}')$, where we use the same notation for the operators as for the bilinear forms, since it will be evident from the context to which we refer.

  Therefore, the variational problems (VP1), (VP2), and (VP3) considered above are uniquely solvable. For (VP3), the data-to-solution mapping $(f, u_0) \mapsto u$, where $u \in \mathcal{X}$ denotes the solution to (VP3), satisfies the stability bound

  $$\|u\|_{\mathcal{X}} \leq \gamma^{-1}\|\ell\|_{\mathcal{Y}'} \leq \gamma^{-1}(\|f\|_{\mathcal{Y}^*} + \|u_0\|_H),$$

  and analogous results hold for (VP1) and (VP2).
- Recall the equivalent norm $|||\cdot|||_{\mathcal{Y}}$ on $\mathcal{Y}$ from Subsection 1.4:

  $$|||v|||_{\mathcal{Y}} = \Big(\|v\|^2_{L_2(J;V)} + \|\partial_t v\|^2_{L_2(J;V^*)} + \|v(0)\|^2_H\Big)^{1/2}.$$

  With this norm the induced operator $b: \mathcal{X} \to \mathcal{Y}'$ is an isometry. To show this, we first emphasize the following identity

  $$|||v|||_{\mathcal{Y}} = \|-\partial_t v + Av\|_{L_2(J;V^*)} \quad \forall v \in \mathcal{Y},$$

and, thus, $|||bw|||_{\mathcal{Y}'} := \sup_{v\in\mathcal{Y}} \frac{|b(w,v)|}{|||v|||_{\mathcal{Y}}} \leq \|w\|_{\mathcal{X}}$ for $w \in \mathcal{X}$ follows from the definition of b. It remains to verify $|||bw|||_{\mathcal{Y}'} \geq \|w\|_{\mathcal{X}}$ or, equivalently, that $\|b^{-1}\ell\|_{\mathcal{X}} \leq |||\ell|||_{\mathcal{Y}'}$ holds for all $\ell \in \mathcal{Y}'$. Since the inf-sup constant γ of b is positive and $|||\cdot|||_{\mathcal{Y}}$ defines an equivalent norm on $\mathcal{Y}$, one knows that the inf-sup constant $\gamma_{|||}$ of b with respect to $|||\cdot|||_{\mathcal{Y}}$ on $\mathcal{Y}$ inherits the positivity. Moreover, $\|b^{-1}\ell\|_{\mathcal{X}} \leq \gamma_{|||}^{-1}|||\ell|||_{\mathcal{Y}'}$ for $\ell \in \mathcal{Y}'$ and the following identity holds

$$\gamma_{|||} = \inf_{w\in\mathcal{X}} \sup_{v\in\mathcal{Y}} \frac{b(w,v)}{\|w\|_{\mathcal{X}}|||v|||_{\mathcal{Y}}} = \inf_{v\in\mathcal{Y}} \sup_{w\in\mathcal{X}} \frac{b(w,v)}{\|w\|_{\mathcal{X}}|||v|||_{\mathcal{Y}}}.$$

For $v \in \mathcal{Y}$ and $w := v - (A^*)^{-1}\partial_t v$ we obtain

$$\begin{aligned} b(w,v) &= \|v\|_{L_2(J;V)}^2 + \|\partial_t v\|_{L_2(J;V^*)}^2 - 2\int_J {}_{V^*}\langle \partial_t v(t), v(t)\rangle_V \,\mathrm{d}t \\ &= |||v|||_{\mathcal{Y}}^2 = \|w\|_{\mathcal{X}}^2 = \|w\|_{\mathcal{X}}|||v|||_{\mathcal{Y}}. \end{aligned}$$

This shows that $\gamma_{|||} \geq 1$ and the isometry property of $b\colon \mathcal{X} \to \mathcal{Y}'$ with respect to the norms $\|\cdot\|_{\mathcal{X}}$ and $|||\cdot|||_{\mathcal{Y}'}$ follows.

Note also that this result in conjunction with the observation that $|||v|||_{\mathcal{Y}} \geq \|v\|_{\mathcal{Y}}$ for all $v \in \mathcal{Y}$ implies the following lower bound for the inf-sup constant of the bilinear form b with respect to the norms $\|\cdot\|_{\mathcal{X}}$ and $\|\cdot\|_{\mathcal{Y}}$:

$$\inf_{w\in S(\mathcal{X})} \sup_{v\in S(\mathcal{Y})} b(w,v) = \inf_{w\in\mathcal{X}} \sup_{v\in\mathcal{Y}} \frac{b(w,v)}{\|w\|_{\mathcal{X}}\|v\|_{\mathcal{Y}}} \geq 1.$$

and, therefore, $\gamma \geq 1$ holds.

- The latter variational approach (VP3) does not pose more assumptions on the source term f and the initial value u_0 in (IVP) than the semigroup approach does, since $\ell \in \mathcal{Y}'$ for every $f \in L_1(J;H)$ and $u_0 \in H$.

3. Stochastic differential equations

In order to introduce the stochastic differential equations of interest, we recall certain notions and concepts from probability theory in Subsection 3.1 first. In addition, we summarize basic definitions and results from Itô integration and Itô calculus in Subsection 3.2. This establishes the framework for defining solutions to stochastic differential equations in Subsection 3.3.

From here on, let $(\Omega, \mathcal{A}, \mathbb{P})$ denote a complete probability space equipped with the filtration $\mathcal{F} := (\mathcal{F}_t, t \in I)$ which satisfies the "ususal conditions", i.e.,

(i) $\mathcal{F}$ is right continuous, i.e., $\mathcal{F}_t = \mathcal{F}_{t^+} := \bigcap_{s>t} \mathcal{F}_s$ for all $t \in I$;
(ii) $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets of $\mathcal{A}$.

For our purposes, the index set I is either the nonnegative part of the real line $I := \{t \in \mathbb{R} : t \geq 0\}$ or the closed finite time interval $I := \bar{J} = [0,T]$.

Throughout this section, we write $s \wedge t := \min\{s,t\}$ for $s,t \in \mathbb{R}$ and we mark equalities which hold $\mathbb{P}$-almost surely with $\mathbb{P}$-a.s.

3.1. Stochastic processes, moments, and covariance functions. In this subsection we present the classes of stochastic processes as well as their characteristics which are of interest for our investigations. We refer the reader, e.g., to [16, §§3, 4] for a more detailed introduction to stochastic processes with values in Hilbert spaces and, in particular, Lévy processes.

Any measurable mapping $Z \colon \Omega \to H$ is called an H-valued random variable and an H-valued stochastic process is defined as a family $(X(t), t \in I)$ of H-valued random variables. For a stochastic process $X := (X(t), t \in I)$ taking values in H one can define the following characteristics:

(i) *integrability*: X is said to be integrable if

$$\|X(t)\|_{L_1(\Omega;H)} := \mathbb{E}[\|X(t)\|_H] < +\infty \quad \forall t \in I,$$

and square-integrable if

$$\|X(t)\|_{L_2(\Omega;H)} := \mathbb{E}[\|X(t)\|_H^2]^{1/2} < +\infty \quad \forall t \in I;$$

(ii) *mean*: if X is integrable, the H-valued mapping

$$m \colon I \to H, \qquad t \mapsto \mathbb{E}[X(t)]$$

is well-defined and it is called the mean or first moment of X;

(iii) *second moment and covariance*: assuming that X is square-integrable, the tensor-space-valued functions $M, C \colon I \times I \to H_2$ defined by

$$\begin{aligned} M(s,t) &:= \mathbb{E}[X(s) \otimes X(t)], \\ C(s,t) &:= \mathbb{E}[(X(s) - \mathbb{E}[X(s)]) \otimes (X(t) - \mathbb{E}[X(t)])], \quad s,t \in I, \end{aligned}$$

are called the second moment and the covariance of X, respectively.

We note that the covariance can be expressed in terms of the second moment and the mean:

$$C(s,t) = M(s,t) - m(s) \otimes m(t).$$

Furthermore, the second moment and the covariance are well-defined mappings to the Hilbert tensor product space H_2 due to the observation

$$\begin{aligned} \|\mathbb{E}[X(s) \otimes X(t)]\|_{H_2} &\le \mathbb{E}[\|X(s) \otimes X(t)\|_{H_2}] = \mathbb{E}[\|X(s)\|_H \|X(t)\|_H] \\ &\le \|X(s)\|_{L_2(\Omega;H)} \|X(t)\|_{L_2(\Omega;H)}, \end{aligned}$$

where the first estimate holds by the properties of the expectation operator for H-valued random variables and the last one is Hölder's inequality.

In the following we present certain classes of stochastic processes which we are going to refer to in the course of the thesis.

3.1.1. *Martingales.* An integrable stochastic process $(X(t), t \in I)$ taking values in H is called an H-valued martingale with respect to $\mathcal{F}$ if it is $\mathcal{F}$-adapted, i.e., $X(t)$ is $\mathcal{F}_t$-measurable for all $t \in I$, and it satisfies the martingale property: the conditional expectation of $X(t)$ with respect to the σ-field $\mathcal{F}_s$ for $s \le t$ is given by $\mathbb{E}[X(t)|\mathcal{F}_s] = X(s)$.

3.1.2. *Lévy processes.* An H-valued stochastic process $L := (L(t), t \in I)$ is said to be a Lévy process if the following conditions are satisfied:

(i) L has independent increments, i.e., for all $0 \le t_0 < t_1 < \ldots < t_n$, $t_1, \ldots, t_n \in I$, the H-valued random variables $L(t_1) - L(t_0)$, $L(t_2) - L(t_1)$, …, $L(t_n) - L(t_{n-1})$ are independent;
(ii) L has stationary increments, i.e., the distribution of $L(t) - L(s)$, $s \le t$, $s, t \in I$, depends only on the difference $t - s$;
(iii) $L(0) = 0$ $\mathbb{P}$-almost surely;
(iv) L is stochastically continuous, i.e.,

$$\lim_{\substack{s \to t \\ s \in I}} \mathbb{P}(\|L(t) - L(s)\|_H > \epsilon) = 0 \quad \forall \epsilon > 0, \quad \forall t \in I.$$

We often assume that a Lévy process L satisfies some or all of the following conditions:

(a) L is adapted with respect to the filtration $\mathcal{F}$;
(b) for $t > s$ the increment $L(t) - L(s)$ is independent of $\mathcal{F}_s$;
(c) L is integrable;
(d) L has mean zero, i.e., $\mathbb{E}[L(t)] = 0$ for all $t \in I$;
(e) L is square-integrable.

Note that the Lévy process L satisfies Assumptions (a)–(b), e.g., for the filtration $\mathcal{F} := (\bar{\mathcal{F}}^L_t, t \in I)$, where $\bar{\mathcal{F}}^L_t$ denotes the smallest σ-field containing the σ-field $\mathcal{F}^L_t := \sigma(L(s) : s \le t)$ generated by L and all $\mathbb{P}$-null sets of $\mathcal{A}$, [16, Remark 4.43]. Under Assumptions (a)–(d) L is a martingale, see [16, Prop. 3.25]. If, in addition, L satisfies Assumption (e), then there exists a self-adjoint, nonnegative trace-class operator $Q \in \mathcal{L}_1^+(H)$ such that for all $s, t \in I$

$$\mathbb{E}\left[(L(s) \otimes L(t), \phi \otimes \psi)_2\right] = (s \wedge t)\,(Q\phi, \psi)_H \quad \forall \phi, \psi \in H,$$

cf. [16, Thm. 4.44]. This operator is also referred to as the covariance operator of the Lévy process L.

In the following we illustrate how covariance functions relate to the concept of tensor product spaces from Subsection 1.2. Suppose that $I = \bar{J} = [0, T]$ and that L is an H-valued Lévy process satisfying the assumptions above. Since L has mean zero, the second moment and the covariance C of L coincide and they satisfy

$$\begin{aligned} \int_{J \times J} (C(s,t), \phi(s) \otimes \psi(t))_2 \,\mathrm{d}s\,\mathrm{d}t &= \int_{J \times J} (s \wedge t)(Q\phi(s), \psi(t))_H \,\mathrm{d}s\,\mathrm{d}t \\ &= \int_{J \times J} \Big(\sum_{i \in \mathbb{N}} (s \wedge t)\gamma_i (e_i \otimes e_i), \phi(s) \otimes \psi(t) \Big)_2 \,\mathrm{d}s\,\mathrm{d}t \end{aligned}$$

for all $\phi, \psi \in \mathcal{W} = L_2(J; H)$, where $\{e_i\}_{i \in \mathbb{N}}$ is an orthonormal basis of H consisting of eigenvectors of Q with corresponding nonnegative eigenvalues $\{\gamma_i\}_{i \in \mathbb{N}}$. Therefore, the covariance of L can be represented as

$$C(s,t) = \sum_{i \in \mathbb{N}} (s \wedge t)\gamma_i (e_i \otimes e_i), \quad s, t \in \bar{J},$$

with convergence of this series in $\mathcal{W}_2$, since $\sum_{i\in\mathbb{N}} \gamma_i < +\infty$. Furthermore, C is even an element of $\mathcal{W}_\pi$ with $\|C\|_\pi \le \frac{1}{2}T^2 \operatorname{tr} Q$.

In order to derive the latter bound, we define the vector $q := \sum_{i\in\mathbb{N}} \gamma_i(e_i\otimes e_i)$ and the real-valued function $w_\wedge(s,t) := s \wedge t$. Then $w_\wedge$ is an element of the projective tensor space $L_2(J;\mathbb{R})_\pi$ with

$$\|w_\wedge\|_\pi = \delta(w_\wedge) = \tfrac{1}{2}T^2,$$

where we used that $w_\wedge$ is a positive semi-definite kernel as well as the characterization of the projective norm derived in Subsection 1.5. In addition, the vector q is an element of H_π, since $\|q\|_\pi = \operatorname{tr}(Q) < +\infty$. Thus, the covariance of L satisfies

$$C = w_\wedge \otimes q \in L_2(J;\mathbb{R})_\pi \otimes H_\pi, \qquad \|C\|_{L_2(J;\mathbb{R})_\pi \hat{\otimes}_\pi H_\pi} = \tfrac{1}{2}T^2 \operatorname{tr}(Q),$$

and the same bound holds for the projective norm on $\mathcal{W}_\pi$ due to

$$L_2(J;\mathbb{R})_\pi \,\hat{\otimes}_\pi\, H_\pi \cong (L_2(J;\mathbb{R}) \,\hat{\otimes}_\pi\, H)_\pi \hookrightarrow (L_2(J;\mathbb{R}) \,\hat{\otimes}_2\, H)_\pi \cong \mathcal{W}_\pi,$$

where the first identification is due to the associativity of the projective tensor product [9, Ch. 33]. The embedding holds, since the projective tensor product inherits the continuous embedding from $L_2(J;\mathbb{R}) \,\hat{\otimes}_\pi\, H \hookrightarrow L_2(J;\mathbb{R}) \,\hat{\otimes}_2\, H$ and the last identification is a consequence of the definitions of the Hilbert tensor product space $L_2(J;\mathbb{R}) \,\hat{\otimes}_2\, H$ and the Bochner space $\mathcal{W} = L_2(J;H)$.

3.1.3. *Wiener processes.* An important subclass of Lévy processes is formed by Wiener processes. Here, a Lévy process $W := (W(t), t \in I)$ is said to be an H-valued Wiener process if it has continuous trajectories in H and mean zero. In the finite-dimensional case $H = \mathbb{R}^n$ the covariance function is given by

$$\mathbb{E}\left[W(s)W(t)^{\mathsf{T}}\right] = (s \wedge t)Q \quad \forall s,t \in I$$

for a symmetric positive semi-definite matrix $Q \in \mathbb{R}^{n\times n}$. If Q equals the identity matrix in $\mathbb{R}^{n\times n}$ then W is called Wiener white noise. In one dimension ($H = \mathbb{R}$) W is called a real-valued Brownian motion if $Q = 1$.

3.2. Stochastic integration. The purpose of this section is to make sense of the stochastic integral

$$\int_0^t \Psi(s)\,\mathrm{d}L(s), \quad t \in \bar{J} = [0,T],$$

where the noise $L = (L(t), t \in \bar{J})$ is a Lévy process taking values in a separable Hilbert space U and Ψ is a stochastic process taking values in an appropriate space of operators mapping from U to the separable Hilbert space H, so that for fixed $t \in \bar{J}$ the stochastic integral itself becomes an H-valued random variable.

To this end, following the lines of [16, §8.2], we first define the stochastic Itô integral for $\mathcal{L}(U;H)$-valued processes Ψ on $\bar{J}$, which are simple, i.e., there exist finite sequences

- of nonnegative numbers $0 = t_0 < t_1 < \ldots < t_m \le T$,
- of operators $\Psi_1, \ldots, \Psi_m \in \mathcal{L}(U;H)$, and
- of events $A_j \in \mathcal{F}_{t_{j-1}}$, $1 \le j \le m$,

such that

$$\Psi(s) = \sum_{j=1}^{m} \mathbb{1}_{A_j} \, \mathbb{1}_{(t_{j-1},t_j]}(s) \, \Psi_j, \quad s \in \bar{J},$$

where $\mathbb{1}_A$ and $\mathbb{1}_{(s,t]}$ denote the indicator functions of the event $A \in \mathcal{A}$ and the interval $(s,t] \subset \bar{J}$, respectively.

Let the U-valued Lévy process L satisfy Assumptions (a)–(e) posed in Subsection 3.1.2. Then the stochastic integral with respect to the simple process Ψ and the Lévy noise L is defined by

$$\int_0^t \Psi(s) \, \mathrm{d}L(s) := \sum_{j=1}^{m} \mathbb{1}_{A_j} \Psi_j (L(t_j \wedge t) - L(t_{j-1} \wedge t)) \quad \forall t \in \bar{J}.$$

The so-constructed stochastic integral is called Itô integral and it satisfies the following important property—the Itô isometry, see [16, Prop. 8.6]:

$$\mathbb{E}\left[\left\| \int_0^t \Psi(s) \, \mathrm{d}L(s) \right\|_H^2 \right] = \mathbb{E}\left[\int_0^t \|\Psi(s) Q^{1/2}\|^2_{\mathcal{L}_2(U;H)} \, \mathrm{d}s \right] \quad \forall t \in \bar{J}.$$

Here, the operator $Q \in \mathcal{L}_1^+(U)$ is the self-adjoint, nonnegative covariance operator of the Lévy process L. Its square root $Q^{1/2}$ and the pseudo inverse $Q^{-1/2}$ can be defined via the spectral expansions

$$Q^{\pm 1/2} x := \sum_{i \in \mathcal{I}} \gamma_i^{\pm 1/2} (x, e_i)_U \, e_i \quad \forall x \in U$$

with respect to an orthonormal basis $\{e_i\}_{i \in \mathbb{N}} \subset U$ consisting of eigenvectors of Q with corresponding nonnegative eigenvalues $\{\gamma_i\}_{i \in \mathbb{N}}$, where the index set is given by $\mathcal{I} := \{i \in \mathbb{N} : \gamma_i \neq 0\}$.

The space $\mathcal{H} := Q^{1/2} U$ equipped with the inner product

$$(\cdot, \cdot)_{\mathcal{H}} := (Q^{-1/2} \cdot, Q^{-1/2} \cdot)_U$$

is a Hilbert space called the reproducing kernel Hilbert space of the Lévy process L, [16, Def. 7.2]. It is common to formulate the Itô isometry in terms of the Hilbert–Schmidt norm with $\mathcal{H}$ instead of U as inverse image space:

$$\mathbb{E}\left[\left\| \int_0^t \Psi(s) \, \mathrm{d}L(s) \right\|_H^2 \right] = \mathbb{E}\left[\int_0^t \|\Psi(s)\|^2_{\mathcal{L}_2(\mathcal{H};H)} \, \mathrm{d}s \right] \quad \forall t \in \bar{J}.$$

In order to extend the space of admissible integrands, we take the closure of the vector space of all $\mathcal{L}(U;H)$-valued simple processes with respect to the following norm:

$$\|\Psi\|^2_{L^2_{\mathcal{H},T}(H)} := \mathbb{E}\left[\int_0^T \|\Psi(s)\|^2_{\mathcal{L}_2(\mathcal{H};H)} \, \mathrm{d}s \right].$$

The resulting Banach space denoted by $L^2_{\mathcal{H},T}(H)$ is in fact a Hilbert space, namely the space of predictable processes Ψ taking values in the space of

Hilbert–Schmidt operators $\mathcal{L}_2(\mathcal{H};H)$ such that the $L^2_{\mathcal{H},T}(H)$-norm defined above is finite, i.e.,

$$\begin{aligned} L^2_{\mathcal{H},T}(H) &:= \{\Psi\colon \Omega\times\bar{J}\to\mathcal{L}_2(\mathcal{H};H) : \Psi \text{ predictable}, \|\Psi\|_{L^2_{\mathcal{H},T}(H)} < +\infty\} \\ &= L_2(\Omega\times\bar{J},\mathcal{P}_{\bar{J}},\mathbb{P}\,\mathrm{d}t;\mathcal{L}_2(\mathcal{H};H)), \end{aligned}$$

where $\mathcal{P}_{\bar{J}}$ denotes the σ-field of all predictable sets in $\Omega\times\bar{J}$, that is the smallest σ-field of subsets of $\Omega\times\bar{J}$ containing all sets of the form $A\times(s,t]$, where $0\le s<t\le T$ and $A\in\mathcal{F}_s$, see [16, Thm. 8.7, Cor. 8.17]. Recall that the process Ψ is called predictable if Ψ is measurable with respect to $\mathcal{P}_{\bar{J}}$.

The construction described above yields a well-defined stochastic integral $\int_0^t \Psi(s)\,\mathrm{d}L(s)\in L_2(\Omega;H)$ for integrands $\Psi\in L^2_{\mathcal{H},T}(H)$ for all $t\in\bar{J}$. Moreover, it is the largest class of integrands satisfying the Itô isometry.

In Paper I we furthermore need the notion of the weak stochastic integral. In order to introduce it, let Ψ be a stochastic process in the space of admissible integrands $L^2_{\mathcal{H},T}(H)$ and $v\in C(\bar{J};H)$ be a deterministic continuous H-valued function. We then define the $\mathcal{L}_2(\mathcal{H};\mathbb{R})$-valued stochastic process Ψ_v for $t\in\bar{J}$ by

$$\Psi_v(t) : z\mapsto (v(t),\Psi(t)z)_H \quad \forall z\in\mathcal{H}.$$

The predictability of Ψ and the continuity of v imply that Ψ_v is predictable. Moreover, it can be derived from the finiteness of $\sup_{t\in\bar{J}}\|v(t)\|_H<+\infty$ and $\|\Psi\|_{L^2_{\mathcal{H},T}(H)}<+\infty$ that Ψ_v is an element in the space of admissible integrands $L^2_{\mathcal{H},T}(\mathbb{R})$. Therefore, we can define the real-valued weak stochastic stochastic integral $\int_0^t (v(s),\Psi(s)\,\mathrm{d}L(s))_H$ as the stochastic integral with respect to the integrand Ψ_v, i.e.,

$$\int_0^t (v(s),\Psi(s)\,\mathrm{d}L(s))_H := \int_0^t \Psi_v(s)\,\mathrm{d}L(s) \quad \forall t\in\bar{J} \quad \mathbb{P}\text{-a.s.}$$

We note that the Itô isometry for the original stochastic integral implies the following isometry for the weak stochastic integral:

$$\mathbb{E}\left[\left|\int_0^t (v(s),\Psi(s)\,\mathrm{d}L(s))_H\right|^2\right] = \mathbb{E}\left[\int_0^t \|\Psi_v(s)\|^2_{\mathcal{L}_2(\mathcal{H};\mathbb{R})}\,\mathrm{d}s\right] \quad \forall t\in\bar{J}.$$

This weak version of the Itô isometry and some of its implications are important for our analysis in Paper I.

3.3. Strong and mild solutions. We consider equations of the form

$$\text{(SDE)} \qquad \mathrm{d}X(t)+AX(t)\mathrm{d}t = G(X(t))\,\mathrm{d}L(t), \quad t\in\bar{J}, \qquad X(0)=X_0$$

in the Hilbert space H, where

- $A\colon\mathcal{D}(A)\to H$ is a self-adjoint, positive definite, possibly unbounded operator with a compact inverse as in Subsection 1.3;
- $G\colon H\to\mathcal{L}_2(\mathcal{H};H)$ is an affine operator, i.e., there exist operators

$$G_1\in\mathcal{L}(H;\mathcal{L}_2(\mathcal{H};H)) \quad\text{and}\quad G_2\in\mathcal{L}_2(\mathcal{H};H)$$

such that

$$G(\varphi) = G_1(\varphi) + G_2 \quad \forall \varphi \in H;$$

- L is a U-valued Lévy process satisfying Assumptions (a)–(e) posed in Subsection 3.1.2 and $\mathcal{H} = Q^{1/2}U$ denotes its reproducing kernel Hilbert space;
- X_0 is an $\mathcal{F}_0$-measurable, square-integrable H-valued random variable.

Equation (SDE) is said to be a stochastic differential equation, SDE for short. More precisely, it is called a stochastic *ordinary* differential equation (SODE) if $H = \mathbb{R}^n$ and $U = \mathbb{R}^m$ for finite dimensions $m, n \in \mathbb{N}$ and a stochastic *partial* differential equation (SPDE) if H is an infinite dimensional function space and A a differential operator. For vanishing G_1, (SDE) is said to have additive noise, otherwise it is called an SDE with multiplicative noise.

The purpose of this section is to make sense of the notion of solutions to these kinds of equations. In fact, as for the deterministic initial value problem (IVP) there are also different solution concepts for (SDE). In the following we present their definitions and how they relate. For an introduction to stochastic ordinary differential equations and stochastic partial differential equations the reader is referred to [10, 14] and to [2, 16], respectively.

An H-valued predictable process $X = (X(t), t \in \bar{J})$ taking values in $\mathcal{D}(A)$ $\mathcal{P}_{\bar{J}}$-a.s. is called a strong solution to (SDE) if

$$\int_J \|X(s)\|_H + \|AX(s)\|_H + \|G(X(s))\|^2_{\mathcal{L}_2(\mathcal{H};H)} \,\mathrm{d}s < +\infty \quad \mathbb{P}\text{-a.s.}$$

and the following integral equation holds for all $t \in \bar{J}$:

$$X(t) = X_0 - \int_0^t AX(s)\,\mathrm{d}s + \int_0^t G(X(s))\,\mathrm{d}L(s) \quad \mathbb{P}\text{-a.s.}$$

We emphasize the close relation in the definitions of a strong solution to the deterministic initial value problem (IVP) on the one hand, and of a strong solution to the stochastic differential equation (SDE) on the other hand.

Since we assume that the operators A and G are linear and affine, respectively, in the case of an SODE with $H = \mathbb{R}^n$ and $U = \mathbb{R}^m$ we have

$$A \in \mathbb{R}^{n\times n}, \quad G_1 \in \mathcal{L}(\mathbb{R}^n; \mathbb{R}^{n\times m}), \quad G_2 \in \mathbb{R}^{n\times m}.$$

It is well-known [14, Thm. 5.2.1] that this SODE when driven by $\mathbb{R}^m$-valued Wiener white noise (cf. Subsection 3.1.3) admits a strong solution. Moreover, this solution is unique up to modification, i.e., if X_1 and X_2 are two strong solutions, then

$$\mathbb{P}(X_1(t) = X_2(t)) = 1 \quad \forall t \in \bar{J}.$$

Under the additional assumption that the mapping $t \mapsto X(t)$ is continuous $\mathbb{P}$-a.s. (t-continuity), the solution is pathwise unique in the sense of [10], i.e.,

$$\mathbb{P}(X_1(t) = X_2(t) \quad \forall t \in \bar{J}) = 1$$

for any two t-continuous strong solutions X_1 and X_2.

As a further illustration of the concept of strong solutions we take as explicit examples the real-valued ($m = n = 1$) model SODEs considered in Paper II: with additive noise

$$\text{(SODE1)} \qquad \mathrm{d}X(t) + \lambda X(t)\,\mathrm{d}t = \mu\,\mathrm{d}W(t), \quad t \in \bar{J}, \quad X(0) = X_0,$$

and with multiplicative noise

$$\text{(SODE2)} \quad \mathrm{d}X(t) + \lambda X(t)\,\mathrm{d}t = \rho X(t)\,\mathrm{d}W(t), \quad t \in \bar{J}, \quad X(0) = X_0$$

for an initial value $X_0 \in L_2(\Omega; \mathbb{R})$ and constant parameters $\lambda, \mu, \rho > 0$.

As stated above, there exist strong solutions to these SODEs. Indeed, for additive noise the so-called Ornstein–Uhlenbeck process defined by

$$X(t) := e^{-\lambda t} X_0 + \mu \int_0^t e^{-\lambda(t-s)}\,\mathrm{d}W(s), \quad t \in \bar{J},$$

and in the multiplicative case the geometric Brownian motion given by

$$X(t) := X_0 e^{-(\lambda + \rho^2/2)t + \rho W(t)}, \quad t \in \bar{J},$$

satisfy the conditions of being strong solutions to the model SODEs above. We note that in both cases the integral equation can be verified by an application of the Itô formula [14, Thm. 4.1.2]: to the process $(e^{\lambda t} X(t), t \in \bar{J})$ in the additive case and to the geometric Brownian motion X itself in the multiplicative case. Moreover, these solutions are the unique t-continuous strong solutions.

Due to the availability of existence and uniqueness results for strong solutions to SODEs of the kind above and more generally with global Lipschitz coefficients (see [10, 14] for Wiener noise and [1] for Lévy noise) this definition is usually sufficient in the finite-dimensional case when $\dim(H) = n < +\infty$ and $\dim(U) = m < +\infty$.

However, as for deterministic initial value problems, it is often unsatisfactory when considering equations in infinite dimensions, since – depending on the operator A – the condition "X takes values in $\mathcal{D}(A)$ $\mathbb{P}$-a.s." may be very restrictive. Recall that for the deterministic problem (IVP) existence of a strong solution is only ensured, if the source term f is Hölder continuous. In the terminology of the deterministic framework, the noise term generated by the Lévy process L in (SDE) takes the role of the source term. Since a Lévy process is in general not pathwise differentiable (e.g., in the case when L is a Wiener process), it is usually irregular with respect to t. For this reason, strong solutions rarely exist and a less restrictive solution concept is needed.

As in the deterministic case, the semigroup $(S(t), t \geq 0)$ generated by $-A$ can be used to define mild solutions of (SDE), see [16, Def. 9.5].

Let $X = (X(t), t \in \bar{J})$ be an H-valued predictable process with

$$\sup_{t \in \bar{J}} \|X(t)\|_{L^2(\Omega;H)} < +\infty.$$

Then X is said to be a mild solution to (SDE) if

$$X(t) = S(t)X_0 + \int_0^t S(t-s)G(X(s))\,\mathrm{d}L(s), \quad \forall t \in \bar{J}, \quad \mathbb{P}\text{-a.s.}$$

In contrast to strong solutions, one knows [16, Theorems 9.15, 9.29] that under the assumptions on A and G made above, there exists a mild solution to (SDE), which is unique up to modification. Moreover, the mild solution has a càdlàg modification, which is pathwise unique. Whenever a strong solution to (SDE) exists, it coincides with the mild solution.

4. Summary of Paper I

In Paper I we pursue the study of [11], where the second moment and the covariance of the mild solution to a parabolic stochastic partial differential equation driven by additive Wiener noise have been described as solutions to well-posed space-time variational problems posed on Hilbert tensor products of Bochner spaces. More precisely, in [11] parabolic stochastic partial differential equations of the form (SDE) are considered, where the noise term is driven by an H-valued Wiener process, satisfying Assumptions (a)–(e) in Subsection 3.1.2 and the operator G equals the identity on H, i.e., $G_1 = 0$, $G_2 = \mathrm{Id}$. In this way, the state space of the Wiener noise and of the mild solution X to (SDE) coincide (both are H).

With the notation and definitions of Hilbert tensor product spaces as well as the vector-valued function spaces $\mathcal{X}$, $\mathcal{Y}$ and the bilinear form $b\colon \mathcal{X}\times\mathcal{Y}\to\mathbb{R}$ from Subsections 1.2, 1.4, and 2.2, the mean m of the mild solution X satisfies the following deterministic variational problem:

$$\text{(VPm)} \qquad \text{Find} \quad m\in\mathcal{X} \quad \text{s.t.} \quad b(m,v) = (\mathbb{E}X_0, v(0))_H \quad \forall v\in\mathcal{Y}.$$

Well-posedness of this variational problem was already observed in the analysis of (VP3) in Subsection 2.2. In [11] the tensorized bilinear form $B\colon \mathcal{X}_2\times\mathcal{Y}_2\to\mathbb{R}$ is introduced as $B := b\otimes b$, or explicitly as

$$B(w,v) := \int_J\int_J {}_{V_2}\langle w(s,t), (-\partial_s + A)\otimes(-\partial_t + A)v(s,t)\rangle_{V_2^*}\,\mathrm{d}s\,\mathrm{d}t$$

and it is proven that the second moment M of the mild solution X satisfies the following deterministic variational problem:

$$(\text{VPM}_+) \qquad \text{Find} \quad M\in\mathcal{X}_2 \quad \text{s.t.} \quad B(M,v) = \ell(v) \quad \forall v\in\mathcal{Y}_2,$$

where $\ell(v) := (\mathbb{E}[X_0\otimes X_0], v(0,0))_2 + \delta_q(v)$ and the functional $\delta_q\colon\mathcal{Y}_2\to\mathbb{R}$ is defined by (recall that $(\cdot,\cdot)_2$ abbreviates the inner product on $H_2 = H\hat{\otimes}H$)

$$\delta_q(v) := \int_J (q, v(t,t))_2\,\mathrm{d}t \quad \forall v\in\mathcal{Y}_2.$$

It is shown that under the assumption that $\mathrm{tr}(AQ) < +\infty$, the functional δ_q and, thus, the right-hand side of the variational problem for M is an element of the dual $(\mathcal{Y}_2)'$. As remarked in Subsection 2.2, the operator $b\in\mathcal{L}(\mathcal{X};\mathcal{Y}')$ is an isomorphism, so that $B = b\otimes b\in\mathcal{L}(\mathcal{X}_2;\mathcal{Y}_2')$ inherits this property and the variational problem for the second moment is well-posed.

In Paper I we prove that also in the case of multiplicative Lévy noise the second moment as well as the covariance of the square-integrable mild solution satisfy deterministic space-time variational problems posed on tensor products of Bochner spaces. In contrast to the case of additive Wiener noise considered in [11], the pair of trial–test spaces is not given by Hilbert tensor product spaces, but projective–injective tensor product spaces. In addition, the resulting bilinear form in the variational problem involves a non-separable form on these tensor spaces. Therefore, well-posedness does not readily follow from the isomorphism property of b and a careful analysis is needed to derive existence and uniqueness of a solution to the derived variational problem.

To specify this non-separable form, we define the bilinear form Δ, referred to as the trace product, on the algebraic tensor product spaces $\mathcal{W} \otimes \mathcal{W}$ and $\mathcal{Y} \otimes \mathcal{Y}$ by

$$\Delta(w,v) := \sum_{k=1}^{N} \sum_{\ell=1}^{M} \int_J (w_k^1(t), v_\ell^1(t))_H (w_k^2(t), v_\ell^2(t))_H \, \mathrm{d}t,$$

where $\sum_{k=1}^{N} w_k^1 \otimes w_k^2$ and $\sum_{\ell=1}^{M} v_\ell^1 \otimes v_\ell^2$ are any representations of $w \in \mathcal{W} \otimes \mathcal{W}$ and $v \in \mathcal{Y} \otimes \mathcal{Y}$. We prove that the trace product Δ admits a unique continuous extension by bilinearity to a bounded bilinear form $\Delta\colon \mathcal{W}_\pi \times \mathcal{Y}_\epsilon \to \mathbb{R}$ and the induced operator $\Delta \in \mathcal{L}(\mathcal{W}_\pi; \mathcal{Y}_\epsilon')$ satisfies $\|\Delta\|_{\mathcal{L}(\mathcal{W}_\pi;\mathcal{Y}_\epsilon')} \leq 1$, where $\mathcal{Y}_\epsilon'$ denotes the dual of the injective tensor product space $\mathcal{Y}_\epsilon$. Having defined the trace product Δ, we introduce

$$\mathcal{B}\colon \mathcal{X}_\pi \times \mathcal{Y}_\epsilon \to \mathbb{R}, \quad \mathcal{B}(w,v) := B(w,v) - \Delta((G_1 \otimes G_1)(w)q, v),$$

where $q := \sum_{i\in\mathbb{N}} \gamma_i (e_i \otimes e_i)$ with the nonnegative eigenvalues $\{\gamma_i\}_{i\in\mathbb{N}}$ and the corresponding orthonormal eigenvectors $\{e_i\}_{i\in\mathbb{N}}$ of the covariance operator Q of the Lévy noise L in (SDE). Continuity of this bilinear form is proven. We show that the second moment of (SDE) with multiplicative Lévy noise satisfies the following deterministic variational problem:

$$(\mathrm{VPM}_*) \qquad \text{Find} \quad M \in \mathcal{X}_\pi \quad \text{s.t.} \quad \mathcal{B}(M,v) = f(v) \quad \forall v \in \mathcal{Y}_\epsilon$$

with the right-hand side (recall $m \in \mathcal{X}$ denotes the mean of X):

$$\begin{aligned} f(v) := (\mathbb{E}[X_0 \otimes X_0], v(0,0))_2 &+ \Delta((G_1(m) \otimes G_2)q, v) \\ &+ \Delta((G_2 \otimes G_1(m))q, v) + \Delta((G_2 \otimes G_2)q, v). \end{aligned}$$

Well-posedness of this problem is proven under an appropriate assumption on the operator G_1. More precisely, the lower bound $\gamma \geq 1$ for the inf-sup constant of the bilinear form $b\colon (\mathcal{X}, \|\cdot\|_{\mathcal{X}}) \times (\mathcal{Y}, \|\cdot\|_{\mathcal{Y}}) \to \mathbb{R}$ discussed in Subsection 2.2 along with the observation that $\mathcal{Y}_\epsilon' = (\mathcal{Y}')_\pi$, see [18, Theorems 2.5 and 5.13], implies that the bilinear form $B\colon \mathcal{X}_\pi \times \mathcal{Y}_\epsilon \to \mathbb{R}$ satisfies the following inf-sup condition

$$\inf_{w\in S(\mathcal{X}_\pi)} \sup_{v\in S(\mathcal{Y}_\epsilon)} B(w,v) \geq 1.$$

Owing to $\|\Delta\|_{\mathcal{L}(\mathcal{W}_\pi;\mathcal{Y}_\epsilon')} \leq 1$ and the definition of q we find

$$\|\mathcal{B}\|_{\mathcal{L}(\mathcal{X}_\pi;\mathcal{Y}_\epsilon')} \leq 1 - \|G_1\|^2_{\mathcal{L}(V;\mathcal{L}_2(\mathcal{H};H))}$$

for the operator $\mathcal{B}\colon \mathcal{X}_\pi \to \mathcal{Y}_\epsilon'$ induced by the bilinear form $\mathcal{B}$ in (VPM_*) and injectivity holds for

$$\|G_1\|_{\mathcal{L}(V;\mathcal{L}_2(\mathcal{H};H))} < 1. \tag{G1}$$

A variational equation for the covariance function of the mild solution X to (SDE), again posed on $\mathcal{X}_\pi$ and $\mathcal{Y}_\epsilon$ as trial–test spaces, follows from the results for the mean and the second moment.

5. Petrov–Galerkin approximations

In this section we collect results for Petrov–Galerkin discretizations of the generic linear variational problem

$$\text{Find} \quad u \in \mathcal{U}: \quad \mathscr{B}(u,v) = \ell(v) \quad \forall v \in \mathcal{V},$$

posed on normed vector spaces $(\mathcal{U}, \|\cdot\|_{\mathcal{U}})$ and $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$, with a continuous linear right-hand side $\ell \in \mathcal{V}'$, and a bilinear form $\mathscr{B}\colon \mathcal{U} \times \mathcal{V} \to \mathbb{R}$. We assume that $\mathscr{B}$ is continuous on $\mathcal{U} \times \mathcal{V}$, so that the operator induced by $\mathscr{B}$ (again denoted by $\mathscr{B}$) is linear and bounded, i.e., $\mathscr{B} \in \mathcal{L}(\mathcal{U}; \mathcal{V}')$. The generality of considering normed vector spaces as trial–test spaces instead of Hilbert spaces will allow us to address the variational problem (VPM_*) satisfied by the second moment of the solution to (SDE) with multiplicative noise – recall that the trial and test spaces $\mathcal{X}_\pi$ and $\mathcal{Y}_\epsilon$ are non-reflexive Banach spaces.

We assume that $\mathcal{U}_h \times \mathcal{V}_h \subset \mathcal{U} \times \mathcal{V}$ is a fixed pair of non-trivial subspaces with equal finite dimension $\dim \mathcal{U}_h = \dim \mathcal{V}_h < +\infty$, and we aim at approximating the solution $u \in \mathcal{U}$ of the variational problem above by a function $u_h \in \mathcal{U}_h$ and quantifying the error $\|u - u_h\|_{\mathcal{U}}$. For this purpose, suppose that the operator $\bar{\mathscr{B}}\colon \mathcal{U} \to \mathcal{V}'$ is an approximation of $\mathscr{B}$, which again is continuous. We introduce the notation

$$\|\ell\|_{\mathcal{V}_h'} := \sup_{v \in S(\mathcal{V}_h)} |\ell(v)|$$

for functionals ℓ which are defined on $\mathcal{V}_h$, and assume that the approximation $\bar{\mathscr{B}}$ admits a constant $\bar{\gamma}_h > 0$ such that

$$\|\bar{\mathscr{B}} w_h\|_{\mathcal{V}_h'} \geq \bar{\gamma}_h \|w_h\|_{\mathcal{U}} \quad \forall w_h \in \mathcal{U}_h.$$

In other words, the corresponding bilinear form $\bar{\mathscr{B}}$ satisfies the following discrete inf-sup condition:

$$\inf_{w_h \in S(\mathcal{U}_h)} \sup_{v_h \in S(\mathcal{V}_h)} \bar{\mathscr{B}}(w_h, v_h) \geq \bar{\gamma}_h > 0.$$

We then define the approximate solution u_h as the solution of the following discrete variational problem:

$$\text{Find} \quad u_h \in \mathcal{U}_h: \quad \bar{\mathscr{B}}(u_h, v_h) = \ell(v_h) \quad \forall v_h \in \mathcal{V}_h.$$

Recall that we only assume that the discrete trial and test spaces $\mathcal{U}_h$ and $\mathcal{V}_h$ are of the same dimension and that they may differ. In this case when $\mathcal{U}_h \neq \mathcal{V}_h$, the discrete variational problem is said to be a Petrov–Galerkin discretization and its solution $u_h \in \mathcal{U}_h$ is called a Petrov–Galerkin approximation.

The following proposition ensures existence and uniqueness of the Petrov–Galerkin approximation. In addition, it quantifies the error $\|u - u_h\|_{\mathcal{U}}$, which is of importance for the convergence analysis of the Petrov–Galerkin discretizations discussed in Paper II.

PROPOSITION 5.1. *Fix $u \in \mathcal{U}$. Under the above assumptions there exists a unique $u_h \in \mathcal{U}_h$ such that*

$$\bar{\mathscr{B}}(u_h, v_h) = \mathscr{B}(u, v_h) \quad \forall v_h \in \mathcal{V}_h.$$

The mapping $u \mapsto u_h$ is linear with $\|u_h\|_{\mathcal{U}} \le \bar{\gamma}_h^{-1} \|\mathscr{B}u\|_{\mathcal{V}_h'}$, and satisfies the quasi-optimality estimate

$$\|u - u_h\|_{\mathcal{U}} \le \left(1 + \bar{\gamma}_h^{-1} \|\bar{\mathscr{B}}\|_{\mathcal{L}(\mathcal{U};\mathcal{V}')}\right) \inf_{w_h \in \mathcal{U}_h} \|u - w_h\|_{\mathcal{U}} + \bar{\gamma}_h^{-1} \|(\mathscr{B} - \bar{\mathscr{B}})u\|_{\mathcal{V}_h'}.$$

PROOF. Injectivity of the operator $\bar{\mathscr{B}}$ on $\mathcal{U}_h$ follows from the discrete inf-sup condition imposed above. Since $\dim \mathcal{U}_h = \dim \mathcal{V}_h$, the operator $\bar{\mathscr{B}} \colon \mathcal{U}_h \to \mathcal{V}_h'$ is an isomorphism and existence and uniqueness of u_h follows.

In order to derive the quasi-optimality estimate, fix $w_h \in \mathcal{U}_h$. By the triangle inequality, $\|u - u_h\|_{\mathcal{U}} \le \|u - w_h\|_{\mathcal{U}} + \|w_h - u_h\|_{\mathcal{U}}$. Due to the discrete inf-sup condition we can bound the second term as follows:

$$\begin{aligned} \bar{\gamma}_h \|w_h - u_h\|_{\mathcal{U}} &\le \sup_{v_h \in \mathcal{V}_h} \bar{\mathscr{B}}(w_h - u_h, v_h) = \sup_{v_h \in \mathcal{V}_h} [\bar{\mathscr{B}}(w_h, v_h) - \mathscr{B}(u, v_h)] \\ &\le \sup_{v_h \in \mathcal{V}_h} \bar{\mathscr{B}}(w_h - u, v_h) + \sup_{v_h \in \mathcal{V}_h} [\bar{\mathscr{B}}(u, v_h) - \mathscr{B}(u, v_h)] \\ &\le \|\bar{\mathscr{B}}\|_{\mathcal{L}(\mathcal{U};\mathcal{V}')} \|u - w_h\|_{\mathcal{U}} + \|(\mathscr{B} - \bar{\mathscr{B}})u\|_{\mathcal{V}_h'}. \end{aligned}$$

Therefore, for arbitrary $w_h \in \mathcal{U}_h$ we may estimate the error $\|u - u_h\|_{\mathcal{U}}$ by

$$\|u - u_h\|_{\mathcal{U}} \le \left(1 + \bar{\gamma}_h^{-1} \|\bar{\mathscr{B}}\|_{\mathcal{L}(\mathcal{U};\mathcal{V}')}\right) \|u - w_h\|_{\mathcal{U}} + \bar{\gamma}_h^{-1} \|(\mathscr{B} - \bar{\mathscr{B}})u\|_{\mathcal{V}_h'}$$

and taking the infimum with respect to $w_h \in \mathcal{U}_h$ proves the assertion. ☐

6. Summary of Paper II

In Paper II we consider the canonical examples of stochastic ODEs with additive or multiplicative Wiener noise, namely the Ornstein–Uhlenbeck process (SODE1) and the geometric Brownian motion (SODE2) from Subsection 3.3.

We first recall the deterministic equations in variational form satisfied by the first and second moments of the solution processes. As already seen in Paper I, the equations for the second moment and the covariance are posed on tensor products of function spaces. In the additive case (VPM$_+$) they can be taken as the Hilbert tensor product spaces $\mathcal{X}_2$ and $\mathcal{Y}_2$ and well-posedness is readily seen, since $B \colon (\mathcal{X}, \|\cdot\|_{\mathcal{X}}) \to (\mathcal{Y}, |||\cdot|||_{\mathcal{Y}})$ is an isometric isomorphism. In the multiplicative case (VPM$_*$), however, the pair $\mathcal{X}_\pi \times \mathcal{Y}_\epsilon$ of projective–injective tensor product spaces as trial–test spaces is required and well-posedness, proven in Paper I under Assumption (G1), is not an immediate consequence anymore due to the presence of the trace product Δ in the operator induced by the bilinear form $\mathcal{B}$. For the considered example (SODE2) with multiplicative noise, we prove well-posedness even beyond the assumption (G1) imposed in Paper I.

Afterwards, we focus on deriving numerical approximations for the mean and the second moment. We start by discussing different Petrov–Galerkin discretizations for the variational problem (VPm) satisfied by the first moment. From these, Petrov–Galerkin discretizations based on tensor product piecewise polynomials are constructed, which are then applied to the variational problems (VPM_+) and (VPM_*) for the second moments. We discuss briefly the discretization of (VPM_+) and focus on the more sophisticated multiplicative case (VPM_*). We prove stability of the discrete solution with respect to the projective norm on $\mathcal{X}_\pi$ and conclude that a discrete inf-sup condition is satisfied. Therefore, Proposition 5.1 is applicable, which yields a quasi-optimality estimate. From this, convergence of the discrete solution to the exact solution in $\mathcal{X}_\pi$ is derived.

These results should be useful for developing numerical methods for moments of solutions to stochastic partial differential evolution equations by investigating the underlying family of ODEs in the spectral space of the differential operator A.

References

[1] D. Applebaum, *Lévy Processes and Stochastic Calculus*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2009.

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

[3] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 5 Evolution Problems I*, Springer, 2000.

[4] N. Dunford and J. Schwartz, *Linear Operators. Part I: General Theory*, Pure and Applied Mathematics, Interscience Publishers, 1958. With the assistance of W. G. Bade and R. G. Bartle.

[5] ——, *Linear Operators. Part II: Spectral theory. Self Adjoint Operators in Hilbert Space*, Pure and Applied Mathematics, Interscience Publishers, 1963.

[6] K.-J. Engel and R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, vol. 194 of Graduate Texts in Mathematics, Springer, 2000.

[7] A. Grothendieck, *La théorie de Fredholm*, Bull. Soc. Math. France, 84 (1956), pp. 319–384.

[8] W. Hackbusch, *Tensor Spaces and Numerical Tensor Calculus*, Springer Series in Computational Mathematics, Springer, 2012.

[9] W. Johnson and J. Lindenstrauss, *Handbook of the Geometry of Banach Spaces, Vol. 2*, Elsevier Science B.V., 2003.

[10] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, Springer, 1998.

[11] A. Lang, S. Larsson, and Ch. Schwab, *Covariance structure of parabolic stochastic partial differential equations*, Stoch. PDE: Anal. Comp., 1 (2013), pp. 351–364.

[12] A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, Progress in Nonlinear Differential Equations and their Applications, Birkhäuser Verlag, 1995.

[13] J. Mercer, *Functions of positive and negative type and their connection with the theory of integral equations*, Philosophical Transactions of the Royal Society A, 209 (1909), pp. 415–446.

[14] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, Universitext, Springer, 6th ed., 2013.

[15] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied Mathematical Sciences, Springer, 1983.

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

[17] R. Ryan, *Introduction to Tensor Products of Banach Spaces*, Springer Monographs in Mathematics, Springer, 2002.

[18] R. Schatten, *A Theory of Cross-Spaces*, Annals of Mathematics Studies, no. 26, Princeton University Press, Princeton, N. J., 1950.

[19] Ch. Schwab and R. Stevenson, *Space-time adaptive wavelet methods for parabolic evolution problems*, Math. Comp., 78 (2009), pp. 1293–1318.

[20] Ch. Schwab and E. Süli, *Adaptive Galerkin approximation algorithms for Kolmogorov equations in infinite dimensions*, Stoch. PDE: Anal. Comp., 1 (2013), pp. 204–239.

[21] K. Yosida, *Functional Analysis*, Classics in Mathematics, Springer, 6th ed., 1995.