Optimal Control Problems Constrained by Stochastic Partial Differential Equations

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

von **M. Sc. Christoph Trautwein**

geb. am **23.12.1989** in Merseburg (Saale)

genehmigt durch die Fakultät für Mathematik

der Otto-von-Guericke-Universität Magdeburg

Gutachter: **Prof. Dr. Peter Benner**

Prof. Dr. Wilfried Grecksch

eingereicht am: **29.10.2018**

Verteidigung am: **19.12.2018**

List of Publications

Preprints

- P. Benner and C. Trautwein. *A Stochastic Maximum Principle for Control Problems Constrained by the Stochastic Navier-Stokes Equations.* arXiv:1810.12119, 2018.
- P. Benner and C. Trautwein. *Optimal Control Problems Constrained by the Stochastic Navier-Stokes Equations with Multiplicative Lévy Noise.* accepted for publication in Math. Nachr., 2018.
- P. Benner and C. Trautwein. *Optimal Distributed and Tangential Boundary Control for the Unsteady Stochastic Stokes Equations.* arXiv:1809.00911, 2018.

Publications

- P. Benner and C. Trautwein. *A linear quadratic control problem for the stochastic heat equation driven by Q-Wiener processes.* J. Math. Anal. Appl., 457(1):776-802, 2018.

Abstract

In this thesis, we solve several optimal control problems constrained by linear as well as nonlinear stochastic partial differential equations by a stochastic maximum principle. We provide some basic concepts from functional analysis and a stochastic calculus to obtain existence and uniqueness results of mild solutions to these equations. For the linear case, we consider two specific examples, where we involve nonhomogeneous boundary conditions using the theory of fractional powers of closed operators. First, we treat the stochastic heat equation with nonhomogeneous Neumann boundary conditions, where controls and additive noise terms appear inside the domain as well as on the boundary. Here, the control problem is described by tracking a desired state at the terminal point of time leading to a convex optimization problem. Using a stochastic maximum principle, we state necessary and sufficient optimality conditions, which we utilize to design explicit formulas for the optimal controls. By a reformulation of these formulas, we finally obtain a feedback law of the optimal controls. Next, we consider the stochastic Stokes equations with nonhomogeneous Dirichlet boundary conditions, where we include a linear multiplicative noise term. Here, controls appear inside the domain as well as on the boundary. The control problem is defined by tracking a desired state through the whole time interval leading to a convex optimization problem. Again, we state necessary and sufficient optimality conditions the optimal controls have to satisfy. The design of these optimal controls is mainly based on a duality principle giving relations between the mild solutions of forward equations and a backward equation. Here, the forward equations are given by the partial Gâteaux derivatives of the stochastic Stokes equations with respect to the controls and the backward equation is characterized by the adjoint equation. To derive this duality principle, an approximation of the mild solutions by strong solutions is required, which we obtain using the resolvent operator. This provides formulas for the optimal controls based on the adjoint equation. As a consequence, it remains to solve a system of coupled forward and backward stochastic partial differential equations. For the nonlinear case, we study the stochastic Navier-Stokes equations with homogeneous Dirichlet boundary conditions, where we include a linear multiplicative noise term. Here, the theory of fractional powers of closed operators gives a treatment of the convection term arising in these equations. In general, it is not possible to define a solution over an arbitrary time interval. We overcome this problem using a local mild solution well defined upto a certain stopping time. Hence, the cost functional related to the control problem has to incorporate this stopping time leading to a nonconvex optimization problem. Thus, a stochastic maximum principle provides only a necessary optimality condition. However, we still design the optimal control based on the adjoint equation using a duality principle. Again, it remains to solve a system of coupled forward and backward stochastic partial differential equations. Furthermore, we show that the optimal control satisfies a sufficient optimality condition based on the second order Fréchet derivative of the cost functional.

Zusammenfassung

In der vorliegenden Arbeit werden verschiedene Optimalsteuerprobleme sowohl für lineare als auch nichtlineare stochastische partielle Differentialgleichungen mittels eines stochastischen Maximumprinzips gelöst. Wir führen einige Grundlagen aus der Funktionalanalysis und ein stochastisches Kalkül ein, um Existenz- und Eindeutigkeitsresultate von milden Lösungen dieser Gleichungen zu erhalten. Im linearen Fall betrachten wir zwei konkrete Beispiele, wobei wir inhomogene Randbedingungen, unter Verwendung der Theorie der abgeschlossenen Operatoren mit gebrochen Exponenten, einbeziehen. Zunächst behandeln wir die stochastische Wärmeleitungsgleichung mit inhomogenen Neumann-Randbedingungen, wobei Steuerungen und additive Rauschterme sowohl im Gebiet als auch auf dem Rand auftreten. Hier wird das Steuerproblem durch die Verfolgung eines gewünschten Zustandes zum Endzeitpunkt beschrieben, was zu einem konvexen Optimierungsproblem führt. Mittels eines stochastischen Maximumprinzips geben wir notwendige und hinreichende Optimalitätsbedingungen an, welche wir verwenden, um explizite Formeln für die optimalen Steuerungen zu konstruieren. Durch eine Umformulierung erhalten wir letztendlich, dass die optimalen Steuerungen als Rückkoppelungssteuerung dargestellt werden kann. Danach betrachten wir die stochastischen Stokes Gleichungen mit inhomogenen Dirichlet-Randbedingungen, wobei wir einen linear-multiplikativen Rauschterm einbeziehen. Hier treten Steuerungen sowohl im Gebiet als auch auf dem Rand auf. Das Steuerproblem besteht aus einer Verfolgung eines gewünschten Zustandes über einem bestimmten Zeitintervall, was zu einem konvexen Optimierungsproblem führt. Wieder geben wir notwendige und hinreichende Optimalitätsbedingungen an, welche die optimalen Steuerungen erfüllen. Die Konstruktion der optimalen Steuerungen basiert vorwiegend auf einem Dualitätsprinzip, welches Zusammenhänge zwischen den milden Lösungen von Vorwärtsgleichungen und einer Rückwärtsgleichung angibt. Die Vorwärstgleichungen sind durch die partiellen Gâteaux-Ableitungen der Lösungen der stochastischen Stokes Gleichungen bezüglich der Steuerungen gegeben und die Rückwartsgleichung ist charakterisiert durch die adjungierte Gleichung. Um dieses Dualitätsprinzip herzuleiten, ist eine Approximation der milden Lösungen durch starke Lösungen erforderlich, welche wir mittels der Resolvente erlangen. Wir erhalten somit Formeln für die optimalen Steuerungen basierend auf der adjungierten Gleichung. Somit bleibt ein System von gekoppelten stochastischen partiellen Vorwärts- und Rückwärtsgleichungen zu lösen. Im nichtlinearen Fall analysieren wir die stochastischen Navier-Stokes Gleichungen mit homogenen Dirichlet-Randbedingungen, wobei wir einen linear-multiplikativen Rauschterm einbeziehen. Hier gibt uns die Theorie der abgeschlossenen Operatoren mit gebrochen Exponenten eine Möglichkeit den Konvektionsterm in diesen Gleichungen handhabbar zu machen. Im Allgemeinen ist es nicht möglich eine Lösung über einem beliebigen Zeitintervall zu definieren. Wir bewältigen dieses Problem, indem wir eine lokale milde Lösung verwenden, welche bis zu einer gewissen Stoppzeit wohldefiniert ist. Demzufolge muss das zum Steuerproblem gehörige Kostenfunktional diese Stoppzeit einbeziehen, was uns zu einem nicht-konvexen Optimierungsproblem führt. Dadurch gibt uns ein stochastisches Maximumprinzip lediglich notwendige Optimalitätsbedingungen. Nichtsdestotrotz konstruieren wir die optimalen Steuerungen basierend auf der adjungierten Gleichung unter Verwendung eines Dualitätsprinzips. Ferner zeigen wir, dass die optimale Steuerung eine hinreichende Optimalitätsbedingung, unter Verwendung der Fréchet Ableitung zweiter Ordnung des Kostenfunktionals, erfüllt.

Contents

Notation

General

$\mathbb{N}$	natural numbers $\{1, 2, ...\}$
$\mathbb{Z}$, $\mathbb{Z}^+$	integers, nonnegative integers
$\mathbb{R}$, $\mathbb{R}^+$, $\mathbb{R}_0^+$	real numbers, nonnegative real numbers, positive real numbers
$\mathbb{R}^{n\times m}$	real matrices with n rows and m columns
$\mathbb{C}$	complex numbers
$\mathcal{B}(\mathcal{X})$	Borel σ-field of a Banach space $\mathcal{X}$
$\mathcal{D}$	bounded domain in $\mathbb{R}^n$, i.e. an open and bounded subset of $\mathbb{R}^n$
$\partial\mathcal{D}$	boundary of $\mathcal{D}$
$\text{Im } z$	imaginary part of $z \in \mathbb{C}$
$\overline{M}$	closure of a set M
$M_1 \cap M_2$	intersection of sets M_1 and M_2
$M_1 \cup M_2$	union of sets M_1 and M_2
$\bigcup_{n=1}^{\infty} M_n$	union of a sequence of sets $(M_n)_{n\in\mathbb{N}}$
$M_1 \backslash M_2$	relative complement of a set M_2 in a set M_1
$\text{Re } z$	real part of $z \in \mathbb{C}$
$s \wedge t$, $s \vee t$	$\min\{s, t\}$, $\max\{s, t\}$ with $s, t \in \mathbb{R}$
$[t_0, t_1]$	closed interval from $t_0 \in \mathbb{R}^+$ to $t_1 \in \mathbb{R}^+$ with $t_0 \leq t_1$
$(t_0, t_1]$, $[t_0, t_1)$	half-closed interval from $t_0 \in \mathbb{R}^+$ to $t_1 \in \mathbb{R}^+$ with $t_0 < t_1$
$\mathcal{X}^n$	n-dimensional vector space of a Banach space $\mathcal{X}$
$\emptyset$	empty set

Operators and Functions

A^*	adjoint of an operator A
A^α	fractional power of an operator A with $\alpha \in \mathbb{R}$
$D(A)$	domain of an operator A
$\det(A)$	determinant of a matrix A
div	divergence of a vector field
$R(\lambda; A)$	resolvent operator of an operator A with $\lambda \in \rho(A)$
$R(\lambda)$	$\lambda R(\lambda; A)$
$(S(t))_{t\geq 0}$, $(e^{At})_{t\geq 0}$	C_0 semigroup generated by an operator A
$Tr(A)$	trace of an operator A
$\rho(A)$	resolvent set of an operator A
$\lVert \cdot \rVert_{\mathcal{X}}$	norm on a Banach space $\mathcal{X}$
$\langle \cdot, \cdot \rangle_{\mathcal{X}}$	inner product on a Hilbert space $\mathcal{X}$
∇, Δ	Nabla operator, Laplace operator
$\mathbb{1}_{M_1}$	indicator function of a subset M_1 of a set M_2

Probability Theory

$(\Omega, \mathcal{F}, \mathbb{P})$	complete probability space
$\mathbb{E}[X]$	expected value of a random variable X
$\mathbb{E}[X\|\mathcal{G}]$	conditional expectation of a random variable X given a σ-field $\mathcal{G}$
$\mathbb{P}(\mathcal{A})$	probability of an event $\mathcal{A} \in \mathcal{F}$
$(\Delta L(t))_{t\geq 0}$	jump process of a Lévy process $(L(t))_{t\geq 0}$
$\sigma(M)$	smallest σ-field containing a set M
$\sigma(X(s) \colon 0 \leq s \leq t)$	smallest σ-field generated by a stochastic process $(X(s))_{s\in\mathcal{I}}$ upto a point of time $t \in \mathcal{I}$, where $\mathcal{I} = [0, T]$ or $\mathcal{I} = \mathbb{R}^+$

Spaces

$C([t_0, t_1]; \mathcal{X})$	continuous functions mapping $[t_0, t_1]$ into a Banach space $\mathcal{X}$
$C(\mathcal{D})$, $C(\partial\mathcal{D})$	continuous real functions on $\mathcal{D}$ or $\partial\mathcal{D}$
$C^\infty(\mathcal{D})$	infinite differentiable functions in $C(\mathcal{D})$ or $C(\partial\mathcal{D})$ with continuous derivatives
$C_0^\infty(\mathcal{D})$	functions in $C^\infty(\mathcal{D})$ with compact support
$H^s(\mathcal{D})$, $H^s(\partial\mathcal{D})$	Sobolev space of square integrable real functions on $\mathcal{D}$ or $\partial\mathcal{D}$ in the sense of Bessel potential spaces with $s \geq 0$
$H_0^s(\mathcal{D})$	functions in $H^s(\mathcal{D})$ with compact support and $s > \frac{1}{2}$
$L^p([t_0, t_1]; \mathcal{X})$	p-integrable functions for $1 \leq p < \infty$ mapping $[t_0, t_1]$ into a Banach space $\mathcal{X}$
$L^\infty([t_0, t_1]; \mathcal{X})$	measurable functions mapping $[t_0, t_1]$ into a Banach space $\mathcal{X}$ such that the essential supremum is finite
$L^p([t_0, t_1])$	$L^p([t_0, t_1]; \mathbb{R})$
$L^p(\mathcal{D})$, $L^p(\partial\mathcal{D})$	p-integrable real functions on $\mathcal{D}$ or $\partial\mathcal{D}$
$L^p(\Omega; \mathcal{X})$	functions mapping Ω into a Banach space $\mathcal{X}$, which are p-integrable with respect to a measure for $1 \leq p < \infty$
$L^p_{\mathcal{F}}(\Omega; L^q([t_0, t_1]; \mathcal{X}))$	stochastic processes in $L^p(\Omega; L^q([t_0, t_1]; \mathcal{X}))$ adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$
$\mathcal{L}(\mathcal{X}; \mathcal{Y})$	linear and bounded operators mapping a Banach space $\mathcal{X}$ into another Banach space $\mathcal{Y}$
$\mathcal{L}(\mathcal{X})$	$\mathcal{L}(\mathcal{X}; \mathcal{X})$
$\mathcal{L}_1(\mathcal{X}; \mathcal{Y})$	nuclear operators mapping $\mathcal{X}$ into $\mathcal{Y}$
$\mathcal{L}_1(\mathcal{X})$	$\mathcal{L}_1(\mathcal{X}; \mathcal{X})$
$\mathcal{L}_1^+(\mathcal{X}; \mathcal{Y})$	self-adjoint and nonnegative operators in $\mathcal{L}_1(\mathcal{X}; \mathcal{Y})$
$\mathcal{L}_1^+(\mathcal{X})$	$\mathcal{L}_1^+(\mathcal{X}; \mathcal{X})$
$\mathcal{L}_{(HS)}(\mathcal{X}; \mathcal{Y})$	Hilbert-Schmidt operators mapping a Banach space $\mathcal{X}$ into another Banach space $\mathcal{Y}$
$\mathcal{L}_{(HS)}(\mathcal{X})$	$\mathcal{L}_{(HS)}(\mathcal{X}; \mathcal{X})$
$Q^{1/2}(\mathcal{X})$	subspace of a Hilbert space $\mathcal{X}$ generated by a self-adjoint nonnegative operator $Q^{1/2} \in \mathcal{L}(\mathcal{X})$
$\mathcal{X} \times \mathcal{Y}$	product space of Banach spaces $\mathcal{X}$ and $\mathcal{Y}$
$\mathcal{X}'$	dual space of a Banach space $\mathcal{X}$

Chapter 1

Introduction

1.1. Stochastic Systems

Unsteady deterministic ordinary differential equations and unsteady deterministic partial differential equations arise as models for many systems in engineering, chemistry, biology and physics. To cover random environmental phenomena affecting theses systems, it is often required to involve noise terms as stochastic processes leading to stochastic systems. Consequently, stochastic systems can always be motivated from the deterministic approach. Furthermore, the state described by such a system is not differentiable with respect to the time variable in general. A possibility to overcome this difficulty is given by reformulating the differential equations as integral equations. This leads us immediately to stochastic differential equations (SDEs) and stochastic partial differential equations (SPDEs), which are symbolic notions describing these integral equations. Basically, one can consider SDEs driven by a Wiener noise, see [54, 55, 69, 66] and the references therein. SPDEs with respect to Wiener noise can be considered as a generalization of SDEs in the sense that these equations are formulated as evolution equations on infinite dimensional spaces. Here, we will focus on infinite dimensional spaces given by separable Hilbert spaces. Thus, the solutions of SPDEs are defined in a generalized sense in various ways. As a direct ansatz, one can formulate a strong solution, see [23, 28, 42, 45]. This often requires too strong regularity properties of the solution to the SPDE and thus, one introduces weaker concepts to obtain an equation well defined on a larger space. In [23, 28, 42, 73], weak solutions are introduced, where the construction is mainly based on the inner product defined on a suitable Hilbert space. Using Gelfand triples, a similar approach is given by variational solutions, see [73, 80]. Mild solutions are often used for problems containing a linear (and possibly unbounded) operator as the generator of a semigroup, see [23, 28, 42]. All of these concepts are based on a given probability space and therefore, they are called (probabilistic) strong solutions. Solutions constructing the probability space are called (probabilistic) weak solutions or martingale solutions, see [23, 28]. For various reasons, a Wiener noise cannot cover all random environmental phenomena. Thus, it is often required to use more general noise terms. One may consider systems including jumps leading to Lévy noise. For SDEs, we refer to [2, 20, 82]. An approach for SPDEs is presented in [3, 18, 71]. In a different direction one can consider a noise term, where the increments are not necessarily independent. Such a noise term can be modeled using a fractional Brownian motion. SDEs driven by fractional Brownian motions are studied in [64]. For SPDEs, an approach is given in [29, 62].

In this thesis, we will mainly concentrate on systems described by SPDEs with Lévy noise. Here, we analyze the following classes in more detail:

(i) linear SPDEs with additive Lévy noise,

(ii) linear SPDEs with multiplicative Lévy noise,

(iii) nonlinear SPDEs with multiplicative Lévy noise.

Based on [71], we prove existence and uniqueness results of these equations, where we incorporate some additional difficulties. On the one hand, we treat a possibility to involve nonhomogeneous boundary conditions appearing in linear SPDEs. As examples, the stochastic heat equation with Neumann boundary condition as well as the stochastic Stokes equations with Dirichlet boundary condition will be analyzed. On the other hand, we consider a nonlinear SPDE, where the nonlinearity does not satisfy the usual assumptions given by a growth condition and a Lipschitz condition. Here, the stochastic Navier-Stokes equations with homogeneous Dirichlet boundary condition will be treated as an example. These systems have in common that they contain a linear and closed operator generating an analytic semigroup such that fractional powers of these operators (possibly with a suitable perturbation) are well defined. We will figure out that the theory of fractional powers of closed operators is useful to overcome the difficulties mentioned above, where the solutions of the SPDEs are defined in a mild sense.

1.2. Stochastic Control

Due to the presence of a noise term, it might be the case that the state of the system reveals an undesired behavior. Thus, it it reasonable to control a system in a certain desired way, where we always assume that the state is completely observable. This immediately leads us to a stochastic control problem (in infinite dimensions), which we consider as an optimization problem for a given cost functional constrained by a SPDE. The minimizer of the cost functional is then called an optimal control. To solve this problem, there exist mainly two approaches:

(i) stochastic maximum principle;

(ii) dynamic programming.

Based on existence and uniqueness results for the solution to a SPDE, one can often reformulate the control problem as a minimization problem on a set of admissible controls given by a suitable Hilbert space or a suitable subset of this Hilbert space. For that reason, the main idea of the stochastic maximum principle is to state necessary and sufficient optimality conditions the optimal control has to satisfy. In general, the necessary optimality condition can be derived using the Gâteaux derivative of the cost functional. Using this necessary optimality condition, one can derive an explicit formula of the optimal control based on the adjoint equation, which is given by a backward stochastic partial differential equation (BSPDE). Sufficient optimality conditions are often stated based on the second order Fréchet derivative of the cost functional. If the control problem is additionally convex, then the necessary optimality condition is also sufficient. For general concepts of optimization problems on Hilbert spaces, we refer to [57, 93]. Closely related is Pontryagin's maximum principle, where one minimizes the Hamiltonian instead of the original control problem. However, one still obtains an explicit formula of the optimal control based on the adjoint equation. As a consequence, it remains to solve the so called Hamiltonian system. For applications, we refer to [14, 36, 47, 67]. In this context, we may also note the general theory for finite dimensional control problems presented in [91]. In contrast to these methods, the dynamic programming principle considers the control problem at different initial times and initial states through the so called value function. This value function is the solution of a nonlinear partial differential equation given by the Hamilton-Jacobi-Bellman equation. If the equation is solvable, then one can obtain a feedback law of the optimal control, see [33]. For applications, we refer to [22, 26, 32, 61, 83, 92]. We also note the finite dimensional approach presented in [35, 91].

The scope of this thesis is to provide a theory for solutions to specific stochastic control problems such that they can be treated numerically. Since sufficient optimality conditions are useful to obtain the convergence to an optimal control, we use a stochastic maximum principle here. This fact is already known for deterministic problems, see [51]. As mentioned above, the design of the optimal controls is based on the adjoint equation given by a BSPDE. To obtain the existence and uniqueness of a solution to the adjoint equation, it is often required to apply a martingale representation theorem. Since a martingale representation theorem is not available for Hilbert space valued Lévy processes in general, we are forced to restrict ourself to the case of

Q-Wiener processes. However, we will state some possibilities to expand the theory to systems governed by SPDE with Lévy noise.

1.3. Outline of the Thesis

This thesis is divided into two main parts. In the first part, we provide foundations from functional analysis and a stochastic calculus in infinite dimensional spaces required for the second part, where we solve certain stochastic control problems via a stochastic maximum principle.

In Chapter 2, we introduce the class of linear (not necessary bounded) operators $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ generating a C_0 semigroup $(S(t))_{t\geq 0}$ on an arbitrary Hilbert space $\mathcal{H}$. We state some basic properties and we introduce the resolvent operator $R(\lambda; A) = (\lambda I - A)^{-1}$ for appropriate $\lambda \in \mathbb{C}$, where the operator I denotes the identity operator on $\mathcal{H}$. Here, we will use the resolvent operator to approximate mild solutions of SPDEs and BSPDEs by strong solutions, which is required to obtain a so called duality principle. Furthermore, we introduce fractional powers of the operator A denoted by A^α with $\alpha \in \mathbb{R}$. If the C_0 semigroup $(S(t))_{t\geq 0}$ is analytic and the operator A is invertible, then we get some additional properties, which enable us to incorporate nonhomogeneous boundary data to SPDEs. Moreover, we get a possible treatment of the convection term arising in the stochastic Navier-Stokes equations. Especially, we will use the following inequality frequently:

$$\|A^\alpha S(t)\|_{\mathcal{L}(\mathcal{H})} \leq M_\alpha t^{-\alpha} e^{-\delta t}$$

for all $\alpha > 0$ and all $t > 0$, where $M_\alpha, \delta > 0$ are constants. Thus, this inequality is the main result of this chapter. Finally, we consider the Laplace operator and the Stokes operator as typical examples for generators of analytic semigroups with their fractional powers (with a possible modification) being well defined.

Chapter 3 is devoted to the stochastic calculus used in the following chapters. We start with some basic definitions and we introduce Lévy processes $(L(t))_{t\geq 0}$ with values in an arbitrary Hilbert space $\mathcal{U}$. In general, a Lévy process has the following decomposition for all $t \geq 0$ and $\mathbb{P}$-almost surely:

$$L(t) = at + W(t) + J(t),$$

where $a \in \mathcal{U}$ represents the drift, $(W(t))_{t\geq 0}$ is the continuous part given by an $\mathcal{U}$-valued Q-Wiener process and $(J(t))_{t\geq 0}$ illustrates the pure jump part characterized by a series of $\mathcal{U}$-valued compound Poisson process. When studying stochastic equations, it is necessary to define a stochastic integral of the form

$$\int_0^t \Psi(s)\, dL(s)$$

for all $t \in [0, T]$ with $T > 0$ and $\mathbb{P}$-almost surely, where $(\Psi(t))_{t\in[0,T]}$ is a stochastic process taking values in a suitable space of Hilbert-Schmidt operators. Here, we assume that the Lévy process $(L(t))_{t\geq 0}$ is square integrable and a martingale with respect to a certain filtration. We will state basic properties of such a stochastic integral, which enables us to prove existence and uniqueness results of mild solutions to SPDEs driven by Lévy processes. For the existence and uniqueness of mild solutions to BSPDEs, a martingale representation theorem is often required. Since such a theorem is not available for Hilbert space valued Lévy processes in general, we will study BSPDEs for the special of Q-Wiener processes. The SPDEs and BSPDEs introduced here are motivated by systems arising in the following chapters. Moreover, we give a comparison of strong, weak and mild solutions to these equations.

In Chapter 4, we consider a control problem constrained by the stochastic heat equation with nonhomogeneous Neumann boundary conditions on a bounded domain $\mathcal{D} \subset \mathbb{R}^n$ with sufficiently smooth boundary

$\partial\mathcal{D}$. Namely, we will treat the following SPDE in $L^2(\mathcal{D})$:

$$\begin{cases} dy(t) = [Ay(t) + Bu(t) + (\lambda - A)Nv(t)]\,dt + G(t)\,dW(t) + (\lambda - A)N\,dW_b(t), \\ y(0) = \xi \end{cases}$$

for $t \in [0,T]$. Here, the operator $A\colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ is the Neumann realization of the Laplace operator generating an analytic semigroup of contractions $\left(e^{At}\right)_{t\geq 0}$. The process $(u(t))_{t\in[0,T]}$ represents a distributed control with values in $L^2(\mathcal{D})$ and B is a linear and bounded operator on $L^2(\mathcal{D})$. The process $(v(t))_{t\in[0,T]}$ describes a boundary control with values in $L^2(\partial\mathcal{D})$ and $N\colon L^2(\partial\mathcal{D}) \to L^2(\mathcal{D})$ denotes the Neumann operator. The real number λ is chosen such that fractional powers of the operator $\lambda - A$ are well defined. The noise terms $(W(t))_{t\geq 0}$ and $(W_b(t))_{t\geq 0}$ are given by Q-Wiener processes with values in $L^2(\mathcal{D})$ and $L^2(\partial\mathcal{D})$, respectively. We denote by $Q \in \mathcal{L}_1^+(L^2(\mathcal{D}))$ and $Q_b \in \mathcal{L}_1^+(L^2(\partial\mathcal{D}))$ the covariance operators of the processes $(W(t))_{t\geq 0}$ or $(W_b(t))_{t\geq 0}$, respectively. The process $(G(t))_{t\in[0,T]}$ takes values in $\mathcal{L}_{(HS)}(Q^{1/2}(L^2(\mathcal{D})); L^2(\mathcal{D}))$. As a consequence, controls and noise terms are defined inside the domain as well as on the boundary. The cost functional related to the control problem is formulated as follows:

$$J(u,v) = \frac{1}{2}\mathbb{E}\|y(T) - y_d\|^2_{L^2(\mathcal{D})} + \frac{\kappa_1}{2}\mathbb{E}\int_0^T \|u(t)\|^2_{L^2(\mathcal{D})}dt + \frac{\kappa_2}{2}\mathbb{E}\int_0^T \|v(t)\|^2_{L^2(\partial\mathcal{D})}dt,$$

where $y_d \in L^2(\mathcal{D})$ is a given desired state and $\kappa_1, \kappa_2 > 0$ are weights. The task is to find optimal controls $\overline{u}$ and $\overline{v}$ minimizing this cost functional. The corresponding optimal state is denoted by $(\overline{y}(t))_{t\in[0,T]}$. Employing a stochastic maximum principle, we will show that the optimal controls satisfy the following feedback law for all $\alpha \in (\frac{1}{2}, \frac{3}{4})$, almost all $t \in [0,T]$ and $\mathbb{P}$-almost surely:

$$\begin{aligned} \overline{u}(t) &= -\frac{1}{\kappa_1}B^*[\mathcal{P}(t)\overline{y}(t) + a(t)], \\ \overline{v}(t) &= -\frac{1}{\kappa_2}\mathcal{G}^*(\lambda - A)^{1-\alpha}[\mathcal{P}(t)\overline{y}(t) + a(t)], \end{aligned}$$

where B^* and $\mathcal{G}^*$ denote the adjoint operators of B and $\mathcal{G} = (\lambda - A)^\alpha N$, respectively. The function $\mathcal{P}\colon [0,T] \to \mathcal{L}(L^2(\mathcal{D}))$ is the mild solution of the Riccati equation

$$\begin{cases} \dfrac{d}{dt}\mathcal{P}(t) = A\mathcal{P}(t) + \mathcal{P}(t)A - \dfrac{1}{\kappa_1}\mathcal{P}(t)BB^*\mathcal{P}(t) - \dfrac{1}{\kappa_2}\mathcal{H}^*(t)\mathcal{G}\mathcal{G}^*\mathcal{H}(t), \\ \mathcal{P}(T) = I, \end{cases}$$

where $\mathcal{H}(t) = (\lambda - A)^{1-\alpha}\mathcal{P}(t)$ and I is the identity operator on $L^2(\mathcal{D})$. The function $a\colon [0,T] \to D((\lambda - A)^{1-\alpha})$ is the unique solution of the deterministic backward integral equation

$$a(t) = \int_t^T e^{A(s-t)}\left(-\frac{1}{\kappa_1}\mathcal{P}(s)BB^* - \frac{1}{\kappa_2}\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*(\lambda - A)^{1-\alpha}\right)a(s)\,ds - e^{A(T-t)}y_d.$$

In Chapter 5, we study a control problem constrained by the stochastic Stokes equation with nonhomogeneous Dirichlet boundary conditions on a bounded domain $\mathcal{D} \subset \mathbb{R}^n$ with sufficiently smooth boundary $\partial\mathcal{D}$. In fact, we will deal with the following SPDE in $H = \left\{y \in (L^2(\mathcal{D}))^n : \operatorname{div} y = 0 \text{ in } \mathcal{D}, y \cdot \eta = 0 \text{ on } \partial\mathcal{D}\right\}$:

$$\begin{cases} dy(t) = [-Ay(t) + Bu(t) + ADv(t)]\,dt + G(y(t))\,dW(t), \\ y(0) = \xi. \end{cases}$$

Above, the operator $A\colon D(A)\subset H\to H$ is the Stokes operator. Fractional powers of A are well defined and denoted by A^α with $\alpha\in\mathbb{R}$. The process $(u(t))_{t\in[0,T]}$ represents a distributed control with values in H and B is a linear and bounded operator on H. The process $(v(t))_{t\in[0,T]}$ describes a boundary control with values in $V^0(\partial\mathcal{D})=\left\{y\in\left(L^2(\partial\mathcal{D})\right)^n : y\cdot\eta=0 \text{ on } \partial\mathcal{D}\right\}$ and $D\colon V^0(\partial\mathcal{D}))\to H$ denotes the Dirichlet operator. The noise term $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in H and covariance operator $Q\in\mathcal{L}_1^+(H)$. The operator $G\colon H\to\mathcal{L}_{(HS)}(Q^{1/2}(H);H)$ is linear and bounded. Here, we will consider the following cost functional:

$$J(u,v)=\frac{1}{2}\,\mathbb{E}\int_0^T\|y(t;u,v)-y_d(t)\|_H^2dt+\frac{\kappa_1}{2}\,\mathbb{E}\int_0^T\|u(t)\|_H^2\,dt+\frac{\kappa_2}{2}\,\mathbb{E}\int_0^T\|v(t)\|_{V^0(\partial\mathcal{D})}^2\,dt,$$

where $y_d\in L^2([0,T];H)$ is a given desired velocity field and $\kappa_1,\kappa_2>0$ are weights. The task is to find optimal controls $\overline{u}$ and $\overline{v}$ as minimizers of this cost functional. Using a stochastic maximum principle, we will obtain that the optimal controls satisfy for all $\alpha\in(0,\frac{1}{4})$, almost all $t\in[0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned}\overline{u}(t)&=-\frac{1}{\kappa_1}\,B^*z^*(t),\\ \overline{v}(t)&=-\frac{1}{\kappa_2}\,K^*A^{1-\alpha}z^*(t),\end{aligned}$$

where B^* and K^* are the adjoint operators of B and $K=A^\alpha D$, respectively. The process $(z^*(t))_{t\in[0,T]}$ is characterized by the adjoint equation given by the following BSPDE in H:

$$\begin{cases}dz^*(t)=-[-Az^*(t)+G^*(\Phi(t))+y(t)-y_d(t)]dt+\Phi(t)\,dW(t),\\ z^*(T)=0,\end{cases}$$

where the operator G^* is the adjoint operator of G and the process $(\Phi(t))_{t\in[0,T]}$ takes values in the space $\mathcal{L}_{(HS)}(Q^{1/2}(H);H)$. As a consequence, it remains to solve a system of coupled forward and backward SPDEs.

In Chapter 6, we treat a control problem governed by the stochastic Navier-Stokes equations with homogeneous Dirichlet boundary conditions on a bounded domain $\mathcal{D}\subset\mathbb{R}^n$ with sufficiently smooth boundary $\partial\mathcal{D}$. Indeed, we will study the following SPDE in $D(A^\alpha)$ for suitable $\alpha>0$:

$$\begin{cases}dy(t)=-[Ay(t)+B(y(t))-Fu(t)]dt+G(y(t))\,dW(t),\\ y(0)=\xi.\end{cases}$$

Again, the operator $A\colon D(A)\subset H\to H$ is the Stokes operator and B is a bilinear operator related to the convection term arising in the Navier-Stokes equations. The operator $A^{-\delta}B$ is well defined as a mapping from $D(A^\alpha)$ into H for certain $\delta\geq 0$. The process $(u(t))_{t\in[0,T]}$ represents a distributed control with values in $D(A^\beta)$ with $\beta\in[0,\alpha]$ and F is a linear and bounded operator on $D(A^\beta)$. The noise term $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in H and covariance operator $Q\in\mathcal{L}_1^+(H)$. The operator $G\colon H\to\mathcal{L}_{(HS)}(Q^{1/2}(H);D(A^\alpha))$ is linear and bounded. Due to the presence of the bilinear operator B, we cannot ensure the existence and uniqueness of a mild solution over an arbitrary time interval $[0,T]$. However, we will show that there exists a unique mild solution upto a stopping time τ_m for fixed $m\in\mathbb{N}$. Thus, the cost functional related to the control problem has to incorporate this stopping time. In fact, the cost functional is given by

$$J_m(u)=\frac{1}{2}\,\mathbb{E}\int_0^{\tau_m}\|A^\gamma(y(t)-y_d(t))\|_H^2\,dt+\frac{1}{2}\,\mathbb{E}\int_0^T\|A^\beta u(t)\|_H^2dt$$

for fixed $m \in \mathbb{N}$, where $y_d \in L^2([0,T]; D(A^\gamma))$ with $\gamma \in [0,\alpha]$ is a given desired state. The task is to find a optimal control $\overline{u}_m$ minimizing this cost functional. By a stochastic maximum principle, we will prove that the optimal control satisfies for almost all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\overline{u}_m(t) = -P_U\left(F^* A^{-2\beta} z_m^*(t)\right),$$

where P_U is a projection onto the set of admissible controls U and F^* is the adjoint operator of F. The process $(z_m^*(t))_{t\in[0,T]}$ is described by the BSPDE in $D(A^\delta)$:

$$\begin{cases} dz_m^*(t) = -\mathbb{1}_{[0,\tau_m)}(t)[-Az_m^*(t) - A^{2\alpha} B_\delta^*\left(y(t), A^\delta z_m^*(t)\right) + G^*(A^{-2\alpha}\Phi_m(t)) \\ \qquad\qquad + A^{2\gamma}(y(t) - y_d(t))]dt + \Phi_m(t)\, dW(t), \\ z_m^*(T) = 0, \end{cases}$$

where the operator $B_\delta^*(y(t),\cdot)$ is the adjoint operator of $A^{-\delta}B(\cdot, y(t))$ for $t \in [0,\tau_m)$. Similarly, the operator G^* is the adjoint operator of G and the process $(\Phi_m(t))_{t\in[0,T]}$ takes values in $\mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$. Again, we can conclude that it remains to solve a system of coupled forward and backward SPDEs.

In the appendix, we provide some useful Gronwall-type inequalities. Moreover, we introduce Bochner integrals as well as nuclear and Hilbert-Schmidt operators. These are the basic foundations to define solutions to SPDEs. Finally, we treat optimization problems in infinite dimensional spaces, which enables us to solve control problems constrained by SPDEs. The results stated in this part are well known. However, we give a brief overview for the convenience of the reader.

Chapter 2

Infinitesimal Generators of Analytic Semigroups

In this chapter, we give some basic properties of strongly continuous semigroups and their infinitesimal generators, see [31, 70, 89]. We mainly focus on infinitesimal generators as closed operators such that their fractional powers can be defined. If the strongly continuous semigroup is analytic, then further regularity results and estimates can be obtained, which we use frequently in the following chapters. Finally, we consider the Laplace operator and the Stokes operator defined on bounded domains with sufficiently smooth boundary. Here, we treat domains as open subsets and the characterization of the boundary as introduced in [48]. We will ascertain that fractional powers of the Laplace operator as well as the Stokes operator are well defined. The results shown here are mainly based on [30, 48, 85, 89].

Throughout this chapter, let $\mathcal{H}$ be a Hilbert space and let I be the identity operator on $\mathcal{H}$. We note that most of the following results remain still true for Banach spaces.

2.1. Strongly Continuous Semigroups and the Resolvent Operator

In this section, we give basic definitions and basic properties of strongly continuous semigroups and their infinitesimal generators. We introduce the resolvent set and resolvent operator of a closed operator. An integral representation of the resolvent operator is provided and we state necessary and sufficient conditions such that the closed operator is the infinitesimal generator of a strongly continuous semigroup of contractions well known as the Hille-Yosida theorem. We start with a formal definition.

Definition 2.1. *A family of linear and bounded operators $(S(t))_{t\geq 0}$ mapping $\mathcal{H}$ into itself is called a* ***semigroup*** *if*

(i) $S(0) = I$;

(ii) $S(t+s) = S(t)S(s)$ *for all* $s, t \geq 0$.

The semigroup $(S(t))_{t\geq 0}$ mapping $\mathcal{H}$ into itself is called a ***strongly continuous semigroup*** *or a* ***C_0 semigroup*** *if for every $x \in \mathcal{H}$*

$$\lim_{t\downarrow 0} \|S(t)x - x\|_{\mathcal{H}} = 0.$$

Theorem 2.2 (Chapter 1, Theorem 2.2, [70])**.** *Let $(S(t))_{t\geq 0}$ be a C_0 semigroup. There exist constants $\theta \in \mathbb{R}$ and $M \geq 1$ such that for all $t \geq 0$*

$$\|S(t)\|_{\mathcal{L}(\mathcal{H})} \leq M e^{\theta t}. \tag{2.1}$$

Remark 2.3. *If $\theta = 0$ in inequality (2.1), then $(S(t))_{t\geq 0}$ is called a uniformly bounded C_0 semigroup. We call $(S(t))_{t\geq 0}$ a C_0 semigroup of contractions if additionally $M = 1$.*

Corollary 2.4 (Chapter 1, Corollary 2.3, [70])**.** *If $(S(t))_{t\geq 0}$ is a C_0 semigroup, then for every $x \in \mathcal{H}$, the mapping $t \mapsto S(t)x$ is a continuous function from $\mathbb{R}^+$ into $\mathcal{H}$.*

Definition 2.5. *An operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ is called the **infinitesimal generator** or simply **generator** of a C_0 semigroup $(S(t))_{t\geq 0}$ if*

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

for every $x \in D(A)$ with

$$D(A) = \left\{ x \in \mathcal{H}\colon \lim_{t\downarrow 0} \frac{S(t)x - x}{t} \text{ exists} \right\}.$$

The set $D(A)$ is called the domain of the operator A.

The generator of a C_0 semigroup is a linear and closed operator but not necessarily bounded. The domain is a dense subset of the underlying Hilbert space.

Theorem 2.6 ([31, 70, 89])**.** *Let $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be the generator of a C_0 semigroup $(S(t))_{t\geq 0}$. Then the following properties hold:*

- *if $x \in D(A)$, then $S(t)x \in D(A)$ and*

$$\frac{d}{dt} S(t)x = AS(t)x = S(t)Ax$$

  for all $t \geq 0$;

- *for all $t \geq 0$ and every $x \in \mathcal{H}$, we have*

$$\int_0^t S(s)x\, ds \in D(A) \quad \text{and} \quad S(t)x - x = A \int_0^t S(s)x\, ds;$$

- *for all $t \geq 0$ and every $x \in D(A)$, we have*

$$S(t)x - S(s)x = \int_s^t AS(r)x\, dr = \int_s^t S(r)Ax\, dr.$$

Using these properties and the closed graph theorem, we get a characterization of uniformly continuous semigroups. Let $\mathcal{L}(\mathcal{H})$ contain all linear and bounded operators on $\mathcal{H}$.

Definition 2.7. *A semigroup $(S(t))_{t\geq 0}$ is called **uniformly continuous** if*

$$\lim_{t\downarrow 0} \|S(t) - I\|_{\mathcal{L}(\mathcal{H})} = 0$$

Corollary 2.8 (Chapter 2, Corollary 1.5, [31])**.** *Let $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be the generator of a C_0 semigroup $(S(t))_{t\geq 0}$. The following assertions are equivalent:*

(a) The operator A is bounded.

(b) The domain of A satisfies $D(A) = \mathcal{H}$.

(c) The domain $D(A)$ is closed in $\mathcal{H}$.

(d) The semigroup $(S(t))_{t\geq 0}$ is uniformly continuous.

In each case, the semigroup is given by

$$S(t) = \sum_{n=0}^{\infty} \frac{t^n A^n}{n!}$$

for all $t \geq 0$.

Let $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be a linear (not necessarily bounded) operator. We introduce the resolvent set $\rho(A)$ containing all complex numbers λ for which $\lambda I - A$ is invertible, i.e.

$$\rho(A) = \{\lambda \in \mathbb{C}\colon (\lambda I - A)^{-1} \text{ exists and belongs to } \mathcal{L}(\mathcal{H})\}.$$

We write $\lambda - A$ instead of $\lambda I - A$ to simplify the notation. For all $\lambda \in \rho(A)$, we define the resolvent operator $R(\lambda; A) \in \mathcal{L}(\mathcal{H})$ by

$$R(\lambda; A) = (\lambda - A)^{-1}.$$

We have the following characterization of elements of the resolvent set and an integral representation of the resolvent operator.

Theorem 2.9 (Chapter 2, Theorem 1.10, [31])**.** *Let* $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ *be the generator of a* C_0 *semigroup* $(S(t))_{t\geq 0}$ *and take constants* $\theta \in \mathbb{R}$ *and* $M \geq 1$ *such that for all* $t \geq 0$

$$\|S(t)\|_{\mathcal{L}(\mathcal{H})} \leq M e^{\theta t}.$$

Then we have the following properties:

(i) If $\lambda \in \mathbb{C}$ *such that* $\int_0^\infty e^{-\lambda t} S(t)\, dt$ *exists for every* $x \in \mathcal{H}$*, then* $\lambda \in \rho(A)$*.*

(ii) If $\operatorname{Re}\lambda > \theta$*, then* $\lambda \in \rho(A)$ *and* $\|R(\lambda; A)\|_{\mathcal{L}(\mathcal{H})} \leq \frac{M}{\operatorname{Re}\lambda - \theta}$*.*

In each case, the resolvent operator is given by

$$R(\lambda; A) = \int_0^\infty e^{-\lambda t} S(t)\, dt.$$

Corollary 2.10 (Chapter 2, Corollary 1.11, [31])**.** *Let* $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ *be the generator of a* C_0 *semigroup* $(S(t))_{t\geq 0}$ *and take constants* $\theta \in \mathbb{R}$ *and* $M \geq 1$ *such that for all* $t \geq 0$

$$\|S(t)\|_{\mathcal{L}(\mathcal{H})} \leq M e^{\theta t}.$$

For all $\lambda \in \mathbb{C}$ *with* $\operatorname{Re}\lambda > \theta$ *and each* $n \in \mathbb{N}$*, we have*

$$\begin{aligned} R(\lambda; A)^n &= \frac{(-1)^{n-1}}{(n-1)!} \frac{d^{n-1}}{d\lambda^{n-1}} R(\lambda; A) \\ &= \frac{1}{(n-1)!} \int_0^\infty t^{n-1} e^{-\lambda t} S(t)\, dt \end{aligned}$$

and $\|R(\lambda; A)^n\|_{\mathcal{L}(\mathcal{H})} \leq \frac{M}{(\operatorname{Re}\lambda - \theta)^n}$*.*

Next, we state necessary and sufficient conditions such that the operator A is the generator of a C_0 semigroup of contractions well known as the Hille-Yosida theorem.

Theorem 2.11 (Chapter 1, Theorem 3.1, [70])**.** *An operator* $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ *is the generator of a* C_0 *semigroup of contractions* $(S(t))_{t\geq 0}$ *if and only if* *(i) A is closed and $D(A)$ is dense in $\mathcal{H}$;*

(ii) the resolvent set $\rho(A)$ contains $\mathbb{R}_0^+$ and for all $\lambda > 0$

$$\|R(\lambda;A)\|_{\mathcal{L}(\mathcal{H})} \leq \frac{1}{\lambda}.$$

The previous theorem and its proof have some simple consequences on convergence results of the so called Yosida approximation.

Corollary 2.12 (Section 1.3, [70])**.** *Let $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be the generator of a C_0 semigroup of contractions $(S(t))_{t\geq 0}$ and let A_λ be the Yosida approximation of A given by*

$$A_\lambda = \lambda A R(\lambda; A).$$

Then A_λ is the generator of an uniformly continuous semigroup of contractions $(e^{A_\lambda t})_{t\geq 0}$ and we have

(i) $\lim_{\lambda\to\infty} \lambda R(\lambda;A)x = x$ *for every $x \in \mathcal{H}$;*

(ii) $\lim_{\lambda\to\infty} A_\lambda x = Ax$ *for every $x \in D(A)$;*

(iii) $\lim_{\lambda\to\infty} e^{A_\lambda t}x = S(t)x$ *for every $x \in \mathcal{H}$ and all $t \geq 0$.*

Remark 2.13. *For general versions of Theorem 2.11 and Corollary 2.12 concerning arbitrary C_0 semigroups, we refer to [31, 70].*

The following dilation theorem gives an important property of C_0 semigroups of contractions.

Theorem 2.14. *Let $(S(t))_{t\geq 0}$ be a C_0 semigroup of contractions and set $S(-t) = S(t)^*$ for all $t > 0$. Then there exists a Hilbert space $\widehat{\mathcal{H}}$ containing $\mathcal{H}$ and a group $(\widehat{S}(t))_{t\in\mathbb{R}}$ on $\widehat{\mathcal{H}}$ such that $S(t) = P_{\mathcal{H}}\widehat{S}(t)$ for all $t \in \mathbb{R}$, where $P_{\mathcal{H}}$ is the orthogonal projection from $\widehat{\mathcal{H}}$ onto $\mathcal{H}$.*

Proof. The claim follows from Theorem 9.22 and Theorem 9.23 in [71]. □

2.2. Analytic Semigroups

In this section, we introduce analytic semigroups and we state conditions such that an operator is the generator of an analytic semigroup. We start with a formal definition. The main idea is to extend the domain of the semigroup operator to regions in the complex plane containing $\mathbb{R}^+$. For $\theta \in (0, \pi]$, we define the sector

$$\Sigma_\theta = \{z \in \mathbb{C} \colon |\arg z| < \theta\}.$$

Definition 2.15. *A C_0 semigroup $(S(t))_{t\geq 0}$ is called **analytic** if there exists $\theta \in (0, \pi]$ and a mapping $\tilde{S}\colon \overline{\Sigma}_\theta \to \mathcal{L}(\mathcal{H})$ such that*

- *$S(t) = \tilde{S}(t)$ for all $t \geq 0$;*
- *$\tilde{S}(z_1 + z_2) = \tilde{S}(z_1)\tilde{S}(z_2)$ for every $z_1, z_2 \in \overline{\Sigma}_\theta$;*
- *the mapping $z \mapsto \tilde{S}(z)$ is analytic in Σ_θ;*
- $\lim_{z\to 0, z\in\overline{\Sigma}_\theta} S(z)x = x$ *for every $x \in \mathcal{H}$.*

To state conditions on an operator to be the generator of an analytic semigroup, we need the concept of differentiable semigroups.

Definition 2.16. *A C_0 semigroup $(S(t))_{t\geq 0}$ is called **differentiable** for $t > t_0$ if for every $x \in \mathcal{H}$, the mapping $t \mapsto S(t)x$ is differentiable for $t > t_0$. The derivative of order $n \in \mathbb{N}$ is denoted by $S^{(n)}(t)$ for $t > t_0$.*

The following lemma provides an useful presentation of the derivatives to a differentiable semigroup.

Lemma 2.17 (Chapter 2, Lemma 4.2, [70])**.** *Let $(S(t))_{t\geq 0}$ be a differentiable C_0 semigroup for $t > t_0$ and let $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be its generator. Then*

- *for $n \in \mathbb{N}$ and $t > nt_0$, we have $S(t)\colon \mathcal{H} \to D(A^n)$ and $S^{(n)}(t) = A^n S(t)$ is a bounded linear operator;*
- *for $n \in \mathbb{N}$ and $t > nt_0$, the operator $S^{(n-1)}(t)$ is continuous in the uniform operator topology.*

We are now able to state basic properties of analytic semigroups.

Theorem 2.18 ([70, 89])**.** *Let $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be the generator of a C_0 semigroup $(S(t))_{t\geq 0}$. If $0 \in \rho(A)$, then the following statements are equivalent:*

(a) The C_0 semigroup $(S(t))_{t\geq 0}$ is uniformly bounded and analytic.

(b) For all $\lambda \in \mathbb{C}$ with $\operatorname{Re}\lambda > 0$ and $\operatorname{Im}\lambda \neq 0$, there exists a constant $C > 0$ such that

$$\|R(\lambda; A)\|_{\mathcal{L}(\mathcal{H})} \leq \frac{C}{|\operatorname{Im}\lambda|}.$$

(c) There exist $\theta \in (0, \frac{\pi}{2})$ and a constant $M > 0$ such that $\Sigma_{\frac{\pi}{2}+\theta} \cup \{0\} \subset \rho(A)$ and

$$\|R(\lambda; A)\|_{\mathcal{L}(\mathcal{H})} \leq \frac{M}{|\lambda|}$$

for $\lambda \in \Sigma_{\frac{\pi}{2}+\theta}$.

(d) The semigroup $(S(t))_{t\geq 0}$ is differentiable for $t > 0$ and there exists a constant $C > 0$ such that for all $t > 0$

$$\|AS(t)\|_{\mathcal{L}(\mathcal{H})} \leq \frac{C}{t}.$$

Under additional assumptions, we can state a further generation theorem of analytic semigroups resulting from the previous theorem. First, we define the adjoint operator of a linear operator. This requires the following preliminary result.

Lemma 2.19 (Lemma 4.1.4, [85])**.** *Let $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be linear and densely defined. Then for every $y \in \mathcal{H}$, there exists at most one element $z \in \mathcal{H}$ such that for every $x \in D(A)$*

$$\langle Ax, y\rangle_{\mathcal{H}} = \langle x, z\rangle_{\mathcal{H}}.$$

Definition 2.20. *Let $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be linear and densely defined. We set*

$$D(A^*) = \{y \in \mathcal{H}\colon \text{there exists } z \in \mathcal{H} \text{ such that } \langle Ax, y\rangle_{\mathcal{H}} = \langle x, z\rangle_{\mathcal{H}} \text{ for every } x \in D(A)\}.$$

*The **adjoint operator** $A^*\colon D(A^*) \subset \mathcal{H} \to \mathcal{H}$ is defined by $A^*y = z$ for every $y \in D(A^*)$.*

By Lemma 2.19, the element $z \in \mathcal{H}$ in the above definition is unique. This justifies the notation $A^*y = z$.

Definition 2.21. *A linear and densely defined operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ is called* ***symmetric*** *if for every $x, y \in D(A)$*

$$\langle Ax, y\rangle_{\mathcal{H}} = \langle x, Ay\rangle_{\mathcal{H}}.$$

The operator A is called ***self-adjoint*** *if $A = A^*$ and $D(A) = D(A^*)$.*

Remark 2.22. *Obviously, a self-adjoint operator is symmetric. The converse is generally not true.*

The following corollary gives some simple requirements on a C_0 semigroup to be analytic.

Corollary 2.23 (Corollary 7.1.1, [89])**.** *If the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ is self-adjoint and the generator of a C_0 semigroup of contractions $(S(t))_{t\geq 0}$, then $(S(t))_{t\geq 0}$ analytic.*

We will often use this corollary to obtain that a C_0 semigroup is analytic.

2.3. Fractional Powers of Closed Operators

In this section, we define fractional powers of closed operators. We give conditions such that these operators are well defined. Moreover, we state some basic properties, which are used frequently in the following chapters. First, we introduce the gamma function given by

$$\Gamma(\alpha) = \int_0^\infty s^{\alpha-1} e^{-s}\, ds$$

for all $\alpha > 0$. A change of variables with $s = ct$ for $c > 0$ gives us

$$c^{-\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-ct}\, dt. \tag{2.2}$$

Let $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be a linear (not necessarily bounded) operator such that $-A$ is the generator of a C_0 semigroup $(S(t))_{t\geq 0}$. It is quite natural to consider equation (2.2) by substituting c with A and e^{-ct} with $S(t)$. Note that we can write at least formally $S(t) = e^{-At}$. We then have the following definition.

Definition 2.24. *For $\alpha > 0$, the operator $A^{-\alpha}\colon D(A^{-\alpha}) \subset \mathcal{H} \to \mathcal{H}$ given by*

$$A^{-\alpha}x = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} S(t)x\, dt$$

for every $x \in D(A^{-\alpha})$ is called the ***fractional power*** *of the operator A with exponent $-\alpha$. The domain of $A^{-\alpha}$ is given by*

$$D(A^{-\alpha}) = \left\{ x \in \mathcal{H}\colon \int_0^\infty t^{\alpha-1} S(t)x\, dt \text{ is convergent} \right\}.$$

For $\alpha = 0$, we set $A^0 = I$ and $D(A^0) = \mathcal{H}$.

The space $D(A^{-\alpha})$ is a linear subspace of $\mathcal{H}$ and $A^{-\alpha}$ is a linear and closed operator for $\alpha \geq 0$. Moreover, we have for $0 \leq \alpha \leq \beta$

$$D(A^{-\beta}) \subset D(A^{-\alpha}).$$

Remark 2.25. *If $-A$ is the generator of a C_0 semigroup $(S(t))_{t\geq 0}$ satisfying for all $t \geq 0$*

$$\|S(t)\|_{\mathcal{L}(\mathcal{H})} \leq Me^{-\theta t}$$

with $\theta > 0$ and $M \geq 1$, then $D(A^{-\alpha}) = \mathcal{H}$ and $A^{-\alpha}$ is a linear and bounded operator. Indeed, one can easily obtain

$$\int_0^\infty t^{\alpha-1}\|S(t)\|_{\mathcal{L}(\mathcal{H})}dt < \infty.$$

The fact that $A^{-\alpha}$ is linear and bounded follows immediately from the closed graph theorem.

In the remaining part of this section, we assume that $-A$ satisfies the assumptions of Remark 2.25.

Remark 2.26. *By definition of the resolvent operator and Corollary 2.10, we have for $\lambda = 0$ and each $n \in \mathbb{N}$*

$$\left(A^{-1}\right)^n = R(0;-A)^n = \frac{1}{(n-1)!}\int_0^\infty t^{n-1}S(t)\,dt.$$

Recall that $\Gamma(n) = (n-1)!$ for each $n \in \mathbb{N}$. Hence, the operator A^{-n} given by Definition 2.24 coincides with the classical representation of the operator $\left(A^{-1}\right)^n$ for each $n \in \mathbb{N}$.

In the following lemma, we state some basic properties.

Lemma 2.27 (Section 7.6, [89])**.** *We have*

(i) $A^{-(\alpha+\beta)} = A^{-\alpha}A^{-\beta}$ *for all* $\alpha, \beta \geq 0$*;*

(ii) for all $\alpha \in (0,1)$

$$A^{-\alpha} = \frac{\sin \pi\alpha}{\pi}\int_0^\infty \lambda^{-\alpha}(\lambda + A)^{-1}d\lambda;$$

(iii) $\lim_{\alpha\downarrow 0} A^{-\alpha}x = x$ *for every* $x \in \mathcal{H}$*;*

(iv) the operator $A^{-\alpha}$ is injective for all $\alpha \geq 0$.

The fact that the operator $A^{-\alpha}$ is injective for all $\alpha \geq 0$ allows us to define fractional powers of the operator A for any positive real number.

Definition 2.28. *Let $A^{-\alpha}$ be the fractional power of the operator A with $\alpha > 0$. We define*

$$A^\alpha = \left(A^{-\alpha}\right)^{-1}.$$

We get the following basic properties.

Theorem 2.29 (Chapter 2, Theorem 6.8, [70])**.** *We have*

(i) $A^\alpha \colon D(A^\alpha) \subset \mathcal{H} \to \mathcal{H}$ *is a closed operator with* $D(A^\alpha) = \mathcal{R}(A^{-\alpha})$ *for all* $\alpha > 0$*, where* $\mathcal{R}(A^{-\alpha})$ *denotes the range of the operator* $A^{-\alpha}$*;*

(ii) $D(A^\beta) \subset D(A^\alpha)$ *for all* $0 \leq \alpha \leq \beta$*;*

(iii) for all $\alpha > 0$, the domain $D(A^\alpha)$ is dense in $\mathcal{H}$;

(iv) $A^{\alpha+\beta}x = A^\alpha A^\beta x$ *for all* $\alpha, \beta \in \mathbb{R}$ *and every* $x \in D(A^\gamma)$ *with* $\gamma = \max\{\alpha, \beta, \alpha + \beta\}$*.*

In general, one can not give an explicit formula for the operator A^α with $\alpha > 0$. Nevertheless, we get the following representation, which is an immediate consequence of Lemma 2.27.

Theorem 2.30 (Theorem 7.6.2, [89]). *If $\alpha \in (0,1)$ and $x \in D(A)$, then*

$$A^\alpha x = \frac{\sin \pi\alpha}{\pi} \int_0^\infty \lambda^{\alpha-1}(\lambda + A)^{-1} Ax \, d\lambda.$$

Next, we state some useful estimates.

Theorem 2.31 (Theorem 7.6.3, [89]). *If $\alpha \in (0,1)$, then there exists a constant $C > 0$ such that for every $x \in D(A)$ and all $\rho > 0$*

$$\|A^\alpha x\|_{\mathcal{H}} \leq C(\rho^\alpha \|x\|_{\mathcal{H}} + \rho^{\alpha-1}\|Ax\|_{\mathcal{H}})$$

and

$$\|A^\alpha x\|_X \leq 2C\|x\|_{\mathcal{H}}^{1-\alpha}\|Ax\|_{\mathcal{H}}^\alpha.$$

Corollary 2.32. *Let $\alpha \in (0,1]$ and let $B\colon D(B) \subset \mathcal{H} \to \mathcal{H}$ be a closed operator with $D(A^\alpha) \subset D(B)$. There exists a constant $C > 0$ such that for every $x \in D(A^\alpha)$*

$$\|Bx\|_{\mathcal{H}} \leq C\|A^\alpha x\|_{\mathcal{H}} \tag{2.3}$$

and in particular

$$\|A^\beta x\|_{\mathcal{H}} \leq C\|A^\alpha x\|_{\mathcal{H}} \tag{2.4}$$

for $0 \leq \beta < \alpha \leq 1$. Moreover, there exists a constant $C_1 > 0$ such that for every $x \in D(A)$ and all $\rho > 0$

$$\|Bx\|_{\mathcal{H}} \leq C_1(\rho^\alpha \|x\|_{\mathcal{H}} + \rho^{\alpha-1}\|Ax\|_{\mathcal{H}}). \tag{2.5}$$

Proof. A proof of inequalities (2.3) and (2.5) can be found in [89, Corollary 7.6.2]. Inequality (2.4) follows from inequality (2.3) and Theorem 2.29 (ii). □

Theorem 2.33. *For $\alpha \geq 0$, the space $D(A^\alpha)$ equipped with the inner product*

$$\langle x, y\rangle_{D(A^\alpha)} = \langle A^\alpha x, A^\alpha y\rangle_{\mathcal{H}}$$

for every $x, y \in D(A^\alpha)$ becomes a Hilbert space.

Proof. The norm on $D(A^\alpha)$ is given by

$$\|x\|_{D(A^\alpha)} = \sqrt{\langle x, x\rangle_{D(A^\alpha)}}$$

for every $x \in D(A^\alpha)$. Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $D(A^\alpha)$. Then $(A^\alpha x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{H}$. Since $\mathcal{H}$ is a Hilbert space, there exists $y \in \mathcal{H}$ such that $\lim_{n\to\infty} \|y - A^\alpha x_n\|_{\mathcal{H}} = 0$. Using inequality (2.4) with $\beta = 0$, we have for each $n, m \in \mathbb{N}$

$$\|x_n - x_m\|_{\mathcal{H}} \leq C\|A^\alpha x_n - A^\alpha x_m\|_{\mathcal{H}}.$$

We conclude that the sequence $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{H}$ and there exists $x \in \mathcal{H}$ such that $\lim_{n\to\infty} \|x - x_n\|_{\mathcal{H}} = 0$. Since A^α is closed, we have $x \in D(A^\alpha)$ and $A^\alpha x = y$. Therefore, we obtain $\lim_{n\to\infty} \|A^\alpha x_n - A^\alpha x\|_{\mathcal{H}} = 0$. Therefore, the sequence $(x_n)_{n\in\mathbb{N}}$ converges in $D(A^\alpha)$. □

Lemma 2.34. *If the operator $-A$ is self-adjoint, then A^α is self-adjoint for all $\alpha \in \mathbb{R}$.*

Proof. First, we show the claim for negative exponents. Recall that the operator $-A$ is the generator of a C_0 semigroup $(S(t))_{t\geq 0}$. Since the operator $-A$ is self-adjoint, the semigroup $(S(t))_{t\geq 0}$ is self-adjoint as well. By Definition 2.24, we get for every $x_1, x_2 \in \mathcal{H}$ and all $\alpha > 0$

$$\begin{aligned} \left\langle A^{-\alpha} x_1, x_2 \right\rangle_{\mathcal{H}} &= \left\langle \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} S(t) x_1 \, dt, x_2 \right\rangle_{\mathcal{H}} = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} \left\langle S(t) x_1, x_2 \right\rangle_{\mathcal{H}} dt \\ &= \left\langle x_1, \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} S(t) x_2 \, dt \right\rangle_{\mathcal{H}} = \left\langle x_1, A^{-\alpha} x_2 \right\rangle_{\mathcal{H}}. \end{aligned} \tag{2.6}$$

Next, we show the claim for positive exponents. Using Theorem 2.29 (iv) and equation (2.6), we obtain for every $x_1, x_2 \in D(A^\alpha)$ and all $\alpha > 0$

$$\langle A^\alpha x_1, x_2 \rangle_{\mathcal{H}} = \left\langle A^\alpha x_1, A^{-\alpha} A^\alpha x_2 \right\rangle_{\mathcal{H}} = \left\langle A^{-\alpha} A^\alpha x_1, A^\alpha x_2 \right\rangle_{\mathcal{H}} = \langle x_1, A^\alpha x_2 \rangle_{\mathcal{H}}.$$

For $\alpha = 0$, the claim is obvious. □

Under additionally requirements, we get the following regularity results and useful estimates.

Theorem 2.35. *Let $-A$ be the generator of an analytic semigroup $(S(t))_{t\geq 0}$ satisfying the assumptions of Remark 2.25. If $0 \in \rho(A)$, then*

(i) $S(t)\colon \mathcal{H} \to D(A^\alpha)$ for all $t > 0$ and all $\alpha \in \mathbb{R}$;

(ii) for every $x \in D(A^\alpha)$ and all $\alpha \in \mathbb{R}$, we have $A^\alpha S(t) x = S(t) A^\alpha x$;

(iii) the operator $A^\alpha S(t)$ is linear and bounded for all $t > 0$ and all $\alpha \in \mathbb{R}$. In addition, there exist constants $M_\alpha, \delta > 0$ such that for all $t > 0$ and all $\alpha > 0$

$$\|A^\alpha S(t)\|_{\mathcal{L}(\mathcal{H})} \leq M_\alpha t^{-\alpha} e^{-\delta t};$$

(iv) for all $\alpha \in (0, 1]$, there exists a constant $C_\alpha > 0$ such that for every $x \in D(A^\alpha)$

$$\|S(t)x - x\|_{\mathcal{H}} \leq C_\alpha t^\alpha \|A^\alpha x\|_{\mathcal{H}}.$$

Proof. The proof can be found in [70, Chapter 2, Theorem 6.13] and [89, Theorem 7.7.2]. □

Remark 2.36. *The previous theorem is the main result of this chapter and used frequently in the following chapters. Hence, we will often require that an analytic semigroup satisfying Remark 2.25. Moreover, the number 0 has to be an element of the resolvent set.*

Corollary 2.37. *Let $R(\lambda; -A)$ be the resolvent operator of $-A$ with a real number $\lambda \in \rho(-A)$ such that $\lambda > 0$. If the assumptions of Theorem 2.35 hold, then we have for every $y \in D(A^\alpha)$ with $\alpha < 1$*

$$A^\alpha R(\lambda; -A) y = R(\lambda; -A) A^\alpha y.$$

Proof. Using Theorem 2.9, we get for every $y \in \mathcal{H}$

$$R(\lambda; -A) y = \int_0^\infty e^{-\lambda t} S(t) y \, dt.$$

First, we show the claim for $\alpha \leq 0$. By Remark 2.25, the operator A^α is linear and bounded. Using Theorem 2.35, we obtain

$$A^\alpha R(\lambda; -A)y = A^\alpha \int_0^\infty e^{-\lambda t} S(t)y \, dt = \int_0^\infty e^{-\lambda t} S(t) A^\alpha y \, dt = R(\lambda; -A) A^\alpha y.$$

Next, let $\alpha \in (0,1)$. By Theorem 2.29 (i), the operator A^α is linear and closed. Due to Theorem 2.35 (iii), we have

$$\int_0^\infty e^{-\lambda t} \|A^\alpha S(t) y\|_{\mathcal{H}} dt \leq M_\alpha \int_0^\infty e^{-\lambda t} t^{-\alpha} dt \|y\|_{\mathcal{H}} = M_\alpha \lambda^{\alpha-1} \Gamma(1-\alpha) \|y\|_{\mathcal{H}} < \infty.$$

Hence, the assumptions of Proposition B.9 are fulfilled. Using additionally Theorem 2.35, we get for every $y \in D(A^\alpha)$

$$A^\alpha R(\lambda; -A)y = A^\alpha \int_0^\infty e^{-\lambda t} S(t)y \, dt = \int_0^\infty e^{-\lambda t} S(t) A^\alpha y \, dt = R(\lambda; -A) A^\alpha y.$$

□

2.4. Friedrichs Extension

To obtain that an operator is the generator of an analytic semigroup, we will frequently use Corollary 2.23. This requires a self-adjoint operator, which is not always given. However, one can often show that a self-adjoint extension exists, which is given by the so called Friedrichs extension. First, we introduce the energy space of a linear (not necessarily bounded) operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$. We start with the definition of semi-bounded operators.

Definition 2.38. *A linear and densely defined operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ is called* ***semi-bounded*** *if there exists a constant $c \in \mathbb{R}$ such that for every $x \in D(A)$*

$$\langle Ax, x \rangle_{\mathcal{H}} \geq c \|x\|_{\mathcal{H}}^2.$$

Theorem 2.39. *A semi-bounded operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ is symmetric.*

Proof. By definition, the domain $D(A)$ is dense in $\mathcal{H}$ and $\langle Ax, x \rangle_{\mathcal{H}}$ is real. The claim follows immediately from [85, Theorem 4.1.5 (d)]. □

Let the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be semi-bounded with $\langle Ax, x \rangle_{\mathcal{H}} \geq c\|x\|_{\mathcal{H}}^2$ for every $x \in D(A)$ and let $\lambda \in \mathbb{R}$ such that $\lambda + c > 0$. We set for every $x, y \in D(A)$

$$[x, y]_\lambda = \langle Ax, y \rangle_{\mathcal{H}} + \lambda \langle x, y \rangle_{\mathcal{H}}. \tag{2.7}$$

One can easily verify that $[\cdot,\cdot]_\lambda$ is an inner product on $D(A)$. Then the norm is defined by $\|x\|_\lambda = \sqrt{[x,x]_\lambda}$ for every $x \in D(A)$. Since the space $D(A)$ is not complete with this norm, we need the following construction of the so called energy space.

Definition 2.40. *Let the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be semi-bounded with $\langle Ax, x \rangle_{\mathcal{H}} \geq c\|x\|_{\mathcal{H}}^2$ for every $x \in D(A)$ and let $\lambda \in \mathbb{R}$ such that $\lambda + c > 0$. The* ***energy space*** *H_λ is defined by*

$$H_\lambda = \Big\{ x \in \mathcal{H} \colon \text{there exists a sequence } (x_n)_{n\in\mathbb{N}} \subset D(A) \text{ such that } \lim_{n\to\infty} \|x - x_n\|_{\mathcal{H}} = 0 \text{ and } \lim_{n,m\to\infty} \|x_n - x_m\|_\lambda = 0 \Big\}.$$

The sequence $(x_n)_{n\in\mathbb{N}}$ is called an ***approximating sequence****.*

Lemma 2.41 (Lemma 4.1.8,[85]). *Let $x, y \in H_\lambda$ with approximating sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$, respectively. Then the limit*

$$\lim_{n\to\infty} [x_n, y_n]_\lambda = [x, y]_\lambda$$

exists and is independent of the choice of the approximating sequences.

Due to the previous lemma, we can define the inner product $[\cdot,\cdot]_\lambda$ on the energy space H_λ. Moreover, we get the following properties.

Theorem 2.42 (Theorem 4.1.8,[85]). *Let the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be semi-bounded satisfying $\langle Ax, x\rangle_{\mathcal{H}} \geq c\|x\|^2_{\mathcal{H}}$ for every $x \in D(A)$ and let $\lambda \in \mathbb{R}$ such that $\lambda + c > 0$. The space H_λ equipped with the inner product $[x, y]_\lambda$ for every $x, y \in H_\lambda$ in the sense of Lemma 2.41 becomes a Hilbert space. The domain $D(A)$ is a dense subset of H_λ. If $\mu \in \mathbb{R}$ satisfies $\mu + c > 0$, then $H_\lambda = H_\mu$ and the corresponding norms $\|\cdot\|_\lambda$ and $\|\cdot\|_\mu$ are equivalent.*

Due to the previous theorem, we can conclude that the energy space H_λ depends only on the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ and not on $\lambda \in \mathbb{R}$. Hence, we shall write H_A instead of H_λ. We are now able to state Friedrichs extension theorem.

Theorem 2.43 (Theorem 4.1.9, [85]). *Let the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be semi-bounded satisfying $\langle Ax, x\rangle_{\mathcal{H}} \geq c\|x\|^2_{\mathcal{H}}$ for every $x \in D(A)$. If H_A is the corresponding energy space, then the operator $A_F\colon D(A_F) \subset \mathcal{H} \to \mathcal{H}$ given by*

$$D(A_F) = H_A \cap D(A^*), \quad A_F x = A^* x \text{ for every } x \in D(A_F)$$

is a self-adjoint extension of the operator A. Moreover, we have for every $x \in D(A_F)$

$$\langle A_F x, x\rangle_{\mathcal{H}} \geq c\|x\|^2_{\mathcal{H}}.$$

Remark 2.44. *Let $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ be a semi-bounded operator. By Theorem 2.39, we get that A is symmetric and hence, we obtain $D(A) \subset D(A^*)$ and $Ax = A^*x$ for every $x \in D(A)$. Also note that $D(A) \subset H_A$ since for every $x \in D(A)$ the sequence $(x_n)_{n\in\mathbb{N}}$ given by $x_n = x$ for each $n \in \mathbb{N}$ is an approximating sequence. Therefore, we can conclude that the operator $A_F\colon D(A_F) \subset \mathcal{H} \to \mathcal{H}$ constructed in the previous theorem is an extension of the operator A.*

2.5. Examples

In this section, we consider some important examples of closed operators generating analytic semigroups such that their fractional powers are well defined. Here, we introduce the Laplace operator as well as the Stokes operator defined on L_2-spaces.

2.5.1. The Laplace Operator

Here, we study the Dirichlet realization as well as the Neumann realization of the Laplace operator. Let $x = (x_1, ..., x_n) \in \mathbb{R}^n$. For functions $y\colon \mathbb{R}^n \to \mathbb{R}$, we introduce the nabla operator ∇ given by

$$\nabla y(x) = \left(\frac{\partial y(x)}{\partial x_1}, ..., \frac{\partial y(x)}{\partial x_n}\right)$$

and we introduce the Laplace operator Δ defined by

$$\Delta y(x) = \sum_{i=1}^{n} \frac{\partial^2 y(x)}{\partial x_i^2}.$$

Often, we will omit the dependence on x for the sake of simplicity. Here, we analyze the Laplace operator as a closed operator on the Hilbert space $\mathcal{H} = L^2(\mathcal{D})$, where $\mathcal{D}$ is a bounded domain with sufficiently smooth boundary $\partial\mathcal{D}$.

The Dirichlet Realization of the Laplace Operator

We assume that $\mathcal{D} \subset \mathbb{R}^n$ is a bounded domain with C^2 boundary $\partial\mathcal{D}$. We set $D(A_0) = C_0^\infty(\mathcal{D})$ and we define the operator $A_0 \colon D(A_0) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ by

$$A_0 y = -\Delta y \tag{2.8}$$

for every $y \in D(A_0)$.

Lemma 2.45. *The operator $A_0 \colon D(A_0) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ defined by equation (2.8) is semi-bounded with $\langle A_0 y, y\rangle_{L^2(\mathcal{D})} \geq c\|y\|^2_{L^2(\mathcal{D})}$ for every $y \in D(A_0)$, where $c > 0$ is a constant.*

Proof. It is well known that $D(A_0)$ is dense in the space $L^2(\mathcal{D})$, see [85, Theorem 1.3.6/2]. Obviously, the operator A_0 is linear. By partial integration and the Poincaré inequality, there exists a constant $c > 0$ such that for every $y \in D(A_0)$

$$\langle A_0 y, y\rangle_{L^2(\mathcal{D})} = -\int_{\mathcal{D}} \Delta y(x) y(x)\, dx = \int_{\mathcal{D}} \nabla y(x) \nabla y(x)\, dx = \|\nabla y\|^2_{L^2(\mathcal{D})} \geq c\|y\|^2_{L^2(\mathcal{D})}.$$

Thus, the operator A_0 is semi-bounded. □

Let $A_0^* \colon D(A_0^*) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ be the adjoint operator of A_0. By the previous lemma, we can apply Theorem 2.43 with the result that the Friedrichs extension of the operator A_0 exists. We denote the Friedrichs extension of the operator A_0 by $A \colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$. Moreover, we get the following properties.

Lemma 2.46. *The operator $A \colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ is self-adjoint and we have for every $y \in D(A)$*

$$\langle Ay, y\rangle_{L^2(\mathcal{D})} \geq c\|y\|^2_{L^2(\mathcal{D})}, \tag{2.9}$$

where the constant $c > 0$ arises from Lemma 2.45. Furthermore, we have

$$D(A) = H_0^1(\mathcal{D}) \cap H^2(\mathcal{D}), \quad Ay = -\Delta y$$

for every $y \in D(A)$.

Proof. By Theorem 2.43, we have

$$D(A) = H_{A_0} \cap D(A_0^*), \quad Ay = A_0^* y$$

for every $y \in D(A)$, where H_{A_0} is the energy space of the operator A_0. The operator A is the self-adjoint extension of A_0 and we have for every $y \in D(A)$

$$\langle Ay, y\rangle_{L^2(\mathcal{D})} \geq c\|y\|^2_{L^2(\mathcal{D})},$$

where the constant $c > 0$ arises from Lemma 2.45.

Next, we determine the domain $D(A)$ explicitly. We start with the space H_{A_0}. For every $y, z \in D(A_0)$, let $[y, z]_1$ be given by equation (2.7) with $\lambda = 1$ and $\|y\|_1 = \sqrt{[y, y]_1}$. By partial integration, we obtain for every $y \in D(A_0)$

$$\|y\|_1^2 = \langle A_0 y, y\rangle_{L^2(\mathcal{D})} + \|y\|^2_{L^2(\mathcal{D})} = \|\nabla y\|^2_{L^2(\mathcal{D})} + \|y\|^2_{L^2(\mathcal{D})}.$$

Hence, the norm $\|\cdot\|_1$ is equal to the norm on $H^1(\mathcal{D})$. It is well known that $D(A_0)$ is dense in the space $H_0^1(\mathcal{D})$, see [85, Theorem 1.5.5/1]. By definition of the energy space, we get $H_{A_0} = H_0^1(\mathcal{D})$. To determine the set $D(A_0^*)$, we first calculate the operator A_0^*. In the sense of distributions, we obtain for every $y \in D(A_0)$ and $z \in D(A_0^*)$

$$\langle A_0^* z, y\rangle_{L^2(\mathcal{D})} = \langle z, A_0 y\rangle_{L^2(\mathcal{D})} = -\int_{\mathcal{D}} z(x)\Delta y(x)\,dx = -\int_{\mathcal{D}} \Delta z(x) y(x)\,dx = \langle -\Delta z, y\rangle_{L^2(\mathcal{D})}\,. \tag{2.10}$$

Hence, we get for every $z \in D(A_0^*)$

$$A_0^* z = -\Delta z$$

in the sense of distributions. Moreover, we have

$$D(A_0^*) \subset \left\{ z \in L^2(\mathcal{D})\colon\, -\Delta z \in L^2(\mathcal{D}) \right\}.$$

Conversely, if $z \in L^2(\mathcal{D})$ such that $-\Delta z \in L^2(\mathcal{D})$ in the sense of distributions, then by equation (2.10), we get $z \in D(A_0^*)$. Therefore, we obtain

$$D(A_0^*) = \left\{ z \in L^2(\mathcal{D})\colon\, -\Delta z \in L^2(\mathcal{D}) \right\}.$$

By definition of the domain of the operator A, we can conclude

$$D(A) = H_{A_0} \cap D(A_0^*) = \left\{ z \in H_0^1(\mathcal{D})\colon\, -\Delta z \in L^2(\mathcal{D}) \right\}.$$

It is well known that we can write equivalently $D(A) = H_0^1(\mathcal{D}) \cap H^2(\mathcal{D})$, see [85, Remark 6.2.2/3] and the references therein. □

Remark 2.47. *There exists another approach to introduce the Laplace operator with Dirichlet boundary condition. For more details, we refer to [30, Chapter 2, Section 3.3.A]. Let $\mathcal{D} \subset \mathbb{R}^n$ be a bounded domain with C^∞ boundary $\partial\mathcal{D}$. We consider the following Dirichlet boundary value problem:*

$$\begin{cases} -\Delta y(x) = z(x) & x \in \mathcal{D}, \\ \quad y(x) = 0 & x \in \partial\mathcal{D}. \end{cases} \tag{2.11}$$

Here, the definition of a solution (often called weak solution) is in a generalized sense as follows: First, we assume that $y \in C_0^\infty(\mathcal{D})$. Multiplying both sides of the equation $-\Delta y = z$ by a function $\phi \in C_0^\infty(\mathcal{D})$ and using partial integration, we get

$$\int_{\mathcal{D}} \nabla y(x)\cdot\nabla\phi(x)\,dx = \int_{\mathcal{D}} z(x)\phi(x)\,dx. \tag{2.12}$$

Obviously, the above equation remains valid for $y, \phi \in H_0^1(\mathcal{D})$ and $z \in L^2(\mathcal{D})$, which can be achieved using density results. We call $y \in H_0^1(\mathcal{D})$ a weak solution of (2.11) if equation (2.12) holds for every $\phi \in H_0^1(\mathcal{D})$. If $z \in L^2(\mathcal{D})$, then there exists a unique weak solution $y \in H_0^1(\mathcal{D})$ of (2.11). Moreover, we get $y \in H^2(\mathcal{D})$, see [15, Theorem 9.25]. Hence, we can introduce an operator $A\colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ given by

$$D(A) = H_0^1(\mathcal{D}) \cap H^2(\mathcal{D}), \quad Ay = -\Delta y$$

for every $y \in D(A)$. The operator A is self-adjoint and there exists a constant $c > 0$ such that the inequality $\langle Ay, y\rangle_{L^2(\mathcal{D})} \geq c\|y\|^2_{L^2(\mathcal{D})}$ holds for every $y \in D(A)$, see [89, Section 4.1].

We proceed with the Friedrichs extension $A\colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$.

Corollary 2.48. *The operator $A\colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ is closed and densely defined.*

Proof. Recall that the operator A is self-adjoint. Hence, we can conclude that the operator A is closed, see [85, Theorem 4.1.5. (c)]. Since $D(A_0) \subset D(A)$ and $D(A_0)$ is dense in $L^2(\mathcal{D})$, the set $D(A)$ is also dense in $L^2(\mathcal{D})$. □

Next, we show the existence of the resolvent operator.

Lemma 2.49. *If $\lambda > 0$, then the resolvent operator $R(\lambda; -A) = (\lambda + A)^{-1}$ exists and we have*

$$\|R(\lambda; -A)\|_{\mathcal{L}(L^2(\mathcal{D}))} \leq \frac{1}{\lambda}. \tag{2.13}$$

Proof. For the operator $\lambda + A$, we define the range by

$$\mathcal{R}(\lambda + A) = \{z \in L^2(\mathcal{D})\colon \text{there exists } y \in D(A) \text{ such that } (\lambda + A)y = z\}$$

and the null space by

$$\mathcal{N}(\lambda + A) = \{y \in D(A)\colon (\lambda + A)y = 0\}.$$

Since the operator A is self-adjoint, we have

$$L^2(\mathcal{D}) = \overline{\mathcal{R}(\lambda + A)} \oplus \mathcal{N}(\lambda + A),$$

where $\oplus$ denotes the direct sum, see [85, Lemma 4.1.6]. First, we determine the set $\mathcal{N}(\lambda + A)$. Using inequality (2.9), we get for every $y \in D(A)$

$$\|(\lambda + A)y\|^2_{L^2(\mathcal{D})} = \lambda^2\|y\|^2_{L^2(\mathcal{D})} + 2\lambda\langle Ay, y\rangle_{L^2(\mathcal{D})} + \|Ay\|^2_{L^2(\mathcal{D})} \geq \lambda^2\|y\|^2_{L^2(\mathcal{D})}. \tag{2.14}$$

Hence, we have for every $y \in \mathcal{N}(\lambda + A)$

$$0 = \|(\lambda + A)y\|^2_{L^2(\mathcal{D})} \geq \lambda^2\|y\|^2_{L^2(\mathcal{D})}.$$

As a consequence, the null space $\mathcal{N}(\lambda + A)$ contains only $0 \in D(A)$. Thus, we have $L^2(\mathcal{D}) = \overline{\mathcal{R}(\lambda + A)}$.

Next, we show that $L^2(\mathcal{D}) = \mathcal{R}(\lambda + A)$. If $z \in L^2(\mathcal{D})$, then there exists a sequence $(z_m)_{m\in\mathbb{N}} \subset \mathcal{R}(\lambda + A)$ such that $\lim_{m\to\infty} z_m = z$ in $L^2(\mathcal{D})$. Moreover, there exists $y_m \in D(A)$ such that $(\lambda + A)y_m = z_m$ for each $m \in \mathbb{N}$. By inequality (2.14), we have for each $m_1, m_2 \in \mathbb{N}$

$$\|y_{m_1} - y_{m_2}\|_{L^2(\mathcal{D})} \leq \frac{1}{\lambda}\|(\lambda + A)(y_{m_1} - y_{m_2})\|_{L^2(\mathcal{D})} = \frac{1}{\lambda}\|z_{m_1} - z_{m_2}\|_{L^2(\mathcal{D})}.$$

We obtain that the sequence $(y_m)_{m\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\mathcal{D})$ and hence, there exists $y \in L^2(\mathcal{D})$ such that $\lim_{m\to\infty} y_m = y$ in $L^2(\mathcal{D})$. Moreover, we get

$$\lim_{m\to\infty} Ay_m = \lim_{m\to\infty} [(\lambda + A)y_m - \lambda y_m] = \lim_{m\to\infty} [z_m - \lambda y_m] = z - \lambda y,$$

where the convergence is in $L^2(\mathcal{D})$. Since the operator A is closed, we can conclude that $y \in D(A)$ and $(\lambda + A)y = z$. Therefore, we have $L^2(\mathcal{D}) = \mathcal{R}(\lambda + A)$.

Next, we consider the operator $\lambda + A\colon D(A) \to L^2(\mathcal{D})$. Let $y_1, y_2 \in D(A)$ satisfy $(\lambda + A)y_1 = z$ and $(\lambda+A)y_2 = z$ for $z \in L^2(\mathcal{D})$. We obtain $y_1-y_2 \in \mathcal{N}(\lambda+A)$ and hence, we get $y_1 = y_2$. Therefore, the operator $\lambda + A$ is injective. Since $L^2(\mathcal{D}) = \mathcal{R}(\lambda + A)$, we infer that the inverse operator $(\lambda + A)^{-1}\colon L^2(\mathcal{D}) \to D(A)$ exists. Due to inequality (2.14), we get for every $y \in L^2(\mathcal{D})$

$$\|R(\lambda; -A)y\|_{L^2(\mathcal{D})} \leq \frac{1}{\lambda}\|(\lambda + A)(\lambda + A)^{-1}y\|_{L^2(\mathcal{D})} = \frac{1}{\lambda}\|y\|_{L^2(\mathcal{D})}$$

and hence, inequality (2.13) holds. □

We are now able to show the main result.

Theorem 2.50. *The operator $-A\colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ is the generator of an analytic semigroup of contractions $(e^{-At})_{t\geq 0}$.*

Proof. Due to Corollary 2.48, the operator $-A\colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ is closed and densely defined. By Lemma 2.49, the resolvent set $\rho(-A)$ contains $\mathbb{R}_0^+$ and for $\lambda > 0$

$$\|R(\lambda; -A)\|_{\mathcal{L}(L^2(\mathcal{D}))} \leq \frac{1}{\lambda}.$$

Thus, we can apply Theorem 2.11 with the result that the operator $-A$ is the generator of a C_0 semigroup of contractions $(e^{-At})_{t\geq 0}$. Due to Lemma 2.46, the operator $-A$ is self-adjoint and thus, the C_0 semigroup $(e^{-At})_{t\geq 0}$ is analytic due to Corollary 2.23. □

As a consequence of the previous theorem and the fact that the operator $-A$ is self-adjoint, there exists a constant $\theta > 0$ such that

$$\|e^{-At}\|_{\mathcal{L}(L^2(\mathcal{D}))} \leq e^{-\theta t}$$

for all $t \geq 0$, see [89, Theorem 7.2.8]. Hence, the assumptions of Remark 2.25 are satisfied with $M = 1$. Therefore, we can define fractional powers of the operator A denoted by A^α with $\alpha \in \mathbb{R}$ according to Section 2.3. Furthermore, if $\partial\mathcal{D}$ is a C^∞ boundary, then we can determine the domain $D(A^\alpha)$ for $\alpha \in (0,1)$ explicitly.

Theorem 2.51 (Theorem 1, [38])**.** *The domain of fractional powers of the operator A is given by*

(i) $D(A^\alpha) = H^{2\alpha}(\mathcal{D})$ *for* $\alpha \in \left(0, \frac{1}{4}\right)$,

(ii) $D(A^{1/4}) \subset H^{1/2}(\mathcal{D})$,

(iii) $D(A^\alpha) = H_0^{2\alpha}(\mathcal{D})$ *for* $\alpha \in \left(\frac{1}{4}, \frac{3}{4}\right)$,

(ii) $D(A^{3/4}) \subset H_0^{3/2}(\mathcal{D})$,

(v) $D(A^\alpha) = H_0^{2\alpha}(\mathcal{D})$ *for* $\alpha \in \left(\frac{3}{4}, 1\right)$.

Remark 2.52. *For general results on the Dirichlet realization of the Laplace operator defined on L^p-spaces, we refer to [89].*

The Neumann Realization of the Laplace Operator

Let $\mathcal{D} \subset \mathbb{R}^n$ be a bounded domain with C^∞ boundary $\partial\mathcal{D}$. We set

$$D(A_0) = \left\{ y \in C^\infty(\mathcal{D})\colon \frac{\partial y}{\partial \eta} = 0 \text{ on } \partial\mathcal{D} \right\},$$

where η is the C^∞ outward normal to $\partial\mathcal{D}$, i.e. the vector field $\eta = (\eta_1, ..., \eta_n)$ is the outward normal to $\partial\mathcal{D}$ with $\eta_1, ..., \eta_n \in C^\infty(\partial\mathcal{D})$. We define the operator $A_0\colon D(A_0) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ by

$$A_0 y = \Delta y \tag{2.15}$$

for every $y \in D(A_0)$.

Lemma 2.53. *The operator $A_0\colon D(A_0) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ defined by equation (2.15) is linear, densely defined and $\langle A_0 y, y\rangle_{L^2(\mathcal{D})} \leq 0$ for every $y \in D(A_0)$.*

Proof. Since $C(\mathcal{D})$ is dense in $L^2(\mathcal{D})$ and $C(\mathcal{D}) \subset D(A_0)$, we have that $D(A_0)$ is dense in $L^2(\mathcal{D})$. Obviously, the operator A_0 is linear. By Green's identity, we get for every $y \in D(A_0)$

$$\langle A_0 y, y\rangle_{L^2(\mathcal{D})} = \int_{\mathcal{D}} \Delta y(x) y(x)\, dx = -\int_{\mathcal{D}} \nabla y(x) \cdot \nabla y(x)\, dx = -\|\nabla y\|^2_{L^2(\mathcal{D})} \leq 0.$$

□

As a consequence, we get that the operator $-A_0$ is semi-bounded with $\langle -A_0 y, y\rangle_{L^2(\mathcal{D})} \geq 0$ for every $y \in D(A_0)$. Let $A_0^*\colon D(A_0^*) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ be the adjoint operator of A_0. Then we can apply Theorem 2.43 with the result that the Friedrichs extension of the operator A_0 exists. In the remaining part, we denote the Friedrichs extension of the operator A_0 by $A\colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$. We get the following properties, which can be derived similarly to Lemma 2.46.

Lemma 2.54 (Theorem 5.31 (ii), [48])**.** *The operator $A\colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ is self-adjoint and $\langle Ay, y\rangle_{L^2(\mathcal{D})} \leq 0$ for every $y \in D(A)$. Furthermore, we have*

$$D(A) = \left\{ y \in H^2(\mathcal{D})\colon \frac{\partial y}{\partial \eta} = 0 \text{ on } \partial\mathcal{D} \right\}, \quad Ay = \Delta y \text{ for every } y \in D(A).$$

Remark 2.55. *For the Neumann realization of the Laplace operator, the number 0 is an eigenvalue with constant functions as the related eigenfunctions, while 0 is an element of the resolvent set of the Dirichlet realization of the Laplace operator, see [48, Theorem 5.31]. This is the main difference of these operators.*

Remark 2.56. *Similarly to Remark 2.47, there exists another approach to introduce the Laplace operator with Neumann boundary conditions. For more details, we refer to [30, Chapter 2, Section 3.3.C]. Let $\mathcal{D} \subset \mathbb{R}^n$ be a bounded domain with C^∞ boundary $\partial\mathcal{D}$. We consider the following Neumann boundary value problem:*

$$\begin{cases} \Delta y(x) = z(x) & x \in \mathcal{D}, \\ \dfrac{\partial y(x)}{\partial \eta} = 0 & x \in \partial\mathcal{D}, \end{cases} \tag{2.16}$$

where η is the outward normal to $\partial\mathcal{D}$. Here, the definition of a solution (often called weak solution) is in a generalized sense as follows: First, we assume that $y \in C^\infty(\overline{\mathcal{D}})$. Multiplying both sides of the equation $\Delta y = z$ by a function $\phi \in C^\infty(\overline{\mathcal{D}})$ and using Green's identity, we get

$$\int_{\mathcal{D}} \nabla y(x) \cdot \nabla \phi(x)\, dx = \int_{\mathcal{D}} z(x)\phi(x)\, dx. \tag{2.17}$$

Obviously, the above equation remains still valid for $y, \phi \in H^1(\mathcal{D})$ and $z \in L^2(\mathcal{D})$, which can be achieved using density results. We call $y \in H^1(\mathcal{D})$ a weak solution of (2.16) if equation (2.17) holds for every $\phi \in H^1(\mathcal{D})$. A weak solution $y \in H^1(\mathcal{D})$ of (2.16) exists and is unique up to a constant if and only if $z \in L^2(\mathcal{D})$ satisfies

$$\int_{\mathcal{D}} z(x)\, dx = 0.$$

Moreover, one can conclude that $y \in H^2(\mathcal{D})$, see [15, Theorem 9.26]. Hence, we can introduce an operator $A\colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ given by

$$D(A) = \left\{ y \in H^2(\mathcal{D})\colon \frac{\partial y}{\partial \eta} = 0 \text{ on } \partial\mathcal{D} \right\}, \quad Ay = \Delta y$$

for every $y \in D(A)$. The operator A is self-adjoint and $\langle Ay, y\rangle_{L^2(\mathcal{D})} \leq 0$ for every $y \in D(A)$, see [89, Section 4.2 and Lemma 1.6.1].

Similarly to the Dirichlet realization of the Laplace operator, we can show that the operator A is closed and densely defined. Moreover, if $\lambda > 0$, then the resolvent operator $R(\lambda; A)$ exists and we have

$$\|R(\lambda; A)\|_{\mathcal{L}(L^2(\mathcal{D}))} \leq \frac{1}{\lambda}.$$

Therefore, we get the following generation theorem, which can be obtained similarly to Theorem 2.50.

Theorem 2.57. *The operator $A\colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ is the generator of an analytic semigroup of contractions $(e^{At})_{t\geq 0}$.*

By Remark 2.55, the number 0 is an eigenvalue of the operator A and hence, we have $0 \notin \rho(A)$. As a consequence, we can not apply Theorem 2.35 directly. Here, we can easily overcome this problem as follows: Let $\lambda > 0$. Due to the previous theorem, the operator $A - \lambda$ is still the generator of an analytic semigroup given by $(e^{-\lambda t} e^{At})_{t\geq 0}$, see [70, Chapter 3, Corollary 2.2]. Hence, the operator $A - \lambda$ satisfies the assumptions of Remark 2.25 with $M = 1$ and $\theta = \lambda$. Therefore, we can define fractional powers of the operator $\lambda - A$ denoted by $(\lambda - A)^\alpha$ with $\alpha \in \mathbb{R}$ according to Section 2.3. Also note that we can apply Theorem 2.35 since $0 \in \rho(A - \lambda)$. Furthermore, we can determine the domain $D((\lambda - A)^\alpha)$ for $\alpha \in (0, 1)$ explicitly.

Theorem 2.58 (Theorem 2, [38])**.** *The domain of fractional powers of the operator $\lambda - A$ is given by*

(i) $D((\lambda - A)^\alpha) = H^{2\alpha}(\mathcal{D})$ *for* $\alpha \in \left(0, \frac{3}{4}\right)$,

(ii) $D((\lambda - A)^{3/4}) \subset H^{3/2}(\mathcal{D})$,

(iii) $D((\lambda - A)^\alpha) = \left\{ y \in H^{2\alpha}(\mathcal{D})\colon \frac{\partial y}{\partial \eta} = 0 \text{ on } \partial\mathcal{D} \right\}$ *for* $\alpha \in \left(\frac{3}{4}, 1\right)$.

2.5.2. The Stokes Operator

Let $\mathcal{D} \subset \mathbb{R}^n$ be a bounded domain with C^2 boundary $\partial\mathcal{D}$ and let

$$C_{0,\sigma}^\infty = \{y \in (C_0^\infty(\mathcal{D}))^n \colon \operatorname{div} y = 0 \text{ in } \mathcal{D}\}.$$

We introduce the following common spaces:

$$\begin{aligned} H &= \text{Completion of } C_{0,\sigma}^\infty \text{ in } (L^2(\mathcal{D}))^n \\ &= \left\{ y \in (L^2(\mathcal{D}))^n \colon \operatorname{div} y = 0 \text{ in } \mathcal{D}, y \cdot \eta = 0 \text{ on } \partial\mathcal{D} \right\}, \\ V &= \text{Completion of } C_{0,\sigma}^\infty \text{ in } \left(H^1(\mathcal{D})\right)^n \\ &= \left\{ y \in \left(H_0^1(\mathcal{D})\right)^n \colon \operatorname{div} y = 0 \text{ in } \mathcal{D} \right\}, \end{aligned}$$

where η denotes the unit outward normal to $\partial\mathcal{D}$. The space H equipped with the inner product

$$\langle y, z\rangle_H = \langle y, z\rangle_{(L^2(\mathcal{D}))^n} = \int_{\mathcal{D}} \sum_{i=1}^n y_i(x) z_i(x)\, dx$$

for every $y = (y_1, ..., y_n), z = (z_1, ..., z_n) \in H$ becomes a Hilbert space. For all $x = (x_1, ..., x_n) \in \mathcal{D}$, we denote $D^j = \frac{\partial^{|j|}}{\partial x_1^{j_1} \cdots \partial x_n^{j_n}}$ with $|j| = \sum_{i=1}^n j_i$. We set $D^j y = (D^j y_1, ..., D^j y_n)$ with $|j| \leq 1$ for every $y = (y_1, ..., y_n) \in V$. Then the space V equipped with the inner product

$$\langle y, z\rangle_V = \sum_{|j|\leq 1} \langle D^j y, D^j z\rangle_{(L^2(\mathcal{D}))^n}$$

for every $y, z \in V$ becomes a Hilbert space. The norms in H and V are denoted by $\|\cdot\|_H$ and $\|\cdot\|_V$, respectively. Moreover, we get the orthogonal Helmholtz decomposition

$$(L^2(\mathcal{D}))^n = H \oplus \{\nabla y : y \in H^1(\mathcal{D})\},$$

where $\oplus$ denotes the direct sum. Then there exists an orthogonal projection $\Pi\colon (L^2(\mathcal{D}))^n \to H$, see [39]. We set $D(A_0) = C^\infty_{0,\sigma}$ and we define the operator $A_0\colon D(A_0) \subset H \to H$ by

$$A_0 y = -\Pi\Delta y \tag{2.18}$$

for every $y \in D(A_0)$, where Δ is the Laplace operator defined for vector functions in the sense that $\Delta y = (\Delta y_1, ..., \Delta y_n)$.

Lemma 2.59. *The operator $A_0\colon D(A_0) \subset H \to H$ given by (2.18) is semi-bounded with $\langle A_0 y, y\rangle_H \geq c\|y\|_H^2$ for every $y \in D(A_0)$, where $c > 0$ is a constant.*

Proof. By definition of the space H, we get that $D(A_0)$ is dense in H. Since the operator $\Pi\colon (L^2(\mathcal{D}))^n \to H$ is an orthogonal projection, the operator Π is linear, self-adjoint and we have $\Pi y = y$ for every $y \in H$. Hence, the operator A_0 is linear and we get for every $y \in D(A_0)$

$$\langle A_0 y, y\rangle_H = -\langle \Pi\Delta y, y\rangle_H = -\langle \Delta y, y\rangle_H.$$

The remaining part of the proof can be obtained similarly to Lemma 2.45. □

Let $A_0^*\colon D(A_0^*) \subset H \to H$ be the adjoint operator of A_0. As a consequence of the previous lemma, we can apply Theorem 2.43 with the result that the Friedrichs extension of the operator A_0 exists. In the remaining part, we denote the Friedrichs extension of the operator A_0 by $A\colon D(A) \subset H \to H$. We get the following properties, which can be derived similarly to Lemma 2.46.

Lemma 2.60. *The operator $A\colon D(A) \subset H \to H$ is self-adjoint and we have for every $y \in D(A)$*

$$\langle Ay, y\rangle_H \geq c\|y\|_H^2,$$

where the constant $c > 0$ arises from Lemma 2.59. Furthermore, we have

$$D(A) = (H^2(\mathcal{D}))^n \cap V, \quad Ay = -\Pi\Delta y \text{ for every } y \in D(A).$$

Similarly to the Dirichlet realization of the Laplace operator, we can show that the operator A is closed and densely defined. Moreover, if $\lambda > 0$, then the resolvent operator $R(\lambda; -A)$ exists and we have

$$\|R(\lambda; -A)\|_{\mathcal{L}(H)} \leq \frac{1}{\lambda}.$$

Therefore, we get the following generation theorem, which can be obtained similarly to Theorem 2.50.

Theorem 2.61. *The operator $-A\colon D(A) \subset H \to H$ is the generator of an analytic semigroup of contractions $(e^{-At})_{t\geq 0}$.*

Due to the previous Theorem and the fact that the operator $-A$ is self-adjoint, there exists a constant $\theta > 0$ such that

$$\|e^{-At}\|_{\mathcal{L}(H)} \leq e^{-\theta t}$$

for all $t \geq 0$, see [89, Remark 7.2.1]. Hence, the assumptions of Remark 2.25 are satisfied with $M = 1$. Therefore, we can define fractional powers of the operator A denoted by A^α with $\alpha \in \mathbb{R}$ according to Section 2.3. Furthermore, if the boundary $\partial\mathcal{D}$ is a C^∞ boundary, then we can determine the domain $D(A^\alpha)$ for $\alpha \in (0,1)$ explicitly. Let the operator $A_D\colon D(A_D) \subset (L^2(\mathcal{D}))^n \to (L^2(\mathcal{D}))^n$ be the Dirichlet realization of the Laplace operator, which we can introduce similarly to Section 2.5.1. Then, we get the following presentation.

Theorem 2.62 ([37, 88]). *For all $\alpha \in (0,1)$, we have*

$$D(A^\alpha) = D(A_D^\alpha) \cap H.$$

As a consequence of the previous theorem and Theorem 2.51, we can determine the domain $D(A^\alpha)$ for $\alpha \in (0,1)$ explicitly.

Corollary 2.63. *The domain of fractional powers of the operator A is given by*

(i) $D(A^\alpha) = \left(H^{2\alpha}(\mathcal{D})\right)^n \cap H$ *for* $0 < \alpha < \frac{1}{4}$,

(ii) $D(A^{1/4}) \subset \left(H^{1/2}(\mathcal{D})\right)^n \cap H$,

(iii) $D(A^\alpha) = \left(H_0^{2\alpha}(\mathcal{D})\right)^n \cap H$ *for* $\frac{1}{4} < \alpha < \frac{3}{4}$,

(ii) $D(A^{3/4}) \subset \left(H_0^{3/2}(\mathcal{D})\right)^n \cap H$,

(v) $D(A^\alpha) = \left(H_0^{2\alpha}(\mathcal{D})\right)^n \cap H$ *for* $\frac{3}{4} < \alpha < 1$.

Remark 2.64. *For general results on the Stokes operator defined on L^p-spaces, we refer to [43].*

Chapter 3

Stochastic Calculus

This chapter is devoted to SPDEs both of forward and of backward type. Forward SPDEs driven by Lévy noise are often stated as stochastic evolution equations on infinite dimensional spaces, see [71]. The theory presented here extends results well known for the case of Wiener noise, see [23, 42, 73]. For stochastic ordinary differential equations with Lévy noise, we refer to [20, 74]. Similarly, backward SPDEs can also be stated as stochastic evolution equations, see [1, 52]. Existence and uniqueness results of these backward equations are mainly based on a martingale representation theorem. Since these representation formulas are not available for infinite dimensional Lévy processes, we have to restrict to the case of to backward SPDEs driven by Wiener noise.

We start with basic notions and definitions concerning random variables and stochastic processes on separable Hilbert spaces. Afterwards, we give an overview on properties of infinite dimensional Lévy processes. For a certain class of Lévy processes, we introduce the stochastic integral and we state basic results, which we use in the following chapters. This allows us to define solutions to SPDEs both of forward and of backward type. The equations considered here are mainly motivated by control problems we discuss in the following chapters. For forward SPDEs, we will figure out that the mild solution is useful to involve nonhomogeneous boundary conditions. In control theory, backward SPDEs characterizes the dynamics of the adjoint equation, which is closely related to the corresponding forward equation. For that reason, we will also consider mild solutions of backward SPDEs. Finally, we state different concepts of solutions and we will give a relationship between these solutions, where we mainly use results shown in [1, 23, 53, 71].

Throughout this chapter, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given probability space. We always assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is complete, i.e. $A \in \mathcal{F}$, $B \subset A$ and $\mathbb{P}(A) = 0$ imply $B \in \mathcal{F}$.

3.1. Preliminaries

Let $\mathcal{U}$ be a separable Hilbert space and let $\mathcal{B}(\mathcal{U})$ denote its Borel σ-field. An $\mathcal{U}$-valued random variable or a random variable with values in $\mathcal{U}$ is any measurable mapping $X \colon \Omega \to \mathcal{U}$, i.e. X maps Ω into $\mathcal{U}$ such that $\{X \in A\} = \{\omega \in \Omega \colon X(\omega) \in A\} \in \mathcal{F}$ for arbitrary $A \in \mathcal{B}(\mathcal{U})$. We denote the law or the distribution of X by $\mathscr{L}(X)(A) = \mathbb{P}(\omega \in \Omega \colon X(\omega) \in A)$ for all $A \in \mathcal{B}(\mathcal{U})$. For an $\mathcal{U}$-valued random variable X, one can introduce its expected value

$$\mathbb{E}[X] = \int_\Omega X(\omega)\, \mathbb{P}(d\omega)$$

in the sense of a Bochner integral as introduced in Section B. The expected value is well defined if

$$\mathbb{E}\, \|X\|_{\mathcal{U}} = \int_\Omega \|X(\omega)\|_{\mathcal{U}}\, \mathbb{P}(d\omega) < \infty. \tag{3.1}$$

A random variable satisfying condition (3.1) is integrable. If $\mathbb{E}\, \|X\|_{\mathcal{U}}^p < \infty$ with $1 \leq p < \infty$, then X is p-integrable. The space of p-integrable random variables with values in $\mathcal{U}$ is denoted by $L^p(\Omega; \mathcal{U})$.

Let $\mathcal{I}$ be a time interval given by all nonnegative real numbers $\mathbb{R}^+$ or a finite interval $[0, T]$ with $T > 0$. A family $(X(t))_{t \in \mathcal{I}}$ of $\mathcal{U}$-valued random variables is called $\mathcal{U}$-valued stochastic process or stochastic process with values in $\mathcal{U}$. The stochastic process $(X(t))_{t\in I}$ is p-integrable if for all $t \in \mathcal{I}$, the random variable $X(t)$ is p-integrable. We set $X(t)(\omega) = X(t, \omega)$ for all $t \in \mathcal{I}$ and $\omega \in \Omega$. The function $X(\cdot, \omega)\colon \mathcal{I} \to \mathcal{U}$ is called trajectory of $(X(t))_{t\in\mathcal{I}}$.

Definition 3.1. *Let $(X(t))_{t\in\mathcal{I}}$ be a stochastic process with values in $\mathcal{U}$. An $\mathcal{U}$-valued stochastic process $(Y(t))_{t\in\mathcal{I}}$ is a **modification** of $(X(t))_{t\in\mathcal{I}}$ if for all $t \in \mathcal{I}$*

$$\mathbb{P}(X(t) = Y(t)) = 1.$$

Next, we introduce different continuity properties of stochastic processes.

Definition 3.2. *An $\mathcal{U}$-valued stochastic process $(X(t))_{t\in\mathcal{I}}$ is*

- ***stochastically continuous** if for all $t \in \mathcal{I}$ and $\varepsilon > 0$*

$$\lim_{s\to t} \mathbb{P}(\|X(t) - X(s)\|_{\mathcal{U}} > \varepsilon) = 0;$$

- ***continuous** (with probability 1) if its trajectories $X(\cdot, \omega)$ are continuous $\mathbb{P}$-almost surely.*

Under additional requirements, we can define the mean square continuity.

Definition 3.3. *An $\mathcal{U}$-valued square integrable stochastic process $(X(t))_{t\in\mathcal{I}}$ is **mean square continuous** or **continuous in mean square** if for all $t \in \mathcal{I}$*

$$\lim_{s\to t} \mathbb{E}\|X(t) - X(s)\|_{\mathcal{U}}^2 = 0.$$

We have the following relationships between these various types of continuity, which are well known for real valued stochastic processes, see [55, 66]. One can easily adapt these results to Hilbert space valued stochastic processes.

Proposition 3.4. *We have the following implications:*

(i) Every continuous stochastic process is stochastically continuous.

(ii) Every mean square continuous stochastic process is stochastically continuous.

In general, there is no relation between the continuity with probability 1 and the continuity in mean square. Furthermore, we obtain that the stochastically continuity is the weakest notion among the continuity properties introduced above. However, to require the stochastically continuity is often sufficient. Obviously, stochastic processes with jumps occurring in the trajectories are not continuous. Therefore, we introduce the concept of càdlàg trajectories.

Definition 3.5. *A stochastic process $(X(t))_{t\in\mathcal{I}}$ taking values in $\mathcal{U}$ is **càdlàg** (continu à droite et limites à gauche) if $\mathbb{P}$-a.s.*

- *$(X(t))_{t\in\mathcal{I}}$ is right-continuous, i.e. $X(t+) = \lim_{s\downarrow t} X(s) = X(t)$ for all $t \in \mathcal{I}$ and*
- *$(X(t))_{t\in\mathcal{I}}$ has left limits, i.e. $X(t-) = \lim_{s\uparrow t} X(s)$ exists for all $t \in \mathcal{I}$.*

Next, we introduce different properties of measurability for stochastic processes. Let the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ be equipped with a filtration $(\mathcal{F}_t)_{t\in\mathcal{I}}$, i.e. $(\mathcal{F}_t)_{t\in\mathcal{I}}$ is a family of increasing σ-fields.

Definition 3.6. *The filtration $(\mathcal{F}_t)_{t\in\mathcal{I}}$ is said to be **normal** if $\mathcal{F}_0$ contains all $A \in \mathcal{F}$ such that $\mathbb{P}(A) = 0$ and we have $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ for all $t \in \mathcal{I}$.*

Definition 3.7. *An $\mathcal{U}$-valued stochastic process $(X(t))_{t\in\mathcal{I}}$ is **$\mathcal{F}_t$-adapted** if for all $t \in \mathcal{I}$, the random variable $X(t)$ is $\mathcal{F}_t$-measurable.*

Let $\mathcal{P}_\mathcal{I}$ denote the smallest σ-field of subsets of $\mathcal{I} \times \Omega$ containing all sets of the form

$$(s,t] \times A \text{ with } s,t \in \mathcal{I}, s < t, A \in \mathcal{F}_s \quad \text{and} \quad \{0\} \times A \text{ with } A \in \mathcal{F}_0.$$

We have the following definition.

Definition 3.8. *An $\mathcal{U}$-valued stochastic process $(X(t))_{t\in\mathcal{I}}$ is called **predictable** if it is a measurable mapping from $(\mathcal{I} \times \Omega, \mathcal{P}_\mathcal{I})$ into $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$.*

Every predictable stochastic process is $\mathcal{F}_t$-adapted. The converse is in general not true. However, the following result is useful to conclude that a stochastic process has a predictable modification.

Proposition 3.9 (Proposition 3.7 (ii),[23])**.** *Assume that the stochastic process $(X(t))_{t\in[0,T]}$ is $\mathcal{F}_t$-adapted and stochastically continuous. Then the process $(X(t))_{t\in[0,T]}$ has a predictable modification.*

Next, we define stopping times, which are necessary for the definition of local solutions to SPDEs.

Definition 3.10. *A random variable $\tau\colon \Omega \to [0,\infty]$ is a **stopping time** (with respect to the filtration $(\mathcal{F}_t)_{t\in\mathcal{I}}$) if for all $t \in \mathcal{I}$*

$$\{\tau \le t\} = \{\omega \in \Omega\colon \tau(\omega) \le t\} \in \mathcal{F}_t.$$

We have the following basic properties, which follow immediately from the previous definition.

Lemma 3.11 (Lemma 9, [82])**.** *Let τ and ρ be stopping times and let $(\tau_m)_{m\in\mathbb{N}}$ be a sequence of stopping times. Then*

(i) $\tau \wedge \rho = \min\{\tau, \rho\}$ and $\tau \vee \rho = \max\{\tau, \rho\}$ are stopping times;

(ii) the limit $\tau = \lim_{m\to\infty} \tau_m$ is a stopping time if $(\tau_m)_{m\in\mathbb{N}}$ is increasing or decreasing.

Lemma 3.12 ([74, 78, 82])**.** *Let the Filtration $(\mathcal{F}_t)_{t\in\mathcal{I}}$ be normal and let $(X(t))_{t\in\mathcal{I}}$ be an $\mathcal{F}_t$-adapted càdlàg process with values in $\mathbb{R}^n$, $n \in \mathbb{N}$. If $\Gamma \in \mathcal{B}(\mathbb{R}^n)$ is open, then*

$$\tau = \inf\{t > 0\colon t \in \mathcal{I}, X(t) \in \Gamma\}$$

is a stopping time. We employ the standard convention that $\inf\{\emptyset\} = +\infty$.

Theorem 3.13. *Let the Filtration $(\mathcal{F}_t)_{t\in\mathcal{I}}$ be normal and let $(X(t))_{t\in\mathcal{I}}$ be an $\mathcal{F}_t$-adapted càdlàg process with values in $\mathcal{U}$. If $c \ge 0$, then*

$$\tau = \inf\{t > 0\colon t \in \mathcal{I}, \|X(t)\|_\mathcal{U} > c\}$$

is a stopping time. We employ the standard convention that $\inf\{\emptyset\} = +\infty$.

Proof. Obviously, the stochastic process $(\|X(t)\|_\mathcal{U})_{t\in\mathcal{I}}$ is $\mathcal{F}_t$-adapted, càdlàg and takes values in $\mathbb{R}$. By Lemma 3.12 with $n = 1$, the claim follows immediately. □

Remark 3.14. *If $\mathcal{I} = [0,T]$, then by Lemma 3.11 (i) and Theorem 3.13, the random variable*

$$\tau = \inf\{t \in (0,T)\colon \|X(t)\|_\mathcal{U} > c\} \wedge T$$

is a stopping time.

Finally, we introduce martingales on Hilbert spaces. Therefor, we need the concept of the conditional expectation. The existence and uniqueness is provided by the following result.

Proposition 3.15 (Proposition 3.13, [71]). *Let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$ and let X be an $\mathcal{U}$-valued integrable random variable. Then, up to a set of $\mathbb{P}$-measure 0, there is a unique integrable $\mathcal{G}$-measurable random variable $\mathbb{E}[X|\mathcal{G}]$ with values in $\mathcal{U}$ such that for all $A \in \mathcal{G}$*

$$\int_A X(\omega)\,\mathbb{P}(d\omega) = \int_A \mathbb{E}[X|\mathcal{G}](\omega)\,\mathbb{P}(d\omega).$$

In the previous proposition, we call $\mathbb{E}[X|\mathcal{G}]$ the conditional expectation of X given $\mathcal{G}$. We have the following basic properties, which are well known for real-valued random variables, see [12].

Proposition 3.16 (Proposition 3.15, [71]). *Let X, Y be $\mathcal{U}$-valued integrable random variable and let $a, b \in \mathbb{R}$. Assume that $\mathcal{G}$ is a sub-σ-field of $\mathcal{F}$. Then the following properties hold $\mathbb{P}$-almost surely:*

(i) $\mathbb{E}[a\,X + b\,Y|\mathcal{G}] = a\,\mathbb{E}[X|\mathcal{G}] + b\,\mathbb{E}[Y|\mathcal{G}]$;

(ii) if $K \in \mathcal{L}(\mathcal{U};\mathcal{H})$ where $\mathcal{H}$ is another separable Hilbert space, then $\mathbb{E}[K\,X|\mathcal{G}] = K\,\mathbb{E}[X|\mathcal{G}]$;

(iii) if X is $\mathcal{G}$-measurable and ξ is a real-valued integrable random variables such that $\xi\,X$ is integrable, then $\mathbb{E}[\xi\,X|\mathcal{G}] = X\,\mathbb{E}[\xi|\mathcal{G}]$;

(iii) if $\mathcal{V}$ is a sub-σ-field of $\mathcal{G}$, then $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\,|\mathcal{V}] = E[X|\mathcal{V}]$;

(iv) if X is independent of $\mathcal{G}$, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$;

(v) if $f\colon \mathbb{R} \to \mathbb{R}$ is a convex function such that the random variable $f(\|X\|_{\mathcal{U}})$ is integrable, then we get $f(\|\mathbb{E}[X|\mathcal{G}]\|_{\mathcal{U}}) \le \mathbb{E}[f(\|X\|_{\mathcal{U}})|\mathcal{G}]$;

(vi) if the σ-fields $(\mathcal{G}_m)_{m\in\mathbb{N}}$ is an increasing sequence satisfying $\mathcal{G} = \sigma(\mathcal{G}_m\colon m \in \mathbb{N})$, then we obtain $\mathbb{E}[X|\mathcal{G}] = \lim_{m\to\infty} \mathbb{E}[X|\mathcal{G}_m]$.

We proceed with the definition of Hilbert space valued martingales.

Definition 3.17. *An $\mathcal{F}_t$-adapted integrable stochastic process $(M(t))_{t\in\mathcal{I}}$ with values in $\mathcal{U}$ is a* ***martingale*** *(with respect to the filtration $(\mathcal{F}_t)_{t\in\mathcal{I}}$) if for all $s, t \in \mathcal{I}$ with $s \le t$ and $\mathbb{P}$-a.s.*

$$\mathbb{E}[M(t)|\mathcal{F}_s] = M(s).$$

Proposition 3.18 (Proposition 3.25, [71]). *Let $(X(t))_{t\in\mathcal{I}}$ be an $\mathcal{F}_t$-adapted integrable stochastic process with values in $\mathcal{U}$. Assume that $X(t) - X(s)$ is independent of $\mathcal{F}_s$ for all $s, t \in \mathcal{I}$ with $t > s$. Then the process $(M(t))_{t\in\mathcal{I}}$ given by $M(t) = X(t) - \mathbb{E}[X(t)]$ for all $t \in \mathcal{I}$ and $\mathbb{P}$-almost surely is a martingale.*

Theorem 3.19 (Theorem 3.41, [71]). *Let $(M(t))_{t\ge 0}$ be a stochastically continuous square integrable martingale with values in $\mathcal{U}$. Then $(M(t))_{t\ge 0}$ has a càdlàg modification (still denoted by $(M(t))_{t\ge 0}$) satisfying for all $T \ge 0$ and all $r > 0$*

$$\mathbb{P}\left(\sup_{t\in[0,T]} \|M(t)\|_{\mathcal{U}} \ge r\right) \le \frac{1}{r^2}\,\mathbb{E}\|M(T)\|_{\mathcal{U}}^2.$$

Moreover, we have for all $T > 0$ and all $k \in (0, 2)$

$$\mathbb{E}\sup_{t\in[0,T]} \|M(t)\|_{\mathcal{U}}^k \le \frac{2}{2-k}\,\left(\mathbb{E}\|M(T)\|_{\mathcal{U}}^2\right)^{k/2}.$$

Let $\mathcal{M}^2(\mathcal{U})$ contain all stochastically continuous square integrable martingales $(M(t))_{t\ge 0}$ with values in $\mathcal{U}$. By Theorem 3.19, we can always assume that the elements of $\mathcal{M}^2(\mathcal{U})$ are càdlàg. Moreover, we have the following result.

Theorem 3.20. *If $M \in \mathcal{M}^2(\mathcal{U})$, then there exists a unique increasing predictable process $(\langle M\rangle_t)_{t\geq 0}$ such that $\langle M\rangle_0 = 0$ and $(\|M(t)\|^2_{\mathcal{U}} - \langle M\rangle_t)_{t\geq 0}$ is a real-valued martingale.*

Proof. The claim follows by applying the Doob-Meyer decomposition theorem to the real valued process $(\|M(t)\|^2_{\mathcal{U}})_{t\geq 0}$, see [71]. □

In the previous theorem, the process $(\langle M\rangle_t)_{t\geq 0}$ is called angle bracket or predictable variation process.

Theorem 3.21. *Let $M \in \mathcal{M}^2(\mathcal{U})$ such that $M(0) = 0$. Then we have for all $T \geq 0$*

$$\mathbb{E} \sup_{t\in[0,T]} \|M(t)\|^2_{\mathcal{U}} \leq 4\,\mathbb{E}\langle M\rangle_T = 4\,\mathbb{E}\|M(T)\|^2_{\mathcal{U}}.$$

Proof. One can deduce the assertion from [63, Theorem 20.6]. □

Finally, for $M \in \mathcal{M}^2(\mathcal{U})$, we introduce the operator angle bracket process denoted by $(\langle\langle M\rangle\rangle_t)_{t\geq 0}$. Let $\mathcal{L}_1(\mathcal{U})$ be the space of all nuclear operators on $\mathcal{U}$ equipped with the nuclear norm and let $\mathcal{L}_1^+(\mathcal{U})$ denote the subspace of $\mathcal{L}_1(\mathcal{U})$ containing all self-adjoint nonnegative nuclear operators. For more details, see Appendix C. For $x, y \in \mathcal{U}$, we define the operator $x \otimes y\colon \mathcal{U} \to \mathcal{U}$ by $x \otimes y(z) = \langle x, z\rangle_{\mathcal{U}} y$ for every $z \in \mathcal{U}$. Then, we have $x \otimes y \in \mathcal{L}_1(\mathcal{U})$ with $\|x \otimes y\|_{\mathcal{L}_1(\mathcal{U})} = \|x\|_{\mathcal{U}}\|y\|_{\mathcal{U}}$. If $M \in \mathcal{M}^2(\mathcal{U})$, then the process $(M(t) \otimes M(t))_{t\geq 0}$ is an $\mathcal{L}_1(\mathcal{U})$-valued right-continuous process such that for all $t \geq 0$

$$\mathbb{E}\|M(t) \otimes M(t)\|_{\mathcal{L}_1(\mathcal{U})} = \mathbb{E}\|M(t)\|^2_{\mathcal{U}}.$$

We have the following result.

Theorem 3.22 (Theorem 8.2, [71])**.** *Let $M \in \mathcal{M}^2(\mathcal{U})$. Then there exists a unique right-continuous increasing predictable process $(\langle\langle M\rangle\rangle_t)_{t\geq 0}$ with values in $\mathcal{L}_1^+(\mathcal{U})$ such that $\langle\langle M\rangle\rangle_0 = 0$ and $(M(t)\otimes M(t) - \langle\langle M\rangle\rangle_t)_{t\geq 0}$ is an $\mathcal{L}_1(\mathcal{U})$-valued martingale. Moreover, there exists a predictable process $(Q(t))_{t\geq 0}$ with values in $\mathcal{L}_1^+(\mathcal{U})$ such that for all $t \geq 0$ and $\mathbb{P}$-a.s.*

$$\langle\langle M\rangle\rangle_t = \int_0^t Q(s)\, d\langle M\rangle_s.$$

3.2. Lévy Processes

In this section, we give an introduction to Lévy processes with values in a separable Hilbert space $\mathcal{U}$. For a comparison with finite dimensional Lévy processes, we refer to [2]. We assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is equipped with a normal filtration $(\mathcal{F}_t)_{t\in\mathcal{I}}$. Let us start with a formal definition.

Definition 3.23. *A stochastic process $(L(t))_{t\geq 0}$ taking values in $\mathcal{U}$ is called a* ***Lévy process*** *if*

- *$\mathbb{P}$-a.s. $L(0) = 0$;*
- *$(L(t))_{t\geq 0}$ has independent and time-homogeneous increments, i.e. for $0 \leq t_0 < t_1 < ... < t_m$, the random variables $L(t_1) - L(t_0), L(t_2) - L(t_1), ..., L(t_m) - L(t_{m-1})$ are independent and for any $s, t \geq 0$ with $s < t$, the law $\mathscr{L}(L(t) - L(s))$ depends only on the difference $t - s$;*
- *$(L(t))_{t\geq 0}$ is stochastically continuous.*

For an $\mathcal{U}$-valued Lévy process $(L(t))_{t\geq 0}$, let μ_t be the law of the random variable $L(t)$ for $t \geq 0$. Then

(i) $\mu_0 = \delta_0$ and $\mu_{t+s} = \mu_t * \mu_s$ for all $s, t \geq 0$;

(ii) $\lim_{t\downarrow 0} \mu_t(\{x \in \mathcal{U} : \|x\|_{\mathcal{U}} < r\}) = 1$ for every $r > 0$.

Here, the measure δ_0 is the unit measure concentrated at $\{0\}$ and $\mu_t * \mu_s$ denotes the convolution of the measures μ_t and μ_s. The family of measures $(\mu_t)_{t\geq 0}$ is called the convolution semigroup of measures. For more details, see [71].

Theorem 3.24 (Theorem 4.3, [71])**.** *Every Lévy process $(L(t))_{t\geq 0}$ has a càdlàg modification.*

Given a càdlàg process $(L(t))_{t\geq 0}$, the process of jumps $(\Delta L(t))_{t\geq 0}$ is defined by $\Delta L(t) = L(t) - L(t-)$ for all $t \geq 0$ and $\mathbb{P}$-almost surely.

Theorem 3.25 (Theorem 4.4, [71])**.** *Assume that $(L(t))_{t\geq 0}$ is an $\mathcal{U}$-valued càdlàg Lévy process with bounded jumps, i.e. there exists $c > 0$ such that $\|\Delta L(t)\|_{\mathcal{U}} \leq c$ for all $t \geq 0$ and $\mathbb{P}$-almost surely. Then we have for all $\beta > 0$ and all $t \geq 0$*

$$\mathbb{E}e^{\beta \|L(t)\|_{\mathcal{U}}} < \infty.$$

For every $\mathcal{U}$-valued càdlàg Lévy process $(L(t))_{t\geq 0}$ with bounded jumps, we have $\mathbb{E}\|L(t)\|^n_{\mathcal{U}} < \infty$ for each $n \in \mathbb{N}$ and all $t \geq 0$ resulting from the previous theorem.

3.2.1. Examples

Compound Poisson Processes

We start with a definition of a Poisson process, which is an increasing Lévy process taking values in $\mathbb{Z}^+$ with jumps of size 1.

Definition 3.26. *A **Poisson process** $(N(t))_{t\geq 0}$ with intensity $a \in (0, \infty)$ is a Lévy process with values in $\mathbb{Z}^+$ such that the random variable $N(t)$ has a Poisson distribution with parameter at for all $t \geq 0$, i.e. for all $t \geq 0$ and each $k \in \mathbb{Z}^+$*

$$\mathbb{P}(N(t) = k) = \frac{(at)^k}{k!} e^{-at}.$$

Recall that a random variable Z with values in $\mathbb{R}^+$ is exponentially distributed with parameter $a \in (0, \infty)$ if $\mathbb{P}(Z > t) = e^{-at}$ for all $t \geq 0$. The following proposition provides the construction of a Poisson process based on a sequence of independent exponentially distributed random variables.

Proposition 3.27 (Proposition 4.9 (i) and (ii), [42])**.** *The process $(N(t))_{t\geq 0}$ is a Poisson process with intensity a if and only if there exists a sequence $(Z_m)_{m\in\mathbb{N}}$ of independent exponentially distributed random variables with parameter a such that for all $t \geq 0$ and $\mathbb{P}$-a.s.*

$$N(t) = \begin{cases} 0 & \text{if } t < Z_1 \\ k & \text{if } t \in [Z_1 + ... + Z_k, Z_1 + ... + Z_{k+1}). \end{cases}$$

As a consequence, a Poisson process takes values in $\mathbb{Z}^+$ with a finite number of jumps on a finite time interval. Due to the following proposition, we get that each jump is of size 1.

Proposition 3.28 (Proposition 4.9 (iv), [42])**.** *An $\mathbb{Z}^+$-valued Lévy process $(N(t))_{t\geq 0}$ is a Poisson process if and only if for all $t \geq 0$*

$$\mathbb{P}(\Delta N(t) = N(t) - N(t-) \in \{0, 1\}) = 1.$$

Next, we use Poisson processes for the construction of compound Poisson processes with values in the separable Hilbert space $\mathcal{U}$. Let ν be a finite measure on $\mathcal{U}$ such that $\nu(\{0\}) = 0$. Moreover, let ν^{*k} denote the k-th convolution of the measure ν and let $\nu^0 = \delta_0$, i.e. ν^0 is the unit measure concentrated at $\{0\}$. Then we have the following definition.

Definition 3.29. *A càdlàg Lévy process $(L(t))_{t\geq 0}$ with values in $\mathcal{U}$ is a* ***compound Poisson process*** *with Lévy measure or jump intensity measure ν if for all $t \geq 0$ and all $\Gamma \in \mathcal{B}(\mathcal{U})$*

$$\mathbb{P}(L(t) \in \Gamma) = e^{-\nu(\mathcal{U})t} \sum_{k=0}^{\infty} \frac{t^k}{k!} \nu^{*k}(\Gamma).$$

We get the following construction of a compound Poisson process given a Lévy measure ν.

Proposition 3.30 (Theorem 4.15 (i) and (ii), [42])**.** *Let ν be a finite measure supported on $\mathcal{U}\backslash\{0\}$ and let $a = \nu(\mathcal{U})$. An $\mathcal{U}$-valued stochastic process $(L(t))_{t\geq 0}$ is a compound Poisson process with Lévy measure ν if and only if there exists a sequence $(Z_m)_{m\in\mathbb{N}}$ of independent $\mathcal{U}$-valued random variables with identical laws equal to $a^{-1}\nu$ and a Poisson process $(N(t))_{t\geq 0}$ with intensity a and independent of $(Z_m)_{m\in\mathbb{N}}$ such that for all $t \geq 0$ and $\mathbb{P}$-a.s.*

$$L(t) = \sum_{k=1}^{N(t)} Z_k.$$

Next, we state important properties of compound Poisson processes. We start with a condition for integrability.

Proposition 3.31 (Proposition 4.18 (i), [42])**.** *An $\mathcal{U}$-valued compound Poisson process $(L(t))_{t\geq 0}$ with Lévy measure ν is integrable if and only if*

$$\int_{\mathcal{U}} \|y\|_{\mathcal{U}} \, \nu(dy) < \infty. \tag{3.2}$$

If the condition (3.2) holds, then we have for all $t \geq 0$

$$\mathbb{E}[L(t)] = t \int_{\mathcal{U}} y \, \nu(dy).$$

If $(L(t))_{t\geq 0}$ is an integrable compound Poisson process with values in $\mathcal{U}$, then one can introduce the compensated compound Poisson process $(\widehat{L}(t))_{t\geq 0}$ given by $\widehat{L}(t) = L(t) - \mathbb{E}[L(t)]$ for all $t \geq 0$ and $\mathbb{P}$-almost surely. We get the following result.

Proposition 3.32. *Let $(\mathcal{F}_t)_{t\geq 0}$ be a normal filtration. If $(L(t))_{t\geq 0}$ is an integrable $\mathcal{F}_t$-adapted compound Poisson process, then the compensated compound Poisson process $(\widehat{L}(t))_{t\geq 0}$ is a martingale.*

Proof. Due to the fact that $(L(t))_{t\geq 0}$ is an $\mathcal{F}_t$-adapted Lévy process, the increment $\widehat{L}(t)-\widehat{L}(s)$ is independent of $\mathcal{F}_s$ for all $t \geq s \geq 0$. By Proposition 3.16 (i) and (iv), we get for all $t \geq s \geq 0$ and $\mathbb{P}$-a.s.

$$\mathbb{E}\left[\widehat{L}(t)\Big|\,\mathcal{F}_s\right] = \mathbb{E}\left[\widehat{L}(t) - \widehat{L}(s)\Big|\,\mathcal{F}_s\right] + \mathbb{E}\left[\widehat{L}(s)\Big|\,\mathcal{F}_s\right] = \mathbb{E}\left[\widehat{L}(t) - \widehat{L}(s)\right] + \mathbb{E}\left[\widehat{L}(s)\Big|\,\mathcal{F}_s\right].$$

Obviously, the process $(\widehat{L}(t))_{t\geq 0}$ is $\mathcal{F}_t$-adapted and $\mathbb{E}\left[\widehat{L}(t)\right] = 0$ for all $t \geq 0$. Since the filtration $(\mathcal{F}_t)_{t\geq 0}$ is normal, we can conclude that the process $(\widehat{L}(t))_{t\geq 0}$ is a martingale. □

We proceed with a condition on (compensated) compound Poisson process to be square integrable.

Proposition 3.33 (Proposition 4.18 (iii), [42])**.** *An $\mathcal{U}$-valued compound Poisson process $(L(t))_{t\geq 0}$ with Lévy measure ν is square integrable if and only if*

$$\int_{\mathcal{U}} \|y\|_{\mathcal{U}}^2 \, \nu(dy) < \infty. \tag{3.3}$$

If the condition (3.3) holds, then we have for all $t \geq 0$

$$\mathbb{E}\|L(t)\|_{\mathcal{U}}^2 = t\int_{\mathcal{U}} \|y\|_{\mathcal{U}}^2\,\nu(dy) + t^2\left\|\int_{\mathcal{U}} y\,\nu(dy)\right\|_{\mathcal{U}}^2 \quad \textit{and} \quad \mathbb{E}\left\|\widehat{L}(t)\right\|_{\mathcal{U}}^2 = t\int_{\mathcal{U}} \|y\|_{\mathcal{U}}^2\,\nu(dy).$$

Moreover, we get for all $t \geq 0$ and every $x, \tilde{x} \in \mathcal{U}$

$$\mathbb{E}\left[\left\langle \widehat{L}(t), x\right\rangle_{\mathcal{U}}\left\langle \widehat{L}(t), \tilde{x}\right\rangle_{\mathcal{U}}\right] = t\int_{\mathcal{U}} \langle x, y\rangle_{\mathcal{U}}\langle \tilde{x}, y\rangle_{\mathcal{U}}\,\nu(dy).$$

Finally, we state the characteristic function of a (compensated) compound Poisson process.

Proposition 3.34 ([71]). *Let $(L(t))_{t\geq 0}$ be an $\mathcal{U}$-valued compound Poisson process with Lévy measure ν. We have for every $x \in \mathcal{U}$, all $z \in \mathbb{C}$ and all $t \geq 0$*

$$\mathbb{E}\exp\left\{z\langle x, L(t)\rangle_{\mathcal{U}}\right\} = \exp\left\{-t\int_{\mathcal{U}}\left(1 - e^{z\langle x,y\rangle_{\mathcal{U}}}\right)\nu(dy)\right\}.$$

For the compensated process $(\widehat{L}(t))_{t\geq 0}$, we have for every $x \in \mathcal{U}$, all $z \in \mathbb{C}$ and all $t \geq 0$

$$\mathbb{E}\exp\left\{z\left\langle x, \widehat{L}(t)\right\rangle_{\mathcal{U}}\right\} = \exp\left\{-t\int_{\mathcal{U}}\left(1 - e^{z\langle x,y\rangle_{\mathcal{U}}} + z\langle x, y\rangle_{\mathcal{U}}\right)\nu(dy)\right\}.$$

Q-Wiener Processes

A Q-Wiener process is a typical example of a continuous Lévy process. The definition of a Q-Wiener process requires to introduce Gaussian measures on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$. First, we recall some basic properties of Gaussian measures on finite dimensional spaces. A Gaussian measure μ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ with $n \in \mathbb{N}$ is either concentrated at one point $\mu = \delta_m$ with $m \in \mathbb{R}^n$ or has the density $f\colon \mathbb{R}^n \to \mathbb{R}$ given by

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det(Q)}}\, e^{-\frac{1}{2}\langle Q^{-1}(x-m), x-m\rangle_{\mathbb{R}^n}}$$

for all $x \in \mathbb{R}^n$, where $m = (m_1, ...m_n) \in \mathbb{R}^n$ and $Q = (q_{j,k})_{j,k=1,..,n} \in \mathbb{R}^{n\times n}$ is a positive symmetric matrix. The characteristic functional $\hat{\mu}\colon \mathbb{R}^n \to \mathbb{R}$ of a Gaussian measure μ is given by

$$\hat{\mu}(\lambda) = \int_{\mathbb{R}^n} e^{i\langle\lambda,x\rangle_{\mathbb{R}^n}}\,\mu(dx) = e^{i\langle\lambda,m\rangle_{\mathbb{R}^n} - \frac{1}{2}\langle Q\lambda,\lambda\rangle_{\mathbb{R}^n}}$$

for all $\lambda \in \mathbb{R}^n$. As a consequence, a Gaussian measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ is uniquely determined by the values of m and Q and hence, we denoted this measure by $\mathcal{N}(m, Q)$. Moreover, we have for each $j, k = 1, ..., n$

$$\int_{\mathbb{R}^n} x_j\,\mathcal{N}(m, Q)(dx) = m_j \quad \text{and} \quad \int_{\mathbb{R}^n} (x_j - m_j)(x_k - m_k)\,\mathcal{N}(m, Q)(dx) = q_{j,k}.$$

Thus, we call m the mean and Q is the covariance matrix. Based on the finite dimensional setting, we can introduce Gaussian measures on Hilbert spaces.

Definition 3.35. *A measure μ on the space $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$ is called **Gaussian** if for every $h \in \mathcal{U}$, there exist $m = m(h) \in \mathbb{R}$ and $q = q(h) \geq 0$ such that for all $A \in \mathcal{B}(\mathbb{R})$*

$$\mu(\{x \in \mathcal{U}\colon \langle h, x\rangle_{\mathcal{U}} \in A\}) = \mathcal{N}(m, q)(A).$$

Proposition 3.36 (Theorem 2.1.2, [73]). *A measure μ on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$ is Gaussian if and only if*

$$\hat{\mu}(\lambda) = \int_{\mathcal{U}} e^{i\langle \lambda, x\rangle_{\mathcal{U}}} \mu(dx) = e^{i\langle \lambda, m\rangle_{\mathcal{U}} - \frac{1}{2}\langle Q\lambda, \lambda\rangle_{\mathcal{U}}}$$

for every $\lambda \in \mathcal{U}$, where $m \in \mathcal{U}$ and $Q \in \mathcal{L}_1^+(\mathcal{U})$. Moreover, we have for every $g, h \in \mathcal{U}$

$$\int_{\mathcal{U}} \langle x, h\rangle_{\mathcal{U}}\, \mu(dx) = \langle m, h\rangle_{\mathcal{U}},$$
$$\int_{\mathcal{U}} \langle x - m, h\rangle_{\mathcal{U}} \langle x - m, g\rangle_{\mathcal{U}}\, \mu(dx) = \langle Qh, g\rangle_{\mathcal{U}},$$
$$\int_{\mathcal{U}} \|x - m\|_{\mathcal{U}}^2\, \mu(dx) = Tr(Q).$$

In the previous proposition, the element $m \in \mathcal{U}$ is called the mean and $Q \in \mathcal{L}_1^+(\mathcal{U})$ is called the covariance operator, which uniquely determine the Gaussian measure on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$. Hence, we denote the Gaussian measure on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$ by $\mathcal{N}(m, Q)$.

Definition 3.37. *A random variable X with values in $\mathcal{U}$ is* ***Gaussian*** *if there exist $m \in \mathcal{U}$ and $Q \in \mathcal{L}_1^+(\mathcal{U})$ such that $\mathscr{L}(X) = \mathcal{N}(m, Q)$. A stochastic process $(X(t))_{t\geq 0}$ with values in $\mathcal{U}$ is* ***Gaussian*** *if for each $n \in \mathbb{N}$ and arbitrary $t_1, ..., t_n \geq 0$, the $\mathcal{U}^n$-valued random variable $(X(t_1), ..., X(t_n))$ is Gaussian.*

Proposition 3.38 (Proposition 2.1.4, [73]). *Let X be an $\mathcal{U}$-valued Gaussian random variable. Then $\langle X, h\rangle_{\mathcal{U}}$ is a real valued Gaussian random variable for every $h \in \mathcal{U}$ and the following statements holds:*

- $\mathbb{E}\langle X, h\rangle_{\mathcal{U}} = \langle m, h\rangle_{\mathcal{U}}$ *for every $h \in \mathcal{U}$;*
- $\mathbb{E}\langle X - m, h\rangle_{\mathcal{U}} \langle X - m, g\rangle_{\mathcal{U}} = \langle Qh, g\rangle_{\mathcal{U}}$ *for every $g, h \in \mathcal{U}$;*
- $\mathbb{E}\|X - m\|_{\mathcal{U}}^2 = Tr(Q)$.

We proceed with the definition and basic properties of a Q-Wiener process.

Definition 3.39. *An $\mathcal{U}$-valued continuous Lévy process $(W(t))_{t\geq 0}$ with mean 0 is called a* ***Q-Wiener process****.*

Proposition 3.40 (Theorem 4.20, [42]). *Let $(W(t))_{t\geq 0}$ be a Q-Wiener process with values in $\mathcal{U}$. Then $(W(t))_{t\geq 0}$ is a Gaussian process such that $\mathbb{E}\|W(t)\|_{\mathcal{U}}^2 < \infty$ for all $t \geq 0$.*

Remark 3.41. *In [23, 73], it is assumed that a Q-Wiener process $(W(t))_{t\geq 0}$ has Gaussian increments, i.e. $\mathscr{L}(W(t) - W(s)) = \mathcal{N}(0, (t - s)Q)$ for all $t \geq s \geq 0$. Here, one can obtain this property as follows: By Proposition 3.40, the process $(W(t))_{t\geq 0}$ is a Gaussian process. Hence, the distribution $\mathscr{L}(W(t))$ is Gaussian with mean 0 for all $t \geq 0$. Moreover, we have for every $g, h \in \mathcal{U}$ and all $s, t \geq 0$*

$$\mathbb{E}\langle W(t), h\rangle_{\mathcal{U}} \langle W(s), g\rangle_{\mathcal{U}} = t \wedge s\, \mathbb{E}\langle W(1), h\rangle_{\mathcal{U}} \langle W(1), g\rangle_{\mathcal{U}} = t \wedge s\, \langle Qh, g\rangle_{\mathcal{U}},$$

where $Q \in \mathcal{L}_1^+(\mathcal{U})$ is the covariance operator of the Gaussian measure $\mathscr{L}(W(1))$ arising from Proposition 3.38. This implies for every $h \in \mathcal{U}$ and all $t \geq s \geq 0$

$$\mathbb{E}\langle W(t) - W(s), h\rangle_{\mathcal{U}}^2 = (t - s)\langle Qh, h\rangle_{\mathcal{U}}.$$

Thus, we get $\mathscr{L}(W(t) - W(s)) = \mathcal{N}(0, (t - s)Q)$ for all $t \geq s \geq 0$.

As a consequence of the previous remark, the covariance operator $Q \in \mathcal{L}_1^+(\mathcal{U})$ completely characterizes the distribution of a Q-Wiener process. By Proposition C.5, there exist an orthonormal basis $(u_n)_{n\in\mathbb{N}}$ of $\mathcal{U}$ and a sequence $(\lambda_n)_{n\in\mathbb{N}}$ of nonnegative real numbers such that $Qu_n = \lambda_n u_n$ for each $n \in \mathbb{N}$. In the following proposition, we provide a series presentation of a Q-Wiener process in term of mutually independent real valued Brownian motions.

Proposition 3.42 (Proposition 4.3, [23])**.** *Assume that $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in $\mathcal{U}$ and covariance operator $Q \in \mathcal{L}_1^+(\mathcal{U})$. Let $(u_n)_{n\in\mathbb{N}}$ be an orthonormal basis of $\mathcal{U}$ and let $(\lambda_n)_{n\in\mathbb{N}}$ be a sequence of nonnegative real numbers such that $Qu_n = \lambda_n u_n$ for each $n \in \mathbb{N}$. Then for all $t \geq 0$ and $\mathbb{P}$-a.s.*

$$W(t) = \sum_{n=1}^{\infty} \sqrt{\lambda_n} w_n(t) u_n, \tag{3.4}$$

where for each $n \in \mathbb{N}$, the processes $(w_n(t))_{t\geq 0}$ are mutually independent real valued Brownian motions given by

$$w_n(t) = \frac{1}{\sqrt{\lambda_n}} \langle W(t), u_n \rangle_{\mathcal{U}}$$

for all $t \geq 0$ and $\mathbb{P}$-almost surely. The series in (3.4) converges in $L^2(\Omega;\mathcal{U})$.

Finally, we state the characteristic function of a Q-Wiener process.

Proposition 3.43. *Let $(W(t))_{t\geq 0}$ be an $\mathcal{U}$-valued Q-Wiener process with covariance operator $Q \in \mathcal{L}_1^+(\mathcal{U})$. We have for every $x \in \mathcal{U}$ and all $t \geq 0$*

$$\mathbb{E}\, e^{i\langle x, W(t)\rangle_{\mathcal{U}}} = e^{-\frac{t}{2}\langle Qx, x\rangle_{\mathcal{U}}}.$$

Proof. Due to Remark 3.41, we have that the distribution $\mathscr{L}(W(t))$ is Gaussian with mean 0 and covariance operator tQ. Using Proposition 3.36, the claim follows. □

3.2.2. Lévy-Khinchin Decomposition

Let $(L(t))_{t\geq 0}$ be a càdlàg Lévy process with values in $\mathcal{U}$. The Lévy-Khinchin decomposition provides an representation of the process $(L(t))_{t\geq 0}$ as the sum of its continuous part and its pure jump part. First, we consider the pure jump part. Let $A \in \mathcal{B}(\mathcal{U})$ be separated from 0, i.e. the set $A \in \mathcal{B}(\mathcal{U})$ satisfies $A \cap \{y \in \mathcal{U} \colon \|y\|_{\mathcal{U}} \leq r\} = \emptyset$ for sufficiently small $r > 0$. Then the process $(L_A(t))_{t\geq 0}$ given by

$$L_A(t) = \sum_{0\leq s\leq t} \mathbb{1}_A(\Delta L(s))\Delta L(s),$$

for all $t \geq 0$ and $\mathbb{P}$-almost surely, is a well defined Lévy process with values in $\mathcal{U}$. We get the following results.

Lemma 3.44. *Let $A \in \mathcal{B}(\mathcal{U})$ be separated from 0. Then the processes $(L_A(t))_{t\geq 0}$ and $(L(t) - L_A(t))_{t\geq 0}$ are independent Lévy processes.*

Proof. A proof can be found in [71, Appendix F]. □

Corollary 3.45. *Let $A, B \in \mathcal{B}(\mathcal{U})$ be disjoint sets that are separated from 0. Then the processes $(L_A(t))_{t\geq 0}$ and $(L_B(t))_{t\geq 0}$ are independent Lévy processes.*

Proof. Obviously, the set $A \cup B \in \mathcal{B}(\mathcal{U})$ is separated from 0. Thus, the process $(L_{A\cup B}(t))_{t\geq 0}$ is a well defined Lévy process with values in $\mathcal{U}$. Since A and B are disjoint, we get for all $t \geq 0$ and $\mathbb{P}$-a.s.

$$L_{A\cup B}(t) = L_A(t) + L_B(t).$$

The claim follows by Lemma 3.44. □

Lemma 3.46 (Lemma 4.25, [71])**.** *For all $A \in \mathcal{B}(\mathcal{U})$ separated from 0 and all $x \in \mathcal{U}$, we have*

$$\mathbb{E}\exp\left\{i\langle x, L_A(t)\rangle_{\mathcal{U}}\right\} = \exp\left\{-t\int_A \left(1 - e^{i\langle x,y\rangle_{\mathcal{U}}}\right)\nu(dy)\right\}.$$

By Proposition 3.34, we can conclude that $(L_A(t))_{t\geq 0}$ is an $\mathcal{U}$-valued compound Poisson process with Lévy measure $\mathbb{1}_A(y)\nu(dy)$. This process is fundamental by constructing the pure jump part of $(L(t))_{t\geq 0}$. Let $(r_n)_{n\in\mathbb{Z}^+}$ be a strictly decreasing null sequence. Then the Lévy process $(L(t))_{t\geq 0}$ has finite many jumps on the set $A_0 = \{y \in \mathcal{U}\colon \|y\|_{\mathcal{U}} \geq r_0\}$ on a finite time interval. The process $(L_{A_0}(t))_{t\geq 0}$ given by

$$L_{A_0}(t) = \sum_{0\leq s\leq t} \mathbb{1}_{A_0}(\Delta L(t))\Delta L(t),$$

for all $t \geq 0$ and $\mathbb{P}$-almost surely, is a well defined $\mathcal{U}$-valued Lévy process containing all jumps larger or equal to r_0 with respect to the norm in $\mathcal{U}$. Due to Lemma 3.46, the process $(L_{A_0}(t))_{t\geq 0}$ is a compound Poisson process with Lévy measure $\mathbb{1}_{\{\|y\|_{\mathcal{U}}\geq r_0\}}(y)\nu(dy)$. To cover the remaining jumps we introduce the sets $A_n = \{y \in \mathcal{U}\colon r_n \leq \|y\|_{\mathcal{U}} < r_{n-1}\}$ for each $n \in \mathbb{N}$. Note that these sets are still separated from 0. Hence, for each $n \in \mathbb{N}$, the processes $(L_{A_n}(t))_{t\geq 0}$ given by

$$L_{A_n}(t) = \sum_{0\leq s\leq t} \mathbb{1}_{A_n}(\Delta L(t))\Delta L(t),$$

for all $t \geq 0$ and $\mathbb{P}$-almost surely, are well defined $\mathcal{U}$-valued Lévy process. By Lemma 3.46, the processes $(L_{A_n}(t))_{t\geq 0}$, $n \in \mathbb{N}$, are compound Poisson processes with Lévy measure $\mathbb{1}_{\{r_n\leq\|y\|_{\mathcal{U}}<r_{n-1}\}}(y)\nu(dy)$. In general, we can not ensure that the series of small jumps $\sum_{n=1}^{\infty} L_{A_n}(t)$ converges on a bounded time interval $[0,T]$. Therefore, for each $n \in \mathbb{N}$, we have to consider the compensated processes of $(L_{A_n}(t))_{t\geq 0}$. Note that the jumps of $(L_{A_n}(t))_{t\geq 0}$ are bounded by r_{n-1} for each $n \in \mathbb{N}$. Hence, we can apply Theorem 3.25 with the result that the processes $(L_{A_n}(t))_{t\geq 0}$ are integrable for each $n \in \mathbb{N}$. Using Proposition 3.31, we have for each $n \in \mathbb{N}$ and all $t \geq 0$

$$\mathbb{E}[L_{A_n}(t)] = t\int_{A_n} y\,\nu(dy).$$

Thus, we can introduce the compensated compound Poisson processes $(L_n(t))_{t\geq 0}$, $n \in \mathbb{N}$, given by

$$L_n(t) = L_{A_n}(t) - \mathbb{E}[L_{A_n}(t)] = L_{A_n}(t) - t\int_{A_n} y\,\nu(dy)$$

for all $t \geq 0$ and $\mathbb{P}$-almost surely. To get a convergence result of the series $\sum_{n=1}^{\infty} L_n(t)$ on a bounded interval $[0,T]$, we need the following preliminary result.

Proposition 3.47 (Theorem 4.23 (i), [71])**.** *If ν is a jump intensity measure corresponding to an $\mathcal{U}$-valued Lévy process $(L(t))_{t\geq 0}$, then*

$$\int_{\mathcal{U}} \left(\|y\|_{\mathcal{U}}^2 \wedge 1\right)\nu(dy) < \infty. \tag{3.5}$$

Lemma 3.48 (Lemma 4.26, [71])**.** *If assumption (3.5) is satisfied, then the series $\sum_{n=1}^{\infty} L_n(t)$ converges $\mathbb{P}$-a.s. uniformly on each bounded interval $[0,T]$.*

Now, we are able to state the Lévy-Khinchin decomposition, where we also characterize the continuous part of a Lévy process.

Theorem 3.49 (Theorem 4.23 (ii), [71]). *Let $(L(t))_{t\geq 0}$ be an $\mathcal{U}$-valued Lévy process with Lévy measure ν. Then we have the following representation for all $t \geq 0$ and $\mathbb{P}$-almost surely:*

$$L(t) = at + W(t) + \sum_{n=1}^{\infty} \left(L_{A_n}(t) - t \int_{A_n} y \, \nu(dy) \right) + L_{A_0}(t), \tag{3.6}$$

where $a \in \mathcal{U}$, $(W(t))_{t\geq 0}$ is an $\mathcal{U}$-valued Q-Wiener process, $(L_{A_0}(t))_{t\geq 0}$ is an $\mathcal{U}$-valued compound Poisson process with Lévy measure $\mathbb{1}_{\{\|y\|_{\mathcal{U}} \geq r_0\}}(y)\nu(dy)$ and $(L_{A_n}(t))_{t\geq 0}$ is an $\mathcal{U}$-valued compound Poisson processes with Lévy measure $\mathbb{1}_{\{r_n \leq \|y\|_{\mathcal{U}} < r_{n-1}\}}(y)\nu(dy)$ for each $n \in \mathbb{N}$. Moreover, all members of the representation are independent processes and the series converges $\mathbb{P}$-a.s. uniformly on each bounded subinterval of $[0, \infty)$.

As a consequence of the previous theorem, we get the following Lévy-Khinchin formula.

Theorem 3.50 (Theorem 4.27, [42]). *Let $(L(t))_{t\geq 0}$ be an $\mathcal{U}$-valued Lévy process with Lévy measure ν and let μ_t be the distribution of $L(t)$ for all $t \geq 0$. We have for every $x \in \mathcal{U}$ and all $t \geq 0$*

$$\mathbb{E}\, e^{i\langle x, L(t)\rangle_{\mathcal{U}}} = \int_{\mathcal{U}} e^{i\langle x,y\rangle_{\mathcal{U}}} \mu_t(dy) = e^{-t\psi(x)},$$

where

$$\psi(x) = -i\langle a, x\rangle_{\mathcal{U}} + \frac{1}{2}\langle Qx, x\rangle_{\mathcal{U}} + \int_{\mathcal{U}} \left(1 - e^{i\langle x,y\rangle_{\mathcal{U}}} + \mathbb{1}_{\{\|y\|_{\mathcal{U}}<1\}}(y)\, i\langle x, y\rangle_{\mathcal{U}}\right) \nu(dy).$$

Here, the value $a \in \mathcal{U}$ and the covariance operator $Q \in \mathcal{L}_1^+(\mathcal{U})$ corresponding to the Q-Wiener process $(W(t))_{t\geq 0}$ arise from Theorem 3.49.

In the previous theorem, the triple (a, Q, ν) is called the characteristics of the Lévy process $(L(t))_{t\geq 0}$, which describes the distribution μ_t for all $t \geq 0$ completely.

3.2.3. Square Integrable Lévy Processes

Let $(L(t))_{t\geq 0}$ be a square integrable Lévy process with values in $\mathcal{U}$. We assume that $(L(t))_{t\geq 0}$ is $\mathcal{F}_t$-adapted such that $L(t) - L(s)$ is independent of $\mathcal{F}_s$ for all $t > s \geq 0$.

Proposition 3.51 (Theorem 4.44, [71]). *There exist $m \in \mathcal{U}$ and $Q \in \mathcal{L}_1^+(\mathcal{U})$ such that*

$$\begin{aligned}
&\mathbb{E}\langle L(t), x\rangle_{\mathcal{U}} = t\,\langle m, x\rangle_{\mathcal{U}},\\
&\mathbb{E}\langle L(t) - mt, x\rangle_{\mathcal{U}}\langle L(s) - ms, y\rangle_{\mathcal{U}} = t \wedge s\,\langle Qx, y\rangle_{\mathcal{U}},\\
&\mathbb{E}\|L(t) - mt\|_{\mathcal{U}}^2 = t\,Tr(Q)
\end{aligned}$$

for all $t, s \geq 0$ and every $x, y \in \mathcal{U}$.

In the previous proposition, the element $m \in \mathcal{U}$ is called the mean and the operator $Q \in \mathcal{L}_1^+(\mathcal{U})$ is called the covariance operator of $(L(t))_{t\geq 0}$. Next, we consider the Lévy-Khinchin decomposition for square integrable Lévy processes. We need the following preliminary result, which is an immediate consequence of Proposition 3.33 and Theorem 3.49.

Corollary 3.52 (Theorem 4.47 (i), [71]). *An $\mathcal{U}$-valued Lévy process with Lévy measure ν is square integrable if and only if*

$$\int_{\mathcal{U}} \|y\|_{\mathcal{U}}^2\, \nu(dy).$$

Using Theorem 3.49, the process $(L(t))_{t\geq 0}$ has the representation (3.6). Recall that $(L_{A_0}(t))_{t\geq 0}$ is an $\mathcal{U}$-valued compound Poisson process with Lévy measure $\mathbb{1}_{\{\|y\|_{\mathcal{U}}\geq r_0\}}(y)\nu(dy)$. By Proposition 3.33 and Corollary 3.52, the process $(L_{A_0}(t))_{t\geq 0}$ is square integrable. Hence, the mean exists and due to Proposition 3.31, we get for all $t \geq 0$

$$\mathbb{E}[L_{A_0}(t)] = t\int_{\{\|y\|_{\mathcal{U}}\geq r_0\}} y\,\nu(dy) = t\int_{A_0} y\,\nu(dy).$$

We get the following result.

Theorem 3.53. *A square integrable Lévy process $(L(t))_{t\geq 0}$ with values in $\mathcal{U}$ has the following decomposition for all $t \geq 0$ and $\mathbb{P}$-almost surely:*

$$L(t) = tb + W(t) + M_J(t), \tag{3.7}$$

where $b \in \mathcal{U}$, $(W(t))_{t\geq 0}$ is an $\mathcal{U}$-valued Q-Wiener process and $(M_J(t))_{t\geq 0}$ is an $\mathcal{U}$-valued square integrable Lévy process containing all jumps of $(L(t))_{t\geq 0}$. Moreover, the processes $(W(t))_{t\geq 0}$ and $(M_J(t))_{t\geq 0}$ are independent martingales with mean 0.

Proof. Using Theorem 3.49, we have for all $t \geq 0$ and $\mathbb{P}$-almost surely

$$L(t) = ta + t\int_{A_0} y\,\nu(dy) + W(t) + \sum_{n=1}^{\infty}\left(L_{A_n}(t) - t\int_{A_n} y\,\nu(dy)\right) + L_{A_0}(t) - t\int_{A_0} y\,\nu(dy).$$

We set $b = a + \int_{A_0} y\,\nu(dy)$ and

$$M_J(t) = \sum_{n=1}^{\infty}\left(L_{A_n}(t) - t\int_{A_n} y\,\nu(dy)\right) + L_{A_0}(t) - t\int_{A_0} y\,\nu(dy)$$

for all $t \geq 0$ and $\mathbb{P}$-almost surely. By Theorem 3.49, the processes $(W(t))_{t\geq 0}$ and $(M_J(t))_{t\geq 0}$ are independent. Since $(L(t))_{t\geq 0}$ is $\mathcal{F}_t$-adapted such that $L(t) - L(s)$ is independent of $\mathcal{F}_s$ for all $t > s > 0$, we can conclude that this property holds for $(W(t))_{t\geq 0}$ as well as for $(M_J(t))_{t\geq 0}$. Hence, the processes $(W(t))_{t\geq 0}$ and $(M_J(t))_{t\geq 0}$ are martingales. By definition, the Q-Wiener process $(W(t))_{t\geq 0}$ has mean 0. Since $(M_J(t))_{t\geq 0}$ is a series of compensated compound Poisson processes, it has mean 0 as well. □

Remark 3.54. *It follows from the proof of the previous theorem that for all $t \geq 0$*

$$\mathbb{E}[L(t)] = tb = ta + t\int_{A_0} y\,\nu(dy),$$

where $a \in \mathcal{U}$ arises from Theorem 3.49 and ν is the Lévy measure corresponding to $(L(t))_{t\geq 0}$.

Assume that $(L(t))_{t\geq 0}$ has the representation (3.7). Let $Q_0 \in \mathcal{L}_1^+(\mathcal{U})$ be the covariance operator of $(W(t))_{t\geq 0}$ and let $Q_1 \in \mathcal{L}_1^+(\mathcal{U})$ be the covariance operator of $(M_J(t))_{t\geq 0}$. Since $(W(t))_{t\geq 0}$ and $(M_J(t))_{t\geq 0}$ are independent, the process $(L(t))_{t\geq 0}$ has the covariance operator $Q = Q_0 + Q_1$. The following theorem provides a characterization of Q_1.

Theorem 3.55 (Theorem 4.47 (ii), [71])**.** *Assume that the process $(L(t))_{t\geq 0}$ has the representation (3.7). Let $Q_1 \in \mathcal{L}_1^+(\mathcal{U})$ be the covariance operator of $(M_J(t))_{t\geq 0}$. Then we have for every $x, z \in \mathcal{U}$*

$$\langle Q_1 x, z\rangle_{\mathcal{U}} = \int_{\mathcal{U}} \langle x, y\rangle_{\mathcal{U}}\langle z, y\rangle_{\mathcal{U}}\,\nu(dy).$$

Another important representation of a Lévy process is the expansion as a series of real-valued Lévy processes. For the remaining part of this section, we assume that the $\mathcal{U}$-valued square integrable Lévy process $(L(t))_{t\geq 0}$ has mean 0 and covariance operator $Q \in \mathcal{L}_1^+(\mathcal{U})$. By Proposition C.5, there exist an orthonormal basis $(u_n)_{n\in\mathbb{N}}$ of $\mathcal{U}$ and a sequence $(\lambda_n)_{n\in\mathbb{N}}$ of nonnegative real numbers such that $Qu_n = \lambda_n u_n$ for each $n \in \mathbb{N}$. We get the following convergence results, which generalizes Proposition 3.42.

Theorem 3.56. *We have for all $t \geq 0$ and $\mathbb{P}$-a.s.*

$$L(t) = \sum_{n=1}^{\infty} L_n(t) u_n, \tag{3.8}$$

where for each $n \in \mathbb{N}$, the processes $(L_n(t))_{t\geq 0}$ are uncorrelated square integrable Lévy processes with values in $\mathbb{R}$ and mean 0 given by

$$L_n(t) = \langle L(t), u_n \rangle_{\mathcal{U}}$$

for all $t \geq 0$ and $\mathbb{P}$-almost surely. The series in (3.8) converges in probability uniformly in t on any compact interval $[0,T]$ and in $L^2(\Omega;\mathcal{U})$ for all $t \geq 0$.

Proof. Since $Q \in \mathcal{L}_1^+(\mathcal{U})$, we have

$$\sum_{n=1}^{\infty} \lambda_n = Tr(Q) < \infty.$$

Obviously, the processes $(L_n(t))_{t\geq 0}$ are real valued Lévy processes for each $n \in \mathbb{N}$. By Proposition 3.51, we get for each $n \in \mathbb{N}$ and all $t \geq 0$

$$\mathbb{E}[L_n(t)] = \mathbb{E}\langle L(t), u_n \rangle_{\mathcal{U}} = 0$$

and

$$\mathbb{E}\left[L_n(t)^2\right] = \mathbb{E}\langle L(t), u_n \rangle_{\mathcal{U}}^2 = t\langle Qu_n, u_n \rangle_{\mathcal{U}} = \lambda_n t.$$

Similarly, we obtain for each $n, m \in \mathbb{N}$ and all $t, s \geq 0$

$$\mathbb{E}[L_n(t)L_m(s)] = \mathbb{E}\langle L(t), u_n \rangle_{\mathcal{U}} \langle L(s), u_m \rangle_{\mathcal{U}} = t \wedge s\langle Qu_n, u_m \rangle_{\mathcal{U}} = (t \wedge s)\lambda_n \delta_{n,m},$$

where $\delta_{n,m}$ denotes the Kronecker delta. Hence, the processes $(L_n(t))_{t\geq 0}$ are square integrable and uncorrelated with mean 0 for each $n \in \mathbb{N}$.

Next, we show the convergence of the series (3.8) in $L^2(\Omega;\mathcal{U})$ for all $t \geq 0$. We set for each $k \in \mathbb{N}$, all $t \geq 0$ and $\mathbb{P}$-a.s.

$$S_k(t) = \sum_{n=1}^{k} L_n(t) u_n.$$

Then it follows that for all $t \geq 0$ and each $j, k \in \mathbb{N}$ with $j < k$

$$\mathbb{E}\left\| S_k(t) - S_j(t) \right\|_{\mathcal{U}}^2 = \mathbb{E}\left\| \sum_{n=j+1}^{k} L_n(t) u_n \right\|_{\mathcal{U}}^2 = \sum_{n,m=j+1}^{k} \mathbb{E}\left[L_n(t)L_m(t)\right] \langle u_n, u_m \rangle_{\mathcal{U}} = t \sum_{n=j+1}^{k} \lambda_n.$$

Since $\sum_{n=1}^{\infty} \lambda_n < \infty$, we can conclude that $(S_k(t))_{k\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega;\mathcal{U})$ for all $t \geq 0$. Hence, the series (3.8) converges in $L^2(\Omega;\mathcal{U})$ for all $t \geq 0$.

Finally, we prove the convergence of the series (3.8) in probability uniformly in t on any compact interval $[0,T]$. Recall that $(L(t))_{t\geq 0}$ is a square integrable Lévy process with mean 0 such that $L(t) - L(s)$ is independent of $\mathcal{F}_s$ for all $t > s \geq 0$. Due to Proposition 3.18, the process $(L(t))_{t\geq 0}$ is a martingale. Hence, we can conclude that the process $(S_k(t))_{t\geq 0}$ is a stochastically continuous square integrable martingale with values in $\mathcal{U}$. Applying Theorem 3.19, we get for all $r > 0$ and each $j, k \in \mathbb{N}$

$$\mathbb{P}\left(\sup_{t\in[0,T]} \| S_k(t) - S_j(t) \|_{\mathcal{U}} \geq r \right) \leq \frac{1}{r^2}\, \mathbb{E}\| S_k(T) - S_j(T) \|_{\mathcal{U}}^2$$

and the claim follows. □

Remark 3.57. *(i) The series (3.8) converges for all* $t \geq 0$ *and* $\mathbb{P}$*-a.s. if and only if*

$$\sum_{n=1}^{\infty} |L_n(t)|^2 < \infty$$

for all $t \geq 0$ *and* $\mathbb{P}$*-almost surely.*
(ii) For a general Lévy process, the series (3.8) converges in probability, uniformly on any compact interval $[0, T]$*, see [71, Theorem 4.39].*

By Proposition 3.18, we can conclude that the Lévy process $(L(t))_{t\geq 0}$ is a martingale. Moreover, we get the following martingale properties.

Proposition 3.58 (Theorem 4.49 (ii),[71])**.** *The processes* $(\|L(t)\|_{\mathcal{U}}^2 - t\,Tr(Q))_{t\geq 0}$ *and* $(L(t) \otimes L(t) - tQ)_{t\geq 0}$ *are martingales with values in* $\mathbb{R}$ *and* $\mathcal{L}_1(\mathcal{U})$*, respectively. Moreover, the process* $(\langle L(t), x\rangle_{\mathcal{U}})_{t\geq 0}$ *is a square integrable real valued martingale for every* $x \in \mathcal{U}$ *and the process* $(\langle L(t), x\rangle_{\mathcal{U}} \langle L(t), y\rangle_{\mathcal{U}} - t\langle Qx, y\rangle_{\mathcal{U}})_{t\geq 0}$ *is a real valued martingale for every* $x, y \in \mathcal{U}$*.*

It follows from the previous proposition that the angle bracket $(\langle L\rangle_t)_{t\geq 0}$ is given by $\langle L\rangle_t = t\,Tr(Q)$ for all $t \geq 0$ due to Theorem 3.20. Moreover, the operator angle bracket process $(\langle\langle L\rangle\rangle_t)_{t\geq 0}$ satisfies $\langle\langle L\rangle\rangle_t = tQ$ for all $t \geq 0$ resulting from Theorem 3.22.

3.3. A Stochastic Integral

In this section, we introduce infinite dimensional stochastic integrals with respect to a square integrable Lévy martingale. The construction is similar to the case of a stochastic integral with respect to a Q-Wiener process, see [23, 42, 73]. Let the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ be equipped with a normal filtration $(\mathcal{F}_t)_{t\geq 0}$. We start with a formal definition of elementary processes. Let $T > 0$ be a fixed terminal point of time and let $\mathcal{U}$ and $\mathcal{H}$ be separable Hilbert spaces.

Definition 3.59. *An* $\mathcal{L}(\mathcal{U}; \mathcal{H})$*-valued stochastic process* $(\Psi(t))_{t\in[0,T]}$ *is called* ***elementary*** *if there exists* $m \in \mathbb{N}$ *such that for all* $t \in [0, T]$ *and* $\mathbb{P}$*-a.s.*

$$\Psi(t) = \sum_{j=0}^{m-1} \Psi_j \mathbb{1}_{(t_j, t_{j+1}]}(t) \tag{3.9}$$

where $0 = t_0 < t_1 < ... < t_m = T$ *and* Ψ_j *are* $\mathcal{F}_{t_j}$*-measurable* $\mathcal{L}(\mathcal{U}; \mathcal{H})$*-valued random variables for* $j = 0, 1, ..., m - 1$*.*

The space of all $\mathcal{L}(\mathcal{U}; \mathcal{H})$-valued elementary processes is denoted by $\mathcal{E}_T$. Moreover, let $(L(t))_{t\geq 0}$ be a square integrable Lévy martingale with values in $\mathcal{U}$, i.e. $(L(t))_{t\geq 0}$ is an $\mathcal{U}$-valued square integrable Lévy process and a martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$. For $\Psi \in \mathcal{E}_T$ with representation (3.9), we define the stochastic integral by

$$I_t^L(\Psi) = \int_0^t \Psi(s)\, dL(s) = \sum_{j=0}^{m-1} \Psi_j \left(L(t_{j+1} \wedge t) - L(t_j \wedge t)\right) \tag{3.10}$$

for all $t \in [0, T]$ and $\mathbb{P}$-almost surely. Obviously, the operator I_t^L is linear on $\mathcal{E}_T$ for all $t \in [0, T]$ and $\mathbb{P}$-almost surely. Next, we extend the definition of the stochastic integral to a larger class of stochastic processes. As a consequence of Theorem 3.53, the process $(L(t))_{t\geq 0}$ is a square integrable Lévy martingale if and only if

$\mathbb{E}[L(t)] = 0$ for all $t \geq 0$. By Proposition 3.51, there exists an covariance operator $Q \in \mathcal{L}_1^+(\mathcal{U})$ of $(L(t))_{t\geq 0}$ such that

$$\mathbb{E}\langle L(t), x\rangle_{\mathcal{U}} \langle L(s), y\rangle_{\mathcal{U}} = t \wedge s \, \langle Qx, y\rangle_{\mathcal{U}}$$

for all $s, t \geq 0$ and every $x, y \in \mathcal{U}$. Let us denote by $\mathcal{L}_{(HS)}(\mathcal{U}; \mathcal{H})$ the space of Hilbert-Schmidt operators from $\mathcal{U}$ to $\mathcal{H}$ as introduced in Appendix C. Using Proposition C.9 and Remark C.10, there exists an operator $Q^{1/2} \in \mathcal{L}_1^+(\mathcal{U})$ such that $Q^{1/2}Q^{1/2} = Q$ and we get $\mathcal{L}(\mathcal{U}; \mathcal{H}) \subset \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U}); \mathcal{H})$. Therefore, we can conclude that elementary processes takes values in $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U}); \mathcal{H})$. For the remaining part of this section, we assume that every $\Psi \in \mathcal{E}_T$ satisfies

$$\mathbb{E} \int_0^T \|\Psi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})} \, dt < \infty.$$

We equip the space $\mathcal{E}_T$ with the inner product

$$\langle \Psi, \Phi\rangle_{\mathcal{E}_T} = \mathbb{E} \int_0^T \langle \Psi(t), \Phi(t)\rangle_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})} dt$$

for every $\Psi, \Phi \in \mathcal{E}_T$. We get the following result.

Theorem 3.60. *Let $(L(t))_{t\geq 0}$ be a square integrable Lévy martingale with values in a separable Hilbert space $\mathcal{U}$ and covariance operator $Q \in \mathcal{L}_1^+(\mathcal{U})$. If $\Psi \in \mathcal{E}_T$, then we have for all $t \in [0, T]$*

$$\mathbb{E} \left\| \int_0^t \Psi(s) \, dL(s) \right\|^2_{\mathcal{H}} = \mathbb{E} \int_0^t \|\Psi(s)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})} \, ds.$$

Proof. The claim results from [71, Proposition 8.6] and Proposition 3.58. □

As a consequence of the previous theorem, we get that the mapping $I_t^L \colon \mathcal{E}_T \to L^2(\Omega; \mathcal{H})$ is linear and bounded for all $t \in [0, T]$ and especially, the mapping $I_T^L \colon \mathcal{E}_T \to L^2(\Omega; \mathcal{H})$ is an isometric transformation. Hence, we can uniquely extend the definition of the stochastic integral to integrands taking values in the completion of $\mathcal{E}_T$ denoted by $\mathcal{L}_T^2$. We still denote the extension from $\mathcal{L}_T^2$ into $L^2(\Omega; \mathcal{H})$ by I_t^L and we write formally

$$I_t^L(\Psi) = \int_0^t \Psi(s) \, dL(s)$$

for all $t \in [0, T]$ and $\mathbb{P}$-almost surely. Moreover, we set

$$\int_r^t \Psi(s) \, dL(s) = \int_0^t \Psi(s) \, dL(s) - \int_0^r \Psi(s) \, dL(s)$$

for all $0 \leq r \leq t \leq T$ and $\mathbb{P}$-almost surely. The following proposition characterizes the space of integrands $\mathcal{L}_T^2$ explicitly.

Proposition 3.61 (Proposition 4.22, [23])**.** *The following statements hold:*

(i) If $(\Psi(t))_{t\in[0,T]}$ is an $\mathcal{L}(\mathcal{U}; \mathcal{H})$-valued predictable process, then $(\Psi(t))_{t\in[0,T]}$ is an $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U}); \mathcal{H})$-valued predictable process. In particular, elementary processes are $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U}); \mathcal{H})$-valued predictable processes.

(ii) *If* $(\Psi(t))_{t\in[0,T]}$ *is an* $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})$*-valued predictable processes such that*

$$\mathbb{E}\int_0^T \|\Psi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}\,dt<\infty,$$

then there exists a sequence $(\Psi_n)_{n\in\mathbb{N}}\subset\mathcal{E}_T$ *such that*

$$\lim_{n\to\infty}\mathbb{E}\int_0^T \|\Psi(t)-\Psi_n(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}\,dt=0.$$

Hence, the space $\mathcal{L}_T^2$ contains all predictable processes $(\Psi(t))_{t\in[0,T]}$ with values in $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})$ such that

$$\mathbb{E}\int_0^T \|\Psi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}\,dt<\infty.$$

The space $\mathcal{L}_T^2$ equipped with the inner product of $\mathcal{E}_T$ becomes a Hilbert space. Next, we provide some basic properties of the stochastic integral.

Theorem 3.62. *Let* $(L(t))_{t\geq 0}$ *be a square integrable Lévy martingale with values in a separable Hilbert space* $\mathcal{U}$ *and covariance operator* $Q\in\mathcal{L}_1^+(\mathcal{U})$ *and let* $\Psi\in\mathcal{L}_T^2$*. Then the following statements hold:*

(i) *If* $0\leq r<t\leq T$*, then* $\mathbb{1}_{(r,t]}\Psi\in\mathcal{L}_T^2$ *and we have* $\mathbb{P}$*-a.s.*

$$\int_r^t \Psi(s)\,dL(s)=\int_0^T \mathbb{1}_{(r,t]}(s)\Psi(s)\,dL(s). \tag{3.11}$$

(ii) *We get for all* $t\in[0,T]$

$$\mathbb{E}\left[\int_0^t \Psi(s)\,dL(s)\right]=0.$$

(iii) *We have for all* $t\in[0,T]$

$$\mathbb{E}\left\|\int_0^t \Psi(s)\,dL(s)\right\|^2_{\mathcal{H}}=\mathbb{E}\int_0^t \|\Psi(s)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}\,ds.$$

(iv) *The process* $(I_t^L(\Psi))_{t\in[0,T]}$ *given by*

$$I_t^L(\Psi)=\int_0^t \Psi(s)\,dL(s)$$

for all $t\in[0,T]$ *and* $\mathbb{P}$*-a.s. is a mean square continuous* $\mathcal{H}$*-valued martingale.*

Proof. It suffices to show the assertions (i), (ii) and (iii) for elementary processes. The generalization for elements in $\mathcal{L}_T^2$ can be obtained using standard density arguments.

First, we show that (i) holds. If $\Psi \in \mathcal{E}_T$, then the process $(\mathbb{1}_{(r,t]}(s)\Psi(s))_{s\in[0,T]}$ is an elementary process with values in $\mathcal{L}(\mathcal{U};\mathcal{H})$ and

$$\mathbb{E}\int_0^T \left\|\mathbb{1}_{(r,t]}(s)\Psi(s)\right\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})} ds \leq \mathbb{E}\int_0^T \|\Psi(s)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})} ds < \infty.$$

Hence, we get $\mathbb{1}_{(r,t]}\Psi \in \mathcal{E}_T$. To obtain equation (3.11), we use the operator notation introduced by equation (3.10). We show some useful preliminary identities. Let $\Phi \in \mathcal{E}_T$ have the following representation for all $s \in [0,T]$ and $\mathbb{P}$-almost surely:

$$\Phi(s) = \sum_{j=0}^{m-1} \Phi_j \mathbb{1}_{(t_j,t_{j+1}]}(s),$$

where $0 = t_0 < t_1 < ... < t_m = T$ and Φ_j are $\mathcal{F}_{t_j}$-measurable $\mathcal{L}(\mathcal{U};\mathcal{H})$-valued random variables for $j = 0,1,...,m-1$. Let $(u_n)_{n\in\mathbb{N}}$ be an orthonormal basis of $\mathcal{U}$. Using Proposition B.7, we get

$$\begin{aligned}
\mathbb{E}\left\langle I_T^L(\Phi), I_t^L(\Phi)\right\rangle_{\mathcal{H}} &= \sum_{j,k=0}^{m-1} \mathbb{E}\langle \Phi_j\left(L(t_{j+1}) - L(t_j)\right), \Phi_k\left(L(t_{k+1}\wedge t) - L(t_k \wedge t)\right)\rangle_{\mathcal{H}} \\
&= \sum_{j,k=0}^{m-1}\sum_{n=1}^{\infty} \mathbb{E}\left[\langle \Phi_k^*\Phi_j\left(L(t_{j+1}) - L(t_j)\right), u_n\rangle_{\mathcal{U}}\langle u_n, L(t_{k+1}\wedge t) - L(t_k\wedge t)\rangle_{\mathcal{U}}\right] \\
&= \sum_{j,k=0}^{m-1}\sum_{n=1}^{\infty}\sum_{l=1}^{\infty} \mathbb{E}\left[\langle L(t_{j+1}) - L(t_j), u_l\rangle_{\mathcal{U}}\langle u_l, \Phi_j^*\Phi_k u_n\rangle_{\mathcal{U}}\langle u_n, L(t_{k+1}\wedge t) - L(t_k\wedge t)\rangle_{\mathcal{U}}\right].
\end{aligned}$$

Note that there exists $j_0 \in \{0,1,...,m-1\}$ such that $t \in (t_{j_0}, t_{j_0+1}]$. Using Proposition 3.16 and the fact that the process $(L(t))_{t\geq 0}$ is $\mathcal{F}_t$-adapted, we obtain for each $k = 0,1,...,j_0$

$$\begin{aligned}
&\mathbb{E}\left[\langle L(t_{j_0+1}) - L(t_{j_0}), u_l\rangle_{\mathcal{U}}\langle u_l, \Phi_{j_0}^*\Phi_k u_n\rangle_{\mathcal{U}}\langle u_n, L(t_{k+1}\wedge t) - L(t_k\wedge t)\rangle_{\mathcal{U}}\right] \\
&= \mathbb{E}\left[\mathbb{E}[\langle L(t_{j_0+1}) - L(t_{j_0}), u_l\rangle_{\mathcal{U}}|\mathcal{F}_t]\langle u_l, \Phi_{j_0}^*\Phi_k u_n\rangle_{\mathcal{U}}\langle u_n, L(t_{k+1}\wedge t) - L(t_k\wedge t)\rangle_{\mathcal{U}}\right] \\
&= \mathbb{E}\left[\langle L(t) - L(t_{j_0}), u_l\rangle_{\mathcal{U}}\langle u_l, \Phi_{j_0}^*\Phi_k u_n\rangle_{\mathcal{U}}\langle u_n, L(t_{k+1}\wedge t) - L(t_k\wedge t)\rangle_{\mathcal{U}}\right] \\
&= \mathbb{E}\left[\langle L(t_{j_0+1}\wedge t) - L(t_{j_0}\wedge t), u_l\rangle_{\mathcal{U}}\langle u_l, \Phi_{j_0}^*\Phi_k u_n\rangle_{\mathcal{U}}\langle u_n, L(t_{k+1}\wedge t) - L(t_k\wedge t)\rangle_{\mathcal{U}}\right].
\end{aligned}$$

If $j_0 < m-1$, then $L(t_{k+1}\wedge t) - L(t_k\wedge t) = 0$ for $k > j_0$. We can conclude

$$\begin{aligned}
&\mathbb{E}\left\langle I_T^L(\Phi), I_t^L(\Phi)\right\rangle_{\mathcal{H}} \\
&= \sum_{j,k=0}^{m-1}\sum_{n=1}^{\infty}\sum_{l=1}^{\infty} \mathbb{E}\left[\langle L(t_{j+1}\wedge t) - L(t_j\wedge t), u_l\rangle_{\mathcal{U}}\langle u_l, \Phi_j^*\Phi_k u_n\rangle_{\mathcal{U}}\langle u_n, L(t_{k+1}\wedge t) - L(t_k\wedge t)\rangle_{\mathcal{U}}\right] \\
&= \sum_{j,k=0}^{m-1}\sum_{n=1}^{\infty} \mathbb{E}\left[\langle \Phi_k^*\Phi_j\left(L(t_{j+1}\wedge t) - L(t_j\wedge t)\right), u_n\rangle_{\mathcal{U}}\langle u_n, L(t_{k+1}\wedge t) - L(t_k\wedge t)\rangle_{\mathcal{U}}\right] \\
&= \sum_{j,k=0}^{m-1} \mathbb{E}\left[\langle \Phi_j\left(L(t_{j+1}\wedge t) - L(t_j\wedge t)\right), \Phi_k\left(L(t_{k+1}\wedge t) - L(t_k\wedge t)\right)\rangle_{\mathcal{H}}\right] \\
&= \mathbb{E}\left\|I_t^L(\Phi)\right\|^2_{\mathcal{H}}.
\end{aligned}$$

Using additionally Theorem 3.60, we obtain

$$
\begin{aligned}
&\mathbb{E}\left\|I_T^L(\mathbb{1}_{[0,t]}\Psi) - I_t^L(\Psi)\right\|_{\mathcal{H}}^2 \\
&\leq 2\,\mathbb{E}\left\|I_T^L(\mathbb{1}_{[0,t]}\Psi) - I_t^L(\mathbb{1}_{[0,t]}\Psi)\right\|_{\mathcal{H}}^2 + 2\,\mathbb{E}\left\|I_t^L(\mathbb{1}_{[0,t]}\Psi) - I_t^L(\Psi)\right\|_{\mathcal{H}}^2 \\
&\leq 2\,\mathbb{E}\left\|I_T^L(\mathbb{1}_{[0,t]}\Psi)\right\|_{\mathcal{H}}^2 - 4\,\mathbb{E}\left\langle I_T^L(\mathbb{1}_{[0,t]}\Psi), I_t^L(\mathbb{1}_{[0,t]}\Psi)\right\rangle_{\mathcal{H}} + 2\,\mathbb{E}\left\|I_t^L(\mathbb{1}_{[0,t]}\Psi)\right\|_{\mathcal{H}}^2 + 2\,\mathbb{E}\left\|I_t^L(\mathbb{1}_{[0,t]}\Psi - \Psi)\right\|_{\mathcal{H}}^2 \\
&\leq 2\,\mathbb{E}\left\|I_T^L(\mathbb{1}_{[0,t]}\Psi)\right\|_{\mathcal{H}}^2 - 2\,\mathbb{E}\left\|I_t^L(\mathbb{1}_{[0,t]}\Psi)\right\|_{\mathcal{H}}^2 + 2\,\mathbb{E}\left\|I_t^L(\mathbb{1}_{[0,t]}\Psi - \Psi)\right\|_{\mathcal{H}}^2 \\
&\leq 2\,\mathbb{E}\int_0^T \left\|\mathbb{1}_{[0,t]}(s)\Psi(s)\right\|_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}^2 ds - 2\,\mathbb{E}\int_0^t \left\|\mathbb{1}_{[0,t]}(s)\Psi(s)\right\|_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}^2 ds \\
&\quad + 2\,\mathbb{E}\int_0^t \left\|\mathbb{1}_{[0,t]}(s)\Psi(s) - \Psi(s)\right\|_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}^2 ds \\
&= 0.
\end{aligned}
$$

Thus, we obtain $\mathbb{P}$-a.s.

$$
\int_r^t \Psi(s)\,dL(s) = \int_0^t \Psi(s)\,dL(s) - \int_0^r \Psi(s)\,dL(s) = I_t^L(\Psi) - I_r^L(\Psi) = I_T^L(\mathbb{1}_{(r,t]}\Psi) = \int_0^T \mathbb{1}_{(r,t]}(s)\Psi(s)\,dL(s).
$$

Next, we prove (ii). Due to (i), it suffices to show the result for $t = T$. We assume that $\Psi \in \mathcal{E}_T$ has the representation (3.9). By definition of the stochastic integral, we get

$$
\mathbb{E}\left[\int_0^T \Psi(s)\,dL(s)\right] = \sum_{j=0}^{m-1} \mathbb{E}\left[\Psi_j\left(L(t_{j+1}) - L(t_j)\right)\right].
$$

Let $(h_n)_{n\in\mathbb{N}}$ and $(u_n)_{n\in\mathbb{N}}$ be orthonormal basis of $\mathcal{H}$ and $\mathcal{U}$, respectively. By Proposition B.7 and Proposition 3.16, we get for $j = 0, 1, ..., m-1$

$$
\begin{aligned}
\mathbb{E}\left[\Psi_j\left(L(t_{j+1}) - L(t_j)\right)\right] &= \mathbb{E}\left[\sum_{n=1}^{\infty}\langle \Psi_j\left(L(t_{j+1}) - L(t_j)\right), h_n\rangle_{\mathcal{H}} h_n\right] \\
&= \sum_{n=1}^{\infty} \mathbb{E}\left[\langle L(t_{j+1}) - L(t_j), \Psi_j^* h_n\rangle_{\mathcal{U}} h_n\right] \\
&= \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \mathbb{E}\left[\langle \Psi_j^* h_n, u_k\rangle_{\mathcal{U}} \langle L(t_{j+1}) - L(t_j), u_k\rangle_{\mathcal{U}}\right] h_n \\
&= \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \mathbb{E}\left[\langle \Psi_j^* h_n, u_k\rangle_{\mathcal{U}} \mathbb{E}\left[\langle L(t_{j+1}) - L(t_j), u_k\rangle_{\mathcal{U}} | \mathcal{F}_{t_j}\right]\right] h_n.
\end{aligned}
$$

Since the Lévy process $(L(t))_{t\in[0,T]}$ in an $\mathcal{U}$-valued martingale, we obtain for $j = 0, 1, ..., m-1$

$$
\mathbb{E}\left[\langle L(t_{j+1}) - L(t_j), u_k\rangle_{\mathcal{U}} | \mathcal{F}_{t_j}\right] = 0
$$

and thus, the claim (ii) holds. Note that (iii) is already stated in Theorem 3.60. A proof of (iv) can be found in [71, Theorem 8.7 (iii)]. □

The following proposition is useful when dealing with a closed operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$.

Proposition 3.63. *Let $\Psi \in \mathcal{L}_T^2$. If $\Psi(t)y \in D(A)$ for every $y \in \mathcal{U}$, all $t \in [0,T]$ and $\mathbb{P}$-almost surely,*

$$\mathbb{E}\int_0^T \|\Psi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}\,dt < \infty \quad \text{and} \quad \mathbb{E}\int_0^T \|A\Psi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}\,dt < \infty,$$

then we have $\mathbb{P}$-a.s. $\int_0^T \Psi(t)dL(t) \in D(A)$ and

$$A\int_0^T \Psi(t)\,dL(t) = \int_0^T A\Psi(t)\,dL(t).$$

Proof. One obtains the result similarly to the case of Q-Wiener processes, see [23, Proposition 4.30]. □

Next, we state a stochastic Fubini theorem. Let λ be a finite measure on a measurable space $(E,\mathcal{E})$. Recall that $\mathcal{P}_T$ denotes the smallest σ-field of subsets of $\Omega_T = [0,T]\times\Omega$ containing all sets of the form

$$(s,t]\times A \text{ with } 0\le s<t\le T, s<t, A\in\mathcal{F}_s \quad \text{and} \quad \{0\}\times A \text{ with } A\in\mathcal{F}_0.$$

Then we get the following result.

Proposition 3.64 (Theorem 8.14, [71])**.** *Assume that the mapping $(t,\omega,x)\mapsto\Psi(t,\omega,x)$ is measurable from $(\Omega_T\times E, \mathcal{P}_T\times\mathcal{B}(E))$ into $(\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H}), \mathcal{B}(\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})))$ and*

$$\int_E \mathbb{E}\int_0^T \|\Psi(t,\omega,x)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}\,dt\,d\lambda(x) < \infty.$$

Then $\mathbb{P}$-a.s.

$$\int_E \left(\int_0^T \Psi(t,\omega,x)\,dL(t)\right) d\lambda(x) = \int_0^T \left(\int_E \Psi(t,\omega,x)\,d\lambda(x)\right) dL(t).$$

Next, we introduce stochastic convolutions. Let $(S(t))_{t\ge0}$ be a C_0-semigroup on $\mathcal{H}$ and let $\Psi\in\mathcal{L}_T^2$. Then the stochastic convolution $(I(t))_{t\in[0,T]}$ given by

$$I(t) = \int_0^t S(t-s)\Psi(s)\,dL(s) \tag{3.12}$$

is well defined for all $t\in[0,T]$ and $\mathbb{P}$-almost surely. Under additional assumptions, we get the following maximal inequality.

Proposition 3.65 (cf. Proposition 1.3, [49])**.** *Let $(S(t))_{t\ge0}$ be a contraction semigroup on $\mathcal{H}$ and assume that $\Psi\in\mathcal{L}_T^2$. Then the following statements hold:*

(i) If $k\in(0,2]$, then

$$\mathbb{E}\sup_{t\in[0,T]}\left\|\int_0^t S(t-s)\Psi(s)\,dL(s)\right\|_{\mathcal{H}}^k \le \widetilde{C}_k\,\mathbb{E}\left(\int_0^T \|\Psi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}\,dt\right)^{k/2},$$

where $\widetilde{C}_k>0$ is a constant.

(ii) Assume that $(L(t))_{t\geq 0}$ has continuous trajectories. If $k \in (0,\infty)$, then

$$\mathbb{E} \sup_{t\in[0,T]} \left\| \int_0^t S(t-s)\Psi(s)\, dL(s) \right\|_{\mathcal{H}}^k \leq c_k^k\, \mathbb{E} \left(\int_0^T \|\Psi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}\, dt \right)^{k/2},$$

where $c_k > 0$ is a constant.

In order to define local mild solutions to SPDEs, we need to introduce a stopped stochastic convolution. Here, we can argue as in [16, Appendix]. Let τ be a stopping time with values in $[0,T]$. We consider the stopped process $(I(t\wedge\tau))_{t\in[0,T]}$. Unfortunately, the formula

$$I(t\wedge\tau) = \int_0^{t\wedge\tau} S(t\wedge\tau - s)\Psi(s)\, dL(s)$$

is not well defined due to the fact that we integrate a process, which is not even $\mathcal{F}_t$-adapted. To overcome this problem, we introduce a process $(I_\tau(t))_{t\in[0,T]}$ given by

$$I_\tau(t) = \int_0^t \mathbb{1}_{[0,\tau)}(s) S(t-s)\Phi(s\wedge\tau)\, dL(s) \tag{3.13}$$

for all $t \in [0,T]$ and $\mathbb{P}$-almost surely. We get the following result.

Lemma 3.66. *Let $(S(t))_{t\geq 0}$ be a C_0-semigroup on $\mathcal{H}$ and let τ be a stopping time with values in $[0,T]$. Assume that the processes $(I(t))_{t\in[0,T]}$ and $(I_\tau(t))_{t\in[0,T]}$ are given by (3.12) and (3.13), respectively. Then, we have for all $t \in [0,T]$ and $\mathbb{P}$-almost surely*

$$S(t - t\wedge\tau)I(t\wedge\tau) = I_\tau(t)$$

and in particular

$$I(t\wedge\tau) = I_\tau(t\wedge\tau).$$

Proof. The processes $(I(t))_{t\in[0,T]}$ and $(I_\tau(t))_{t\in[0,T]}$ have càdlàg modifications by [71, Theorem 9.24]. The remaining part of the proof can be obtained similarly to [16, Lemma A.1]. □

Finally, we state a product formula for infinite dimensional stochastic processes. Here, we assume that the Lévy process $(L(t))_{t\geq 0}$ is given by a Q-Wiener process. To be consistent with the notation introduced in Section 3.2.1, we denote this process by $(W(t))_{t\geq 0}$. We have the following Itô formula.

Proposition 3.67 (Theorem 4.32, [23])**.** *Assume that X^0 is an $\mathcal{F}_0$-measurable $\mathcal{H}$-valued random variable, $(f(t))_{t\in[0,T]}$ is an $\mathcal{H}$-valued $\mathcal{F}_t$-adapted process such that $\mathbb{E}\int_0^T \|f(t)\|_{\mathcal{H}} dt < \infty$ and $\Psi \in \mathcal{L}^2_T$. Let the process $(X(t))_{t\in[0,T]}$ be given by*

$$X(t) = X^0 + \int_0^t f(s)\, ds + \int_0^t \Psi(s)\, dW(s)$$

for all $t \in [0,T]$ and $\mathbb{P}$-almost surely. Assume that the function $J\colon [0,T]\times\mathcal{H} \to \mathbb{R}$ is continuous and its partial Fréchet derivatives denoted by J_t, J_x, J_{xx} are uniformly continuous on bounded subsets of $[0,T]\times\mathcal{H}$.

Then we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$J\left(t, X(t)\right) = J\left(0, X^0\right) + \int_0^t \left\langle J_x\left(s, X(s)\right), \Psi(s)\, dW(s)\right\rangle_{\mathcal{H}}$$
$$+ \int_0^t \left[J_t\left(s, X(s)\right) + \left\langle J_x\left(s, X(s)\right), f(s)\right\rangle_{\mathcal{H}} + \frac{1}{2} Tr\Big(J_{xx}\left(s, X(s)\right)(\Psi(s)Q^{1/2})(\Psi(s)Q^{1/2})^*\Big)\right] ds.$$

Remark 3.68. *For further versions of the Itô formula for infinite dimensional stochastic processes, we refer to [45, Section 2.5].*

Corollary 3.69. *For $i = 1, 2$, assume that X_i^0 are $\mathcal{F}_0$-measurable $\mathcal{H}$-valued random variables, $(f_i(t))_{t\in[0,T]}$ are $\mathcal{H}$-valued $\mathcal{F}_t$-adapted process such that $\mathbb{E}\int_0^T \|f_i(t)\|_{\mathcal{H}} dt < \infty$ and $\Psi_i \in \mathcal{L}_T^2$. For $i = 1, 2$, assume that the processes $(X_i(t))_{t\in[0,T]}$ satisfy for all $t \in [0,T]$ and $\mathbb{P}$-a.s.*

$$X_i(t) = X_i^0 + \int_0^t f_i(s)\, ds + \int_0^t \Psi_i(s)\, dW(s).$$

Then we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\langle X_1(t), X_2(t)\rangle_{\mathcal{H}} = \left\langle X_1^0, X_2^0\right\rangle_{\mathcal{H}} + \int_0^t \Big[\langle X_1(s), f_2(s)\rangle_{\mathcal{H}} + \langle X_2(s), f_1(s)\rangle_{\mathcal{H}} + \langle \Psi_1(s), \Psi_2(s)\rangle_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{U});\mathcal{H})}\Big]\, ds$$
$$+ \int_0^t \langle X_1(s), \Psi_2(s)\, dW(s)\rangle_{\mathcal{H}} + \int_0^t \langle X_2(s), \Psi_1(s)\, dW(s)\rangle_{\mathcal{H}}.$$

Proof. The claim follows from Proposition 3.67 with $J\colon \mathcal{H} \times \mathcal{H} \to \mathbb{R}$ given by $J(x_1, x_2) = \langle x_1, x_2\rangle_{\mathcal{H}}$. □

3.4. Stochastic Partial Differential Equations

In this section, we prove existence and uniqueness results for SPDEs both of forward and backward type, which we deal with in the following chapters. Here, we will mainly concentrate on mild solutions to SPDEs. For forward equations, the proof of the existence and uniqueness of mild solutions is based on the Banach fixed point theorem, see [23, 42, 71, 73]. Mild solutions of backward equations require a martingale representation theorem, see [52]. Furthermore, we will also show the relationship to different concepts of solutions. Throughout this section, we assume that the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is equipped with a normal filtration $(\mathcal{F}_t)_{t\geq 0}$.

3.4.1. Forward Stochastic Partial Differential Equations

Here, we prove existence and uniqueness results of forward SPDEs motivated by systems arising in stochastic control problems. We study systems on bounded domain with sufficiently smooth boundary, where we also involve nonhomogeneous boundary data. Therefore, we introduce two separable Hilbert spaces $\mathcal{H}$ and $\mathcal{H}_b$, where $\mathcal{H}$ refers to data defined inside the domain and $\mathcal{H}_b$ refers to data defined on the boundary. We start with the following linear system in $\mathcal{H}$:

$$\begin{cases} dy(t) = [Ay(t) + Bu(t) + (\lambda - A)N_1 v(t)]\, dt + G(t)\, dL(t) + (\lambda - A)N_2\, dL_b(t), \\ y(0) = \xi. \end{cases} \tag{3.14}$$

We assume that

- the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ is the generator of an analytic semigroup of contractions $(e^{At})_{t\geq 0}$ and $\lambda > 0$ is an element of the resolvent set $\rho(A)$;
- the process $(u(t))_{t\in[0,T]}$ is $\mathcal{F}_t$-adapted and takes values in $\mathcal{H}$ such that
$$\mathbb{E}\int_0^T \|u(t)\|^2_{\mathcal{H}}\,dt < \infty;$$
- $B \in \mathcal{L}(\mathcal{H})$;
- $(L(t))_{t\geq 0}$ is an $\mathcal{H}$-valued square integrable Lévy martingale with covariance operator $Q \in \mathcal{L}_1^+(\mathcal{H})$;
- $(G(t))_{t\in[0,T]}$ is a predictable process with values in $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})$ such that
$$\mathbb{E}\int_0^T \|G(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}dt < \infty;$$
- the process $(v(t))_{t\in[0,T]}$ is $\mathcal{F}_t$-adapted and takes values in $\mathcal{H}_b$ such that
$$\mathbb{E}\int_0^T \|v(t)\|^2_{\mathcal{H}_b}\,dt < \infty;$$
- $(L_b(t))_{t\geq 0}$ is an $\mathcal{H}_b$-valued square integrable Lévy martingale with covariance operator $Q_b \in \mathcal{L}_1^+(\mathcal{H}_b)$;
- $N_1, N_2 \in \mathcal{L}(\mathcal{H}_b; D((\lambda - A)^\alpha))$ for $\alpha \in (0, \frac{3}{4})$;
- ξ is an $\mathcal{F}_0$-measurable random variable with values in $\mathcal{H}$.

Note that the operator $A - \lambda$ is still generator of an analytic semigroup given by $(e^{-\lambda t}e^{At})_{t\geq 0}$, see [70, Chapter 3, Corollary 2.2]. Hence, the operator $A - \lambda$ satisfies the assumptions of Remark 2.25 with $M = 1$ and $\theta = \lambda$. Therefore, we can define fractional powers of the operator $\lambda - A$ denoted by $(\lambda - A)^\alpha$ with $\alpha \in \mathbb{R}$ according to Section 2.3, respectively. Moreover, we have $0 \in \rho(A - \lambda)$. Thus, we get the following properties.

Corollary 3.70. *We have*

- $(\lambda - A)^{\alpha+\beta}y = (\lambda - A)^\alpha(\lambda - A)^\beta y$ *for all* $\alpha, \beta \in \mathbb{R}$ *and every* $y \in D(A^\gamma)$ *with* $\gamma = \max\{\alpha, \beta, \alpha + \beta\}$*;*
- $e^{At}\colon \mathcal{H} \to D((\lambda - A)^\alpha)$ *for all* $t > 0$ *and all* $\alpha \in \mathbb{R}$*;*
- $(\lambda - A)^\alpha e^{At}y = e^{At}(\lambda - A)^\alpha y$ *for every* $y \in D((\lambda - A)^\alpha)$ *and all* $\alpha \in \mathbb{R}$*;*
- *the operator* $(\lambda - A)^\alpha e^{At}$ *is linear and bounded for all* $t > 0$ *and all* $\alpha \in \mathbb{R}$*. In addition, there exist constants* $M_\alpha, \delta > 0$ *such that for all* $t > 0$ *and all* $\alpha > 0$
$$\|(\lambda - A)^\alpha e^{At}\|_{\mathcal{H}} \leq M_\alpha t^{-\alpha}e^{-\delta t}.$$

Proof. The assertions follow immediately from Theorem 2.29 (iv) and Theorem 2.35. □

Remark 3.71. *In control theory, system (3.14) arises for the controlled stochastic heat equation with nonhomogeneous Neumann boundary conditions. Then the operator A refers the Neumann realization of the Laplace operator introduced in Section 2.5.1. Moreover, the term $u(t)$ is a distributed control and $L(t)$ is a Lévy noise defined inside the domain. Similarly, the term $v(t)$ is a boundary control and $L_b(t)$ is a Lévy noise defined on the boundary. The operators N_1, N_2 belong to the Neumann operator mapping the boundary data inside the domain. Typically, we have $N_1 = N_2$.*

Definition 3.72. *A predictable process $(y(t))_{t\in[0,T]}$ with values in $\mathcal{H}$ is called a* ***mild solution of system (3.14)*** *if*

$$\sup_{t\in[0,T]} \mathbb{E}\|y(t)\|_{\mathcal{H}}^2 < \infty$$

and for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} y(t) = e^{At}\xi + \int_0^t e^{A(t-s)}Bu(s)\,ds + \int_0^t (\lambda - A)e^{A(t-s)}N_1 v(s)\,ds + \int_0^t e^{A(t-s)}G(s)\,dL(s) \\ + \int_0^t (\lambda - A)e^{A(t-s)}N_2\,dL_b(s). \end{aligned}$$

Theorem 3.73. *Let $(u(t))_{t\in[0,T]}$ and $(v(t))_{t\in[0,T]}$ be fixed. For any $\xi \in L^2(\Omega;\mathcal{H})$, there exists a unique mild solution $(y(t))_{t\in[0,T]}$ of system (3.14). Moreover, the process $(y(t))_{t\in[0,T]}$ is mean square continuous.*

Proof. By definition, the mild solution of system (3.14) is unique. Next, we show that $(y(t))_{t\in[0,T]}$ takes values in $\mathcal{H}$ such that $\sup_{t\in[0,T]} \mathbb{E}\|y(t)\|_{\mathcal{H}}^2 < \infty$. We define for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\psi_1(t) = e^{At}\xi + \int_0^t e^{A(t-s)}Bu(s)\,ds, \quad \psi_2(t) = \int_0^t (\lambda - A)e^{A(t-s)}N_1 v(s)\,ds + \int_0^t (\lambda - A)e^{A(t-s)}N_2\,dL_b(s),$$

$$\psi_3(t) = \int_0^t e^{A(t-s)}G(s)\,dL(s).$$

Recall that $\left\|e^{-At}\right\|_{\mathcal{L}(\mathcal{H})} \leq 1$ for all $t \geq 0$ and $B \in \mathcal{L}(\mathcal{H})$. Hence, the process $(\psi_1(t))_{t\in[0,T]}$ takes values in $\mathcal{H}$ and there exists a constant $C_1 > 0$ such that

$$\begin{aligned} \sup_{t\in[0,T]} \mathbb{E}\,\|\psi_1(t)\|_{\mathcal{H}}^2 &\leq 2 \sup_{t\in[0,T]} \mathbb{E}\left\|e^{At}\xi\right\|_{\mathcal{H}}^2 + 2 \sup_{t\in[0,T]} \mathbb{E} \int_0^t \left\|e^{A(t-s)}Bu(s)\right\|_{\mathcal{H}}^2 ds \\ &\leq C_1 \left[\mathbb{E}\|\xi\|_{\mathcal{H}}^2 + \mathbb{E}\int_0^T \|u(t)\|_{\mathcal{H}}^2\,dt \right]. \end{aligned}$$

Since $N_1, N_2 \in \mathcal{L}(\mathcal{H}_b; D((\lambda - A)^\alpha))$ for $\alpha \in (0, \frac{3}{4})$, we get $(\lambda - A)^\alpha N_1, (\lambda - A)^\alpha N_2 \in \mathcal{L}(\mathcal{H}_b;\mathcal{H})$ by the closed graph theorem. By Theorem 3.62 (iii), Corollary 3.70 and the Cauchy-Schwarz inequality, the process

$(\psi_2(t))_{t\in[0,T]}$ takes values in $\mathcal{H}$ and there exists a constant $C_2 > 0$ such that for all $\alpha \in (\frac{1}{2}, \frac{3}{4})$

$$\begin{aligned}
\sup_{t\in[0,T]} \mathbb{E}\,\|\psi_2(t)\|_{\mathcal{H}}^2 &\leq 2 \sup_{t\in[0,T]} \mathbb{E} \left(\int_0^t \left\| (\lambda - A)^{1-\alpha} e^{A(t-s)} (\lambda - A)^{\alpha} N_1 v(s) \right\|_{\mathcal{H}} ds \right)^2 \\
&\quad + 2 \sup_{t\in[0,T]} \mathbb{E} \left\| \int_0^t (\lambda - A)^{1-\alpha} e^{A(t-s)} (\lambda - A)^{\alpha} N_2 \, dL_b(s) \right\|_{\mathcal{H}}^2 \\
&\leq 2M_{1-\alpha}^2 \sup_{t\in[0,T]} \mathbb{E} \left(\int_0^t (t-s)^{\alpha-1} \|(\lambda - A)^{\alpha} N_1 v(s)\|_{\mathcal{H}} \, ds \right)^2 \\
&\quad + 2M_{1-\alpha}^2 \|(\lambda - A)^{\alpha} N_2\|^2_{\mathcal{L}_{(HS)}(Q_b^{1/2}(\mathcal{H}_b);\mathcal{H})} \sup_{t\in[0,T]} \int_0^t (t-s)^{2\alpha-2} ds \\
&\leq C_2 \left[1 + \mathbb{E} \int_0^T \|v(t)\|^2_{\mathcal{H}_b} \, dt \right].
\end{aligned}$$

Using Theorem 3.62 (iii) and Fubini's theorem, the process $(\psi_3(t))_{t\in[0,T]}$ takes values in $\mathcal{H}$ such that

$$\sup_{t\in[0,T]} \mathbb{E}\,\|\psi_3(t)\|_{\mathcal{H}}^2 \leq \sup_{t\in[0,T]} \mathbb{E} \int_0^t \left\| e^{A(t-s)} G(s) \right\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} ds \leq \mathbb{E} \int_0^T \|G(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} \, dt.$$

Next, we prove that the process $(y(t))_{t\in[0,T]}$ is mean square continuous. We assume w.l.o.g. $0 \leq t_0 \leq t \leq T$. Let I be the identity operator on $\mathcal{H}$. By the Cauchy-Schwarz inequality, there exists a constant $c_1 > 0$ such that

$$\begin{aligned}
\mathbb{E}\,\|\psi_1(t) - \psi_1(t_0)\|_{\mathcal{H}}^2 &\leq 3\,\mathbb{E} \left\| \left(e^{A(t-t_0)} - I \right) e^{At_0} \xi \right\|_{\mathcal{H}}^2 + 3\,\mathbb{E} \left\| \int_0^{t_0} \left(e^{A(t-t_0)} - I \right) e^{A(t_0-s)} Bu(s) \, ds \right\|_{\mathcal{H}}^2 \\
&\quad + 3\,\mathbb{E} \left\| \int_{t_0}^t e^{A(t-s)} Bu(s) \, ds \right\|_{\mathcal{H}}^2 \\
&\leq 3\,\mathbb{E} \left\| \left(e^{A(t-t_0)} - I \right) e^{At_0} \xi \right\|_{\mathcal{H}}^2 + 3\,\mathbb{E} \left\| \left(e^{A(t-t_0)} - I \right) \int_0^{t_0} e^{A(t_0-s)} Bu(s) \, ds \right\|_{\mathcal{H}}^2 \\
&\quad + c_1 (t - t_0)\, \mathbb{E} \int_0^T \|u(t)\|_{\mathcal{H}}^2 \, dt.
\end{aligned}$$

Due to Corollary 3.70, Theorem 3.62 (i) and (iii) and the Cauchy-Schwarz inequality, there exists a constant

$c_2 > 0$ such that

$$\begin{aligned}
&\mathbb{E}\,\|\psi_2(t)-\psi_2(t_0)\|^2_{\mathcal{H}} \\
&\leq 4\,\mathbb{E}\left\|\int_0^{t_0}\left(e^{A(t-t_0)}-I\right)(\lambda-A)e^{A(t_0-s)}N_1v(s)\,ds\right\|^2_{\mathcal{H}} + 4\,\mathbb{E}\left(\int_{t_0}^{t}\left\|(\lambda-A)e^{A(t-s)}N_1v(s)\right\|_{\mathcal{H}}ds\right)^2 \\
&\quad + 4\,\mathbb{E}\left\|\int_0^{t_0}\left(e^{A(t-t_0)}-I\right)(\lambda-A)e^{A(t_0-s)}N_2\,dL_b(s)\right\|^2_{\mathcal{H}} + 4\,\mathbb{E}\left\|\int_{t_0}^{t}(\lambda-A)e^{A(t-s)}N_2\,dL_b(s)\right\|^2_{\mathcal{H}} \\
&\leq 4\,\mathbb{E}\left\|\left(e^{A(t-t_0)}-I\right)\int_0^{t_0}(\lambda-A)e^{A(t_0-s)}N_1v(s)\,ds\right\|^2_{\mathcal{H}} + c_2(t-t_0)^{2\alpha-1}\mathbb{E}\int_0^T\|v(t)\|^2_{\mathcal{H}_b}\,dt \\
&\quad + 4\,\mathbb{E}\left\|\left(e^{A(t-t_0)}-I\right)\int_0^{t_0}(\lambda-A)e^{A(t_0-s)}N_2\,dL_b(s)\right\|^2_{\mathcal{H}} + c_2(t-t_0)^{2\alpha-1}.
\end{aligned}$$

Let $(h_n)_{n\in\mathbb{N}}$ be an orthonormal basis in $\mathcal{H}$. Using Theorem 3.62 (i) and (iii), we obtain

$$\begin{aligned}
&\mathbb{E}\,\|\psi_3(\tilde{y})(t)-\psi_3(\tilde{y})(t_0)\|^2_{\mathcal{H}} \\
&\leq 2\,\mathbb{E}\left\|\int_0^{t_0}\left(e^{A(t-t_0)}-I\right)e^{A(t_0-s)}G(s)\,dL(s)\right\|^2_{\mathcal{H}} + 2\,\mathbb{E}\left\|\int_{t_0}^{t}e^{A(t-s)}G(s)\,dL(s)\right\|^2_{\mathcal{H}} \\
&\leq 2\,\mathbb{E}\left\|\left(e^{A(t-t_0)}-I\right)\int_0^{t_0}e^{A(t_0-s)}G(s)\,dL(s)\right\|^2_{\mathcal{H}} + 2\,\mathbb{E}\int_{t_0}^{t}\|G(s)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\,ds.
\end{aligned}$$

Note that $\lim_{t\to t_0}\|e^{-A(t-t_0)}h-h\|_{\mathcal{H}}=0$ holds for every $h\in\mathcal{H}$. Using Corollary B.6 and Proposition B.7, we can infer that the process $(y(t))_{t\in[0,T]}$ is mean square continuous. Moreover, the process $(y(t))_{t\in[0,T]}$ is obviously $\mathcal{F}_t$-adapted. Hence, the process $(y(t))_{t\in[0,T]}$ has a predictable modification resulting from Proposition 3.9. □

Remark 3.74. *Let the process $(G(t))_{t\in[0,T]}$ be time independent, i.e. $G(t)=G$ for all $t\in[0,T]$ and $\mathbb{P}$-almost surely, where G is a square integrable random variable with values in $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})$. If $\alpha\in(\frac{1}{2},\frac{3}{4})$, then the mild solution $(y(t))_{t\in[0,T]}$ of system (3.14) takes values in $D((\lambda-A)^\beta)$ with $\beta\in[0,\frac{3}{4}-\alpha)$ such that*

$$\sup_{t\in[0,T]}\mathbb{E}\|y(t)\|^2_{D((\lambda-A)^\beta)}<\infty.$$

Remark 3.75. *The mild solution $(y(t))_{t\in[0,T]}$ of system (3.14) has also a càdlàg modification. One can argue as in [71, Theorem 9.24]. Since $(e^{At})_{t\geq 0}$ is a C_0 semigroup of contractions, we can apply Theorem 2.14. Thus, there exists a Hilbert space $\widehat{\mathcal{H}}$ containing $\mathcal{H}$ and a group $(\widehat{S}(t))_{t\in\mathbb{R}}$ on $\widehat{\mathcal{H}}$ such that $e^{At}=P_{\mathcal{H}}\widehat{S}(t)$ for all $t\in\mathbb{R}$, where $P_{\mathcal{H}}$ is the orthogonal projection from $\widehat{\mathcal{H}}$ onto $\mathcal{H}$. Note that $P_{\mathcal{H}}\widehat{S}(t)\colon\mathcal{H}\to D((\lambda-A)^\alpha)$*

for all $t > 0$ and all $\alpha \in \mathbb{R}$ due to Corollary 3.70. Using Proposition 3.63, we get for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\int_0^t e^{A(t-s)} G(s)\, dL(s) = P_{\mathcal{H}} \widehat{S}(t) \int_0^t \widehat{S}(-s) G(s)\, dL(s),$$

$$\int_0^t (\lambda - A) e^{A(t-s)} N_2\, dL_b(s) = (\lambda - A)^{1-\alpha} P_{\mathcal{H}} \widehat{S}(t) \int_0^t \widehat{S}(-s)(\lambda - A)^{\alpha} N_2\, dL_b(s).$$

We set $X(t) = \int_0^t \widehat{S}(-s)G(s)\, dL(s)$ and $X_b(t) = \int_0^t \widehat{S}(-s)(\lambda - A)^{\alpha} N_2\, dL_b(s)$ for all $t \in [0,T]$ and $\mathbb{P}$-almost surely. By Theorem 3.62 (iv), the processes $(X(t))_{t\in[0,T]}$ and $(X_b(t))_{t\in[0,T]}$ are mean square continuous $\widehat{\mathcal{H}}$-valued martingales. Therefore, the processes $(X(t))_{t\in[0,T]}$ and $(X_b(t))_{t\in[0,T]}$ have càdlàg modifications as a consequence of Theorem 3.19. Since the mapping $t \mapsto \widehat{S}(t)x$ is continuous from $\mathbb{R}\backslash\{0\}$ into $\widehat{\mathcal{H}}$ for every $x \in \widehat{\mathcal{H}}$, we can conclude that the processes $(P_{\mathcal{H}}\widehat{S}(t)X(t))_{t\in[0,T]}$ and $((\lambda - A)^{1-\alpha} P_{\mathcal{H}} \widehat{S}(t) X_b(t))_{t\in[0,T]}$ have càdlàg modifications.

Remark 3.76. *Let $(W(t))_{t\geq 0}$ be a Q-Wiener process and let $\Psi \in \mathcal{L}_T^2$. Then it is well known that the process $(X(t))_{t\in[0,T]}$ given by*

$$X(t) = \int_0^t \Psi(s)\, dW(s)$$

for all $t \in [0,T]$ and $\mathbb{P}$-a.s. is continuous, see [23, Section 4.2]. Therefore, the mild solution $(y(t))_{t\in[0,T]}$ of system (3.14) has a continuous modification if $(L(t))_{t\geq 0}$ and $(L_b(t))_{t\geq 0}$ are Q-Wiener processes. The assertion can be obtained similarly to the previous Remark.

Next, we consider the following linear system on $D(A^{\alpha})$:

$$\begin{cases} dy(t) = [-Ay(t) + Bu(t) + ADv(t)]\, dt + G(y(t))\, dL(t), \\ y(0) = \xi. \end{cases} \tag{3.15}$$

We assume that

- the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ is linear and closed such that $-A$ is the generator of an analytic semigroup of contractions $(e^{-At})_{t\geq 0}$ and 0 is an element of the resolvent set $\rho(A)$;
- the process $(u(t))_{t\in[0,T]}$ is predictable and takes values in $\mathcal{H}$ such that
$$\mathbb{E}\int_0^T \|u(t)\|_{\mathcal{H}}^2\, dt < \infty;$$
- $B \in \mathcal{L}(\mathcal{H})$;
- $(L(t))_{t\geq 0}$ is an $\mathcal{H}$-valued square integrable Lévy martingale with covariance operator $Q \in \mathcal{L}_1^+(\mathcal{H})$;
- $G\colon \mathcal{H} \to \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})$ is linear and bounded;
- the process $(v(t))_{t\in[0,T]}$ is predictable and takes values in $\mathcal{H}_b$ such that
$$\mathbb{E}\int_0^T \|v(t)\|_{\mathcal{H}_b}^2\, dt < \infty;$$

- $D \in \mathcal{L}(\mathcal{H}_b; D(A^\beta))$ for $\beta \in (0, \frac{1}{4})$;
- ξ is an $\mathcal{F}_0$-measurable random variable with values in $\mathcal{H}$.

Remark 3.77. *In control theory, system (3.15) arises for the controlled stochastic Stokes equations with nonhomogeneous Dirichlet boundary conditions. Then the operator A refers to the Stokes operator introduced in Section 2.5.2. Moreover, the term $u(t)$ is a distributed control and $L(t)$ is a Lévy noise defined inside the domain. The term $v(t)$ is a boundary control and D denotes the Dirichlet operator mapping the boundary data inside the domain.*

Definition 3.78. *A predictable process $(y(t))_{t\in[0,T]}$ with values in $D(A^\alpha)$ is called a* ***mild solution of system (3.15)*** *if*

$$\mathbb{E}\int_0^T \|y(t)\|^2_{D(A^\alpha)} dt < \infty$$

and for $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$y(t) = e^{-At}\xi + \int_0^t e^{-A(t-s)}Bu(s)\,ds + \int_0^t Ae^{-A(t-s)}Dv(s)\,ds + \int_0^t e^{-A(t-s)}G(y(s))\,dL(s).$$

Theorem 3.79. *Let $(u(t))_{t\in[0,T]}$ and $(v(t))_{t\in[0,T]}$ be fixed. If $\alpha \in [0, \frac{1}{4})$ and $\beta \in (0, \frac{1}{4} - \alpha)$, then for any $\xi \in L^2(\Omega; D(A^\alpha))$, there exists a unique mild solution $(y(t))_{t\in[0,T]}$ of system (3.15).*

Proof. For all $t_0, t_1 \in [0,T]$ with $t_0 < t_1$, let the space $\mathcal{Z}_{[t_0,t_1]}$ contain all predictable processes $(\tilde{y}(t))_{t\in[t_0,t_1]}$ with values in $D(A^\alpha)$ such that $\mathbb{E}\int_{t_0}^{t_1} \|\tilde{y}(t)\|^2_{D(A^\alpha)} dt < \infty$. The space $\mathcal{Z}_{[t_0,t_1]}$ equipped with the inner product

$$\langle \tilde{y}_1, \tilde{y}_2\rangle_{\mathcal{Z}_{[t_0,t_1]}} = \mathbb{E}\int_{t_0}^{t_1} \langle \tilde{y}_1(t), \tilde{y}_2(t)\rangle_{D(A^\alpha)} dt$$

for every $\tilde{y}_1, \tilde{y}_2 \in \mathcal{Z}_{[t_0,t_1]}$ becomes a Hilbert space. We define for $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\mathcal{J}(\tilde{y})(t) = e^{-At}\xi + \int_0^t e^{-A(t-s)}Bu(s)\,ds + \int_0^t Ae^{-A(t-s)}Dv(s)\,ds + \int_0^t e^{-A(t-s)}G(\tilde{y}(s))\,dL(s).$$

Let $T_1 \in (0,T]$ and let us denote by $\mathcal{Z}_{T_1}$ the space $\mathcal{Z}_{[0,T_1]}$. First, we prove that $\mathcal{J}$ maps $\mathcal{Z}_{T_1}$ into itself. We define for $t \in [0,T_1]$ and $\mathbb{P}$-a.s.

$$\psi_1(t) = e^{-At}\xi + \int_0^t e^{-A(t-s)}Bu(s)\,ds, \quad \psi_2(t) = \int_0^t Ae^{-A(t-s)}Dv(s)\,ds,$$

$$\psi_3(\tilde{y})(t) = \int_0^t e^{-A(t-s)}G(\tilde{y}(s))\,dL(s).$$

Recall that $\left\|e^{-At}\right\|_{\mathcal{L}(\mathcal{H})} \leq 1$ for all $t \geq 0$ and $B \in \mathcal{L}(\mathcal{H})$. Using Theorem 2.35, Proposition B.9 and the Cauchy-Schwarz inequality, the process $(\psi_1(t))_{t\in[0,T_1]}$ takes values in $D(A^\alpha)$ and there exists a constant

$C_1 > 0$ such that

$$\begin{aligned}\mathbb{E}\int_0^{T_1} \|\psi_1(t)\|^2_{D(A^\alpha)}\,dt &\le 2\,\mathbb{E}\int_0^{T_1} \left\|e^{-At}A^\alpha\xi\right\|^2_{\mathcal{H}}\,dt + 2\,\mathbb{E}\int_0^{T_1}\left(\int_0^t \left\|A^\alpha e^{-A(t-s)}Bu(s)\right\|_{\mathcal{H}}\,ds\right)^2 dt\\ &\le 2T_1\,\mathbb{E}\,\|\xi\|^2_{D(A^\alpha)} + 2M_\alpha^2\,\mathbb{E}\int_0^{T_1}\left(\int_0^t (t-s)^{-\alpha}\,\|Bu(s)\|_{\mathcal{H}}\,ds\right)^2 dt\\ &\le C_1\left[\mathbb{E}\|\xi\|^2_{D(A^\alpha)} + \mathbb{E}\int_0^{T_1}\|u(t)\|^2_{\mathcal{H}}\,dt\right].\end{aligned}$$

Since $D \in \mathcal{L}(\mathcal{H}_b; D(A^{\alpha+\beta}))$ for $\alpha+\beta \in (0, \frac{1}{4})$, we get $A^{\alpha+\beta}D \in \mathcal{L}(\mathcal{H}_b;\mathcal{H})$ by the closed graph theorem. By Theorem 2.29 (iv), Theorem 2.35, Proposition B.9 and Young's inequality for convolutions, the process $(\psi_2(t))_{t\in[0,T_1]}$ takes values in $D(A^\alpha)$ and there exists a constant $C_2 > 0$ such that

$$\begin{aligned}\mathbb{E}\int_0^{T_1}\|\psi_2(t)\|^2_{D(A^\alpha)}\,dt &\le 2\,\mathbb{E}\int_0^{T_1}\left(\int_0^t \left\|A^{1-\beta}e^{-A(t-s)}A^{\alpha+\beta}Dv(s)\right\|_{\mathcal{H}}\,ds\right)^2 dt\\ &\le M_{1-\beta}^2\mathbb{E}\int_0^{T_1}\left(\int_0^t (t-s)^{\beta-1}\left\|A^{\alpha+\beta}Dv(s)\right\|_{\mathcal{H}}\,ds\right)^2 dt\\ &\le M_{1-\beta}^2\left(\int_0^{T_1} t^{\beta-1}dt\right)^2\,\mathbb{E}\int_0^{T_1}\left\|A^{\alpha+\beta}Dv(t)\right\|^2_{\mathcal{H}}\,dt\\ &\le C_2\,\mathbb{E}\int_0^{T_1}\|v(t)\|^2_{\mathcal{H}_b}\,dt.\end{aligned}$$

Due to Theorem 2.35 and since $G\colon \mathcal{H}\to\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})$ is linear and bounded, one can verify the assumptions of Proposition 3.63. Hence, the process $(\psi_3(\tilde{y})(t))_{t\in[0,T_1]}$ takes values in $D(A^\alpha)$. Using Fubini's theorem, Theorem 3.62 (iii), Theorem 2.35, Young's inequality for convolutions and Corollary 2.32, there exists a constant $C_3 > 0$ such that

$$\begin{aligned}\mathbb{E}\int_0^{T_1}\|\psi_3(\tilde{y})(t)\|^2_{D(A^\alpha)}\,dt &= \int_0^{T_1}\mathbb{E}\left\|\int_0^t A^\alpha e^{-A(t-s)}G(\tilde{y}(s))\,dL(s)\right\|^2_{\mathcal{H}} dt\\ &\le M_\alpha^2\,\mathbb{E}\int_0^{T_1}\int_0^t (t-s)^{-2\alpha}\,\|G(\tilde{y}(s))\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\,ds\,dt\\ &\le C_3T_1^{1-2\alpha}\,\mathbb{E}\int_0^{T_1}\|\tilde{y}(t)\|^2_{D(A^\alpha)}dt. \end{aligned}\tag{3.16}$$

Hence, we can conclude that for fixed $\tilde{y} \in \mathcal{Z}_{T_1}$, the process $(\mathcal{J}(\tilde{y})(t))_{t\in[0,T_1]}$ takes values in $D(A^\alpha)$ such that $\mathbb{E}\int_0^{T_1}\|\mathcal{J}(\tilde{y})(t)\|^2_{D(A^\alpha)}dt < \infty$. Obviously, the process $(\mathcal{J}(\tilde{y})(t))_{t\in[0,T_1]}$ is predictable. Therefore, we can infer that $\mathcal{J}$ maps $\mathcal{Z}_{T_1}$ into itself.

Next, we show that $\mathcal{J}$ is a contraction on $\mathcal{Z}_{T_1}$. Recall that $G\colon \mathcal{H} \to \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})$ is linear. Using inequality (3.16), we get for every $\tilde{y}_1, \tilde{y}_2 \in \mathcal{Z}_{T_1}$

$$\mathbb{E}\int_0^{T_1} \|\mathcal{J}(\tilde{y}_1)(t) - \mathcal{J}(\tilde{y}_2)(t)\|^2_{D(A^\alpha)}\, dt = \mathbb{E}\int_0^{T_1} \|\psi_3(\tilde{y}_1 - \tilde{y}_2)(t)\|^2_{D(A^\alpha)}\, dt \le C_3 T_1^{1-2\alpha}\, \mathbb{E}\int_0^{T_1} \|\tilde{y}_1(t) - \tilde{y}_2(t)\|^2_{D(A^\alpha)} dt.$$

We choose $T_1 \in (0,T]$ such that $C_3 T_1^{1-2\alpha} < 1$. Applying the Banach fixed point theorem, we get a unique element $y \in \mathcal{Z}_{T_1}$ such that for $t \in [0,T_1]$ and $\mathbb{P}$-a.s. $y(t) = \mathcal{J}(y)(t)$.

Next, we consider for $t \in [T_1,T]$ and $\mathbb{P}$-a.s.

$$\mathcal{J}(\tilde{y})(t) = e^{-A(t-T_1)} y(T_1) + \int_{T_1}^t e^{-A(t-s)} Bu(s)\, ds + \int_{T_1}^t A e^{-A(t-s)} Dv(s)\, ds + \int_{T_1}^t e^{-A(t-s)} G(\tilde{y}(s))\, dL(s).$$

Again, for a certain $T_2 \in [T_1,T]$, there exists a unique fixed point of $\mathcal{J}$ on $\mathcal{Z}_{[T_1,T_2]}$. By continuing the method, we get the existence and uniqueness of a predictable process $(y(t))_{t\in[0,T]}$ satisfying for $t \in [0,T]$ and $\mathbb{P}$-a.s. $y(t) = \mathcal{J}(y)(t)$. □

Next, we consider the following nonlinear system in $D(A^\alpha)$:

$$\begin{cases} dy(t) = -\left[Ay(t) + B(y(t)) - Fu(t)\right] dt + G(y(t))\, dL(t), \\ y(0) = \xi. \end{cases} \tag{3.17}$$

We assume that

- the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ is linear and closed such that $-A$ is the generator of an analytic semigroup of contractions $(e^{-At})_{t\ge 0}$ and 0 is an element of the resolvent set $\rho(A)$;
- there exists $\alpha, \delta \in [0,1)$ and a constant $C > 0$ such that for every $y, z \in D(A^\alpha)$

$$\|A^{-\delta} B(y)\|_{\mathcal{H}} \le C\|y\|_{D(A^\alpha)}, \tag{3.18}$$

$$\|A^{-\delta}(B(y) - B(z))\|_{\mathcal{H}} \le C\|y - z\|_{D(A^\alpha)}; \tag{3.19}$$

- the process $(u(t))_{t\in[0,T]}$ is $\mathcal{F}_t$-adapted and takes values in $D(A^\beta)$, $\beta \in [0,\alpha]$, such that

$$\mathbb{E}\int_0^T \|u(t)\|^2_{D(A^\beta)}\, dt < \infty;$$

- $F \in \mathcal{L}(D(A^\beta))$;
- $(L(t))_{t\ge 0}$ is a square integrable Lévy martingale with values in $\mathcal{H}$ and covariance operator $Q \in \mathcal{L}_1^+(\mathcal{H})$;
- $G\colon \mathcal{H} \to \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); D(A^\alpha))$ satisfies for every $y, z \in \mathcal{H}$

$$\|G(y)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});D(A^\alpha))} \le \widehat{C}\|y\|_{\mathcal{H}}, \tag{3.20}$$

$$\|G(y) - G(z)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});D(A^\alpha))} \le \widehat{C}\|y - z\|_{\mathcal{H}}, \tag{3.21}$$

  where $\widehat{C} > 0$ is a constant;

- ξ is an $\mathcal{F}_0$-measurable random variable with values in $\mathcal{H}$.

Remark 3.80. *In control theory, system (3.17) arises for the controlled stochastic Navier-Stokes equations with homogeneous Dirichlet boundary conditions. Then the operator A refers to the Stokes operator introduced in Section 2.5.2. The operator B is related to the convection term. Moreover, the term $u(t)$ is a distributed control and $L(t)$ is a Lévy noise defined inside the domain.*

Definition 3.81. *A predictable process $(y(t))_{t\in[0,T]}$ with values in $D(A^\alpha)$ is called a* ***mild solution of system (3.17)*** *if*

$$\mathbb{E} \sup_{t\in[0,T]} \|y(t)\|^2_{D(A^\alpha)} < \infty \tag{3.22}$$

and for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$y(t) = e^{-At}\xi - \int_0^t A^\delta e^{-A(t-s)} A^{-\delta} B(y(s))\, ds + \int_0^t e^{-A(t-s)} Fu(s)\, ds + \int_0^t e^{-A(t-s)} G(y(s))\, dL(s).$$

The main difficulty is the case $\alpha+\delta > \frac{1}{2}$. For that reason, the strong regularity property (3.22) is required.

Theorem 3.82. *Let the parameters $\alpha, \delta \in [0,1)$ satisfy $\alpha + \delta < 1$. Moreover, let $(u(t))_{t\in[0,T]}$ be fixed with $\beta \in [0,\alpha)$ such that $\alpha - \beta < \frac{1}{2}$. Then for any $\xi \in L^2(\Omega; D(A^\alpha))$, there exists a unique mild solution $(y(t))_{t\in[0,T]}$ of system (3.17). Moreover, the process $(y(t))_{t\in[0,T]}$ is mean square continuous.*

Proof. For all $t_0, t_1 \in [0,T]$ with $t_0 < t_1$, let the space $\mathcal{Z}_{[t_0,t_1]}$ contain all predictable processes $(\tilde{y}(t))_{t\in[t_0,t_1]}$ with values in $D(A^\alpha)$ such that $\mathbb{E}\sup_{t\in[t_0,t_1]} \|\tilde{y}(t)\|^2_{D(A^\alpha)} < \infty$. The space $\mathcal{Z}_{[t_0,t_1]}$ equipped with the norm

$$\|\tilde{y}\|^2_{\mathcal{Z}_{[t_0,t_1]}} = \mathbb{E} \sup_{t\in[t_0,t_1]} \|\tilde{y}(t)\|^2_{D(A^\alpha)}$$

for every $\tilde{y} \in \mathcal{Z}_{[t_0,t_1]}$ becomes a Banach space. We define for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\mathcal{J}(\tilde{y})(t) = e^{-At}\xi - \int_0^t A^\delta e^{-A(t-s)} A^{-\delta} B(\tilde{y}(s))\, ds + \int_0^t e^{-A(t-s)} Fu(s)\, ds + \int_0^t e^{-A(t-s)} G(\tilde{y}(s))\, dL(s).$$

Let $T_1 \in (0,T]$ and let us denote by $\mathcal{Z}_{T_1}$ the space $\mathcal{Z}_{[0,T_1]}$. First, we prove that $\mathcal{J}$ maps $\mathcal{Z}_{T_1}$ into itself. We define for all $t \in [0,T_1]$ and $\mathbb{P}$-a.s.

$$\psi_1(t) = e^{-At}\xi + \int_0^t e^{-A(t-s)} Fu(s)\, ds, \quad \psi_2(\tilde{y})(t) = \int_0^t A^\delta e^{-A(t-s)} A^{-\delta} B(\tilde{y}(s))\, ds,$$

$$\psi_3(\tilde{y})(t) = \int_0^t e^{-A(t-s)} G(\tilde{y}(s))\, dL(s).$$

Recall that $\left\|e^{-At}\right\|_{\mathcal{L}(\mathcal{H})} \leq 1$ for all $t \geq 0$ and $F \in \mathcal{L}(\mathcal{H})$. Using Theorem 2.35, Proposition B.9 and the Cauchy-Schwarz inequality, we get that the process $(\psi_1(t))_{t\in[0,T_1]}$ takes values in $D(A^\alpha)$ and there exists a constant $C_1 > 0$ such that

$$\begin{aligned}
\mathbb{E} \sup_{t\in[0,T_1]} \|\psi_1(t)\|^2_{D(A^\alpha)} &\le 2\,\mathbb{E} \sup_{t\in[0,T_1]} \left\|e^{-At}A^\alpha\xi\right\|^2_{\mathcal{H}} + 2\,\mathbb{E} \sup_{t\in[0,T_1]} \left(\int_0^t \left\|A^{\alpha-\beta}e^{-A(t-s)}A^\beta Fu(s)\right\|_{\mathcal{H}} ds\right)^2 \\
&\le 2\,\mathbb{E}\,\|\xi\|^2_{D(A^\alpha)} + 2M^2_{\alpha-\beta}\,\mathbb{E} \sup_{t\in[0,T_1]} \left(\int_0^t (t-s)^{\beta-\alpha}\,\|Fu(s)\|_{D(A^\beta)}\,ds\right)^2 \\
&\le C_1 \left[\mathbb{E}\|\xi\|^2_{D(A^\alpha)} + \mathbb{E}\int_0^{T_1} \|u(t)\|^2_{D(A^\beta)}\,dt\right].
\end{aligned}$$

By Theorem 2.29 (iv), Theorem 2.35, Proposition B.9 and inequality (3.18), the process $(\psi_2(\tilde{y})(t))_{t\in[0,T_1]}$ takes values in $D(A^\alpha)$ and there exists a constant $C_2 > 0$ such that

$$\begin{aligned}
\mathbb{E} \sup_{t\in[0,T_1]} \|\psi_2(\tilde{y})(t)\|^2_{D(A^\alpha)} &\le \mathbb{E} \sup_{t\in[0,T_1]} \left(\int_0^t \left\|A^{\alpha+\delta}e^{-A(t-s)}A^{-\delta}B(\tilde{y}(s))\right\|_{\mathcal{H}} ds\right)^2 \\
&\le M^2_{\alpha+\delta}C^2\,\mathbb{E} \sup_{t\in[0,T_1]} \left(\int_0^t (t-s)^{-\alpha-\delta}\,\|\tilde{y}(s)\|_{D(A^\alpha)}\,ds\right)^2 \\
&\le C_2\,\mathbb{E} \sup_{t\in[0,T_1]} \|\tilde{y}(t)\|^2_{D(A^\alpha)}.
\end{aligned}$$

Due to Theorem 2.35, Corollary 2.32 inequality (3.20), one can verify the assumptions of Proposition 3.63 and hence, the process $(\psi_3(\tilde{y})(t))_{t\in[0,T_1]}$ takes values in $D(A^\alpha)$. Using additionally Proposition 3.65 (i) with $k = 2$, there exists a constant $C_3 > 0$ such that

$$\begin{aligned}
\mathbb{E} \sup_{t\in[0,T_1]} \|\psi_3(\tilde{y})(t)\|^2_{D(A^\alpha)} &= \mathbb{E} \sup_{t\in[0,T_1]} \left\|\int_0^t e^{-A(t-s)}A^\alpha G(\tilde{y}(s))\,dL(s)\right\|^2_{\mathcal{H}} \\
&\le \widetilde{C}_2\,\mathbb{E}\int_0^{T_1} \|G(\tilde{y}(t))\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});D(A^\alpha))}\,dt \\
&\le C_3\,\mathbb{E} \sup_{t\in[0,T_1]} \|\tilde{y}(t)\|^2_{D(A^\alpha)}.
\end{aligned}$$

Hence, we can conclude that for fixed $\tilde{y} \in \mathcal{Z}_{T_1}$, the processes $(\mathcal{J}(\tilde{y})(t))_{t\in[0,T]}$ takes values in $D(A^\alpha)$ such that $\mathbb{E}\sup_{t\in[0,T]} \|\mathcal{J}(\tilde{y})(t)\|^2_{D(A^\alpha)} < \infty$. To conclude that $\mathcal{J}$ maps $\mathcal{Z}_{T_1}$ into itself, it remains to show that the process $(\mathcal{J}(\tilde{y})(t))_{t\in[0,T_1]}$ is predictable. We first prove that the process $(\mathcal{J}(\tilde{y})(t))_{t\in[0,T_1]}$ is mean square continuous. Note that similarly to Theorem 3.73, we obtain that the processes $(\psi_1(t))_{t\in[0,T_1]}$ and $(\psi_3(\tilde{y})(t))_{t\in[0,T_1]}$ are mean square continuous for fixed $\tilde{y} \in \mathcal{Z}_{T_1}$. We assume w.l.o.g. $0 \le t_0 \le t \le T_1$. Let I be the identity operator on $\mathcal{H}$. From Theorem 2.29 (iv), Theorem 2.35 and inequality (3.18), there exists a constant $\tilde{c} > 0$

such that

$$\begin{aligned}\mathbb{E}\left\|\psi_2(\tilde{y})(t)-\psi_2(\tilde{y})(t_0)\right\|^2_{D(A^\alpha)} &\le 2\,\mathbb{E}\left\|\int_0^{t_0}\left(e^{-A(t-t_0)}-I\right)A^{\alpha+\delta}e^{-A(t_0-s)}A^{-\delta}B(\tilde{y}(s))\,ds\right\|^2_{\mathcal{H}}\\ &\quad+2\,\mathbb{E}\left(\int_{t_0}^{t}\left\|A^{\alpha+\delta}e^{-A(t-s)}A^{-\delta}B(\tilde{y}(s))\right\|_{\mathcal{H}}ds\right)^2\\ &\le 2\,\mathbb{E}\left\|\left(e^{-A(t-t_0)}-I\right)\int_0^{t_0}A^{\alpha+\delta}e^{-A(t_0-s)}A^{-\delta}B(\tilde{y}(s))\,ds\right\|^2_{\mathcal{H}}\\ &\quad+\tilde{c}(t-t_0)^{2-2\alpha-2\delta}\,\mathbb{E}\sup_{t\in[0,T_1]}\|\tilde{y}(t)\|^2_{D(A^\alpha)}.\end{aligned}$$

Since $\lim_{t\to t_0}\|e^{-A(t-t_0)}h-h\|_{\mathcal{H}}=0$ holds for every $h\in\mathcal{H}$ and using Proposition B.7, we can infer that the process $(\psi_2(\tilde{y})(t))_{t\in[0,T_1]}$ is mean square continuous for fixed $\tilde{y}\in\mathcal{Z}_{T_1}$. Thus, we can conclude that the process $(\mathcal{J}(\tilde{y})(t))_{t\in[0,T_1]}$ is mean square continuous for fixed $\tilde{y}\in\mathcal{Z}_{T_1}$. Since $(\mathcal{J}(\tilde{y})(t))_{t\in[0,T_1]}$ is $\mathcal{F}_t$-adapted, we can apply Proposition 3.9. Hence, the process $(\mathcal{J}(\tilde{y})(t))_{t\in[0,T_1]}$ has a predictable modification for fixed $\tilde{y}\in\mathcal{Z}_{T_1}$.

Next, we show that $\mathcal{J}$ is a contraction on $\mathcal{Z}_{T_1}$. Using Theorem 2.29 (iv), Theorem 2.35 and inequality (3.19), there exists a constant $c_1>0$ such that for every $\tilde{y}_1,\tilde{y}_2\in\mathcal{Z}_{T_1}$

$$\begin{aligned}\mathbb{E}\sup_{t\in[0,T_1]}\|\psi_2(\tilde{y}_1)(t)-\psi_2(\tilde{y}_2)(t)\|^2_{D(A^\alpha)} &\le \mathbb{E}\sup_{t\in[0,T_1]}\left(\int_0^t\left\|A^{\alpha+\delta}e^{-A(t-s)}A^{-\delta}\left[B(\tilde{y}_1(s))-B(\tilde{y}_2(s))\right]\right\|_{\mathcal{H}}ds\right)^2\\ &\le c_1T_1^{2-2\alpha-2\delta}\,\mathbb{E}\sup_{t\in[0,T_1]}\|\tilde{y}_1(t)-\tilde{y}_2(t)\|^2_{D(A^\alpha)}.\end{aligned}$$

By Theorem 2.35, Proposition 3.65 and inequality (3.21), there exists a constant $c_2>0$ such that for every $\tilde{y}_1,\tilde{y}_2\in\mathcal{Z}_{T_1}$

$$\begin{aligned}\mathbb{E}\sup_{t\in[0,T_1]}\|\psi_3(\tilde{y}_1)(t)-\psi_3(\tilde{y}_2)(t)\|^2_{D(A^\alpha)} &= \mathbb{E}\sup_{t\in[0,T_1]}\left\|\int_0^t e^{-A(t-s)}A^\alpha\left[G(\tilde{y}_1(s))-G(\tilde{y}_2(s))\right]dL(s)\right\|^2_{\mathcal{H}}\\ &\le c_2T_1\,\mathbb{E}\sup_{t\in[0,T_1]}\|\tilde{y}_1(t)-\tilde{y}_2(t)\|^2_{D(A^\alpha)}.\end{aligned}$$

Consequently, we obtain for every $\tilde{y}_1,\tilde{y}_2\in\mathcal{Z}_{T_1}$

$$\mathbb{E}\sup_{t\in[0,T_1]}\|\mathcal{J}(\tilde{y}_1)(t)-\mathcal{J}(\tilde{y}_2)(t)\|^2_{D(A^\alpha)}\le K_1\,\mathbb{E}\sup_{t\in[0,T_1]}\|\tilde{y}_1(t)-\tilde{y}_2(t)\|^2_{D(A^\alpha)},$$

where $K_1=2c_1T_1^{2-2\alpha-2\delta}+2c_2T_1$. We chose $T_1\in[0,T]$ such that $K_1<1$. Applying the Banach fixed point theorem, we get a unique element $y\in\mathcal{Z}_{T_1}$ such that for all $t\in[0,T_1]$ and $\mathbb{P}$-a.s. $y(t)=\mathcal{J}(y)(t)$.

Next, we consider for all $t\in[T_1,T]$ and $\mathbb{P}$-a.s.

$$\mathcal{J}(\tilde{y})(t)=e^{-A(t-T_1)}y(T_1)-\int_{T_1}^{t}A^\delta e^{-A(t-s)}A^{-\delta}B(\tilde{y}(s))\,ds+\int_{T_1}^{t}e^{-A(t-s)}Fu(s)\,ds+\int_{T_1}^{t}e^{-A(t-s)}G(\tilde{y}(s))\,dL(s).$$

Again, for a certain $T_2 \in [T_1, T]$, there exists a unique fixed point of $\mathcal{J}$ on $\mathcal{Z}_{[T_1,T_2]}$. By continuing the method, we get the existence and uniqueness of a predictable process $(y(t))_{t\in[0,T]}$ satisfying for all $t \in [0, T]$ and $\mathbb{P}$-a.s. $y(t) = \mathcal{J}(y)(t)$. □

Remark 3.83. *Similarly to Remark 3.75, one can conclude that the mild solution of system 3.17 has a càdlàg modification. If $(L(t))_{t\geq 0}$ is a Q-Wiener process, then there exists a continuous modification, where we can argue as in Remark 3.76.*

3.4.2. Backward Stochastic Partial Differential Equations

Existence and uniqueness results of mild solutions to backward SPDEs are mainly based on a martingale representation theorem. These theorems are not available for infinite dimensional Lévy processes in general. Here, we will restrict to the case of Q-Wiener processes. Let $\mathcal{H}$ be a separable Hilbert space. Throughout this section, we assume that $(W(t))_{t\geq 0}$ is an $\mathcal{H}$-valued Q-Wiener process with covariance operator $Q \in \mathcal{L}_1^+(\mathcal{H})$. First, we provide a martingale representation theorem. A more general result is given in [42, Theorem 2.5]. By Proposition C.5, there exists an orthonormal basis $(h_n)_{n\in\mathbb{N}}$ of $\mathcal{H}$ and a sequence of nonnegative real numbers $(\lambda_n)_{n\in\mathbb{N}}$ such that $Qh_n = \lambda_n h_n$ for each $n \in \mathbb{N}$. Due to Proposition 3.42, we have the following expansion for arbitrary $t \geq 0$:

$$W(t) = \sum_{n=1}^{\infty} \sqrt{\lambda_n} w_n(t) h_n,$$

where $(w_n(t))_{t\geq 0}$, $n \in \mathbb{N}$, are mutually independent real valued Brownian motions. For the remaining part of this section, we assume that the complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is endowed with the filtration $(\mathcal{F}(t))_{t\geq 0}$ given by $\mathcal{F}_t = \sigma\{\bigcup_{n=1}^{\infty} \mathcal{F}_t^n\}$ for all $t \geq 0$, where $\mathcal{F}_t^n = \sigma\{w_n(s) : 0 \leq s \leq t\}$. We need the following auxiliary results, where $T > 0$ is fixed.

Lemma 3.84. *For each $n \in \mathbb{N}$, the linear span of the random variables*

$$\left\{ \exp\left\{ \int_0^T h(t)\, dw_n(t) - \frac{1}{2}\int_0^T (h(t))^2\, dt \right\} : h \in L^2([0,T]) \text{ deterministic} \right\}$$

is dense in $L^2(\Omega, \mathcal{F}_T^n, \mathbb{P})$.

Proof. The claim follows immediately from [66, Lemma 4.3.2]. □

Lemma 3.85. *Let the process $(m(t))_{t\geq 0}$ be a continuous real valued $\mathcal{F}_t$-martingale such that $\mathbb{E}|m(t)|^2 < \infty$ for all $t \geq 0$. Then there exists a unique sequence of predictable real valued processes $(\phi_n(t))_{t\in[0,T]}$, $n \in \mathbb{N}$, such that for all $t \in [0, T]$ and $\mathbb{P}$-a.s.*

$$m(t) = \mathbb{E}[m(0)] + \sum_{n=1}^{\infty} \int_0^t \sqrt{\lambda_n} \phi_n(s)\, dw_n(s),$$

where $\sum_{n=1}^{\infty} \lambda_n \mathbb{E} \int_0^T |\phi_n(t)|^2 dt < \infty$.

Proof. By definition, we get

$$L^2(\Omega, \mathcal{F}_T, \mathbb{P}) = \bigoplus_{n=1}^{\infty} L^2(\Omega, \mathcal{F}_T^n, \mathbb{P}).$$

As a consequence of Lemma 3.84, the linear span of the random variables

$$\left\{\exp\left\{\int\limits_0^T h(t)\,dw_n(t) - \frac{1}{2}\int\limits_0^T (h(t))^2\,dt\right\} : h \in L^2([0,T]) \text{ deterministic}, n \in \mathbb{N}\right\}$$

is dense in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. For the remaining part, we can adopt the proof of [66, Theorem 4.3.4]. □

We have the following martingale representation theorem in $(\Omega, \mathcal{F}_T, \mathbb{P})$. A proof can be found in [42]. For the convenience of the reader, we will adopt this proof here.

Theorem 3.86. *Let the process $(M(t))_{t\geq 0}$ be a continuous $\mathcal{F}_t$-martingale with values in $\mathcal{H}$ such that $\mathbb{E}\|M(t)\|^2_{\mathcal{H}} < \infty$ for all $t \geq 0$. Then there exists a unique predictable process $(\Phi(t))_{t\in[0,T]}$ with values in $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})$ such that $\mathbb{E}\int_0^T \|\Phi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dt < \infty$ and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.*

$$M(t) = \mathbb{E}[M(0)] + \int\limits_0^t \Phi(s)\,dW(s).$$

Proof. Recall that $(h_m)_{m\in\mathbb{N}}$ is an orthonormal basis of $\mathcal{H}$. Using Lemma 3.85 for each $m \in \mathbb{N}$, there exists a unique sequence of predictable real valued processes $(\phi_n^m(t))_{t\in[0,T]}$, $n \in \mathbb{N}$, such that for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\langle M(t), h_m\rangle_{\mathcal{H}} = \mathbb{E}[\langle M(0), h_m\rangle_{\mathcal{H}}] + \sum_{n=1}^{\infty}\int\limits_0^t \sqrt{\lambda_n}\phi_n^m(s)\,dw_n(s),$$

where $\sum_{n=1}^{\infty}\lambda_n\,\mathbb{E}\int_0^T |\phi_n^m(t)|^2 dt < \infty$. Note that for all $t \in [0,T]$

$$\mathbb{E}\sum_{m=1}^{\infty}\langle M(t), h_m\rangle^2_{\mathcal{H}} = \mathbb{E}\|M(t)\|^2_{\mathcal{H}} < \infty.$$

Hence, we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$M(t) = \sum_{m=1}^{\infty}\langle M(t), h_m\rangle_{\mathcal{H}} h_m$$

and using Proposition B.7, we obtain

$$\sum_{m=1}^{\infty}\mathbb{E}[\langle M(0), h_m\rangle_{\mathcal{H}} h_m] = \mathbb{E}\left[\sum_{m=1}^{\infty}\langle M(0), h_m\rangle_{\mathcal{H}} h_m\right] = \mathbb{E}[M(0)].$$

Therefore, we get for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} M(t) &= \sum_{m=1}^{\infty}\mathbb{E}[\langle M(0), h_m\rangle_{\mathcal{H}} h_m] + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\int\limits_0^t \sqrt{\lambda_n}\phi_n^m(s) h_m\,dw_n(s) \\ &= \mathbb{E}[M(0)] + \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\int\limits_0^t \sqrt{\lambda_n}\phi_n^m(s) h_m\,dw_n(s). \end{aligned}$$

This representation and the assumptions on the process $(M(t))_{t\geq 0}$ justifies the interchanging of summations with the result that

$$M(t) = \mathbb{E}[M(0)] + \sum_{n=1}^{\infty} \sqrt{\lambda_n} \sum_{m=1}^{\infty} \int_0^t \phi_n^m(s) h_m \, dw_n(s).$$

Next, let the process $(\Phi(t))_{t\in[0,T]}$ be defined for every $x \in Q^{1/2}(\mathcal{H})$, every $y \in \mathcal{H}$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\langle \Phi(t)x, y\rangle_{\mathcal{H}} = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} \lambda_n \langle y, h_m\rangle_{\mathcal{H}} \langle x, h_n\rangle_{Q^{1/2}(\mathcal{H})} \phi_n^m(t),$$

where the inner product in $Q^{1/2}(\mathcal{H})$ is defined in Remark C.10. Then the process $(\Phi(t))_{t\in[0,T]}$ is predictable with values in $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})$ such that $\mathbb{E}\int_0^T \|\Phi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dt < \infty$ and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\int_0^t \Phi(s)\, dW(s) = \sum_{n=1}^{\infty} \sqrt{\lambda_n} \sum_{m=1}^{\infty} \int_0^t \phi_n^m(s) h_m \, dw_n(s),$$

which completes the proof. □

Remark 3.87. *Here, we recall a martingale representation theorem, where the filtration is generated by a real-valued Lévy process. For more details, we refer to [65]. Let $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ be a complete probability space and let $(L(t))_{t\geq 0}$ be a real-valued Lévy process, where we assume that we are using the càdlàg modification. We endow the probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ with the filtration $(\overline{\mathcal{F}_t})_{t\geq 0}$ given by $\overline{\mathcal{F}_t} = \sigma\{\mathcal{G}_t \cup \mathcal{N}\}$ for all $t \geq 0$, where $\mathcal{G}_t = \sigma\{L(s) : 0 \leq s \leq t\}$ and $\mathcal{N}$ contains all sets $A \in \mathcal{F}$ with $\overline{\mathbb{P}}(A) = 0$. The characteristic function of $(L(t))_{t\geq 0}$ is given by*

$$\mathbb{E}\, e^{i\theta L(t)} = e^{-t\psi(\theta)}$$

for every $\theta \in \mathbb{R}$ and all $t \geq 0$, where

$$\psi(\theta) = -ia\theta + \frac{\sigma^2}{2}\theta^2 + \int_{\mathbb{R}} \left(1 - e^{i\theta x} + \mathbb{1}_{\{|x|<1\}}(x)\, i\theta x\right) \nu(dx)$$

with $a \in \mathbb{R}$, $\sigma^2 \geq 0$ and ν is a measure on $\mathbb{R}\backslash\{0\}$ with $\int_{\mathbb{R}}(1 \wedge x^2)\nu(dx) < \infty$. This formula is the well known Lévy-Khinchin formula, which is also stated in Theorem 3.50 for an infinite dimensional Lévy process. We assume that for some $\varepsilon > 0$ and $\lambda > 0$

$$\int_{(-\varepsilon,\varepsilon)^c} e^{\lambda|x|}\, \nu(dx) < \infty,$$

which implies especially that $\mathbb{E}|L(t)|^n < \infty$ for each $n \in \mathbb{N}$ and all $t \geq 0$. Let $(\Delta L(t))_{t\geq 0}$ be the process of jumps given by $\Delta L(t) = L(t) - L(t-)$ for all $t \geq 0$ and $\mathbb{P}$-almost surely. We also introduce the power jump processes $(X^k(t))_{t\geq 0}$ with $k \in \mathbb{N}$ defined by

$$X^k(t) = \begin{cases} L(t) & \text{if } k = 1 \\ \sum\limits_{0<s\leq t} (\Delta L(s))^k & \text{if } k \geq 2 \end{cases}$$

for all $t \geq 0$ and $\mathbb{P}$-almost surely. Then the processes $(X^k(t))_{t\geq 0}$ with $k \in \mathbb{N}$ are again Lévy processes and we get for each $k \in \mathbb{N}$ and all $t \geq 0$

$$\mathbb{E}[X^k(t)] = tm_k$$

with $m_1 = \mathbb{E}[L(1)]$ *and* $m_k = \int_{\mathbb{R}} x^k \nu(dx)$ *for* $k \geq 2$*. We denote by* $(Y^k(t))_{t\geq 0}$ *with* $k \in \mathbb{N}$ *the compensated power jump processes given by* $Y^k(t) = X^k(t) - tm_k$ *for each* $k \in \mathbb{N}$*, all* $t \geq 0$ *and* $\mathbb{P}$*-almost surely. The processes* $(Y^k(t))_{t\geq 0}$ *with* $k \in \mathbb{N}$ *are martingales. We introduce the stochastic processes* $(H^k(t))_{t\geq 0}$ *with* $k \in \mathbb{N}$ *as linear combinations of the stochastic processes* $(Y^j(t))_{t\geq 0}$ *for* $j = 1, ..., k$ *with the leading coefficient equal to 1, i.e. we have for each* $k \in \mathbb{N}$*, all* $t \geq 0$ *and* $\mathbb{P}$*-a.s.*

$$H^k(t) = Y^k(t) + a_{k,k-1}Y^{k-1}(t) + ... + a_{k,1}Y^1(t),$$

where $a_{k,j} \in \mathbb{R}$ *for all* $j = 1, ..., k-1$*. The processes* $(H^k(t))_{t\geq 0}$ *with* $k \in \mathbb{N}$ *are again martingales. Furthermore, the coefficients* $a_{k,j} \in \mathbb{R}$ *with* $k \in \mathbb{N}$ *and* $j = 1, ..., k-1$ *are chosen such that the processes* $(H^k(t))_{t\geq 0}$ *with* $k \in \mathbb{N}$ *are pairwise strongly orthogonal, i.e. we have for each* $k, l \in \mathbb{N}$

$$\lim_{n\to\infty} \sup_{t\geq 0} \int_{\{|H^k(t)H^l(t)|\geq n\}} |H^k(t)H^l(t)| \, \mathbb{P}(d\omega) = 0.$$

We get the following martingale representation theorem, see [65, Remark 2]: If $(m(t))_{t\geq 0}$ *is a square integrable real-valued* $\overline{\mathcal{F}_t}$*-martingale satisfying* $\sup_{t\geq 0} \mathbb{E}|m(t)|^2 < \infty$*, then there exist predictable processes* $(\phi^k(t))_{t\geq 0}$ *with* $k \in \mathbb{N}$ *such that* $\mathbb{E}\int_0^\infty |\phi^k(t)|^2 dt < \infty$ *for each* $k \in \mathbb{N}$ *and we have for all* $t \geq 0$ *and* $\mathbb{P}$*-a.s.*

$$m(t) = \mathbb{E}[m(0)] + \sum_{k=1}^{\infty} \int_0^t \phi^k(s) \, dH^k(s).$$

We also note that further martingale representation theorems for filtration generated by real-valued square integrable Lévy processes can be found in [2, 77].

Remark 3.88. *The previous remark enables us to state a martingale representation theorem for an infinite dimensional martingale as follows: Again, we assume that the complete probability space* $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ *is endowed with the filtration* $(\overline{\mathcal{F}_t})_{t\geq 0}$ *as introduced in the previous remark. Let* $(M(t))_{t\geq 0}$ *be a square integrable* $\overline{\mathcal{F}_t}$*-martingale with values in* $\mathcal{H}$ *satisfying* $\sup_{t\geq 0} \mathbb{E}\|M(t)\|^2_{\mathcal{H}} < \infty$*. Let* $(h_n)_{n\in\mathbb{N}}$ *be an orthonormal basis in* $\mathcal{H}$*. The processes* $(\langle M(t), h_n\rangle_{\mathcal{H}})_{t\geq 0}$ *are square integrable real-valued* $\overline{\mathcal{F}_t}$*-martingale such that* $\sup_{t\geq 0} \mathbb{E}|\langle M(t), h_n\rangle_{\mathcal{H}}|^2 < \infty$ *for each* $n \in \mathbb{N}$*. Hence, for each* $n \in \mathbb{N}$*, there exist predictable processes* $(\phi_n^k(t))_{t\geq 0}$ *with* $k \in \mathbb{N}$ *such that for all* $t \geq 0$ *and* $\mathbb{P}$*-a.s.*

$$\langle M(t), h_n\rangle_{\mathcal{H}} = \mathbb{E}[\langle M(0), h_n\rangle_{\mathcal{H}}] + \sum_{k=1}^{\infty} \int_0^t \phi_n^k(s) \, dH^k(s).$$

Since $\sup_{t\geq 0} \mathbb{E}\sum_{n=1}^\infty \langle M(t), h_n\rangle^2_{\mathcal{H}} = \sup_{t\geq 0} \mathbb{E}\|M(t)\|^2_{\mathcal{H}} < \infty$*, we obtain for all* $t \geq 0$ *and* $\mathbb{P}$*-a.s.*

$$M(t) = \sum_{n=1}^{\infty} \langle M(t), h_n\rangle_{\mathcal{H}} h_n = \sum_{n=1}^{\infty} \mathbb{E}[\langle M(0), h_n\rangle_{\mathcal{H}} h_n] + \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \int_0^t \phi_n^k(s) h_n \, dH^k(s).$$

Using Proposition B.7, we get for all $t \geq 0$ *and* $\mathbb{P}$*-a.s.*

$$M(t) = \mathbb{E}[M(0)] + \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \int_0^t \phi_n^k(s) h_n \, dH^k(s).$$

Remark 3.89. *Note that the martingale representation theorem derived in the previous remark is based on the filtration* $(\overline{\mathcal{F}_t})_{t\geq 0}$ *generated by the real-valued Lévy process* $(L(t))_{t\geq 0}$*. Following the proof of Theorem*

3.86, a general martingale representation theorem with a filtration generated by an infinite dimensional Lévy process requires a series expansion with mutually independent real-valued Lévy processes. According to Theorem 3.56 such a series expansion is only available with uncorrelated Lévy processes. For that reason, we are forced to restrict the martingale representation theorem stated in the previous remark to the case of a filtration generated by a real-valued Lévy process.

Next, we introduce the following system in $\mathcal{H}$:

$$\begin{cases} dz(t) = -[-Az(t) + G(z(t), \Phi(t)) + g(t)]dt + \Phi(t)\, dW(t), \\ z(T) = Z. \end{cases} \tag{3.23}$$

We assume that

- the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ is linear and closed such that $-A$ is the generator of a C_0 semigroup $(e^{-At})_{t\geq 0}$;
- $G\colon \mathcal{H} \times \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H}) \to \mathcal{H}$ satisfies for every $y, z \in \mathcal{H}$ and every $\Phi, \Psi \in L_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})$
$$\|G(y, \Phi)\|_{\mathcal{H}} \leq \widehat{C}\left[\|y\|_{\mathcal{H}} + \|\Phi\|_{L_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\right], \tag{3.24}$$
$$\|G(y, \Phi) - G(z, \Psi)\|_{\mathcal{H}} \leq \widehat{C}\left[\|y - z\|_{\mathcal{H}} + \|\Phi - \Psi\|_{L_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\right], \tag{3.25}$$
where $\widehat{C} > 0$ is a constant;
- $(g(t))_{t\in[0,T]}$ is a predictable process with values in $\mathcal{H}$ such that
$$\mathbb{E}\int_0^T \|g(t)\|_{\mathcal{H}}^2\, dt < \infty;$$
- Z is an $\mathcal{F}_T$-measurable random variable with values in $\mathcal{H}$.

Remark 3.90. *In control theory, system (3.23) arises for the adjoint equation of the controlled stochastic Stokes equations.*

Definition 3.91. *A pair of predictable processes $(z(t), \Phi(t))_{t\in[0,T]}$ with values in $\mathcal{H} \times \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})$ is called a **mild solution of system (3.23)** if*

$$\sup_{t\in[0,T]} \mathbb{E}\, \|z(t)\|_{\mathcal{H}}^2 < \infty, \qquad \mathbb{E}\int_0^T \|\Phi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dt < \infty$$

and we have for all $t \in [0, T]$ and $\mathbb{P}$-a.s.

$$z(t) = e^{-A(T-t)}Z + \int_t^T e^{-A(s-t)}[G(z(s), \Phi(s)) + g(s)]\, ds - \int_t^T e^{-A(s-t)}\Phi(s)\, dW(s). \tag{3.26}$$

An existence and uniqueness result is mainly based on the following lemma.

Lemma 3.92 (Lemma 2.1,[52]). *Let $\zeta \in L^2(\Omega;\mathcal{H})$ be $\mathcal{F}_T$-measurable and let $(f(t))_{t\in[0,T]}$ be a predictable process with values in $\mathcal{H}$ such that $\mathbb{E}\int_0^T \|f(t)\|_{\mathcal{H}}^2 dt < \infty$. Then there exists a unique pair of predictable processes $(\varphi(t),\phi(t))_{t\in[0,T]}$ with values in $\mathcal{H}\times\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})$ such that for all $t\in[0,T]$ and $\mathbb{P}$-a.s.*

$$\varphi(t) = e^{-A(T-t)}\zeta + \int_t^T e^{-A(s-t)} f(s)\, ds - \int_t^T e^{-A(s-t)}\phi(s)\, dW(s).$$

Moreover, there exists a constant $c>0$ such that for all $t\in[0,T]$

$$\mathbb{E}\,\|\varphi(t)\|_{\mathcal{H}}^2 \le c\left[\mathbb{E}\,\|\zeta\|_{\mathcal{H}}^2 + (T-t)\,\mathbb{E}\int_t^T \|f(s)\|_{\mathcal{H}}^2 ds\right], \tag{3.27}$$

$$\mathbb{E}\int_t^T \|\phi(s)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} ds \le c\left[\mathbb{E}\,\|\zeta\|_{\mathcal{H}}^2 + (T-t)\,\mathbb{E}\int_t^T \|f(s)\|_{\mathcal{H}}^2 ds\right]. \tag{3.28}$$

Existence and uniqueness results of mild solutions to backward SPDEs with cylindrical Wiener processes can be found in [52]. Similarly, we get the existence of a unique mild solution to system (3.23).

Theorem 3.93. *Let $(g(t))_{t\in[0,T]}$ be fixed. For any $Z\in L^2(\Omega;\mathcal{H})$, there exists a unique mild solution $(z(t),\Phi(t))_{t\in[0,T]}$ of system (3.23).*

Proof. Let $\mathcal{Z}_T^1$ contain all $\mathcal{H}$-valued predictable processes $(\tilde z(t))_{t\in[0,T]}$ such that $\sup_{t\in[0,T]}\mathbb{E}\|\tilde z(t)\|_{\mathcal{H}}^2 < \infty$. The space $\mathcal{Z}_T^1$ equipped with the norm

$$\|\tilde z\|^2_{\mathcal{Z}_T^1} = \sup_{t\in[0,T]} \mathbb{E}\|\tilde z(t)\|^2_{\mathcal{H}}$$

for every $\tilde z\in\mathcal{Z}_T^1$ becomes a Banach space. Similarly, let $\mathcal{Z}_T^2$ denote the space of all predictable processes $(\tilde\Phi(t))_{t\in[0,T]}$ with values in $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})$ such that $\mathbb{E}\int_0^T \|\tilde\Phi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dt < \infty$. The space $\mathcal{Z}_T^2$ equipped with the inner product

$$\left\langle \tilde\Phi_1,\tilde\Phi_2\right\rangle_{\mathcal{Z}_T^2} = \mathbb{E}\int_0^T \left\langle \tilde\Phi_1(t),\tilde\Phi_2(t)\right\rangle_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dt$$

for every $\tilde\Phi_1,\tilde\Phi_2\in\mathcal{Z}_T^2$ becomes a Hilbert space.

Next, we define a sequence $(z_n,\Phi_n)_{n\in\mathbb{N}}\subset\mathcal{Z}_T^1\times\mathcal{Z}_T^2$ satisfying for each $n\in\mathbb{N}$, all $t\in[0,T]$ and $\mathbb{P}$-a.s.

$$z_n(t) = e^{-A(T-t)}Z + \int_t^T e^{-A(s-t)}[G(z_{n-1}(s),\Phi_{n-1}(s)) + g(s)]ds - \int_t^T e^{-A(s-t)}\Phi_n(s)\, dW(s), \tag{3.29}$$

where $z_0(t)=0$ and $\Phi_0(t)=0$ for all $t\in[0,T]$. Note that by Lemma 3.92 and inequality (3.24), one can easily verify that $(z_n,\Phi_n)_{n\in\mathbb{N}}\subset\mathcal{Z}_T^1\times\mathcal{Z}_T^2$. Furthermore, we obtain for each $n\in\mathbb{N}$, all $t\in[0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} z_{n+1}(t) - z_n(t) &= \int_t^T e^{-A(s-t)}[G(z_n(s),\Phi_n(s)) - G(z_{n-1}(s),\Phi_{n-1}(s))]ds \\ &\quad - \int_t^T e^{-A(s-t)}\left[\Phi_{n+1}(s) - \Phi_n(s)\right] dW(s). \end{aligned} \tag{3.30}$$

Using inequality (3.25) and Fubini's theorem, there exists a constant $C > 0$ such that for each $n \in \mathbb{N}$

$$\mathbb{E}\int_0^T \|G(z_n(t),\Phi_n(t)) - G(z_{n-1}(t),\Phi_{n-1}(t))\|^2_{\mathcal{H}}\,dt$$
$$\leq C\left[\sup_{t\in[0,T]}\mathbb{E}\|z_n(t)-z_{n-1}(t)\|^2_{\mathcal{H}} + \mathbb{E}\int_0^T \|\Phi_n(t)-\Phi_{n-1}(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\,dt\right]. \tag{3.31}$$

Hence, equation (3.30) satisfies the assumptions of Lemma 3.92. Let $T_1 \in [0,T)$. Due to the inequalities (3.27), (3.28) and (3.31), there exists a constant $C^* > 0$ such that for each $n \in \mathbb{N}$

$$\sup_{t\in[T_1,T]}\mathbb{E}\|z_{n+1}(t)-z_n(t)\|^2_{\mathcal{H}} + \mathbb{E}\int_{T_1}^T \|\Phi_{n+1}(t)-\Phi_n(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\,dt$$
$$\leq C^*(T-T_1)\left[\sup_{t\in[T_1,T]}\mathbb{E}\|z_n(t)-z_{n-1}(t)\|^2_{\mathcal{H}} + \mathbb{E}\int_{T_1}^T \|\Phi_n(t)-\Phi_{n-1}(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\,dt\right].$$

Therefore, we find for each $n \in \mathbb{N}$

$$\sup_{t\in[T_1,T]}\mathbb{E}\|z_{n+1}(t)-z_n(t)\|^2_{\mathcal{H}} + \mathbb{E}\int_{T_1}^T \|\Phi_{n+1}(t)-\Phi_n(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\,dt$$
$$\leq \left(C^*(T-T_1)\right)^n\left[\sup_{t\in[T_1,T]}\mathbb{E}\|z_1(t)\|^2_{\mathcal{H}} + \mathbb{E}\int_{T_1}^T \|\Phi_1(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\,dt\right].$$

We choose $T_1 \in [0,T)$ such that $C^*(T-T_1) < 1$. Thus, we can conclude that $(z_n,\Phi_n)_{n\in\mathbb{N}} \subset \mathcal{Z}_T^1 \times \mathcal{Z}_T^2$ is a Cauchy sequence on the interval $[T_1,T]$. Using equation (3.30), we have for each $n \in \mathbb{N}$, all $t \in [0,T_1]$ and $\mathbb{P}$-a.s.

$$z_{n+1}(t)-z_n(t) = e^{-A(T_1-t)}[z_{n+1}(T_1)-z_n(T_1)] + \int_t^{T_1} e^{-A(s-t)}[G(z_n(s),\Phi_n(s)) - G(z_{n-1}(s),\Phi_{n-1}(s))]ds$$
$$- \int_t^{T_1} e^{-A(s-t)}\left[\Phi_{n+1}(s)-\Phi_n(s)\right]dW(s).$$

Again, we find $T_2 \in [0,T_1]$ such that the sequence $(z_n,\Phi_n)_{n\in\mathbb{N}} \subset \mathcal{Z}_T^1 \times \mathcal{Z}_T^2$ is a Cauchy sequence on the interval $[T_2,T_1]$. By continuing this method, we can infer that $(z_n,\Phi_n)_{n\in\mathbb{N}} \subset \mathcal{Z}_T^1 \times \mathcal{Z}_T^2$ is a Cauchy sequence on the interval $[0,T]$. Hence, there exist $z \in \mathcal{Z}_T^1$ and $\Phi \in \mathcal{Z}_T^2$ such that

$$z = \lim_{n\to\infty} z_n, \quad \Phi = \lim_{n\to\infty}\Phi_n.$$

Using equation (3.29), one can easily verify that the pair of stochastic processes $(z(t),\Phi(t))_{t\in[0,T]}$ satisfy equation (3.26). □

We introduce the following system in $D(A^\delta)$:

$$\begin{cases} dz(t) = -[-Az(t) + A^\alpha B(z(t)) + A^\beta G(\Phi(t)) + A^\gamma g(t)]dt + \Phi(t)\, dW(t), \\ z(T) = Z. \end{cases} \tag{3.32}$$

We assume that

- the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ is linear and closed such that $-A$ is the generator of an analytic semigroup of contractions $(e^{-At})_{t\geq 0}$ and 0 is an element of the resolvent set $\rho(A)$;
- $B\colon D(A^\delta) \to \mathcal{H}$ satisfies for every $y, z \in \mathcal{H}$

$$\|B(y)\|_{\mathcal{H}} \leq \widetilde{C}\|y\|_{D(A^\delta)}, \tag{3.33}$$

$$\|B(y) - B(z)\|_{\mathcal{H}} \leq \widetilde{C}\|y - z\|_{D(A^\delta)}, \tag{3.34}$$

  where $C^* > 0$ is a constant;
- $G\colon \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H}) \to \mathcal{H}$ satisfies for every $\Phi, \Psi \in L_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})$

$$\|G(\Phi)\|_{\mathcal{H}} \leq \widehat{C}\|\Phi\|_{L_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}, \tag{3.35}$$

$$\|G(\Phi) - G(\Psi)\|_{\mathcal{H}} \leq \widehat{C}\|\Phi - \Psi\|_{L_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}, \tag{3.36}$$

  where $\widehat{C} > 0$ is a constant;
- $(g(t))_{t\in[0,T]}$ is a predictable process with values in $\mathcal{H}$ such that

$$\mathbb{E}\int_0^T \|g(t)\|^2_{\mathcal{H}}\, dt < \infty;$$

- Z is an $\mathcal{F}_T$-measurable random variable with values in $\mathcal{H}$.

Remark 3.94. *In control theory, system (3.32) arises for the adjoint equation of the controlled stochastic Navier-Stokes equations.*

Definition 3.95. *A pair of predictable processes $(z(t), \Phi(t))_{t\in[0,T]}$ with values in $D(A^\delta)\times\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})$ is called a* ***mild solution of system (3.32)*** *if*

$$\mathbb{E}\sup_{t\in[0,T]} \|z(t)\|^2_{D(A^\delta)} < \infty, \qquad \mathbb{E}\int_0^T \|\Phi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dt < \infty$$

and we have for all $t \in [0, T]$ and $\mathbb{P}$-a.s.

$$z(t) = e^{-A(T-t)}Z + \int_t^T A^\alpha e^{-A(s-t)}B(z(s))\, ds + \int_t^T A^\beta e^{-A(s-t)}G(\Phi(s))\, ds + \int_t^T A^\gamma e^{-A(s-t)}g(s)\, ds$$
$$- \int_t^T e^{-A(s-t)}\Phi(s)\, dW(s). \tag{3.37}$$

An existence and uniqueness result requires a generalization of Lemma 3.92 as follows.

Lemma 3.96. *Let $\delta, \varepsilon \in [0, \frac{1}{2})$ satisfy $\delta + \varepsilon < \frac{1}{2}$. Furthermore, let $\zeta \in L^2(\Omega; D(A^\delta))$ be $\mathcal{F}_T$-measurable and let $(f(t))_{t\in[0,T]}$ be a predictable process with values in $\mathcal{H}$ such that $\mathbb{E}\int_0^T \|f(t)\|_{\mathcal{H}}^2 dt < \infty$. Then there exists a unique pair of predictable processes $(\varphi(t), \phi(t))_{t\in[0,T]}$ with values in $D(A^\delta) \times \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); D(A^\varepsilon))$ such that for all $t \in [0,T]$ and $\mathbb{P}$-a.s.*

$$\varphi(t) = e^{-A(T-t)}\zeta + \int_t^T A^\varepsilon e^{-A(s-t)} f(s)\, ds - \int_t^T e^{-A(s-t)} A^\varepsilon \phi(s)\, dW(s). \tag{3.38}$$

Moreover, there exists a constant $c > 0$ such that for all $t \in [0,T]$

$$\mathbb{E} \sup_{s\in[t,T]} \|\varphi(s)\|_{D(A^\delta)}^2 \leq c \left[\mathbb{E}\, \|\zeta\|_{D(A^\delta)}^2 + (T-t)^{1-2\delta-2\varepsilon}\, \mathbb{E} \int_t^T \|f(s)\|_{\mathcal{H}}^2 ds \right], \tag{3.39}$$

$$\mathbb{E} \int_t^T \|\phi(s)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); D(A^\varepsilon))}^2 ds \leq c \left[\mathbb{E}\, \|\zeta\|_{D(A^\delta)}^2 + (T-t)^{1-2\varepsilon}\, \mathbb{E} \int_t^T \|f(s)\|_{\mathcal{H}}^2 ds \right]. \tag{3.40}$$

Proof. Let the process $(\varphi(t))_{t\in[0,T]}$ satisfy for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\varphi(t) = \mathbb{E}\left[e^{-A(T-t)}\zeta + \int_t^T A^\varepsilon e^{-A(s-t)} f(s)\, ds \,\middle|\, \mathcal{F}_t \right].$$

Due to Proposition 3.16, we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\varphi(t) = e^{-A(T-t)} \mathbb{E}\left[\zeta | \mathcal{F}_t\right] + \int_t^T A^\varepsilon e^{-A(s-t)} \mathbb{E}\left[f(s) | \mathcal{F}_t\right] ds. \tag{3.41}$$

Using Theorem 2.35 and the Cauchy-Schwarz inequality, we can conclude that the process $(\varphi(t))_{t\in[0,T]}$ takes values in $D(A^\delta)$ such that $\mathbb{E}\sup_{t\in[0,T]} \|\varphi(t)\|_{D(A^\delta)}^2 < \infty$. Moreover, the process $(\varphi(t))_{t\in[0,T]}$ is predictable, which can be obtained similarly to Theorem 3.73. By Proposition 3.86, there exists a unique predictable process $(J(r))_{r\in[0,T]}$ with values in $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})$ such that $\mathbb{E}\int_0^T \|J(r)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}^2 dr < \infty$ and we get for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\mathbb{E}\left[\zeta | \mathcal{F}_t\right] = \mathbb{E}\left[\zeta\right] + \int_0^t J(r)\, dW(r).$$

Thus, we get for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\mathbb{E}\left[\zeta | \mathcal{F}_t\right] = \zeta - \int_t^T J(r)\, dW(r). \tag{3.42}$$

We first assume that $(f(s))_{s\in[0,T]}$ is predictable and continuous such that $\sup_{s\in[0,T]} \mathbb{E}\|f(s)\|_{\mathcal{H}}^2 < \infty$. Using Proposition 3.86, for all $s \in [0,T]$, there exist a unique predictable process $(K(s,r))_{r\in[0,T]}$ with values in

$\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})$ such that $\mathbb{E}\int_0^T \|K(s,r)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}dr < \infty$ and we get for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\mathbb{E}\left[f(s)|\mathcal{F}_t\right] = \mathbb{E}\left[f(s)\right] + \int_0^t K(s,r)\,dW(r).$$

Since $(f(s))_{s\in[0,T]}$ is predictable, one can conclude that $K(s,r) = 0$ for all $s \in [0,T]$ and almost all $r \in [s,T]$. Moreover, we have for all $s,t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\mathbb{E}\left[f(s)|\mathcal{F}_t\right] = f(s) - \int_t^s K(s,r)\,dW(r). \tag{3.43}$$

Using equations (3.41) - (3.43), we obtain for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\varphi(t) = e^{-A(T-t)}\left[\zeta - \int_t^T J(r)\,dW(r)\right] + \int_t^T A^{\varepsilon}e^{-A(s-t)}\left[f(s) - \int_t^s K(s,r)\,dW(r)\right]ds.$$

Applying Proposition 3.64, we get for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\varphi(t) = e^{-A(T-t)}\zeta + \int_t^T A^{\varepsilon}e^{-A(s-t)}f(s)\,ds - \int_t^T A^{\varepsilon}e^{-A(r-t)}\phi(r)\,dW(r), \tag{3.44}$$

where for almost all $r \in [0,T]$ and $\mathbb{P}$-a.s.

$$\phi(r) = A^{-\varepsilon}e^{-A(T-r)}J(r) + \int_r^T e^{-A(s-r)}K(s,r)\,ds.$$

By Theorem 2.35 and the Cauchy-Schwarz inequality, we can conclude that the process $(\phi(t))_{t\in[0,T]}$ takes values in $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});D(A^{\varepsilon}))$. Since the processes $(J(r))_{r\in[0,T]}$ and $(K(s,r))_{r\in[0,T]}$ are predictable for all $s \in [0,T]$, the process $(\phi(r))_{r\in[0,T]}$ is predictable as well.

Next, we show that inequality (3.39) and inequality (3.40) hold. By equation (3.41), Proposition 3.16, Theorem 2.35 and the Cauchy-Schwarz inequality, there exists a constant $c > 0$ such that for all $t \in [0,T]$

$$\begin{aligned}\mathbb{E}\sup_{s\in[t,T]}\|\varphi(s)\|^2_{D(A^{\delta})} &\leq 2\,\mathbb{E}\sup_{s\in[t,T]}\|e^{-A(T-s)}A^{\delta}\zeta\|^2_{\mathcal{H}} + 2\,\mathbb{E}\sup_{s\in[t,T]}\left(\int_s^T \left\|A^{\delta+\varepsilon}e^{-A(r-s)}f(r)\right\|_{\mathcal{H}}dr\right)^2\\ &\leq c\left[\mathbb{E}\,\|\zeta\|^2_{D(A^{\delta})} + (T-t)^{1-2\delta-2\varepsilon}\,\mathbb{E}\int_t^T \|f(s)\|^2_{\mathcal{H}}ds\right].\end{aligned}$$

Using equation (3.42) and Theorem 3.62 (i) and (iii), we get for all $t \in [0,T]$

$$\mathbb{E}\int_t^T \|J(r)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\,dr \leq 4\,\mathbb{E}\,\|\zeta\|^2_{\mathcal{H}}\,.$$

Similarly, by equation (3.43) and Theorem 3.62 (i) and (iii), we get for all $s,t \in [0,T]$

$$\mathbb{E}\int_t^s \|K(s,r)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\, dr \leq 4\,\mathbb{E}\,\|f(s)\|^2_{\mathcal{H}}\,.$$

Due to Theorem 2.35, the Cauchy-Schwarz inequality, Fubini's theorem and Corollary 2.32, there exists a constant $c > 0$ such that for all $t \in [0,T]$

$$\begin{aligned}
&\mathbb{E}\int_t^T \|A^\varepsilon \phi(r)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dr \\
&\leq 2\,\mathbb{E}\int_t^T \|e^{-A(T-r)}J(r)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dr + 2\,\mathbb{E}\int_t^T \left(\int_r^T \left\|A^\varepsilon e^{-A(s-r)}K(s,r)\right\|_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} ds\right)^2 dr \\
&\leq 2\,\mathbb{E}\int_t^T \|J(r)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dr + \frac{2M_\varepsilon^2(T-t)^{1-2\varepsilon}}{1-2\varepsilon}\,\mathbb{E}\int_t^T\int_t^s \|K(s,r)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\, dr\, ds \\
&\leq c\left[\mathbb{E}\,\|\zeta\|^2_{D(A^\delta)} + (T-t)^{1-2\varepsilon}\,\mathbb{E}\int_t^T \|f(s)\|^2_{\mathcal{H}} ds\right].
\end{aligned}$$

Note that equation (3.44), inequality (3.39) and inequality (3.40) also hold for an $\mathcal{H}$-valued predictable process $(f(t))_{t\in[0,T]}$ such that $\mathbb{E}\int_0^T \|f(t)\|^2_{\mathcal{H}} dt < \infty$, which is an immediate consequence of the fact that $C([0,T];L^2(\Omega;\mathcal{H}))$ is dense in $L^2([0,T];L^2(\Omega,\mathcal{H}))$.

Finally, we prove that the pair of stochastic processes $(\varphi(t),\phi(t))_{t\in[0,T]}$ is unique. Let $(\varphi_1(t),\phi_1(t))_{t\in[0,T]}$ and $(\varphi_2(t),\phi_2(t))_{t\in[0,T]}$ satisfy equation (3.38). Then we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\varphi_1(t) - \varphi_2(t) = -\int_t^T e^{-A(s-t)}A^\varepsilon(\phi_1(s) - \phi_2(s))\, dW(s).$$

We obtain that the pair of processes $(\varphi_1(t) - \varphi_2(t))_{t\in[0,T]}$ and $(\phi_1(t) - \phi_2(t))_{t\in[0,T]}$ fulfills equation (3.38) with $\zeta = 0$ and $f = 0$. By inequality (3.39), we have $\varphi_1(t) = \varphi_2(t)$ for all $t \in [0,T]$ and $\mathbb{P}$-almost surely. Using inequality (3.40), we get $\phi_1(t) = \phi_2(t)$ for almost all $t \in [0,T]$ and $\mathbb{P}$-almost surely. ☐

Corollary 3.97. *Let $\delta \in [0,1)$ and $\varepsilon \in [0,\frac{1}{2})$ satisfy $\delta + \varepsilon < 1$. Furthermore, let $\zeta \in L^2(\Omega;D(A^\delta))$ be an $\mathcal{F}_T$-measurable random variable and let $(f(t))_{t\in[0,T]}$ be a predictable stochastic process with values in $\mathcal{H}$ such that $\mathbb{E}\sup_{t\in[0,T]}\|f(t)\|^2_{\mathcal{H}} < \infty$. Then there exists a unique pair of predictable processes $(\varphi(t),\phi(t))_{t\in[0,T]}$ with values in $D(A^\delta) \times \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});D(A^\varepsilon))$ such that for all $t \in [0,T]$ and $\mathbb{P}$-a.s.*

$$\varphi(t) = e^{-A(T-t)}\zeta + \int_t^T A^\varepsilon e^{-A(s-t)}f(s)\, ds - \int_t^T e^{-A(s-t)}A^\varepsilon\phi(s)\, dW(s).$$

Moreover, there exists a constant $\hat{c} > 0$ such that for all $t \in [0,T]$

$$\mathbb{E}\sup_{s\in[t,T]}\|\varphi(s)\|^2_{D(A^\delta)} \leq \hat{c}\left[\mathbb{E}\,\|\zeta\|^2_{D(A^\delta)} + (T-t)^{2-2\delta-2\varepsilon}\,\mathbb{E}\sup_{s\in[t,T]}\|f(s)\|^2_{\mathcal{H}}\right], \tag{3.45}$$

$$\mathbb{E}\int_t^T \|\phi(s)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});D(A^\varepsilon))} ds \leq \hat{c}\left[\mathbb{E}\,\|\zeta\|^2_{D(A^\delta)} + (T-t)^{2-2\varepsilon}\,\mathbb{E}\sup_{s\in[t,T]}\|f(s)\|^2_{\mathcal{H}}\right]. \tag{3.46}$$

Proof. The proof can be obtained similarly to Lemma 3.96. □

Based on the above results, we are able to prove the existence and uniqueness of the mild solution to system (3.32).

Theorem 3.98. *Let $\alpha,\beta,\gamma,\delta \in [0,\frac{1}{2})$ satisfy $\beta+\delta<\frac{1}{2}$ and $\gamma+\delta<\frac{1}{2}$. Moreover, let $(g(t))_{t\in[0,T]}$ be fixed. For any $Z \in L^2(\Omega;D(A^\delta))$, there exists a unique mild solution $(z(t),\Phi(t))_{t\in[0,T]}$ of system (3.32).*

Proof. Let $\mathcal{Z}^1_T$ contain all $D(A^\delta)$-valued predictable processes $(\tilde{z}(t))_{t\in[0,T]}$ with $\mathbb{E}\sup_{t\in[0,T]}\|\tilde{z}(t)\|^2_{D(A^\delta)}<\infty$. The space $\mathcal{Z}^1_T$ equipped with the norm

$$\|\tilde{z}\|^2_{\mathcal{Z}^1_T} = \mathbb{E}\sup_{t\in[0,T]}\|\tilde{z}(t)\|^2_{D(A^\delta)}$$

for every $\tilde{z}\in\mathcal{Z}^1_T$ becomes a Banach space. Similarly, let $\mathcal{Z}^2_T$ denote the space of all predictable processes $(\tilde{\Phi}(t))_{t\in[0,T]}$ with values in $\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})$ such that $\mathbb{E}\int_0^T\|\tilde{\Phi}(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}dt<\infty$. The space $\mathcal{Z}^2_T$ equipped with the inner product

$$\left\langle\tilde{\Phi}_1,\tilde{\Phi}_2\right\rangle_{\mathcal{Z}^2_T} = \mathbb{E}\int_0^T\left\langle\tilde{\Phi}_1(t),\tilde{\Phi}_2(t)\right\rangle_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dt$$

for every $\tilde{\Phi}_1,\tilde{\Phi}_2\in\mathcal{Z}^2_T$ becomes a Hilbert space.

Next, we define a sequence $(z_n,\Phi_n)_{n\in\mathbb{N}}\subset\mathcal{Z}^1_T\times\mathcal{Z}^2_T$ satisfying for each $n\in\mathbb{N}$, all $t\in[0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} z_n(t) = e^{-A(T-t)}Z &+ \int_t^T A^\alpha e^{-A(s-t)}B(z_{n-1}(s))\,ds + \int_t^T A^\beta e^{-A(s-t)}G(\Phi_{n-1}(s))\,ds + \int_t^T A^\gamma e^{-A(s-t)}g(s)\,ds \\ &- \int_t^T e^{-A(s-t)}\Phi_n(s)\,dW(s), \end{aligned} \tag{3.47}$$

where $z_0(t)=0$ and $\Phi_0(t)=0$ for all $t\in[0,T]$. Note that by Lemma 3.96, Corollary 3.97, inequality (3.33) and inequality (3.35), one can easily verify that $(z_n,\Phi_n)_{n\in\mathbb{N}}\subset\mathcal{Z}^1_T\times\mathcal{Z}^2_T$. Furthermore, we obtain for each $n\in\mathbb{N}$, all $t\in[0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} z_{n+1}(t)-z_n(t) &= \int_t^T A^\alpha e^{-A(s-t)}[B(z_n(s))-B(z_{n-1}(s))]\,ds + \int_t^T A^\beta e^{-A(s-t)}[G(\Phi_n(s))-G(\Phi_{n-1}(s))]\,ds \\ &\quad - \int_t^T e^{-A(s-t)}[\Phi_{n+1}(s)-\Phi_n(s)]\,dW(s). \end{aligned} \tag{3.48}$$

Using inequality (3.34) and inequality (3.36), we have for each $n\in\mathbb{N}$

$$\mathbb{E}\sup_{t\in[0,T]}\|B(z_n(t))-B(z_{n-1}(t))\|^2_{\mathcal{H}} \leq \widetilde{C}^2\,\mathbb{E}\sup_{t\in[0,T]}\|z_n(t)-z_{n-1}(t)\|^2_{D(A^\delta)},$$

$$\mathbb{E}\int_0^T\|G(\Phi_n(t))-G(\Phi_{n-1}(t))\|^2_{\mathcal{H}}\,dt \leq \widehat{C}^2\,\mathbb{E}\int_0^T\|\Phi_n(t)-\Phi_{n-1}(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\,dt.$$

Hence, equation (3.48) satisfies the assumptions of Lemma 3.96 and Corollary 3.97. Let $T_1 \in [0, T)$. Due to inequality (3.39) and inequality (3.45), there exist constants $C_1, C_2 > 0$ such that for each $n \in \mathbb{N}$

$$\begin{aligned}
\mathbb{E} \sup_{t\in[T_1,T]} \|z_{n+1}(t) - z_n(t)\|^2_{D(A^\delta)} &\leq \hat{c}(T - T_1)^{2-2\alpha-2\delta}\, \mathbb{E} \sup_{t\in[T_1,T]} \|B(z_n(t)) - B(z_{n-1}(t))\|^2_{\mathcal{H}} \\
&\quad + c(T - T_1)^{1-2\beta-2\delta}\, \mathbb{E} \int_{T_1}^{T} \|G(\Phi_n(t)) - G(\Phi_{n-1}(t))\|^2_{\mathcal{H}}\, dt \\
&\leq C_1(T - T_1)^{2-2\alpha-2\delta}\, \mathbb{E} \sup_{t\in[T_1,T]} \|z_n(t) - z_{n-1}(t)\|^2_{D(A^\delta)} \\
&\quad + C_2(T - T_1)^{1-2\beta-2\delta}\, \mathbb{E} \int_{T_1}^{T} \|\Phi_n(t) - \Phi_{n-1}(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\, dt.
\end{aligned}$$

Using inequality (3.40) and inequality (3.46), we get for each $n \in \mathbb{N}$

$$\begin{aligned}
\mathbb{E} \int_{T_1}^{T} \|\Phi_{n+1}(s) - \Phi_n(s)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\, dt &\leq \hat{c}(T - T_1)^{2-2\alpha}\, \mathbb{E} \sup_{t\in[T_1,T]} \|B(z_n(t)) - B(z_{n-1}(t))\|^2_{\mathcal{H}} \\
&\quad + c(T - T_1)^{1-2\beta}\, \mathbb{E} \int_{T_1}^{T} \|G(\Phi_n(t)) - G(\Phi_{n-1}(t))\|^2_{\mathcal{H}}\, dt \\
&\leq C_1(T - T_1)^{2-2\alpha}\, \mathbb{E} \sup_{t\in[T_1,T]} \|z_n(t) - z_{n-1}(t)\|^2_{D(A^\delta)} \\
&\quad + C_2(T - T_1)^{1-2\beta}\, \mathbb{E} \int_{T_1}^{T} \|\Phi_n(t) - \Phi_{n-1}(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\, dt.
\end{aligned}$$

Hence, we obtain for each $n \in \mathbb{N}$

$$\begin{aligned}
&\mathbb{E} \sup_{t\in[T_1,T]} \|z_{n+1}(t) - z_n(t)\|^2_{D(A^\delta)} + \mathbb{E} \int_{T_1}^{T} \|\Phi_{n+1}(s) - \Phi_n(s)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\, dt \\
&\leq C \left[\mathbb{E} \sup_{t\in[T_1,T]} \|z_n(t) - z_{n-1}(t)\|^2_{D(A^\delta)} + \mathbb{E} \int_{T_1}^{T} \|\Phi_n(t) - \Phi_{n-1}(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\, dt \right],
\end{aligned}$$

where $C = \max\{C_1((T - T_1)^{2-2\alpha-2\delta} + (T - T_1)^{2-2\alpha}), C_2((T - T_1)^{1-2\beta-2\delta} + (T - T_1)^{1-2\beta})\}$. Therefore, we find for each $n \in \mathbb{N}$

$$\begin{aligned}
&\mathbb{E} \sup_{t\in[T_1,T]} \|z_{n+1}(t) - z_n(t)\|^2_{D(A^\delta)} + \mathbb{E} \int_{T_1}^{T} \|\Phi_{n+1}(s) - \Phi_n(s)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\, dt \\
&\leq C^n \left[\mathbb{E} \sup_{t\in[T_1,T]} \|z_1(t)\|^2_{D(A^\delta)} + \mathbb{E} \int_{T_1}^{T} \|\Phi_1(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\, dt \right].
\end{aligned}$$

We choose $T_1 \in [0, T)$ such that $C < 1$. Thus, we can conclude that $(z_n, \Phi_n)_{n\in\mathbb{N}} \subset \mathcal{Z}_T^1 \times \mathcal{Z}_T^2$ is a Cauchy sequence on the interval $[T_1, T]$. Using equation (3.48), we have for each $n \in \mathbb{N}$, all $t \in [0, T_1]$ and $\mathbb{P}$-a.s.

$$z_{n+1}(t) - z_n(t) = e^{-A(T_1-t)}[z_{n+1}(T_1) - z_n(T_1)] + \int_t^{T_1} A^\alpha e^{-A(s-t)}[B(z_n(s)) - B(z_{n-1}(s))]\, ds$$
$$+ \int_t^{T_1} A^\beta e^{-A(s-t)}[G(\Phi_n(s)) - G(\Phi_{n-1}(s))]\, ds - \int_t^{T_1} e^{-A(s-t)}[\Phi_{n+1}(s) - \Phi_n(s)]\, dW(s).$$

Again, we find $T_2 \in [0, T_1]$ such that $(z_n, \Phi_n)_{n\in\mathbb{N}} \subset \mathcal{Z}_T^1 \times \mathcal{Z}_T^2$ is a Cauchy sequence on the interval $[T_2, T_1]$. By continuing this method, we can conclude that $(z_n, \Phi_n)_{n\in\mathbb{N}} \subset \mathcal{Z}_T^1 \times \mathcal{Z}_T^2$ is a Cauchy sequence on the interval $[0, T]$. Hence, there exist $z \in \mathcal{Z}_T^1$ and $\Phi \in \mathcal{Z}_T^2$ such that

$$z = \lim_{n\to\infty} z_n, \quad \Phi = \lim_{n\to\infty} \Phi_n.$$

Using equation (3.47), one can easily verify that the pair of stochastic processes $(z(t), \Phi(t))_{t\in[0,T]}$ satisfy equation (3.37). □

Remark 3.99. *Note that the proofs of Theorem 3.93 and Theorem 3.98 are mainly based on the martingale representation theorem stated in Proposition 3.86. According to Remark 3.88, one can also consider backward SPDEs driven by a real-valued Lévy process. However, we will focus on backward SPDEs driven by a Q-Wiener process due to the fact that we can model noise terms dependent on a spatial variable, which is more suitable for applications.*

3.4.3. A Comparison of Strong, Weak and Mild Solutions

In this section, we give a comparison between different concepts of solution to SPDEs, where the noise term $(W(t))_{t\geq 0}$ is an $\mathcal{H}$-valued Q-Wiener process with covariance operator $Q \in \mathcal{L}_1^+(\mathcal{H})$. We start with forward SPDEs. Let us consider the following nonlinear system in $\mathcal{H}$:

$$\begin{cases} dy(t) = [Ay(t) + B(y(t)) + f(t)]\, dt + G(y(t))\, dW(t), \\ y(0) = \xi. \end{cases} \tag{3.49}$$

We assume that

- the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ is the generator of a C_0 semigroup $(e^{At})_{t\geq 0}$;
- $B\colon \mathcal{H} \to \mathcal{H}$ satisfies for every $y, z \in \mathcal{H}$

  $$\|B(y)\|_{\mathcal{H}} \leq C\|y\|_{\mathcal{H}},$$
  $$\|B(y) - B(z)\|_{\mathcal{H}} \leq C\|y - z\|_{\mathcal{H}},$$

  where $C > 0$ is a constant;
- the process $(f(t))_{t\in[0,T]}$ is $\mathcal{F}_t$-adapted and takes values in $\mathcal{H}$ such that

  $$\mathbb{E}\int_0^T \|f(t)\|_{\mathcal{H}}^2\, dt < \infty;$$

- $G\colon \mathcal{H} \to \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})$ satisfies for every $y, z \in \mathcal{H}$

$$\|G(y)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} \leq \widehat{C}\|y\|_{\mathcal{H}},$$
$$\|G(y) - G(z)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} \leq \widehat{C}\|y - z\|_{\mathcal{H}},$$

  where $\widehat{C} > 0$ is a constant;

- ξ is an $\mathcal{F}_0$-measurable random variable with values in $\mathcal{H}$.

Next, we introduce several concepts of a solution to system (3.49).

Definition 3.100. *A predictable process $(y(t))_{t\in[0,T]}$ with values in $\mathcal{H}$ is called a **strong solution of system (3.49)** if $(y(t))_{t\in[0,T]}$ takes values in $D(A)$ for almost all $t \in [0, T]$ such that $\mathbb{P}$-a.s. $\int_0^T \|Ay(t)\|_{\mathcal{H}} dt < \infty$,*

$$\sup_{t\in[0,T]} \mathbb{E}\|y(t)\|_{\mathcal{H}}^2 < \infty$$

and for all $t \in [0, T]$ and $\mathbb{P}$-a.s.

$$y(t) = \xi + \int_0^t [Ay(s) + B(y(s)) + f(s)]\, ds + \int_0^t G(y(s))\, dW(s).$$

Definition 3.101. *A predictable process $(y(t))_{t\in[0,T]}$ with values in $\mathcal{H}$ is called a **weak solution of system (3.49)** if*

$$\sup_{t\in[0,T]} \mathbb{E}\|y(t)\|_{\mathcal{H}}^2 < \infty$$

and for every $\psi \in D(A^*)$, all $t \in [0, T]$ and $\mathbb{P}$-a.s.

$$\langle y(t), \psi\rangle_{\mathcal{H}} = \langle \xi, \psi\rangle_{\mathcal{H}} + \int_0^t [\langle y(s), A^*\psi\rangle_{\mathcal{H}} + \langle B(y(s)), \psi\rangle_{\mathcal{H}} + \langle f(s), \psi\rangle_{\mathcal{H}}]\, ds + \int_0^t \langle G(y(s))\, dW(s), \psi\rangle_{\mathcal{H}}.$$

Definition 3.102. *A predictable process $(y(t))_{t\in[0,T]}$ with values in $\mathcal{H}$ is a **mild solution of system (3.49)** if*

$$\sup_{t\in[0,T]} \mathbb{E}\|y(t)\|_{\mathcal{H}}^2 < \infty$$

and for all $t \in [0, T]$ and $\mathbb{P}$-a.s.

$$y(t) = e^{At}\xi + \int_0^t e^{A(t-s)}\, [B(y(s)) + f(s)]\, ds + \int_0^t e^{A(t-s)}G(y(s))\, dW(s).$$

Existence and uniqueness results of these types of solution can be found in [23, 42, 73]. The following theorem gives relationships between these solutions.

Theorem 3.103. *If $(y(t))_{t\in[0,T]}$ is a strong solution of system (3.49), then it is a weak solution. The process $(y(t))_{t\in[0,T]}$ is a weak solution of system (3.49) if and only if it is a mild solution.*

Proof. The fact that a strong solution is also a weak solution follows immediately from the definitions. A proof of the equivalence of weak and mild solutions can be found in [71, Theorem 9.15]. □

Remark 3.104. *Alternatively, one can show that a strong solution of system (3.49) is a mild solution, see [53, Proposition 2.1]. Furthermore, the previous theorem can be shown for SPDEs driven by Lévy noise, see [71, Theorem 9.15].*

To obtain the converse of the previous theorem, additional assumptions are required.

Theorem 3.105. *Let $(y(t))_{t\in[0,T]}$ be a weak solution of system (3.49). If $(y(t))_{t\in[0,T]}$ takes values in $D(A)$ for almost all $t \in [0,T]$ such that $\mathbb{P}$-a.s. $\int_0^T \|Ay(t)\|_{\mathcal{H}} dt < \infty$, then it is a strong solution.*

Proof. The proof follows immediately from the definitions. □

Alternatively, one can give conditions on a mild solution to be a strong solution.

Theorem 3.106 (Proposition 2.3,[53])**.** *Let $(y(t))_{t\in[0,T]}$ be a mild solution of system (3.49). Suppose that*

- *ξ is an $D(A)$-valued random variable and $(f(t))_{t\in[0,T]}$ takes values in $D(A)$ for almost all $t \in [0,T]$;*
- *for all $t \in (0,T]$ and every $y \in \mathcal{H}$, we have $e^{At}B(y) \in D(A)$ and*

$$\|Ae^{At}B(y)\|_{\mathcal{H}} \leq g_1(t)\|y\|_{\mathcal{H}},$$

  where $g_1 \in L^1([0,T])$;
- *for all $t \in (0,T]$ and every $y \in \mathcal{H}$, we have $e^{At}G(y) \in \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); D(A))$ and*

$$\|Ae^{At}G(y)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} \leq g_2(t)\|y\|_{\mathcal{H}},$$

  where $g_2 \in L^2([0,T])$.

Then $(y(t))_{t\in[0,T]}$ is also a strong solution of system (3.49).

Next, we consider the following backward SPDE in $\mathcal{H}$:

$$\begin{cases} dz(t) = -[Az(t) + G(z(t), \Phi(t)) + g(t)]dt + \Phi(t)\, dW(t), \\ z(T) = Z. \end{cases} \tag{3.50}$$

We assume that

- the operator $A\colon D(A) \subset \mathcal{H} \to \mathcal{H}$ is the generator of a C_0 semigroup $(e^{At})_{t\geq 0}$;
- $G\colon \mathcal{H} \times \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H}) \to \mathcal{H}$ satisfies for every $y, z \in \mathcal{H}$ and every $\Phi, \Psi \in L_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})$

$$\begin{aligned} \|G(y,\Phi)\|_{\mathcal{H}} &\leq \widehat{C}\left[\|y\|_{\mathcal{H}} + \|\Phi\|_{L_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\right], \\ \|G(y,\Phi) - G(z,\Psi)\|_{\mathcal{H}} &\leq \widehat{C}\left[\|y-z\|_{\mathcal{H}} + \|\Phi - \Psi\|_{L_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\right], \end{aligned}$$

  where $\widehat{C} > 0$ is a constant;
- $(g(t))_{t\in[0,T]}$ is a predictable process with values in $\mathcal{H}$ such that

$$\mathbb{E}\int_0^T \|g(t)\|_{\mathcal{H}}^2\, dt < \infty;$$

- Z is an $\mathcal{F}_T$-measurable random variable with values in $\mathcal{H}$.

Again, we introduce several concepts of a solution to system (3.50).

Definition 3.107. *A pair of predictable processes $(z(t), \Phi(t))_{t\in[0,T]}$ with values in $\mathcal{H} \times \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})$ is called a* ***strong solution of system (3.50)*** *if $(z(t))_{t\in[0,T]}$ takes values in $D(A)$ for almost all $t \in [0,T]$ such that $\mathbb{P}$-a.s. $\int_0^T \|Az(t)\|_{\mathcal{H}} dt < \infty$,*

$$\sup_{t\in[0,T]} \mathbb{E}\,\|z(t)\|^2_{\mathcal{H}} < \infty, \qquad \mathbb{E}\int_0^T \|\Phi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dt < \infty$$

and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$z(t) = Z + \int_t^T [Az(s) + G(z(s), \Phi(s)) + g(s)]\, ds - \int_t^T \Phi(s)\, dW(s).$$

Definition 3.108. *A pair of predictable processes $(z(t), \Phi(t))_{t\in[0,T]}$ with values in $\mathcal{H} \times \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})$ is called a* ***weak solution of system (3.50)*** *if*

$$\sup_{t\in[0,T]} \mathbb{E}\,\|z(t)\|^2_{\mathcal{H}} < \infty, \qquad \mathbb{E}\int_0^T \|\Phi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dt < \infty$$

and we have for every $\psi \in D(A^*)$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\langle z(t), \psi\rangle_{\mathcal{H}} = Z + \int_t^T [\langle z(s), A^*\psi\rangle_{\mathcal{H}} + \langle G(z(s), \Phi(s)), \psi\rangle_{\mathcal{H}} + \langle g(s), \psi\rangle_{\mathcal{H}}]\, ds - \int_t^T \langle \Phi(s)\, dW(s), \psi\rangle_{\mathcal{H}}.$$

Definition 3.109. *A pair of predictable processes $(z(t), \Phi(t))_{t\in[0,T]}$ with values in $\mathcal{H} \times \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})$ is called a* ***mild solution of system (3.50)*** *if*

$$\sup_{t\in[0,T]} \mathbb{E}\,\|z(t)\|^2_{\mathcal{H}} < \infty, \qquad \mathbb{E}\int_0^T \|\Phi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})} dt < \infty$$

and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$z(t) = e^{A(T-t)} Z + \int_t^T e^{A(s-t)} [G(z(s), \Phi(s)) + g(s)]\, ds - \int_t^T e^{A(s-t)} \Phi(s)\, dW(s).$$

The following theorem gives relationships between these solutions.

Theorem 3.110 (Theorem 3.4,[1])**.** *If $(z(t), \Phi(t))_{t\in[0,T]}$ is a strong solution of system (3.50), then it is a weak solution. The pair of processes $(z(t), \Phi(t))_{t\in[0,T]}$ is a weak solution of system (3.50) if and only if it is a mild solution.*

To obtain the converse of the previous theorem, additional assumptions are required.

Theorem 3.111 (Theorem 4.1,[1]). *Let $(z(t), \Phi(t))_{t\in[0,T]}$ be a weak solution of system (3.50). If $(z(t))_{t\in[0,T]}$ takes values in $D(A)$ for almost all $t \in [0,T]$ such that $\mathbb{P}$-a.s. $\int_0^T \|Az(t)\|_{\mathcal{H}} dt < \infty$, then it is a strong solution.*

Alternatively, one can give conditions on a mild solution to be a strong solution.

Theorem 3.112. *Let $(z(t), \Phi(t))_{t\in[0,T]}$ be a mild solution of system (3.50). Suppose that*

- *Z is an $D(A)$-valued random variable and $(g(t))_{t\in[0,T]}$ take values in $D(A)$ for almost all $t \in [0,T]$;*
- *for all $t \in (0,T]$, every $z \in \mathcal{H}$ and every $\Phi \in \mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})$, we have $e^{At}G(y,\Phi) \in D(A)$ and*

$$\|Ae^{At}G(z,\Phi)\|_{\mathcal{H}} \leq h(t)\left[\|z\|_{\mathcal{H}} + \|\Phi\|_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H});\mathcal{H})}\right],$$

  where $h \in L^1([0,T])$.

Then $(z(t), \Phi(t))_{t\in[0,T]}$ is also a strong solution of system (3.50).

Proof. Since the process $(z(t))_{t\in[0,T]}$ is predictable, we get for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$z(t) = \mathbb{E}\left[e^{A(T-t)}Z \,\middle|\, \mathcal{F}_t\right] + \mathbb{E}\left[\int_t^T e^{A(s-t)}[G(z(s),\Phi(s)) + g(s)]\,ds \,\middle|\, \mathcal{F}_t\right].$$

Using the assumptions, one can easily verify that the terms on the right hand side takes values in $D(A)$. Thus, the process $(z(t))_{t\in[0,T]}$ takes values in $D(A)$ such that

$$\mathbb{E}\int_0^T \|Az(t)\|_{\mathcal{H}} dt < \infty.$$

Therefore, we can apply Theorem 3.111 and the claim follows. □

Remark 3.113. *Under stronger assumptions, the previous theorem was also proven in [1, Theorem 4.2].*

Existence and uniqueness results of mild solution can be found in [52]. Theorem 3.110 and Theorem 3.112 gives requirements such that unique weak solutions as well as unique strong solutions exists.

Chapter 4

Optimal Control of Uncertain Heat Distributions

In this chapter, we consider a control problem constrained by the stochastic heat equation with nonhomogeneous Neumann boundary conditions. Here, controls and noise terms are defined inside the domain as well as on the boundary. We first recall some well known facts of the deterministic unsteady heat equation with nonhomogeneous Neumann boundary conditions studied in [9]. The main idea is to reformulate this equation as an evolution equation in a suitable Hilbert space using the theory of fractional powers to closed operators provided in Section 2.3. This approach gives us a motivation how to involve noise terms inside the domain as well as on the boundary. The existence and uniqueness of a mild solution to the stochastic heat equation can be obtained using results shown in Section 3.4.1. Consequently, we are able to solve uniquely a linear quadratic control problem through a stochastic maximum principle, which gives us explicit formulas for the optimal controls. By a reformulation of these formulas, we finally obtain that the optimal controls satisfy a certain feedback law. Here, we mainly use the results shown in [5].

Throughout this chapter, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given complete probability space endowed with a normal filtration $(\mathcal{F}_t)_{t\geq 0}$.

4.1. Motivation

In this section, we introduce the deterministic unsteady heat equation with nonhomogeneous Neumann boundary data. For more details, see [9, Part IV]. Through a reformulation as an evolution equation, we get a motivation to involve noise terms in this equation. Let $\mathcal{D} \subset \mathbb{R}^n$, $n \in \mathbb{N}$, be a bounded domain with C^∞ boundary $\partial\mathcal{D}$ and let $T > 0$. We introduce the following controlled partial differential equation with nonhomogeneous Neumann boundary data:

$$\begin{cases} \dfrac{\partial}{\partial t} y(t,x) = \Delta y(t,x) + b(x)u(t,x) & \text{in } (0,T) \times \mathcal{D}, \\ y(0,x) = \xi(x) & \text{in } \mathcal{D}, \\ \dfrac{\partial}{\partial \eta} y(t,x) = v(t,x) & \text{on } (0,T) \times \partial\mathcal{D}, \end{cases} \tag{4.1}$$

where $y(t,x) \in \mathbb{R}$ is a heat distribution with initial value $\xi(x) \in \mathbb{R}$, $u(t,x) \in \mathbb{R}$ represents a distributed control and $v(t,x) \in \mathbb{R}$ denotes the boundary control. The operator Δ is the Laplace operator in $L^2(\mathcal{D})$ and η represents the outward normal to $\partial\mathcal{D}$.

Remark 4.1. *In application, we often have*

$$b(x) = \begin{cases} 1 & \text{if } x \in \mathcal{S} \\ 0 & \text{if } x \in \mathcal{D} \backslash \mathcal{S}, \end{cases}$$

where $\mathcal{S}$ is a subset of $\mathcal{D}$.

Next, we state a solution to system (4.1). According to Section 2.5.1, we introduce the Neumann realization of the Laplace operator $A\colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ by

$$D(A) = \left\{ y \in H^2(\mathcal{D}) : \frac{\partial y}{\partial \eta} = 0 \text{ on } \partial\mathcal{D} \right\}, \quad Ay = \Delta y$$

for every $y \in D(A)$. By Theorem 2.57, the operator $A\colon D(A) \subset L^2(\mathcal{D}) \to L^2(\mathcal{D})$ is the generator of an analytic semigroup of contractions $\left(e^{At}\right)_{t\geq 0}$. As discussed in the previous chapters, Theorem 2.35 is the main auxiliary result to involve nonhomogeneous boundary conditions. This theorem requires especially that 0 is an element of the resolvent set of A, which does not hold due to Remark 2.55. However, the operator $A - \lambda$ with $\lambda > 0$ is still the generator of an analytic semigroup $\left(e^{-\lambda t}e^{At}\right)_{t\geq 0}$ such that fractional powers denoted by $(\lambda - A)^\alpha$ with $\alpha \in \mathbb{R}$ are well defined. Moreover, we have $0 \in \rho(A - \lambda)$. For the convenience of the reader, we will give an overview of main properties of the operator $(\lambda - A)^\alpha$, which follows immediately from results stated in Section 2.3.

Corollary 4.2. *We have*

(i) for all $\alpha > 0$, the domain $D((\lambda - A)^\alpha)$ equipped with the inner product

$$\langle y, z\rangle_{D((\lambda-A)^\alpha)} = \langle (\lambda - A)^\alpha y, (\lambda - A)^\alpha z\rangle_{L^2(\mathcal{D})}$$

for every $y, z \in D((\lambda - A)^\alpha)$ becomes a Hilbert space;

(ii) $(\lambda - A)^{\alpha+\beta} y = (\lambda - A)^\alpha(\lambda - A)^\beta y$ for all $\alpha, \beta \in \mathbb{R}$ and every $y \in D(A^\gamma)$ with $\gamma = \max\{\alpha, \beta, \alpha + \beta\}$;

(iii) $\|y\|_{L^2(\mathcal{D})} \leq C\|(\lambda - A)^\alpha y\|_{L^2(\mathcal{D})}$ for every $y \in D((\lambda - A)^\alpha)$, where $C > 0$ is a constant;

(iv) $e^{At}\colon L^2(\mathcal{D}) \to D((\lambda - A)^\alpha)$ for all $t > 0$ and all $\alpha \in \mathbb{R}$;

(v) $(\lambda - A)^\alpha e^{At} y = e^{At}(\lambda - A)^\alpha y$ for every $y \in D((\lambda - A)^\alpha)$ and all $\alpha \in \mathbb{R}$;

(vi) the operator $(\lambda - A)^\alpha e^{At}$ is linear and bounded for all $t > 0$ and all $\alpha \in \mathbb{R}$. In addition, there exist constants $M_\alpha, \delta > 0$ such that for all $t > 0$ and all $\alpha > 0$

$$\|(\lambda - A)^\alpha e^{At}\|_{L^2(\mathcal{D})} \leq M_\alpha t^{-\alpha} e^{-\delta t};$$

(vii) the operator $(\lambda - A)^\alpha$ is self-adjoint for all $\alpha \in \mathbb{R}$.

Furthermore, we define the Neumann operator $N\colon L^2(\partial\mathcal{D}) \to L^2(\mathcal{D})$ by $g = Nh$ with

$$\Delta g(x) = \lambda g(x) \quad \text{in } \mathcal{D}, \quad \frac{\partial}{\partial \eta} g(x) = h(x) \quad \text{on } \partial\mathcal{D},$$

where $\lambda > 0$ is introduced above. In [60, Chapter 2], the result $N \in \mathcal{L}\left(L^2(\partial\mathcal{D}); H^{3/2}(\mathcal{D})\right)$ was proven. Due to Theorem 2.58, we can conclude $N \in \mathcal{L}(L^2(\partial\mathcal{D}); D((\lambda - A)^\alpha))$ if $\alpha \in \left(0, \frac{3}{4}\right)$ and by the closed graph theorem, we have $(\lambda - A)^\alpha N \in \mathcal{L}(L^2(\partial\mathcal{D}); L^2(\mathcal{D}))$. As shown in [9], we can rewrite system (4.1) in the following form:

$$\begin{cases} \dfrac{d}{dt} y(t) = Ay(t) + Bu(t) + (\lambda - A)Nv(t), \\ y(0) = \xi, \end{cases} \tag{4.2}$$

where $y(t)(x) = y(t, x)$, $Bu(t)(x) = b(x)u(t, x)$ and $v(t)(x) = v(t, x)$ are interpreted as abstract functions. For more details on abstract functions, we refer to [86, Section 3.4.1]. We note that the operator B is linear and bounded on $L^2(\mathcal{D})$. One can show that system (4.2) has a unique solution $y \in C([0,T]; L^2(\mathcal{D}))$ given by

$$y(t) = e^{At}\xi + \int_0^t e^{A(t-s)}Bu(s)\,ds + \int_0^t (\lambda - A)e^{A(t-s)}Nv(s)\,ds.$$

Remark 4.3. *In [9], the existence and uniqueness of the solution to system (4.2) is proved for the special cases $u = 0$ as well as $v = 0$. Hence, the existence of a unique solution to system (4.2) in the general setting follows immediately.*

4.2. A Controlled Linear Stochastic Heat Equation

In this section, we introduce the stochastic heat equation and we show some basic properties. Motivated by system (4.2), we consider the following SPDE in $L^2(\mathcal{D})$:

$$\begin{cases} dy(t) = [Ay(t) + Bu(t) + (\lambda - A)Nv(t)]\,dt + G(t)\,dW(t) + (\lambda - A)N\,dW_b(t), \\ y(0) = \xi, \end{cases} \tag{4.3}$$

where the initial value $\xi \in L^2(\Omega; L^2(\mathcal{D}))$ is $\mathcal{F}_0$-measurable. The set of admissible distributed controls U contains all $\mathcal{F}_t$-adapted processes $(u(t))_{t\in[0,T]}$ with values in $L^2(\mathcal{D})$ such that

$$\mathbb{E}\int_0^T \|u(t)\|^2_{L^2(\mathcal{D})}dt < \infty.$$

The space U equipped with the inner product of $L^2(\Omega; L^2([0,T]; L^2(\mathcal{D})))$ becomes a Hilbert space. Similarly, the set of admissible boundary controls V contains all $\mathcal{F}_t$-adapted processes $(v(t))_{t\in[0,T]}$ with values in $L^2(\partial\mathcal{D})$ such that

$$\mathbb{E}\int_0^T \|v(t)\|^2_{L^2(\partial\mathcal{D})}dt < \infty.$$

The space V equipped with the inner product of $L^2(\Omega; L^2([0,T]; L^2(\partial\mathcal{D})))$ becomes a Hilbert space. The stochastic processes $(W(t))_{t\geq 0}$ and $(W_b(t))_{t\geq 0}$ are Q-Wiener processes with values in $L^2(\mathcal{D})$ and $L^2(\partial\mathcal{D})$, respectively. We denote by $Q \in \mathcal{L}_1^+(L^2(\mathcal{D}))$ and $Q_b \in \mathcal{L}_1^+(L^2(\partial\mathcal{D}))$ the covariance operators of the processes $(W(t))_{t\geq 0}$ and $(W_b(t))_{t\geq 0}$, respectively. The process $(G(t))_{t\in[0,T]}$ is a predictable process with values in $\mathcal{L}_{(HS)}(Q^{1/2}(L^2(\mathcal{D})); L^2(\mathcal{D}))$ such that

$$\mathbb{E}\int_0^T \|G(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(L^2(\mathcal{D})); L^2(\mathcal{D}))}dt < \infty.$$

Motivated by Section 4.1, we use a mild solution to system (4.3) in the sense of Definition 3.72 with $\mathcal{H} = L^2(\mathcal{D})$ and $\mathcal{H}_b = L^2(\partial\mathcal{D})$. As a consequence of Theorem 3.73, there exists a unique mild solution $(y(t))_{t\in[0,T]}$ of system (4.3) for any $\xi \in L^2(\Omega; L^2(\mathcal{D}))$ and fixed controls $u \in U$ and $v \in V$. Hence, the process $(y(t))_{t\in[0,T]}$ takes values in $L^2(\mathcal{D})$ such that

$$\sup_{t\in[0,T]} \mathbb{E}\|y(t)\|^2_{L^2(\mathcal{D})} < \infty$$

and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$y(t) = e^{At}\xi + \int_0^t e^{A(t-s)}Bu(s)\,ds + \int_0^t (\lambda - A)e^{A(t-s)}Nv(s)\,ds + \int_0^t e^{A(t-s)}G(s)\,dW(s) + \int_0^t (\lambda - A)e^{A(t-s)}N\,dW_b(s). \tag{4.4}$$

For the remaining part of this chapter, let the initial value $\xi \in L^2(\Omega; L^2(\mathcal{D}))$ be fixed. To illustrate the dependence on the controls $u \in U$ and $v \in V$, we denote by $(y(t;u,v))_{t\in[0,T]}$ the mild solution of system (4.3). We get the following properties.

Lemma 4.4. *Let $(y(t;u,v))_{t\in[0,T]}$ be the mild solution of system (4.3) corresponding to the controls $u \in U$ and $v \in V$. Then*

(i) $y(t;u,v)$ is affine linear in both $u \in U$ and $v \in V$ for all $t \in [0,T]$ and $\mathbb{P}$-almost surely;

(ii) there exist constants $c_1, c_2 > 0$ such that for every $u_1, u_2 \in U$ and every $v_1, v_2 \in V$

$$\sup_{t\in[0,T]} \mathbb{E}\|y(t;u_1,v_1) - y(t;u_2,v_2)\|^2_{L^2(\mathcal{D})} \leq c_1\,\mathbb{E}\int_0^T \|u_1(t) - u_2(t)\|^2_{L^2(\mathcal{D})}dt + c_2\,\mathbb{E}\int_0^T \|v_1(t) - v_2(t)\|^2_{L^2(\partial\mathcal{D})}dt.$$

Proof. The claim (i) is an immediate consequence of equation (4.4). It remains to prove (ii). Let $u_1, u_2 \in U$ and $v_1, v_2 \in V$ be arbitrary. Recall that $B \in \mathcal{L}(L^2(\mathcal{D}))$ and $N \in \mathcal{L}(L^2(\partial\mathcal{D}); D((\lambda - A)^\alpha))$ if $\alpha \in \left(0, \frac{3}{4}\right)$. By definition, we get for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$y(t;u_1,v_1) - y(t;u_2,v_2) = \int_0^t e^{A(t-s)}B[u_1(s) - u_2(s)]\,ds + \int_0^t (\lambda - A)e^{A(t-s)}N[v_1(s) - v_2(s)]\,ds.$$

Using Corollary 4.2 and the Cauchy-Schwarz inequality, there exist constants $c_1, c_2 > 0$ such that for all $\alpha \in \left(\frac{1}{2}, \frac{3}{4}\right)$

$$\begin{aligned}
&\sup_{t\in[0,T]} \mathbb{E}\|y(t;u_1,v_1) - y(t;u_2,v_2)\|^2_{L^2(\mathcal{D})} \\
&\leq 2 \sup_{t\in[0,T]} \mathbb{E}\int_0^t \|e^{A(t-s)}B[u_1(s) - u_2(s)]\|^2_{L^2(\mathcal{D})}ds \\
&+ 2 \sup_{t\in[0,T]} \mathbb{E}\left(\int_0^t \|(\lambda - A)^{1-\alpha}e^{A(t-s)}(\lambda - A)^\alpha N[v_1(s) - v_2(s)]\|_{L^2(\mathcal{D})}ds\right)^2 \\
&\leq c_1\,\mathbb{E}\int_0^T \|u_1(t) - u_2(t)\|^2_{L^2(\mathcal{D})}dt + M^2_{1-\alpha} \sup_{t\in[0,T]} \mathbb{E}\left(\int_0^t (t-s)^{\alpha-1}\|(\lambda - A)^\alpha N[v_1(s) - v_2(s)]\|_{L^2(\mathcal{D})}ds\right)^2 \\
&\leq c_1\,\mathbb{E}\int_0^T \|u_1(t) - u_2(t)\|^2_{L^2(\mathcal{D})}dt + c_2\,\mathbb{E}\int_0^T \|v_1(t) - v_2(t)\|^2_{L^2(\partial\mathcal{D})}dt,
\end{aligned}$$

which completes the proof. □

4.3. A Tracking Problem of the Terminal State

In this section, we introduce the control problem. We state necessary and sufficient optimality conditions, which we use to derive explicit formulas for the optimal controls. Let us introduce the cost functional $J\colon U \times V \to \mathbb{R}$ as follows:

$$J(u,v) = \frac{1}{2}\mathbb{E}\|y(T;u,v) - y_d\|^2_{L^2(\mathcal{D})} + \frac{\kappa_1}{2}\mathbb{E}\int_0^T \|u(t)\|^2_{L^2(\mathcal{D})}dt + \frac{\kappa_2}{2}\mathbb{E}\int_0^T \|v(t)\|^2_{L^2(\partial\mathcal{D})}dt, \tag{4.5}$$

where $(y(t;u,v))_{t\in[0,T]}$ is the mild solution of system (4.3) corresponding to the controls $u \in U$ and $v \in V$. The function $y_d \in L^2(\mathcal{D})$ is a given desired state and $\kappa_1, \kappa_2 > 0$ are weights. The task is to find controls $\overline{u} \in U$ and $\overline{v} \in V$ such that

$$J(\overline{u},\overline{v}) = \inf_{u\in U, v\in V} J(u,v). \tag{4.6}$$

Then $\overline{u} \in U$ and $\overline{v} \in V$ are called optimal controls. The corresponding optimal state is denoted by $(\overline{y}(t))_{t\in[0,T]}$. Using Lemma 4.4, we can conclude that the cost functional J is coercive, strictly convex and continuous in both $u \in U$ and $v \in V$. Hence, the control problem (4.6) is formulated as a convex optimization problem on the Hilbert space $U \times V$. By Corollary D.13, we get the existence and uniqueness of the optimal controls $\overline{u} \in U$ and $\overline{v} \in V$.

Note that the cost functional J is partial Fréchet differentiable. Indeed, this results follows from the fact that J is a sum of squared norms and Lemma 4.4. The partial Fréchet derivatives of J can be obtained similarly to Remark D.6 and thus, we get

$$d_u^F J(u,v)[\tilde{u}] = \mathbb{E}\int_0^T \left\langle y(T;u,v) - y_d, e^{A(T-t)}B\tilde{u}(t)\right\rangle_{L^2(\mathcal{D})} dt + \kappa_1\,\mathbb{E}\int_0^T \langle u(t), \tilde{u}(t)\rangle_{L^2(\mathcal{D})}\,dt, \tag{4.7}$$

$$d_v^F J(u,v)[\tilde{v}] = \mathbb{E}\int_0^T \left\langle y(T;u,v) - y_d, (\lambda - A)e^{A(T-t)}N\tilde{v}(t)\right\rangle_{L^2(\mathcal{D})} dt + \kappa_2\,\mathbb{E}\int_0^T \langle v(t), \tilde{v}(t)\rangle_{L^2(\partial\mathcal{D})}\,dt, \tag{4.8}$$

where $\tilde{u} \in U$ and $\tilde{v} \in V$. Since the sets of admissible distributed controls U and boundary controls V are Hilbert spaces, we can apply Proposition D.14. Therefore, the optimal controls $\overline{u} \in U$ and $\overline{v} \in V$ satisfy the following necessary and sufficient optimality conditions:

$$d_u^F J(\overline{u},\overline{v})[\tilde{u}] = 0, \tag{4.9}$$

$$d_v^F J(\overline{u},\overline{v})[\tilde{v}] = 0 \tag{4.10}$$

for every $\tilde{u} \in U$ and every $\tilde{v} \in V$. In the remaining part of this section, we use these necessary and sufficient optimality conditions to derive explicit formulas for the optimal distributed control $\overline{u} \in U$ and the optimal boundary control $\overline{v} \in V$.

Theorem 4.5. *Let the cost functional $J\colon U \times V \to \mathbb{R}$ be given by (4.5). Then the optimal distributed control $\overline{u} \in U$ satisfies for almost all $t \in [0,T]$ and $\mathbb{P}$-a.s.*

$$\overline{u}(t) = -\frac{1}{\kappa_1}B^* e^{A(T-t)}\left(\mathbb{E}\left[\overline{y}(T)|\mathcal{F}_t\right] - y_d\right), \tag{4.11}$$

where $B^* \in \mathcal{L}(L^2(\mathcal{D}))$ denotes the adjoint operator of $B \in \mathcal{L}(L^2(\mathcal{D}))$.

Proof. Since the operator A is self-adjoint, the semigroup $(e^{At})_{t\geq 0}$ is self-adjoint as well. Using Fubini's theorem and Proposition 3.16, we obtain for every $\tilde{u} \in U$

$$\begin{aligned}\mathbb{E}\int_0^T \left\langle y(T;u,v) - y_d, e^{A(T-t)}B\tilde{u}(t)\right\rangle_{L^2(\mathcal{D})} dt &= \int_0^T \mathbb{E}\left[\mathbb{E}\left[\left\langle y(T;u,v) - y_d, e^{A(T-t)}B\tilde{u}(t)\right\rangle_{L^2(\mathcal{D})} \middle| \mathcal{F}_t\right]\right] dt \\ &= \mathbb{E}\int_0^T \left\langle \mathbb{E}\left[y(T;u,v)|\mathcal{F}_t\right] - y_d, e^{A(T-t)}B\tilde{u}(t)\right\rangle_{L^2(\mathcal{D})} dt \\ &= \mathbb{E}\int_0^T \left\langle B^*e^{A(T-t)}\left(\mathbb{E}\left[y(T;u,v)|\mathcal{F}_t\right] - y_d\right), \tilde{u}(t)\right\rangle_{L^2(\mathcal{D})} dt.\end{aligned}$$

By equation (4.7), we get

$$d_u^F J(u,v)[\tilde{u}] = \mathbb{E}\int_0^T \left\langle B^*e^{A(T-t)}\left(\mathbb{E}\left[y(T;u,v)\middle|\mathcal{F}_t\right] - y_d\right) + \kappa_1 u(t), \tilde{u}(t)\right\rangle_{L^2(\mathcal{D})} dt.$$

Using condition (4.9), the optimal distributed control $\overline{u} \in U$ satisfies equation (4.11) for almost all $t \in [0,T]$ and $\mathbb{P}$-almost surely. □

Theorem 4.6. *Let the cost functional $J\colon U \times V \to \mathbb{R}$ be given by (4.5). Then the optimal boundary control satisfies for almost all $t \in [0,T]$ and $\mathbb{P}$-a.s.*

$$\overline{v}(t) = -\frac{1}{\kappa_2}\mathcal{G}^*(\lambda - A)^{1-\alpha}e^{A(T-t)}\left(\mathbb{E}\left[\overline{y}(T)|\mathcal{F}_t\right] - y_d\right), \tag{4.12}$$

where $\mathcal{G}^* \in \mathcal{L}(L^2(\mathcal{D}); L^2(\partial\mathcal{D}))$ denotes the adjoint operator of $\mathcal{G} = (\lambda - A)^\alpha N \in \mathcal{L}(L^2(\partial\mathcal{D}); L^2(\mathcal{D}))$ with $\alpha \in (\frac{1}{2}, \frac{3}{4})$.

Proof. First, we prove the existence of an approximating sequence $(\tilde{y}_i(T;u,v))_{i\in\mathbb{N}} \subset L^2(\Omega; D(A))$ of the random variable $y(T;u,v) - y_d \in L^2(\Omega; L^2(\mathcal{D}))$ for fixed controls $u \in U$ and $v \in V$. Let z be a $L^2(\mathcal{D})$-valued simple random variable, i.e. there exist functions $f_j \in L^2(\mathcal{D})$ for $j = 1, 2, ..., N$ such that $\mathbb{P}$-a.s.

$$z = \sum_{j=1}^N f_j \mathbb{1}_{\mathcal{A}_j},$$

where $\mathbb{1}_{\mathcal{A}_j}$ denotes the indicator function of $\mathcal{A}_j \in \mathcal{F}$. Since $D(A)$ is dense in $L^2(\mathcal{D})$, there exists a sequence $\left(f_j^i\right)_{i\in\mathbb{N}} \subset D(A)$ for each $j \in \{1, 2, ..., N\}$ such that

$$\lim_{i\to\infty} \left\|f_j - f_j^i\right\|_{L^2(\mathcal{D})} = 0.$$

We set $\mathbb{P}$-a.s. $z_i = \sum\limits_{j=1}^N f_j^i \mathbb{1}_{\mathcal{A}_j}$. Then we obtain

$$\lim_{i\to\infty} \mathbb{E}\left\|z - z_i\right\|_{L^2(\mathcal{D})}^2 = 0.$$

Furthermore, it is well known that every random variable with values in $L^2(\mathcal{D})$ can be approximated by a sequence of $L^2(\mathcal{D})$-valued simple random variables. Therefore, for $y(T;u,v) - y_d \in L^2(\Omega; L^2(\mathcal{D}))$, there exists a sequence $(\tilde{y}_i(T;u,v))_{i\in\mathbb{N}} \subset L^2(\Omega; D(A))$ such that for fixed controls $u \in U$ and $v \in V$

$$\lim_{i\to\infty} \mathbb{E}\|y(T;u,v) - y_d - \tilde{y}_i(T;u,v)\|_{L^2(\mathcal{D})}^2 = 0. \tag{4.13}$$

Recall that the semigroup $(e^{At})_{t\geq 0}$ is self-adjoint. Using Fubini's theorem and Corollary 4.2, we have for every $\tilde{v} \in V$ and each $i \in \mathbb{N}$

$$\begin{aligned}
&\mathbb{E}\int_0^T \left\langle \tilde{y}_i(T;u,v), (\lambda - A)e^{A(T-t)}N\tilde{v}(t)\right\rangle_{L^2(\mathcal{D})} dt \\
&= \int_0^T \mathbb{E}\left[\left\langle \tilde{y}_i(T;u,v), (\lambda - A)^{1-\alpha}e^{A(T-t)}(\lambda - A)^{\alpha}N\tilde{v}(t)\right\rangle_{L^2(\mathcal{D})}\right] dt \\
&= \int_0^T \mathbb{E}\left[\mathbb{E}\left[\left\langle \tilde{y}_i(T;u,v), (\lambda - A)^{1-\alpha}e^{A(T-t)}\mathcal{G}\tilde{v}(t)\right\rangle_{L^2(\mathcal{D})}\Big|\mathcal{F}_t\right]\right] dt \\
&= \mathbb{E}\int_0^T \left\langle \mathbb{E}\left[\tilde{y}_i(T;u,v)|\mathcal{F}_t\right], (\lambda - A)^{1-\alpha}e^{A(T-t)}\mathcal{G}\tilde{v}(t)\right\rangle_{L^2(\mathcal{D})} dt \\
&= \mathbb{E}\int_0^T \left\langle e^{A(T-t)}(\lambda - A)^{1-\alpha}\mathbb{E}\left[\tilde{y}_i(T;u,v)|\mathcal{F}_t\right], \mathcal{G}\tilde{v}(t)\right\rangle_{L^2(\mathcal{D})} dt \\
&= \mathbb{E}\int_0^T \left\langle \mathcal{G}^*(\lambda - A)^{1-\alpha}e^{A(T-t)}\mathbb{E}\left[\tilde{y}_i(T;u,v)|\mathcal{F}_t\right], \tilde{v}(t)\right\rangle_{L^2(\partial\mathcal{D})} dt.
\end{aligned}$$

Next, let the operator $\mathcal{M}(t)\colon L^2(\mathcal{D}) \to L^2(\partial\mathcal{D})$ be defined by

$$\mathcal{M}(t) = \mathcal{G}^*(\lambda - A)^{1-\alpha}e^{A(T-t)}$$

for all $t \in (0,T]$. Since the operator $\mathcal{G}^*\colon L^2(\mathcal{D}) \to L^2(\partial\mathcal{D})$ is linear and bounded and using Corollary 4.2 (iv) and (vi), the operator $\mathcal{M}(t)$ is linear and there exists a constant $C > 0$ such that for all $t \in (0,T]$ and every $g \in L^2(\mathcal{D})$

$$\|\mathcal{M}(t)g\|_{L^2(\partial\mathcal{D})} \leq CM_{1-\alpha}(T-t)^{\alpha-1}\|g\|_{L^2(\mathcal{D})}. \tag{4.14}$$

By inequality (4.14), Fubini's theorem and Proposition 3.16, we obtain for every $u \in U$, every $v \in V$ and each $i \in \mathbb{N}$

$$\begin{aligned}
&\mathbb{E}\int_0^T \|\mathcal{M}(t)\left(\mathbb{E}\left[y(T;u,v)|\mathcal{F}_t\right] - y_d\right) - \mathcal{M}(t)\mathbb{E}\left[\tilde{y}_i(T;u,v)|\mathcal{F}_t\right]\|^2_{L^2(\partial\mathcal{D})} dt \\
&\leq C^2M^2_{1-\alpha}\mathbb{E}\int_0^T (T-t)^{2\alpha-2}\left\|\mathbb{E}\left[y(T;u,v)|\mathcal{F}_t\right] - y_d - \mathbb{E}\left[\tilde{y}_i(T;u,v)|\mathcal{F}_t\right]\right\|^2_{L^2(\mathcal{D})} dt \\
&\leq C^2M^2_{1-\alpha}\int_0^T (T-t)^{2\alpha-2}\mathbb{E}\left[\mathbb{E}\left[\|y(T;u,v) - y_d - \tilde{y}_i(T;u,v)\|^2_{L^2(\mathcal{D})}\Big|\mathcal{F}_t\right]\right] dt \\
&= \frac{C^2M^2_{1-\alpha}T^{2\alpha-1}}{2\alpha-1}\mathbb{E}\left\|y(T;u,v) - y_d - \tilde{y}_i(T;u,v)\right\|^2_{L^2(\mathcal{D})}.
\end{aligned}$$

Due to equation (4.13), we can conclude that for fixed controls $u \in U$ and $v \in V$

$$\lim_{i\to\infty}\mathbb{E}\int_0^T \|\mathcal{M}(t)\left(\mathbb{E}\left[y(T;u,v)|\mathcal{F}_t\right] - y_d\right) - \mathcal{M}(t)\mathbb{E}\left[\tilde{y}_i(T;u,v)|\mathcal{F}_t\right]\|_{L^2(\partial\mathcal{D})}dt = 0.$$

Therefore, we have for fixed controls $u \in U$, fixed $v \in V$ and every $\tilde{v} \in V$

$$\begin{aligned}\mathbb{E}\int_0^T \left\langle y(T;u,v) - y_d, (\lambda - A)e^{A(T-t)}N\tilde{v}(t)\right\rangle_{L^2(\mathcal{D})} dt &= \lim_{i\to\infty} \mathbb{E}\int_0^T \left\langle \tilde{y}_i(T;u,v), (\lambda - A)e^{A(T-t)}N\tilde{v}(t)\right\rangle_{L^2(\mathcal{D})} dt \\ &= \lim_{i\to\infty} \mathbb{E}\int_0^T \left\langle \mathcal{M}(t)\mathbb{E}\left[\tilde{y}_i(T;u,v)|\mathcal{F}_t\right], \tilde{v}(t)\right\rangle_{L^2(\partial\mathcal{D})} dt \\ &= \mathbb{E}\int_0^T \left\langle \mathcal{M}(t)\left(\mathbb{E}\left[y(T;u,v)|\mathcal{F}_t\right] - y_d\right), \tilde{v}(t)\right\rangle_{L^2(\partial\mathcal{D})} dt.\end{aligned}$$

Using equation (4.8), we find for every $\tilde{v} \in V$

$$d_v J(u,v)[\tilde{v}] = \mathbb{E}\int_0^T \left\langle \mathcal{M}(t)\left(\mathbb{E}\left[y(T;u,v)|\mathcal{F}_t\right] - y_d\right) + \kappa_2 v(t), \tilde{v}(t)\right\rangle_{L^2(\partial\mathcal{D})} dt.$$

Applying condition (4.10), we can infer that the optimal boundary control satisfies for almost all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\overline{v}(t) = -\frac{1}{\kappa_2}\mathcal{M}(t)\left(\mathbb{E}[\overline{y}(T)|\mathcal{F}_t] - y_d\right).$$

This implies equation (4.12) and proves the theorem. □

Due to the previous theorem, we will always assume that $\alpha \in (\frac{1}{2}, \frac{3}{4})$.

Remark 4.7. *Note that the previous results can be easily obtained if system (4.3) is driven by a square integrable Lévy martingales as introduced in Section 3.3.*

4.4. Design of a Feedback Law

Based on Theorem 4.5 and Theorem 4.6, the optimal controls can be determined by calculating $\mathbb{E}\left[\overline{y}(T)|\mathcal{F}_t\right]$. Since this leads to serious problems in applications, we avoid the calculation of the conditional expectation by using the martingale representation theorem according to Theorem 3.86. Here, we assume that the Q-Wiener processes $(W(t))_{t\geq 0}$ and $(W_b(t))_{t\geq 0}$ are independent. First, we apply Proposition 3.42 to obtain series expansions of the Q-Wiener processes $(W(t))_{t\geq 0}$ and $(W_b(t))_{t\geq 0}$. Let $(u_k)_{k\in\mathbb{N}}$ be an orthonormal basis in $L^2(\mathcal{D})$. For arbitrary $t \geq 0$, we have the following expansions:

$$W(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k} w_k(t) u_k,$$

where $(w_k(t))_{t\geq 0}$, $k \in \mathbb{N}$, are mutually independent real-valued Brownian motions and the sequence of nonnegative real numbers $(\lambda_k)_{k\in\mathbb{N}}$ satisfies $Qu_k = \lambda_k u_k$ for each $k \in \mathbb{N}$. Similarly, let $(u_k^b)_{k\in\mathbb{N}}$ be an orthonormal basis in $L^2(\partial\mathcal{D})$. For arbitrary $t \geq 0$, we have the following expansions:

$$W_b(t) = \sum_{k=1}^{\infty} \sqrt{\lambda_k^b} w_k^b(t) u_k^b,$$

where $(w_k^b(t))_{t\geq 0}$, $k \in \mathbb{N}$, are mutually independent real-valued Brownian motions and the sequence of nonnegative real numbers $(\lambda_k^b)_{k\in\mathbb{N}}$ satisfies $Q_b u_k^b = \lambda_k^b u_k^b$ for each $k \in \mathbb{N}$. On the probability space

$(\Omega, \mathcal{F}, \mathbb{P})$, we assume that the filtration $(\mathcal{F}_t)_{t\geq 0}$ is given by $\mathcal{F}_t = \sigma\left\{\mathcal{F}_t^1 \cup \mathcal{F}_t^2\right\}$ for all $t \geq 0$, where $\mathcal{F}_t^1 = \sigma\left\{\bigcup_{k=1}^{\infty} \sigma\{w_k(s) : 0 \leq s \leq t\}\right\}$ and $\mathcal{F}_t^2 = \sigma\left\{\bigcup_{k=1}^{\infty} \sigma\{w_k^b(s) : 0 \leq s \leq t\}\right\}$ for all $t \geq 0$. Moreover, we set $\mathcal{F} = \mathcal{F}_T$. Obviously, the process $(\mathbb{E}\left[\overline{y}(T)|\mathcal{F}_t\right])_{t\in[0,T]}$ is a continuous square integrable $\mathcal{F}_t$-martingale. As a consequence of Theorem 3.86, there exist predictable processes $(\Phi(t))_{t\in[0,T]}$ and $(\Phi_b(t))_{t\in[0,T]}$ with values in $\mathcal{L}_{(HS)}\left(Q^{1/2}(L^2(\mathcal{D})); L^2(\mathcal{D})\right)$ and $\mathcal{L}_{(HS)}\left(Q_b^{1/2}(L^2(\partial\mathcal{D})); L^2(\mathcal{D})\right)$, respectively, such that

$$\mathbb{E}\int_0^T \|\Phi(t)\|^2_{\mathcal{L}_{(HS)}\left(Q^{1/2}(L^2(\mathcal{D}));L^2(\mathcal{D})\right)} dt < \infty,$$

$$\mathbb{E}\int_0^T \|\Phi_b(t)\|^2_{\mathcal{L}_{(HS)}\left(Q_b^{1/2}(L^2(\partial\mathcal{D}));L^2(\mathcal{D})\right)} dt < \infty$$

and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\mathbb{E}\left[\overline{y}(T)|\mathcal{F}_t\right] = \mathbb{E}[\overline{y}(T)] + \int_0^t \Phi(s)\, dW(s) + \int_0^t \Phi_b(s)\, dW_b(s). \tag{4.15}$$

Remark 4.8. *From the proof of Theorem 3.86, we can easily obtain that the representation (4.15) holds. Especially, it is necessary to assume that the Q-Wiener processes $(W(t))_{t\geq 0}$ and $(W_b(t))_{t\geq 0}$ are independent.*

Let the process $(q(t))_{t\in[0,T]}$ satisfy for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$q(t) = e^{A(T-t)}(\mathbb{E}\left[\overline{y}(T)|\mathcal{F}_t\right] - y_d). \tag{4.16}$$

Next, we introduce the adjoint state $(p(t))_{t\in[0,T]}$ satisfying for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$p(t) = e^{A(T-t)}(\overline{y}(T) - y_d) - \int_t^T e^{A(s-t)}\Phi^{(T)}(s)\, dW(s) - \int_t^T e^{A(s-t)}\Phi_b^{(T)}(s)\, dW_b(s),$$

where $\Phi^{(T)}(s) = e^{A(T-s)}\Phi(s)$ and $\Phi_b^{(T)}(s) = e^{A(T-s)}\Phi_b(s)$. Then we obtain for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$q(t) = \mathbb{E}\left[p(t)|\mathcal{F}_t\right] \tag{4.17}$$

and by equation (4.11) and equation (4.12), the optimal controls $\overline{u} \in U$ and $\overline{v} \in V$ satisfy for almost all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\overline{u}(t) = -\frac{1}{\kappa_1}B^*q(t), \tag{4.18}$$

$$\overline{v}(t) = -\frac{1}{\kappa_2}\mathcal{G}^*(\lambda - A)^{1-\alpha}q(t). \tag{4.19}$$

In the remaining part of this section, we reformulate the process $(q(t))_{t\in[0,T]}$ to obtain a feedback law of the optimal controls. Therefor, we introduce the function $\mathcal{P}\colon [0,T] \to \mathcal{L}(L^2(\mathcal{D}))$, which fulfills the following Riccati equation:

$$\begin{cases} \dfrac{d}{dt}\mathcal{P}(t) = A\mathcal{P}(t) + \mathcal{P}(t)A - \dfrac{1}{\kappa_1}\mathcal{P}(t)BB^*\mathcal{P}(t) - \dfrac{1}{\kappa_2}\mathcal{H}^*(t)\mathcal{G}\mathcal{G}^*\mathcal{H}(t), \\ \mathcal{P}(T) = I, \end{cases} \tag{4.20}$$

where $\mathcal{H}(t) = (\lambda - A)^{1-\alpha}\mathcal{P}(t)$ and I is the identity operator on $L^2(\mathcal{D})$.

Definition 4.9. *We call $\mathcal{P}\colon [0,T] \to \mathcal{L}(L^2(\mathcal{D}))$ a* ***mild solution of system (4.20)*** *if for all $t \in [0,T]$ and every $h \in L^2(\mathcal{D})$*

$$\mathcal{P}(t)h = e^{A(T-t)}e^{A(T-t)}h - \frac{1}{\kappa_1}\int_t^T e^{A(s-t)}\mathcal{P}(s)BB^*\mathcal{P}(s)e^{A(s-t)}h\,ds$$
$$- \frac{1}{\kappa_2}\int_t^T e^{A(s-t)}\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*\mathcal{H}(s)e^{A(s-t)}h\,ds. \tag{4.21}$$

Remark 4.10. *In [9, Part IV], the existence and uniqueness of mild solutions to the Riccati equations*

$$\begin{cases} \dfrac{d}{dt}\mathcal{P}(t) = A\mathcal{P}(t) + \mathcal{P}(t)A - \mathcal{P}(t)BB^*\mathcal{P}(t), \\ \mathcal{P}(T) = I \end{cases}$$

and

$$\begin{cases} \dfrac{d}{dt}\mathcal{P}(t) = A\mathcal{P}(t) + \mathcal{P}(t)A - \mathcal{H}^*(t)\mathcal{G}\mathcal{G}^*\mathcal{H}(t), \\ \mathcal{P}(T) = I \end{cases}$$

are proved. Since equation (4.20) is a generalization of these special cases, an existence and uniqueness result can be easily obtained.

In the following remark, we state some important properties of the function $\mathcal{P}\colon [0,T] \to \mathcal{L}(L^2(\mathcal{D}))$.

Remark 4.11. *Recall that $\alpha \in (\frac{1}{2}, \frac{3}{4})$. According to [9, Part IV], we have*

- *$\mathcal{P}(t)h \in D((\lambda - A)^{1-\alpha})$ for every $h \in L^2(\mathcal{D})$ and all $t \in [0,T)$;*
- *$t \mapsto (\lambda - A)^{1-\alpha}\mathcal{P}(t)$ is a continuous function from $[0,T)$ into $\mathcal{L}(L^2(\mathcal{D}))$;*
- *$\mathcal{P}(t) \in \mathcal{L}(L^2(\mathcal{D}))$ is self-adjoint for all $t \in [0,T]$.*

Lemma 4.12. *Let $\mathcal{P}\colon [0,T] \to \mathcal{L}(L^2(\mathcal{D}))$ be the mild solution of system (4.20). If $z \in D((\lambda - A)^{1-\alpha})$, then we have for all $t \in [0,T]$*

$$\mathcal{P}(t)(\lambda - A)^{1-\alpha}z = (\lambda - A)^{1-\alpha}\mathcal{P}(t)z.$$

Proof. The claim follows immediately from Corollary 4.2 (vii) and the fact that $\mathcal{P}(t)$ is self-adjoint for all $t \in [0,T]$. □

Next, we introduce the function $a\colon [0,T] \to D((\lambda - A)^{1-\alpha})$ satisfying the following deterministic backward integral equation for $t \in [0,T]$:

$$a(t) = \int_t^T e^{A(s-t)}\left(-\frac{1}{\kappa_1}\mathcal{P}(s)BB^* - \frac{1}{\kappa_2}\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*(\lambda - A)^{1-\alpha}\right)a(s)\,ds - e^{A(T-t)}y_d. \tag{4.22}$$

We have the following existence and uniqueness result.

Theorem 4.13. *There exists a unique solution $a\colon [0,T] \to (\lambda - A)^{1-\alpha}$ of equation (4.22) such that*

$$\int_0^T \|a(t)\|^2_{D((\lambda-A)^{1-\alpha})}dt < \infty.$$

Proof. For all $t \in [0,T]$, let us introduce the operator $\mathcal{M}(t)\colon D\left((\lambda - A)^{1-\alpha}\right) \to L^2(\mathcal{D})$ defined by

$$\mathcal{M}(t) = -\frac{1}{\kappa_1}\mathcal{P}(t)BB^* - \frac{1}{\kappa_2}\mathcal{H}^*(t)\mathcal{G}\mathcal{G}^*(\lambda - A)^{1-\alpha}.$$

Then clearly $\mathcal{M}(t)$ is linear and closed for all $t \in [0,T]$. Recall that the operators $\mathcal{P}(t), B, \mathcal{H}(t)$ and $\mathcal{G}$ are linear and bounded on $L^2(\mathcal{D})$ for all $t \in [0,T)$. By Corollary 4.2 (iii), there exists a constant $c > 0$ such that for all $t \in [0,T)$ and every $y \in D((\lambda - A)^{1-\alpha})$

$$\|\mathcal{M}(t)y\|_{L^2(\mathcal{D})} \le \frac{1}{\kappa_1}\|\mathcal{P}(t)BB^*y\|_{L^2(\mathcal{D})} + \frac{1}{\kappa_2}\|\mathcal{H}^*(t)\mathcal{G}\mathcal{G}^*(\lambda - A)^{1-\alpha}y\|_{L^2(\mathcal{D})} \le c\|(\lambda - A)^{1-\alpha}y\|_{L^2(\mathcal{D})}. \quad (4.23)$$

Next, we define for $t \in [0,T]$

$$\mathcal{J}(\tilde{a})(t) = \int_t^T e^{A(s-t)}\mathcal{M}(s)\tilde{a}(s)\,ds - e^{A(T-t)}y_d.$$

Let $T_1 \in [0,T)$. We have that $\mathcal{J}$ maps $L^2([T_1,T];D((\lambda - A)^{1-\alpha}))$ into itself, which follows immediately from Proposition B.9, Corollary 4.2, inequality (4.23) and Young's inequality for convolutions. Indeed, we obtain

$$\begin{aligned}
&\int_{T_1}^T \|\mathcal{J}(\tilde{a})(t)\|^2_{D((\lambda-A)^{1-\alpha})}dt \\
&\le 2\int_{T_1}^T \left(\int_t^T \left\|(\lambda - A)^{1-\alpha}e^{A(s-t)}\mathcal{M}(s)\tilde{a}(s)\right\|_{L^2(\mathcal{D})} ds\right)^2 dt + 2\int_{T_1}^T \|(\lambda - A)^{1-\alpha}e^{A(T-t)}y_d\|^2_{L^2(\mathcal{D})}dt \\
&\le 2M_{1-\alpha}^2 \left[\int_{T_1}^T \left(\int_t^T (s-t)^{\alpha-1}\|\mathcal{M}(s)\tilde{a}(s)\|_{L^2(\mathcal{D})}\,ds\right)^2 dt + \int_{T_1}^T (T-t)^{2\alpha-2}dt\,\|y_d\|^2_{L^2(\mathcal{D})}\right] \\
&\le 2M_{1-\alpha}^2 \left[\frac{c^2(T-T_1)^\alpha}{\alpha}\int_{T_1}^T \|\tilde{a}(t)\|^2_{D((\lambda-A)^{1-\alpha})}dt + \frac{(T-T_1)^{2\alpha-1}}{2\alpha-1}\|y_d\|^2_{L^2(\mathcal{D})}\right].
\end{aligned}$$

Next, we show that $\mathcal{J}$ is a contraction on $L^2([T_1,T];D((\lambda - A)^{1-\alpha}))$. Let $\tilde{a}_1, \tilde{a}_2 \in L^2([T_1,T];D((\lambda - A)^{1-\alpha}))$. Similar as above, we get

$$\begin{aligned}
\int_{T_1}^T \|\mathcal{J}(\tilde{a}_1)(t) - \mathcal{J}(\tilde{a}_2)(t)\|^2_{D((\lambda-A)^{1-\alpha})}dt &\le \int_{T_1}^T \left(\int_t^T \left\|(\lambda - A)^{1-\alpha}e^{A(s-t)}\mathcal{M}(s)[\tilde{a}_1(s) - \tilde{a}_1(s)]\right\|_{L^2(\mathcal{D})} ds\right)^2 dt \\
&\le \frac{M_{1-\alpha}^2 c^2(T-T_1)^\alpha}{\alpha}\int_{T_1}^T \|\tilde{a}_1(t) - \tilde{a}_2(t)\|^2_{D((\lambda-A)^{1-\alpha})}dt.
\end{aligned}$$

We chose $T_1 \in [0,T)$ such that $\frac{M_{1-\alpha}^2 c^2(T-T_1)^\alpha}{\alpha} < 1$. Applying the Banach fixed point theorem, we get a unique element $a \in L^2([T_1,T];D((\lambda - A)^{1-\alpha}))$ such that $\mathcal{J}(a)(t) = a(t)$ for $t \in [T_1,T]$. Next, we consider for $t \in [0,T_1]$

$$\mathcal{J}(\tilde{a})(t) = e^{A(T_1-t)}a(T_1) + \int_t^{T_1} e^{A(s-t)}\mathcal{M}(s)\tilde{a}(s)\,ds.$$

Again, we find $T_2 \in [0, T_1]$ such that there exists a unique fixed point of $\mathcal{J}$ on $L^2([T_2, T_1]; D((\lambda - A)^{1-\alpha}))$. By continuing the method, we get the existence and uniqueness of a function $a\colon [0,T] \to (\lambda - A)^{1-\alpha}$ satisfying $\mathcal{J}(a)(t) = a(t)$ for $t \in [0,T]$. □

This enables us to prove the following representation theorem, where we closely relate to the proof of [21, Theorem 7.8].

Theorem 4.14. *Let the process $(q(t))_{t\in[0,T]}$ be given by (4.17). Then we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.*

$$q(t) = \mathcal{P}(t)\overline{y}(t) + a(t), \tag{4.24}$$

where $\mathcal{P}\colon [0,T] \to \mathcal{L}(L^2(\mathcal{D}))$ is the mild solution of system (4.20) and $a\colon [0,T] \to D((\lambda - A)^{1-\alpha})$ is the unique solution of equation (4.22).

Proof. Let $t \in [0,T]$ be arbitrary. Substituting equations (4.18) and (4.19) in equation (4.4), we find for all $s \in [t,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned}\overline{y}(s) = e^{A(s-t)}\overline{y}(t) - \frac{1}{\kappa_1}\int_t^s e^{A(s-r)}BB^*q(r)\,dr - \frac{1}{\kappa_2}\int_t^s (\lambda - A)^{1-\alpha}e^{A(s-r)}\mathcal{G}\mathcal{G}^*(\lambda - A)^{1-\alpha}q(r)\,dr\\ + \int_t^s e^{A(s-r)}G(r)\,dW(r) + \int_t^s (\lambda - A)e^{A(s-r)}N\,dW_b(r).\end{aligned}$$

Next, we define for all $r \in [t,T]$ and $\mathbb{P}$-a.s.

$$\tilde{q}(r) = \mathbb{E}[p(r)|\mathcal{F}_t]. \tag{4.25}$$

Then by equation (4.17), we get $\mathbb{P}$-a.s. $q(t) = \tilde{q}(t)$ and

$$\mathbb{E}[q(r)|\mathcal{F}_t] = \mathbb{E}\left[\mathbb{E}[p(r)|\mathcal{F}_r]\middle|\mathcal{F}_t\right] = \mathbb{E}[p(r)|\mathcal{F}_t] = \tilde{q}(r)$$

resulting from Proposition 3.16. Thus, we have for all $s \in [t,T]$ and $\mathbb{P}$-a.s.

$$\mathbb{E}[\overline{y}(s)|\mathcal{F}_t] = e^{A(s-t)}\overline{y}(t) - \frac{1}{\kappa_1}\int_t^s e^{A(s-r)}BB^*q(r)\,dr - \frac{1}{\kappa_2}\int_t^s (\lambda - A)^{1-\alpha}e^{A(s-r)}\mathcal{G}\mathcal{G}^*(\lambda - A)^{1-\alpha}q(r)\,dr. \tag{4.26}$$

Using equation (4.16) and Corollary 4.2, we obtain $\mathbb{P}$-a.s.

$$\begin{aligned}q(t) = e^{A(T-t)}e^{A(T-t)}\overline{y}(t) - \frac{1}{\kappa_1}\int_t^T e^{A(T-t)}e^{A(T-s)}BB^*\tilde{q}(s)\,ds\\ - \frac{1}{\kappa_2}\int_t^T (\lambda - A)^{1-\alpha}e^{A(T-t)}e^{A(T-s)}\mathcal{G}\mathcal{G}^*(\lambda - A)^{1-\alpha}\tilde{q}(s)\,ds - e^{A(T-t)}y_d.\end{aligned}$$

By equation (4.21) with $h = \overline{y}(t)$, we find $\mathbb{P}$-a.s.

$$\begin{aligned}q(t) &= \mathcal{P}(t)\overline{y}(t) - e^{A(T-t)}y_d + \frac{1}{\kappa_1}\int_t^T \left[e^{A(s-t)}\mathcal{P}(s)BB^*\mathcal{P}(s)e^{A(s-t)}\overline{y}(t) - e^{A(T-t)}e^{A(T-s)}BB^*\tilde{q}(s)\right]ds\\ &\quad + \frac{1}{\kappa_2}\int_t^T e^{A(s-t)}\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*\mathcal{H}(s)e^{A(s-t)}\overline{y}(t)\,ds - \frac{1}{\kappa_2}\int_t^T (\lambda - A)^{1-\alpha}e^{A(T-t)}e^{A(T-s)}\mathcal{G}\mathcal{G}^*(\lambda - A)^{1-\alpha}\tilde{q}(s)\,ds\\ &= \mathcal{P}(t)\overline{y}(t) - e^{A(T-t)}y_d + \mathcal{I}_1(t) + \mathcal{I}_2(t),\end{aligned} \tag{4.27}$$

where

$$\mathcal{I}_1(t) = \frac{1}{\kappa_1}\int_t^T \Big[e^{A(s-t)}\mathcal{P}(s)BB^*\mathcal{P}(s)e^{A(s-t)}\overline{y}(t) - e^{A(T-t)}e^{A(T-s)}BB^*\tilde{q}(s)\Big]\,ds,$$

$$\mathcal{I}_2(t) = \frac{1}{\kappa_2}\int_t^T e^{A(s-t)}\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*\mathcal{H}(s)e^{A(s-t)}\overline{y}(t)\,ds$$
$$- \frac{1}{\kappa_2}\int_t^T (\lambda - A)^{1-\alpha}e^{A(T-t)}e^{A(T-s)}\mathcal{G}\mathcal{G}^*(\lambda - A)^{1-\alpha}\tilde{q}(s)\,ds.$$

Using again equation (4.21) with $h = BB^*\tilde{q}(s)$, we get $\mathbb{P}$-a.s.

$$\mathcal{I}_1(t) = \frac{1}{\kappa_1}\int_t^T e^{A(s-t)}\mathcal{P}(s)BB^*\Big[\mathcal{P}(s)e^{A(s-t)}\overline{y}(t) - \tilde{q}(s)\Big]\,ds$$
$$- \frac{1}{\kappa_1^2}\int_t^T\int_s^T e^{A(r-t)}\mathcal{P}(r)BB^*\mathcal{P}(r)e^{A(r-s)}BB^*\tilde{q}(s)\,dr\,ds$$
$$- \frac{1}{\kappa_1\kappa_2}\int_t^T\int_s^T e^{A(r-t)}\mathcal{H}^*(r)\mathcal{G}\mathcal{G}^*\mathcal{H}(r)e^{A(r-s)}BB^*\tilde{q}(s)\,dr\,ds.$$

By Fubini's theorem, we have $\mathbb{P}$-a.s.

$$\mathcal{I}_1(t) = \frac{1}{\kappa_1}\int_t^T e^{A(s-t)}\mathcal{P}(s)BB^*\Big[\mathcal{P}(s)e^{A(s-t)}\overline{y}(t) - \tilde{q}(s)\Big]\,ds$$
$$- \frac{1}{\kappa_1^2}\int_t^T e^{A(r-t)}\mathcal{P}(r)BB^*\mathcal{P}(r)\int_t^r e^{A(r-s)}BB^*\tilde{q}(s)\,ds\,dr$$
$$- \frac{1}{\kappa_1\kappa_2}\int_t^T e^{A(r-t)}\mathcal{H}^*(r)\mathcal{G}\mathcal{G}^*\mathcal{H}(r)\int_t^r e^{A(r-s)}BB^*\tilde{q}(s)\,ds\,dr.$$

Through interchanging the integration variables in the last two integrals, we find $\mathbb{P}$-a.s.

$$\mathcal{I}_1(t) = \frac{1}{\kappa_1}\int_t^T e^{A(s-t)}\mathcal{P}(s)BB^*\left[\mathcal{P}(s)e^{A(s-t)}\overline{y}(t) - \tilde{q}(s) - \frac{1}{\kappa_1}\mathcal{P}(s)\int_t^s e^{A(s-r)}BB^*\tilde{q}(r)\,dr\right]ds$$
$$- \frac{1}{\kappa_1\kappa_2}\int_t^T e^{A(s-t)}\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*\mathcal{H}(s)\int_t^s e^{A(s-r)}BB^*\tilde{q}(r)\,dr\,ds. \tag{4.28}$$

Next, we reformulate $\mathcal{I}_2(t)$. Corollary 4.2, Proposition B.9 and equation (4.21) with $h = \tilde{z}$ for an arbitrary

$\tilde{z} \in D((\lambda - A)^{1-\alpha})$ yields for all $s \in [t, T]$

$$(\lambda - A)^{1-\alpha} e^{A(T-s)} e^{A(T-s)} \tilde{z} = (\lambda - A)^{1-\alpha} \mathcal{P}(s) \tilde{z} + \frac{1}{\kappa_1} \int_s^T (\lambda - A)^{1-\alpha} e^{A(r-s)} \mathcal{P}(r) B B^* \mathcal{P}(r) e^{A(r-s)} \tilde{z} \, dr$$
$$+ \frac{1}{\kappa_2} \int_s^T (\lambda - A)^{1-\alpha} e^{A(r-s)} \mathcal{H}^*(r) \mathcal{G}\mathcal{G}^* \mathcal{H}(r) e^{A(r-s)} \tilde{z} \, dr.$$

Similarly, by using additionally Lemma 4.12, we obtain for all $s \in [t, T]$

$$e^{A(T-s)} e^{A(T-s)} (\lambda - A)^{1-\alpha} \tilde{z} = (\lambda - A)^{1-\alpha} \mathcal{P}(s) \tilde{z} + \frac{1}{\kappa_1} \int_s^T e^{A(r-s)} \mathcal{P}(r) B B^* \mathcal{P}(r) (\lambda - A)^{1-\alpha} e^{A(r-s)} \tilde{z} \, dr$$
$$+ \frac{1}{\kappa_2} \int_s^T e^{A(r-s)} \mathcal{H}^*(r) \mathcal{G}\mathcal{G}^* \mathcal{H}(r) (\lambda - A)^{1-\alpha} e^{A(r-s)} \tilde{z} \, dr.$$

By Corollary 4.2, we get

$$(\lambda - A)^{1-\alpha} e^{A(T-s)} e^{A(T-s)} \tilde{z} = e^{A(T-s)} e^{A(T-s)} (\lambda - A)^{1-\alpha} \tilde{z}.$$

Hence, we can conclude for all $s \in [t, T]$

$$\begin{aligned} &\frac{1}{\kappa_1} \int_s^T (\lambda - A)^{1-\alpha} e^{A(r-s)} \mathcal{P}(r) B B^* \mathcal{P}(r) e^{A(r-s)} \tilde{z} \, dr + \frac{1}{\kappa_2} \int_s^T (\lambda - A)^{1-\alpha} e^{A(r-s)} \mathcal{H}^*(r) \mathcal{G}\mathcal{G}^* \mathcal{H}(r) e^{A(r-s)} \tilde{z} \, dr \\ &= \frac{1}{\kappa_1} \int_s^T e^{A(r-s)} \mathcal{P}(r) B B^* \mathcal{P}(r) (\lambda - A)^{1-\alpha} e^{A(r-s)} \tilde{z} \, dr \\ &\quad + \frac{1}{\kappa_2} \int_s^T e^{A(r-s)} \mathcal{H}^*(r) \mathcal{G}\mathcal{G}^* \mathcal{H}(r) (\lambda - A)^{1-\alpha} e^{A(r-s)} \tilde{z} \, dr. \end{aligned} \tag{4.29}$$

Due to the fact that $D((\lambda - A)^{1-\alpha})$ is dense in $L^2(\mathcal{D})$, the previous equation holds for every $\tilde{z} \in L^2(\mathcal{D})$. Applying equation (4.21) with $h = \mathcal{G}\mathcal{G}^* (\lambda - A)^{1-\alpha} \tilde{q}(s)$, we get $\mathbb{P}$-a.s.

$$\begin{aligned} \mathcal{I}_2(t) = &\frac{1}{\kappa_2} \int_t^T e^{A(s-t)} \left[\mathcal{H}^*(s) \mathcal{G}\mathcal{G}^* \mathcal{H}(s) e^{A(s-t)} \overline{y}(t) - \mathcal{H}(s) \mathcal{G}\mathcal{G}^* (\lambda - A)^{1-\alpha} \tilde{q}(s) \right] ds \\ &- \frac{1}{\kappa_1 \kappa_2} \int_t^T \int_s^T (\lambda - A)^{1-\alpha} e^{A(r-t)} \mathcal{P}(r) B B^* \mathcal{P}(r) e^{A(r-s)} \mathcal{G}\mathcal{G}^* (\lambda - A)^{1-\alpha} \tilde{q}(s) \, dr \, ds \\ &- \frac{1}{\kappa_2^2} \int_t^T \int_s^T (\lambda - A)^{1-\alpha} e^{A(r-t)} \mathcal{H}^*(r) \mathcal{G}\mathcal{G}^* \mathcal{H}(r) e^{A(r-s)} \mathcal{G}\mathcal{G}^* (\lambda - A)^{1-\alpha} \tilde{q}(s) \, dr \, ds, \end{aligned}$$

where we also used Corollary 4.2. By equation (4.29) and Fubini's theorem, we have $\mathbb{P}$-a.s.

$$\mathcal{I}_2(t) = \frac{1}{\kappa_2}\int_t^T e^{A(s-t)}\left[\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*\mathcal{H}(s)e^{A(s-t)}\overline{y}(t) - \mathcal{H}(s)\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(s)\right] ds$$
$$- \frac{1}{\kappa_1\kappa_2}\int_t^T e^{A(r-t)}\mathcal{P}(r)BB^*\mathcal{P}(r)\int_t^r (\lambda-A)^{1-\alpha}e^{A(r-s)}\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(s)\,ds\,dr$$
$$- \frac{1}{\kappa_2^2}\int_t^T e^{A(r-t)}\mathcal{H}^*(r)\mathcal{G}\mathcal{G}^*\mathcal{H}(r)\int_t^r (\lambda-A)^{1-\alpha}e^{A(r-s)}\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(s)\,ds\,dr.$$

Through interchanging the integration variables in the last two integrals, we find $\mathbb{P}$-a.s.

$$\mathcal{I}_2(t) = \frac{1}{\kappa_2}\int_t^T e^{A(s-t)}\left[\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*\mathcal{H}(s)e^{A(s-t)}\overline{y}(t) - \mathcal{H}(s)\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(s)\right] ds$$
$$- \frac{1}{\kappa_1\kappa_2}\int_t^T e^{A(s-t)}\mathcal{P}(s)BB^*\mathcal{P}(s)\int_t^s (\lambda-A)^{1-\alpha}e^{A(s-r)}\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(r)\,dr\,ds$$
$$- \frac{1}{\kappa_2^2}\int_t^T e^{A(s-t)}\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*\mathcal{H}(s)\int_t^s (\lambda-A)^{1-\alpha}e^{A(s-r)}\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(r)\,dr\,ds. \tag{4.30}$$

Using equations (4.28) and (4.30), we obtain $\mathbb{P}$-a.s.

$$\mathcal{I}_1(t)+\mathcal{I}_2(t) = \frac{1}{\kappa_1}\int_t^T e^{A(s-t)}\mathcal{P}(s)BB^*\Bigg[\mathcal{P}(s)e^{A(s-t)}\overline{y}(t) - \tilde{q}(s) - \frac{1}{\kappa_1}\mathcal{P}(s)\int_t^s e^{A(s-r)}BB^*\tilde{q}(r)\,dr$$
$$- \frac{1}{\kappa_2}\mathcal{P}(s)\int_t^s (\lambda-A)^{1-\alpha}e^{A(s-r)}\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(r)\,dr\Bigg] ds$$
$$+ \frac{1}{\kappa_2}\int_t^T e^{A(s-t)}\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\mathcal{P}(s)\Bigg[e^{A(s-t)}\overline{y}(t) - \frac{1}{\kappa_1}\int_t^s e^{A(s-r)}BB^*\tilde{q}(r)\,dr$$
$$- \frac{1}{\kappa_2}\int_t^s (\lambda-A)^{1-\alpha}e^{A(s-r)}\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(r)\,dr\Bigg] ds$$
$$- \frac{1}{\kappa_2}\int_t^T e^{A(s-t)}\mathcal{H}(s)\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(s)\,ds.$$

By equation (4.26), we get $\mathbb{P}$-a.s.

$$\begin{aligned}\mathcal{I}_1(t)+\mathcal{I}_2(t) &= \frac{1}{\kappa_1}\int_t^T e^{A(s-t)}\mathcal{P}(s)BB^*\left[\mathcal{P}(s)\mathbb{E}[\overline{y}(s)|\mathcal{F}_t]-\tilde{q}(s)\right]ds \\ &+\frac{1}{\kappa_2}\int_t^T e^{A(s-t)}\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\left[\mathcal{P}(s)\mathbb{E}[\overline{y}(s)|\mathcal{F}_t]-\tilde{q}(s)\right]ds \\ &+\frac{1}{\kappa_2}\int_t^T e^{A(s-t)}\left(\mathcal{H}^*(s)-\mathcal{H}(s)\right)\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(s)\,ds.\end{aligned}$$

Due to Corollary 4.2 and Lemma 4.12, we have for every $\tilde{z}\in L^2(\mathcal{D})$ and $\mathbb{P}$-a.s.

$$\begin{aligned}&\left\langle \int_t^T e^{A(s-t)}\left(\mathcal{H}^*(s)-\mathcal{H}(s)\right)\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(s)ds,\tilde{z}\right\rangle_{L^2(\mathcal{D})} \\ &=\int_t^T \left\langle \mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(s),(\lambda-A)^{1-\alpha}\mathcal{P}(s)e^{A(s-t)}\tilde{z}\right\rangle_{L^2(\mathcal{D})}ds \\ &\quad-\int_t^T \left\langle \mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(s),\mathcal{P}(s)(\lambda-A)^{1-\alpha}e^{A(s-t)}\tilde{z}\right\rangle_{L^2(\mathcal{D})}ds=0.\end{aligned}$$

Hence, we can conclude that $\mathbb{P}$-a.s.

$$\int_t^T e^{A(s-t)}\left(\mathcal{H}^*(s)-\mathcal{H}(s)\right)\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\tilde{q}(s)\,ds=0.$$

Therefore, we have $\mathbb{P}$-a.s.

$$\begin{aligned}\mathcal{I}_1(t)+\mathcal{I}_2(t) &= \frac{1}{\kappa_1}\int_t^T e^{A(s-t)}\mathcal{P}(s)BB^*\left[\mathcal{P}(s)\mathbb{E}[\overline{y}(s)|\mathcal{F}_t]-\tilde{q}(s)\right]ds \\ &+\frac{1}{\kappa_2}\int_t^T e^{A(s-t)}\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\left[\mathcal{P}(s)\mathbb{E}[\overline{y}(s)|\mathcal{F}_t]-\tilde{q}(s)\right]ds. \qquad (4.31)\end{aligned}$$

Next, we set $a(s)=\tilde{q}(s)-\mathcal{P}(s)\mathbb{E}[\overline{y}(s)|\mathcal{F}_t]$ for $s\in[t,T]$ and $\mathbb{P}$-almost surely. Then for $s=t$, we get $a(t)=q(t)-\mathcal{P}(t)\overline{y}(t)$ resulting from equation (4.17) and equation (4.25). Therefore, we obtain equation (4.24). Moreover, we get that $a(t)$ satisfies the following deterministic backward integral equation:

$$a(t)=\int_t^T e^{A(s-t)}\left(-\frac{1}{\kappa_1}\mathcal{P}(s)BB^*-\frac{1}{\kappa_2}\mathcal{H}^*(s)\mathcal{G}\mathcal{G}^*(\lambda-A)^{1-\alpha}\right)a(s)\,ds-e^{A(T-t)}y_d$$

as a consequence of equation (4.27) and equation (4.31). □

Remark 4.15. *As a consequence of equation (4.18), equation (4.19) and the previous theorem, the optimal controls* $\overline{u} \in U$ *and* $\overline{v} \in V$ *satisfy the following feedback laws for almost all* $t \in [0,T]$ *and* $\mathbb{P}$*-almost surely:*

$$\begin{aligned}\overline{u}(t) &= -\frac{1}{\kappa_1} B^*[\mathcal{P}(t)\overline{y}(t) + a(t)],\\ \overline{v}(t) &= -\frac{1}{\kappa_2} \mathcal{G}^*(\lambda - A)^{1-\alpha}[\mathcal{P}(t)\overline{y}(t) + a(t)],\end{aligned}$$

where the function $\mathcal{P}\colon [0,T] \to \mathcal{L}(L^2(\mathcal{D}))$ *is the mild solution of system (4.20) and* $a\colon [0,T] \to D((\lambda - A)^{1-\alpha})$ *is the unique solution of equation (4.22).*

Remark 4.16. *If system (4.3) is driven by Lévy processes, then one can obtain the optimal controls stated in the previous remark as follows:*
Note that the design of the feedback law presented in this section is based on a martingale representation theorem. Similarly to Remark 3.88 such a martingale representation theorem can only be derived if the filtration is generated by independent real-valued Lévy processes. Hence, a feedback law of the optimal controls can be derived if system (4.3) is driven by real-valued Lévy processes.

Chapter 5

Optimal Control of Uncertain Stokes Flows

In this chapter, we consider a control problem constrained by the unsteady stochastic Stokes equations with nonhomogeneous Dirichlet boundary conditions. Here controls appear inside the domain as distributed controls and on the boundary as tangential controls. Motivated by [76], we first analyze the deterministic unsteady Stokes equations with nonhomogeneous Dirichlet boundary conditions. Similarly to the previous chapter, we reformulate these equations as an evolution equation in a suitable Hilbert space such that the existence and uniqueness of a solution can be obtained using fractional powers of closed operators introduced in Section 2.3. Based on this approach, we extend the Stokes equations by an additional noise term. An existence and uniqueness result of a mild solution to the stochastic Stokes equations is provided in Section 3.4.1. This enables us to solve uniquely a tracking problem using a stochastic maximum principle, which gives us necessary and sufficient optimality conditions the optimal controls have to satisfy. Through a duality principle, we can utilize these optimality conditions to calculate the optimal controls. As a consequence, it remains to solve a coupled system of forward and backward SPDEs. The results presented here are mainly based on [8].

Throughout this chapter, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given complete probability space endowed with a normal filtration $(\mathcal{F}_t)_{t\geq 0}$.

5.1. Motivation

In this section, we consider the deterministic Stokes equations with nonhomogeneous Dirichlet boundary conditions. Here, we restrict the problem to tangential boundary conditions. A general formulation can be found in [76]. Let $\mathcal{D} \subset \mathbb{R}^n$, $n \geq 2$, be a connected and bounded domain with C^2 boundary $\partial\mathcal{D}$ and let $T > 0$. We introduce the Stokes equations with nonhomogeneous Dirichlet boundary conditions:

$$\begin{cases} \dfrac{\partial}{\partial t} y(t,x) - \Delta y(t,x) + \nabla p(t,x) = f(t,x) & \text{in } (0,T) \times \mathcal{D}, \\ \text{div } y(t,x) = 0 & \text{in } (0,T) \times \mathcal{D}, \\ y(t,x) = g(t,x) & \text{on } (0,T) \times \partial\mathcal{D}, \\ y(0,x) = \xi(x) & \text{in } \mathcal{D}, \end{cases} \tag{5.1}$$

where $y(t,x) \in \mathbb{R}^n$ denotes the velocity field with initial value $\xi(x) \in \mathbb{R}^n$, $p(t,x) \in \mathbb{R}$ describes the pressure of the fluid and $f(t,x) \in \mathbb{R}^n$ is the external force. The boundary condition $g(t,x) \in \mathbb{R}^n$ is assumed to be tangential, i.e.

$$g(t,x) \cdot \eta(x) = 0 \quad \text{on } (0,T) \times \partial\mathcal{D}$$

in the sense of the inner product in $\mathbb{R}^n$, where η denotes the unit outward normal to $\partial\mathcal{D}$. Next, we reformulate system (5.1) as an evolution equation. According to Section 2.5.2, let us introduce the following Hilbert spaces:

$$\begin{aligned} H &= \left\{ y \in (L^2(\mathcal{D}))^n : \text{div } y = 0 \text{ in } \mathcal{D}, y \cdot \eta = 0 \text{ on } \partial\mathcal{D} \right\}, \\ V &= \left\{ y \in \left(H_0^1(\mathcal{D})\right)^n : \text{div } y = 0 \text{ in } \mathcal{D} \right\} \end{aligned}$$

and let $A\colon D(A) \subset H \to H$ be the Stokes operator given by

$$D(A) = (H^2(\mathcal{D}))^n \cap V, \quad Ay = -\Pi\Delta y$$

for every $y \in D(A)$, where the operator $\Pi\colon (L^2(\mathcal{D}))^n \to H$ is an orthogonal projection. By Theorem 2.61, the operator $-A$ is the generator of an analytic semigroup of contractions $(e^{-At})_{t\geq 0}$. Hence, we can introduce fractional powers of A denoted by A^α with $\alpha \in \mathbb{R}$ according to Section 2.3. Furthermore, let us define the following spaces for $s \geq 0$:

$$V^s(\mathcal{D}) = \{y \in (H^s(\mathcal{D}))^n : \operatorname{div} y = 0 \text{ in } \mathcal{D}, y \cdot \eta = 0 \text{ on } \partial\mathcal{D}\},$$
$$V^s(\partial\mathcal{D}) = \{y \in (H^s(\partial\mathcal{D}))^n : y \cdot \eta = 0 \text{ on } \partial\mathcal{D}\}.$$

For $s < 0$, the space $V^s(\partial\mathcal{D})$ is the dual space of $V^{-s}(\partial\mathcal{D})$ with $V^0(\partial\mathcal{D})$ as pivot space. Moreover, let $H^s(\mathcal{D})/\mathbb{R}$ with $s \geq 0$ be the quotient space of $H^s(\mathcal{D})$ by $\mathbb{R}$, i.e. $H^s(\mathcal{D})/\mathbb{R} = \{y + c\colon y \in H^s(\mathcal{D}), c \in \mathbb{R}\}$. We set $\|y\|_{H^s(\mathcal{D})/\mathbb{R}} = \inf_{c\in\mathbb{R}} \|y + c\|_{H^s(\mathcal{D})}$ for every $y \in H^s(\mathcal{D})/\mathbb{R}$. The dual space is denoted by $(H^s(\mathcal{D})/\mathbb{R})'$ with $H^0(\mathcal{D})/\mathbb{R}$ as pivot space.

Next, let us consider the system

$$\begin{cases} -\Delta w + \nabla\pi = 0 \quad \text{and} \quad \operatorname{div} w = 0 & \text{in } \mathcal{D}, \\ w = g & \text{on } \partial\mathcal{D}. \end{cases} \tag{5.2}$$

We have the following existence and uniqueness results.

Proposition 5.1 (cf. Theorem IV.6.1 (a),[40])**.** *If we assume that $g \in V^{3/2}(\partial\mathcal{D})$, then there exists a unique solution $(w, \pi) \in V^2(\mathcal{D}) \times H^1(\mathcal{D})/\mathbb{R}$ of system (5.2) and the following estimate holds:*

$$\|w\|_{V^2(\mathcal{D})} + \|\pi\|_{H^1(\mathcal{D})/\mathbb{R}} \leq C^* \|g\|_{V^{3/2}(\partial\mathcal{D})},$$

where $C^* > 0$ is a constant.

Proposition 5.2 (cf. [41, 76])**.** *If we assume that $g \in V^{-1/2}(\partial\mathcal{D})$, then there exists a unique solution $(w, \pi) \in V^0(\mathcal{D}) \times \left(H^1(\mathcal{D})/\mathbb{R}\right)'$ of system (5.2) and the following estimate holds:*

$$\|w\|_{V^0(\mathcal{D})} + \|\pi\|_{(H^1(\mathcal{D})/\mathbb{R})'} \leq C^* \|g\|_{V^{-1/2}(\partial\mathcal{D})},$$

where $C^* > 0$ is a constant.

We introduce the Dirichlet operators D and D_p defined by

$$Dg = w \quad \text{and} \quad D_p g = \pi,$$

where (w, π) is the solution of system (5.2). We get the following properties of the Dirichlet operators, which is an immediate consequence of Proposition 5.1 and Proposition 5.2.

Corollary 5.3 (Corollary A.1, [76])**.** *The operator D is linear and continuous from $V^s(\partial\mathcal{D})$ into $V^{s+1/2}(\mathcal{D})$ for all $-\frac{1}{2} \leq s \leq \frac{3}{2}$. If $-\frac{1}{2} \leq s < \frac{1}{2}$, then the operator D_p is linear and continuous from $V^s(\partial\mathcal{D})$ into $\left(H^{1/2-s}(\mathcal{D})/\mathbb{R}\right)'$ and if $\frac{1}{2} \leq s \leq \frac{3}{2}$, then the operator D_p is linear and continuous from $V^s(\partial\mathcal{D})$ into $H^{s-1/2}(\mathcal{D})/\mathbb{R}$.*

As a consequence of Corollary 2.63 and Corollary 5.3, we get $D \in \mathcal{L}\left(V^0(\partial\mathcal{D}); D(A^\beta)\right)$ for $\beta \in \left(0, \frac{1}{4}\right)$. By the closed graph theorem, we have $A^\beta D \in \mathcal{L}\left(V^0(\partial\mathcal{D}); V^0(\mathcal{D})\right)$. We note that $V^0(\mathcal{D}) = H$. Furthermore, system (5.1) can be rewritten in the following form:

$$\begin{cases} \dfrac{d}{dt} y(t) = -Ay(t) + ADg(t) + \Pi f(t), \\ y(0) = \Pi\xi. \end{cases} \tag{5.3}$$

For the sake of simplicity, we assume $f(t), \xi \in H$ for $t \in [0,T]$. Hence, we obtain a linear evolution equation and the solution is given by

$$y(t) = e^{-At}\xi + \int_0^t Ae^{-A(t-s)}Dg(s)\,ds + \int_0^t e^{-A(t-s)}f(s)\,ds.$$

For more details about linear evolution equations, see [9]. The following existence and uniqueness result is stated in [76] for more general boundary conditions and $f = 0$.

Theorem 5.4. *Let $g \in L^2([0,T]; V^0(\partial\mathcal{D}))$ and $f \in L^2([0,T]; H)$. If $\alpha \in [0, \frac{1}{4})$, then for any $\xi \in D(A^\alpha)$, there exists a unique solution $y \in L^2([0,T]; D(A^\alpha))$ of system (5.3) and the following estimate holds:*

$$\|y\|_{L^2([0,T];D(A^\alpha))} \leq C^* \left(\|\xi\|_{D(A^\alpha)} + \|g\|_{L^2([0,T];V^0(\partial\mathcal{D}))} + \|f\|_{L^2([0,T];H)} \right),$$

where $C^* > 0$ is a constant.

5.2. The Controlled Stochastic Stokes Equations

In this section, we consider the controlled stochastic Stokes equations. Here, controls appear as distributed controls inside the domain as well as tangential controls on the boundary. We assume that the external force $f(t)$ in equation (5.3) can be decomposed as the sum of a control term and a noise term dependent on the velocity field $y(t)$. Using the spaces and operators introduced in Section 5.1, we obtain the stochastic Stokes equations in $D(A^\alpha)$:

$$\begin{cases} dy(t) = [-Ay(t) + Bu(t) + ADv(t)]\,dt + G(y(t))\,dW(t), \\ y(0) = \xi, \end{cases} \tag{5.4}$$

where the initial value $\xi \in L^2(\Omega; D(A^\alpha))$ is $\mathcal{F}_0$-measurable and the process $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in H and covariance operator $Q \in \mathcal{L}_1^+(H)$. The set of admissible distributed controls U contains all predictable processes $(u(t))_{t\in[0,T]}$ with values in H such that

$$\mathbb{E}\int_0^T \|u(t)\|_H^2\,dt < \infty.$$

The space U equipped with the inner product of $L^2(\Omega; L^2([0,T]; H))$ becomes a Hilbert space. Similarly, the set of admissible boundary controls V contains all predictable processes $(v(t))_{t\in[0,T]}$ with values in $V^0(\partial\mathcal{D})$ such that

$$\mathbb{E}\int_0^T \|v(t)\|_{V^0(\partial\mathcal{D})}^2\,dt < \infty.$$

The space V equipped with the inner product of $L^2(\Omega; L^2([0,T]; V^0(\partial\mathcal{D})))$ becomes a Hilbert space. The operators $B\colon H \to H$ and $G\colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H); H)$ are linear and bounded. Motivated by Section 5.1, we use a mild solution to system (5.4) in the sense of Definition 3.78 with $\mathcal{H} = H$ and $\mathcal{H}_b = V^0(\partial\mathcal{D})$. As a consequence of Theorem 3.79, there exists a unique mild solution $(y(t))_{t\in[0,T]}$ of system (5.4) for any $\xi \in L^2(\Omega; D(A^\alpha))$ and fixed controls $u \in U$ and $v \in V$. Hence, the process $(y(t))_{t\in[0,T]}$ takes values in $D(A^\alpha)$ with $\alpha \in [0, \frac{1}{4})$ such that

$$\mathbb{E}\int_0^T \|y(t)\|_{D(A^\alpha)}^2 dt < \infty \tag{5.5}$$

and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$y(t) = e^{-At}\xi + \int_0^t e^{-A(t-s)} Bu(s)\,ds + \int_0^t Ae^{-A(t-s)} Dv(s)\,ds + \int_0^t e^{-A(t-s)} G(y(s))\,dW(s).$$

In this chapter, it suffices to require that $(y(t))_{t\in[0,T]}$ satisfies condition (5.5) with $\alpha = 0$ and we assume that the initial value $\xi \in L^2(\Omega; H)$ is fixed. To illustrate the dependence on the controls $u \in U$ and $v \in V$, let us denote by $(y(t;u,v))_{t\in[0,T]}$ the mild solution of system (5.4). Whenever this process is considered for fixed controls, we use the notation introduced above. We get the following properties.

Lemma 5.5. *Let $(y(t;u,v))_{t\in[0,T]}$ be the mild solution of system (5.4) corresponding to the controls $u \in U$ and $v \in V$. Then the process $(y(t;u,v))_{t\in[0,T]}$ is affine linear with respect to u and v and we have for every $u_1, u_2 \in U$ and every $v_1, v_2 \in V$*

$$\mathbb{E}\int_0^T \|y(t;u_1,v_1) - y(t;u_2,v_2)\|_H^2 dt \leq \widehat{C}\left[\mathbb{E}\int_0^T \|u_1(t) - u_2(t)\|_H^2\,dt + \mathbb{E}\int_0^T \|v_1(t) - v_2(t)\|_{V^0(\partial\mathcal{D})}^2\,dt\right], \quad (5.6)$$

where $\widehat{C} > 0$ is a constant.

Proof. First, we show that $(y(t;u,v))_{t\in[0,T]}$ is affine linear with respect to $u \in U$. We assume that $\xi = 0$ and $v = 0$. Moreover, let $a, b \in \mathbb{R}$ and $u_1, u_2 \in U$. Recall that $B\colon H \to H$ and $G\colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H);H)$ are linear and bounded. Moreover, we have $\left\|e^{-At}\right\|_{\mathcal{L}(H)} \leq 1$ for all $t \geq 0$. Using Theorem 3.62 (iii) and Fubini's theorem, there exists a constant $C^* > 0$ such that for $t \in [0,T]$

$$\begin{aligned}
&\mathbb{E}\,\|y(t;a\,u_1 + b\,u_2, 0) - a\,y(t;u_1,0) - b\,y(t;u_2,0)\|_H^2 \\
&\leq \mathbb{E}\left\|\int_0^t e^{-A(t-s)} G(y(t;a\,u_1 + b\,u_2,0) - a\,y(t;u_1,0) - b\,y(t;u_2,0))\,dW(s)\right\|_H^2 \\
&\leq C^* \int_0^t \mathbb{E}\,\|y(s;a\,u_1 + b\,u_2,0) - a\,y(s;u_1,0) - b\,y(s;u_2,0))\|_H^2\,ds.
\end{aligned}$$

By Proposition A.1, we have

$$\mathbb{E}\|y(t;a\,u_1 + b\,u_2,0) - a\,y(t;u_1,0) - b\,y(t;u_2,0)\|_H^2 = 0$$

for $t \in [0,T]$ and thus, we get

$$\mathbb{E}\int_0^T \|y(t;a\,u_1 + b\,u_2,0) - a\,y(t;u_1,0) - b\,y(t;u_2,0)\|_H^2 dt = 0$$

resulting from Fubini's theorem. We obtain that $(y(t;u,0))_{t\in[0,T]}$ with initial value $\xi = 0$ is linear with respect to $u \in U$. For arbitrary $\mathcal{F}_0$-measurable $\xi \in L^2(\Omega;H)$ and arbitrary $v \in V$, we can conclude that $(y(t;u,v))_{t\in[0,T]}$ is affine linear with respect to $u \in U$. Similarly, we obtain that $(y(t;u,v))_{t\in[0,T]}$ is affine linear with respect to $v \in V$.

Next, we show that inequality (5.6) holds. Let $u_1, u_2 \in U$ and $v_1, v_2 \in V$. Recall that $A^\alpha D\colon V^0(\partial\mathcal{D}) \to H$ is linear and bounded for all $\alpha \in \left(0, \frac{1}{4}\right)$. Due to Theorem 2.29 (iv), Theorem 2.35, Theorem 3.62 (iii) and

Fubini's theorem, there exist constants $C_1, C_2, C_3 > 0$ such that for $t \in [0,T]$

$$\begin{aligned}
&\mathbb{E}\,\|y(t;u_1,v_1) - y(t;u_2,v_2)\|_H^2 \\
&\leq 3\,\mathbb{E}\int_0^t \left\|e^{-A(t-s)}B[u_1(s)-u_2(s)]\right\|_H^2 ds + 3\,\mathbb{E}\left(\int_0^t \left\|A^{1-\alpha}e^{-A(t-s)}A^{\alpha}D[v_1(s)-v_2(s)]\right\|_H ds\right)^2 \\
&\quad + 3\,\mathbb{E}\left\|\int_0^t e^{-A(t-s)}G(y(s;u_1,v_1)-y(s;u_2,v_2))\,dW(s)\right\|_H^2 \\
&\leq C_1\,\mathbb{E}\int_0^t \|u_1(s)-u_2(s)\|_H^2\,ds + C_2\,\mathbb{E}\left(\int_0^t (t-s)^{\alpha-1}\|v_1(s)-v_2(s)\|_{V^0(\partial\mathcal{D})}\,ds\right)^2 \\
&\quad + C_3\int_0^t \mathbb{E}\,\|y(s;u_1,v_1)-y(s;u_2,v_2)\|_H^2 ds.
\end{aligned}$$

Using Corollary A.4, Fubini's theorem and Young's inequality for convolutions, we get for $t \in [0,T]$

$$\begin{aligned}
&\mathbb{E}\,\|y(t;u_1,v_1) - y(t;u_2,v_2)\|_H^2 \\
&\leq C_1\,\mathbb{E}\int_0^t \|u_1(s)-u_2(s)\|_H^2\,ds + C_2\,\mathbb{E}\left(\int_0^t (t-s)^{\alpha-1}\|v_1(s)-v_2(s)\|_{V^0(\partial\mathcal{D})}\,ds\right)^2 \\
&\quad + C_3\int_0^t e^{C_3(t-s)}\left[C_1\,\mathbb{E}\int_0^s \|u_1(r)-u_2(r)\|_H^2\,dr + C_2\,\mathbb{E}\left(\int_0^s (s-r)^{\alpha-1}\|v_1(r)-v_2(r)\|_{V^0(\partial\mathcal{D})}\,dr\right)^2\right] ds \\
&\leq C_1\left(1 + C_3 e^{C_3 t} t\right)\,\mathbb{E}\int_0^t \|u_1(s)-u_2(s)\|_H^2\,ds + C_2\,\mathbb{E}\left(\int_0^t (t-s)^{\alpha-1}\|v_1(s)-v_2(s)\|_{V^0(\partial\mathcal{D})}\,ds\right)^2 \\
&\quad + \frac{C_2 C_3 e^{C_3 t} t^{2\alpha}}{\alpha^2}\,\mathbb{E}\int_0^t \|v_1(s)-v_2(s)\|_{V^0(\partial\mathcal{D})}^2\,ds.
\end{aligned}$$

By Fubini's theorem and Young's inequality for convolutions, there exists a constant $\widehat{C} > 0$ such that

$$\mathbb{E}\int_0^T \|y(t;u_1,v_1)-y(t;u_2,v_2)\|_H^2 dt \leq \widehat{C}\left[\mathbb{E}\int_0^T \|u_1(t)-u_2(t)\|_H^2\,dt + \mathbb{E}\int_0^T \|v_1(t)-v_2(t)\|_{V^0(\partial\mathcal{D})}^2\,dt\right].$$

□

5.3. A Tracking Problem

The control problem considered here is motivated by [4, 17, 56, 75, 79]. In this section, we state necessary and sufficient optimality conditions the optimal controls have to satisfy. Let us introduce the following cost functional:

$$J(u,v) = \frac{1}{2}\,\mathbb{E}\int_0^T \|y(t;u,v) - y_d(t)\|_H^2 dt + \frac{\kappa_1}{2}\,\mathbb{E}\int_0^T \|u(t)\|_H^2\, dt + \frac{\kappa_2}{2}\,\mathbb{E}\int_0^T \|v(t)\|_{V^0(\partial\mathcal{D})}^2\, dt, \tag{5.7}$$

where $(y(t;u,v))_{t\in[0,T]}$ is the mild solution of system (5.4) corresponding to the controls $u \in U$ and $v \in V$. The function $y_d \in L^2([0,T];H)$ is a given desired velocity field and $\kappa_1, \kappa_2 > 0$ are weights. The task is to find controls $\overline{u} \in U$ and $\overline{v} \in V$ such that

$$J(\overline{u},\overline{v}) = \inf_{u\in U, v\in V} J(u,v).$$

The controls $\overline{u} \in U$ and $\overline{v} \in V$ are called optimal controls. Note that the control problem is formulated as an unbounded optimization problem constrained by a SPDE. The functional $J\colon U \times V \to \mathbb{R}$ given by equation (5.7) is coercive, strictly convex and continuous, which is a consequence of Lemma 5.5. Hence, we get the existence and uniqueness of optimal controls resulting from Corollary D.13.

Next, let us introduce the following systems in H:

$$\begin{cases} dz_1(t) = [-Az_1(t) + Bu(t)]\,dt + G(z_1(t))\,dW(t), \\ z_1(0) = 0, \end{cases} \tag{5.8}$$

$$\begin{cases} dz_2(t) = [-Az_2(t) + ADv(t)]\,dt + G(z_2(t))\,dW(t), \\ z_2(0) = 0, \end{cases} \tag{5.9}$$

where $u \in U$, $v \in V$ and $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in H and covariance operator $Q \in \mathcal{L}_1^+(H)$. The operators A, B, D, G and the spaces U, V are introduced in Section 5.1 and Section 5.2, respectively. Again, we use a mild solution to system (5.8) in the sense of Definition 3.78 with $\mathcal{H} = H$, $\mathcal{H}_b = V^0(\partial\mathcal{D})$ and $v = 0$. As a consequence of Theorem 3.79 with $\alpha = 0$, there exists a unique mild solution $(z_1(t))_{t\in[0,T]}$ of system (5.8) for fixed control $u \in U$. Hence, the process $(z_1(t))_{t\in[0,T]}$ takes values in H such that

$$\mathbb{E}\int_0^T \|z_1(t)\|_H^2 dt < \infty$$

and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$z_1(t) = \int_0^t e^{-A(t-s)}Bu(s)\,ds + \int_0^t e^{-A(t-s)}G(z_1(s))\,dW(s).$$

Similarly, there exists a unique mild solution $(z_2(t))_{t\in[0,T]}$ of system (5.9) for fixed control $v \in V$. The process $(z_2(t))_{t\in[0,T]}$ takes values in H such that

$$\mathbb{E}\int_0^T \|z_2(t)\|_H^2 dt < \infty$$

and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$z_2(t) = \int_0^t Ae^{-A(t-s)}Dv(s)\,ds + \int_0^t e^{-A(t-s)}G(z_2(s))\,dW(s).$$

Remark 5.6. *Resulting from Theorem 3.82, the mild solution of system (5.8) satisfies even stronger regularity conditions. Indeed the process* $(z_1(t))_{t\in[0,T]}$ *takes values in* $D(A^\alpha)$ *with* $\alpha \in [0, \frac{1}{2})$ *such that*

$$\mathbb{E} \sup_{t\in[0,T]} \|z_1(t)\|^2_{D(A^\alpha)} < \infty.$$

To illustrate the dependence on the controls $u \in U$ and $v \in V$, let us denote by $(z_1(t;u))_{t\in[0,T]}$ and $(z_2(t;v))_{t\in[0,T]}$ the mild solutions of system (5.8) and system (5.9), respectively. Whenever these processes are considered for fixed controls, we use the notation introduced above. Similarly to Lemma 5.5, we get the following result.

Lemma 5.7. *Let* $(z_1(t;u))_{t\in[0,T]}$ *and* $(z_2(t;v))_{t\in[0,T]}$ *be the mild solutions of system (5.8) and system (5.9) corresponding to the controls* $u \in U$ *and* $v \in V$*, respectively. Then the process* $(z_1(t;u))_{t\in[0,T]}$ *is linear with respect to* u *and the process* $(z_2(t;v))_{t\in[0,T]}$ *is linear with respect to* v*. Moreover, we have for every* $u_1, u_2 \in U$ *and every* $v_1, v_2 \in V$

$$\mathbb{E}\int_0^T \|z_1(t;u_1) - z_1(t;u_2)\|^2_H dt \le \widehat{C}\,\mathbb{E}\int_0^T \|u_1(t) - u_2(t)\|^2_H \,dt,$$

$$\mathbb{E}\int_0^T \|z_2(t;v_1) - z_2(t;v_2)\|^2_H dt \le \widehat{C}\,\mathbb{E}\int_0^T \|v_1(t) - v_2(t)\|^2_{V^0(\partial\mathcal{D})} \,dt,$$

where $\widehat{C} > 0$ *is a constant.*

This enables us to calculate the partial Fréchet derivative of the mild solution to system (5.4).

Theorem 5.8. *Let* $(y(t;u,v))_{t\in[0,T]}$*,* $(z_1(t;u))_{t\in[0,T]}$ *and* $(z_2(t;v))_{t\in[0,T]}$ *be the mild solutions of systems (5.4), (5.8) and (5.9) corresponding to the controls* $u \in U$ *and* $v \in V$*, respectively. Then the partial Fréchet derivative of* $y(t;u,v)$ *at* $u \in U$ *in direction* $\tilde{u} \in U$ *satisfies for fixed* $v \in V$*,* $t \in [0,T]$ *and* $\mathbb{P}$*-a.s.*

$$d^F_u y(t;u,v)[\tilde{u}] = z_1(t;\tilde{u}).$$

The partial Fréchet derivative of $y(t;u,v)$ *at* $v \in V$ *in direction* $\tilde{v} \in V$ *satisfies for fixed* $u \in U$*,* $t \in [0,T]$ *and* $\mathbb{P}$*-a.s.*

$$d^F_v y(t;u,v)[\tilde{v}] = z_2(t;\tilde{v}).$$

Proof. First, we calculate the Fréchet derivative of $y(t;u,v)$ at $u \in U$ in direction $\tilde{u} \in U$. Let $v \in V$ be fixed. Recall that the operators $B\colon H \to H$ and $G\colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H);H)$ are linear and bounded. Moreover, we have $\left\|e^{-At}\right\|_{\mathcal{L}(H)} \le 1$ for all $t \ge 0$. Using Theorem 3.62 (iii) and Fubini's theorem, there exists a constant $C^* > 0$ such that for $t \in [0,T]$

$$\begin{aligned}\mathbb{E}\left\|y(t;u+\tilde{u},v) - y(t;u,v) - z_1(t;\tilde{u})\right\|^2_H &= \mathbb{E}\left\|\int_0^t e^{-A(t-s)}G(y(s;u+\tilde{u},v) - y(s;u,v) - z_1(s;\tilde{u}))\,dW(s)\right\|^2_H \\ &\le C^*\int_0^t \mathbb{E}\left\|y(s;u+\tilde{u},v) - y(s;u,v) - z_1(s;\tilde{u})\right\|^2_H ds.\end{aligned}$$

By Proposition A.1, we have $\mathbb{E}\|y(t;u+\tilde{u},v) - y(t;u,v) - z_1(t;\tilde{u})\|^2_H = 0$ and hence, we obtain

$$\mathbb{E}\int_0^T \|y(t;u+\tilde{u},v) - y(t;u,v) - z_1(t;\tilde{u})\|^2_H dt = 0$$

as a consequence of Fubini's theorem. Therefore, the partial Fréchet derivative of $y(t;u,v)$ at $u \in U$ in direction $\tilde{u} \in U$ satisfies for every $v \in V$, $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$d_u^F y(t;u,v)[\tilde{u}] = z_1(t;\tilde{u}).$$

Due to Lemma 5.7, the operator $d_u^F y(t;u,v)$ is linear and bounded on U. Similarly, we obtain the partial Fréchet derivative of $y(t;u,v)$ at $v \in V$ in direction $\tilde{v} \in V$. □

As a consequence of Remark D.6 and Theorem 5.8, we can calculate the partial Fréchet derivatives of the cost functional (5.7). Indeed, the Fréchet derivative at $u \in U$ in direction $\tilde{u} \in U$ for fixed $v \in V$ satisfies

$$d_u^F J(u,v)[\tilde{u}] = \mathbb{E} \int_0^T \langle y(t;u,v) - y_d(t), z_1(t;\tilde{u}) \rangle_H \, dt + \kappa_1 \, \mathbb{E} \int_0^T \langle u(t), \tilde{u}(t) \rangle_H \, dt, \tag{5.10}$$

where $(z_1(t;\tilde{u}))_{t\in[0,T]}$ is the mild solution of system (5.8) corresponding to the control $\tilde{u} \in U$. The partial Fréchet derivative at $v \in V$ in direction $\tilde{v} \in V$ for fixed $u \in U$ satisfies

$$d_v^F J(u,v)[\tilde{v}] = \mathbb{E} \int_0^T \langle y(t;u,v) - y_d(t), z_2(t;\tilde{v}) \rangle_H \, dt + \kappa_2 \, \mathbb{E} \int_0^T \langle v(t), \tilde{v}(t) \rangle_{V^0(\partial\mathcal{D})} \, dt, \tag{5.11}$$

where $(z_2(t;\tilde{v}))_{t\in[0,T]}$ is the mild solution of system (5.9) corresponding to the control $\tilde{v} \in V$. Since the cost functional $J \colon U \times V \to \mathbb{R}$ given by (5.7) is convex, we can apply Proposition D.14. Hence, the optimal controls $\overline{u} \in U$ and $\overline{v} \in V$ satisfy the following necessary and sufficient optimality conditions:

$$d_u^F J(\overline{u},\overline{v})[\tilde{u}] = 0, \tag{5.12}$$
$$d_v^F J(\overline{u},\overline{v})[\tilde{v}] = 0 \tag{5.13}$$

for every $\tilde{u} \in U$ and every $\tilde{v} \in V$.

Remark 5.9. *Note that the necessary and sufficient optimality conditions (5.12) and (5.13) can be easily obtained if system (5.4) is driven by a square integrable Lévy martingale as introduced in Section 3.3.*

5.4. The Adjoint Equation

We use the optimality conditions (5.12) and (5.13) to derive explicit formulas for the optimal controls $\overline{u} \in U$ and $\overline{v} \in V$. Therefor, we need a duality principle, which gives us a relation between the Fréchet derivatives of the mild solution to system (5.4) and the adjoint equation, which is given by the following backward SPDE in H:

$$\begin{cases} dz^*(t) = -[-Az^*(t) + G^*(\Phi(t)) + y(t) - y_d(t)]dt + \Phi(t)\, dW(t), \\ z^*(T) = 0, \end{cases} \tag{5.14}$$

where $(y(t))_{t\in[0,T]}$ is the mild solution of system (5.4) and $y_d \in L^2([0,T];H)$ is the desired velocity field. The process $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in H and covariance operator $Q \in \mathcal{L}_1^+(H)$ and the operator $G^* \colon \mathcal{L}_{(HS)}(Q^{1/2}(H);H) \to H$ is linear and bounded. A precise meaning is given in the following remark.

Remark 5.10. *Recall that the operator $G \colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H);H)$ is linear and bounded. Therefore, there exists a linear and bounded operator $G^* \colon \mathcal{L}_{(HS)}(Q^{1/2}(H);H) \to H$ satisfying for every $h \in H$ and every $\Phi \in \mathcal{L}_{(HS)}(Q^{1/2}(H);H)$*

$$\langle G(h), \Phi \rangle_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} = \langle h, G^*(\Phi) \rangle_H. \tag{5.15}$$

Here, we use a mild solution to system (5.14) in the sense of Definition 3.91 with $\mathcal{H} = H$. Recall that there exists a unique mild solution $(y(t))_{t\in[0,T]}$ of system (5.4) for fixed controls $u \in U$ and $v \in V$. As a consequence of Theorem 3.93, we can conclude that there exists a unique mild solution $(z^*(t), \Phi(t))_{t\in[0,T]}$ of system (5.14) for fixed controls $u \in U$ and $v \in V$. Hence, the pair of processes $(z^*(t), \Phi(t))_{t\in[0,T]}$ takes values in $H \times \mathcal{L}_{(HS)}(Q^{1/2}(H); H)$ such that

$$\sup_{t\in[0,T]} \mathbb{E}\|z^*(t)\|_H^2 < \infty, \qquad \mathbb{E}\int_0^T \|\Phi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dt < \infty$$

and we have for all $t \in [0, T]$ and $\mathbb{P}$-a.s.

$$z^*(t) = \int_t^T e^{-A(s-t)} G^*(\Phi(s))\, ds + \int_t^T e^{-A(s-t)} \left(y(s) - y_d(s)\right) ds - \int_t^T e^{-A(s-t)} \Phi(s)\, dW(s).$$

Furthermore, note that the mild solution of system (5.4) depends on the controls $u \in U$ and $v \in V$. Thus, we get this property for the mild solution of system (5.14) as well. To illustrate the dependence on the controls $u \in U$ and $v \in V$, let us denote by $(z^*(t; u, v), \Phi(t; u, v))_{t\in[0,T]}$ the mild solution of system (5.14). Whenever these processes are considered for fixed controls, we use the notation introduced above. For the process $(z^*(t; u, v))_{t\in[0,T]}$, one can show another important regularity property. Therefor, we need a modification of Young's inequality for convolutions.

Lemma 5.11. *Let $f \in L^p([0, T])$ and $g \in L^q([0, T])$ be arbitrary. We set for $t \in [0, T]$*

$$h(t) = \int_t^T f(s - t) g(s)\, ds.$$

If $p, q, r \geq 1$ satisfy $\frac{1}{p} + \frac{1}{q} = \frac{1}{r} + 1$, then $h \in L^r([0, T])$ and

$$\|h\|_{L^r([0,T])} \leq \|f\|_{L^p([0,T])} \|g\|_{L^q([0,T])}.$$

Proof. The proof can be obtained similarly to the classical version of Young's inequality for convolutions, see [13, Theorem 3.9.4]. □

Proposition 5.12. *Let $(z^*(t; u, v), \Phi(t; u, v))_{t\in[0,T]}$ be the mild solution of system (5.14) corresponding to the controls $u \in U$ and $v \in V$. Then $(z^*(t; u, v))_{t\in[0,T]}$ takes values in $D(A^\varepsilon)$ with $\varepsilon \in [0, 1)$ such that*

$$\mathbb{E}\int_0^T \|z^*(t; u, v)\|^2_{D(A^\varepsilon)} dt < \infty.$$

Proof. For the sake of simplicity, we omit the dependence on the controls. Since $(z^*(t))_{t\in[0,T]}$ is predictable, we get for $t \in [0, T]$ and $\mathbb{P}$-a.s.

$$z^*(t) = \mathbb{E}\left[\int_t^T e^{-A(s-t)} G^*(\Phi(s))\, ds + \int_t^T e^{-A(s-t)} \left(y(s) - y_d(s)\right) ds \,\middle|\, \mathcal{F}_t \right].$$

Recall that the operator $G^*\colon \mathcal{L}_{(HS)}(Q^{1/2}(H);H) \to H$ is bounded. Using Theorem 2.35, Proposition B.9 and Lemma 5.11, the process $(z^*(t))_{t\in[0,T]}$ takes values in $D(A^\varepsilon)$ with $\varepsilon \in [0,1)$ and there exists a constant $C^* > 0$ such that

$$\begin{aligned}
&\mathbb{E}\int_0^T \|z^*(t)\|^2_{D(A^\varepsilon)} dt \\
&\leq 2\,\mathbb{E}\int_0^T \left(\int_t^T \|A^\varepsilon e^{-A(s-t)} G^*(\Phi(s))\|_H ds\right)^2 dt + 2\,\mathbb{E}\int_0^T \left(\int_t^T \|A^\varepsilon e^{-A(s-t)}\left(y(s) - y_d(s)\right)\|^2_H ds\right)^2 dt \\
&\leq 2M_\varepsilon^2\,\mathbb{E}\int_0^T \left(\int_t^T (s-t)^{-\varepsilon} \|G^*(\Phi(s))\|_H ds\right)^2 dt + 2M_\varepsilon^2\,\mathbb{E}\int_0^T \left(\int_t^T (s-t)^{-\varepsilon} \|y(s) - y_d(s)\|_H ds\right)^2 dt \\
&\leq C^* \left[\mathbb{E}\int_0^T \|\Phi(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dt + \mathbb{E}\int_0^T \|y(t)\|^2_H dt + \int_0^T \|y_d(t)\|^2_H dt\right].
\end{aligned}$$

□

5.5. Approximation by a Strong Formulation

In general, a duality principle of solutions to forward and backward SPDEs can be obtained by applying an Itô product formula. For Q-Wiener processes, this is provided by Corollary 3.69, which is not applicable to solutions in a mild sense. Hence, we need to approximate the mild solutions of systems (5.8), (5.9) and (5.14) by strong formulations. One method is given by introducing the Yosida approximation of the operator A, see [23]. For applications regarding duality principles, see [36, 84]. However, we apply the method introduced in [45, 53]. The basic idea is to formulate a mild solution with values in $D(A)$ by using the resolvent operator introduced in Section 2.1. Thus, we get the required convergences and the mild solutions coincides with the strong solutions using results stated in Section 3.4.3. In this section, we omit the dependence on the controls for the sake of simplicity. According to Section 2.1, let us denote by $R(\lambda; -A) \in \mathcal{L}(H)$ the resolvent operator of $-A$ with $\lambda \in \rho(-A)$. We introduce the operator $R(\lambda) \in \mathcal{L}(H)$ given by

$$R(\lambda) = \lambda R(\lambda; -A) \tag{5.16}$$

for all $\lambda \in \rho(-A)$. Then we get the following properties.

Lemma 5.13. *Let the operator $R(\lambda) \in \mathcal{L}(H)$ be given by equation (5.16). Then we have*

(i) $R(\lambda)y \in D(A)$ *for every* $y \in H$*;*

(ii) $\|R(\lambda)\|_{\mathcal{L}(H)} \leq 1$ *for all* $\lambda > 0$*;*

(iii) $\lim_{\lambda\to\infty} R(\lambda)y = y$ *for every* $y \in H$*;*

(iv) $A^\alpha R(\lambda)y = R(\lambda)A^\alpha y$ *for every* $y \in D(A^\alpha)$ *with* $\alpha < 1$*;*

(v) $R(\lambda)$ *is self-adjoint on* H*.*

Proof. The assertion (i) is an immediate consequence of the definition of the resolvent operators $R(\lambda; -A)$. Recall that $(e^{-At})_{t\geq 0}$ is an analytic semigroup of contractions. Hence, we obtain (ii) by Theorem 2.11 and

(iii) results from Corollary 2.12. Using Corollary 2.37, we get (iv). It remains to show (v). By Theorem 2.9, we have

$$R(\lambda; A) = \int_0^\infty e^{-\lambda t} e^{-At}\, dt.$$

Since the operator A is self-adjoint, we can conclude that the semigroup $(e^{-At})_{t\geq 0}$ is self-adjoint as well. Thus, we get the result. □

5.5.1. The Forward Equations

Here, we provide approximations of the mild solutions to system (5.8) and system (5.9). We introduce the following systems in $D(A)$:

$$\begin{cases} dz_1(t,\lambda) = [-Az_1(t,\lambda) + R(\lambda)Bu(t)]\, dt + R(\lambda)G(R(\lambda)z_1(t,\lambda))\, dW(t), \\ z_1(0,\lambda) = 0, \end{cases} \tag{5.17}$$

$$\begin{cases} dz_2(t,\lambda) = [-Az_2(t,\lambda) + AR(\lambda)Dv(t)]\, dt + R(\lambda)G(R(\lambda)z_2(t,\lambda))\, dW(t), \\ z_2(0,\lambda) = 0, \end{cases} \tag{5.18}$$

where $u \in U$ and $v \in V$. The process $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in H and covariance operator $Q \in \mathcal{L}_1^+(H)$. The operators A, B, D, G and the spaces U, V are introduced in Section 5.1 and Section 5.2, respectively. The operator $R(\lambda)$ is given by equation (5.16) with $\lambda > 0$.

Remark 5.14. *Note that the approximation scheme provided in [45, 53] differs to the approximation scheme introduced by system (5.17) or system (5.18). Here, the additional operator $R(\lambda)$ is necessary to obtain a duality principle.*

Similarly to Section 5.3, we introduce mild solutions to system (5.17) and system (5.18).

Definition 5.15. *a) A predictable process $(z_1(t,\lambda))_{t\in[0,T]}$ with values in $D(A)$ is called a* ***mild solution of system (5.17)*** *if*

$$\mathbb{E} \int_0^T \|z_1(t,\lambda)\|^2_{D(A)} dt < \infty$$

and we have for $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$z_1(t,\lambda) = \int_0^t e^{-A(t-s)} R(\lambda)Bu(s)\, ds + \int_0^t e^{-A(t-s)} R(\lambda)G(R(\lambda)z_1(s,\lambda))\, dW(s).$$

b) A predictable process $(z_2(t,\lambda))_{t\in[0,T]}$ with values in $D(A)$ is called a ***mild solution of system (5.18)*** *if*

$$\mathbb{E} \int_0^T \|z_2(t,\lambda)\|^2_{D(A)} dt < \infty$$

and we have for $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$z_2(t,\lambda) = \int_0^t e^{-A(t-s)} AR(\lambda)Dv(s)\, ds + \int_0^t e^{-A(t-s)} R(\lambda)G(R(\lambda)z_2(s,\lambda))\, dW(s).$$

Recall that the operator $R(\lambda)$ is linear and bounded on H. As a consequence of Lemma 5.13 (i) and the closed graph theorem, the operator $AR(\lambda)$ is linear and bounded on H as well. Hence, existence and uniqueness results of mild solutions to system (5.17) and system (5.18) with fixed $\lambda > 0$ can be obtained similarly to Theorem 3.79. The following lemma provides strong formulations of the mild solutions to system (5.17) and system (5.18).

Lemma 5.16. *Let $(z_1(t,\lambda))_{t\in[0,T]}$ and $(z_2(t,\lambda))_{t\in[0,T]}$ be the mild solutions of system (5.17) and system (5.18), respectively. Then we have for fixed $\lambda > 0$, $t \in [0,T]$ and $\mathbb{P}$-a.s.*

$$z_1(t,\lambda) = \int_0^t (-A)z_1(s,\lambda) + R(\lambda)Bu(s)\,ds + \int_0^t R(\lambda)G(R(\lambda)z_1(s,\lambda))\,dW(s),$$
$$z_2(t,\lambda) = \int_0^t (-A)z_2(s,\lambda) + AR(\lambda)Dv(s)\,ds + \int_0^t R(\lambda)G(R(\lambda)z_2(s,\lambda))\,dW(s).$$

Proof. The claim follows immediately from Theorem 2.35, Theorem 3.106 and Lemma 5.13. □

We have the following convergence results.

Lemma 5.17. *(i) Let $(z_1(t))_{t\in[0,T]}$ and $(z_1(t,\lambda))_{t\in[0,T]}$ be the mild solutions of system (5.8) and system (5.17), respectively. Then we have*

$$\lim_{\lambda\to\infty} \mathbb{E}\int_0^T \|z_1(t) - z_1(t,\lambda)\|_H^2 dt = 0.$$

(ii) Let $(z_2(t))_{t\in[0,T]}$ and $(z_2(t,\lambda))_{t\in[0,T]}$ be the mild solutions of system (5.9) and system (5.18), respectively. Then we have

$$\lim_{\lambda\to\infty} \mathbb{E}\int_0^T \|z_2(t) - z_2(t,\lambda)\|_H^2 dt = 0.$$

Proof. First, we show part (i). Let I be the identity operator on H. Recall that $G\colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H);H)$ is linear and bounded. By definition, we have for all $\lambda > 0$, $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} z_1(t) - z_1(t,\lambda) = &\int_0^t e^{-A(t-s)}[I - R(\lambda)]Bu(s)ds \\ &+ \int_0^t e^{-A(t-s)}G([I-R(\lambda)]z_1(s))\,dW(s) \\ &+ \int_0^t e^{-A(t-s)}[I-R(\lambda)]G(R(\lambda)z_1(s))\,dW(s) \\ &+ \int_0^t e^{-A(t-s)}R(\lambda)G(R(\lambda)\,[z_1(s) - z_1(s,\lambda)])\,dW(s). \end{aligned}$$

The remaining part of the proof can be obtained similarly to [53, Lemma 3.1] using Corollary A.4.

Next, we prove part (ii). By definition, we obtain for all $\lambda > 0$, $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$
\begin{aligned}
z_2(t) - z_2(t,\lambda) = &\int_0^t A e^{-A(t-s)}[I - R(\lambda)]Dv(s)\,ds \\
&+ \int_0^t e^{-A(t-s)} G([I - R(\lambda)]z_2(s))\,dW(s) \\
&+ \int_0^t e^{-A(t-s)}[I - R(\lambda)]G(R(\lambda)z_2(s))\,dW(s) \\
&+ \int_0^t e^{-A(t-s)} R(\lambda) G(R(\lambda)\,[z_2(s) - z_2(s,\lambda)])\,dW(s).
\end{aligned}
$$

Thus, we get for all $\lambda > 0$ and $t \in [0,T]$

$$
\mathbb{E}\,\|z_2(t) - z_2(t,\lambda)\|_H^2 \leq 4\,\mathcal{I}_1(t,\lambda) + 4\,\mathcal{I}_2(t,\lambda) + 4\,\mathcal{I}_3(t,\lambda), \tag{5.19}
$$

where

$$
\begin{aligned}
\mathcal{I}_1(t,\lambda) &= \mathbb{E}\left\|\int_0^t A e^{-A(t-s)}[I - R(\lambda)]Dv(s)\,ds\right\|_H^2, \\
\mathcal{I}_2(t,\lambda) &= \mathbb{E}\left\|\int_0^t e^{-A(t-s)} G([I - R(\lambda)]z_2(s))\,dW(s)\right\|_H^2 \\
&\quad + \mathbb{E}\left\|\int_0^t e^{-A(t-s)}[I - R(\lambda)]G(R(\lambda)z_2(s))\,dW(s)\right\|_H^2, \\
\mathcal{I}_3(t,\lambda) &= \mathbb{E}\left\|\int_0^t e^{-A(t-s)} R(\lambda) G(R(\lambda)\,[z_2(s) - z_2(s,\lambda)])\,dW(s)\right\|_H^2.
\end{aligned}
$$

Recall that $D\colon V^0(\partial\mathcal{D}) \to D(A^\alpha)$ for all $\alpha \in \left(0, \frac{1}{4}\right)$. Using Theorem 2.29 (iv), Theorem 2.35, Lemma 5.13 (iv), Fubini's theorem and Young's inequality for convolutions, there exists a constant $C_1 > 0$ such that for all $\lambda > 0$ and all $t \in [0,T]$

$$
\begin{aligned}
\int_0^t \mathcal{I}_1(s,\lambda)\,ds &\leq \mathbb{E}\int_0^t \left(\int_0^s \left\|A^{1-\alpha} e^{-A(s-r)}[I - R(\lambda)]A^\alpha Dv(r)\right\|_H dr\right)^2 ds \\
&\leq C_1\,\mathbb{E}\int_0^T \|[I - R(\lambda)]A^\alpha Dv(t)\|_H^2\,dt.
\end{aligned} \tag{5.20}
$$

Recall that $\left\|e^{-At}\right\|_{\mathcal{L}(H)} \leq 1$ for all $t \geq 0$. Due to Theorem 3.62 (iii) and Fubini's theorem, there exists a constant $C_2 > 0$ such that for all $\lambda > 0$ and all $t \in [0,T]$

$$\begin{aligned}
&\int_0^t \mathcal{I}_2(s,\lambda)\,ds \\
&\leq \int_0^t \mathbb{E}\int_0^s \left\| e^{-A(s-r)} G([I-R(\lambda)]z_2(r)) \right\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dr\,ds \\
&\quad + \int_0^t \mathbb{E}\int_0^s \left\| e^{-A(s-r)} [I-R(\lambda)] G(R(\lambda)z_2(r)) \right\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dr\,ds \\
&\leq C_2 \left[\mathbb{E}\int_0^T \|[I-R(\lambda)]z_2(t)\|_H^2\,dt + \mathbb{E}\int_0^T \|[I-R(\lambda)]G(R(\lambda)z_2(t))\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)}\,dt \right]. \qquad (5.21)
\end{aligned}$$

By Theorem 3.62 (iii), Lemma 5.13 (ii) and Fubini's theorem, there exists a constant $C_3 > 0$ such that for all $\lambda > 0$ and all $t \in [0,T]$

$$\mathcal{I}_3(t,\lambda) \leq C_3 \int_0^t \mathbb{E}\,\|z_2(s) - z_2(s,\lambda)\|_H^2\,ds.$$

Due to inequality (5.19), we get for all $\lambda > 0$ and $t \in [0,T]$

$$\mathbb{E}\,\|z_2(t) - z_2(t,\lambda)\|_H^2 \leq 4\,\mathcal{I}_1(t,\lambda) + 4\,\mathcal{I}_2(t,\lambda) + 4C_3 \int_0^t \mathbb{E}\,\|z_2(s) - z_2(s,\lambda)\|_H^2\,ds.$$

Applying Corollary A.4, we obtain for all $\lambda > 0$ and $t \in [0,T]$

$$\mathbb{E}\,\|z_2(t) - z_2(t,\lambda)\|_H^2 \leq 4\,\mathcal{I}_1(t,\lambda) + 4\,\mathcal{I}_2(t,\lambda) + 16C_3 e^{4C_3 t} \left[\int_0^t \mathcal{I}_1(s,\lambda)\,ds + \int_0^t \mathcal{I}_2(s,\lambda)\,ds \right]. \qquad (5.22)$$

Using equation (5.22), Fubini's theorem, inequality (5.20) and inequality (5.21), there exists a constant $C^* > 0$ such that for all $\lambda > 0$

$$\begin{aligned}
\mathbb{E}\int_0^T \|z_2(t) - z_2(t,\lambda)\|_H^2\,dt &\leq C^*\,\mathbb{E}\int_0^T \|[I-R(\lambda)]A^\alpha Dv(t)\|_H^2\,dt + C^*\,\mathbb{E}\int_0^T \|[I-R(\lambda)]z_2(t)\|_H^2\,dt \\
&\quad + C^*\,\mathbb{E}\int_0^T \|[I-R(\lambda)]G(R(\lambda)z_2(t))\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)}\,dt.
\end{aligned}$$

By Lemma 5.13 (iii) and Proposition B.7, we can infer

$$\lim_{\lambda\to\infty} \mathbb{E}\int_0^T \|z_2(t) - z_2(t,\lambda)\|_H^2\,dt = 0.$$

□

5.5.2. The Backward Equation

Here we provide an approximation of the mild solution to system (5.14). We introduce the following backward SPDE:

$$\begin{cases} dz^*(t,\lambda) = -[-Az^*(t,\lambda) + R(\lambda)G^*(R(\lambda)\Phi(t,\lambda)) + R(\lambda)(y(t)-y_d(t))]\,dt + \Phi(t,\lambda)\,dW(t), \\ z^*(T,\lambda) = 0, \end{cases} \tag{5.23}$$

where $\lambda > 0$. The process $(y(t))_{t\in[0,T]}$ is the mild solution of system (5.4) and $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in H and covariance operator $Q \in \mathcal{L}_1^+(H)$. The function $y_d \in L^2([0,T];H)$ is the desired velocity field. The operators A and G^* are introduced in Section 5.1 and Section 5.4, respectively. The operator $R(\lambda)$ is given by equation (5.16) with $\lambda > 0$. Similarly to Section 5.4, we introduce a mild solution to system (5.23).

Definition 5.18. *A pair of predictable processes $(z^*(t,\lambda),\Phi(t,\lambda))_{t\in[0,T]}$ with values in the product space $D(A)\times\mathcal{L}_{(HS)}(Q^{1/2}(H);H)$ is called a* ***mild solution of system (5.23)*** *if*

$$\sup_{t\in[0,T]} \mathbb{E}\,\|z^*(t,\lambda)\|_{D(A)}^2 < \infty, \qquad \mathbb{E}\int_0^T \|\Phi(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dt < \infty,$$

and we have for all $t\in[0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} z^*(t,\lambda) = &\int_t^T e^{-A(s-t)}R(\lambda)G^*(R(\lambda)\Phi(s,\lambda))\,ds + \int_t^T e^{-A(s-t)}R(\lambda)\,(y(s)-y_d(s))\,ds \\ &- \int_t^T e^{-A(s-t)}\Phi(s,\lambda)\,dW(s). \end{aligned}$$

Recall that the operators $R(\lambda)$ and $AR(\lambda)$ are linear and bounded on H. Hence, existence and uniqueness results of the mild solution to system (5.23) can be obtained similarly to Theorem 3.93. The following lemma states a strong formulation of the mild solution to system (5.23).

Lemma 5.19. *Let the pair of stochastic processes $(z^*(t,\lambda),\Phi(t,\lambda))_{t\in[0,T]}$ be the mild solution of system (5.23). Then we have for fixed $\lambda > 0$, all $t\in[0,T]$ and $\mathbb{P}$-a.s.*

$$z^*(t,\lambda) = \int_t^T (-A)z^*(s,\lambda) + R(\lambda)G^*(R(\lambda)\Phi(s,\lambda)) + R(\lambda)\,(y(s)-y_d(s))\,ds - \int_t^T \Phi(s,\lambda)\,dW(s).$$

Proof. The claim follows from Theorem 2.35, Theorem 3.112 and Lemma 5.13. □

We have the following convergence results.

Lemma 5.20. *Let $(z^*(t),\Phi(t))_{t\in[0,T]}$ and $(z^*(t,\lambda),\Phi(t,\lambda))_{t\in[0,T]}$ be the mild solutions of system (5.14) and system (5.23), respectively. Then we have*

$$\lim_{\lambda\to\infty}\sup_{t\in[0,T]} \mathbb{E}\,\|z^*(t)-z^*(t,\lambda)\|_H^2 = 0, \qquad \lim_{\lambda\to\infty}\mathbb{E}\int_0^T \|\Phi(t)-\Phi(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dt = 0.$$

Proof. Let I be the identity operator on H. By definition, we have for all $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$
\begin{aligned}
z^*(t) - z^*(t,\lambda) &= \int_t^T e^{-A(s-t)}[G^*(\Phi(s)) - R(\lambda)G^*(R(\lambda)\Phi(s,\lambda))]\,ds \\
&\quad + \int_t^T e^{-A(s-t)}[I - R(\lambda)]\,(y(s) - y_d(s))\,ds - \int_t^T e^{-A(s-t)}[\Phi(s) - \Phi(s,\lambda)]\,dW(s). \qquad (5.24)
\end{aligned}
$$

Recall that the operator $G^*\colon \mathcal{L}_{(HS)}(Q^{1/2}(H);H) \to H$ is linear and bounded. Hence, we get for all $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$
\begin{aligned}
z^*(t) - z^*(t,\lambda) &= \int_t^T e^{-A(s-t)}G^*([I - R(\lambda)]\Phi(s))\,ds + \int_t^T e^{-A(s-t)}[I - R(\lambda)]G^*(R(\lambda)\Phi(s))\,ds \\
&\quad + \int_t^T e^{-A(s-t)}R(\lambda)G^*(R(\lambda)[\Phi(s) - \Phi(s,\lambda)])\,ds + \int_t^T e^{-A(s-t)}[I - R(\lambda)]\,(y(s) - y_d(s))\,ds \\
&\quad - \int_t^T e^{-A(s-t)}[\Phi(s) - \Phi(s,\lambda)]\,dW(s).
\end{aligned}
$$

Note that the assumptions of Lemma 3.92 are fulfilled. Thus, inequalities (3.27) and (3.28) hold. Let $T_1 \in [0,T)$. We obtain for all $\lambda > 0$

$$
\sup_{t\in[T_1,T]} \mathbb{E}\,\|z^*(t) - z^*(t,\lambda)\|_H^2 \le 4c(T - T_1)\,[\mathcal{I}_1(\lambda) + \mathcal{I}_2(\lambda)]\,, \qquad (5.25)
$$

$$
\mathbb{E}\int_{T_1}^T \|\Phi(t) - \Phi(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)}dt \le 4c(T - T_1)\,[\mathcal{I}_1(\lambda) + \mathcal{I}_2(\lambda)]\,, \qquad (5.26)
$$

where

$$
\begin{aligned}
\mathcal{I}_1(\lambda) &= \mathbb{E}\int_{T_1}^T \left[\|G^*([I - R(\lambda)]\Phi(t))\|_H^2 + \|[I - R(\lambda)]G^*(R(\lambda)\Phi(t))\|_H^2 + \|[I - R(\lambda)]\,(y(t) - y_d(t))\,\|_H^2\right] dt, \\
\mathcal{I}_2(\lambda) &= \mathbb{E}\int_{T_1}^T \|R(\lambda)G^*(R(\lambda)[\Phi(t) - \Phi(t,\lambda)])\|_H^2 dt.
\end{aligned}
$$

Using Lemma 5.13 (iii) and Proposition B.7, we can conclude

$$
\lim_{\lambda\to\infty} \mathcal{I}_1(\lambda) = 0. \qquad (5.27)
$$

By Lemma 5.13 (ii), there exists a constant $C^* > 0$ such that for all $\lambda > 0$

$$
\mathcal{I}_2(\lambda) \le C^*\,\mathbb{E}\int_{T_1}^T \|\Phi(t) - \Phi(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)}dt. \qquad (5.28)
$$

Due to inequality (5.26) and inequality (5.28), we get for all $\lambda > 0$

$$\begin{aligned}&\mathbb{E}\int_{T_1}^{T} \|\Phi(t) - \Phi(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dt \\ &\le 4c(T-T_1)\,\mathcal{I}_1(\lambda) + 4c\,C^*(T-T_1)\,\mathbb{E}\int_{T_1}^{T} \|\Phi(t) - \Phi(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dt.\end{aligned}$$

We chose $T_1 \in [0,T)$ such that $4c\,C^*(T-T_1) < 1$. Thus, we have for all $\lambda > 0$

$$\mathbb{E}\int_{T_1}^{T} \|\Phi(t) - \Phi(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dt \le \frac{4c(T-T_1)\,\mathcal{I}_1(\lambda)}{1-4c\,C^*(T-T_1)}.$$

Due to equation (5.27), we can conclude

$$\lim_{\lambda\to\infty} \mathbb{E}\int_{T_1}^{T} \|\Phi(t) - \Phi(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dt = 0. \tag{5.29}$$

Using inequality (5.25), inequality (5.28), equation (5.27) and equation (5.29), we have

$$\lim_{\lambda\to\infty} \sup_{t\in[T_1,T]} \mathbb{E}\,\|z^*(t) - z^*(t,\lambda)\|^2_H = 0.$$

By equation (5.24), we get for all $\lambda > 0$, all $t \in [0,T_1]$ and $\mathbb{P}$-a.s.

$$\begin{aligned}z^*(t) - z^*(t,\lambda) &= e^{-A(T_1-t)}[z^*(T_1) - z^*(T_1,\lambda)] + \int_t^{T_1} e^{-A(s-t)}[G^*(\Phi(s)) - R(\lambda)G^*(R(\lambda)\Phi(s,\lambda))]\,ds \\ &\quad + \int_t^{T_1} e^{-A(s-t)}[I-R(\lambda)]\,(y(s)-y_d(s))\,ds - \int_t^{T_1} e^{-A(s-t)}[\Phi(s) - \Phi(s,\lambda)]\,dW(s).\end{aligned}$$

Again, we find $T_2 \in [0,T_1]$ such that

$$\lim_{\lambda\to\infty} \sup_{t\in[T_2,T_1]} \mathbb{E}\,\|z^*(t) - z^*(t,\lambda)\|^2_H dt = 0, \qquad \lim_{\lambda\to\infty} \mathbb{E}\int_{T_2}^{T_1} \|\Phi(t) - \Phi(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dt = 0.$$

By continuing the method, we obtain the result. □

5.6. Design of the Optimal Controls

Based on the results provided in the previous sections, we are able to show a duality principle. Since we formulated a control problem with simultaneous distributed controls and boundary controls, we obtain two equations. The first equation gives us a relation between the mild solution of system (5.8) and the mild solution of the adjoint equation (5.14). The second equation provides a relation between the mild solution of system (5.9) and the mild solution of the adjoint equation (5.14).

Theorem 5.21. *Let $(y(t;u,v))_{t\in[0,T]}$ and $(z^*(t;u,v),\Phi(t;u,v))_{t\in[0,T]}$ be the mild solutions of system (5.4) and system (5.14) corresponding to the distributed control $u \in U$ and the boundary control $v \in V$, respectively. Moreover, let $(z_1(t;\tilde{u}))_{t\in[0,T]}$ and $(z_2(t;\tilde{v}))_{t\in[0,T]}$ be the mild solutions of system (5.8) and system (5.9) corresponding to the controls $\tilde{u} \in U$ and $\tilde{v} \in V$, respectively. Then we have for all $\alpha \in (0, \frac{1}{4})$*

$$\mathbb{E}\int_0^T \langle y(t;u,v) - y_d(t), z_1(t;\tilde{u})\rangle_H \, dt = \mathbb{E}\int_0^T \langle z^*(t;u,v), B\tilde{u}(t)\rangle_H \, dt, \tag{5.30}$$

$$\mathbb{E}\int_0^T \langle y(t;u,v) - y_d(t), z_2(t;\tilde{v})\rangle_H \, dt = \mathbb{E}\int_0^T \left\langle A^{1-\alpha} z^*(t;u,v), A^{\alpha} D\tilde{v}(t)\right\rangle_H dt. \tag{5.31}$$

Proof. For the sake of simplicity, we omit the dependence on the controls. First, we prove the result for the approximations derived in Section 5.5. Let $(z_1(t,\lambda))_{t\in[0,T]}$ and $(z_2(t,\lambda))_{t\in[0,T]}$ be the mild solutions of system (5.17) and system (5.18), respectively. Using Lemma 5.16, we have for all $\lambda > 0$, $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$z_1(t,\lambda) = \int_0^t (-A)z_1(s,\lambda) + R(\lambda)B\tilde{u}(s) \, ds + \int_0^t R(\lambda)G(R(\lambda)z_1(s,\lambda)) \, dW(s), \tag{5.32}$$

$$z_2(t,\lambda) = \int_0^t (-A)z_2(s,\lambda) + AR(\lambda)D\tilde{v}(s) \, ds + \int_0^t R(\lambda)G(R(\lambda)z_2(s,\lambda)) \, dW(s). \tag{5.33}$$

Next, let the pair of stochastic processes $(z^*(t,\lambda),\Phi(t,\lambda))_{t\in[0,T]}$ be the mild solution of system (5.23). Due to Lemma 5.19, we get for all $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$z^*(t,\lambda) = \int_t^T (-A)z^*(s,\lambda) + R(\lambda)G^*(R(\lambda)\Phi(s,\lambda)) + R(\lambda)\left(y(s) - y_d(s)\right) ds - \int_t^T \Phi(s,\lambda) \, dW(s). \tag{5.34}$$

By definition, the process $(z^*(t,\lambda))_{t\in[0,T]}$ is predictable. Using Proposition 3.16, we have for all $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} z^*(t,\lambda) = \mathbb{E}&\left[\int_0^T (-A)z^*(s,\lambda) + R(\lambda)G^*(R(\lambda)\Phi(s,\lambda)) + R(\lambda)\left(y(s) - y_d(s)\right) ds \,\middle|\, \mathcal{F}_t\right] \\ &- \int_0^t (-A)z^*(s,\lambda) + R(\lambda)G^*(R(\lambda)\Phi(s,\lambda)) + R(\lambda)\left(y(s) - y_d(s)\right) ds. \end{aligned}$$

Due to the martingale representation theorem given by Theorem 3.86 with $(M(t))_{t\in[0,T]}$ satisfying for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$M(t) = \mathbb{E}\left[\int_0^T (-A)z^*(s,\lambda) + R(\lambda)G^*(R(\lambda)\Phi(s,\lambda)) + R(\lambda)\left(y(s) - y_d(s)\right) ds \,\middle|\, \mathcal{F}_t\right],$$

there exists a unique predictable process $(\Psi(t,\lambda))_{t\in[0,T]}$ with values in $\mathcal{L}_{(HS)}(Q^{1/2}(H);H)$ such that for all

$\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$z^*(t,\lambda) = \mathbb{E}\left[\int_0^T (-A)z^*(s,\lambda) + R(\lambda)G^*(R(\lambda)\Phi(s,\lambda)) + R(\lambda)\left(y(s) - y_d(s)\right) ds\right]$$
$$- \int_0^t (-A)z^*(s,\lambda) + R(\lambda)G^*(R(\lambda)\Phi(s,\lambda)) + R(\lambda)\left(y(s) - y_d(s)\right) ds + \int_0^t \Psi(s,\lambda)\, dW(s). \quad (5.35)$$

Since the pair $(z^*(t,\lambda), \Phi(t,\lambda))_{t\in[0,T]}$ satisfies equation (5.34) uniquely, we can conclude $\Psi(t,\lambda) = \Phi(t,\lambda)$ for all $\lambda > 0$, almost all $t \in [0,T]$ and $\mathbb{P}$-almost surely. Applying the Itô product formula given by Corollary 3.69 to equation (5.32) and equation (5.35), we get for all $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\langle z_1(t,\lambda), z^*(t,\lambda)\rangle_H = \mathcal{I}_1(t,\lambda) + \mathcal{I}_2(t,\lambda) + \mathcal{I}_3(t,\lambda) + \mathcal{I}_4(t,\lambda),$$

where

$$\mathcal{I}_1(t,\lambda) = \int_0^t \left[\langle z_1(s,\lambda), Az^*(s,\lambda)\rangle_H - \langle z^*(s,\lambda), Az_1(s,\lambda)\rangle_H\right] ds,$$

$$\mathcal{I}_2(t,\lambda) = \int_0^t \left[\langle R(\lambda)G(R(\lambda)z_1(s,\lambda)), \Phi(s,\lambda)\rangle_{\mathcal{L}_{(HS)}(Q^{1/2}(H),H)} - \langle z_1(s,\lambda), R(\lambda)G^*(R(\lambda)\Phi(s,\lambda))\rangle_H\right] ds,$$

$$\mathcal{I}_3(t,\lambda) = \int_0^t \langle z^*(s,\lambda), R(\lambda)B\tilde{u}(s)\rangle_H\, ds - \int_0^t \langle z_1(s,\lambda), R(\lambda)\left(y(s) - y_d(s)\right)\rangle_H\, ds,$$

$$\mathcal{I}_4(t,\lambda) = \int_0^t \langle z_1(s,\lambda), \Phi(s,\lambda)\, dW(s)\rangle_H + \int_0^t \langle z^*(s,\lambda), R(\lambda)G(R(\lambda)z_1(s,\lambda))\, dW(s)\rangle_H.$$

By definition, we have $z^*(T,\lambda) = 0$ for all $\lambda > 0$ and $\mathbb{P}$-almost surely. Hence, we obtain for all $\lambda > 0$ and $\mathbb{P}$-a.s.

$$0 = \mathcal{I}_1(T,\lambda) + \mathcal{I}_2(T,\lambda) + \mathcal{I}_3(T,\lambda) + \mathcal{I}_4(T,\lambda). \quad (5.36)$$

Since the operator A is self-adjoint, we have for all $\lambda > 0$ and $\mathbb{P}$-a.s.

$$\mathcal{I}_1(T,\lambda) = 0. \quad (5.37)$$

Using Lemma 5.13 (v) and equation (5.15), we obtain for all $\lambda > 0$ and $\mathbb{P}$-a.s.

$$\mathcal{I}_2(T,\lambda) = 0. \quad (5.38)$$

By equations (5.36) – (5.38) and $\mathbb{E}\,\mathcal{I}_4(T,\lambda) = 0$ for all $\lambda > 0$, we get for all $\lambda > 0$

$$0 = \mathbb{E}\,\mathcal{I}_3(T,\lambda).$$

Hence, we have for all $\lambda > 0$

$$\mathbb{E}\int_0^T \langle R(\lambda)z_1(t,\lambda), y(t) - y_d(t)\rangle_H\, dt = \mathbb{E}\int_0^T \langle R(\lambda)z^*(t,\lambda), B\tilde{u}(t)\rangle_H\, dt. \quad (5.39)$$

Next, we show that the left hand side and the right hand side of equation (5.39) converges as $\lambda \to \infty$. By the Cauchy-Schwarz inequality and Lemma 5.13 (ii), we have for all $\lambda > 0$

$$\begin{aligned}
&\left| \mathbb{E} \int_0^T \langle z_1(t), y(t) - y_d(t) \rangle_H \, dt - \mathbb{E} \int_0^T \langle R(\lambda) z_1(t,\lambda), y(t) - y_d(t) \rangle_H \, dt \right|^2 \\
&\leq 2 \left| \mathbb{E} \int_0^T \langle [I - R(\lambda)] z_1(t), y(t) - y_d(t) \rangle_H \, dt \right|^2 + 2 \left| \mathbb{E} \int_0^T \langle R(\lambda)(z_1(t) - z_1(t,\lambda)), y(t) - y_d(t) \rangle_H \, dt \right|^2 \\
&\leq 4 \left(\mathbb{E} \int_0^T \|y(t)\|_H^2 \, dt + \int_0^T \|y_d(t)\|_H^2 \, dt \right) \left(\mathbb{E} \int_0^T \|[I - R(\lambda)] z_1(t)\|_H^2 \, dt + \mathbb{E} \int_0^T \|z_1(t) - z_1(t,\lambda)\|_H^2 \, dt \right).
\end{aligned}$$

Using Lemma 5.13 (iii), Proposition B.7 and Lemma 5.17, we can conclude

$$\lim_{\lambda \to \infty} \mathbb{E} \int_0^T \langle R(\lambda) z_1(t,\lambda), y(t) - y_d(t) \rangle_H \, dt = \mathbb{E} \int_0^T \langle z_1(t), y(t) - y_d(t) \rangle_H \, dt. \tag{5.40}$$

Recall that the operator $B \colon H \to H$ is bounded. Similarly as above, there exists a constant $C^* > 0$ such that for all $\lambda > 0$

$$\begin{aligned}
&\left| \mathbb{E} \int_0^T \langle z^*(t), B\tilde{u}(t) \rangle_H \, dt - \mathbb{E} \int_0^T \langle R(\lambda) z^*(t,\lambda), B\tilde{u}(t) \rangle_H \, dt \right|^2 \\
&\leq 2 \left| \mathbb{E} \int_0^T \langle [I - R(\lambda)] z^*(t), B\tilde{u}(t) \rangle_H \, dt \right|^2 + 2 \left| \mathbb{E} \int_0^T \langle R(\lambda)(z^*(t) - z^*(t,\lambda)), B\tilde{u}(t) \rangle_H \, dt \right|^2 \\
&\leq C^* \left(\mathbb{E} \int_0^T \|\tilde{u}(t)\|_H^2 \, dt \right) \left(\mathbb{E} \int_0^T \|[I - R(\lambda)] z^*(t)\|_H^2 \, dt + \sup_{t \in [0,T]} \mathbb{E} \, \|z^*(t) - z^*(t,\lambda)\|_H^2 \right).
\end{aligned}$$

By Lemma 5.13 (iii), Proposition B.7 and Lemma 5.20, we can infer

$$\lim_{\lambda \to \infty} \mathbb{E} \int_0^T \langle R(\lambda) z^*(t,\lambda), B\tilde{u}(t) \rangle_H \, dt = \mathbb{E} \int_0^T \langle z^*(t), B\tilde{u}(t) \rangle_H \, dt.$$

We conclude that the left hand side and the right hand side of equation (5.39) converges as $\lambda \to \infty$ and equation (5.30) holds.

Next, we show that equation (5.31) holds. Again, we apply Corollary 3.69 to equation (5.33) and equation (5.35). Similarly to equation (5.39), we find for all $\lambda > 0$ and all $\alpha \in (0, \frac{1}{4})$

$$\mathbb{E} \int_0^T \langle R(\lambda) z_2(t,\lambda), y(t) - y_d(t) \rangle_H \, dt = \mathbb{E} \int_0^T \left\langle R(\lambda) A^{1-\alpha} z^*(t,\lambda), A^\alpha D\tilde{v}(t) \right\rangle_H dt. \tag{5.41}$$

Similarly to equation (5.40), we can conclude

$$\lim_{\lambda \to \infty} \mathbb{E} \int_0^T \langle R(\lambda) z_2(t,\lambda), y(t) - y_d(t) \rangle_H \, dt = \mathbb{E} \int_0^T \langle z_2(t), y(t) - y_d(t) \rangle_H \, dt.$$

Recall that the operator $A^\alpha D\colon V^0(\partial\mathcal{D}) \to H$ is bounded for all $\alpha \in (0, \frac{1}{4})$. Hence, the stochastic process $(A^\alpha D\tilde{v}(t))_{t\in[0,T]}$ takes values in H such that $\mathbb{E}\int_0^T \|A^\alpha D\tilde{v}(t)\|_H^2 dt < \infty$. Since $D(A^{1-\alpha})$ is dense in H, there exists a sequence of processes $(v_m(t))_{t\in[0,T]}$, $m \in \mathbb{N}$, taking values in $D(A^{1-\alpha})$ such that $\mathbb{E}\int_0^T \|v_m(t)\|^2_{D(A^{1-\alpha})} dt < \infty$ for each $m \in \mathbb{N}$ and

$$\lim_{m\to\infty} \mathbb{E}\int_0^T \|A^\alpha D\tilde{v}(t) - v_m(t)\|_H^2 dt = 0.$$

Due to Proposition 5.12, the process $(z^*(t))_{t\in[0,T]}$ takes values in $D(A^{1-\alpha})$ for all $\alpha \in (0, \frac{1}{4})$. By Lemma 5.13 (ii) and (iv), Lemma 2.34, the Cauchy-Schwarz inequality and Fubini's theorem, there exists a constant $C^* > 0$ such that for all $\lambda > 0$, all $\alpha \in (0, \frac{1}{4})$ and each $m \in \mathbb{N}$

$$\begin{aligned}
&\left|\mathbb{E}\int_0^T \left\langle A^{1-\alpha} z^*(t), v_m(t)\right\rangle_H dt - \mathbb{E}\int_0^T \left\langle R(\lambda)A^{1-\alpha} z^*(t,\lambda), v_m(t)\right\rangle_H dt\right|^2 \\
&\leq 2\left|\mathbb{E}\int_0^T \left\langle [I - R(\lambda)]z^*(t), A^{1-\alpha} v_m(t)\right\rangle_H dt\right|^2 + 2\left|\mathbb{E}\int_0^T \left\langle R(\lambda)(z^*(t) - z^*(t,\lambda)), A^{1-\alpha} v_m(t)\right\rangle_H dt\right|^2 \\
&\leq C^* \left(\mathbb{E}\int_0^T \|v_m(t)\|^2_{D(A^{1-\alpha})} dt\right)\left(\mathbb{E}\int_0^T \|[I - R(\lambda)]z^*(t)\|_H^2 dt + \sup_{t\in[0,T]} \mathbb{E}\,\|z^*(t) - z^*(t,\lambda)\|_H^2\right).
\end{aligned}$$

Using Lemma 5.13 (iii), Proposition B.7 and Lemma 5.20, we can infer for each $m \in \mathbb{N}$

$$\lim_{\lambda\to\infty} \mathbb{E}\int_0^T \left\langle R(\lambda)A^{1-\alpha} z^*(t,\lambda), v_m(t)\right\rangle_H dt = \mathbb{E}\int_0^T \left\langle A^{1-\alpha} z^*(t), v_m(t)\right\rangle_H dt.$$

Due to the Moore-Osgood theorem [81, Theorem 7.11], we get

$$\begin{aligned}
\lim_{\lambda\to\infty} \mathbb{E}\int_0^T \left\langle R(\lambda)A^{1-\alpha} z^*(t,\lambda), A^\alpha D\tilde{v}(t)\right\rangle_H dt &= \lim_{\lambda\to\infty}\lim_{m\to\infty} \mathbb{E}\int_0^T \left\langle R(\lambda)A^{1-\alpha} z^*(t,\lambda), v_m(t)\right\rangle_H dt \\
&= \lim_{m\to\infty}\lim_{\lambda\to\infty} \mathbb{E}\int_0^T \left\langle R(\lambda)A^{1-\alpha} z^*(t,\lambda), v_m(t)\right\rangle_H dt \\
&= \mathbb{E}\int_0^T \left\langle A^{1-\alpha} z^*(t), A^\alpha D\tilde{v}(t)\right\rangle_H dt.
\end{aligned}$$

We conclude that the left hand side and the right hand side of equation (5.41) converges as $\lambda \to \infty$ and equation (5.31) holds. ☐

Based on the optimality conditions given by equation (5.12) and equation (5.13), we deduce formulas for the optimal controls using the duality principle derived in the previous theorem.

Theorem 5.22. *Let $(z^*(t;u,v),\Phi(t;u,v))_{t\in[0,T]}$ be the mild solution of system (5.14) corresponding to the controls $u\in U$ and $v\in V$. Then the optimal controls $\overline{u}\in U$ and $\overline{v}\in V$ satisfy for all $\alpha\in(0,\frac{1}{4})$, almost all $t\in[0,T]$ and $\mathbb{P}$-a.s.*

$$\overline{u}(t) = -\frac{1}{\kappa_1}\,B^* z^*(t;\overline{u},\overline{v}), \tag{5.42}$$

$$\overline{v}(t) = -\frac{1}{\kappa_2}\,K^* A^{1-\alpha} z^*(t;\overline{u},\overline{v}), \tag{5.43}$$

where $B^*\in\mathcal{L}(H)$ and $K^*\in\mathcal{L}(H;V^0(\partial\mathcal{D}))$ are the adjoint operators of the operators $B\in\mathcal{L}(H)$ and $K=A^\alpha D\in\mathcal{L}(V^0(\partial\mathcal{D});H)$, respectively.

Proof. Let $(y(t;u,v))_{t\in[0,T]}$ and $(z_1(t;u))_{t\in[0,T]}$ be the mild solutions of system (5.4) and system (5.8) corresponding to the controls $u\in U$ and $v\in V$, respectively. Using equation (5.10) and equation (5.12), the optimal control $\overline{u}\in U$ satisfies for every $\tilde{u}\in U$

$$\mathbb{E}\int_0^T \langle y(t;\overline{u},\overline{v})-y_d(t),z_1(t;\tilde{u})\rangle_H\,dt+\kappa_1\,\mathbb{E}\int_0^T\langle\overline{u}(t),\tilde{u}(t)\rangle_H\,dt=0.$$

By equation (5.30), we obtain for every $\tilde{u}\in U$

$$\mathbb{E}\int_0^T \langle z^*(t;\overline{u},\overline{v}),B\tilde{u}(t)\rangle_H\,dt+\kappa_1\,\mathbb{E}\int_0^T\langle\overline{u}(t),\tilde{u}(t)\rangle_H\,dt=0.$$

Hence, we get for every $\tilde{u}\in U$

$$\mathbb{E}\int_0^T \langle B^*z^*(t;\overline{u},\overline{v})+\kappa_1\,\overline{u}(t),\tilde{u}(t)\rangle_H\,dt=0.$$

Therefore, the optimal control $\overline{u}\in U$ satisfies equation (5.42) for almost all $t\in[0,T]$ and $\mathbb{P}$-almost surely.

Let $(z_2(t;v))_{t\in[0,T]}$ be the mild solution of system (5.9) corresponding to the control $v\in V$. Due to equation (5.11) and equation (5.13), the optimal control $\overline{v}\in V$ fulfills the following equation for every $\tilde{v}\in V$:

$$\mathbb{E}\int_0^T \langle y(t;\overline{u},\overline{v})-y_d(t),z_2(t;\tilde{v})\rangle_H\,dt+\kappa_2\,\mathbb{E}\int_0^T\langle\overline{v}(t),\tilde{v}(t)\rangle_{V^0(\partial\mathcal{D})}\,dt=0.$$

By equation (5.31), we have for all $\alpha\in(0,\frac{1}{4})$ and every $\tilde{v}\in V$

$$\mathbb{E}\int_0^T \left\langle A^{1-\alpha}z^*(t;\overline{u},\overline{v}),A^\alpha D\tilde{v}(t)\right\rangle_H\,dt+\kappa_2\,\mathbb{E}\int_0^T\langle\overline{v}(t),\tilde{v}(t)\rangle_{V^0(\partial\mathcal{D})}\,dt=0.$$

Hence, we get for all $\alpha\in(0,\frac{1}{4})$ and every $\tilde{v}\in V$

$$\mathbb{E}\int_0^T \left\langle K^*A^{1-\alpha}z^*(t;\overline{u},\overline{v})+\kappa_2\,\overline{v}(t),\tilde{v}(t)\right\rangle_{V^0(\partial\mathcal{D})}\,dt=0.$$

Therefore, the optimal control $\overline{v}\in V$ satisfies equation (5.43) for all $\alpha\in(0,\frac{1}{4})$, almost all $t\in[0,T]$ and $\mathbb{P}$-almost surely. □

Remark 5.23. *Let us denote by $(\overline{y}(t))_{t\in[0,T]}$ and $(\overline{z}^*(t), \overline{\Phi}(t))_{t\in[0,T]}$ the mild solutions of system (5.4) and system (5.14) corresponding to the optimal controls $\overline{u} \in U$ and $\overline{v} \in V$, respectively. As a consequence of the previous theorem, the optimal velocity field $(\overline{y}(t))_{t\in[0,T]}$ can be computed by solving the stochastic boundary value problem imposed by the following system of coupled forward-backward SPDEs:*

$$\begin{cases} d\,\overline{y}(t) = \left[-A\overline{y}(t) - \dfrac{1}{\kappa_1} BB^*\overline{z}^*(t) - \dfrac{1}{\kappa_2} ADK^*A^{1-\alpha}\overline{z}^*(t)\right] dt + G(\overline{y}(t))\, dW(t), \\ d\,\overline{z}^*(t) = -\left[-A\overline{z}^*(t) + G^*\left(\overline{\Phi}(t)\right) + \overline{y}(t) - y_d(t)\right] dt + \overline{\Phi}(t)\, dW(t), \\ \quad \overline{y}(0) = \xi, \quad \overline{z}^*(T) = 0. \end{cases}$$

Remark 5.24. *If system (4.3) is driven by a Lévy process, then one can obtain the optimal controls stated in the previous theorem as follows:*
We assume that system (4.3) is driven by an additive Lévy noise, i.e. the Hilbert-Schmidt operator G does not depend on the velocity field. Hence, the partial Fréchet derivatives are given by system (5.8) and system (5.9), whereby the diffusion term vanishes. Furthermore, the adjoint equation (5.14) has a deterministic structure in the sense that $\Phi(t) = 0$. The duality principle stated in Theorem 5.21 is then a consequence of a suitable product formula. The derivation of the optimal controls follows immediately from the previous theorem.

Chapter 6

Optimal Control of Uncertain Fluid Flows

In this chapter, we consider a control problem constrained by the unsteady stochastic Navier-Stokes equations with homogeneous Dirichlet boundary conditions. Motivated by [44], we first analyze the deterministic Navier-Stokes equations with homogeneous Dirichlet boundary conditions, which are usually considered as no-slip boundary conditions. Similarly to the previous chapters, we reformulate these equations as an evolution equation in a suitable Hilbert space such that the existence and uniqueness of a solution can be obtained using fractional powers of closed operators introduced in Section 2.3. Based on this approach, we extend the Navier-Stokes equations by an additional noise term. Due to the properties of a bilinear operator related to the convection term arising in the Navier-Stokes equations, we get a restriction on the dimension of the domain. However, we will figure out that an existence and uniqueness result of a local mild solution to the stochastic Navier-Stokes equations holds especially in two-dimensional as well as three-dimensional domains. The control problem considered here is motivated by common control strategies such as tracking a desired velocity field or minimizing the enstrophy, see [22, 46, 50, 59, 68, 83, 87]. We provide the existence and uniqueness of the optimal controls, which enables us to solve uniquely the control problem through a stochastic maximum principle. Since the control problem is formulated as a nonconvex optimization problem, we only obtain a necessary optimality condition as a variational inequality. However, we still can utilize this necessary optimality condition using a duality principle to derive the optimal controls. As a consequence, it remains to solve a coupled system of forward and backward SPDEs. Finally, we show that the optimal control satisfies a sufficient optimality condition. The results presented here are mainly based on [6, 7].

Throughout this chapter, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given complete probability space endowed with a normal filtration $(\mathcal{F}_t)_{t\geq 0}$.

6.1. Motivation

Throughout this chapter, let $\mathcal{D} \subset \mathbb{R}^n$, $n \geq 2$, be a connected and bounded domain with C^∞ boundary $\partial\mathcal{D}$. We consider the following Navier-Stokes equations with homogeneous Dirichlet boundary condition:

$$\begin{cases} \dfrac{\partial}{\partial t} y(t,x) + (y(t,x)\cdot\nabla)y(t,x) + \nabla p(t,x) - \nu\Delta y(t,x) = f(t,x) & \text{in } (0,T)\times\mathcal{D}, \\ \operatorname{div} y(t,x) = 0 & \text{in } (0,T)\times\mathcal{D}, \\ y(t,x) = 0 & \text{on } (0,T)\times\partial\mathcal{D}, \\ y(0,x) = \xi(x) & \text{in } \mathcal{D}, \end{cases} \tag{6.1}$$

where $y(t,x) \in \mathbb{R}^n$ denotes the velocity field with initial value $\xi(x) \in \mathbb{R}^n$ and $p(t,x) \in \mathbb{R}$ describes the pressure of the fluid. The parameter $\nu > 0$ is the viscosity parameter (for the sake of simplicity, we assume $\nu = 1$) and $f(t,x) \in \mathbb{R}^n$ is the external force.

Next, we reformulate system (6.1) as an evolution equation. For more details, we refer to [44]. According to Section 2.5.2, let us introduce the following Hilbert spaces:

$$H = \left\{ y \in (L^2(\mathcal{D}))^n \colon \operatorname{div} y = 0 \text{ in } \mathcal{D}, y\cdot\eta = 0 \text{ on } \partial\mathcal{D} \right\},$$

$$V = \left\{ y \in \left(H_0^1(\mathcal{D})\right)^n : \operatorname{div} y = 0 \text{ in } \mathcal{D} \right\}$$

and let $A\colon D(A) \subset H \to H$ be the Stokes operator given by

$$D(A) = (H^2(\mathcal{D}))^n \cap V, \quad Ay = -\Pi \Delta y$$

for every $y \in D(A)$, where the operator $\Pi\colon (L^2(\mathcal{D}))^n \to H$ is an orthogonal projection. By Theorem 2.61, the operator $-A$ is the generator of an analytic semigroup of contractions $(e^{-At})_{t\geq 0}$. Hence, we can introduce fractional powers of A denoted by A^α with $\alpha \in \mathbb{R}$ according to Section 2.3. Furthermore, we define $B(y,z) = \Pi(y \cdot \nabla)z$ for some $y, z \in H$. If $y = z$, we write $B(y) = B(y,y)$. Applying the projection Π, system (6.1) can be formulated in the following abstract form:

$$\begin{cases} \dfrac{d}{dt} y(t) = -Ay(t) - B(y(t)) + \Pi f(t), \\ \quad y(0) = \Pi \xi. \end{cases} \tag{6.2}$$

For the sake of simplicity, we assume $f(t), \xi \in H$ for all $t \in [0,T]$. We consider this equation in integral form

$$y(t) = e^{-At}\xi - \int_0^t e^{-A(t-s)} [B(y(s)) - f(s)] \, ds$$

for all $t \in [0,T]$. Moreover, we have the following properties of the nonlinear term in equation (6.2).

Lemma 6.1 (cf. Lemma 2.2,[44])**.** *Let $0 \leq \delta < \frac{1}{2} + \frac{n}{4}$. If $y \in D(A^{\alpha_1})$ and $z \in D(A^{\alpha_2})$, then we have*

$$\left\| A^{-\delta} B(y,z) \right\|_H \leq \widetilde{M} \left\| A^{\alpha_1} y \right\|_H \left\| A^{\alpha_2} z \right\|_H,$$

with some constant $\widetilde{M} = \widetilde{M}_{\delta,\alpha_1,\alpha_2}$, provided that $\alpha_1, \alpha_2 > 0$, $\delta + \alpha_2 > \frac{1}{2}$ and $\delta + \alpha_1 + \alpha_2 \geq \frac{n}{4} + \frac{1}{2}$.

Corollary 6.2. *Let α_1, α_2 and δ be as in Lemma 6.1. If $y, z \in D(A^\beta)$, $\beta = \max\{\alpha_1, \alpha_2\}$, then we have*

$$\left\| A^{-\delta}(B(y) - B(z)) \right\|_H \leq \widetilde{M} (\left\| A^{\alpha_1} y \right\|_H \left\| A^{\alpha_2}(y-z) \right\|_H + \left\| A^{\alpha_1}(y-z) \right\|_H \left\| A^{\alpha_2} z \right\|_H).$$

Proof. Using Lemma 6.1, we get

$$\begin{aligned} \left\| A^{-\delta}(B(y) - B(z)) \right\|_H &= \left\| A^{-\delta}(B(y, y-z) + B(y-z, z)) \right\|_H \\ &\leq \left\| A^{-\delta} B(y, y-z) \right\|_H + \left\| A^{-\delta} B(y-z, z) \right\|_H \\ &\leq \widetilde{M} (\left\| A^{\alpha_1} y \right\|_H \left\| A^{\alpha_2}(y-z) \right\|_H + \left\| A^{\alpha_1}(y-z) \right\|_H \left\| A^{\alpha_2} z \right\|_H). \end{aligned}$$

□

As a consequence, the bilinear operator B satisfies a growth condition and a Lipschitz condition only locally. Thus, we cannot prove the existence and uniqueness of a solution to system (6.2) over the whole time interval $[0,T]$ in general. However, we have the following local result.

Theorem 6.3. *Let $\alpha \in (0,1)$ and $\delta \in [0,1)$ be given parameters such that $1 > \delta + \alpha > \frac{1}{2}$ and $\delta + 2\alpha \geq \frac{n}{4} + \frac{1}{2}$. Furthermore, let the function $f\colon [0,T] \to D(A^\alpha)$ satisfy*

$$\int_0^T \| f(t) \|^2_{D(A^\beta)} \, dt < \infty$$

with $\beta \in [0, \alpha]$ such that $\alpha - \beta < \frac{1}{2}$. Then for any $\xi \in D(A^\alpha)$, there exists a unique continuous solution $y\colon [0, T] \to D(A^\alpha)$ of system (6.2) satisfying

$$\sup_{t \in [0,\tau]} \|y(t)\|_{D(A^\alpha)} < \infty$$

for a certain point of time $\tau \in [0, T]$.

Proof. The proof can be obtained similarly to [44, Theorem 2.3]. □

6.2. The Controlled Stochastic Navier-Stokes Equations

In this section, we introduce the controlled stochastic Navier-Stokes equation with homogeneous Dirichlet boundary conditions. We show that there exists a unique mild solution up to a certain stopping time and we state some basic properties. Let us assume that the external force $f(t)$ in equation (6.2) can be decomposed as the sum of a control term and a noise term dependent on the velocity field $y(t)$. Hence, we obtain the stochastic Navier-Stokes equations in $D(A^\alpha)$:

$$\begin{cases} dy(t) = -[Ay(t) + B(y(t)) - Fu(t)]dt + G(y(t))\, dW(t), \\ y(0) = \xi, \end{cases} \tag{6.3}$$

where $\xi \in L^2(\Omega; D(A^\alpha))$ is $\mathcal{F}_0$-measurable and $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in H and covariance operator $Q \in \mathcal{L}_1^+(H)$. We introduce the space $L^k_{\mathcal{F}}(\Omega; L^r([0, T]; D(A^\beta)))$ containing all $\mathcal{F}_t$-adapted processes $(u(t))_{t\in[0,T]}$ with values in $D(A^\beta)$ such that $\mathbb{E}(\int_0^T \|u(t)\|^r_{D(A^\beta)}\, dt)^{k/r} < \infty$ with $k, r \in [0, \infty)$ and $\beta \in [0, \alpha]$. The space $L^k_{\mathcal{F}}(\Omega; L^r([0, T]; D(A^\beta)))$ equipped with the norm

$$\|u\|^k_{L^k_{\mathcal{F}}(\Omega;L^r([0,T];D(A^\beta)))} = \mathbb{E}\left(\int_0^T \|u(t)\|^r_{D(A^\beta)}\, dt\right)^{k/r}$$

for every $u \in L^k_{\mathcal{F}}(\Omega; L^r([0, T]; D(A^\beta)))$ becomes a Banach space. The set of admissible controls U is a closed, bounded and convex subset of the Hilbert space $L^2_{\mathcal{F}}(\Omega; L^2([0, T]; D(A^\beta)))$ such that $0 \in U$. Moreover, we assume that the operators $F\colon D(A^\beta) \to D(A^\beta)$ and $G\colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$ are linear and bounded. In general, we can not ensure the existence and uniqueness of a mild solution over an arbitrary time interval $[0, T]$ since the nonlinear operator B satisfies a growth condition and a Lipschitz condition only locally, which is a consequence of Lemma 6.1 and Corollary 6.2. Thus, we need the following definition of a local mild solution.

Definition 6.4 (cf. Definition 3.2, [25])**.** *Let τ be a stopping time taking values in $(0, T]$ and $(\tau_m)_{m\in\mathbb{N}}$ be an increasing sequence of stopping times taking values in $[0, T]$ satisfying $\lim_{m\to\infty} \tau_m = \tau$. A predictable process $(y(t))_{t\in[0,\tau)}$ with values in $D(A^\alpha)$ is called a* ***local mild solution of system (6.3)*** *if for fixed $m \in \mathbb{N}$*

$$\mathbb{E} \sup_{t\in[0,\tau_m)} \|y(t)\|^2_{D(A^\alpha)} < \infty$$

and we have for each $m \in \mathbb{N}$, all $t \in [0, T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} y(t \wedge \tau_m) = e^{-A(t\wedge\tau_m)}\xi - \int_0^{t\wedge\tau_m} A^\delta e^{-A(t\wedge\tau_m - s)} A^{-\delta} B(y(s))\, ds + \int_0^{t\wedge\tau_m} e^{-A(t\wedge\tau_m-s)} Fu(s)\, ds \\ + I_{\tau_m}(G(y))(t \wedge \tau_m), \end{aligned}$$

where

$$I_{\tau_m}(G(y))(t) = \int_0^t \mathbb{1}_{[0,\tau_m)}(s) e^{-A(t-s)} G(y(s \wedge \tau_m))\, dW(s). \tag{6.4}$$

Remark 6.5. *Note that the stopped stochastic convolution $(I_{\tau_m}(G(y))(t \wedge \tau_m))_{t\in[0,T]}$ is well defined due to Lemma 3.66.*

The proof of the existence and uniqueness of a local mild solution to system (6.3) can be shown in two steps. First, we consider a modified system to get a mild solution well defined over the whole time interval $[0, T]$. Second, we introduce suitable stopping times such that the mild solution of the modified system and the local mild solution of system (6.3) coincides. We introduce the following auxiliary system in $D(A^\alpha)$:

$$\begin{cases} dy_m(t) = -[Ay_m(t) + B(\pi_m(y_m(t))) - Fu(t)]\, dt + G(y_m(t))\, dW(t), \\ y_m(0) = \xi, \end{cases} \tag{6.5}$$

where $m \in \mathbb{N}$ and $\pi_m \colon D(A^\alpha) \to D(A^\alpha)$ is defined by

$$\pi_m(y) = \begin{cases} y & \|y\|_{D(A^\alpha)} \le m, \\ m\|y\|^{-1}_{D(A^\alpha)} y & \|y\|_{D(A^\alpha)} > m. \end{cases} \tag{6.6}$$

Then we have for every $y, z \in D(A^\alpha)$

$$\|\pi_m(y)\|_{D(A^\alpha)} \le \min\{m, \|y\|_{D(A^\alpha)}\}, \tag{6.7}$$

$$\|\pi_m(y) - \pi_m(z)\|_{D(A^\alpha)} \le 2\|y - z\|_{D(A^\alpha)}. \tag{6.8}$$

We get the following existence and uniqueness result.

Theorem 6.6. *Let the parameters $\alpha \in (0, 1)$ and $\delta \in [0, 1)$ satisfy $1 > \delta + \alpha > \frac{1}{2}$ and $\delta + 2\alpha \ge \frac{n}{4} + \frac{1}{2}$. Furthermore, let $u \in L^2_{\mathcal{F}}(\Omega; L^2([0, T]; D(A^\beta)))$ be fixed for $\beta \in [0, \alpha]$ such that $\alpha - \beta < \frac{1}{2}$. Then for fixed $m \in \mathbb{N}$ and any $\xi \in L^2(\Omega; D(A^\alpha))$, there exists a unique mild solution $(y_m(t))_{t\in[0,T]}$ of system (6.5) in the sense of Definition 3.81. Moreover, the process $(y_m(t))_{t\in[0,T]}$ has a continuous modification.*

Proof. Using Lemma 6.1, Lemma 6.2, inequality (6.7) and inequality (6.8), we have for every $y, z \in D(A^\alpha)$

$$\|A^{-\delta} B(\pi_m(y))\|_H \le m\widetilde{M}\, \|y\|_{D(A^\alpha)},$$

$$\|A^{-\delta}[B(\pi_m(y)) - B(\pi_m(z))]\|_H \le 2m\widetilde{M}\, \|y - z\|_{D(A^\alpha)}.$$

Therefore, the assumptions of Theorem 3.82 hold and we get the existence and uniqueness of a mild solution to system (6.5). The fact that the process $(y_m(t))_{t\in[0,T]}$ has a continuous modification is a consequence of Remark 3.83. □

As a consequence of the previous theorem, the mild solution $(y_m(t))_{t\in[0,T]}$ of system (6.5) takes values in $D(A^\alpha)$ such that

$$\mathbb{E} \sup_{t\in[0,T]} \|y_m(t)\|^2_{D(A^\alpha)} < \infty$$

and we have for all $t \in [0, T]$ and $\mathbb{P}$-a.s.

$$y_m(t) = e^{-At}\xi - \int_0^t A^\delta e^{-A(t-s)} A^{-\delta} B(\pi_m(y_m(s)))\, ds + \int_0^t e^{-A(t-s)} Fu(s)\, ds + \int_0^t e^{-A(t-s)} G(y_m(s))\, dW(s). \tag{6.9}$$

Thus, we obtain the following existence and uniqueness result of a local mild solution to system (6.3).

Theorem 6.7. *Let the parameters $\alpha \in (0,1)$ and $\delta \in [0,1)$ satisfy $1 > \delta + \alpha > \frac{1}{2}$ and $\delta + 2\alpha \geq \frac{n}{4} + \frac{1}{2}$. Furthermore, let $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ be fixed for $\beta \in [0,\alpha]$ such that $\alpha - \beta < \frac{1}{2}$. Then for any $\xi \in L^2(\Omega; D(A^\alpha))$, there exists a unique local mild solution $(y(t))_{t\in[0,\tau)}$ of system (6.3). Moreover, the process $(y(t))_{t\in[0,\tau)}$ has a continuous modification.*

Proof. Due to Theorem 6.6, we get the existence and uniqueness of a mild solution $(y_m(t))_{t\in[0,T]}$ to system (6.5), which has a continuous modification. Next, we define a sequence of stopping times $(\tau_m)_{m\in\mathbb{N}}$ by

$$\tau_m = \inf\{t \in (0,T) : \|y_m(t)\|_{D(A^\alpha)} > m\} \wedge T \tag{6.10}$$

$\mathbb{P}$-a.s. with the usual convention that $\inf \emptyset = +\infty$. The fact that τ_m is a stopping time results from Remark 3.14. By definition of the map π_m, we get $\pi_m(y_m(t)) = y_m(t)$ for all $t \in [0,\tau_m)$ and $\mathbb{P}$-almost surely. Using equation (6.9) and Lemma 3.66, we obtain for fixed $m \in \mathbb{N}$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} y_m(t \wedge \tau_m) = e^{-A(t\wedge\tau_m)}\xi - \int_0^{t\wedge\tau_m} A^\delta e^{-A(t\wedge\tau_m - s)} A^{-\delta} B(y_m(s))\, ds + \int_0^{t\wedge\tau_m} e^{-A(t\wedge\tau_m - s)} F u(s)\, ds \\ + I_{\tau_m}(G(y))(t \wedge \tau_m), \end{aligned}$$

where $\mathcal{I}_{\tau_m}(G(y))(t)$ is given by equation (6.4). Since the sequence of stopping times $(\tau_m)_{m\in\mathbb{N}}$ is increasing and bounded, there exists a stopping time τ such that $\tau = \lim_{m\to\infty} \tau_m$ resulting from Lemma 3.11 (ii). Moreover, we have $\mathbb{P}$-a.s. $0 < \tau \leq T$. We set for each $m \in \mathbb{N}$, all $t \in [0,\tau_m)$ and $\mathbb{P}$-a.s.

$$y(t) = y_m(t). \tag{6.11}$$

Then the process $(y(t))_{t\in[0,\tau)}$ is the unique local mild solution of system (6.3). □

Remark 6.8. *(i) Note that the previous theorem is especially valid for $n = 2$ and $n = 3$. Hence, we get the existence and uniqueness of a solution to the stochastic Navier-Stokes equations for two-dimensional as well as three-dimensional domains up to a certain stopping time.*
(ii) In case of additive noise in system (6.3), i.e. $G(y) \equiv G$, we have

$$\mathbb{E} \sup_{t\in[0,\rho]} \|y(t)\|^2_{D(A^\alpha)} < \infty$$

for a certain stopping time ρ with values in $[0,T]$ and independent of $m \in \mathbb{N}$. The proof can be found in [10, 34].
(iii) If, in addition to the assumptions of Theorem 6.7, we require

$$\mathbb{E} \sup_{t\in[0,\tau)} \int_0^t (t-s)^{-n/4} |\nabla y(s)|\, ds < \infty,$$

then the solution of system (6.3) is a global mild solution in the sense that $\mathbb{P}(\tau = T) = 1$, see [25].

In the remaining part of this chapter, we always assume that the parameters $\alpha \in (0,1)$, $\delta \in [0,1)$ and $\beta \in [0,\alpha]$ satisfy the assumptions of Theorem 6.7 and the stopping times $(\tau_m)_{m\in\mathbb{N}}$ are given by equation (6.10). Moreover, we assume that the initial value $\xi \in L^2(\Omega; D(A^\alpha))$ is fixed. To illustrate the dependence on the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, let us denote by $(y_m(t;u))_{t\in[0,T]}$ and $(y(t;u))_{t\in[0,\tau^u)}$ the mild solution of system (6.5) and the local mild solution of system (6.3), respectively. Note that the stopping times $(\tau^u_m)_{m\in\mathbb{N}}$ depend on the control as well. Whenever these processes and the stopping times are considered for fixed control, we use the notation introduced above. We have the following continuity property.

Lemma 6.9. *For fixed $m \in \mathbb{N}$, let $(y_m(t;u))_{t\in[0,T]}$ be the mild solution of system (6.5) corresponding to the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. If $u_1, u_2 \in L^k_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ for $k \geq 2$, then there exists a constant $c > 0$ such that*

$$\mathbb{E} \sup_{t\in[0,T]} \|y_m(t;u_1) - y_m(t;u_2)\|^k_{D(A^\alpha)} \leq c \, \|u_1 - u_2\|^k_{L^k_{\mathcal{F}}(\Omega;L^2([0,T];D(A^\beta)))}.$$

Proof. Recall that the operators $F\colon D(A^\beta) \to D(A^\beta)$ and $G\colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$ are linear and bounded. By definition, we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} y_m(t;u_1) - y_m(t;u_2) = & - \int_0^t A^\delta e^{-A(t-s)} A^{-\delta} \left[B(\pi_m(y_m(s;u_1))) - B(\pi_m(y_m(s;u_2))) \right] ds \\ & + \int_0^t e^{-A(t-s)} F[u_1(s) - u_2(s)] \, ds \\ & + \int_0^t e^{-A(t-s)} G(y_m(s;u_1) - y_m(s;u_2)) \, dW(s). \end{aligned}$$

Let $T_{1,m} \in (0,T]$. Using Theorem 2.35, Corollary 6.2, inequalities (6.7) and (6.8), the Cauchy-Schwarz inequality and Proposition 3.65 (ii), there exist constants $C_1, C_2, C_3 > 0$ such that

$$\begin{aligned} & \mathbb{E} \sup_{t\in[0,T_{1,m}]} \|y_m(t;u_1) - y_m(t;u_2)\|^k_{D(A^\alpha)} \\ & \leq 3^{k-1} \, \mathbb{E} \sup_{t\in[0,T_{1,m}]} \left(\int_0^t \left\| A^{\alpha+\delta} e^{-A(t-s)} A^{-\delta} \left[B(\pi_m(y_m(s;u_1))) - B(\pi_m(y_m(s;u_2))) \right] \right\|_H ds \right)^k \\ & \quad + 3^{k-1} \, \mathbb{E} \sup_{t\in[0,T_{1,m}]} \left(\int_0^t \left\| A^{\alpha-\beta} e^{-A(t-s)} A^\beta F[u_1(s) - u_2(s)] \right\|_H ds \right)^k \\ & \quad + 3^{k-1} \, \mathbb{E} \sup_{t\in[0,T_{1,m}]} \left\| \int_0^t e^{-A(t-s)} A^\alpha G(y_m(s;u_1) - y_m(s;u_2)) \, dW(s) \right\|^k_H \\ & \leq \left(C_1 T_{1,m}^{k(1-\alpha-\delta)} + C_2 T_{1,m}^{k/2} \right) \mathbb{E} \sup_{t\in[0,T_{1,m}]} \|y_m(t;u_1) - y_m(t;u_2)\|^k_{D(A^\alpha)} \\ & \quad + C_3 \mathbb{E} \left(\int_0^T \|u_1(t) - u_2(t)\|^2_{D(A^\beta)} \, dt \right)^{k/2}. \end{aligned}$$

We chose $T_{1,m} \in (0,T]$ such that $C_1 T_{1,m}^{k(1-\alpha-\delta)} + C_2 T_{1,m}^{k/2} < 1$. Hence, we get

$$\mathbb{E} \sup_{t\in[0,T_{1,m}]} \|y_m(t;u_1) - y_m(t;u_2)\|^k_{D(A^\alpha)} \leq c_{1,m} \mathbb{E} \left(\int_0^T \|u_1(t) - u_2(t)\|^2_{D(A^\beta)} \, dt \right)^{k/2},$$

where $c_{1,m} = \frac{C_3}{1-C_1 T_{1,m}^{k(1-\alpha-\delta)} - C_2 T_{1,m}^{k/2}}$. Next, we consider for all $t \in [T_{1,m}, T]$, $\mathbb{P}$-a.s. and for $i = 1, 2$

$$y_m(t; u_i) = e^{-A(t-T_{1,m})} y_m(T_{1,m}; u_i) - \int_{T_{1,m}}^{t} A^{\delta} e^{-A(t-s)} A^{-\delta} B(\pi_m(y_m(s; u_i)))\, ds + \int_{T_{1,m}}^{t} e^{-A(t-s)} F u_i(s)\, ds$$
$$+ \int_{T_{1,m}}^{t} e^{-A(t-s)} G(y_m(s; u_i))\, dW(s).$$

Again, we find $T_{2,m} \in [T_{1,m}, T]$ and a constant $c_{2,m} > 0$ such that

$$\mathbb{E} \sup_{t \in [T_{1,m}, T_{2,m}]} \|y_m(t; u_1) - y_m(t; u_2)\|_{D(A^\alpha)}^k \leq c_{2,m} \mathbb{E} \left(\int_0^T \|u_1(t) - u_2(t)\|_{D(A^\beta)}^2\, dt \right)^{k/2}.$$

By continuing this method, we obtain the result. □

Remark 6.10. *By definition, we have for all $t \in [0, \tau_m^u)$ and $\mathbb{P}$-a.s. $y(t; u) = y_m(t; u)$. Hence, a similar result of the previous lemma holds for the local mild solution of system (6.3).*

In the following lemmas, we show some useful properties of the stopping times.

Lemma 6.11. *For fixed $m \in \mathbb{N}$, let $(y_m(t; u))_{t \in [0,T]}$ be the mild solution of system (6.5) corresponding to the control $u \in L_{\mathcal{F}}^2(\Omega; L^2([0, T]; D(A^\beta)))$ and let the stopping time τ_m^u be given by equation (6.10). Then we have*

$$\lim_{u_1 \to u_2} \mathbb{P}\left(\tau_m^{u_1} \neq \tau_m^{u_2}\right) = 0.$$

Proof. By the extended version of Markov's inequality and Lemma 6.9 with $k = 2$, we get for all $\varepsilon > 0$

$$\mathbb{P}\left(\sup_{t \in [0,T]} \|y_m(t; u_1) - y_m(t; u_2)\|_{D(A^\alpha)} \geq \varepsilon \right) \leq \frac{1}{\varepsilon^2} \mathbb{E} \sup_{t \in [0,T]} \|y_m(t; u_1) - y_m(t; u_2)\|_{D(A^\alpha)}^2$$
$$\leq \frac{c}{\varepsilon^2} \mathbb{E} \int_0^T \|u_1(t) - u_2(t)\|_{D(A^\beta)}^2\, dt. \tag{6.12}$$

Next, we assume $\lim_{u_1 \to u_2} \mathbb{P}\left(\tau_m^{u_1} < \tau_m^{u_2}\right) > 0$. Due to the definition of the stopping times, we can conclude

$$\lim_{u_1 \to u_2} \mathbb{P}\left(\{\|y_m(\tau_m^{u_1}; u_1)\|_{D(A^\alpha)} > m\} \cap \{\|y_m(\tau_m^{u_1}; u_2)\|_{D(A^\alpha)} \leq m\} \right) > 0.$$

Therefore, there exists $\varepsilon_0 > 0$ such that

$$\lim_{u_1 \to u_2} \mathbb{P}\left(\|y_m(\tau_m^{u_1}; u_1)\|_{D(A^\alpha)} - \|y_m(\tau_m^{u_1}; u_2)\|_{D(A^\alpha)} \geq \varepsilon_0 \right) > 0.$$

This implies that $\lim_{u_1 \to u_2} \mathbb{P}\left(\|y_m(\tau_m^{u_1}; u_1) - y_m(\tau_m^{u_1}; u_2)\|_{D(A^\alpha)} \geq \varepsilon_0 \right) > 0$, which is a contradiction to inequality (6.12). We get $\lim_{u_1 \to u_2} \mathbb{P}\left(\tau_m^{u_1} < \tau_m^{u_2}\right) = 0$. Similarly, we obtain $\lim_{u_1 \to u_2} \mathbb{P}\left(\tau_m^{u_1} > \tau_m^{u_2}\right) = 0$. □

Lemma 6.12. *For fixed $m \in \mathbb{N}$, let $(y_m(t; u))_{t \in [0,T]}$ be the mild solution of system (6.5) corresponding to the control $u \in L_{\mathcal{F}}^2(\Omega; L^2([0, T]; D(A^\beta)))$ and let the stopping time τ_m^u be given by equation (6.10). If $u_1, u_2 \in L_{\mathcal{F}}^{k+1}(\Omega; L^2([0, T]; D(A^\beta)))$ for $k \geq 1$, then*

$$\lim_{\theta \to 0} \frac{\mathbb{P}\left(\tau_m^{u_1} \neq \tau_m^{u_1 + \theta u_2}\right)}{\theta^k} = 0.$$

Proof. The claim follows similarly to Lemma 6.11. □

6.3. A Generalized Control Problem

In this section, we introduce the control problem. We provide the existence and uniqueness of an optimal control and we calculate the Gâteaux derivatives of the cost functional related to the control problem. This requires the Gâteaux derivative of the local mild solution to system (6.3), which is given by the local mild solution of the linearized stochastic Navier-Stokes equations. Moreover, we show that the Gâteaux derivatives of the cost functional up to order two coincide with the Fréchet derivatives up to order two. As a consequence, we can obtain necessary and sufficient optimality conditions. Let us introduce the cost functional $J_m\colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to \mathbb{R}$ given by

$$J_m(u) = \frac{1}{2}\,\mathbb{E}\int_0^{\tau_m^u} \|A^\gamma(y(t;u) - y_d(t))\|_H^2 \, dt + \frac{1}{2}\,\mathbb{E}\int_0^T \|A^\beta u(t)\|_H^2 dt, \tag{6.13}$$

where $m \in \mathbb{N}$ is fixed and $\gamma \in [0, \alpha]$. Moreover, the process $(y(t;u))_{t\in[0,\tau^u)}$ is the local mild solution of system (6.3) corresponding to the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ and $y_d \in L^2([0,T]; D(A^\gamma))$ is a given desired velocity field. The task is to find a control $\overline{u}_m \in U$ such that

$$J_m(\overline{u}_m) = \inf_{u\in U} J_m(u). \tag{6.14}$$

The control $\overline{u}_m \in U$ is called an optimal control. Note that for $\gamma = 0$, the formulation coincides with a tracking problem, for more details see [50, 59, 68, 87]. For $\gamma = \frac{1}{2}$ and $y_d \equiv 0$, we minimize the enstrophy, see [22, 46, 83]. Hence, we are dealing with a generalized cost functional, which incorporates common control problems in fluid dynamics.

Remark 6.13. *(i) Note that by definition of the local mild solution, we only can ensure that the first addend of the cost functional given by (6.13) is well defined up to the stopping time τ_m^u for fixed $m \in \mathbb{N}$.*
(ii) In case of additive noise in system (6.3), we can replace the stopping time τ_m^u in equation (6.13) by a certain stopping time ρ^u independent of $m \in \mathbb{N}$.
(iii) If the assumptions of Remark 6.8 (iii) are fulfilled, then we can replace the stopping time τ_m^u in equation (6.13) by the deterministic terminal point of time T.

6.3.1. Existence and Uniqueness of the Optimal Control

Since the velocity field as well as the stopping times are nonconvex with respect to the control, we formulated the control problem as a nonconvex optimization problem. To obtain the existence and uniqueness of the optimal control $\overline{u}_m \in U$, we show that Corollary D.18 can be applied. For that purpose, we first show the following continuity result.

Lemma 6.14. *Let $(y(t;u))_{t\in[0,\tau^u)}$ be the local mild solution of system (6.3) corresponding to the control $u \in U$, where the stopping times $(\tau_m^u)_{m\in\mathbb{N}}$ are defined by equation (6.10). Then for fixed $m \in \mathbb{N}$, the functional*

$$f_m(u) = \mathbb{E}\int_0^{\tau_m^u} \|A^\gamma(y(t;u) - y_d(t))\|_H^2 \, dt$$

is continuous with respect to the control $u \in U$.

Proof. Let the process $(y_m(t;u))_{t\in[0,T]}$ be the mild solution of system (6.5) corresponding to the control $u \in U$ and let $u_1, u_2 \in U$. We define the stopping times $\underline{\tau_m} = \tau_m^{u_1} \wedge \tau_m^{u_2}$ and $\overline{\tau_m} = \tau_m^{u_1} \vee \tau_m^{u_2}$. Moreover, let the control $\overline{u} \in U$ be given by

$$\overline{u} = \begin{cases} u_1 & \text{if } \overline{\tau_m} = \tau_m^{u_1}, \\ u_2 & \text{if } \overline{\tau_m} = \tau_m^{u_2}. \end{cases}$$

Using Corollary 2.32, equation (6.11) and the Cauchy-Schwarz inequality, there exists a constant $\widetilde{K} > 0$ such that

$$\begin{aligned}
&|f_m(u_1) - f_m(u_2)| \\
&= \left| \mathbb{E} \int_0^{\tau_m^{u_1}} \|A^\gamma(y_m(t;u_1) - y_d(t))\|_H^2 \, dt - \mathbb{E} \int_0^{\tau_m^{u_2}} \|A^\gamma(y_m(t;u_2) - y_d(t))\|_H^2 \, dt \right| \\
&\leq \mathbb{E} \int_0^{\underline{\tau_m}} \Big| \|A^\gamma(y_m(t;u_1) - y_d(t))\|_H^2 - \|A^\gamma(y_m(t;u_2) - y_d(t))\|_H^2 \Big| \, dt + \mathbb{E} \int_{\underline{\tau_m}}^{\overline{\tau_m}} \|A^\gamma(y_m(t;\overline{u}) - y_d(t))\|_H^2 \, dt \\
&\leq \widetilde{K} \left(\mathbb{E} \sup_{t\in[0,T]} \|y_m(t;u_1) - y_m(t;u_2)\|_{D(A^\alpha)}^2 \right)^{1/2} \\
&\quad + 2 \int_0^T \mathbb{P}(\tau_m^{u_1} \wedge \tau_m^{u_2} \leq t < \tau_m^{u_1} \vee \tau_m^{u_2}) \left(C^2 m^2 + \|y_d(t))\|_{D(A^\gamma)}^2 \right) dt.
\end{aligned}$$

Due to Lemma 6.9 with $k = 2$, we have $\lim_{u_1\to u_2} \mathbb{E} \sup_{t\in[0,T]} \|y_m(t;u_1) - y_m(t;u_2)\|_{D(A^\alpha)}^2 = 0$. By Lemma 6.11, we get $\lim_{u_1\to u_2} \mathbb{P}(\tau_m^{u_1} \wedge \tau_m^{u_2} \leq t < \tau_m^{u_1} \vee \tau_m^{u_2}) = 0$. Using Proposition B.7, we obtain

$$\lim_{u_1\to u_2} |f_m(u_1) - f_m(u_2)| = 0.$$

Hence, the functional $f_m(u)$ is continuous with respect to the control $u \in U$. □

As a consequence, we get the following existence and uniqueness result.

Theorem 6.15. *Let the functional $J_m \colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to \mathbb{R}$ be given by (6.13). Then there exists a unique optimal control $\overline{u}_m \in U$.*

Proof. The space $L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ is a Hilbert space and thus, a uniformly convex Banach space and by definition, the set of admissible controls $U \subset L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ is bounded and closed such that $0 \in U$. Due to Lemma 6.14, the functional

$$f_m(u) = \mathbb{E} \int_0^{\tau_m^u} \|A^\gamma(y(t;u) - y_d(t))\|_H^2 \, dt$$

is continuous and obviously, we have $f_m(u) \geq 0$ for every $u \in U$. Applying Corollary D.18 with $p = 2$, the claim follows. □

Remark 6.16. *As shown in [7], the previous theorem can be proven for the stochastic Navier-Stokes equations with multiplicative Lévy noise, i.e. in system (6.3), we replace the Q-Wiener process by a square integrable Lévy martingale as introduced in Section 3.3.*

6.3.2. The Linearized Stochastic Navier-Stokes Equations

We introduce the following system in $D(A^\alpha)$:

$$\begin{cases} dz(t) = -[Az(t) + B(z(t), y(t)) + B(y(t), z(t)) - Fv(t)] \, dt + G(z(t)) \, dW(t), \\ z(0) = 0, \end{cases} \tag{6.15}$$

where $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, the process $(y(t))_{t\in[0,\tau)}$ is the local mild solution of system (6.3) and the process $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in H and covariance operator $Q \in \mathcal{L}_1^+(H)$. The operators A, B, F, G are introduced in Section 6.1 and Section 6.2, respectively.

Similarly to Section 6.2, we first consider the following modified system in $D(A^\alpha)$:

$$\begin{cases} dz_m(t) = -[Az_m(t) + B(z_m(t), \pi_m(y_m(t))) + B(\pi_m(y_m(t)), z_m(t)) - Fv(t)]\, dt + G(z_m(t))\, dW(t), \\ z_m(0) = 0, \end{cases} \tag{6.16}$$

where the process $(y_m(t))_{t\in[0,T]}$ is the mild solution of system (6.5) and $\pi_m\colon D(A^\alpha) \to D(A^\alpha)$ is given by equation (6.6). By Theorem 6.6, we get the existence and uniqueness of the mild solution $(y_m(t))_{t\in[0,T]}$ to system (6.5) for fixed $m \in \mathbb{N}$ and fixed control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Recall that the initial value $\xi \in L^2(\Omega; D(A^\alpha))$ is fixed as well. Similarly to Theorem 6.6, we obtain the existence and uniqueness of a mild solution $(z_m(t))_{t\in[0,T]}$ to system (6.16) for fixed $m \in \mathbb{N}$ and fixed $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. As a consequence, the process $(z_m(t))_{t\in[0,T]}$ takes values in $D(A^\alpha)$ such that

$$\mathbb{E} \sup_{t\in[0,T]} \|z_m(t)\|^2_{D(A^\alpha)} < \infty$$

and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} z_m(t) = &- \int_0^t A^\delta e^{-A(t-s)} A^{-\delta} \left[B(z_m(s), \pi_m(y_m(s))) + B(\pi_m(y_m(s)), z_m(s))\right] ds + \int_0^t e^{-A(t-s)} Fv(s)\, ds \\ &+ \int_0^t e^{-A(t-s)} G(z_m(s))\, dW(s). \end{aligned}$$

Due to Theorem 6.7, we get the existence and uniqueness of the local mild solution $(y(t))_{t\in[0,\tau)}$ to system (6.5) for fixed control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Similarly to Theorem 6.7, we obtain the existence and uniqueness of a local mild solution $(z(t))_{t\in[0,\tau)}$ to system (6.15) for fixed $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, where the stopping times $(\tau_m)_{m\in\mathbb{N}}$ are given by equation (6.10). Therefore, the process $(z(t))_{t\in[0,\tau)}$ takes values in $D(A^\alpha)$ such that for fixed $m \in \mathbb{N}$

$$\mathbb{E} \sup_{t\in[0,\tau_m)} \|z(t)\|^2_{D(A^\alpha)} < \infty$$

and we have for each $m \in \mathbb{N}$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} z(t\wedge\tau_m) = &- \int_0^{t\wedge\tau_m} A^\delta e^{-A(t\wedge\tau_m - s)} A^{-\delta} \left[B(z(s), y(s)) + B(y(s), z(s))\right] ds + \int_0^{t\wedge\tau_m} e^{-A(t\wedge\tau_m - s)} Fv(s)\, ds \\ &+ I_{\tau_m}(G(z))(t\wedge\tau_m), \end{aligned}$$

where

$$I_{\tau_m}(G(z))(t) = \int_0^t \mathbb{1}_{[0,\tau_m)}(s) e^{-A(t-s)} G(z(s\wedge\tau_m))\, dW(s).$$

Next, we show some useful properties. Note that the mild solution of system (6.5) depends on the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Hence, the mild solution of system (6.16) depends on the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ as well as on the control $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Let us denote by $(z_m(t;u,v))_{t\in[0,T]}$ the mild solution of system (6.16). Similarly, we indicate by $(z(t;u,v)))_{t\in[0,\tau^u)}$ the local mild solution of system (6.15) corresponding to the controls $u, v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Whenever these processes are considered for fixed controls, we use the notation introduced above.

Lemma 6.17. *For fixed $m \in \mathbb{N}$, let $(z_m(t;u,v))_{t\in[0,T]}$ be the mild solution of system (6.16) corresponding to the controls $u,v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. If $v \in L^k_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ for $k \geq 2$, then there exists a constant $\tilde{c} > 0$ such that*

$$\mathbb{E} \sup_{t\in[0,T]} \|z_m(t;u,v)\|^k_{D(A^\alpha)} \leq \tilde{c}\, \|v\|^k_{L^k_{\mathcal{F}}(\Omega;L^2([0,T];D(A^\beta)))}. \tag{6.17}$$

Proof. Let the stochastic process $(y_m(t;u))_{t\in[0,T]}$ be the mild solution of system (6.5) corresponding to the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Recall that the operator $F\colon D(A^\beta) \to D(A^\beta)$ and the operator $G\colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$ are bounded. Let $T_{1,m} \in (0,T]$. By Theorem 2.35, Lemma 6.1, Proposition 3.65 (ii), inequality (6.7) and the Cauchy-Schwarz inequality, there exist constants $C_1, C_2, C_3 > 0$ such that

$$\begin{aligned}
&\mathbb{E} \sup_{t\in[0,T_{1,m}]} \|z_m(t;u,v)\|^k_{D(A^\alpha)} \\
&\leq 4^{k-1}\mathbb{E} \sup_{t\in[0,T_{1,m}]} \left(\int_0^t \left\| A^{\alpha+\delta} e^{-A(t-s)} A^{-\delta} B(z_m(s;u,v), \pi_m(y_m(s;u))) \right\|_H ds \right)^k \\
&\quad + 4^{k-1}\mathbb{E} \sup_{t\in[0,T_{1,m}]} \left(\int_0^t \left\| A^{\alpha+\delta} e^{-A(t-s)} A^{-\delta} B(\pi_m(y_m(s;u)), z_m(s;u,v)) \right\|_H ds \right)^k \\
&\quad + 4^{k-1}\mathbb{E} \sup_{t\in[0,T_{1,m}]} \left(\int_0^t \left\| A^{\alpha-\beta} e^{-A(t-s)} A^{\beta} F v(s) \right\|_H ds \right)^k \\
&\quad + 4^{k-1}\mathbb{E} \sup_{t\in[0,T_{1,m}]} \left\| \int_0^t e^{-A(t-s)} A^\alpha G(z_m(s;u,v))\, dW(s) \right\|^k_H \\
&\leq \left(C_1 T_{1,m}^{k(1-\alpha-\delta)} + C_2 T_{1,m}^{k/2} \right) \mathbb{E} \sup_{t\in[0,T_{1,m}]} \|z_m(t;u,v)\|^k_{D(A^\alpha)} + C_3\, \mathbb{E} \left(\int_0^T \|v(t)\|^2_{D(A^\beta)}\, dt \right)^{k/2}.
\end{aligned}$$

We chose $T_{1,m} \in (0,T]$ such that $C_1 T_{1,m}^{k(1-\alpha-\delta)} + C_2 T_{1,m}^{k/2} < 1$. Then we have

$$\mathbb{E} \sup_{t\in[0,T_{1,m}]} \|z_m(t;u,v)\|^k_{D(A^\alpha)} \leq c_{1,m}\, \mathbb{E} \left(\int_0^T \|v(t)\|^2_{D(A^\beta)}\, dt \right)^{k/2},$$

where $c_{1,m} = \frac{C_3}{1 - C_1 T_{1,m}^{k(1-\alpha-\delta)} - C_2 T_{1,m}^{k/2}}$. By definition, we have for all $t \in [T_{1,m}, T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned}
z_m(t;u,v) = {} & e^{-A(t-T_{1,m})} z_m(T_{1,m};u,v) \\
& - \int_{T_{1,m}}^t A^\delta e^{-A(t-s)} A^{-\delta} \left[B(z_m(s;u,v), \pi_m(y_m(s;u))) + B(\pi_m(y_m(s;u)), z_m(s;u,v)) \right] ds \\
& + \int_{T_{1,m}}^t e^{-A(t-s)} F v(s)\, ds + \int_{T_{1,m}}^t e^{-A(t-s)} G(z_m(s;u,v))\, dW(s).
\end{aligned}$$

Again, we find $T_{2,m} \in [T_{1,m}, T]$ such that

$$\mathbb{E} \sup_{t\in[T_{1,m},T_{2,m}]} \|z_m(t;u,v)\|^k_{D(A^\alpha)} \leq c_{2,m}\mathbb{E}\left(\int_0^T \|v(t)\|^2_{D(A^\beta)}\,dt\right)^{k/2},$$

where $c_{2,m} > 0$ is a constant. By continuing the method, we obtain inequality (6.17). □

Lemma 6.18. *For fixed $m \in \mathbb{N}$, let $(z_m(t;u,v))_{t\in[0,T]}$ be the mild solution of system (6.16) corresponding to the controls $u, v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Then we have for every $u, v_1, v_2 \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, all $a, b \in \mathbb{R}$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.*

$$z_m(t;u,a\,v_1 + b\,v_2) = a\,z_m(t;u,v_1) + b\,z_m(t;u,v_2).$$

Proof. Let the process $(y_m(t;u))_{t\in[0,T]}$ be the mild solution of system (6.5) corresponding to the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. To simplify the notation, we set for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\tilde{z}_m(t) = z_m(t;u,a\,v_1 + b\,v_2) - a\,z_m(t;u,v_1) - b\,z_m(t;u,v_2).$$

Recall that the operators $F\colon D(A^\beta) \to D(A^\beta)$ and $G\colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$ are linear and bounded. Let $T_{1,m} \in (0,T]$. By Theorem 2.35, Lemma 6.1, Proposition 3.65 (ii) with $k = 2$ and inequality (6.7), there exist constants $C_1, C_2 > 0$ such that

$$\begin{aligned}
\mathbb{E} \sup_{t\in[0,T_{1,m}]} \|\tilde{z}_m(t)\|^2_{D(A^\alpha)} &\leq 3\,\mathbb{E} \sup_{t\in[0,T_{1,m}]} \left(\int_0^t \left\|A^{\alpha+\delta}e^{-A(t-s)}A^{-\delta}B(\tilde{z}_m(s), \pi_m(y_m(s;u)))\right\|_H ds\right)^2 \\
&\quad + 3\,\mathbb{E} \sup_{t\in[0,T_{1,m}]} \left(\int_0^t \left\|A^{\alpha+\delta}e^{-A(t-s)}A^{-\delta}B(\pi_m(y_m(s;u)), \tilde{z}_m(s))\right\|_H ds\right)^2 \\
&\quad + 3\,\mathbb{E} \sup_{t\in[0,T_{1,m}]} \left\|\int_0^t e^{-A(t-s)}A^\alpha G(\tilde{z}_m(s))\,dW(s)\right\|^2_H \\
&\leq \left(C_1 T_{1,m}^{2-2\alpha-2\delta} + C_2 T_{1,m}\right)\mathbb{E} \sup_{t\in[0,T_{1,m}]} \|\tilde{z}_m(t)\|^2_{D(A^\alpha)}.
\end{aligned}$$

We chose $T_{1,m} \in (0,T]$ such that $C_1 T_{1,m}^{2-2\alpha-2\delta} + C_2 T_{1,m} < 1$. Then we have

$$\mathbb{E} \sup_{t\in[0,T_{1,m}]} \|\tilde{z}_m(t)\|^2_{D(A^\alpha)} = \mathbb{E} \sup_{t\in[0,T_{1,m}]} \|z_m(t;u,a\,v_1 + b\,v_2) - a\,z_m(t;u,v_1) - b\,z_m(t;u,v_2)\|^2_{D(A^\alpha)} = 0.$$

Similarly to Lemma 6.17, we can conclude that the result holds for the whole time interval $[0,T]$. □

Lemma 6.19. *For fixed $m \in \mathbb{N}$, let $(z_m(t;u,v))_{t\in[0,T]}$ be the mild solution of system (6.16) corresponding to the controls $u, v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Then there exists a constant $\overline{c} > 0$ such that for every $u_1, u_2 \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ and every $v \in L^4_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$*

$$\mathbb{E} \sup_{t\in[0,T]} \|z_m(t;u_1,v) - z_m(t;u_2,v)\|^2_{D(A^\alpha)} \leq \overline{c}\,\|v\|^2_{L^4_{\mathcal{F}}(\Omega;L^2([0,T];D(A^\beta)))}\|u_1 - u_2\|_{L^2_{\mathcal{F}}(\Omega;L^2([0,T];D(A^\beta)))}. \tag{6.18}$$

Proof. To simplify the notation, we set for every $y, z \in D(A^\alpha)$

$$\widetilde{B}(y,z) = B(z,y) + B(y,z).$$

Since the operator B is bilinear on $D(A^\alpha) \times D(A^\alpha)$, the operator $\widetilde{B}$ is bilinear as well and using Lemma 6.1, we get for every $y, z \in D(A^\alpha)$

$$\left\| A^{-\delta} \widetilde{B}(y,z) \right\|_H \leq 2\widetilde{M} \|y\|_{D(A^\alpha)} \|z\|_{D(A^\alpha)}. \tag{6.19}$$

Let the stochastic process $(y_m(t;u_i))_{t\in[0,T]}$ be the mild solution of system (6.5) corresponding to the control $u_i \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ for $i = 1,2$. Recall that the operator $G\colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$ is linear and bounded. Let $T_{1,m} \in (0,T]$. By Theorem 2.35, the inequalities (6.7), (6.8) and (6.19), Proposition 3.65 (ii) with $k = 2$ and the Cauchy-Schwarz inequality, there exist constants $C_1, C_2, C_3 > 0$ such that

$$\begin{aligned}
&\mathbb{E} \sup_{t\in[0,T_{1,m}]} \|z_m(t;u_1,v) - z_m(t;u_2,v)\|^2_{D(A^\alpha)} \\
&\leq 3\,\mathbb{E} \sup_{t\in[0,T_{1,m}]} \left(\int_0^t \left\| A^{\alpha+\delta} e^{-A(t-s)} A^{-\delta} \widetilde{B}(\pi_m(y_m(s;u_1)), z_m(s;u_1,v) - z_m(s;u_2,v)) \right\|_H ds \right)^2 \\
&\quad + 3\,\mathbb{E} \sup_{t\in[0,T_{1,m}]} \left(\int_0^t \left\| A^{\alpha+\delta} e^{-A(t-s)} A^{-\delta} \widetilde{B}(\pi_m(y_m(s;u_1)) - \pi_m(y_m(s;u_2)), z_m(s;u_2,v)) \right\|_H ds \right)^2 \\
&\quad + 3\,\mathbb{E} \sup_{t\in[0,T_{1,m}]} \left\| \int_0^t e^{-A(t-s)} A^\alpha G(z_m(s;u_1,v) - z_m(s;u_2,v))\, dW(s) \right\|^2_H \\
&\leq \left(C_1 T_{1,m}^{2-2\alpha-2\delta} + C_2 T_{1,m} \right) \mathbb{E} \sup_{t\in[0,T_{1,m}]} \|z_m(t;u_1,v) - z_m(t;u_2,v)\|^2_{D(A^\alpha)} \\
&\quad + C_3 \left(\mathbb{E} \sup_{t\in[0,T_{1,m}]} \|z_m(t;u_2,v)\|^4_{D(A^\alpha)} \right)^{1/2} \left(\mathbb{E} \sup_{t\in[0,T_{1,m}]} \|y_m(t;u_1) - y_m(t;u_2)\|^2_{D(A^\alpha)} \right)^{1/2}.
\end{aligned}$$

Using Lemma 6.9 with $k = 2$ and Lemma 6.17 with $k = 4$, we can conclude that there exists a constant $C_3^* > 0$ such that

$$\begin{aligned}
&\mathbb{E} \sup_{t\in[0,T_{1,m}]} \|z_m(t;u_1,v) - z_m(t;u_2,v)\|^2_{D(A^\alpha)} \\
&\leq \left(C_1 T_{1,m}^{2-2\alpha-2\delta} + C_2 T_{1,m} \right) \mathbb{E} \sup_{t\in[0,T_{1,m}]} \|z_m(t;u_1,v) - z_m(t;u_2,v)\|^2_{D(A^\alpha)} \\
&\quad + C_3^* \left(\mathbb{E} \left[\int_0^T \|v(t)\|^2_{D(A^\beta)} dt \right]^2 \right)^{1/2} \left(\mathbb{E} \int_0^T \|u_1(t) - u_2(t)\|^2_{D(A^\beta)} dt \right)^{1/2}.
\end{aligned}$$

We chose $T_{1,m} \in (0,T]$ such that $C_1 T_{1,m}^{2-2\alpha-2\delta} + C_2 T_{1,m} < 1$. Then we infer

$$\begin{aligned}
&\mathbb{E} \sup_{t\in[0,T_{1,m}]} \|z_m(t;u_1,v) - z_m(t;u_2,v)\|^2_{D(A^\alpha)} \\
&\leq c_{1,m} \left(\mathbb{E} \left[\int_0^T \|v(t)\|^2_{D(A^\beta)} dt \right]^2 \right)^{1/2} \left(\mathbb{E} \int_0^T \|u_1(t) - u_2(t)\|^2_{D(A^\beta)} dt \right)^{1/2},
\end{aligned}$$

where $c_{1,m} = \frac{C_3^*}{1-C_1 T_{1,m}^{2-2\alpha-2\delta} - C_2 T_{1,m}}$. Similarly to Lemma 6.17, we can conclude that the result holds for the whole time interval $[0, T]$. $\square$

Remark 6.20. *By definition, we have for all $t \in [0, \tau_m^u)$ and $\mathbb{P}$-a.s. $z(t; u, v) = z_m(t; u, v)$. Hence, one can easily obtain similar results for the local mild solution of system (6.15).*

6.3.3. The Derivatives of the Cost Functional

First, we show that the local mild solution of system (6.15) is the Gâteaux derivative of the local mild solution to system (6.3) with respect to the control variable.

Theorem 6.21. *Let $(y(t; u))_{t\in[0,\tau^u)}$ and $(z(t; u, v))_{t\in[0,\tau^u)}$ be the local mild solution of system (6.3) and system (6.15) corresponding to the controls $u, v \in L^2_{\mathcal{F}}(\Omega; L^2([0, T]; D(A^\beta)))$, respectively. Then for fixed $m \in \mathbb{N}$, the Gâteaux derivative of $y(t; u)$ at $u \in L^2_{\mathcal{F}}(\Omega; L^2([0, T]; D(A^\beta)))$ in direction $v \in L^2_{\mathcal{F}}(\Omega; L^2([0, T]; D(A^\beta)))$ satisfies for all $t \in [0, \tau_m^u)$ and $\mathbb{P}$-a.s.*

$$d_u^G y(t; u)[v] = z(t; u, v).$$

Proof. First, we assume that $u, v \in L^4_{\mathcal{F}}(\Omega; L^2([0, T]; D(A^\beta)))$. Since the operator B is bilinear on the space $D(A^\alpha) \times D(A^\alpha)$ and the operators $F\colon D(A^\beta) \to D(A^\beta)$ and $G\colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$ are linear, we find for all $\theta \in \mathbb{R}\backslash\{0\}$, all $t \in [0, \tau_m^u \wedge \tau_m^{u+\theta v})$ and $\mathbb{P}$-a.s.

$$\begin{aligned}
&\frac{1}{\theta}[y(t; u + \theta v) - y(t; u)] - z(t; u, v) \\
&= -\int_0^t A^\delta e^{-A(t-s)} A^{-\delta} B\left(y(s; u + \theta v), \frac{1}{\theta}[y(s; u + \theta v) - y(s; u)] - z(s; u, v)\right) ds \\
&\quad - \int_0^t A^\delta e^{-A(t-s)} A^{-\delta} B\left(\frac{1}{\theta}[y(s; u + \theta v) - y(s; u)] - z(s; u, v), y(s; u)\right) ds \\
&\quad - \int_0^t A^\delta e^{-A(t-s)} A^{-\delta} B(y(s; u + \theta v) - y(s; u), z(s; u, v))\, ds \\
&\quad + \int_0^t e^{-A(t-s)} G\left(\frac{1}{\theta}[y(s; u + \theta v) - y(s; u)] - z(s; u, v)\right) dW(s).
\end{aligned} \tag{6.20}$$

Next, let $0 = T_{0,m} < T_{1,m} < ... < T_{l,m} = T$ be a partition of the time interval $[0, T]$, which we specify below. Since the stopping time $\tau_m^u \wedge \tau_m^{u+\theta v}$ takes values in $[0, T]$, we have for almost all $\omega \in \Omega$ and all $\theta \in \mathbb{R}\backslash\{0\}$

$$\mathbb{1}_{\tau_m^u \wedge \tau_m^{u+\theta v} \in [0,T_{1,m}]}(\omega) + \sum_{j=1}^{l-1} \mathbb{1}_{\tau_m^u \wedge \tau_m^{u+\theta v} \in (T_{j,m},T_{j+1,m}]}(\omega) = 1. \tag{6.21}$$

To simplify notation, we set $\mathbb{1}_0 = \mathbb{1}_{\tau_m^u \wedge \tau_m^{u+\theta v} \in [0,T_{1,m}]}$ and $\mathbb{1}_j = \mathbb{1}_{\tau_m^u \wedge \tau_m^{u+\theta v} \in (T_{j,m},T_{j+1,m}]}$ for $j = 1, ..., l-1$. Furthermore, let $(y_m(t; u^*))_{t\in[0,T]}$ and $(z_m(t; u^*, v^*))_{t\in[0,T]}$ be the mild solutions of system (6.5) and system (6.16) corresponding to arbitrary controls $u^*, v^* \in L^2_{\mathcal{F}}(\Omega; L^2([0, T]; D(A^\beta)))$, respectively. By definition, we have for every $u^*, v^* \in L^2_{\mathcal{F}}(\Omega; L^2([0, T]; D(A^\beta)))$, all $t \in [0, \tau_m^{u^*})$ and $\mathbb{P}$-a.s. $y(t; u^*) = y_m(t; u^*)$ and $z(t; u^*, v^*) = z_m(t; u^*, v^*)$. Recall that $G\colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$ is bounded. By equation (6.20), Theorem 2.35, Lemma 6.1, Proposition 3.65 (ii) with $k = 2$ and the Cauchy-Schwarz inequality, there exist constants $C_1, C_2, C_3 > 0$ such that for all $\theta \in \mathbb{R}\backslash\{0\}$ and for $j = 1, ..., l-1$

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{1}_j \sup_{t\in[0,T_{1,m}]}\left\|\frac{1}{\theta}[y(t;u+\theta v)-y(t;u)]-z(t;u,v)\right\|^2_{D(A^\alpha)}\right]\\
&\leq \left(C_1 T_{1,m}^{2-2\alpha-2\delta}+C_2T_{1,m}\right)\mathbb{E}\left[\mathbb{1}_j \sup_{t\in[0,T_{1,m}]}\left\|\frac{1}{\theta}[y(t;u+\theta v)-y(t;u)]-z(t;u,v)\right\|^2_{D(A^\alpha)}\right]\\
&\quad + C_3\left(\mathbb{E}\sup_{t\in[0,T_{1,m}]}\|z_m(t;u,v)\|^4_{D(A^\alpha)}\right)^{1/2}\left(\mathbb{E}\sup_{t\in[0,T_{1,m}]}\|y_m(t;u+\theta v)-y_m(t;u)\|^4_{D(A^\alpha)}\right)^{1/2}.
\end{aligned}$$

We chose $T_{1,m} \in (0,T]$ such that $C_1 T_{1,m}^{2-2\alpha-2\delta} + C_2 T_{1,m} < 1$. Then we find for all $\theta \in \mathbb{R}\backslash\{0\}$ and for $j = 1, ..., l-1$

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{1}_j \sup_{t\in[0,T_{1,m}]}\left\|\frac{1}{\theta}[y(t;u+\theta v)-y(t;u)]-z(t;u,v)\right\|^2_{D(A^\alpha)}\right]\\
&\leq c_{1,m}\left(\mathbb{E}\sup_{t\in[0,T_{1,m}]}\|z_m(t;u,v)\|^4_{D(A^\alpha)}\right)^{1/2}\left(\mathbb{E}\sup_{t\in[0,T_{1,m}]}\|y_m(t;u+\theta v)-y_m(t;u)\|^4_{D(A^\alpha)}\right)^{1/2},
\end{aligned}$$

where $c_{1,m} = \frac{C_3}{1-C_1T_{1,m}^{2-2\alpha-2\delta}-C_2T_{1,m}}$. Using Lemma 6.9 with $k = 4$ and Lemma 6.17 with $k = 4$, we can conclude for $j = 1, ..., l-1$

$$\lim_{\theta\to 0}\mathbb{E}\left[\mathbb{1}_j \sup_{t\in[0,T_{1,m}]}\left\|\frac{1}{\theta}[y(t;u+\theta v)-y(t;u)]-z(t;u,v)\right\|^2_{D(A^\alpha)}\right]=0. \tag{6.22}$$

Similarly, we get

$$\lim_{\theta\to 0}\mathbb{E}\left[\mathbb{1}_0 \sup_{t\in[0,\tau_m^u\wedge\tau_m^{u+\theta v})}\left\|\frac{1}{\theta}[y(t;u+\theta v)-y(t;u)]-z(t;u,v)\right\|^2_{D(A^\alpha)}\right]=0.$$

By definition, we have for all $t \in [T_{1,m}, T]$, $\mathbb{P}$-a.s. and for $i = 1, 2$

$$\begin{aligned}
y(t\wedge\tau_m^{u_i};u_i) &= e^{-A(t\wedge\tau_m^{u_i}-T_{1,m}\wedge\tau_m^{u_i})}\left[y(T_{1,m}\wedge\tau_m^{u_i};u_i)-I_{\tau_m^{u_i}}(G(y))(T_{1,m}\wedge\tau_m^{u_i})\right]\\
&\quad-\int_{T_{1,m}\wedge\tau_m^{u_i}}^{t\wedge\tau_m^{u_i}} A^\delta e^{-A(t\wedge\tau_m^{u_i}-s)}A^{-\delta}B(y(s;u_i))\,ds+\int_{T_{1,m}\wedge\tau_m^{u_i}}^{t\wedge\tau_m^{u_i}} e^{-A(t\wedge\tau_m^{u_i}-s)}Fu_i(s)\,ds\\
&\quad+I_{\tau_m^{u_i}}(G(y))(t\wedge\tau_m^{u_i}),
\end{aligned}$$

where $u_1 = u + \theta v$ and $u_2 = u$ and

$$\begin{aligned}
z(t\wedge\tau_m^{u};u,v) &= e^{-A(t\wedge\tau_m^{u}-T_{1,m}\wedge\tau_m^{u})}\left[z(T_{1,m}\wedge\tau_m^{u};u,v)-I_{\tau_m^{u}}(G(z))(T_{1,m}\wedge\tau_m^{u})\right]\\
&\quad-\int_{T_{1,m}\wedge\tau_m^{u}}^{t\wedge\tau_m^{u}} A^\delta e^{-A(t\wedge\tau_m^{u}-s)}A^{-\delta}\left[B(z(s;u,v),y(s;u))+B(y(s;u),z(s;u,v))\right]ds\\
&\quad+\int_{T_{1,m}\wedge\tau_m^{u}}^{t\wedge\tau_m^{u}} e^{-A(t\wedge\tau_m^{u}-s)}Fv(s)\,ds+I_{\tau_m^{u}}(G(z))(t\wedge\tau_m^{u}).
\end{aligned}$$

Again, we find $T_{2,m} \in [T_{1,m}, T]$ such that for $j = 2..., l-1$

$$\lim_{\theta\to 0} \mathbb{E}\left[\mathbb{1}_j \sup_{t\in[T_{1,m},T_{2,m}]} \left\| \frac{1}{\theta}[y(t;u+\theta v) - y(t;u)] - z(t;u,v)\right\|^2_{D(A^\alpha)}\right] = 0$$

and

$$\lim_{\theta\to 0} \mathbb{E}\left[\mathbb{1}_1 \sup_{t\in[T_{1,m},\tau_m^u\wedge\tau_m^{u+\theta v})} \left\| \frac{1}{\theta}[y(t;u+\theta v) - y(t;u)] - z(t;u,v)\right\|^2_{D(A^\alpha)}\right] = 0.$$

Using equation (6.22) for $j = 1$, we obtain

$$\lim_{\theta\to 0} \mathbb{E}\left[\mathbb{1}_1 \sup_{t\in[0,\tau_m^u\wedge\tau_m^{u+\theta v})} \left\| \frac{1}{\theta}[y(t;u+\theta v) - y(t;u)] - z(t;u,v)\right\|^2_{D(A^\alpha)}\right] = 0.$$

By continuing the method, we obtain for $j = 0, 1, ..., l-1$

$$\lim_{\theta\to 0} \mathbb{E}\left[\mathbb{1}_j \sup_{t\in[0,\tau_m^u\wedge\tau_m^{u+\theta v})} \left\| \frac{1}{\theta}[y(t;u+\theta v) - y(t;u)] - z(t;u,v)\right\|^2_{D(A^\alpha)}\right] = 0.$$

Due to equation (6.21), we have

$$\begin{aligned}&\lim_{\theta\to 0} \mathbb{E} \sup_{t\in[0,\tau_m^u\wedge\tau_m^{u+\theta v})} \left\| \frac{1}{\theta}[y(t;u+\theta v) - y(t;u)] - z(t;u,v)\right\|^2_{D(A^\alpha)} \\ &= \sum_{j=0}^{l-1} \lim_{\theta\to 0} \mathbb{E}\left[\mathbb{1}_j \sup_{t\in[0,\tau_m^u\wedge\tau_m^{u+\theta v})} \left\| \frac{1}{\theta}[y(t;u+\theta v) - y(t;u)] - z(t;u,v)\right\|^2_{D(A^\alpha)}\right] = 0.\end{aligned}$$

Therefore, the Gâteaux derivative of the velocity field $(y(t;u))_{t\in[0,\tau^u)}$ at $u \in L^4_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ in direction $v \in L^4_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ satisfies for all $t \in [0, \tau_m^u \wedge \tau_m^{u+\theta v})$ and $\mathbb{P}$-a.s.

$$d_u^G y(t;u)[v] = z(t;u,v). \tag{6.23}$$

Note that by Lemma 6.11, we have

$$\lim_{\theta\to 0} \mathbb{P}(\tau_m^u \neq \tau_m^{u+\theta v}) = 0.$$

Moreover, the operator $d_u^G y(t;u)$ is linear and bounded due to Lemma 6.17 with $k = 4$ and Lemma 6.18. Since the space $L^4_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ is dense in $L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, the equation (6.23) holds for $u, v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, which is a consequence of Lemma 6.17 with $k = 2$, Lemma 6.18 and Lemma 6.19. □

This enables us to derive the Gâteaux derivative of the cost functional.

Theorem 6.22. *Let $J_m \colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to \mathbb{R}$ be defined by (6.13). Then the Gâteaux derivative at $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ in direction $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ satisfies*

$$d^G J_m(u)[v] = \mathbb{E}\int_0^{\tau_m^u} \langle A^\gamma(y(t;u) - y_d(t)), A^\gamma z(t;u,v)\rangle_H \, dt + \mathbb{E}\int_0^T \left\langle A^\beta u(t), A^\beta v(t)\right\rangle_H dt,$$

where the process $(z(t;u,v))_{t\in[0,\tau^u)}$ is the local mild solution of system (6.15) corresponding to the controls $u, v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$.

Proof. We define the functionals $\Phi_1, \Phi_2 \colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to \mathbb{R}$ by

$$\Phi_1(u) = \frac{1}{2}\,\mathbb{E}\int_0^{\tau_m^u} \|A^\gamma(y(t;u) - y_d(t))\|_H^2\,dt, \qquad \Phi_2(u) = \frac{1}{2}\,\mathbb{E}\int_0^T \|A^\beta u(t)\|_H^2 dt.$$

First, we derive the Gâteaux derivative of the functional Φ_1 at $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ in direction $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. We get for all $\theta \in \mathbb{R}\backslash\{0\}$

$$\begin{aligned}&\left|\frac{1}{\theta}[\Phi_1(u+\theta v) - \Phi_1(u)] - \mathbb{E}\int_0^{\tau_m^u} \langle A^\gamma(y(t;u) - y_d(t)), A^\gamma z(t;u,v)\rangle_H\,dt\right|\\ &\leq \mathcal{I}_1(\theta) + \mathcal{I}_2(\theta) + \mathcal{I}_3(\theta) + \mathcal{I}_4(\theta) + \mathcal{I}_5(\theta),\end{aligned} \tag{6.24}$$

where

$$\mathcal{I}_1(\theta) = \left|\frac{1}{2\theta}\,\mathbb{E}\int_0^{\tau_m^u \wedge \tau_m^{u+\theta v}} \|A^\gamma(y(t;u+\theta v) - y(t;u))\|_H^2\,dt\right|,$$

$$\mathcal{I}_2(\theta) = \left|\mathbb{E}\int_0^{\tau_m^u \wedge \tau_m^{u+\theta v}} \left\langle A^\gamma(y(t;u) - y_d(t)), A^\gamma\left(\frac{1}{\theta}[y(t;u+\theta v) - y(t;u)] - z(t;u,v)\right)\right\rangle_H dt\right|,$$

$$\mathcal{I}_3(\theta) = \left|\frac{1}{2\theta}\,\mathbb{E}\int_{\tau_m^u \wedge \tau_m^{u+\theta v}}^{\tau_m^{u+\theta v}} \|A^\gamma(y(t;u+\theta v) - y_d(t))\|_H^2\,dt\right|,$$

$$\mathcal{I}_4(\theta) = \left|\frac{1}{2\theta}\,\mathbb{E}\int_{\tau_m^u \wedge \tau_m^{u+\theta v}}^{\tau_m^u} \|A^\gamma(y(t;u) - y_d(t))\|_H^2\,dt\right|,$$

$$\mathcal{I}_5(\theta) = \left|\mathbb{E}\int_{\tau_m^u \wedge \tau_m^{u+\theta v}}^{\tau_m^u} \langle A^\gamma(y(t;u) - y_d(t)), A^\gamma z(t;u,v)\rangle_H\,dt\right|.$$

Let the process $(y_m(t;u^*))_{t\in[0,T]}$ be the mild solution of system (6.5) corresponding to an arbitrary control $u^* \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. We have for every $u^* \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, all $t \in [0, \tau_m^{u^*})$ and $\mathbb{P}$-a.s. $y(t;u^*) = y_m(t;u^*)$ and $\|y(t;u^*)\|_{D(A^\alpha)} \leq m$. Using Corollary 2.32, we obtain for all $\theta \in \mathbb{R}\backslash\{0\}$

$$\mathcal{I}_1(\theta) \leq \left|\frac{CT}{2\theta}\,\mathbb{E}\sup_{t\in[0,T]} \|y_m(t;u+\theta v) - y_m(t;u)\|_{D(A^\alpha)}^2\right|.$$

Due to Lemma 6.9 with $k = 2$, we can conclude

$$\lim_{\theta\to 0} \mathcal{I}_1(\theta) = 0. \tag{6.25}$$

Using the Cauchy-Schwarz inequality and Corollary 2.32, there exists a constant $C^* > 0$ such that for all $\theta \in \mathbb{R}\backslash\{0\}$

$$\mathcal{I}_2(\theta) \leq C^*\left(\mathbb{E}\sup_{t\in[0,\tau_m^u \wedge \tau_m^{u+\theta v})} \left\|\frac{1}{\theta}[y(t;u+\theta v) - y(t;u)] - z(t;u,v)\right\|_{D(A^\alpha)}^2\right)^{1/2}.$$

Due to Theorem 6.21, we can infer

$$\lim_{\theta\to 0} \mathcal{I}_2(\theta) = 0. \tag{6.26}$$

Using Corollary 2.32 and Fubini's theorem, we get for all $\theta \in \mathbb{R}\backslash\{0\}$

$$\mathcal{I}_3(\theta) \leq \left| \int_0^T \frac{1}{2\theta} \mathbb{P}\left(\tau_m^u \wedge \tau_m^{u+\theta v} \leq t < \tau_m^{u+\theta v}\right) \left(2C^2 m^2 + 2\left\|y_d(t)\right\|_{D(A^\gamma)}^2\right) dt \right|.$$

Due to Lemma 6.12 with $k = 1$, we have

$$\lim_{\theta\to 0} \frac{1}{\theta} \mathbb{P}\left(\tau_m^u \wedge \tau_m^{u+\theta v} \leq t < \tau_m^{u+\theta v}\right) = 0$$

for all $t \in [0, T]$. By Proposition B.7, we can infer

$$\lim_{\theta\to 0} \mathcal{I}_3(\theta) = 0. \tag{6.27}$$

Similarly, we find

$$\lim_{\theta\to 0} \mathcal{I}_4(\theta) + \lim_{\theta\to 0} \mathcal{I}_5(\theta) = 0. \tag{6.28}$$

Using inequality (6.24) and equations (6.25) – (6.28), we get

$$\lim_{\theta\to 0} \left| \frac{1}{\theta}[\Phi_1(u+\theta v) - \Phi_1(u)] - \mathbb{E} \int_0^{\tau_m^u} \langle A^\gamma(y(t;u) - y_d(t)), A^\gamma z(t;u,v)\rangle_H \, dt \right| = 0.$$

Therefore, the Gâteaux derivative of $\Phi_1 \colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to \mathbb{R}$ at $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ in direction $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ is given by

$$d^G\Phi_1(u)[v] = \mathbb{E} \int_0^{\tau_m^u} \langle A^\gamma(y(t;u) - y_d(t)), A^\gamma z(t;u,v)\rangle_H \, dt. \tag{6.29}$$

Let the stochastic process $(z_m(t;u,v))_{t\in[0,T]}$ be the mild solution of system (6.16) corresponding to the controls $u, v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. We have for all $t \in [0, \tau_m^u)$ and $\mathbb{P}$-a.s. $z(t;u,v) = z_m(t;u,v)$. Using Lemma 6.18, the functional $d^G\Phi_1(u)$ is linear. Moreover, by the Cauchy-Schwarz inequality, Corollary 2.32, Lemma 6.17 with $k = 2$, there exists a constant $C^* > 0$ such that

$$\begin{aligned}\left|d^G\Phi_1(u)[v]\right|^2 &\leq \left(2\,\mathbb{E}\int_0^{\tau_m^u} \|A^\gamma y(t;u)\|_H^2 dt + 2\int_0^T \|A^\gamma y_d(t)\|_H^2 dt\right) \mathbb{E}\int_0^{\tau_m^u} \|A^\gamma z(t;u,v)\|_H^2 dt \\ &\leq C^* \left\|v\right\|^2_{L^2_{\mathcal{F}}(\Omega;L^2([0,T];D(A^\beta)))}.\end{aligned}$$

Hence, the functional $d^G\Phi_1(u)$ is bounded.

Note that the functional $\Phi_2 \colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to \mathbb{R}$ is given by the squared norm in the Hilbert space $L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Similarly to Remark D.6 (ii), we get that the Gâteaux derivative of Φ_2 at $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ in direction $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ is given by

$$d^G\Phi_2(u)[v] = \mathbb{E} \int_0^T \left\langle A^\beta u(t), A^\beta v(t)\right\rangle_H dt. \tag{6.30}$$

Obviously, the functional $d^G\Phi_2(u)$ is linear and bounded.

Using equation (6.29) and equation (6.30), the Gâteaux derivative of J_m at $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ in direction $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ is given by

$$\begin{aligned} d^G J_m(u)[v] &= d^G\Phi_1(u)[v] + d^G\Phi_2(u)[v] \\ &= \mathbb{E}\int_0^{\tau_m^u} \langle A^\gamma(y(t;u) - y_d(t)), A^\gamma z(t;u,v)\rangle_H \, dt + \mathbb{E}\int_0^T \left\langle A^\beta u(t), A^\beta v(t)\right\rangle_H dt. \end{aligned}$$

Since $d^G\Phi_1(u)$ and $d^G\Phi_2(u)$ are linear and bounded, the functional $d^G J_m(u)$ is linear and bounded as well. □

Recall that the set of admissible controls U is a closed, bounded and convex subset of the Hilbert space $L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ such that $0 \in U$. Moreover, the cost functional $J_m \colon U \to \mathbb{R}$ given by equation (6.13) satisfies the assumptions of Proposition D.19. Hence, the optimal control $\overline{u}_m \in U$ satisfies the necessary optimality condition

$$d^G J_m(\overline{u}_m)[u - \overline{u}_m] \geq 0 \tag{6.31}$$

for fixed $m \in \mathbb{N}$ and every $u \in U$. Due to Theorem 6.22, we get the variational inequality

$$\mathbb{E}\int_0^{\tau_m^{\overline{u}_m}} \langle A^\gamma(y(t;\overline{u}_m) - y_d(t)), A^\gamma z(t;\overline{u}_m, u - \overline{u}_m)\rangle_H \, dt + \mathbb{E}\int_0^T \left\langle A^\beta \overline{u}_m(t), A^\beta(u(t) - \overline{u}_m(t))\right\rangle_H dt \geq 0 \tag{6.32}$$

for fixed $m \in \mathbb{N}$ and every $u \in U$. We will use this inequality to derive an explicit formula for the optimal control $\overline{u}_m \in U$.

Remark 6.23. *If system (6.3) is driven by a square integrable Lévy martingale (introduced in Section 3.3), then one may obtain the necessary optimality condition (6.31) as follows:*
Note that it is also necessary to replace the Q-Wiener process by a square integrable Lévy martingale in system (6.15). Especially, a generalization of Lemma 6.17 with $k = 4$ is required as an auxiliary result to obtain Theorem 6.21. To prove this lemma, we need a generalization of the maximal inequality stated in Proposition 3.65 (i) for $k = 4$. Such an inequality can be found in [49, Proposition 1.3 (i)], which requires to calculate the quadratic variation of a square integrable Lévy martingale. In general, the quadratic variation of a right-continuous square integrable martingale $(M(t))_{t\geq 0}$ with values in an arbitrary Hilbert space $\mathcal{H}$ is defined by

$$[M]_t = \lim_{j\to\infty} \sum_{t_i \in P_j} \|M(t_{i+1} \wedge t) - M(t_i \wedge t)\|^2_{\mathcal{H}} \tag{6.33}$$

for all $t \geq 0$ and $\mathbb{P}$-almost surely, where P_j contains the points of time $0 < t_0 < t_1 < \ldots < t_j$ satisfying $\lim_{i\to\infty} t_i = \infty$ and $\lim_{j\to\infty} \delta(P_j) = 0$ with $\delta(P_j) = \sup_{t_i \in P_j}(t_{i+1} - t_i)$. The convergence of equation (6.33) is in $L^1(\Omega, \mathcal{F}, \mathbb{P})$. Let $(\langle M\rangle_t)_{t\geq 0}$ be the predictable variation of $(M(t))_{t\geq 0}$ introduced in Theorem 3.20. If $(M(t))_{t\geq 0}$ is real-valued and continuous, then

$$[M]_t = \langle M\rangle_t$$

for all $t \geq 0$ and $\mathbb{P}$-almost surely, see [63, Theorem 18.6]. One can easily adopt the proof to obtain this result for a Hilbert space valued process. However, for a square integrable Lévy martingale, the quadratic variation and the predictable variation do not coincide in general. Hence, the determination of the quadratic variation to a square integrable Lévy martingale might be the most challenging task here.

Next, we state the second order Gâteaux derivative of the cost functional (6.13). Moreover, we show that the Gâteaux derivatives and the Fréchet derivatives coincide, which will enable us to obtain a sufficient optimality condition.

Corollary 6.24. *Let* $J_m\colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to \mathbb{R}$ *be defined by (6.13). Then the Gâteaux derivative of order two at* $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ *in directions* $v_1, v_2 \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ *satisfies*

$$d^G(J_m(u))^2[v_1,v_2] = \mathbb{E}\int_0^{\tau_m^u} \langle A^\gamma z(t;u,v_1), A^\gamma z(t;u,v_2)\rangle_H \, dt + \mathbb{E}\int_0^T \left\langle A^\beta v_1(t), A^\beta v_2(t)\right\rangle_H dt,$$

where the processes $(z(t;u,v_i))_{t\in[0,\tau^u)}$ *form the local mild solution of system (6.15) corresponding to the controls* $u, v_i \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ *for* $i = 1,2$.

Proof. The result can be obtained similarly to Theorem 6.22. ☐

Corollary 6.25. *Let* $J_m\colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to \mathbb{R}$ *be defined by (6.13). Then the Fréchet derivative at* $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ *in direction* $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ *satisfies*

$$d^F J_m(u)[v] = \mathbb{E}\int_0^{\tau_m^u} \langle A^\gamma (y(t;u) - y_d(t)), A^\gamma z(t;u,v)\rangle_H \, dt + \mathbb{E}\int_0^T \left\langle A^\beta u(t), A^\beta v(t)\right\rangle_H dt,$$

where the process $(z(t;u,v))_{t\in[0,\tau^u)}$ *is the local mild solution of system (6.15) corresponding to the controls* $u, v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. *Moreover, the functional* $d^F J_m(u)[v]$ *is continuous with respect to* u.

Proof. Using Theorem 6.22, we have that the Gâteaux derivative at $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ in direction $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ satisfies

$$d^G J_m(u)[v] = \mathbb{E}\int_0^{\tau_m^u} \langle A^\gamma (y(t;u) - y_d(t)), A^\gamma z(t;u,v)\rangle_H \, dt + \mathbb{E}\int_0^T \left\langle A^\beta u(t), A^\beta v(t)\right\rangle_H dt.$$

If $v \in L^4_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, then the process $(z(t;u,v))_{t\in[0,\tau_m^u)}$ is continuous with respect to the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ resulting from Lemma 6.19. By Lemma 6.17 with $k = 2$, Lemma 6.18 and the fact that the space $L^4_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ is dense in $L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, we can conclude that the process $(z(t;u,v))_{t\in[0,\tau_m^u)}$ is continuous with respect to $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ for $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. The fact that $(y(t;u))_{t\in[0,\tau_m^u)}$ is continuous with respect to the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ is an immediate consequence of Lemma 6.9 with $k = 2$. Using additionally Lemma 6.11, one can show that $u \mapsto d^G J_m(u)$ is a continuous mapping from $L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ into $\mathcal{L}(L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))); \mathbb{R})$. Hence, we can apply Corollary D.5 and the claim follows. ☐

Corollary 6.26. *Let* $J_m\colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to \mathbb{R}$ *be defined by (6.13). Then the Fréchet derivative of order two at* $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ *in directions* $v_1, v_2 \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ *satisfies*

$$d^F(J_m(u))^2[v_1,v_2] = \mathbb{E}\int_0^{\tau_m^u} \langle A^\gamma z(t;u,v_1), A^\gamma z(t;u,v_2)\rangle_H \, dt + \mathbb{E}\int_0^T \left\langle A^\beta v_1(t), A^\beta v_2(t)\right\rangle_H dt,$$

where the processes $(z(t;u,v_i))_{t\in[0,\tau^u)}$ *are the local mild solution of system (6.15) corresponding to the controls* $u, v_i \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ *for* $i = 1,2$. *Moreover, the functional* $d^F(J_m(u))^2[v_1,v_2]$ *is continuous with respect to* u.

Proof. The result can be obtained similarly to Corollary 6.25. ☐

6.4. The Adjoint Equation

We will use the necessary optimality condition (6.32) to derive an explicit formula the optimal control $\overline{u}_m \in U$ has to satisfy. Therefor, we need a duality principle, which provides a relation between the local mild solution to system (6.15) and the corresponding adjoint equation, which is given by the following backward SPDE in $D(A^\delta)$:

$$\begin{cases} dz_m^*(t) = -\mathbb{1}_{[0,\tau_m)}(t)[-Az_m^*(t) - A^{2\alpha} B_\delta^* \left(y(t), A^\delta z_m^*(t)\right) + G^*(A^{-2\alpha}\Phi_m(t)) \\ \qquad\qquad + A^{2\gamma}(y(t) - y_d(t))]\, dt + \Phi_m(t)\, dW(t), \\ z_m^*(T) = 0, \end{cases} \tag{6.34}$$

where $m \in \mathbb{N}$ and the process $(y(t))_{t\in[0,\tau)}$ is the local mild solution of system (6.3). The stopping times $(\tau_m)_{m\in\mathbb{N}}$ are defined by equation (6.10) and $y_d \in L^2([0,T]; D(A^\gamma))$ is the given desired velocity field. The operator A and its fractional powers are introduced in Section 6.1. The process $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in H and covariance operator $Q \in \mathcal{L}_1^+(H)$. The operators $B_\delta^*(y(t), \cdot) : H \to D(A^\alpha)$ for $t \in [0, \tau_m)$ and $G^* \colon \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha)) \to H$ are linear and bounded. A precise meaning is given in the following remark.

Remark 6.27. *(i) By Lemma 6.1, we obtain that the operator $A^{-\delta}[B(\cdot, y) + B(y, \cdot)] \colon D(A^\alpha) \to H$ is linear and bounded for every $y \in D(A^\alpha)$ such that $\|y\|_{D(A^\alpha)} \leq m$. Therefore, there exists a linear and bounded operator $B_\delta^*(y, \cdot) : H \to D(A^\alpha)$ satisfying for every $h \in H$ and every $z \in D(A^\alpha)$*

$$\langle A^{-\delta}[B(z, y) + B(y, z)], h\rangle_H = \langle z, B_\delta^*(y, h)\rangle_{D(A^\alpha)}.$$

We can rewrite this equivalently as

$$\langle A^{-\delta}[B(z, y) + B(y, z)], h\rangle_H = \langle A^\alpha z, A^\alpha B_\delta^*(y, h)\rangle_H \tag{6.35}$$

for every $h \in H$ and every $z \in D(A^\alpha)$. By the closed graph theorem, the operator $A^\alpha B_\delta^*(y, \cdot) : H \to H$ is linear and bounded.
(ii) Recall that $\|y(t)\|_{D(A^\alpha)} \leq m$ for all $t \in [0, \tau_m)$ and $\mathbb{P}$-almost surely.
(iii) Due to the fact that the operator $G \colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$ is linear and bounded, there exists a linear and bounded operator $G^* \colon \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha)) \to H$ satisfying for every $h \in H$ and every $\Phi \in \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$

$$\langle G(h), \Phi\rangle_{\mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))} = \langle h, G^*(\Phi)\rangle_H.$$

We can rewrite this equivalently as

$$\langle A^\alpha G(h), A^\alpha \Phi\rangle_{\mathcal{L}_{(HS)}(Q^{1/2}(H); H)} = \langle h, G^*(\Phi)\rangle_H \tag{6.36}$$

for every $h \in H$ and every $\Phi \in \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$.

Here, we use a mild solution to system (6.34) in the sense of Definition 3.95 with $\mathcal{H} = H$. By Theorem 6.7, we get the existence and uniqueness of a local mild solution $(y(t))_{t\in[0,\tau)}$ to system (6.3) for fixed control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. As a consequence of Theorem 3.98, we can conclude that there exists a unique mild solution $(z_m^*(t), \Phi_m(t))_{t\in[0,T]}$ of system (6.34) for fixed control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ and fixed $m \in \mathbb{N}$. Hence, the pair of processes $(z_m^*(t), \Phi_m(t))_{t\in[0,T]}$ takes values in $D(A^\delta) \times \mathcal{L}_{(HS)}(Q^{1/2}(H); H)$ such that

$$\mathbb{E} \sup_{t\in[0,T]} \|z_m^*(t)\|^2_{D(A^\delta)} < \infty,$$

$$\mathbb{E} \int_0^T \|\Phi_m(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(\mathcal{H}); \mathcal{H})} dt < \infty$$

and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} z_m^*(t) = &- \int_t^T \mathbb{1}_{[0,\tau_m)}(s) A^\alpha e^{-A(s-t)} A^\alpha B_\delta^* \left(y(s\wedge\tau_m), A^\delta z_m^*(s)\right) ds \\ &+ \int_t^T \mathbb{1}_{[0,\tau_m)}(s) e^{-A(s-t)} G^*(A^{-2\alpha}\Phi_m(s))\, ds \\ &+ \int_t^T \mathbb{1}_{[0,\tau_m)}(s) A^\gamma e^{-A(s-t)} A^\gamma \left(y(s\wedge\tau_m) - y_d(s)\right) ds \\ &- \int_t^T e^{-A(s-t)}\Phi_m(s)\, dW(s). \end{aligned}$$

Since the local mild solution of system (6.3) depends on the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, we get this property for the mild solution of system (6.34) as well. To illustrate the dependence on the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, let us denote by $(z_m^*(t;u), \Phi_m(t;u))_{t\in[0,T]}$ the mild solution of system (6.34). Whenever these processes are considered for fixed controls, we use the notation introduced above.

Remark 6.28. *(i) As a consequence of Theorem 3.98, we get the additional restrictions $\alpha, \delta < \frac{1}{2}$ and $\gamma + \delta < \frac{1}{2}$. Therefore, we can not solve the control problem (6.14) for $\gamma = \frac{1}{2}$ through a stochastic maximum principle directly. However, we will show in Section 6.7 that we easily overcome this problem in the case of additive noise in system (6.3).*
(ii) If $y_d \in L^\infty([0,T]; D(A^\gamma))$, then the restriction $\gamma + \delta < \frac{1}{2}$ vanishes.

Lemma 6.29. *Let $(z_m^*(t), \Phi_m(t))_{t\in[0,T]}$ be the mild solution of system (6.34). Then we have for fixed $m \in \mathbb{N}$*

$$\mathbb{E} \sup_{t\in[\tau_m,T]} \|z_m^*(t)\|^2_{D(A^\delta)} = 0 \quad \textit{and} \quad \mathbb{E}\int_{\tau_m}^T \|\Phi_m(t)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)}\, dt = 0.$$

Proof. By definition, we obtain for all $t \in [\tau_m, T]$ and $\mathbb{P}$-a.s.

$$z_m^*(t) = -\int_t^T e^{-A(s-t)}\Phi_m(s)\, dW(s).$$

The claim follows by Lemma 3.96. $\square$

6.5. Approximation by a Strong Formulation

As shown in Section 5.6, through a duality principle one can utilize an optimality condition to obtain an explicit formula the optimal control has to satisfy. A duality principle of solutions to forward and backward SPDEs can be obtained by applying an Itô product formula. Since this formula is not applicable to solutions in a mild sense, we derive an approximation similarly to Section 5.5. Recall that the operator $R(\lambda) \in \mathcal{L}(H)$ is given by

$$R(\lambda) = \lambda R(\lambda; -A) \tag{6.37}$$

for all $\lambda \in \rho(-A)$, where $\lambda R(\lambda; -A)$ is the resolvent operator of $-A$ introduced in Section 2.1. Especially, we use Lemma 5.13, which is directly applicable here. Furthermore, we omit the dependence on the controls in this section.

6.5.1. The Forward Equation

Here, we provide an approximation of the mild solution to system (6.16). We introduce the following system in $D(A^{1+\alpha})$:

$$\begin{cases} dz_m(t,\lambda) = -[Az_m(t,\lambda) + R(\lambda)B(R(\lambda)z_m(t,\lambda), \pi_m(y_m(t))) \\ \qquad\qquad + R(\lambda)B(\pi_m(y_m(t)), R(\lambda)z_m(t,\lambda)) - R(\lambda)Fv(t)]\,dt + R(\lambda)G(R(\lambda)z_m(t,\lambda))\,dW(t), \\ z_m(0,\lambda) = 0, \end{cases} \tag{6.38}$$

where $m \in \mathbb{N}$ and $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. The operator $R(\lambda)$ is given by equation (6.37) with $\lambda > 0$ and the operators A, B, F, G are introduced in Section 6.1 and Section 6.2, respectively. The mapping $\pi_m\colon D(A^\alpha) \to D(A^\alpha)$ is given by (6.6) and the process $(y_m(t))_{t\in[0,T]}$ is the mild solution of system (6.5). The process $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in H and covariance operator $Q \in \mathcal{L}_1^+(H)$.

Definition 6.30. *A predictable process $(z_m(t,\lambda))_{t\in[0,T]}$ with values in $D(A^{1+\alpha})$ is called a* ***mild solution of system (6.38)*** *if*

$$\mathbb{E} \sup_{t\in[0,T]} \|z_m(t,\lambda)\|^2_{D(A^{1+\alpha})} < \infty$$

and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$z_m(t,\lambda) = -\int_0^t A^\delta e^{-A(t-s)} R(\lambda) A^{-\delta} \left[B(R(\lambda)z_m(s,\lambda), \pi_m(y_m(s))) + B(\pi_m(y_m(s)), R(\lambda)z_m(s,\lambda))\right] ds$$
$$+ \int_0^t e^{-A(t-s)} R(\lambda) F v(s)\, ds + \int_0^t e^{-A(t-s)} R(\lambda) G(R(\lambda) z_m(s,\lambda))\, dW(s).$$

Recall that the operators $R(\lambda)$ and $AR(\lambda)$ are linear and bounded on H. Hence, for fixed $m \in \mathbb{N}$ and fixed $\lambda > 0$, an existence and uniqueness result of a mild solution to system (6.38) can be obtained similarly to the mild solution of system (6.16). In the following lemma, we state a strong formulation of the mild solution to system (6.38).

Lemma 6.31. *Let $(z_m(t,\lambda))_{t\in[0,T]}$ be the mild solution of system (6.38). Then we have for fixed $m \in \mathbb{N}$, fixed $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.*

$$z_m(t,\lambda) = -\int_0^t A z_m(s,\lambda) + A^\delta R(\lambda) A^{-\delta} \left[B(R(\lambda)z_m(s,\lambda), \pi_m(y_m(s))) + B(\pi_m(y_m(s)), R(\lambda)z_m(s,\lambda))\right] ds$$
$$+ \int_0^t R(\lambda) F v(s)\, ds + \int_0^t R(\lambda) G(R(\lambda) z_m(s,\lambda))\, dW(s).$$

Proof. The claim follows immediately from Theorem 2.35, Theorem 3.106 and Lemma 5.13. □

We get the following convergence result.

Lemma 6.32. *Let $(z_m(t))_{t\in[0,T]}$ and $(z_m(t,\lambda))_{t\in[0,T]}$ be the mild solutions of system (6.16) and system (6.38), respectively. Then we have for fixed $m \in \mathbb{N}$*

$$\lim_{\lambda\to\infty} \mathbb{E} \sup_{t\in[0,T]} \|z_m(t) - z_m(t,\lambda)\|^2_{D(A^\alpha)} = 0.$$

Proof. Let I be the identity operator on H. We define the operator

$$\widetilde{B}(y,z) = B(z,y) + B(y,z)$$

for every $y, z \in D(A^\alpha)$. Since B is bilinear on $D(A^\alpha) \times D(A^\alpha)$, the operator $\widetilde{B}$ is bilinear as well and using Lemma 6.1, we get for every $y, z \in D(A^\alpha)$

$$\left\| A^{-\delta} \widetilde{B}(y,z) \right\|_H \leq 2\widetilde{M} \|y\|_{D(A^\alpha)} \|z\|_{D(A^\alpha)}. \tag{6.39}$$

Recall that the operator $G \colon H \to \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$ is linear and bounded. By definition, we find for all $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned}
z_m(t) - z_m(t,\lambda) = &- \int_0^t A^\delta e^{-A(t-s)} A^{-\delta} \widetilde{B}(\pi_m(y_m(s)), [I - R(\lambda)] z_m(s))\, ds \\
&- \int_0^t A^\delta e^{-A(t-s)} [I - R(\lambda)] A^{-\delta} \widetilde{B}(\pi_m(y_m(s)), R(\lambda) z_m(s))\, ds \\
&- \int_0^t A^\delta e^{-A(t-s)} R(\lambda) A^{-\delta} \widetilde{B}(\pi_m(y_m(s)), R(\lambda)\, [z_m(s) - z_m(s,\lambda)])\, ds \\
&+ \int_0^t e^{-A(t-s)} [I - R(\lambda)] F v(s)\, ds \\
&+ \int_0^t e^{-A(t-s)} G([I - R(\lambda)] z_m(s))\, dW(s) \\
&+ \int_0^t e^{-A(t-s)} [I - R(\lambda)] G(R(\lambda) z_m(s))\, dW(s) \\
&+ \int_0^t e^{-A(t-s)} R(\lambda) G(R(\lambda)\, [z_m(s) - z_m(s,\lambda)])\, dW(s).
\end{aligned}$$

Let $T_{1,m} \in (0,T]$. Then we get for all $\lambda > 0$

$$\mathbb{E} \sup_{t \in [0,T_{1,m}]} \|z_m(t) - z_m(t,\lambda)\|^2_{D(A^\alpha)} \leq 7\, \mathcal{I}_1(\lambda) + 7\, \mathcal{I}_2(\lambda) + 7\, \mathcal{I}_3(\lambda), \tag{6.40}$$

where

$$\begin{aligned}
\mathcal{I}_1(\lambda) = &\,\mathbb{E} \sup_{t \in [0,T_{1,m}]} \left\| \int_0^t A^\delta e^{-A(t-s)} R(\lambda) A^{-\delta} \widetilde{B}(\pi_m(y_m(s)), R(\lambda)\, [z_m(s) - z_m(s,\lambda)])\, ds \right\|^2_{D(A^\alpha)} \\
&+ \mathbb{E} \sup_{t \in [0,T_{1,m}]} \left\| \int_0^t e^{-A(t-s)} R(\lambda) G(R(\lambda)\, [z_m(s) - z_m(s,\lambda)])\, dW(s) \right\|^2_{D(A^\alpha)},
\end{aligned}$$

$$\mathcal{I}_2(\lambda) = \mathbb{E} \sup_{t\in[0,T_{1,m}]} \left\| \int_0^t A^{\delta} e^{-A(t-s)} A^{-\delta} \widetilde{B}(\pi_m(y_m(s)), [I-R(\lambda)] z_m(s))\, ds \right\|^2_{D(A^\alpha)}$$

$$+ \mathbb{E} \sup_{t\in[0,T_{1,m}]} \left\| \int_0^t A^{\delta} e^{-A(t-s)} [I-R(\lambda)] A^{-\delta} \widetilde{B}(\pi_m(y_m(s)), R(\lambda) z_m(s))\, ds \right\|^2_{D(A^\alpha)}$$

$$+ \mathbb{E} \sup_{t\in[0,T_{1,m}]} \left\| \int_0^t e^{-A(t-s)} [I-R(\lambda)] F v(s)\, ds \right\|^2_{D(A^\alpha)},$$

$$\mathcal{I}_3(\lambda) = \mathbb{E} \sup_{t\in[0,T_{1,m}]} \left\| \int_0^t e^{-A(t-s)} G([I-R(\lambda)] z_m(s))\, dW(s) \right\|^2_{D(A^\alpha)}$$

$$+ \mathbb{E} \sup_{t\in[0,T_{1,m}]} \left\| \int_0^t e^{-A(t-s)} [I-R(\lambda)] G(R(\lambda) z_m(s))\, dW(s) \right\|^2_{D(A^\alpha)}.$$

By Theorem 2.35, Corollary 2.32, Lemma 5.13, Proposition 3.65 (ii) with $k = 2$, inequality (6.7) and inequality (6.39), there exist constants $C_1, C_2 > 0$ such that for all $\lambda > 0$

$$\mathcal{I}_1(\lambda) \le \left(C_1 T_{1,m}^{2-2\alpha-2\delta} + C_2 T_{1,m} \right) \mathbb{E} \sup_{t\in[0,T_{1,m}]} \| z_m(t) - z_m(t,\lambda) \|^2_{D(A^\alpha)}. \tag{6.41}$$

Similarly, there exists a constant $C^* > 0$ such that for all $\lambda > 0$

$$\mathcal{I}_2(\lambda) \le C^* \, \mathbb{E} \sup_{t\in[0,T_{1,m}]} \| [I-R(\lambda)] A^\alpha z_m(t) \|^2_H + C^* \, \mathbb{E} \sup_{t\in[0,T_{1,m}]} \left\| [I-R(\lambda)] A^{-\delta} \widetilde{B}(\pi_m(y_m(t)), R(\lambda) z_m(t)) \right\|^2_H$$

$$+ C^* \, \mathbb{E} \int_0^{T_{1,m}} \left\| [I-R(\lambda)] A^\beta F v(t) \right\|^2_H dt,$$

$$\mathcal{I}_3(\lambda) \le C^* \, \mathbb{E} \int_0^{T_{1,m}} \| [I-R(\lambda)] z_m(t) \|^2_H \, dt + C^* \, \mathbb{E} \int_0^{T_{1,m}} \| [I-R(\lambda)] A^\alpha G(R(\lambda) z_m(t)) \|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} \, dt.$$

Using Lemma 5.13 (iii) and Proposition B.7, we can conclude

$$\lim_{\lambda\to\infty} \mathcal{I}_2(\lambda) + \lim_{\lambda\to\infty} \mathcal{I}_3(\lambda) = 0. \tag{6.42}$$

Due to inequality (6.40) and inequality (6.41), we find for all $\lambda > 0$

$$\mathbb{E} \sup_{t\in[0,T_{1,m}]} \| z_m(t) - z_m(t,\lambda) \|^2_{D(A^\alpha)} \le K_{1,m} \, \mathbb{E} \sup_{t\in[0,T_{1,m}]} \| z_m(t) - z_m(t,\lambda) \|^2_{D(A^\alpha)} + 7\, \mathcal{I}_2(\lambda) + 7\, \mathcal{I}_3(\lambda),$$

where $K_{1,m} = 7 C_1 T_{1,m}^{2-2\alpha-2\delta} + 7 C_2 T_{1,m}$. We chose $T_{1,m} \in (0, T]$ such that $K_{1,m} < 1$. Then we obtain for all $\lambda > 0$

$$\mathbb{E} \sup_{t\in[0,T_{1,m}]} \| z_m(t) - z_m(t,\lambda) \|^2_{D(A^\alpha)} \le \frac{7\, \mathcal{I}_2(\lambda) + 7\, \mathcal{I}_3(\lambda)}{1 - K_{1,m}}.$$

By equation (6.42), we get

$$\lim_{\lambda\to\infty} \mathbb{E} \sup_{t\in[0,T_{1,m}]} \|z_m(t) - z_m(t,\lambda)\|^2_{D(A^\alpha)} = 0.$$

Similarly to Lemma 6.17, we can conclude that the result holds for the whole time interval $[0,T]$. □

6.5.2. The Backward Equation

Here we provide an approximation of the mild solution to system (6.34). We introduce the following backward SPDE in $D(A^{1+\delta})$:

$$\begin{cases} dz^*_m(t,\lambda) = -\mathbb{1}_{[0,\tau_m)}(t) \left[-Az^*_m(t,\lambda) - A^\alpha R(\lambda) A^\alpha B^*_\delta \left(y(t), R(\lambda)A^\delta z^*_m(t,\lambda) \right) \right. \\ \qquad\qquad \left. + R(\lambda) G^*(A^{-2\alpha} R(\lambda)\Phi_m(t,\lambda)) + A^\gamma R(\lambda) A^\gamma \left(y(t) - y_d(t) \right) \right] dt + \Phi_m(t,\lambda)\, dW(t), \\ z^*_m(T,\lambda) = 0, \end{cases} \tag{6.43}$$

where $m \in \mathbb{N}$. The operator $R(\lambda)$ is given by equation (6.37) with $\lambda > 0$ and the operators A, B^*_δ, G^* are introduced in Section 6.1 and Section 6.4, respectively. The process $(y(t))_{t\in[0,\tau)}$ is the local mild solution of system (6.3) with stopping times $(\tau_m)_{m\in\mathbb{N}}$ defined by (6.10) and $y_d \in L^2([0,T]; D(A^\gamma))$ is the given desired velocity field. The process $(W(t))_{t\geq 0}$ is a Q-Wiener process with values in H and covariance operator $Q \in \mathcal{L}^+_1(H)$.

Definition 6.33. *We call a pair of predictable processes $(z^*_m(t,\lambda), \Phi_m(t,\lambda))_{t\in[0,T]}$ with values in the space $D(A^{1+\delta}) \times \mathcal{L}_{(HS)}(Q^{1/2}(H); H)$ a **mild solution of system (6.43)** if*

$$\mathbb{E} \sup_{t\in[0,T]} \|z^*_m(t,\lambda)\|^2_{D(A^{1+\delta})} < \infty, \quad \mathbb{E} \int_0^T \|\Phi_m(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dt < \infty$$

and we have for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} z^*_m(t,\lambda) = &- \int_t^T \mathbb{1}_{[0,\tau_m)}(s) A^\alpha e^{-A(s-t)} R(\lambda) A^\alpha B^*_\delta \left(y(s\wedge\tau_m), R(\lambda)A^\delta z^*_m(s,\lambda) \right) ds \\ &+ \int_t^T \mathbb{1}_{[0,\tau_m)}(s) e^{-A(s-t)} R(\lambda) G^*(A^{-2\alpha} R(\lambda)\Phi_m(s,\lambda))\, ds \\ &+ \int_t^T \mathbb{1}_{[0,\tau_m)}(s) A^\gamma e^{-A(s-t)} R(\lambda) A^\gamma \left(y(s\wedge\tau_m) - y_d(s) \right) ds - \int_t^T e^{-A(s-t)} \Phi_m(s,\lambda)\, dW(s). \end{aligned}$$

Recall that the operators $R(\lambda)$ and $AR(\lambda)$ are linear and bounded on H. Hence, an existence and uniqueness result of a mild solution to system (6.43) can be obtained similarly to the mild solution of system (6.34) for fixed $m \in \mathbb{N}$ and fixed $\lambda > 0$. Moreover, we get the following result.

Lemma 6.34. *Let the pair of processes $(z^*_m(t,\lambda), \Phi_m(t,\lambda))_{t\in[0,T]}$ be the mild solution of system (6.43). Then we have for fixed $m \in \mathbb{N}$ and fixed $\lambda > 0$*

$$\mathbb{E} \sup_{t\in[\tau_m,T]} \|z^*_m(t,\lambda)\|^2_{D(A^{1+\delta})} = 0 \quad \textit{and} \quad \mathbb{E} \int_{\tau_m}^T \|\Phi_m(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dt = 0.$$

Proof. The claim follows similarly to Lemma 6.29. □

The following lemma gives us a strong formulation of the mild solution to system (6.43).

Lemma 6.35. *Let the pair of processes $(z_m^*(t,\lambda), \Phi_m(t,\lambda))_{t\in[0,T]}$ be the mild solution of system (6.43). Then we have for fixed $m \in \mathbb{N}$, fixed $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.*

$$\begin{aligned} z_m^*(t,\lambda) = &-\int_t^T \mathbb{1}_{[0,\tau_m)}(s)\left[Az_m^*(s,\lambda) + A^\alpha R(\lambda)A^\alpha B_\delta^*\left(y(s\wedge\tau_m), R(\lambda)A^\delta z_m^*(s,\lambda)\right)\right] ds \\ &+\int_t^T \mathbb{1}_{[0,\tau_m)}(s)R(\lambda)G^*(A^{-2\alpha}R(\lambda)\Phi_m(s,\lambda))\, ds \\ &+\int_t^T \mathbb{1}_{[0,\tau_m)}(s)A^\gamma R(\lambda)A^\gamma\left(y(s\wedge\tau_m) - y_d(s)\right) ds \\ &-\int_t^T \Phi_m(s,\lambda)\, dW(s). \end{aligned}$$

Proof. The claim follows from Theorem 2.35, Theorem 3.112 and Lemma 5.13. □

We get the following convergence results.

Lemma 6.36. *Let $(z_m^*(t), \Phi_m(t))_{t\in[0,T]}$ and $(z_m^*(t,\lambda), \Phi_m(t,\lambda))_{t\in[0,T]}$ be the mild solutions of system (6.34) and system (6.43), respectively. Then we have for fixed $m \in \mathbb{N}$*

$$\lim_{\lambda\to\infty} \mathbb{E} \sup_{t\in[0,T]} \|z_m^*(t) - z_m^*(t,\lambda)\|_{D(A^\delta)}^2 = 0,$$

$$\lim_{\lambda\to\infty} \mathbb{E} \int_0^T \|\Phi_m(t) - \Phi_m(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)} dt = 0.$$

Proof. Let I be the identity operator on H. By definition, we have for all $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} &z_m^*(t) - z_m^*(t,\lambda) \\ &= -\int_t^T \mathbb{1}_{[0,\tau_m)}(s)A^\alpha e^{-A(s-t)}[A^\alpha B_\delta^*\left(y(s\wedge\tau_m), A^\delta z_m^*(s)\right) - R(\lambda)A^\alpha B_\delta^*\left(y(s\wedge\tau_m), R(\lambda)A^\delta z_m^*(s,\lambda)\right)] ds \\ &+\int_t^T \mathbb{1}_{[0,\tau_m)}(s)e^{-A(s-t)}[G^*(A^{-2\alpha}\Phi_m(s)) - R(\lambda)G^*(A^{-2\alpha}R(\lambda)\Phi_m(s,\lambda))] ds \\ &+\int_t^T \mathbb{1}_{[0,\tau_m)}(s)A^\gamma e^{-A(s-t)}[I - R(\lambda)]A^\gamma\left(y(s\wedge\tau_m) - y_d(s)\right) ds \\ &-\int_t^T e^{-A(s-t)}[\Phi_m(s) - \Phi_m(s,\lambda)]\, dW(s). \end{aligned}$$

Recall that the operators $A^\alpha B_\delta^*(y(t),\cdot) : H \to H$ for $t \in [0,\tau_m)$ and $G^* \colon \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha)) \to H$ are linear and bounded. Hence, we find for all $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$
\begin{aligned}
z_m^*(t) - z_m^*(t,\lambda) = &- \int_t^T \mathbb{1}_{[0,\tau_m)}(s) A^\alpha e^{-A(s-t)} A^\alpha B_\delta^* \left(y(s\wedge\tau_m), [I-R(\lambda)]A^\delta z_m^*(s)\right) ds \\
&- \int_t^T \mathbb{1}_{[0,\tau_m)}(s) A^\alpha e^{-A(s-t)} [I-R(\lambda)] A^\alpha B_\delta^* \left(y(s\wedge\tau_m), R(\lambda)A^\delta z_m^*(s)\right) ds \\
&- \int_t^T \mathbb{1}_{[0,\tau_m)}(s) A^\alpha e^{-A(s-t)} R(\lambda) A^\alpha B_\delta^* \left(y(s\wedge\tau_m), R(\lambda)A^\delta [z_m^*(s) - z_m^*(s,\lambda)]\right) ds \\
&+ \int_t^T \mathbb{1}_{[0,\tau_m)}(s) e^{-A(s-t)} G^*(A^{-2\alpha}[I-R(\lambda)]\Phi_m(s))\, ds \\
&+ \int_t^T \mathbb{1}_{[0,\tau_m)}(s) e^{-A(s-t)} [I-R(\lambda)] G^*(A^{-2\alpha}R(\lambda)\Phi_m(s))\, ds \\
&+ \int_t^T \mathbb{1}_{[0,\tau_m)}(s) e^{-A(s-t)} R(\lambda) G^*(A^{-2\alpha}R(\lambda)[\Phi_m(s) - \Phi_m(s,\lambda)])\, ds \\
&+ \int_t^T \mathbb{1}_{[0,\tau_m)}(s) A^\gamma e^{-A(s-t)} [I-R(\lambda)] A^\gamma \left(y(s\wedge\tau_m) - y_d(s)\right) ds \\
&- \int_t^T e^{-A(s-t)} [\Phi_m(s) - \Phi_m(s,\lambda)]\, dW(s).
\end{aligned}
$$

Note that each integrand of the Bochner integrals on the right hand side satisfies the assumptions of Lemma 3.96 and Corollary 3.97, respectively. Let $T_{1,m} \in [0,T)$. Using inequality (3.39) and inequality (3.45), we get for all $\lambda > 0$

$$
\mathbb{E} \sup_{t\in[T_{1,m},T]} \|z_m^*(t) - z_m^*(t,\lambda)\|^2_{D(A^\delta)} \leq 7\,\mathcal{I}_1(\lambda) + 7\,\mathcal{I}_2(\lambda) + 7\,\mathcal{I}_3(\lambda), \tag{6.44}
$$

where

$$
\begin{aligned}
\mathcal{I}_1(\lambda) = &\,\hat{c}(T-T_{1,m})^{2-2\alpha-2\delta}\, \mathbb{E} \sup_{t\in[T_{1,m},T]} \left[\mathbb{1}_{[0,\tau_m)}(t) \left\|R(\lambda)A^\alpha B_\delta^*\left(y(t\wedge\tau_m), R(\lambda)A^\delta[z_m^*(t) - z_m^*(t,\lambda)]\right)\right\|_H^2\right] \\
&+ c(T-T_{1,m})^{1-2\delta}\, \mathbb{E} \int_{T1,m}^T \mathbb{1}_{[0,\tau_m)}(t) \left\|R(\lambda)G^*(A^{-2\alpha}R(\lambda)[\Phi_m(t) - \Phi_m(t,\lambda)])\right\|_H^2 dt, \\
\mathcal{I}_2(\lambda) = &\,\hat{c}(T-T_{1,m})^{2-2\alpha-2\delta}\, \mathbb{E} \sup_{t\in[T_{1,m},T]} \left[\mathbb{1}_{[0,\tau_m)}(t) \left\|A^\alpha B_\delta^*\left(y(t\wedge\tau_m), [I-R(\lambda)]A^\delta z_m^*(t)\right)\right\|_H^2\right] \\
&+ \hat{c}(T-T_{1,m})^{2-2\alpha-2\delta}\, \mathbb{E} \sup_{t\in[T_{1,m},T]} \left[\mathbb{1}_{[0,\tau_m)}(t) \left\|[I-R(\lambda)]A^\alpha B_\delta^*\left(y(t\wedge\tau_m), R(\lambda)A^\delta z_m^*(t)\right)\right\|_H^2\right],
\end{aligned}
$$

$$\mathcal{I}_3(\lambda) = c(T - T_{1,m})^{1-2\delta}\, \mathbb{E} \int_{T1,m}^{T} \mathbb{1}_{[0,\tau_m)}(t) \left\| G^*(A^{-2\alpha}[I - R(\lambda)]\Phi_m(t)) \right\|_H^2 dt$$
$$+ c(T - T_{1,m})^{1-2\delta}\, \mathbb{E} \int_{T1,m}^{T} \mathbb{1}_{[0,\tau_m)}(t) \left\| [I - R(\lambda)]G^*(A^{-2\alpha} R(\lambda)\Phi_m(t)) \right\|_H^2 dt$$
$$+ c(T - T_{1,m})^{1-2\gamma-2\delta}\, \mathbb{E} \int_{T1,m}^{T} \mathbb{1}_{[0,\tau_m)}(t) \left\| [I - R(\lambda)]A^{\gamma} \left(y(t \wedge \tau_m) - y_d(t)\right) \right\|_H^2 dt.$$

By Lemma 5.13 (ii), there exist constants $C_1, C_2 > 0$ such that for all $\lambda > 0$

$$\mathcal{I}_1(\lambda) \le C_1(T - T_{1,m})^{2-2\alpha-2\delta}\, \mathbb{E} \sup_{t\in[T_{1,m},T]} \|z_m^*(t) - z_m^*(t,\lambda)\|_{D(A^\delta)}^2$$
$$+ C_2(T - T_{1,m})^{1-2\delta}\, \mathbb{E} \int_{T_{1,m}}^{T} \|\Phi_m(t) - \Phi_m(t,\lambda)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)}^2 dt. \tag{6.45}$$

Moreover, there exists a constant $C^* > 0$ such that for all $\lambda > 0$

$$\mathcal{I}_2(\lambda) \le C^*\, \mathbb{E} \sup_{t\in[T_{1,m},T]} \left\| [I - R(\lambda)]A^\delta z_m^*(t) \right\|_H^2$$
$$+ C^*\, \mathbb{E} \sup_{t\in[T_{1,m},T]} \left[\mathbb{1}_{[0,\tau_m)}(t) \left\| [I - R(\lambda)]A^\alpha B_\delta^* \left(y(t\wedge\tau_m), R(\lambda)A^\delta z_m^*(t)\right) \right\|_H^2 \right],$$
$$\mathcal{I}_3(\lambda) \le C^*\, \mathbb{E} \int_{T1,m}^{T} \|[I - R(\lambda)]\Phi_m(t)\|_H^2 dt + C^*\, \mathbb{E} \int_{T1,m}^{T} \left\| [I - R(\lambda)]G^*(A^{-2\alpha} R(\lambda)\Phi_m(t)) \right\|_H^2 dt$$
$$+ C^*\, \mathbb{E} \int_{T1,m}^{T} \mathbb{1}_{[0,\tau_m)}(t) \left\| [I - R(\lambda)]A^{\gamma} \left(y(t \wedge \tau_m) - y_d(t)\right) \right\|_H^2 dt.$$

Using Lemma 5.13 (iii) and Proposition B.7, we can conclude

$$\lim_{\lambda\to\infty} \mathcal{I}_2(\lambda) + \lim_{\lambda\to\infty} \mathcal{I}_3(\lambda) = 0. \tag{6.46}$$

Due to inequality (3.40) and inequality (3.46), we get for all $\lambda > 0$

$$\mathbb{E} \int_{T_{1,m}}^{T} \|\Phi_m(t) - \Phi_m(t,\lambda)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)}^2 dt \le 7\, \mathcal{I}_4(\lambda) + 7\, \mathcal{I}_5(\lambda) + 7\, \mathcal{I}_6(\lambda), \tag{6.47}$$

where

$$\mathcal{I}_4(\lambda) = \hat{c}(T - T_{1,m})^{2-2\alpha}\, \mathbb{E} \sup_{t\in[T_{1,m},T]} \left[\mathbb{1}_{[0,\tau_m)}(t) \left\| R(\lambda)A^\alpha B_\delta^* \left(y(t\wedge\tau_m), R(\lambda)A^\delta [z_m^*(t) - z_m^*(t,\lambda)]\right) \right\|_H^2 \right]$$
$$+ c(T - T_{1,m})\, \mathbb{E} \int_{T_{1,m}}^{T} \mathbb{1}_{[0,\tau_m)}(t) \|R(\lambda)G^*(A^{-2\alpha} R(\lambda)[\Phi_m(t) - \Phi_m(t,\lambda)])\|_H^2 dt,$$

$$\mathcal{I}_5(\lambda) = \hat{c}(T - T_{1,m})^{2-2\alpha}\,\mathbb{E}\sup_{t\in[T_{1,m},T]}\Big[\mathbb{1}_{[0,\tau_m)}(t)\left\|A^{\alpha}B_{\delta}^{*}\left(y(t\wedge\tau_m),[I-R(\lambda)]A^{\delta}z_m^{*}(t)\right)\right\|_H^2\Big]$$
$$+\hat{c}(T - T_{1,m})^{2-2\alpha}\,\mathbb{E}\sup_{t\in[T_{1,m},T]}\Big[\mathbb{1}_{[0,\tau_m)}(t)\left\|[I-R(\lambda)]A^{\alpha}B_{\delta}^{*}\left(y(t\wedge\tau_m),R(\lambda)A^{\delta}z_m^{*}(t)\right)\right\|_H^2\Big],$$
$$\mathcal{I}_6(\lambda) = c(T-T_{1,m})\,\mathbb{E}\int_{T1,m}^{T}\mathbb{1}_{[0,\tau_m)}(t)\left\|G^{*}(A^{-2\alpha}[I-R(\lambda)]\Phi_m(t))\right\|_H^2\,dt$$
$$+c(T-T_{1,m})\,\mathbb{E}\int_{T1,m}^{T}\mathbb{1}_{[0,\tau_m)}(t)\left\|[I-R(\lambda)]G^{*}(A^{-2\alpha}R(\lambda)\Phi_m(t))\right\|_H^2\,dt$$
$$+c(T-T_{1,m})^{1-2\gamma}\,\mathbb{E}\int_{T1,m}^{T}\mathbb{1}_{[0,\tau_m)}(t)\left\|[I-R(\lambda)]A^{\gamma}\left(y(t\wedge\tau_m)-y_d(t)\right)\right\|_H^2\,dt.$$

Again, there exist constants $C_1, C_2 > 0$ such that for all $\lambda > 0$

$$\mathcal{I}_4(\lambda) \le C_1(T-T_{1,m})^{2-2\alpha}\,\mathbb{E}\sup_{t\in[T_{1,m},T]}\|z_m^{*}(t)-z_m^{*}(t,\lambda)\|_{D(A^{\delta})}^2$$
$$+C_2(T-T_{1,m})\,\mathbb{E}\int_{T_{1,m}}^{T}\|\Phi_m(t)-\Phi_m(t,\lambda)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)}^2\,dt. \tag{6.48}$$

Similarly to equation (6.46), we get

$$\lim_{\lambda\to\infty}\mathcal{I}_5(\lambda)+\lim_{\lambda\to\infty}\mathcal{I}_6(\lambda)=0. \tag{6.49}$$

By inequalities (6.44), (6.45), (6.47) and (6.48), we have for all $\lambda > 0$

$$\mathbb{E}\sup_{t\in[T_{1,m},T]}\|z_m^{*}(t)-z_m^{*}(t,\lambda)\|_{D(A^{\delta})}^2+\mathbb{E}\int_{T_{1,m}}^{T}\|\Phi_m(t)-\Phi_m(t,\lambda)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)}^2\,dt$$
$$\le K_{1,m}\left(\mathbb{E}\sup_{t\in[T_{1,m},T]}\|z_m^{*}(t)-z_m^{*}(t,\lambda)\|_{D(A^{\delta})}^2+\mathbb{E}\int_{T_{1,m}}^{T}\|\Phi_m(t)-\Phi_m(t,\lambda)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)}^2\,dt\right)$$
$$+7\,\mathcal{I}_2(\lambda)+7\,\mathcal{I}_3(\lambda)+7\,\mathcal{I}_5(\lambda)+7\,\mathcal{I}_6(\lambda),$$

where

$$K_{1,m}=\max\left\{C_1(T-T_{1,m})^{2-2\alpha-2\delta}+C_1(T-T_{1,m})^{2-2\alpha},C_2(T-T_{1,m})^{1-2\delta}+C_2(T-T_{1,m})\right\}.$$

We chose $T_{1,m}\in[0,T)$ such that $K_{1,m}<1$. Thus, we get for all $\lambda > 0$

$$\mathbb{E}\sup_{t\in[T_{1,m},T]}\|z_m^{*}(t)-z_m^{*}(t,\lambda)\|_{D(A^{\delta})}^2+\mathbb{E}\int_{T_{1,m}}^{T}\|\Phi_m(t)-\Phi_m(t,\lambda)\|_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)}^2\,dt$$
$$\le\frac{7\,\mathcal{I}_2(\lambda)+7\,\mathcal{I}_3(\lambda)+7\,\mathcal{I}_5(\lambda)+7\,\mathcal{I}_6(\lambda)}{1-K_{1,m}}.$$

Due to equation (6.46) and equation (6.49), we have

$$\lim_{\lambda\to\infty} \mathbb{E} \sup_{t\in[T_{1,m},T]} \|z_m^*(t) - z_m^*(t,\lambda)\|^2_{D(A^\delta)} = 0,$$

$$\lim_{\lambda\to\infty} \mathbb{E} \int_{T_{1,m}}^{T} \|\Phi_m(t) - \Phi_m(t,\lambda)\|^2_{\mathcal{L}_{(HS)}(Q^{1/2}(H);H)}\, dt = 0.$$

Similarly to Lemma 6.17, we can conclude that the result holds for the whole time interval $[0,T]$. □

6.6. Design of the Optimal Control

Based on the results provided in the previous sections, we are able to show a duality principle, which gives us a relation between the local mild solution of system (6.15) and the mild solution of system (6.34).

Theorem 6.37. *Let the processes $(y(t;u))_{t\in[0,\tau^u)}$ and $(z(t;u,v))_{t\in[0,\tau^u)}$ be the local mild solutions of system (6.3) and system (6.15), respectively, corresponding to the controls $u,v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Moreover, let the process $(z_m^*(t;u), \Phi_m(t;u))_{t\in[0,T]}$ be the mild solution of system (6.34) corresponding to the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Then we have for fixed $m\in\mathbb{N}$*

$$\mathbb{E} \int_0^{\tau_m^u} \langle A^\gamma(y(t;u) - y_d(t)), A^\gamma z(t;u,v)\rangle_H \, dt = \mathbb{E} \int_0^{\tau_m^u} \langle z_m^*(t;u), Fv(t)\rangle_H \, dt. \tag{6.50}$$

Proof. For the sake of simplicity, we omit the dependence on the controls. First, we prove the result for the approximations derived in Section 6.5. Let $(z_m(t,\lambda))_{t\in[0,T]}$ be the mild solution of system (6.38). Using Lemma 6.31, we have for all $\lambda > 0$, all $t\in[0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} z_m(t,\lambda) = &-\int_0^t A z_m(s,\lambda) + A^\delta R(\lambda) A^{-\delta}\left[B(R(\lambda) z_m(s,\lambda), \pi_m(y_m(s))) + B(\pi_m(y_m(s)), R(\lambda) z_m(s,\lambda))\right] ds \\ &+ \int_0^t R(\lambda) F v(s)\, ds + \int_0^t R(\lambda) G(R(\lambda) z_m(s,\lambda))\, dW(s). \end{aligned} \tag{6.51}$$

Next, let the pair of stochastic processes $(z_m^*(t,\lambda), \Phi_m(t,\lambda))_{t\in[0,T]}$ be the mild solution of system (6.43). By Lemma 6.35, we get for all $\lambda > 0$, all $t\in[0,T]$ and $\mathbb{P}$-a.s.

$$\begin{aligned} z_m^*(t,\lambda) = &-\int_t^T \mathbb{1}_{[0,\tau_m)}(s) \left[A z_m^*(s,\lambda) + A^\alpha R(\lambda) A^\alpha B_\delta^*\left(y(s\wedge\tau_m), R(\lambda) A^\delta z_m^*(s,\lambda)\right)\right] ds \\ &+ \int_t^T \mathbb{1}_{[0,\tau_m)}(s) R(\lambda) G^*(A^{-2\alpha} R(\lambda)\Phi_m(s,\lambda))\, ds + \int_t^T \mathbb{1}_{[0,\tau_m)}(s) A^\gamma R(\lambda) A^\gamma \left(y(s\wedge\tau_m) - y_d(s)\right) ds \\ &- \int_t^T \Phi_m(s,\lambda)\, dW(s). \end{aligned} \tag{6.52}$$

By definition, the process $(z_m^*(t,\lambda))_{t\in[0,T]}$ is predictable. Using Proposition 3.16, we find for all $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$
\begin{aligned}
&z_m^*(t,\lambda)\\
&= -\mathbb{E}\left[\int_0^T \mathbb{1}_{[0,\tau_m)}(s)\left[Az_m^*(s,\lambda) + A^\alpha R(\lambda) A^\alpha B_\delta^*\left(y(s\wedge\tau_m), R(\lambda)A^\delta z_m^*(s,\lambda)\right)\right] ds \middle| \mathcal{F}_t\right]\\
&\quad + \mathbb{E}\left[\int_0^T \mathbb{1}_{[0,\tau_m)}(s) R(\lambda) G^*(A^{-2\alpha}R(\lambda)\Phi_m(s,\lambda))\, ds + \int_0^T \mathbb{1}_{[0,\tau_m)}(s) A^\gamma R(\lambda) A^\gamma \left(y(s\wedge\tau_m) - y_d(s)\right) ds \middle| \mathcal{F}_t\right]\\
&\quad + \int_0^t \mathbb{1}_{[0,\tau_m)}(s)\left[Az_m^*(s,\lambda) + A^\alpha R(\lambda) A^\alpha B_\delta^*\left(y(s\wedge\tau_m), R(\lambda)A^\delta z_m^*(s,\lambda)\right)\right] ds\\
&\quad - \int_0^t \mathbb{1}_{[0,\tau_m)}(s) R(\lambda) G^*(A^{-2\alpha}R(\lambda)\Phi_m(s,\lambda))\, ds - \int_0^t \mathbb{1}_{[0,\tau_m)}(s) A^\gamma R(\lambda) A^\gamma \left(y(s\wedge\tau_m) - y_d(s)\right) ds.
\end{aligned}
$$

By Theorem 3.86 with $(M(t))_{t\in[0,T]}$ satisfying for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$
\begin{aligned}
&M(t)\\
&= -\mathbb{E}\left[\int_0^T \mathbb{1}_{[0,\tau_m)}(s)\left[Az_m^*(s,\lambda) + A^\alpha R(\lambda) A^\alpha B_\delta^*\left(y(s\wedge\tau_m), R(\lambda)A^\delta z_m^*(s,\lambda)\right)\right] ds \middle| \mathcal{F}_t\right]\\
&\quad + \mathbb{E}\left[\int_0^T \mathbb{1}_{[0,\tau_m)}(s) R(\lambda) G^*(A^{-2\alpha}R(\lambda)\Phi_m(s,\lambda))\, ds + \int_0^T \mathbb{1}_{[0,\tau_m)}(s) A^\gamma R(\lambda) A^\gamma \left(y(s\wedge\tau_m) - y_d(s)\right) ds \middle| \mathcal{F}_t\right],
\end{aligned}
$$

there exists a unique predictable process $(\Psi_m(t,\lambda))_{t\in[0,T]}$ with values in $\mathcal{L}_{(HS)}(Q^{1/2}(H);H)$ such that for all $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$
\begin{aligned}
&z_m^*(t,\lambda)\\
&= -\mathbb{E}\left[\int_0^T \mathbb{1}_{[0,\tau_m)}(s)\left[Az_m^*(s,\lambda) + A^\alpha R(\lambda) A^\alpha B_\delta^*\left(y(s\wedge\tau_m), R(\lambda)A^\delta z_m^*(s,\lambda)\right)\right] ds\right]\\
&\quad + \mathbb{E}\left[\int_0^T \mathbb{1}_{[0,\tau_m)}(s) R(\lambda) G^*(A^{-2\alpha}R(\lambda)\Phi_m(s,\lambda))\, ds + \int_0^T \mathbb{1}_{[0,\tau_m)}(s) A^\gamma R(\lambda) A^\gamma \left(y(s\wedge\tau_m) - y_d(s)\right) ds\right]\\
&\quad + \int_0^t \mathbb{1}_{[0,\tau_m)}(s)\left[Az_m^*(s,\lambda) + A^\alpha R(\lambda) A^\alpha B_\delta^*\left(y(s\wedge\tau_m), R(\lambda)A^\delta z_m^*(s,\lambda)\right)\right] ds\\
&\quad - \int_0^t \mathbb{1}_{[0,\tau_m)}(s) R(\lambda) G^*(A^{-2\alpha}R(\lambda)\Phi_m(s,\lambda))\, ds - \int_0^t \mathbb{1}_{[0,\tau_m)}(s) A^\gamma R(\lambda) A^\gamma \left(y(s\wedge\tau_m) - y_d(s)\right) ds\\
&\quad + \int_0^t \Psi_m(s,\lambda)\, dW(s). \qquad (6.53)
\end{aligned}
$$

Since the pair $(z_m^*(t,\lambda), \Phi_m(t,\lambda))_{t\in[0,T]}$ satisfies equation (6.52) uniquely, we have $\Psi_m(t,\lambda) = \Phi_m(t,\lambda)$ for all $\lambda > 0$, almost all $t \in [0,T]$ and $\mathbb{P}$-almost surely. Applying Corollary 3.69 to equation (6.51) and equation (6.53), we get for all $\lambda > 0$, all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\langle z_m(t,\lambda), z_m^*(t,\lambda)\rangle_H = \mathcal{I}_1(t,\lambda) + \mathcal{I}_2(t,\lambda) + \mathcal{I}_3(t,\lambda) + \mathcal{I}_4(t,\lambda) + \mathcal{I}_5(t,\lambda),$$

where

$$\mathcal{I}_1(t,\lambda) = \int_0^t \mathbb{1}_{[0,\tau_m)}(s)\,\langle z_m(s,\lambda), Az_m^*(s,\lambda)\rangle_H\,ds - \int_0^t \langle z_m^*(s,\lambda), Az_m(s,\lambda)\rangle_H\,ds,$$

$$\begin{aligned}\mathcal{I}_2(t,\lambda) = &\int_0^t \mathbb{1}_{[0,\tau_m)}(s)\left\langle z_m(s,\lambda), A^\alpha R(\lambda) A^\alpha B_\delta^*\left(y(s\wedge\tau_m), R(\lambda)A^\delta z_m^*(s,\lambda)\right)\right\rangle_H ds\\ &- \int_0^t \left\langle z_m^*(s,\lambda), A^\delta R(\lambda) A^{-\delta}\left[B(R(\lambda)z_m(s,\lambda), \pi_m(y_m(s))) + B(\pi_m(y_m(s)), R(\lambda)z_m(s,\lambda))\right]\right\rangle_H ds,\end{aligned}$$

$$\begin{aligned}\mathcal{I}_3(t,\lambda) = &\int_0^t \langle R(\lambda)G(R(\lambda)z_m(s,\lambda)), \Phi_m(s,\lambda)\rangle_{\mathcal{L}_{(HS)}(Q^{1/2}(H),H)}\,ds\\ &- \int_0^t \mathbb{1}_{[0,\tau_m)}(s)\left\langle z_m(s,\lambda), R(\lambda)G^*(A^{-2\alpha}R(\lambda)\Phi_m(s,\lambda))\right\rangle_H ds,\end{aligned}$$

$$\mathcal{I}_4(t,\lambda) = \int_0^t \langle z_m^*(s,\lambda), R(\lambda)Fv(s)\rangle_H\,ds - \int_0^t \mathbb{1}_{[0,\tau_m)}(s)\,\langle z_m(s,\lambda), A^\gamma R(\lambda)A^\gamma\left(y(s\wedge\tau_m) - y_d(s)\right)\rangle_H\,ds,$$

$$\mathcal{I}_5(t,\lambda) = \int_0^t \langle z_m(s,\lambda), \Phi_m(s,\lambda)\,dW(s)\rangle_H + \int_0^t \langle z_m^*(s,\lambda), R(\lambda)G(R(\lambda)z_m(s,\lambda))\,dW(s)\rangle_H\,.$$

By Lemma 6.34, we obtain for all $\lambda > 0$ and $\mathbb{P}$-a.s.

$$0 = \mathcal{I}_1(\tau_m,\lambda) + \mathcal{I}_2(\tau_m,\lambda) + \mathcal{I}_3(\tau_m,\lambda) + \mathcal{I}_4(\tau_m,\lambda) + \mathcal{I}_5(\tau_m,\lambda). \tag{6.54}$$

Since the operator A is self-adjoint, we have for all $\lambda > 0$ and $\mathbb{P}$-a.s.

$$\mathcal{I}_1(\tau_m,\lambda) = 0. \tag{6.55}$$

Recall that $y(t) = \pi_m(y_m(t))$ for all $t \in [0,\tau_m)$ and $\mathbb{P}$-almost surely. Using Lemma 2.34, Lemma 5.13 and equation (6.35), we find for all $\lambda > 0$ and $\mathbb{P}$-a.s.

$$\mathcal{I}_2(\tau_m,\lambda) = 0. \tag{6.56}$$

Due to Lemma 2.29 (iv), Lemma 2.34 and equation (6.36), we obtain for all $\lambda > 0$ and $\mathbb{P}$-a.s.

$$\mathcal{I}_3(\tau_m,\lambda) = 0. \tag{6.57}$$

By equations (6.54) - (6.57) and the fact that $\mathbb{E}\,I_5(\tau_m,\lambda) = 0$, we get for all $\lambda > 0$

$$0 = \mathbb{E}\,\mathcal{I}_4(\tau_m,\lambda).$$

Hence, we have for all $\lambda > 0$

$$\mathbb{E}\int_0^{\tau_m} \langle R(\lambda)A^\gamma z_m(t,\lambda), A^\gamma\left(y(t)-y_d(t)\right)\rangle_H\,dt = \mathbb{E}\int_0^{\tau_m} \langle R(\lambda)z_m^*(t,\lambda), Fv(t)\rangle_H\,dt. \tag{6.58}$$

Next, we show that the right hand side and the left hand side of equation (6.58) converges as $\lambda \to \infty$. Let $(y_m(t))_{t\in[0,T]}$ and $(z_m(t))_{t\in[0,T]}$ be the mild solutions of system (6.5) and system (6.16), respectively. By definition, we have for all $t \in [0,\tau_m)$ and $\mathbb{P}$-a.s. $y(t) = y_m(t)$, $\|y_m(t)\|_{D(A^\alpha)} \leq m$ and $z(t) = z_m(t)$. Using Lemma 6.32, we obtain

$$\lim_{\lambda\to\infty} \mathbb{E}\sup_{t\in[0,\tau_m)} \|z(t)-z_m(t,\lambda)\|^2_{D(A^\alpha)} = 0. \tag{6.59}$$

By the Cauchy-Schwarz inequality, Lemma 5.13 (ii) and Corollary 2.32, there exists a constant $C^* > 0$ such that for all $\lambda > 0$

$$\begin{aligned}
&\left|\mathbb{E}\int_0^{\tau_m} \langle A^\gamma z(t), A^\gamma\left(y(t)-y_d(t)\right)\rangle_H\,dt - \mathbb{E}\int_0^{\tau_m} \langle R(\lambda)A^\gamma z_m(t,\lambda), A^\gamma\left(y(t)-y_d(t)\right)\rangle_H\,dt\right|^2\\
&\leq 2\left|\mathbb{E}\int_0^{\tau_m} \langle [I-R(\lambda)]A^\gamma z(t), A^\gamma\left(y(t)-y_d(t)\right)\rangle_H\,dt\right|^2\\
&\quad + 2\left|\mathbb{E}\int_0^{\tau_m} \langle R(\lambda)A^\gamma(z(t)-z_m(t,\lambda)), A^\gamma\left(y(t)-y_d(t)\right)\rangle_H\,dt\right|^2\\
&\leq C^*\left(\mathbb{E}\int_0^{\tau_m} \|[I-R(\lambda)]A^\gamma z(t)\|_H^2\,dt + \mathbb{E}\sup_{t\in[0,\tau_m)} \|z(t)-z_m(t,\lambda)\|^2_{D(A^\alpha)}\right).
\end{aligned}$$

Using Lemma 5.13 (iii), equation (6.59) and Proposition B.7, we can conclude

$$\lim_{\lambda\to\infty}\mathbb{E}\int_0^{\tau_m} \langle R(\lambda)A^\gamma z_m(t,\lambda), A^\gamma\left(y(t)-y_d(t)\right)\rangle_H\,dt = \mathbb{E}\int_0^{\tau_m} \langle A^\gamma z(t), A^\gamma\left(y(t)-y_d(t)\right)\rangle_H\,dt.$$

Recall that the operator $F\colon D(A^\beta) \to D(A^\beta)$ is bounded. Similarly as above, there exist constants $C^* > 0$ such that for all $\lambda > 0$

$$\begin{aligned}
&\left|\mathbb{E}\int_0^{\tau_m} \langle z_m^*(t), Fv(t)\rangle_H\,dt - \mathbb{E}\int_0^{\tau_m} \langle R(\lambda)z_m^*(t,\lambda), Fv(t)\rangle_H\,dt\right|^2\\
&\leq 2\left|\mathbb{E}\int_0^{\tau_m} \langle [I-R(\lambda)]z_m^*(t), Fv(t)\rangle_H\,dt\right|^2 + 2\left|\mathbb{E}\int_0^{\tau_m} \langle R(\lambda)(z_m^*(t)-z_m^*(t,\lambda)), Fv(t)\rangle_H\,dt\right|^2\\
&\leq C^*\left(\mathbb{E}\int_0^{T} \|[I-R(\lambda)]z_m^*(t)\|_H^2\,dt + \mathbb{E}\sup_{t\in[0,T]} \|z_m^*(t)-z_m^*(t,\lambda)\|^2_{D(A^\delta)}\right).
\end{aligned}$$

By Lemma 5.13 (iii), Lemma 6.36 and Proposition B.7, we can infer

$$\lim_{\lambda\to\infty}\mathbb{E}\int_0^{\tau_m} \langle R(\lambda)z_m^*(t,\lambda), Fv(t)\rangle_H\,dt = \mathbb{E}\int_0^{\tau_m} \langle z_m^*(t), Fv(t)\rangle_H\,dt.$$

We conclude that the right hand side and the left hand side of equation (6.58) converges as $\lambda \to \infty$ and equation (6.50) holds. □

Based on the necessary optimality condition formulated as the variational inequality (6.32) and the duality principle derived in the previous theorem, we are able to deduce a formula the optimal control has to satisfy.

Theorem 6.38. *Let $(z_m^*(t;u), \Phi_m(t;u))_{t\in[0,T]}$ be the mild solution of system (6.34) corresponding to the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Then for fixed $m \in \mathbb{N}$, the optimal control $\overline{u}_m \in U$ satisfies for almost all $t \in [0,T]$ and $\mathbb{P}$-a.s.*

$$\overline{u}_m(t) = -P_U\left(F^* A^{-2\beta} z_m^*(t;\overline{u}_m)\right), \tag{6.60}$$

where $P_U \colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to U$ is the projection onto U and $F^* \in \mathcal{L}(D(A^\beta))$ is the adjoint operator of $F \in \mathcal{L}(D(A^\beta))$.

Proof. Using inequality (6.32) and Theorem 6.37, the optimal control $\overline{u}_m \in U$ satisfies for every $u \in U$

$$\mathbb{E}\int_0^{\tau_m^{\overline{u}_m}} \langle z_m^*(t;\overline{u}_m), F(u(t)-\overline{u}_m(t))\rangle_H \, dt + \mathbb{E}\int_0^T \left\langle A^\beta \overline{u}_m(t), A^\beta(u(t)-\overline{u}_m(t))\right\rangle_H dt \geq 0.$$

By Lemma 6.29, we have $\mathbb{1}_{[0,\tau_m^{\overline{u}_m})}(t) z_m^*(t;\overline{u}_m) = z_m^*(t;\overline{u}_m)$ for all $t \in [0,T]$ and $\mathbb{P}$-almost surely. Due to Lemma 2.29 (iv), Lemma 2.34, we obtain for every $u \in U$

$$\begin{aligned}
\mathbb{E}\int_0^{\tau_m^{\overline{u}_m}} \langle z_m^*(t;\overline{u}_m), F(u(t)-\overline{u}_m(t))\rangle_H \, dt &= \mathbb{E}\int_0^T \left\langle \mathbb{1}_{[0,\tau_m^{\overline{u}_m})}(t) z_m^*(t;\overline{u}_m), F(u(t)-\overline{u}_m(t))\right\rangle_H dt \\
&= \mathbb{E}\int_0^T \left\langle A^\beta A^{-2\beta} z_m^*(t;\overline{u}_m), A^\beta F(u(t)-\overline{u}_m(t))\right\rangle_H dt \\
&= \mathbb{E}\int_0^T \left\langle A^\beta F^* A^{-2\beta} z_m^*(t;\overline{u}_m), A^\beta(u(t)-\overline{u}_m(t))\right\rangle_H dt.
\end{aligned}$$

Hence, we find for every $u \in U$

$$\mathbb{E}\int_0^T \left\langle -F^* A^{-2\beta} z_m^*(t;\overline{u}_m) - \overline{u}_m(t), u(t)-\overline{u}_m(t)\right\rangle_{D(A^\beta)} dt \leq 0.$$

By Proposition D.21, we obtain equation (6.60). We note that the mild solution of system (6.34) is a pair of predictable processes $(z_m^*(t;u), \Phi_m(t;u))_{t\in[0,T]}$ such that especially $\mathbb{E}\sup_{t\in[0,T]} \|z_m^*(t;u)\|^2_{D(A^\delta)} < \infty$ for every $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Hence, we get $F^* A^{-2\beta} z_m^*(\cdot;\overline{u}_m) \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, which justifies the application of the projection operator P_U. □

Remark 6.39. *Let us denote by $(\overline{y}(t))_{t\in[0,\overline{\tau})}$ and $(\overline{z}_m^*(t), \overline{\Phi}_m(t))_{t\in[0,T]}$ the local mild solutions of system (6.3) and the mild solution of system (6.34), respectively, corresponding to the optimal control $\overline{u}_m \in U$. As a consequence of the previous theorem, the optimal velocity field $(\overline{y}(t))_{t\in[0,,\overline{\tau})}$ can be computed by solving the following system of coupled forward-backward SPDEs:*

$$\begin{cases}
d\overline{y}(t) = -[A\overline{y}(t) + B(\overline{y}(t)) + FP_U\left(F^*A^{-2\beta}\overline{z}_m^*(t)\right)]\,dt + G(\overline{y}(t))\,dW(t), \\
d\overline{z}_m^*(t) = -\mathbb{1}_{[0,\tau_m)}(t)[-A\overline{z}_m^*(t) - A^{2\alpha}B_\delta^*\left(\overline{y}(t), A^\delta \overline{z}_m^*(t)\right) + G^*(A^{-2\alpha}\overline{\Phi}_m(t)) + A^{2\gamma}\left(\overline{y}(t) - y_d(t)\right)]\,dt \\
\qquad\qquad + \overline{\Phi}_m(t)\,dW(t), \\
\overline{y}(0) = \xi, \quad \overline{z}_m^*(T) = 0.
\end{cases}$$

Corollary 6.40. *Let the control $\overline{u}_m \in U$ be given by equation (6.60). Then we have for fixed $m \in \mathbb{N}$*

$$\mathbb{E} \int_{\tau_m^{\overline{u}_m}}^{T} \|\overline{u}_m(t)\|^2_{D(A^\beta)} dt = 0.$$

Proof. Let $(z_m^*(t;\overline{u}_m), \Phi_m(t;\overline{u}_m))_{t\in[0,T]}$ be the mild solution of system (6.34) corresponding to the optimal control $\overline{u}_m \in U$. By Lemma 6.29, we have $\mathbb{E}\sup_{t\in[\tau_m^{\overline{u}_m},T]} \|z_m^*(t;\overline{u}_m)\|^2_{D(A^\delta)} = 0$. Moreover, note that the operators in equation (6.60) are linear and bounded. Using Corollary 2.32, there exists a constant $C^* > 0$ such that

$$\mathbb{E} \int_{\tau_m^{\overline{u}_m}}^{T} \|\overline{u}_m(t)\|^2_{D(A^\beta)} dt = \mathbb{E} \int_{\tau_m^{\overline{u}_m}}^{T} \left\| P_U \left(F^* A^{-2\beta} z_m^*(t;\overline{u}_m)\right)\right\|^2_{D(A^\beta)} dt \le C^* \, \mathbb{E} \sup_{t\in[\tau_m^{\overline{u}_m},T]} \|z_m^*(t;\overline{u}_m)\|^2_{D(A^\delta)} = 0.$$

□

Finally, we show that the optimal control $\overline{u}_m \in U$ given by equation (6.60) satisfies a sufficient optimality condition.

Theorem 6.41. *Let $\overline{u}_m \in U$ be given by equation (6.60). Then $\overline{u}_m \in U$ is an optimal control of the control problem (6.14).*

Proof. Note that the set U is a convex subset of the Hilbert space $L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. By Corollary 6.26, the cost functional J_m given by equation (6.13) is twice continuous Fréchet differentiable. Recall that $\overline{u}_m \in U$ satisfies the necessary optimality condition (6.31), which are also valid for the Fréchet derivative due to Theorem 6.22 and Corollary 6.25. Moreover, we have for every $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$

$$d^F(J_m(\overline{u}_m))^2[v,v] = \mathbb{E} \int_0^{\tau_m^{\overline{u}_m}} \|A^\gamma z(t;\overline{u}_m, v)\|^2_H \, dt + \mathbb{E} \int_0^T \|A^\beta v(t)\|^2_H dt \ge \mathbb{E} \int_0^T \|v(t)\|^2_{D(A^\beta)} dt.$$

Hence, the assumptions of Proposition D.22 are fulfilled and the optimal control $\overline{u}_m \in U$ given by equation (6.60) is a local minimum of the cost functional J_m. Due to Theorem 6.15, we can conclude that this minimum is also global. □

6.7. The Case of Additive Noise

As described in Remark 6.28, the control problem (6.14) for $\gamma = \frac{1}{2}$ can not be solved through a stochastic maximum principle directly. Here, we give a possible simplification, which enables us to overcome this problem.

We introduce the stochastic Navier-Stokes equations with additive noise in $D(A^\alpha)$:

$$\begin{cases} dy(t) = -[Ay(t) + B(y(t)) - Fu(t)]\, dt + G\, dW(t), \\ y(0) = \xi, \end{cases} \tag{6.61}$$

where $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, $\xi \in L^2(\Omega; D(A^\alpha))$ is $\mathcal{F}_0$ measurable and the process $(W(t))_{t\ge 0}$ is a Q-Wiener process with values in H and covariance operator $Q \in \mathcal{L}_1^+(H)$. The operators A, B, F are introduced in Section 6.1 and Section 6.2, respectively. Moreover, we assume that $G \in \mathcal{L}_{(HS)}(Q^{1/2}(H); D(A^\alpha))$. Note that system (6.61) is a special case of system (6.3). Hence, the existence and uniqueness of a local mild solution to system (6.61) is an immediate consequence of Theorem 6.7. We denote by $(y(t;u))_{t\in[0,\tau^u)}$ the local mild solution of system (6.61) to illustrate the dependence on the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$.

Remark 6.42. *Similarly to Section 6.2, one first considers the following system in $D(A^\alpha)$:*

$$\begin{cases} dy_m(t) = -[Ay_m(t) + B(\pi_m(y_m(t))) - Fu(t)]\,dt + G\,dW(t), \\ y_m(0) = \xi, \end{cases} \tag{6.62}$$

where $m \in \mathbb{N}$ and $\pi_m\colon D(A^\alpha) \to D(A^\alpha)$ is given by equation (6.6). If the parameters $\alpha \in (0,1)$ and $\delta \in [0,1)$ and $\beta \in [0,\alpha]$ satisfy the assumptions of Theorem 6.6, then there exists a unique mild solution of system (6.62) for fixed $m \in \mathbb{N}$, fixed $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ and any $\xi \in L^2(\Omega; D(A^\alpha))$ in the sense of Definition 3.81. Similarly to the proof of Theorem 6.7, the sequence of stopping times $(\tau_m)_{m\in\mathbb{N}}$ is defined by

$$\tau_m = \inf\{t \in (0,T) : \|y_m(t)\|_{D(A^\alpha)} > m\} \wedge T \tag{6.63}$$

$\mathbb{P}$-a.s. and the stopping time τ is given by $\tau = \lim_{m\to\infty} \tau_m$.

Following Section 6.3, we consider again the cost functional $J_m\colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to \mathbb{R}$ given by

$$J_m(u) = \frac{1}{2}\,\mathbb{E}\int_0^{\tau_m^u} \|A^\gamma(y(t;u) - y_d(t))\|_H^2\,dt + \frac{1}{2}\,\mathbb{E}\int_0^T \|u(t)\|^2_{D(A^\beta)} dt, \tag{6.64}$$

where $m \in \mathbb{N}$ is fixed and $\gamma \in [0,\alpha]$. Moreover, the process $(y(t;u))_{t\in[0,\tau^u)}$ is the local mild solution of system (6.61) corresponding to the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ and $y_d \in L^2([0,T]; D(A^\gamma))$ is a given desired velocity field. The task is to find a control $\overline{u}_m \in U$ such that

$$J_m(\overline{u}_m) = \inf_{u\in U} J_m(u),$$

where the set of admissible controls U is a closed, bounded and convex subset of the Hilbert space $L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ such that $0 \in U$. The control $\overline{u}_m \in U$ is called an optimal control. The existence and uniqueness of the optimal control follows from Theorem 6.15.

Next, we state the necessary optimality condition. Similarly to Section 6.3, we first introduce the following linearized system in $D(A^\alpha)$:

$$\begin{cases} dz(t) = -[Az(t) + B(z(t), y(t)) + B(y(t), z(t)) - Fv(t)]\,dt, \\ z(0) = 0, \end{cases} \tag{6.65}$$

where $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$, the process $(y(t))_{t\in[0,\tau)}$ is the local mild solution of system (6.61). The operators A, B, F are introduced in Section 6.1 and Section 6.2, respectively. Note that system (6.65) is a special case of system (6.15). Hence, the existence and uniqueness of a local mild solution to system (6.65) follows immediately. We denote by $(z(t;u,v))_{t\in[0,\tau^u)}$ the local mild solution of system (6.61) to illustrate the dependence on the control $u, v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Furthermore, for fixed $m \in \mathbb{N}$, the Gâteaux derivative of $y(t;u)$ at $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ in direction $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ satisfies for all $t \in [0, \tau_m^u)$ and $\mathbb{P}$-a.s.

$$d_u^G y(t;u)[v] = z(t;u,v),$$

which can be obtained similarly to Theorem 6.21. Therefore, the Gâteaux derivative of the cost functional $J_m\colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to \mathbb{R}$ given by (6.64) at $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ in direction $v \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$ satisfies

$$d^G J_m(u)[v] = \mathbb{E}\int_0^{\tau_m^u} \langle A^\gamma(y(t;u) - y_d(t)), A^\gamma z(t;u,v)\rangle_H\,dt + \mathbb{E}\int_0^T \left\langle A^\beta u(t), A^\beta v(t)\right\rangle_H dt.$$

This result follows immediately from Theorem 6.22. The Gâteaux derivative of order two as well as the Fréchet derivatives can be obtained from Corollaries 6.24 - 6.26. According to inequality (6.31), the optimal control $\overline{u}_m \in U$ satisfies the following necessary optimality condition for fixed $m \in \mathbb{N}$ and every $u \in U$:

$$d^G J_m(\overline{u}_m)[u - \overline{u}_m] \geq 0.$$

To utilize this necessary optimality condition, we first introduce the adjoint equation in $D(A^\delta)$:

$$\begin{cases} dz_m^*(t) = -\mathbb{1}_{[0,\tau_m)}(t)[-Az_m^*(t) - A^{2\alpha} B_\delta^* \left(y(t), A^\delta z_m^*(t)\right) + A^{2\gamma}(y(t) - y_d(t))]\, dt, \\ z_m^*(T) = 0, \end{cases} \tag{6.66}$$

where $m \in \mathbb{N}$ and the process $(y(t))_{t\in[0,\tau)}$ is the local mild solution of system (6.61). The sequence of stopping times $(\tau_m)_{m\in\mathbb{N}}$ is defined by equation (6.63) and $y_d \in L^2([0,T]; D(A^\gamma))$ is the given desired velocity field. The operator A and its fractional powers are introduced in Section 6.1. Moreover, the operator $B_\delta^*(y(t), \cdot) : H \to D(A^\alpha)$ for $t \in [0, \tau_m)$ is linear and bounded. A precise meaning is given by Remark 6.27 (i). The existence and uniqueness of a mild solution to system (6.66) can be obtained using the Banach fixed point theorem without additional restrictions on the parameters $\alpha \in (0, 1)$ and $\delta \in [0, 1)$ and $\beta, \gamma \in [0, \alpha]$ if we require that $y_d \in L^\infty([0,T]; D(A^\gamma))$. Let us denote by $(z_m^*(t;u))_{t\in[0,T]}$ the mild solution of system (6.66) to illustrate the dependence on the control $u \in L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta)))$. Note that the process $(z_m^*(t;u))_{t\in[0,T]}$ is not $\mathcal{F}_t$-adapted. Furthermore, we can easily derive a duality principle similarly to Theorem 6.37. As a consequence, for fixed $m \in \mathbb{N}$, the optimal control $\overline{u}_m \in U$ satisfies for almost all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\overline{u}_m(t) = -P_U \left(F^* A^{-2\beta} \mathbb{E}\left[z_m^*(t; \overline{u}_m) | \mathcal{F}_t\right]\right), \tag{6.67}$$

where $P_U \colon L^2_{\mathcal{F}}(\Omega; L^2([0,T]; D(A^\beta))) \to U$ is the projection onto U and $F^* \in \mathcal{L}(D(A^\beta))$ is the adjoint operator of $F \in \mathcal{L}(D(A^\beta))$. This result follows immediately from Theorem 6.38. As described in Section 6.6, the optimal control $\overline{u}_m \in U$ given by equation (6.67) satisfies also a sufficient optimality condition.

Remark 6.43. *Note that we can easily generalize these results if system (6.3) is driven by a square integrable Lévy martingale as introduced in Section 3.3.*

Appendix

A. Some Gronwall-type Inequalities

In this section, we state the Gronwall inequality and their modifications. Let $T > 0$. We start with the classical version.

Proposition A.1 (Corollary 6.60, [69])**.** *Let $a\colon [0,T] \to [0,\infty)$ be an increasing function and let the functions $x, b\colon [0,T] \to \mathbb{R}$ be integrable such that $b(t) \geq 0$ for almost all $t \in [0,T]$ and*

$$\int_0^T b(t)|x(t)|\, dt < \infty.$$

If

$$x(t) \leq a(t) + \int_0^t b(s)x(s)\, ds$$

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

$$x(t) \leq a(t) \exp\left\{ \int_0^t b(s)\, ds \right\}$$

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

As a consequence of the previous Proposition, we get the following Gronwall inequality of backward type. A proof for Stieltjes integrals can be found in [69, Corollary 6.61].

Corollary A.2. *Let $a\colon [0,T] \to [0,\infty)$ be a decreasing function and let the functions $x, b\colon [0,T] \to \mathbb{R}$ be integrable such that $b(t) \geq 0$ for almost all $t \in [0,T]$ and*

$$\int_0^T b(t)|x(t)|\, dt < \infty.$$

If

$$x(t) \leq a(t) + \int_t^T b(s)x(s)\, ds$$

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

$$x(t) \leq a(t) \exp\left\{ \int_t^T b(s)\, ds \right\}$$

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

Proof. By a change of variables, we get for $t \in [0,T]$

$$x(T-t) \leq a(T-t) + \int_0^t b(T-s)x(T-s)\,ds.$$

We set $\tilde{x}(t) = x(T-t)$, $\tilde{a}(t) = a(T-t)$ and $\tilde{b}(t) = b(T-t)$ for $t \in [0,T]$. Applying Proposition A.1, we obtain for $t \in [0,T]$

$$\tilde{x}(t) \leq \tilde{a}(t) \exp\left\{ \int_0^t \tilde{b}(s)\,ds \right\}.$$

Thus, we have for $t \in [0,T]$

$$x(T-t) \leq a(T-t) \exp\left\{ \int_0^t b(T-s)\,ds \right\} = a(T-t) \exp\left\{ \int_{T-t}^T b(s)\,ds \right\}.$$

Therefore, the claim follows. □

The following inequality is applicable for nonmonotonic functions.

Proposition A.3 (Theorem 1, [90])**.** *Let $a, x\colon [0,T] \to [0,\infty)$ be integrable functions such that $a(t), x(t) \geq 0$ for $t \in [0,T]$ and let the function $b\colon [0,T] \to [0,\infty)$ be nondecreasing and continuous such that $0 \leq b(t) \leq M$ for all $t \in [0,T]$. Suppose that $\beta > 0$. If*

$$x(t) \leq a(t) + b(t) \int_0^t (t-s)^{\beta-1} x(s)\,ds$$

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

$$x(t) \leq a(t) + \int_0^t \left[\sum_{n=1}^{\infty} \frac{(b(t)\Gamma(\beta))^n}{\Gamma(n\beta)} (t-s)^{n\beta-1} a(s) \right] ds$$

for $t \in [0,T]$, where $\Gamma(\cdot)$ is the gamma function.

The previous result has some useful consequences.

Corollary A.4. *Let $a, x\colon [0,T] \to [0,\infty)$ be integrable functions such that $a(t), x(t) \geq 0$ for $t \in [0,T]$ and let $b \geq 0$. If*

$$x(t) \leq a(t) + b \int_0^t x(s)\,ds$$

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

$$x(t) \leq a(t) + b \int_0^t e^{b(t-s)} a(s)\,ds$$

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

Proof. Using Proposition A.3 with $b(t) = b$ for all $t \in [0, T]$ and $\beta = 1$, we get for $t \in [0, T]$

$$x(t) \leq a(t) + \int_0^t \left[\sum_{n=1}^{\infty} \frac{(b\Gamma(1))^n}{\Gamma(n)} (t-s)^{n-1} a(s) \right] ds.$$

Since $\Gamma(n) = (n-1)!$ for each $n \in \mathbb{N}$, the claim follows. □

Corollary A.5. *Let $a, x \colon [0, T] \to [0, \infty)$ be integrable functions such that $a(t), x(t) \geq 0$ for $t \in [0, T]$ and let $b \geq 0$. If*

$$x(t) \leq a(t) + b \int_t^T x(s)\, ds$$

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

$$x(t) \leq a(t) + b \int_t^T e^{b(s-t)} a(s)\, ds$$

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

Proof. The claim can be obtained similarly to Corollary A.2. □

B. The Bochner Integral

Here, we introduce the Bochner integral and we will state some basic properties. For more details, we refer to [27, 58]. Throughout this section, let X be a Banach space and let (Ω, Σ, μ) be a measure space with finite measure μ.

Let $f \colon \Omega \to X$ be a simple function, i.e.

$$f = \sum_{k=1}^{n} x_k \mathbb{1}_{A_k},$$

where $(A_k)_{k=1,\dots,n} \subset \Sigma$ is a partition of Ω and $(x_k)_{k=1,\dots,n} \subset X$. Then the Bochner integral is defined by

$$\int_\Omega f\, d\mu = \sum_{k=1}^{n} x_k\, \mu(A_k)$$

and for every $A \in \Sigma$, we set

$$\int_A f\, d\mu = \int_\Omega \mathbb{1}_A f\, d\mu.$$

Next, we extend the definition of the Bochner integral.

Definition B.1. *a) A function $f \colon \Omega \to X$ is called (strongly)* ***measurable*** *if there exists a sequence of simple functions $(f_n)_{n\in\mathbb{N}}$ such that μ-a.e. $\lim_{n\to\infty} \|f - f_n\|_X = 0$.*
b) We call $f \colon \Omega \to X$ ***weakly measurable*** *if for every $x' \in X'$, the real-valued function $\omega \mapsto \langle x', f(\omega)\rangle$ is measurable.*

On separable Banach spaces, we have the following equivalences, which are consequences of Pettis measurability theorem.

Proposition B.2 (Corollary 3.10.5,[27])**.** *Let X be a separable Banach space and let $f\colon \Omega \to X$. Then the following conditions are equivalent:*

(i) f is measurable;

(ii) f is weakly measurable;

(iii) for every open (or closed) set A in X, we have $f^{-1}(A) \in \Sigma$.

Remark B.3. *Whenever we are dealing with separable Banach spaces, we use property (iii) of the previous Proposition to characterize that a function is measurable.*

Definition B.4. *A measurable function $f\colon \Omega \to X$ is called **Bochner integrable** if there exists a sequence of simple functions $(f_n)_{n\in\mathbb{N}}$ such that*

$$\lim_{n\to\infty} \int_\Omega \|f - f_n\|_X \, d\mu = 0.$$

*We call the sequence $(f_n)_{n\in\mathbb{N}}$ an **approximating sequence**.*

For every Bochner integrable function $f\colon \Omega \to X$ with approximating sequence $(f_n)_{n\in\mathbb{N}}$, we define for every $A \in \Sigma$

$$\int_A f \, d\mu = \lim_{n\to\infty} \int_A f_n \, d\mu.$$

The definition is independent of the approximating sequence. Furthermore, a characterization of Bochner integrable functions is given as follows.

Theorem B.5 (Theorem 3.10.9, [27])**.** *A measurable function $f\colon \Omega \to X$ is Bochner integrable if and only if $\int_\Omega \|f\|_X \, d\mu < \infty$.*

Corollary B.6 (Corollary 3.10.10 and Corollary 3.10.11, [27])**.** *If $f\colon \Omega \to X$ is Bochner integrable, then*

$$\left\| \int_\Omega f \, d\mu \right\|_X \leq \int_\Omega \|f\|_X \, d\mu$$

and the Bochner integral of f is absolutely continuous with respect to the measure μ, i.e.

$$\lim_{\mu(A)\to 0} \int_A f \, d\mu = 0.$$

We denote by $L^1(\Omega, \Sigma, \mu)$ the space of integrable real-valued functions with respect to the measure space (Ω, Σ, μ). Then we get the following dominated convergence theorem for Bochner integrals.

Proposition B.7 (Theorem 3.10.12, [27])**.** *Let $f_n\colon \Omega \to X$, $n \in \mathbb{N}$, be Bochner integrable functions and $f\colon \Omega \to X$ such that $\lim_{n\to\infty} \mu\{\|f_n - f\|_X > \varepsilon\} = 0$ for all $\varepsilon > 0$ and there exists $g \in L^1(\Omega, \Sigma, \mu)$ such that μ-a.e. $\|f_n\|_X \leq g$ for each $n \in \mathbb{N}$. Then $f\colon \Omega \to X$ is Bochner integrable and*

$$\lim_{n\to\infty} \int_\Omega f_n \, d\mu = \int_\Omega f \, d\mu.$$

In this context, we also recall the monotone convergence theorem for nonnegative real valued functions.

Proposition B.8 (Theorem 2.2.6, [27])**.** *Let $f_n\colon \Omega \to [0, +\infty]$, $n \in \mathbb{N}$, be an increasing sequence of measurable functions converging to a function $f\colon \Omega \to [0, +\infty]$. Then f is measurable and*

$$\lim_{n\to\infty} \int_\Omega f_n \, d\mu = \int_\Omega f \, d\mu.$$

The following property is useful, when dealing with closed linear operators $A\colon D(A) \subset X \to Y$, where Y is another Banach space.

Proposition B.9 (Theorem 3.10.16, [27]). *If $f\colon \Omega \to X$ and $Af\colon \Omega \to Y$ are Bochner integrable, then*

$$A \int_\Omega f \, d\mu = \int_\Omega Af \, d\mu.$$

Finally, we introduce L^p-spaces for Banach space valued functions. For $p \in [1, \infty)$, we set

$$\mathcal{L}^p(\Omega; X) = \left\{ f\colon \Omega \to X\colon f \text{ is Bochner integrable such that } \int_\Omega \|f\|_X^p d\mu < \infty \right\}$$

and

$$\mathcal{L}^\infty(\Omega; X) = \{ f\colon \Omega \to X\colon f \text{ is measurable and there exists } M > 0 \text{ such that } \mu\text{-a.e. } \|f\|_X \leq M \}.$$

Moreover, let $\mathcal{N} = \{ f\colon \Omega \to X\colon \mu\text{-a.e. } f = 0\}$. We define the L^p-space as quotient spaces as follows:

$$L^p(\Omega; X) = \mathcal{L}^p(\Omega; X)/\mathcal{N} \text{ for } p \in [1, \infty), \qquad L^\infty(\Omega; X) = \mathcal{L}^\infty(\Omega; X)/\mathcal{N}.$$

If we equip the space $L^p(\Omega; X)$ for $p \in [1, \infty)$ with the norm

$$\|f\|_{L^p(\Omega;X)} = \left(\int_\Omega \|f\|_X^p d\mu \right)^{1/p},$$

then $L^p(\Omega; X)$ becomes a Banach space. Similarly, if we equip the space $L^\infty(\Omega; X)$ with the norm

$$\|f\|_{L^\infty(\Omega;X)} = \inf\{M > 0\colon \mu\{\|f\|_X > M\} = 0\},$$

then $L^\infty(\Omega; X)$ becomes a Banach space. Furthermore, simple functions are dense in $L^p(\Omega; X)$ for $p \in [1, \infty)$.

C. Nuclear and Hilbert-Schmidt Operators

In this section, we state some basic facts of linear and bounded operators on Hilbert spaces. We will mainly focus on nuclear operators and Hilbert-Schmidt operators. Here, we closely follow [23, 45, 71, 73]. Let $\mathcal{U}$ and $\mathcal{H}$ be two separable Hilbert spaces.

The space of all linear and bounded (or continuous) operators is denoted by $\mathcal{L}(\mathcal{U}; \mathcal{H})$. Then $\mathcal{L}(\mathcal{U}; \mathcal{H})$ equipped with the operator norm

$$\|T\|_{\mathcal{L}(\mathcal{U};\mathcal{H})} = \sup_{x \in \mathcal{U}, x \neq 0} \frac{\|Tx\|_\mathcal{H}}{\|x\|_\mathcal{U}}$$

for every $T \in \mathcal{L}(\mathcal{U}; \mathcal{H})$ becomes a Banach space. For the sake of simplicity, we set $\mathcal{L}(\mathcal{U}) = \mathcal{L}(\mathcal{U}; \mathcal{U})$.

Remark C.1. *Note that we can define linear and bounded operators even if $\mathcal{U}$ and $\mathcal{H}$ are Banach spaces.*

The adjoint operator of $T \in \mathcal{L}(\mathcal{U}; \mathcal{H})$ is denoted by $T^* \in \mathcal{L}(\mathcal{H}; \mathcal{U})$. It is uniquely determined by the following equation for every $x \in \mathcal{U}$ and every $y \in \mathcal{H}$:

$$\langle Tx, y \rangle_\mathcal{H} = \langle x, T^* y \rangle_\mathcal{U}.$$

We call an operator $T \in \mathcal{L}(\mathcal{U})$ self-adjoint if $T = T^*$. The operator $T \in \mathcal{L}(\mathcal{U})$ is nonnegative (semidefinite) if $\langle Tx, x \rangle_\mathcal{U} \geq 0$ for every $x \in \mathcal{U}$.

Definition C.2. *An operator $T \in \mathcal{L}(\mathcal{U};\mathcal{H})$ is called **nuclear** or **trace class** if it has the representation*

$$Tx = \sum_{n=1}^{\infty} a_n \langle b_n, x\rangle_{\mathcal{U}}$$

for every $x \in \mathcal{U}$, where the sequences $(a_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ and $(b_n)_{n\in\mathbb{N}} \subset \mathcal{U}$ satisfy $\sum_{n=1}^{\infty} \|a_n\|_{\mathcal{H}}\|b_n\|_{\mathcal{U}} < \infty$.

The space of all nuclear operators is denoted by $\mathcal{L}_1(\mathcal{U};\mathcal{H})$. Similarly as above, we set $\mathcal{L}_1(\mathcal{U}) = \mathcal{L}_1(\mathcal{U};\mathcal{U})$. The space $\mathcal{L}_1(\mathcal{U};\mathcal{H})$ equipped with the nuclear norm

$$\|T\|_{\mathcal{L}_1(\mathcal{U};\mathcal{H})} = \inf\left\{\sum_{n=1}^{\infty} \|a_n\|_{\mathcal{H}}\|b_n\|_{\mathcal{U}} \colon Tx = \sum_{n=1}^{\infty} a_n\langle b_n, x\rangle_{\mathcal{U}}\right\}$$

for every $T \in \mathcal{L}_1(\mathcal{U};\mathcal{H})$ becomes a separable Banach space. We get the following basic properties.

Proposition C.3 (Proposition A.4, [71])**.** *Let $\mathcal{V}$ be another separable Hilbert space.*

(i) If $S \in \mathcal{L}_1(\mathcal{U};\mathcal{H})$ and $T \in \mathcal{L}(\mathcal{H};\mathcal{V})$, then $TS \in \mathcal{L}_1(\mathcal{U};\mathcal{V})$ and

$$\|TS\|_{\mathcal{L}_1(\mathcal{U};\mathcal{V})} \le \|S\|_{\mathcal{L}_1(\mathcal{U};\mathcal{H})}\|T\|_{\mathcal{L}(\mathcal{H};\mathcal{V})}.$$

(ii) If $S \in \mathcal{L}(\mathcal{U};\mathcal{H})$ and $T \in \mathcal{L}_1(\mathcal{H};\mathcal{V})$, then $TS \in \mathcal{L}_1(\mathcal{U};\mathcal{V})$ and

$$\|TS\|_{\mathcal{L}_1(\mathcal{U};\mathcal{V})} \le \|S\|_{\mathcal{L}(\mathcal{U};\mathcal{H})}\|T\|_{\mathcal{L}_1(\mathcal{H};\mathcal{V})}.$$

For $T \in \mathcal{L}_1(\mathcal{U})$, we can introduce the trace of T by

$$Tr(T) = \sum_{n=1}^{\infty} \langle Tu_n, u_n\rangle_{\mathcal{U}},$$

where $(u_n)_{n\in\mathbb{N}}$ is an orthonormal basis of $\mathcal{U}$.

Proposition C.4 (Remark B.0.4, [73])**.** *If $T \in \mathcal{L}_1(\mathcal{U})$, then $Tr(T)$ is well defined and independent on the choice of the orthonormal basis. Moreover, we have*

$$|Tr(T)| \le \|T\|_{\mathcal{L}_1(\mathcal{U})}.$$

We denote by $\mathcal{L}_1^+(\mathcal{U})$ the subspace of $\mathcal{L}_1(\mathcal{U})$ containing all self-adjoint nonnegative nuclear operators. We have the following result, which is especially valid for all $T \in \mathcal{L}_1^+(\mathcal{U})$.

Proposition C.5 (Proposition 2.1.5, [73])**.** *If $T \in \mathcal{L}(\mathcal{U})$ is a self-adjoint nonnegative operator such that $Tr(T) < \infty$, then there exist an orthonormal basis $(u_n)_{n\in\mathbb{N}}$ of $\mathcal{U}$ and a sequence $(\lambda_n)_{n\in\mathbb{N}}$ of nonnegative real numbers such that for each $n \in \mathbb{N}$*

$$Tu_n = \lambda_n u_n$$

and 0 is the only accumulation point of $(\lambda_n)_{n\in\mathbb{N}}$.

Definition C.6. *An operator $T \in \mathcal{L}(\mathcal{U};\mathcal{H})$ is called **Hilbert-Schmidt** if*

$$\sum_{n=1}^{\infty} \|Tu_n\|_{\mathcal{H}}^2 < \infty,$$

where $(u_n)_{n\in\mathbb{N}}$ is an orthonormal basis of $\mathcal{U}$.

The space of all Hilbert-Schmidt operators is denoted by $\mathcal{L}_{(HS)}(\mathcal{U};\mathcal{H})$ and we set $\mathcal{L}_{(HS)}(\mathcal{U}) = \mathcal{L}_{(HS)}(\mathcal{U};\mathcal{U})$. Let $(u_n)_{n\in\mathbb{N}}$ be an orthonormal basis of $\mathcal{U}$. The space $\mathcal{L}_{(HS)}(\mathcal{U};\mathcal{H})$ equipped with the inner product

$$\langle S,T\rangle_{\mathcal{L}_{(HS)}(\mathcal{U};\mathcal{H})} = \sum_{n=1}^{\infty}\langle Su_n, Tu_n\rangle_{\mathcal{H}}$$

for every $S,T \in \mathcal{L}_{(HS)}(\mathcal{U};\mathcal{H})$ becomes a separable Hilbert space. If $(h_n)_{n\in\mathbb{N}}$ is an orthonormal basis of $\mathcal{H}$, then $(h_n \otimes u_m)_{n,m\in\mathbb{N}}$ defined by $h_n \otimes u_m = h_n\langle u_m, \cdot\rangle_{\mathcal{U}}$ is a complete orthonormal basis of $\mathcal{L}_{(HS)}(\mathcal{U};\mathcal{H})$.

Proposition C.7 (Proposition A.3, [71])**.** *The norm corresponding to the inner product on $\mathcal{L}_{(HS)}(\mathcal{U};\mathcal{H})$ is independent on the choice of the orthonormal basis of $\mathcal{U}$. Moreover, we have $T \in \mathcal{L}_{(HS)}(\mathcal{U};\mathcal{H})$ if and only if $T^* \in \mathcal{L}_{(HS)}(\mathcal{H};\mathcal{U})$. In this case, it holds that*

$$\|T\|_{\mathcal{L}_{(HS)}(\mathcal{U};\mathcal{H})} = \|T^*\|_{\mathcal{L}_{(HS)}(\mathcal{H};\mathcal{U})}.$$

Proposition C.8 (Proposition C.4, [23])**.** *Let $\mathcal{V}$ be another separable Hilbert space. If $S \in \mathcal{L}_{(HS)}(\mathcal{U};\mathcal{H})$ and $T \in \mathcal{L}_{(HS)}(\mathcal{H};\mathcal{V})$, then $TS \in \mathcal{L}_1(\mathcal{U};\mathcal{V})$ and*

$$\|TS\|_{\mathcal{L}_1(\mathcal{U};\mathcal{V})} \leq \|S\|_{\mathcal{L}_{(HS)}(\mathcal{U};\mathcal{H})}\|T\|_{\mathcal{L}_{(HS)}(\mathcal{H};\mathcal{V})}.$$

Proposition C.9 (Proposition 2.3.4, [73])**.** *If $T \in \mathcal{L}(\mathcal{U})$ is self-adjoint and nonnegative, then there exists a unique self-adjoint nonnegative operator $T^{1/2} \in \mathcal{L}(\mathcal{U})$ such that $T^{1/2}T^{1/2} = T$. If additionally $Tr(T) < \infty$, then $T^{1/2} \in \mathcal{L}_{(HS)}(\mathcal{U})$ with $\|T^{1/2}\|_{L_{(HS)}(\mathcal{U})} = Tr(T)$ and $ST^{1/2} \in \mathcal{L}_{(HS)}(\mathcal{U};\mathcal{H})$ for every $S \in \mathcal{L}(\mathcal{U};\mathcal{H})$.*

Remark C.10. *Let $T \in \mathcal{L}(\mathcal{U})$ be a self-adjoint nonnegative operator such that $Tr(T) < \infty$. Due to Proposition C.5, there exist an orthonormal basis $(u_n)_{n\in\mathbb{N}}$ of $\mathcal{U}$ and a sequence $(\lambda_n)_{n\in\mathbb{N}}$ of nonnegative real numbers such that $Tu_n = \lambda_n u_n$ for each $n \in \mathbb{N}$. Using Proposition C.9, there exists a unique self-adjoint nonnegative operator $T^{1/2} \in \mathcal{L}(\mathcal{U})$ such that $T^{1/2}T^{1/2} = T$. We obtain $T^{1/2}u_n = \sqrt{\lambda_n}u_n$ for each $n \in \mathbb{N}$. Hence, we can conclude that the subspace $T^{1/2}(\mathcal{U})$ of $\mathcal{U}$ equipped with the inner product*

$$\langle x,y\rangle_{T^{1/2}(\mathcal{U})} = \sum_{n=1}^{\infty}\frac{1}{\lambda_n}\langle x,u_n\rangle_{\mathcal{U}}\langle y,u_n\rangle_{\mathcal{U}}$$

for every $x,y \in T^{1/2}(\mathcal{U})$ becomes a Hilbert space and $(\sqrt{\lambda_n}u_n)_{n\in\mathbb{N}}$ is an orthonormal basis of $T^{1/2}(\mathcal{U})$. By definition, we obtain that the inner product on $\mathcal{L}_{(HS)}(T^{1/2}(\mathcal{U});\mathcal{H})$ is given by

$$\langle R,S\rangle_{\mathcal{L}_{(HS)}(T^{1/2}(\mathcal{U});\mathcal{H})} = \sum_{n=1}^{\infty}\lambda_n\langle Ru_n, Su_n\rangle_{\mathcal{H}} = \langle RT^{1/2}, ST^{1/2}\rangle_{\mathcal{L}_{(HS)}(\mathcal{U};\mathcal{H})}$$

for every $R,S \in \mathcal{L}_{(HS)}(T^{1/2}(\mathcal{U});\mathcal{H})$. Moreover, we have $\mathcal{L}(\mathcal{U};\mathcal{H}) \subset \mathcal{L}_{(HS)}(T^{1/2}(\mathcal{U});\mathcal{H})$ as a consequence of Proposition C.9.

D. Optimization in Infinite Dimension

In this section, we consider convex as well as nonconvex optimization problems of functionals defined on Banach spaces. We introduce the concepts of Gâteaux and Fréchet derivatives. Moreover, we state results on the existence of unique extrema, which represents a solution of an optimization problem. Finally we state necessary and sufficient optimality conditions such a extrema has to satisfy. For more details, we refer to [11, 51, 57, 93]. Throughout this section, let X, Y and Z be Banach spaces.

Differential Calculus in Banach Spaces

We start with a formal the definition.

Definition D.1. *Let $f\colon M \subset X \to Y$ be an operator with $M \neq \emptyset$ open.*

*(i) We call f **Gâteaux differentiable** at $x \in M$ if the limit*

$$d^G f(x)[h] = \lim_{t \to 0} \frac{f(x+th) - f(x)}{t}$$

*exists for all $h \in X$ and $d^G f(x) \in \mathcal{L}(X;Y)$. We then call $d^G f(x)[h]$ the **Gâteaux derivative** of f at $x \in M$ in direction $h \in X$.*

*(ii) We call f **Fréchet differentiable** at $x \in M$ if the limit*

$$\lim_{\|h\|_X \to 0} \frac{\|f(x+h) - f(x) - d^F f(x)[h]\|_Y}{\|h\|_X} = 0$$

*exists and $d^F f(x) \in \mathcal{L}(X;Y)$. We then call $d^F f(x)[h]$ the **Fréchet derivative** of f at $x \in M$ in direction $h \in X$.*

*(iii) We call f **Gâteaux/Fréchet differentiable** on M if f is Gâteaux/Fréchet differentiable at every $x \in M$.*

Remark D.2. *(i) Let $f\colon M \subset X \to Y$ be an operator with $M \neq \emptyset$ open. The Gâteaux derivative of order $n \geq 2$ denoted by $d^G(f(x))^n$ with $x \in M$ is defined as the Gâteaux derivative of $d^G(f(x))^{n-1}$, whenever it exists. Similarly, we define the Fréchet derivative of order $n \geq 2$.*
(ii) Let $M_X \subset X$, $M_Y \subset Y$ be nonempty and open and let $f\colon M_X \times M_Y \to Z$ be an operator. For fixed $y \in M_Y$, the partial Gâteaux derivatives of f at $x \in M_X$ in direction $h \in X$ is defined by

$$d_x^G f(x,y)[h] = \lim_{t \to 0} \frac{f(x+th, y) - f(x,y)}{t},$$

whenever the limit exists. The partial Gâteaux derivatives of f at $y \in M_Y$ is defined analogously. Similarly, we define the partial Fréchet derivatives.

In contrast to the Gâteaux derivative, there exists a chain rule for the Fréchet derivative. This is the main difference of both type of derivatives.

Proposition D.3 (Theorem 4.1.1,[57])**.** *Let $g\colon M_X \subset X \to Y$ and $f\colon M_Y \subset Y \to Z$ be Fréchet differentiable on the open sets M_X and M_Y, respectively. Then the Fréchet derivative of the composition $f \circ g$ at $x \in M_X$ in direction $h \in X$ is given by*

$$d^F f \circ g(x)[h] = d^F f(g(x)) \left[d^F g(x)[h] \right].$$

Obviously, every Fréchet differentiable operator is Gâteaux differentiable. The converse is in general not true. However, we can state conditions such that the Gâteaux derivative and the Fréchet derivative coincides. Therefor, we need the following preliminary result known as the mean value theorem.

Proposition D.4 (Theorem 4.1.2 (b),[57])**.** *Let $M \subset X$ open and let $f\colon M \subset X \to Y$ be Gâteaux differentiable on the interval*

$$[x, x+h] = \{x + th \colon t \in [0,1]\} \subset M.$$

If $z \mapsto d^G f(z)$ is continuous from $[x, x+h]$ into $\mathcal{L}(X;Y)$, then

$$\|f(x+h) - f(x)\|_Y \leq \sup_{t\in[0,1]} \|d^G f(x+th)\|_{\mathcal{L}(X;Y)} \|h\|_X$$

and for every $T \in \mathcal{L}(X;Y)$

$$\|f(x+h) - f(x) - Th\|_Y \leq \sup_{t\in[0,1]} \|d^G f(x+th) - T\|_{\mathcal{L}(X;Y)} \|h\|_X.$$

Note that the previous proposition holds especially for $T = d^G f(x)$. Thus, we get immediately the following result.

Corollary D.5 (Corollary 4.1.1,[57])**.** *Let $M \subset X$ open and let $f\colon M \subset X \to Y$ be continuous and Gâteaux differentiable on M. If $x \mapsto d^G f(x)$ is continuous from M into $\mathcal{L}(X;Y)$, then f is Fréchet differentiable on M and for every $x \in M$*

$$d^F f(x) = d^G f(x).$$

Remark D.6. *Here, we show some classical examples of Fréchet differentiable operators:*

(i) Let $g\colon X \to Y$ be defined by

$$g(x) = Tx + y,$$

where $T \in \mathcal{L}(X;Y)$ and $y \in Y$ is fixed. Then we get for every $x, h \in X$

$$g(x+h) - g(x) = Th.$$

Thus, we can conclude that for every $x \in X$

$$\lim_{\|h\|_X \to 0} \frac{\|g(x+h) - g(x) - Th\|_Y}{\|h\|_X} = 0.$$

Therefore, the operator g is Fréchet differentiable on X and the Fréchet derivative of g at $x \in X$ in direction $h \in X$ is given by

$$d^F g(x)[h] = Th.$$

(ii) Let Y be a Hilbert space and let $f\colon Y \to \mathbb{R}$ be given by

$$f(y) = \|y\|_Y^2.$$

We obtain for every $x, h \in X$

$$f(y+h) - f(y) = \langle y+h, y+h\rangle_Y - \langle y, y\rangle_Y = 2\langle y, h\rangle_Y + \|h\|_Y^2.$$

Hence, we can infer that for every $y \in Y$

$$\lim_{\|h\|_Y \to 0} \frac{|f(y+h) - f(y) - 2\langle y, h\rangle_Y|}{\|h\|_Y} = 0.$$

Therefore, the functional f is Fréchet differentiable on Y and the Fréchet derivative of f at $y \in Y$ in direction $h \in Y$ is given by

$$d^F f(y)[h] = 2\langle y, h\rangle_Y.$$

(iii) Let Y be a Hilbert space. Moreover, let $g\colon X \to Y$ and $f\colon Y \to \mathbb{R}$ be as in (i) and (ii), respectively. Thus, the composition $f \circ g\colon X \to \mathbb{R}$ is given by

$$f \circ g(x) = \|Tx + y\|_Y^2.$$

Using Proposition D.3, the functional $f \circ g$ is Fréchet differentiable on X and the Fréchet derivative of $f \circ g$ at $x \in X$ in direction $h \in X$ is given by

$$d^F f \circ g(x)[h] = d^F f(g(x)) \left[d^F g(x)[h]\right] = 2\langle g(x), Th\rangle_Y.$$

Convex Optimization Problems

Let $f\colon M \subset X \to \mathbb{R}$ be a functional with $M \neq \emptyset$. We consider the following optimization problem:

$$f(\overline{x}) = \inf_{x \in M} f(x), \tag{D.1}$$

where $\overline{x} \in M$ is called the minimum of f. First, we state conditions on the optimization problem ensuring the existence and uniqueness of such a minimum $\overline{x} \in M$. Therefor, we introduce the concept of lower semi-continuous functionals.

Definition D.7. *Let $f\colon M \subset X \to \mathbb{R}$ be a functional. Then:*

(i) f is called ***lower semi-continuous*** *at $x \in M$ if*

$$f(x) \leq \liminf_{y \to x} f(y).$$

(ii) f is called ***sequentially lower semi-continuous*** *at $x \in M$ if for every sequence $(x_n)_{n\in\mathbb{N}} \subset M$ such that $\lim_{n\to\infty} x_n = x$, we have*

$$f(x) \leq \liminf_{n \to \infty} f(x_n);$$

(iii) f is called ***weak sequentially lower semi-continuous*** *at $x \in M$ if for every sequence $(x_n)_{n\in\mathbb{N}} \subset M$ such that $\lim_{n\to\infty}\langle x', x_n\rangle = \langle x', x\rangle$ for every element $x' \in X'$, we have*

$$f(x) \leq \liminf_{n \to \infty} f(x_n).$$

Remark D.8. *Note that every continuous functional $f\colon M \subset X \to \mathbb{R}$ is lower semi-continuous. Moreover, we can conclude that f is lower semi-continuous if and only if f is sequentially lower semi-continuous.*

Under additional assumptions, we get the following equivalence.

Lemma D.9 (Proposition 38.7 (2), [93])**.** *Let $f\colon M \subset X \to \mathbb{R}$ be convex with M closed and convex. Then f is (sequentially) lower semi-continuous if and only if f is weak sequentially lower semi-continuous.*

We have the following existence result on a solution to problem (D.1).

Proposition D.10 (Theorem 7.3.5, [57])**.** *Let X be a reflexive Banach space and suppose that the functional $f\colon M \subset X \to \mathbb{R}$ is weak sequentially lower semi-continuous with M bounded, closed and convex. Then there exists $\overline{x} \in M$ such that*

$$f(\overline{x}) = \inf_{x \in M} f(x).$$

If the functional is defined on an unbounded set, then we can overcome this problems as follows.

Definition D.11. *A functional $f\colon M \subset X \to \mathbb{R}$ is* ***coercive*** *over M if*

$$\lim_{\|x\|_X \to \infty} |f(x)| = \infty,$$

where $x \in M$.

Proposition D.12 (Theorem 7.3.7, [57])**.** *Let X be a reflexive Banach space and suppose that the functional $f\colon M \subset X \to \mathbb{R}$ is coercive and weak sequentially lower semi-continuous with M closed and convex. Then there exists $\overline{x} \in M$ such that*

$$f(\overline{x}) = \inf_{x \in M} f(x).$$

The uniqueness of a solution to problem (D.1) can be achieved under additional requirements.

Corollary D.13. *Let X be a reflexive Banach space and suppose that the functional $f\colon M \subset X \to \mathbb{R}$ is coercive, strictly convex and weak sequentially lower semi-continuous with M closed and convex. Then there exists a unique element $\overline{x} \in M$ such that*

$$f(\overline{x}) = \inf_{x \in M} f(x).$$

Proof. Using Proposition D.12, we get immediately the existence of a minimum. To prove uniqueness, we assume that $\overline{x}_1, \overline{x}_2 \in M$ with $\overline{x}_1 \neq \overline{x}_2$ satisfy $f(\overline{x}_1) = f(\overline{x}_2) = \inf_{x \in M} f(x)$. For all $t \in (0,1)$, we have $t\overline{x}_1 + (1-t)\overline{x}_2 \in M$ due to the fact that M is convex. Since f is strictly convex, we get for all $t \in (0,1)$

$$f(\overline{x}_2) = tf(\overline{x}_1) + (1-t)f(\overline{x}_2) > f(t\overline{x}_1 + (1-t)\overline{x}_2),$$

which is a contradiction to the assumption. □

Note that the previous corollary remains still true if $M = X$. In this case, we get the following necessary and sufficient optimality conditions.

Proposition D.14 (Proposition 42.10, [93])**.** *If $f\colon X \to \mathbb{R}$ is convex and Gâteaux differentiable on X, then $\overline{x} \in X$ is a minimum of f if and only if*

$$d^G f(\overline{x})[h] = 0$$

for every $h \in X$.

Nonconvex Optimization Problems

Let $f\colon M \subset X \to \mathbb{R}$ be a functional with $M \neq \emptyset$. We study again the optimization problem

$$f(\overline{x}) = \inf_{x \in M} f(x),$$

where $\overline{x} \in M$ is the minimum of f. In Corollary D.13, note that the uniqueness of a minimum is mainly based on the assumption that the functional f is strictly convex. In contrast to this result, we consider here more general problems, where f is not necessarily convex. However, the existence and uniqueness of a minimum can still be obtained for a certain class of optimization problems.

First, we introduce uniformly convex Banach spaces. These spaces were first introduced in [19]. It is also shown that the function spaces L^p and the sequence spaces l^p for $p \in (1,\infty)$ are specific examples.

Definition D.15. *A Banach space X is called* ***uniformly convex*** *if for every $\varepsilon \in (0,2]$, there exists $\delta(\varepsilon) > 0$ such that for every $x, y \in X$*

$$\|x\|_X = \|y\|_X = 1, \|x-y\|_X \geq \varepsilon \quad \Rightarrow \quad \left\|\frac{x+y}{2}\right\|_X \leq 1 - \delta(\varepsilon).$$

This definition has a simple geometric interpretation. It states that the mid-point of two elements of the unit sphere cannot approach the surface of the sphere unless the distance of these elements goes to zero.

Remark D.16. *We have the following basic results:*

(i) The Milman–Pettis theorem states that every uniformly convex Banach space is reflexive, see [72]. The converse is in general not true, see [24].

(ii) *Let X be a Hilbert space. If $x, y \in X$ satisfy $\|x\|_X = \|y\|_X = 1$ and $\|x - y\| \geq \varepsilon$ with $\varepsilon \in (0, 2]$, then by the parallelogram law, we have*

$$\begin{aligned}\|x+y\|_X^2 &= \|x+y\|_X^2 + \|x-y\|_X^2 - \|x-y\|_X^2 \\ &\leq 2\|x\|_X + 2\|y\|_X^2 - \|x-y\|_X^2 \\ &\leq 4 - \varepsilon^2.\end{aligned}$$

We set $\delta(\varepsilon) = 1 - \frac{1}{2}\sqrt{4 - \varepsilon^2}$. Then we get $\left\|\frac{x+y}{2}\right\|_X \leq 1 - \delta(\varepsilon)$. Therefore, every Hilbert space is uniformly convex.

We have the following existence and uniqueness result of a minimum to the optimization problem introduced above, where we assume that the functional $f\colon M \subset X \to \mathbb{R}$ is given by

$$f(x) = g(x) + \|x\|_X^p \tag{D.2}$$

with $g\colon M \subset X \to \mathbb{R}$ and $p \geq 1$.

Proposition D.17 (Partie (A), Théorème 4.2,[11])**.** *Let X be an uniformly convex Banach space and let $M \subset X$ be bounded and closed. Moreover, let $g\colon M \subset X \to \mathbb{R}$ be a lower semi-continuous functional, which is bounded from below. Then there exists a dense subset $M_0 \subset M$ such that for every $y \in M_0$ and all $p \geq 1$ we get the existence of an element $x(y) \in M$ satisfying*

$$g(x(y)) + \|x(y) - y\|_X^p = \inf_{x \in M} \left(g(x) + \|x - y\|_X^p\right).$$

If $p > 1$, then $x(y)$ is unique. Furthermore, the mapping $y \mapsto x(y)$ is continuous on M_0.

The previous proposition has a simple consequence.

Corollary D.18. *Let X be a uniformly convex Banach space and let $M \subset X$ be bounded and closed such that $0 \in M$. Moreover, let $g\colon M \subset X \to \mathbb{R}$ be a continuous functional, which is bounded from below. Then for all $p \geq 1$, there exists $\overline{x} \in M$ such that*

$$g(\overline{x}) + \|\overline{x}\|_X^p = \inf_{x \in M} \left(g(x) + \|x\|_X^p\right).$$

If $p > 1$, then $\overline{x}$ is unique.

Proof. Using Proposition D.17, there exists a dense subset $M_0 \subset M$ such that for every $y \in M_0$ and all $p \geq 1$ we get the existence of an element $x(y) \in M$ satisfying

$$g(x(y)) + \|x(y) - y\|_X^p = \inf_{x \in M} \left(g(x) + \|x - y\|_X^p\right).$$

Furthermore, the mapping $y \mapsto x(y)$ is continuous. Since M_0 is a dense subset of M and $0 \in M$, there exists a sequence $(y_n)_{n \in \mathbb{N}} \subset M_0$ such that $\lim_{n\to\infty} \|y_n\|_X = 0$. Let $\overline{x} \in M$ be given by

$$\overline{x} = \lim_{n\to\infty} x(y_n). \tag{D.3}$$

Due to the continuity properties, we get

$$\begin{aligned}g(\overline{x}) + \|\overline{x}\|_X^p &= \lim_{n\to\infty} \left(g(x(y_n)) + \|x(y_n) - y_n\|_X^p\right) \\ &= \lim_{n\to\infty} \inf_{x \in M} \left(g(x) + \|x - y_n\|_X^p\right) \\ &= \inf_{x \in M} \left(g(x) + \|x\|_X^p\right).\end{aligned}$$

The uniqueness of $\overline{x}$ is a consequence of Proposition D.17 and the uniqueness of the limit in (D.3). □

Next, we state a necessary optimality condition.

Proposition D.19 (Theorem 1.46, [51])**.** *Let $M \neq \emptyset$ be a convex subset of X and let $f\colon M \subset X \to \mathbb{R}$ be well defined on an open neighborhood of M. Then $\overline{x} \in M$ is a minimum of f if*

$$d^G f(\overline{x})[x - \overline{x}] \geq 0 \tag{D.4}$$

for every $x \in M$, whenever f is Gâteaux differentiable at $\overline{x} \in M$.

Remark D.20. *In the previous proposition, we can consider $d^G f(\overline{x})$ as an element of the dual space X'. In the sense of dual pairing, inequality (D.4) can be rewritten equivalently by*

$$\langle d^G f(\overline{x}), x - \overline{x} \rangle \geq 0$$

for every $x \in M$. Thus, inequality (D.4) is often called variational inequality.

If X is a Hilbert space, then the variational inequality (D.4) can often be solved using a projection operator. Therefor, we need the following result.

Proposition D.21 (Lemma 1.10 (b), [51])**.** *Let M be a closed and convex subset of the Hilbert space X and let $P\colon X \to M$ be the projection on M, i.e.*

$$\|P(x) - x\|_X = \min_{y \in M} \|y - x\|_X$$

for every $x \in X$. Then $z = P(x)$ for $x \in X$ if and only if for every $y \in M$

$$\langle x - z, y - z \rangle_X \leq 0.$$

Finally, we state a sufficient optimality condition.

Proposition D.22 (Theorem 4.23, [86])**.** *Let M be a convex subset of the Banach space X. Moreover, let the functional $f\colon X \to \mathbb{R}$ be twice Fréchet differentiable on an open subset containing $\overline{x} \in M$ such that the mapping $x \mapsto d^F(f(x))^2$ is continuous at $\overline{x}$. If $\overline{x}$ satisfies*

$$d^F f(\overline{x})[x - \overline{x}] \geq 0$$

for every $x \in M$ and there exists a constant $\delta > 0$ such that

$$d^F(f(\overline{x}))^2[h, h] \geq \delta \|h\|_X^2$$

for every $h \in X$, then there exist constants $\varepsilon, \sigma > 0$ such that

$$f(x) \geq f(\overline{x}) + \sigma \|x - \overline{x}\|_X^2$$

for every $x \in M$ with $\|x - \overline{x}\|_X \leq \varepsilon$.

Remark D.23. *From the previous proposition, we only obtain that the minimum $\overline{x} \in M$ is local. If we require additionally that $f\colon X \to \mathbb{R}$ is given by equation (D.2) with $p > 1$ such that the assumptions of Corollary D.18 hold, then we can conclude that the minimum is also global.*

References

[1] A. Al-Hussein. *Strong, mild and weak solutions of backward stochastic evolution equations*. Random Oper. Stoch. Equ., 13(2):129–138, 2005.

[2] D. Applebaum. *Lévy Processes and Stochastic Calculus.* Cambridge University Press, 2009.

[3] D. Applebaum and J.-L. Wu. *Stochastic partial differential equations driven by Lévy space–time white noise*. Random Oper. Stoch. Equ., 8(3):245–259, 2000.

[4] O. Axelsson, S. Farouq, and M. Neytcheva. *A preconditioner for optimal control problems, constrained by Stokes equation with a time–harmonic control*. J. Comput. Appl. Math., 310:5–18, 2017.

[5] P. Benner and C. Trautwein. *A linear quadratic control problem for the stochastic heat equation driven by Q–Wiener processes*. J. Math. Anal. Appl., 457(1):776–802, 2018.

[6] P. Benner and C. Trautwein. *A Stochastic Maximum Principle for Control Problems Constrained by the Stochastic Navier–Stokes Equations*. arXiv:1810.12119, 2018.

[7] P. Benner and C. Trautwein. *Optimal Control Problems Constrained by the Stochastic Navier–Stokes Equations with Multiplicative Lévy Noise*. accepted for publication in Math. Nachr., 2018.

[8] P. Benner and C. Trautwein. *Optimal Distributed and Tangential Boundary Control for the Unsteady Stochastic Stokes Equations*. arXiv:1809.00911, 2018.

[9] A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter. *Representation and Control of Infinite Dimensional Systems.* Birkhäuser, 2007.

[10] A. Bensoussan and J. Frehse. *Local Solutions for Stochastic Navier Stokes Equations*. M2AN Math. Model. Numer. Anal., 34(2):241–273, 2000.

[11] M. F. Bidaut. *Théorèmes d'existence et d'existence "en général" d'un contrôle optimal pour des systèmes régis par des équations aux dérivées partielles non linéaires.* PhD thesis, Université Paris VI, 1973.

[12] P. Billingsley. *Probability and Measure.* John Wiley & Sons Inc., New York, 1995.

[13] V. Bogachev. *Measure Theory.* Springer, Berlin, 2007.

[14] H. Breckner. *Approximation and optimal control of the stochastic Navier–Stokes equation.* PhD thesis, Martin Luther University Halle–Wittenberg, 1999.

[15] H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations.* Springer, New York, 2011.

[16] Z. Brzeżniak, B. Maslowski, and J. Seidler. *Stochastic nonlinear beam equations*. Probab. Theory Relat. Fields, 132:119–149, 2005.

[17] P. Chen, A. Quarteroni, and G. Rozza. *Multilevel and weighted reduced basis method for stochastic optimal control problems constrained by Stokes equations*. Numer. Math., 133:67–102, 2016.

[18] A. Chojnowska-Michalik. *On processes of Ornstein–Uhlenbeck type in Hilbert space*. Stochastics, 21:251–286, 1987.

[19] J. A. Clarkson. *Uniformly convex spaces*. Trans. Amer. Math. Soc., 40(3):396–414, 1936.

[20] R. Cont and P. Tankov. *Financial Modelling with Jump Processes.* Chapman and Hall/CRC, 2004.

[21] R. F. Curtain and A. J. Pritchard. *Infinite Dimensional Linear Systems Theory.* Lecture Notes in Control and Information Sciences, vol. 8, Springer, Berlin, 1978.

[22] G. Da Prato and A. Debussche. *Dynamic programming for the stochastic Navier–Stokes equations*. M2AN Math. Model. Numer. Anal., 34(2):459–475, 2000.

[23] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions.* Cambridge University Press, 2014.

[24] M. M. Day. *Reflexive Banach spaces not isomorphic to uniformly convex spaces*. Bull. Amer. Math. Soc., 47:313–317, 1941.

[25] L. Debbi. *Well–Posedness of the Multidimensional Fractional Stochastic Navier–Stokes Equations on the Torus and on Bounded Domains*. J. Math. Fluid Mech., 18:25–69, 2016.

[26] A. Debussche, M. Fuhrman, and G. Tessitore. *Optimal control of a stochastic heat equation with boundary–noise and boundary–control*. ESAIM Control Optim. Calc. Var., 13:178–205, 2007.

[27] Z. Denkowski, S. Migorski, and N. S. Papageorgiou. *An Introduction to Nonlinear Analysis: Theory.* Kluwer Academic/Pienum, New York, 2003.

[28] J. Duan and W. Wang. *Effective Dynamics of Stochastic Partial Differential Equations.* Elsevier, 2014.

[29] T. E. Duncan, B. Maslowski, and B. Pasik-Duncan. *Stochastic equations in Hilbert space with a multiplicative fractional Gaussian noise*. Stochastic Process. Appl., 115(8):1357–1383, 2005.

[30] Y. V. Egorov and M. A. Shubin. *Foundations of the Classical Theory of Partial Differential Equations.* Springer, Berlin, 1998.

[31] K.-J. Engel and R. Nagel. *A Short Course on Operator Semigroups.* Springer, New York, 2006.

[32] G. Fabbri and B. Goldys. *An LQ problem for the heat equation on the halfline with Dirichlet boundary control and noise*. SIAM J. Control Optim., 48:1473–1488, 2009.

[33] G. Fabbri, F. Gozzi, and A. Święch. *Stochastic Optimal Control in Infinite Dimension: Dynamic Programming and HJB Equations.* Springer, Berlin, 2017.

[34] B. P. W. Fernando, B. Rüdiger, and S. S. Sritharan. *Mild solutions of stochastic Navier–Stokes equation with jump noise in* $\mathbb{L}^p$*–spaces*. Math. Nachr., 288:1615–1621, 2015.

[35] W. H. Fleming and R. W. Rishel. *Deterministic and Stochastic Optimal Control.* Springer, New York, 1975.

[36] M. Fuhrman and C. Orrieri. *Stochastic maximum principle for optimal control of a class of nonlinear SPDEs with dissipative drift*. SIAM J. Control Optim., 54(1):341–371, 2016.

[37] H. Fujita and H. Morimoto. *On fractional powers of the Stokes operator*. Proc. Japan Acad., 46:1141–1143, 1970.

[38] D. Fujiwara. *Concrete characterization of the domains of fractional powers of some elliptic differential operators of the second order*. Proc. Japan Acad., 43:82–86, 1967.

[39] D. Fujiwara and H. Morimoto. *An L_r–theorem of the Helmholtz decomposition of vector fields*. J. Fac. Sci. Univ. Tokyo Sec. 1 A, 24(3):685–700, 1977.

[40] G. P. Galdi. *An Introduction to the Mathematical Theory of the Navier–Stokes Equations: Steady–State Problems.* Springer, New York, 2011.

[41] G. P. Galdi, C. G. Simader, and H. Sohr. *A class of solutions to the stationary Stokes and Navier–Stokes equations with boundary data in $W^{-1/q,q}$*. Math. Ann., 331:41–74, 2005.

[42] L. Gawarecki and V. Mandrekar. *Stochastic Differential Equations in Infinite Dimensions: with Applications to Stochastic Partial Differential Equations.* Springer, Berlin, 2011.

[43] Y. Giga. *Analyticity of the Semigroup Generated by the Stokes Operator in L_r Spaces*. Math. Z., 178:297–329, 1981.

[44] Y. Giga and T. Miyakawa. *Solutions in L_r of the Navier–Stokes initial value problem*. Arch. Ration. Mech. Anal., 89(3):267–281, 1985.

[45] T. E. Govindan. *Yosida Approximations of Stochastic Differential Equations in Infinite Dimensions and Applications.* Springer, New York, 2016.

[46] F. Gozzi, S. S. Sritharan, and A. Święch. *Bellman Equations Associated to the Optimal Feedback Control of Stochastic Navier–Stokes Equations*. Comm. Pure Appl. Math., 58:671–700, 2005.

[47] G. Guatteri and F. Masiero. *On the existence of optimal controls for SPDEs with boundary–noise and boundary–control*. SIAM J. Control Optim., 51(3):1909–1939, 2013.

[48] D. D. Haroske and H. Triebel. *Distributions, Sobolev Spaces, Elliptic Equations.* European Math. Soc., Zürich, 2008.

[49] E. Hausenblas and J. Seidler. *Stochastic Convolutions Driven by Martingales: Maximal Inequalities and Exponential Integrability*. Stoch. Anal. Appl., 26:98–119, 2008.

[50] M. Hintermüller and M. Hinze. *A SQP–Semismooth Newton–type Algorithm applied to Control of the instationary Navier–Stokes System Subject to Control Constraints*. SIAM J. Optim., 16(4):1177–1200, 2006.

[51] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints.* Springer, New York, 2009.

[52] Y. Hu and S. Peng. *Adapted solution of a backward semilinear stochastic evolution equation*. Stoch. Anal. Appl., 9(4):445–459, 1991.

[53] A. Ichikawa. *Stability of Semilinear Stochastic Evolution Equations*. J. Math. Anal. Appl., 90:12–44, 1982.

[54] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, New York, 1998.

[55] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer, Berlin, 1992.

[56] W. Krendl, V. Simoncini, and W. Zulehner. *Efficient Preconditioning for an Optimal Control Problem with the Time–Periodic Stokes Equations*, volume 103 of *A. Abdulle, S. Deparis, D. Kressner, F. Nobile, M. Picasso (eds) Numerical Mathematics and Advanced Applications - ENUMATH 2013. Lecture Notes in Computational Science and Engineering, vol. 103*, pages 479–487. Springer, Cham, 2015.

[57] A. J. Kurdila and M. Zabarankin. *Convex Functional Analysis.* Birkhäuser, 2005.

[58] S. Lang. *Real and Functional Analysis.* Springer, New York, 1993.

[59] H. Liang, L. Hou, and J. Ming. *The velocity tracking problem for Wick–stochastic Navier–Stokes flows using Wiener chaos expansion*. J. Comput. Appl. Math., 307:25–36, 2016.

[60] J. L. Lions and E. Magenes. *Non–Homogeneous Boundary Value Problems and Applications*, volume 1. Springer, Berlin, 1972.

[61] F. Masiero. *A Stochastic Optimal Control Problem for the Heat Equation on the Halfline with Dirichlet Boundary–Noise and Boundary–Control*. Appl. Math. Optim., 62:253–294, 2010.

[62] B. Maslowski and D. Nualart. *Evolution equations driven by a fractional Brownian motion*. J. Funct. Anal., 202:277–305, 2003.

[63] M. Métivier. *Semimartingales: A Course on Stochastic Processes.* De Gruyter, Berlin, New York, 1982.

[64] Y. Mishura. *Stochastic Calculus for Fractional Brownian Motion and Related Processes.* Springer, Berlin, 2008.

[65] D. Nualart and W. Schoutens. *Chaotic and predictable representations for Lévy processes*. Stochastic Process. Appl., 90:109–122, 2000.

[66] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications.* Springer, Berlin, 2003.

[67] B. Øksendal. *Optimal Control of Stochastic Partial Differential Equations*. Stoch. Anal. Appl., 23(1):165–179, 2005.

[68] Y. R. Ou. *Design of Feedback Compensators for Viscous Flow.* In: Optimal Control of Viscous Flow, pages 151–180. SIAM, Philadelphia, 1998.

[69] E. Pardoux and A. Răşcanu. *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations.* Springer, Cham, 2014.

[70] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations.* Springer, New York, 1983.

[71] S. Peszat and J. Zabczyk. *Stochastic Partial Differential Equations with Lévy Noise: An Evolution Equation Approach.* Cambridge University Press, 2007.

[72] B. J. Pettis. *A proof that every uniformly convex space is reflexive*. Duke Math. J., 5(2):249–253, 1939.

[73] C. Prévôt and M. Röckner. *A Concise Course on Stochastic Partial Differential Equations.* Springer, Berlin, 2007.

[74] P. E. Protter. *Stochastic Integration and Differential Equations.* Springer, New York, 2004.

[75] J.-P. Raymond. *Feedback boundary stabilization of the two–dimensional Navier–Stokes equations*. SIAM J. Control Optim., 45(3):790–828, 2006.

[76] J.-P. Raymond. *Stokes and Navier–Stokes equations with nonhomogeneous boundary conditions*. Ann. Inst. H. Poincaré Anal. Non Linéaire, 24:921–951, 2007.

[77] J.-F. Renaud and B. Rémillard. *Malliavin calculus and Clark–Ocone formula for functionals of a square–integrable Lévy process*. Technical report, G–2009–67, 2009.

[78] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion.* Springer, Berlin, 1999.

[79] A. Rösch and B. Vexler. *Optimal control of the Stokes equations: A priori error analysis for finite element discretization with postprocessing*. SIAM J. Numer. Anal., 44(5):1903–1920, 2006.

[80] B. L. Rozovskii. *Stochastic Evolution Systems: Linear Theory and Applications to Non–linear Filtering.* Kluwer Academic, Dordrecht, 1990.

[81] W. Rudin. *Principles of Mathematical Analysis.* McGraw–Hill, New York, 1964.

[82] R. Situ. *Theory of Stochastic Differential Equations with Jumps and Applications: Mathematical and Analytical Techniques with Applications to Engineering.* Springer, New York, 2005.

[83] S. S. Sritharan. *An Introduction to Deterministic and Stochastic Control of Viscous Flow.* In: Optimal Control of Viscous Flow, pages 1–42. SIAM, Philadelphia, 1998.

[84] G. Tessitore. *Existence, uniqueness and space regularity of the adapted solutions of a backward SPDE*. Stochastic Anal. Appl., 14(4):461–486, 1996.

[85] H. Triebel. *Higher Analysis.* Barth, Leipzig, 1992.

[86] F. Tröltzsch. *Optimal Control of Partial Differential Equations: Theory, Methods and Applications.* American Mathematical Society, 2010.

[87] M. Ulbrich. *Constrained optimal control of Navier–Stokes flow by semismooth Newton methods*. Systems Control Lett., 48:297–311, 2003.

[88] W. von Wahl. *The Equations of Navier–Stokes and Abstract Parabolic Equations.* Vieweg + Teubner Verlag, 1985.

[89] I. Vrabie. *C_0–Semigroups and Applications.* Elsevier, Amsterdam, 2003.

[90] H. Ye, J. Gao, and Y. Ding. *A generalized Gronwall inequality and its application to a fractional differential equation*. J. Math. Anal. Appl., 328:1075–1081, 2007.

[91] J. Yong and X. Y. Zhou. *Stochastic Controls: Hamiltonian Systems and HJB Equations.* Springer, New York, 1999.

[92] H. Yu and B. Liu. *Properties of value function and existence of viscosity solution of HJB equation for stochastic boundary control problems*. J. Franklin Inst., 348:2108–2127, 2011.

[93] E. Zeidler. *Nonlinear Functional Analysis and its Applications III: Variational Methods and Optimization.* Springer, New York, 1985.

Ehrenerklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; verwendete fremde und eigene Quellen sind als solche kenntlich gemacht.

Ich habe insbesondere nicht wissentlich:

- Ergebnisse erfunden oder widersprüchliche Ergebnisse verschwiegen,
- statistische Verfahren absichtlich missbraucht, um Daten in ungerechtfertigter Weise zu interpretieren,
- fremde Ergebnisse oder Veröffentlichungen plagiiert oder verzerrt wiedergegeben.

Mir ist bekannt, dass Verstöße gegen das Urheberrecht Unterlassungs- und Schadenersatzansprüche des Urhebers sowie eine strafrechtliche Ahndung durch die Strafverfolgungsbehörden begründen kann.

Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form als Dissertation eingereicht und ist als Ganzes auch noch nicht veröffentlicht.

Magdeburg, 29.10.2018

Christoph Trautwein