Stochastic Evolution Equations with Lévy Noise and Applications to Delay Equations

Dissertation

zur Erlangung des

Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

Der

Naturwissenschaftlichen Fakultät II

der Martin-Luther-Universität

Halle-Wittenberg

vorgelegt von

Herrn Frank Wusterhausen

geboren am 12. April 1983 in Halle (Saale)

Gutachter:
Herr Prof. Dr. Dr. h. c. Wilfried Grecksch (Martin-Luther-Universität Halle-Wittenberg)
Herr Prof. Dr. Jan Prüß (Martin-Luther-Universität Halle-Wittenberg)
Herr Prof. Dr. Vo Anh (Queensland University of Technology, Brisbane, Australien)

Tag der Verteidigung: 06. November 2015

Diese Dissertation wurde teilweise von der Friedrich-Naumann-Stiftung für die Freiheit mit Mitteln des Bundesministeriums für Bildung und Forschung gefördert.

Acknowledgements. In the first place, I would like to express my gratitude to my supervisors, Prof. Dr. Dr. h. c. Wilfried Grecksch and Prof. Dr. Jan Prüß. They were excellent teachers to me, who allowed me to pursue my ideas for this thesis freely and were always open for discussing problems with me. I am also indebted to Prof. Dr. Dr. h. c. Wilfried Grecksch for the participation in national and international conferences and workshops. Furthermore I am grateful to my colleagues at the department, especially to the members of the working groups "Stochastic" and "Applied Analysis". Finally, I want to thank my parents for their constant support throughout my entire education and to my brother, whose help and advice proved useful many times.

Contents

Introduction 1

Chapter 1. Abstract Delay Equation 10
- 1.1. Deterministic linear case 10
- 1.2. Stochastic semilinear case with Lévy noise 24
- 1.3. Notation: $(S_D - ACP)$ and $(S - ACP)$ 35

Chapter 2. Approximation of the Mild Solution of $(S - ACP)$ 36
- 2.1. Yosida approximation, the generalized contraction case 36
- 2.2. Dissipativity of the operator $\mathscr{A}$ 42
- 2.3. Yosida approximation, the general case 48

Chapter 3. Transformation Formula 53
- 3.1. Transformation formula for Lévy processes 53
- 3.2. The space of transformation function 62
- 3.3. Transformation formula for mild solutions 64
- 3.4. The adjoint operator of $\mathscr{A}$ 78

Chapter 4. A Filtering Problem 83
- 4.1. A product formula 83
- 4.2. Filtering problem 86

Appendix A. Basics of Semigroup Theory 95

Appendix B. Stochastic Calculus 102
- B.1. Itô integral with respect to a Q-Wiener process 102
- B.2. Hilbert space-valued Lévy processes 108
- B.3. Stochastic integral with respect to a square integrable Lévy martingale 113
- B.4. Stochastic integral w.r.t. a compensated Poisson random measure 117
- B.5. Stochastic abstract Cauchy problem with Lévy noise 120

List of Figures 122

Bibliography 123

Index 127

Introduction

When modeling real world phenomena mathematically one usually starts by building a deterministic model ignoring all random occurrences. However, in many situations a deterministic view of the problem in question is not satisfactory. Just think of the financial sector, where from simple observations one finds that stock prices, interest rates, and other financial products have a significant random component. But the same remains true for technical and natural phenomena. An electronic signal is corrupted by noise, the water level of a river and the growth of a population change over time with some random effects. Even the expansion of heat does not follow the deterministic heat equation. Therefore, the need to add a stochastic component in order to make the models more realistic and applicable seems promising.

If a system is modeled by differential equations, the extension to include random effects was done very successfully by Itô [**Itô44, Itô51**], Stratonovič [**Str64**] and Shorohod [**Sko75**], who each developed a stochastic integration calculus based on the Wiener process. Those works build the foundations of the research area of stochastic differential equations. For the theory of stochastic differential equations with Wiener noise we refer to the monographs of Øksendal [**Øks03**] and Gihman & Skorohod [**GWMS14**] and references therein.

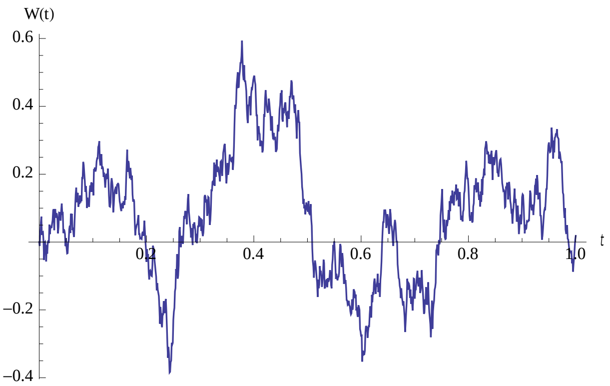

FIGURE 1. A random path of a Wiener process.

However, since the basic process to all these calculi is the Wiener process which is a pathwise continuous Markov process, it becomes inadequate as soon as chronological dependencies or jumps appear. In order to capture long-range or short-range dependence the concept of fractional Brownian motion was introduced by Mandelbrot & van Ness [**MVN68**]. Following this fundamental work a rich integration theory for fractional Brownian motion was developed by many authors over the years. A detailed introduction into the theory of stochastic calculus for fractional Brownian motion can be found in the monograph of Mishura [**Mis08**].

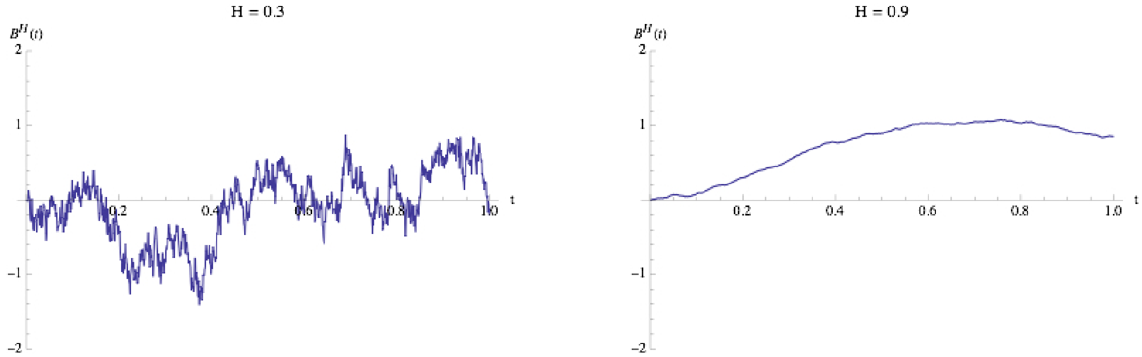

FIGURE 2. Sample paths of fractional Brownian motions with short-range dependence (left) and long-range dependence (right).

Another way to extend the original stochastic calculus by Itô is to allow the basic stochastic process to jump. This can be achieved, for example, by requesting the stochastic process to be only stochastically continuous. Then, a natural class of processes which fulfills this assumption and additionally has some nice properties are Lévy processes. The Wiener process is one example of a Lévy process. But also pure jump processes like the Poisson process or the compound Poisson process are Lévy processes. Compared to the continuous noise of the Wiener process or the fractional Brownian motion, jump noise has some fundamentally different properties. That is why, in some situations one has to separate continuous noise from jump noise in order to be able to deal with it mathematically. We will see this for example in Chapter 3 of this thesis. The extension of stochastic calculus to Lévy noise, to be more precise to square integrable martingales, is due to Kunita & Watanabe [**KW67**].

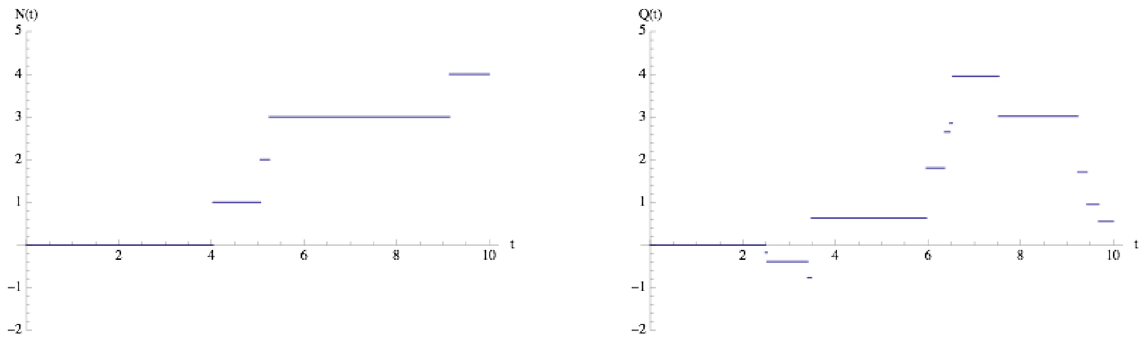

FIGURE 3. Sample paths of a Poisson process (left) and a compound Poisson process (right).

A recent example where Lévy processes are needed in order to build sound mathematical models are electricity prices. Veraart & Veraart [**VV14**] showed using data from the European Energy Exchange that the day-ahead prices indeed jump. For the theory of stochastic

calculus and differential equations with Lévy noise we refer to the monograph of Applebaum [**App09**] and references therein.
Note that the theory can be extended even further in this direction, if one considers a semimartingale as the basic process to build a stochastic calculus, as was first done by Meyer [**Mey76**]. An introduction into this more general theory can be found in the monographs of Métivier [**Mét82**] and Protter [**Pro05**].

Now we return to the idea from the beginning of building sound mathematical models with the help of deterministic differential equations. The standard mathematical curriculum consists of equations that are local in time. This means the dynamic depends only on the current state of the system. The path the system took to reach this state has no influence on its future development at all. Quite often those models can only be seen as a first approximation and have to be improved by allowing terms to depend on past states, in order to make them more accurate. Just think of population models in biology, delayed reaction models in chemistry, implementations of control theory models, where a feedback control is always delayed, and the incubation period, when modeling the spread of a disease. Intuitively it seems reasonable that those systems have a dependences on past states which are relevant for their future development. Allowing the evolution of a quantity to depend on its past states leads to the mathematical concept of differential delay equations or more generally to functional differential equations. We briefly discuss a simple example here to demonstrate the effect a delay can have on a system compared to the undelayed model. Therefore, consider the following logistic growth model

$$\begin{cases} u'(t) &= u(t)(1-u(t)), \\ u(0) &= 0.1. \end{cases}$$

It is well-known that the solution of this nonlinear ordinary differential equation has an exponentially damped stable equilibrium which is in our case $u(t) \equiv 1$.

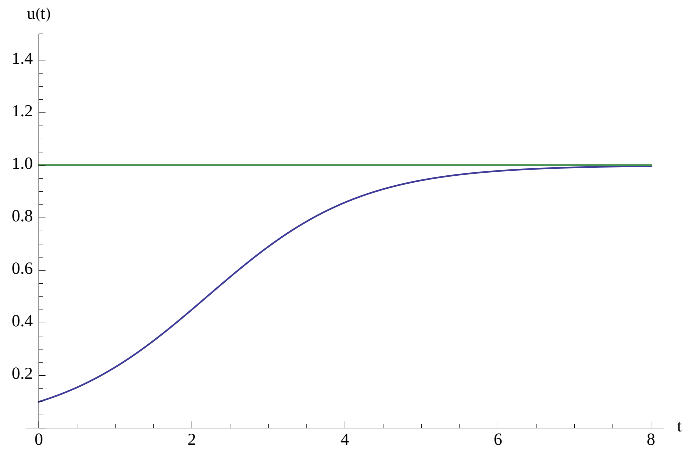

FIGURE 4. Solution of logistic growth model (blue) and stable state (green).

Let us now consider a delayed version of the logistic growth model which is discussed in Wu [**Wu96**] and reads as follows

$$\begin{cases} v'(t) &= v(t)(1 - v(t - \tau)), \\ v(s) &= 0.1, \end{cases}$$

where $s \in [-\tau, 0]$. Here, $\tau > 0$ is the delay which for example could be the duration of gestation. Depending on the size of the delay τ the system has either an exponentially damped stable equilibrium as in Figure 4, an oscillatorily damped stable equilibrium or a stable limit cycle as shown in Figure 5.

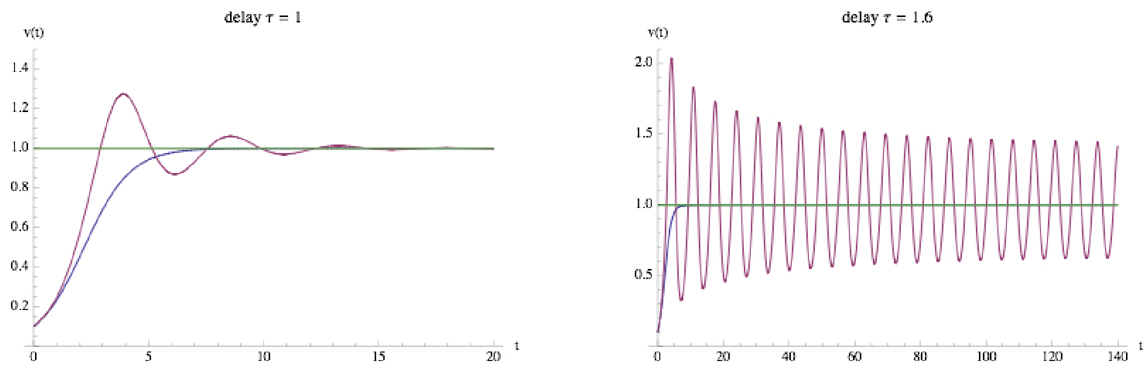

FIGURE 5. Solution of logistic growth model (blue) and stable state (green). In red the solution of the delayed logistic growth model: oscillatorily damped stable equilibrium (left) and stable limit cycle (right).

Hence, we see that a delay can have a dramatic influence onto the dynamics of a system. Historically the first general differential equations with dependence on past states of the system were investigated by Volterra [**Vol28, Vol90**]. Later on Krasovskiĭ [**Kra63**] pointed out that it may be convenient to investigate delay problems in functional spaces even though the state variable is a finite-dimensional vector. This turned out to be a very fruitful approach in developing a strong theory, that has since found application in many feels, see for example Roussel [**Rou96**] and Epstein [**Eps90**] in chemistry, Szydłowski & Krawiec [**SK01**] in economics, Alexander & Moghadas & Röst & Wu [**AMRW08**] in biology, and Makroglou & Li & Kuang [**MLK06**] in medicine. For an introduction into the general theory of functional differential equations with a finite delay and more examples we refer to Hale & Verduyn Lunel [**HVL93**] and Diekmann & van Gils & Verduyn Lunel & Walther [**DvGVLW95**] and the references therein. For differential equation with an infinite delay see Hino & Murakami & Naito [**HMN91**].

In this thesis we are interested in dynamics modeled by partial differential equations (PDE), to be more precise in PDEs which can be rewritten as evolution equations. The basic idea of evolution equations is to consider a PDE as an ordinary differential equation (ODE) in an infinite dimensional functional space. From this point of view the differential operators become linear unbounded operators in the functional space. Hence, this is one of the main differences to ODE theory, where the linear operators are all bounded. However, one can still build a rich theory which is capable of covering many important examples, where the matrix exponential from ODE theory is replaced by a one-parameter semigroup of bounded linear operators and therefore this theory is called semigroup theory. For an introduction we refer to the monographs of Engel & Nagel [**EN00**] and Pazy [**Paz83**] and the references therein.

As mentioned above, it is a suitable approach to consider delay differential equations as an evolution in a functional space. This idea also works well, if a delay appears in a PDE. There are different approaches to realize this idea. Two of them rely on semigroup theory. The first one is to consider the evolution in the history space. In this approach the history space is most often chosen to be the space of continuous functions from the delay interval to the state space of the PDE. Then, the delay problem can be rewritten as an undelayed evolution equation. We refer to Wu [**Wu96**] and the references therein for a detailed introduction into this approach. However, it turns out that this approach is not the most natural one, especially for the stochastic case we want to consider. That is why we favor a different semigroup approach, where the evolution of the delay equation is considered in the product space of the state space and the history space. Then again, the delay problem can be rewritten as an undelayed evolution equation, but this time with the following advantages: first, the structure of the equation remains intact and second, the problem can be formulated in a Hilbert space setting. For an introduction into this theory we refer to the monograph of Bátkai & Piazzera [**BP05**] and the references therein. Examples of PDEs with delay can be found in the two monographs we already mentioned. For population models with diffusion and delay we refer to Fragnelli & Tonetto [**FT04**] and Fragnelli & Idrissi & Maniar [**FIM07**].

Similar to the finite dimensional case we can extend the theory from PDEs to stochastic partial differential equations (SPDE). If one applies the same idea as in the deterministic setting, that is lifting the SPDE to a stochastic evolution equation and this way treat the SPDE as a stochastic differential equation in an infinite dimensional functional space, a rich theory was developed with the help of semigroup theory which is still an active field of research today. The first results were obtained for the Wiener noise case. Those results are summarized in the monographs of Da Prato & Zabczyk [**DPZ92**] and Gawarecki & Mandrekar [**GM11a**], where one can also find several examples of SPDEs driven by Wiener noise. However, for the same reasons as in the finite dimensional case the need of more freedom in the choice of the noise term is apparent. For example Benth & Krühner [**BK14**] recently modeled forward prices in commodity markets with the help of an infinite dimensional stochastic evolution equation driven by Lévy noise. In order to be able to treat those kind of models, an extension of the existing Wiener theory to Lévy noise is needed. One can find those results in the monograph of Peszat & Zabczyk [**PZ07**] and in the references therein. The book is also an excellent starting point for the young theory of SPDEs driven by Lévy noise which is an active field of research today, as can be seen for example in the works of Mandrekar & Wang [**MW11**], Barth & Lang [**BL12**], Albeverio & Mastrogiacomo & Smii [**AMS13**], and Hausenblas & Giri

[**HG13**]. However, the theory is not yet that far developed as it is for the Wiener noise case. For example there is no integration theory for cylindrical Lévy processes yet. First steps have been undertaken recently by Riedle [**Rie14, Rie15**] to develop this theory.

In this thesis we investigate stochastic evolution equations with Lévy noise and in this context treat stochastic partial differential delay equations driven by Lévy noise with the help of semigroup theory. As mentioned above, the deterministic approach presented by Wu [**Wu96**] is not the natural choice for a stochastic setting, especially if one considers Lévy noise, where the paths of the solution can be discontinuous. However, in the case of Wiener noise there are some results in the finite dimensional setting, see for example van Neerven & Riedle [**vNR07**]. But since the approach presented by Bátkai & Piazzera [**BP05**] has the advantage that the problem can be formulated in Hilbert spaces, it is a natural starting point for a stochastic theory. Yet, not much work has been done so far in a stochastic setting using this approach for delay equations. In the finite dimensional case with Wiener noise Chojnowska-Michalik [**CM78**] showed that a stochastic differential delay equation can be transformed into a stochastic evolution equation. For SPDEs with delay Cox & Górajski [**CG11**] proved the same equivalence for Wiener noise in a Banach space setting. Bierkens & van Gaans & Lunel [**BvGL09**] investigate the existence of an invariant measure for solutions of stochastic evolution equations with Wiener noise and mention SPDEs with delays as one example. To date, applications of stochastic differential equations with a delay are mostly found in the finite dimensional setting, for example Lu & Ding [**LD14**] in biology and Appleby & Riedle & Swords [**ARS13**] in finance. However, there is a large interest in an applicable theory of SPDEs with a delay, especially in diabetes research, where one wants to extend the ODE models with a delay, see for example Makroglou & Li & Kuang [**MLK06**], to SPDE models with a delay. Hence, mathematical research in stochastic delay equations is ongoing until today, see for example in the monographs of Mao [**Mao94**], Liu [**Liu06**], Mao & Yuan [**MY06**], Kushner [**Kus08**] and for more recent results the works of Scheutzow [**Sch13**], Xu & Pei & Li [**XPL14**], Górajski [**Gór14**], Zang & Li [**ZL14**], and Zhang & Ye & Li [**ZYL14**].

Our objective in this thesis is to provide more mathematical results for delayed and undelayed SPDEs driven by Lévy noise. Therefore, we build on the results from Bátkai & Piazzera [**BP05**], Chojnowska-Michalik [**CM78**], and Cox & Górajski [**CG11**] and show for the first time that a delayed stochastic evolution equation can be transformed into an undelayed stochastic evolution equation, if the driving noise is an infinite dimensional square integrable Lévy martingale. Thus, the semigroup approach also works well, if we consider jump noise. However, due to the stochastic term involved in the problem only weaker solution concepts work. That is why, we consider mild solutions. Yet, there is no natural definition of a mild solution, if we consider a delayed stochastic evolution equation. We resolve this problem by providing an equivalent solution concept which we call integrated solution. This way we are able to provide a natural and stochastic meaningful definition of a solution for a delayed stochastic evolution equation and link it to the mild solution of the undelayed stochastic evolution equation. Additionally, we extend the setting to include infinite delays. We achieve this by showing that an infinite delay can be naturally included in the Hilbert space setting. Therefore, we can treat delayed stochastic evolution equations with finite or infinite delay driven by Lévy noise as a special case of an undelayed stochastic evolution equation driven by Lévy noise. Hence, by well-known results from Peszat & Zabczyk [**PZ07**] the existence and uniqueness of the solution is guaranteed.

Since we are forced to work with mild solutions, we encounter the problem that they have no stochastic differentials and may not have a càdlàg modification. In other words, the mild solution of a stochastic differential equation may be a well-defined stochastic process, but it is not regular enough to apply standard tools from stochastic analysis. To overcome this problem we provide approximations of the mild solution, where each member of the approximating sequence has the desired properties and hence, is regular enough to apply stochastic tools like Itô's formula. We present two different ideas for an approximation which are both based on Yosida approximation. The general idea for the first one is to smooth all terms such that they lie in the domain of the driving linear operator, whereas in the second approximation one takes the opposite approach and approximates the mild solution with elements from the state space. While the first approximation scheme has been considered in the literature before, the proofs are incomplete. We fill this gap by giving a rigoros proof. To the best of our knowledge, the second approximation represents a new result.

In stochastic analysis the Itô formula which is called transformation formula, if one considers Lévy noise, is one of the most important tools. As mentioned above, mild solutions are not regular enough to apply classical results like the transformation formula to them. Therefore, one is in need of a generalized version which can still be applied to mild solutions. This generalized version is the main result of this thesis. That is, we prove a rigorous transformation formula for mild solutions of stochastic evolution equations driven by Lévy noise. In order to achieve this, we pick up an idea from Ahmed & Fuhrman & Zabczyk [**AFZ97**], who compensate the missing regularity of the mild solution by requesting more regularity of the transformation function. Ahmed & Fuhrman & Zabczyk mention that this idea works for the Wiener noise case, but do not give a proof. We generalize their idea to the Lévy noise case and provide all proofs, where the main tool is one of the approximation schemes we showed before. Doing so, we apply a classical transformation formula to each member of the approximating sequence. Taking the limit yields the desired transformation formula for mild solutions.

As mentioned above, the transformation formula is one of the most important tools in stochastic analysis. Hence, with the new generalized version for mild solutions at hand we can apply it to solutions of stochastic evolution equations driven by Lévy noise, too. One classical application in stochastic analysis, where the transformation formula is needed, is filtering theory. There have been numerous works on several different settings in filtering theory. The field of filtering problems with Lévy noise remains an active field of research until today, as one can see for example in Ahn & Feldman [**AF00**], Sornette & Ide [**SI01**], Meyer-Brandis & Proske [**MBP04**], Popa & Sritharan [**PS09**], Grigelionis & Mikulevicius [**GM11b**], and Frey & Schmidt & Xu [**FSX13**]. We consider a linear filtering problem with additive Lévy noise. For the finite dimensional Wiener noise case Åström [**Åst70**] showed that the Itô formula can be used to transform the filtering problem into a deterministic optimal control problem. This idea has been applied to several other situations, for example Grecksch & Tudor [**GT08**] showed an analog result in infinite dimensions with fractional Brownian motion as the noise term. We apply the idea from Åström to the case of Lévy noise in infinite dimensions. With the help of the transformation formula for mild solutions we are able to adjust the argument to our setting and hence, we show that the filtering problem in the Lévy noise case is also equivalent to a deterministic optimal control problem.

This thesis is structured as follows. In Chapter 1 we first generalize the result that a deterministic linear evolution equation with a finite delay can be transformed into an undelayed evolution equation from Bátkai & Piazzera [**BP05**] to the case of an infinite delay. Next, we show that the analog result remains true in the stochastic case with Lévy noise, where we consider a semilinear delayed stochastic evolution equation. Since stochastic strong solutions are impossible for delay equations, we consider mild solutions and equivalent formulations. From Chapter 1 we conclude that we can treat delayed stochastic evolution equations in the setting of undelayed stochastic evolution equations. That is why, in the subsequent chapters we prove results for undelayed stochastic evolution equations and then apply them to our case of interest.

In Chapter 2 we provide the two different approximations for the mild solution of a stochastic evolution equation driven by Lévy noise, we mentioned above. The first approximation covers the case, when the semigroup is a generalized contraction. In this situation the solution has a càdlàg modification. Since this continuity property is desirable when dealing with jump noise, we treat this case separately. It is well-known that generalized contraction semigroups are characterized by the Lumer-Phillips theorem to have a γ-dissipative generator. Therefore, we show for different delay types, if they are γ-dissipative or not. We end Chapter 2 with the proof of the approximation theorem for the general case. It is noteworthy as explained above, that the approximation idea as well as the proof differ considerably from the special case of a generalized contraction semigroup.

In Chapter 3 we prove the main theorem of this thesis. That is a transformation formula for mild solutions of stochastic evolution equations with Lévy noise. We proceed in three steps. At first we deduce a transformation formula for well-defined Lévy processes. As mentioned above, the mild solution is not sufficiently regular to be a well-defined Lévy process and therefore, we need to ask for more regularity of the transformation function to compensate for this lacking. That is, if A is the driving linear operator of the stochastic evolution equation we request from the transformation function $\phi : H \to \mathbb{R}$, that $\phi'(h) \in \mathscr{D}(A^*)$ for all $h \in H$, where H is a separable Hilbert space, A^* is the adjoint operator of A, and ϕ' denotes the Fréchet derivative of ϕ. We provide a sufficient criterion for that. In the final step, we prove the transformation formula for mild solutions of stochastic evolution equations with Lévy noise using the first approximations from Chapter 2. Doing so we show, if Y is the mild solution of

$$\begin{cases} dY(t) &= AY(t)dt + F(t, Y(t))dt + G_0(t, Y(t))dW_{Q_0}(t) \\ & \quad + \int_U G_1(t, Y(t))x\tilde{N}(t, dx), \quad t \geqslant 0, \\ Y(0) &= y, \end{cases}$$

then the following formula holds by Theorem 3.15 $\mathbb{P}$-a.s.

$$\begin{aligned} \phi(Y(t)) = \; & \phi(y) + \int_0^t \langle A^*\phi'(Y(s-)), Y(s-)\rangle_H ds + \int_0^t \langle \phi'(Y(s-)), F(s, Y(s-))\rangle_H ds \\ & + \int_0^t \langle \phi'(Y(s-)), G_0(s, Y(s-))dW_{Q_0}(s)\rangle_H \\ & + \frac{1}{2}\int_0^t \operatorname{tr}\left[\phi''(Y(s-))\big(G_0(s, Y(s-))Q_0^{1/2}\big)\big(G_0(s, Y(s-))Q_0^{1/2}\big)^*\right]ds \\ & + \int_0^t \int_U \phi(Y(s-) + G_1(s, Y(s-))x) - \phi(Y(s-))\tilde{N}(ds, dx) \\ & + \int_0^t \int_U \phi(Y(s-) + G_1(s, Y(s-))x) - \phi(Y(s-)) - \langle \phi'(Y(s-)), G_1(s, Y(s-))x\rangle_H \nu(dx)ds. \end{aligned}$$

Since the adjoint operator A^* of the driving linear operator A appears in the transformation formula, we close Chapter 3 by calculating the adjoint operators for the most important delay cases.

In the final Chapter 4 we apply the results from Chapter 3 to solve a linear filtering problem with additive Lévy noise. Therefore, we first prove a product formula with the help of the transformation formula for mild solutions. Then, we use this product formula to show that the filtering problem is equivalent to a deterministic optimal control problem and therefore, by classical results, has a unique solution.

Finally we decided to collect the most important results from semigroup theory and stochastic calculus we use throughout the thesis in an appendix. Therefore, a fluent reading of the thesis should be possible, without constantly consulting further literature.

CHAPTER 1

Abstract Delay Equation

In this chapter we introduce the problem which motivates the theory developed later. We start by posting a deterministic linear abstract delay equation with infinite delay. We show that, if we consider classical solutions, it can be equivalently written as an abstract Cauchy problem in a suitable product space. For the case of a finite delay this was already shown in [**BP05**]. We build on those results and use very similar arguments. The question of well-posedness is then reduced to the question of whether the driving linear operator of the abstract Cauchy problem generates a C_0-semigroup.
After that we consider a stochastic semilinear delay equation with Lévy noise. Again, we will show that it is equivalent to a stochastic abstract Cauchy problem. But this time the solution concept is that of mild solutions, which is a much weaker solution concept. This is necessary due to the stochastic term, since a strong solution of the stochastic Cauchy problem would already be deterministic. At the end of the chapter we discuss the question of well-posedness for the stochastic case.

1.1. Deterministic linear case

In [**BP05**] the authors treat an abstract linear delay equation with a finite delay. We will extend those results to the case of an infinite delay in this section. For more flexibility in the choice of history functions we introduce a time weighted history space. In particular, we will reformulate the delay problem as an abstract Cauchy problem. For the questions of well-posedness we will only cite results from [**BP05**], because the proofs are identical to the finite delay case. Since the generalization from the finite delay case to the infinite delay is not that major, we mainly repeat the argumentation of [**BP05**] and only adjust it when necessary. We also make heavy use of their notation. Our intention of this detailed review is that it provides a solid analytical foundation for the stochastic theory presented later.

1.1.1. Setting

In order to formulate the delay problem we introduce some notation and the standing hypotheses.

Definition 1.1. *With I we denote the time interval where the delay is affecting the dynamics of the delay equation. It can be either finite, that is $I = [-\tau, 0]$ for some fixed $\tau > 0$, or infinite, that is $I = \mathbb{R}_- := (-\infty, 0]$.*

Remark 1.2. *In case of a finite delay we can assume without lost of generality that $I = [-1, 0]$, by simply scaling the time.*

DEFINITION 1.3. *Let X be a Banach space and consider a function $u : I \cup \mathbb{R}_+ \to X$, where $\mathbb{R}_+ := (0, \infty)$. For each $t \geqslant 0$, we call the function*

$$u_t : I \ni \sigma \mapsto u(t + \sigma) \in X$$

history segment with respect to $t \geqslant 0$.

DEFINITION 1.4 (history function). *The history function of u is then the function*

$$h_u : t \mapsto u_t$$

on $\mathbb{R}_+$.

In order to obtain more freedom in the choice for the decay of the history function we introduce the following measure.

DEFINITION 1.5 (measure μ). *Let $\varrho \in C^1(I)$ with $\varrho > \epsilon_\varrho > 0$ and for some $T > 0$*

$$\forall \tau \in I \ \forall s \in [0, T] : \frac{\varrho(\tau - s)}{\varrho(\tau)} \leqslant C_\varrho < \infty. \tag{1.1}$$

We define the measure μ by

$$d\mu = \varrho dt,$$

where dt is the standard Lebesgue measure.

EXAMPLE 1.6. *If we set $\varrho \equiv 1$, then μ is the standard Lebesgue measure. Further examples are exponentials like $\varrho(t) = e^{-t}$ and polynomials like $\varrho(t) = 1 + (-t)^m$, where $m \geqslant 1$.*

REMARK 1.7. *Condition* (1.1) *guaranties that the weight function doesn't oscillate too much. This is natural, since normally the impact of the delay becomes smaller the further it lies in the past.*

Now we introduce the standing hypotheses which build the foundations for the theory presented here. Assume that

(H_1) X is a Banach space;
(H_2) $B : \mathscr{D}(B) \subset X \longrightarrow X$ is a closed, densely defined, linear operator;
(H_3) Z is a Banach space, such that $\mathscr{D}(B) \overset{d}{\hookrightarrow} Z \overset{d}{\hookrightarrow} X$ (where $\overset{d}{\hookrightarrow}$ means densely, continuously embedded);
(H_4) $1 \leqslant p < \infty$, $f \in L_p(I; Z; d\mu)$ and $x \in X$;
(H_5) $\Phi : W_p^1(I; Z; d\mu) \longrightarrow X$ is a bounded linear operator, called the delay operator; and
(H_6) $\mathscr{E}_p := X \times L_p(I; Z; d\mu)$.

Under these hypotheses, and for given elements $x \in X$ and $f \in L_p(I; Z; d\mu)$, the following initial value problem will be called an (abstract) delay equation (with history parameter $1 \leqslant p < \infty$)

$$(DE_p) \begin{cases} u'(t) & = & Bu(t) + \Phi u_t, \quad t \geqslant 0, \\ u(0) & = & x, \\ u_0 & = & f. \end{cases}$$

Note, that the main difference of (DE_p) to a common abstract Cauchy problem is that there are two driving linear operators. In addition to the usual operator B which acts on the state space X, there is the linear operator Φ which acts on the history space $W^1_p(I;Z;d\mu)$.
Here is the natural notation of a classical solution to (DE_p).

DEFINITION 1.8 (classical solution of (DE_p)). *We say that a function $u : I \cup \mathbb{R}_+ \longrightarrow X$ is a classical solution of (DE_p) if*

(i) $u \in C(I \cup \mathbb{R}_+; X) \ \cap \ C^1(\mathbb{R}_+; X)$,
(ii) $u(t) \in \mathscr{D}(B)$ *and* $u_t \in W^1_p(I;Z;d\mu)$ *for all* $t \geqslant 0$,
(iii) u *satisfies* (DE_p) *for all* $t \geqslant 0$.

REMARK 1.9. *Due to the properties of the weight function, ϱ, it is clear that only in the case of an infinite delay the choice of ϱ matters. For the case of a finite delay the history space is invariant of ϱ, since the norms are equivalent. This is why we always set $\varrho \equiv 1$, when the delay is finite.*

REMARK 1.10. *The abstract setting allows us to include the two most important kinds of delays. On the one hand, the discrete delay $\Phi u_t := \sum_{i=1}^n u(t-h_i)$ and on the other hand, the average over the delay: $\int_I u(t+\tau)dg(\tau)$, where g is of bounded variation. We will cover this in more detail in Section 1.1.3.*

1.1.2. Reformulation of the problem

We start by proving a lemma which follows from well-known facts about shift semigroups. It is going to be the essential tool to rewrite the delay equation (DE_p) as an abstract Cauchy problem. For the case of a finite delay one can find the results in [**BP05**, Section 3.1]. Here we prove all the results for the case of an infinite delay.

LEMMA 1.11. *In the case of an infinite delay, that is $I = \mathbb{R}_-$, let $u : \mathbb{R} \to Z$ be a function such that for all $a \in \mathbb{R}$ u belongs to $W^1_p((-\infty, a); Z; d\mu)$. Then, the history function $h_u : t \to u_t$ of u is continuously differentiable from $\mathbb{R}_+$ into $L_p(I;Z;d\mu)$ with derivative*

$$\frac{d}{dt}h_u(t) = \frac{d}{d\sigma}u_t.$$

PROOF.

Let $(A, \mathscr{D}(A))$ be the generator of the left shift semigroup $(T(t))_{t\geqslant 0}$ on the space $L_p(\mathbb{R};Z;d\mu)$, that is $\mathscr{D}(A) = W^1_p(\mathbb{R};Z;d\mu)$ and $A = \frac{d}{d\sigma}$. Let $t \in \mathbb{R}_+$ and fix $T > 1$.We extend $u_{|(-\infty,t+T]}$ to a function $v \in W^1_p(\mathbb{R};Z;d\mu) = \mathscr{D}(A)$, such that

$$\frac{d}{dt}T(t)v = AT(t)v.$$

Note that for $\sigma \in I = \mathbb{R}_-$ the identity

$$(T(s)v)(\sigma) = v(s+\sigma) = u(s+\sigma) = u_s(\sigma) = h_u(s)(\sigma)$$

holds if and only if $s + \sigma \in (-\infty, t + T]$. Therefore the identity holds for all $s \in \mathbb{R}$ with $-\infty < s - t \leqslant T$, in particular it holds for $|s - t| < T$. This way we find

$$\begin{aligned}
0 &= \lim_{h\to 0} \left\| \frac{T(t+h)v - T(t)v}{h} - AT(t)v \right\|^p_{L_p(\mathbb{R};Z;d\mu)} \\
&= \lim_{h\to 0} \left\| \frac{T(t+h)v - T(t)v}{h} - \frac{d}{d\sigma} T(t)v \right\|^p_{L_p(\mathbb{R};Z;d\mu)} \\
&\geqslant \lim_{h\to 0} \left\| \frac{T(t+h)v - T(t)v}{h} - \frac{d}{d\sigma} T(t)v \right\|^p_{L_p(I;Z;d\mu)} \\
&= \lim_{h\to 0} \left\| \frac{h_u(t+h) - h_u(t)}{h} - \frac{d}{d\sigma} u_t \right\|^p_{L_p(I;Z;d\mu)},
\end{aligned}$$

which implies

$$\frac{d}{dt} h_u(t) = \frac{d}{d\sigma} u_t.$$

Moreover, the map $t \mapsto \frac{d}{d\sigma} u_t = \frac{d}{dt} h_u(t)$ is continuous from $\mathbb{R}_+$ into $L_p(I; Z; d\mu)$, since the map $t \mapsto AT(t)v = T(t)Av$ is continuous from $\mathbb{R}_+$ into $L_p(\mathbb{R}; Z; d\mu)$.

□

As a direct consequence of Lemma 1.11 we can transform classical solutions of (DE_p) into classical solutions of an abstract Cauchy problem.

COROLLARY 1.12. *Let $u : I \cup \mathbb{R}_+ \to X$ be a classical solution of (DE_p). Then the function*

$$\mathscr{U} : t \mapsto \begin{pmatrix} u(t) \\ u_t \end{pmatrix} \in \mathscr{E}_p$$

from $\mathbb{R}_+$ into $\mathscr{E}_p$ is continuously differentiable with derivative

$$\dot{\mathscr{U}}(t) = \mathscr{A}\mathscr{U}(t),$$

where

$$\mathscr{A} := \begin{pmatrix} B & \Phi \\ 0 & \frac{d}{d\sigma} \end{pmatrix},$$

where $\frac{d}{d\sigma}$ denotes the distributional derivative with domain

$$\mathscr{D}(\mathscr{A}) := \left\{ \begin{pmatrix} x \\ f \end{pmatrix} \in \mathscr{D}(B) \times W^1_p(I; Z; d\mu) : f(0) = x \right\}.$$

Thus every classical solution u of (DE_p) yields a classical solution of the abstract Cauchy problem

$$(ACP_p) \begin{cases} \mathscr{U}'(t) &= \mathscr{A}\mathscr{U}(t), \quad t \geqslant 0, \\ \mathscr{U}(0) &= \begin{pmatrix} x \\ f \end{pmatrix}, \end{cases}$$

on $\mathscr{E}_p$.

REMARK 1.13. *In the definition of the domain of $\mathscr{A}$ we find the expression $f(0) = x$. This means that we must be able to evaluate f at zero. This is possible since $W^1_p(\mathbb{R}_-; Z)$ is embedded in $C_0(\mathbb{R}_-; Z)$. With the properties of the measure μ (continuity and strict positivity of ϱ) one easily finds that $W^1_p(\mathbb{R}_-; Z; d\mu)$ is also embedded in $C_0(\mathbb{R}_-; Z)$.*

REMARK 1.14. *In the first line of the operator matrix $\mathscr{A}$ we find the delay equation $(DE)_p$. Since the history function is a shift in time the derivative in the second line is natural and we have shown this in Lemma 1.11. But if one thinks in terms of partial differential equations, the second line could be interpreted as a transport equation in time.*

Corollary 1.12 shows that every solution of (DE_p) gives us a solution of (ACP_p). Next, we show that $(DE)_p$ and (ACP_p) are equivalent in the sense that conversely every classical solution $t \mapsto \mathscr{U}(t)$ of (ACP_p) is of the form

$$\mathscr{U}(t) = \begin{pmatrix} u(t) \\ u_t \end{pmatrix},$$

where the function u is a classical solution of (DE_p). We fix the Banach space setting for (ACP_p) by adding the following to our standing hypotheses:

(H_7) $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ is the operator on $\mathscr{E}_p$ defined as

$$\mathscr{A} := \begin{pmatrix} B & \Phi \\ 0 & \frac{d}{d\sigma} \end{pmatrix},$$

with domain

$$\mathscr{D}(\mathscr{A}) := \left\{ \begin{pmatrix} x \\ f \end{pmatrix} \in \mathscr{D}(B) \times W^1_p(I; Z; d\mu) : f(0) = x \right\}.$$

Since we want to apply standard arguments from semigroup theory concerning the connection of well-posedness of abstract Cauchy problems and generators of C_0-semigroups (see Appendix A in particular Theorem A.9), we need to show the closeness of the operator $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$.

LEMMA 1.15. *Under Hypotheses $(H_1) - (H_7)$, the operator $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ is closed and densely defined on $\mathscr{E}_p$.*

PROOF.

First we prove the closedness. Let $\binom{x_n}{f_n} \subset \mathscr{D}(\mathscr{A})$ be a sequence such that $\binom{x_n}{f_n}$ converges to $\binom{x}{f} \in \mathscr{E}_p$ and $\mathscr{A}\binom{x_n}{f_n} = \binom{Bx_n + \Phi f_n}{\frac{d}{d\sigma} f_n}$ converges to $\binom{y}{g} \in \mathscr{E}_p$.
In particular, the sequence (f_n) converges to f in the norm topology of the Sobolev space $W^1_p(I; Z; d\mu)$. Hence, we have that $f \in W^1_p(I; Z; d\mu)$ and $\frac{d}{d\sigma} f = g$. Since the operator $\Phi : W^1_p(I : Z; d\mu) \to X$ is bounded, we have that $\Phi f_n \longrightarrow \Phi f$.
Moreover, by the closedness of B, we have $x \in \mathscr{D}(B)$ and $Bx = y - \Phi f$. Finally, since the space $W^1_p(I; Z, d\mu)$ is embedded into $C(I,Z)$ (see Remark 1.13), the sequence $x_n = f_n(0)$ converges to $f(0)$. Hence, $f(0) = x$ and $\binom{x}{f} \in \mathscr{D}(\mathscr{A})$, $\mathscr{A}\binom{x}{f} = \binom{y}{g}$ and the operator $\mathscr{A}$ is closed.

Now, we prove the density of $\mathscr{D}(\mathscr{A})$. Let $\binom{y}{g} \in \mathscr{E}_p$ and $\varepsilon > 0$. First note that $\mathscr{D}(B)$ is dense in X, thus we can find a $x \in \mathscr{D}(B)$ such that

$$\|x - y\|_X < \varepsilon.$$

Next, we show that $W^1_{p,0}(I;Z;d\mu)$ is dense in $L_p(I;Z;d\mu)$, where $W^1_{p,0}(I;Z;d\mu)$ is the closure of $C_c^\infty(I;Z)$ in $W^1_p(I;Z;d\mu)$. Now let $u \in L_p(I;Z;d\mu)$. This is equivalent to $u\varrho^{1/p} \in L_p(I;Z)$. Since $C_c^\infty(I;Z)$ is dense in $L_p(I;Z)$ we find a sequence $u_n \in C_c^\infty(I;Z)$ such that

$$u_n \longrightarrow u\varrho^{1/p} \text{ in } L_p(I;Z).$$

This implies that the sequence $\tilde{u}_n := u_n\varrho^{-1/p}$ converges to u in $L_p(I;Z;d\mu)$. Since $\varrho \in C^1$ and is always positive we find that $\tilde{u}_n \in C^1_c(I;Z)$. All that is left to show now is that each $\tilde{u}_n \in W^1_{p,0}(I;Z;d\mu)$. But this is easy to see since

$$(\tilde{u}_n(t))' = (u_n(t)\varrho^{-1/p}(t))' = u_n'(t)\varrho^{-1/p}(t) - \frac{1}{p}u_n(t)\varrho^{-1/p}(t)\varrho^{-1}(t)\varrho'(t)$$

and all u_n's and their derivatives have compact support. Thus, we can find $\tilde{g} \in W^1_{p,0}(I;Z;d\mu)$, such that

$$\|\tilde{g} - g\|_p < \varepsilon.$$

Let now $h \in W^1_p(I;Z;d\mu)$ such that $h(0) = x$ and let $k \in W^1_{p,0}(I;Z;d\mu)$ such that $\|k-h\|_p < \varepsilon$. Finally, let $f := \tilde{g} + h - k$. We obtain $f \in W^1_p(I;Z;d\mu)$, $f(0) = x$, and

$$\left\| \begin{pmatrix} y \\ g \end{pmatrix} - \begin{pmatrix} x \\ f \end{pmatrix} \right\| \leqslant \left\| \begin{pmatrix} y - x \\ 0 \end{pmatrix} \right\| + \left\| \begin{pmatrix} 0 \\ g - \tilde{g} \end{pmatrix} \right\| + \left\| \begin{pmatrix} 0 \\ h - k \end{pmatrix} \right\| < 3\varepsilon$$

Therefore, the domain $\mathscr{D}(\mathscr{A})$ is dense.

□

In view of Lemma 1.15, we can formulate the following corollary that is a straightforward consequence of semigroup theory (see Theorem A.9).

Corollary 1.16. *The abstract Cauchy problem (ACP_p) associated to the operator $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ on the space $\mathscr{E}_p$ is well-posed if and only if $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ is the generator of a C_0-semigroup $(\mathscr{T}(t))_{t\geqslant 0}$ on $\mathscr{E}_p$.*
In this case, the classical and mild solutions of (ACP_p) are given by the functions

$$\mathscr{U}(t) = \mathscr{T}(t)\begin{pmatrix} x \\ f \end{pmatrix}$$

for $t \geqslant 0$.

Now, we introduce the following notation.

Definition 1.17. *If $\mathscr{U}$ is an element of $\mathscr{E}_p$ we denote its first component by $\mathscr{U}_1 \in X$. Therefore, one can interpret $\mathscr{U}_1$ as the canonical projection from $\mathscr{E}_p$ onto X.*
Similarly, by $\mathscr{U}_2 \in L_p(I;Z;d\mu)$ we denote the second component of $\mathscr{U}$. Again, this could be interpreted as the canonical projection from $\mathscr{E}_p$ onto $L_p(I;Z;d\mu)$.

With this notation in place we can write an element $\mathscr{U}$ *of* $\mathscr{E}_p$ *in the following way*

$$\mathscr{U} = \begin{pmatrix} \mathscr{U}_1 \\ \mathscr{U}_2 \end{pmatrix}.$$

PROPOSITION 1.18. *For every classical solution* $\mathscr{U}$ *of* (ACP_p)*, the function*

$$u(t) := \begin{cases} \mathscr{U}_1(t) & \text{if } t \geqslant 0 \\ f(t) & \text{if } t \in I \end{cases} \tag{1.2}$$

is a classical solution of (DE_p) *and* $\mathscr{U}_2(t) = u_t$ *for all* $t \geqslant 0$.

PROOF.

Since $\mathscr{U}$ is a classical solution of (ACP_p), i.e. $\mathscr{U} \in C^1(\mathbb{R}_+; \mathscr{E}_p) \cap C(\mathbb{R}_+; \mathscr{D}(\mathscr{A}))$, it follows that $\mathscr{U}_2$ is in $C^1(\mathbb{R}_+; L_p(I; Z; d\mu))$ and is a classical solution of the problem

$$\begin{cases} \frac{d}{dt}\mathscr{U}_2(t) &= \frac{d}{d\sigma}\mathscr{U}_2(t), \quad t \geqslant 0, \\ \mathscr{U}_2(t)(0) &= \mathscr{U}_1(t), \qquad t \geqslant 0, \\ \mathscr{U}_2(0) &= f \end{cases} \tag{1.3}$$

in the space $L_p(I; Z; d\mu)$. In particular, since $L_p(I; Z; d\mu) \overset{d}{\hookrightarrow} L_p(I; X; d\mu)$, the function $\mathscr{U}_2$ is in $C^1(\mathbb{R}_+; L_p(I; X; d\mu))$ and is a classical solution of the problem in Equation (1.3) in the space $L_p(I; X; d\mu)$. Note that since

$$\mathscr{U}(t) \in \mathscr{D}(\mathscr{A}) = \left\{ \begin{pmatrix} x \\ f \end{pmatrix} \in \mathscr{D}(B) \times W_p^1(I; Z; d\mu) : f(0) = x \right\}.$$

for all $t \geqslant 0$ it follows that $\mathscr{U}_2(t)(0) = \mathscr{U}_1(t)$.
Now we observe that by definition

$$u_t(\sigma) = u(t+\sigma) = \begin{cases} \mathscr{U}_1(t+\sigma) & \text{for } t+\sigma \geqslant 0, \\ f(t+\sigma) & \text{for } t+\sigma < 0, \end{cases}$$

where $f \in W_p^1(I; Z; d\mu) \overset{d}{\hookrightarrow} W_p^1(I; X; d\mu)$, and $f(0) = x = \mathscr{U}_1(0)$ by assumption. Hence, in the case of an infinite delay $u \in W_p^1((-\infty, a); X; d\mu)$ for all $a \in \mathbb{R}$. We can extend u to a function in $W_{p,loc}^1(\mathbb{R}; X; d\mu)$ and by Lemma 1.11 we have

$$\frac{d}{dt}h_u(t) = \frac{d}{d\sigma}u_t \qquad \text{for all } t \geqslant 0$$

in the space $L_p(I; X; d\mu)$. Moreover, by definition of u_t we have

$$u_t(0) = u(t) = \mathscr{U}_1(t) \qquad \text{for all } t \geqslant 0,$$

and

$$u_0 = f.$$

Hence, the map $t \mapsto u_t$ is also a classical solution of the problem (1.3) in the space $L_p(I; X; d\mu)$.

Now we define $w(t) := u_t - \mathscr{U}_2(t)$ for $t \geqslant 0$. Then w is a classical solution of the problem

$$\left\{\begin{array}{lcll} \frac{d}{dt}w(t) & = & \frac{d}{d\sigma}w(t), & t \geqslant 0, \\ w(t)(0) & = & 0, & t \geqslant 0, \\ w(0) & = & 0 & \end{array}\right. \tag{1.4}$$

in the space $L_p(I; X; d\mu)$. Since Equation (1.4) is the abstract Cauchy problem associated to the generator of the (nilpotent) left shift semigroup on $L_p(I; X; d\mu)$ with initial value zero, we have that $w(t) = 0$ for all $t \geqslant 0$. Therefore, $u_t = \mathscr{U}_2(t) \in W_p^1(I; Z; d\mu)$ and $\mathscr{U}(t) = \begin{pmatrix} u(t) \\ u_t \end{pmatrix}$ for all $t \geqslant 0$, and u is a classical solution of (DE_p).

□

The equivalence of (DE_p) and (ACP_p) established above enables us to use methods and results of semigroup theory in order to deal with the delay problem (DE_p).
At present, we transfer the notions of well-posedness and of mild solution, known from abstract Cauchy problems and semigroups, to (DE_p) (see Definition A.10 and Definition A.12).

DEFINITION 1.19 (well-posedness and mild solution of (DE_p)).

(i) *The problem (DE_p) is called well-posed if (ACP_p) is well-posed, that is if $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ generates a C_0-semigroup on $\mathscr{E}_p$.*

(ii) *Suppose (DE_p) is well-posed and let $(\mathscr{T}(t))_{t\geqslant 0}$ be the semigroup generated by the operator $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ on $\mathscr{E}_p$. Then for every $x \in X$ and every $f \in L_p(I; Z; d\mu)$ the function u defined by Equation (1.2) is called a mild solution of (DE_p).*

The following proposition is the equivalent for (DE_p) to Proposition A.13 for (ACP).

PROPOSITION 1.20. *Let u be a mild solution of (DE_p). Then u satisfies $\int_0^t u(s)ds \in \mathscr{D}(B)$, $\int_0^t u_s ds \in W_p^1(I; Z; d\mu)$, and the integral equation*

$$u(t) = \left\{\begin{array}{ll} x + B\int_0^t u(s)ds + \Phi \int_0^t u_s ds & \textit{for} \quad t \geqslant 0, \\ f(t) & \textit{for} \quad \textit{a.e. } t \in I. \end{array}\right. \tag{1.5}$$

PROOF.
(1) First, we show that

$$u_t = \left(\mathscr{T}(t)\begin{pmatrix} x \\ f \end{pmatrix}\right)_2 \tag{1.6}$$

for every $\begin{pmatrix} x \\ f \end{pmatrix} \in \mathscr{E}_p$ and every $t \geqslant 0$.
For $\begin{pmatrix} x \\ f \end{pmatrix} \in \mathscr{D}(\mathscr{A})$ the Identity (1.6) holds by Proposition 1.18. Now, take $\begin{pmatrix} x \\ f \end{pmatrix} \in \mathscr{E}_p$ and a sequence $\begin{pmatrix} x_n \\ f_n \end{pmatrix} \in \mathscr{D}(\mathscr{A})$ converging to $\begin{pmatrix} x \\ f \end{pmatrix}$. Since the semigroup $(\mathscr{T}(t))_{t\geqslant 0}$ is strongly continuous, the sequence $\mathscr{T}(t)\begin{pmatrix} x_n \\ f_n \end{pmatrix}$ converges to $\mathscr{T}(t)\begin{pmatrix} x \\ f \end{pmatrix}$ in $\mathscr{E}_p$ uniformly for t in compact subsets of $[0, \infty)$.
Now, let

$$u_n(t) := \left\{\begin{array}{lll} \left(\mathscr{T}(t)\begin{pmatrix} x_n \\ f_n \end{pmatrix}\right)_1 & \text{if} & t \geqslant 0, \\ f_n(t) & \text{if} & t \in I. \end{array}\right.$$

Since $\binom{x_n}{f_n} \in \mathscr{D}(\mathscr{A})$, by Proposition 1.18 we have $(u_n)_t = \left(\mathscr{T}(t)\binom{x_n}{f_n}\right)_2$.

For fixed $t \geqslant 0$ and $\sigma \in [-t, 0]$, we have that

$$(u_n)_t(\sigma) = u_n(t+\sigma) = \left(\mathscr{T}(t+\sigma)\begin{pmatrix} x_n \\ f_n \end{pmatrix}\right)_1 \tag{1.7}$$

converges to $\left(\mathscr{T}(t+\sigma)\binom{x}{f}\right)_1 = u_t(\sigma)$ uniformly for $\sigma \in [-t, 0]$. Hence, $(u_n)_t$ converges to u_t in $L_p([-t,0]; X; d\mu)$ and we have

$$u_t = \lim_{n\to\infty}(u_n)_t = \lim_{n\to\infty}\left(\mathscr{T}(t)\begin{pmatrix} x_n \\ f_n \end{pmatrix}\right)_2 = \left(\mathscr{T}(t)\begin{pmatrix} x \\ f \end{pmatrix}\right)_2.$$

In particular, $u_t \in L_p([-t,0]; Z; d\mu)$. From Equation (1.2) we find that for any $t \geqslant 0$

$$(u_n)_t(\sigma) := \begin{cases} \left(\mathscr{T}(t+\sigma)\binom{x_n}{f_n}\right)_1 & \text{for} \quad \sigma \in [-t, 0], \\ f_n(t+\sigma) & \text{for} \quad \sigma \in (-\infty, -t). \end{cases}$$

This formula in conjunction with the calculation given in Equation (1.7) implies that $(u_n)_t(\sigma)$ converges to $\left(\mathscr{T}(t+\sigma)\binom{x}{f}\right)_1$ uniformly for $\sigma \in [-t, 0]$. Moreover, by assumption, $(u_n)_t$ converges to $f(t+\cdot)$ in $L_p((-\infty,-t); Z; d\mu)$. Hence, $(u_n)_t$ converges to u_t in $L^p(I; X; d\mu)$ and, by the same argument as above, $u_t = \left(\mathscr{T}(t)\binom{x}{f}\right)_2$.

(2) Take the first component of the identity

$$\mathscr{T}(t)\begin{pmatrix} x \\ f \end{pmatrix} - \begin{pmatrix} x \\ f \end{pmatrix} = \mathscr{A}\int_0^t \mathscr{T}(s)\begin{pmatrix} x \\ f \end{pmatrix} ds, \quad t \geqslant 0,$$

to obtain Equation (1.5).

□

THEOREM 1.21. *The following assertions are equivalent:*

(i) (DE_p) *is well-posed.*

(ii) For every $\binom{x}{f} \in \mathscr{D}(\mathscr{A})$,

- *(a) there is a unique (classical) solution* $u(x, f, \cdot)$ *of* (DE_p) *and*
- *(b) the solutions depend continuously on the initial values, that is, if a sequence* $\binom{x_n}{f_n}$ *in* $\mathscr{D}(\mathscr{A})$ *converges to* $\binom{x}{f} \in \mathscr{D}(\mathscr{A})$ *in the space* $\mathscr{E}_p = X \times L_p(I; Z; d\mu)$, *then* $u(x_n, f_n, t)$ *converges to* $u(x, f, t)$ *in* X *uniformly for t in compact intervals.*

PROOF.

First we show $(ii) \implies (i)$. Assume that for every $\binom{x_n}{f_n} \in \mathscr{D}(\mathscr{A})$, Equation (DE_p) has a unique solution u. Then, Corollary 1.12 guarantees that for every $\binom{x_n}{f_n} \in \mathscr{D}(\mathscr{A})$ the abstract Cauchy problem (ACP_p) has a classical solution which is unique. It is easy to see that the solution depends continuously on the initial values. Finally, by Lemma 1.15, $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ is closed and densely defined. Therefore, it generates a C_0-semigroup on $\mathscr{E}_p$ by Theorem A.9.

Conversely, if $\mathscr{A}$ is a generator, we have by Corollary 1.16 and Proposition 1.18 that for every initial value $\binom{x_n}{f_n} \in \mathscr{D}(\mathscr{A})$ there is a unique solution u of (DE_p) that is given by Equation (1.2). This implies that the solution depends continuously on the initial values.

□

1.1.3. Well-posedness for the linear deterministic delay problem

We know from the previous section that the question of well-posedness is equivalent to the question:

When does the operator $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ generates a C_0-semigroup on $\mathscr{E}_p$?

Criteria that answer this question are given in [**BP05**, Section 3.3 and Section 3.4] for the case of a finite delay. Since the proofs don't change for the case of an infinite delay we will only state the results here and provide citations for the proofs.

First, we consider the case where the operator B is bounded or, in particular, the space X is finite-dimensional. Furthermore, we assume that $Z = X$ and therefore have a bounded operator in the delay term. Which leads us to the following theorem.

THEOREM 1.22. *If $B \in L(X)$ and $\Phi : W^{1,p}(I; X; d\mu) \longrightarrow X$ is a bounded operator, then the operator matrix*

$$\mathscr{A} := \begin{pmatrix} B & \Phi \\ 0 & \frac{d}{d\sigma} \end{pmatrix}$$

with domain

$$\mathscr{D}(\mathscr{A}) := \left\{ \begin{pmatrix} x \\ f \end{pmatrix} \in X \times W_p^1(I; X; d\mu) : f(0) = x \right\}$$

generates a C_0-semigroup on the space $\mathscr{E}_p$ for all $1 \leqslant p < \infty$.

PROOF.

See [**BP05**, Throrem 3.23.].

Now, we consider the case where the operator B is unbounded and generates the C_0-semigroup $(S(t))_{t\geqslant 0}$ on X. In the delay term we keep a bounded operator for now, that is $Z = X$. We need some more notation before we can formulate the theorem. We denote by $(T_0(t))_{t\geqslant 0}$ the nilpotent left shift semigroup on $L_p(I; X; d\mu)$ and $S_t : X \to L_p(I; X; d\mu)$ is defined by

$$(S_t x)(\tau) := \begin{cases} S(t+\tau)x & \text{if} \quad -t < \tau \leqslant 0, \\ 0 & \text{if} \quad \tau \leqslant -t. \end{cases}$$

Now, we are able to formulate the following theorem.

THEOREM 1.23. *Suppose the operator* $(B, \mathscr{D}(B))$ *generates the* C_0*-semigroup* $(S(t))_{t \geqslant 0}$ *on* X *and let* $\Phi : W_p^1(I; X; d\mu)) \to X$ *be a delay operator, where* $1 \leqslant p < \infty$*. Furthermore, assume that there exist constants* $t_0 > 0$ *and* $0 \leqslant q < 1$ *such that*

$$\int_0^{t_0} \|\Phi(S_r x + T_0(r) f)\|_X dr \leqslant q \left\| \begin{pmatrix} x \\ f \end{pmatrix} \right\|_{\mathscr{E}_p}$$

for all $\begin{pmatrix} x \\ f \end{pmatrix} \in \mathscr{D}(\mathscr{A})$*. Then, the operator* $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ *is the generator of a* C_0*-semigroup on* $\mathscr{E}_p$ *and* (DE_p) *is well-posed.*

PROOF.
See [**BP05**, Throrem 3.26.].

A large class of delay operators is cover in the following important application of Theorem 1.23. Therefore, let $\eta : I \to L(X)$ be of bounded variation and let $\Phi : C(I; X) \to X$ be the bounded linear operator given by the Riemann-Stieltjes Integral,

$$\Phi(f) := \int_I d\eta f. \tag{1.8}$$

Since $W_p^1(I; X; d\mu)$ is continuously embedded in $C(I, X)$, we may note that Φ defines a bounded operator from $W_p^1(I; X; d\mu)$ to X.

THEOREM 1.24. *Let* $(B, \mathscr{D}(B))$ *be the generator of a* C_0*-semigroup* $(S(t))_{t \geqslant 0}$ *on* X *and let* Φ *be given by* (1.8)*, where* $1 \leqslant p < \infty$*. Then, the operator* $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ *is the generator of a* C_0*-semigroup on* $\mathscr{E}_p$ *and* (DE_p) *is well-posed.*

PROOF.
See [**BP05**, Throrem 3.29.].

REMARK 1.25. *From Theorem 1.23 and Theorem 1.24 we see that if we introduce a delay into a well-posed linear nondelayed Cauchy problem it remains well-posed as long as the delay operator has a nice structure.*

From Theorem 1.23 and Theorem 1.24 we deduce, like mentioned in the beginning of the section, that the most important delay operators are included in the setting.

EXAMPLE 1.26. *Let* $B_k \in L(X)$ *and* $h_k \in [-1, 0]$ *for each* $k = 1, \ldots, n$*. When the discrete delay operator defined by*

$$\Phi(f) := \sum_{k=1}^{n} B_k f(-h_k), \;\; f \in W_p^1([-1, 0]; X),$$

is of the form (1.8)*. Thus, Theorem 1.24 can be applied.*

If $h \in L^q(I; L(X))$, $1 < q \leqslant \infty$, the averaging delay operator Φ (also called distributed delay) is given by

$$\Phi(f) := \int_I h(\sigma) f(\sigma) d\sigma.$$

Then Φ is a bounded operator from $L^p(I; X) \to X$ and Theorem 1.23 can be applied.

Finally, we consider the case of an unbounded delay operator. That is Φ is an bounded operator from $W_p^1(I; Z; d\mu)$ to X with $Z \subsetneq X$. We have to strengthen the assumption on the operator B in this case. In particular, we assume that

$$(B, \mathscr{D}(B)) \text{ generates an analytic semigroup } (S(t))_{t \geqslant 0} \text{ on } X \tag{1.9}$$

and that for some $\delta > \omega_0(B)$ (the growth bound of B), there exists $\vartheta < \frac{1}{p}$, such that

$$\mathscr{D}((-B + \delta)^{\vartheta}) \overset{d}{\hookrightarrow} Z \overset{d}{\hookrightarrow} X. \tag{1.10}$$

Then, we can formulate the counterpart to Theorem 1.23.

THEOREM 1.27. *Suppose the operator $(B, \mathscr{D}(B))$ fulfills conditions* (1.9) *and* (1.10) *and let $\Phi : W_p^1(I; Z; d\mu)) \to X$ be a delay operator, where $1 \leqslant p < \infty$. Assume that there exist constants $t_0 > 0$ and $0 \leqslant q < 1$ such that*

$$\int_0^{t_0} \|\Phi(S_r x + T_0(r) f)\|_X dr \leqslant q \left\| \begin{pmatrix} x \\ f \end{pmatrix} \right\|_{\mathscr{E}_p}$$

for all $\begin{pmatrix} x \\ f \end{pmatrix} \in \mathscr{D}(\mathscr{A})$. Then the operator $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ is the generator of a C_0-semigroup on $\mathscr{E}_p$ and (DE_p) is well-posed.

PROOF.
See [**BP05**, Throrem 3.34.].

Similar to the case of a bounded delay operator we can show an important special case, which is the counter part to Theorem 1.24.

THEOREM 1.28. *Let $1 \leqslant p < \infty$ ad let $\eta : I \to L(Z, X)$ be of bounded variation. Let $\Phi : C(I, Z) \to X$ be the bounded linear operator given by the Rieman-Stieltjes Integral*

$$\Phi(f) := \int_I d\eta f.$$

Suppose the operator $(B, \mathscr{D}(B))$ fulfills conditions (1.9) *and* (1.10). *Then the operator $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ is the generator of a C_0-semigroup on $\mathscr{E}_p$ and (DE_p) is well-posed.*

PROOF.
See [**BP05**, Throrem 3.35.].

We close this section by providing some examples of delayed differential equations from [**BP05**, Section 3.1 and Section 3.3] that fit into the setting.

Example 1.29. *First we consider an example in finite space dimensions. Let* $X = \mathbb{C}$ *and consider the initial value problem*

$$\begin{cases} u'(t) &= \int_I h(\sigma)u'(t+\sigma)d\sigma \text{ for } t \geqslant 0, \\ u(0) &= x, \\ u_0 &= f, \end{cases}$$

where

- $u : I \cup \mathbb{R}_+ \to \mathbb{C}$ *is a function,*
- $1 \leqslant p < \infty$ *and* $1 < q \leqslant \infty$ *are such that* $\frac{1}{p} + \frac{1}{q} = 1$,
- $h \in L^q(I)$
- $x \in \mathbb{C}$ *and* $f \in L^p(I)$.

This equation is well-posed by Theorem 1.22.

Next, we consider a heat equation on an open and bounded domain $G \subset \mathbb{R}^n$ *with smooth boundary and Neumann boundary conditions.*

$$\begin{cases} \partial_t u(t,s) &= \Delta u(t,s) + \sum_{i=1}^n c_i \partial_i u(t-h_i,s), & t \geqslant 0,\ s \in G, \\ \frac{\partial u}{\partial \nu}(t,s) &= 0, & t \geqslant 0, s \in \partial G, \\ u(t,s) &= f(t,s), & t \in [-1,0], s \in G, \end{cases} \tag{1.11}$$

for some constants $c_i \in \mathbb{R}$ *and* $h_i \in [0,1]$. *Moreover, assume that* $f \in L_2(I \times G)$. *In order to write* (1.11) *as an abstract delay problem, we introduce*

- $X := L_2(G)$,
- *the operator* B *is defined by* $\mathscr{D}(B) := \{g \in W_2^1(G) : \Delta g \in L_2(G) \text{ and } \frac{\partial g}{\partial \nu} = 0 \text{ on } \partial G\}$ *and* $Bf := \Delta f$,
- *the space* $Z := W_2^1(G)$,
- *the delay operator* $\Phi : W_2^1([-1,0]; Z) \to X$ *defined as*

  $$\Phi f := \sum_{i=1^n} c_i \partial_i f(-h_i),$$

  and
- $x := f(0,\cdot)$.

Then equation (1.11) *is well-posed by Theorem 1.28.*

The next example is a linear diffusion equation with delayed reaction term in an open and bounded domain $G \subset \mathbb{R}^n$ *with smooth boundary.*

$$\begin{cases} \partial_t u(t,s) &= \Delta u(t,s) + c\int_I u(t+\sigma,s)dg(\sigma), & t \geqslant 0,\ s \in G, \\ u(t,s) &= 0, & t \geqslant 0, s \in \partial G, \\ u(t,s) &= f(t,s), & t \in I, s \in G, \end{cases} \tag{1.12}$$

where c is a constant and $g : I \to [0, 1]$, is a function of bounded variation. Moreover, assume that $f \in L_2(I \times G)$. In order to write (1.12) as an abstract delay problem, we introduce

- $X := L_2(G) =: Z$,
- *the operator $(B, \mathscr{D}(B))$ as the variational Laplacian with Dirichlet boundary conditions,*
- *the delay operator $\Phi : W_2^1(I; X) \to X$ defined as*

$$\Phi f := c \int_I f(\sigma) dg(\sigma).$$

Then, equation (1.12) is well-posed by Theorem 1.24.

Finally, we present a second-order equation with delay. We consider the following one-dimensional wave equation on $(0, 1)$. For simplicity we write $H_0^1(0, 1)$ instead of $W_{2,0}^1(0, 1)$.

$$\begin{cases} \partial_t^2 u(t, s) & = \Delta u(t, s) + c_1 \partial_s u(t - h_1, s) + c_2 \partial_t u(t - h_2, s), & t \geqslant 0,\ s \in (0, 1), \\ u(t, s) & = f(t, s),\ \partial_t u(t, s) = g(t, s), & t \in [0, -1], s \in (0, 1), \\ u(t, 0) & = u(t, 1) = 0, & t \in I, s \in G, \end{cases} \tag{1.13}$$

where we assume the following for the initial data

- $f(0, \cdot) \in H_0^1(0, 1)$ *and* $g(0, \cdot) \in L_2(0, 1)$
- *the map $t \mapsto f(t, \cdot)$ is in $L_2([-1, 0]; H_0^1(0, 1))$, and*
- *the map $t \mapsto g(t, \cdot)$ is in $L_2([-1, 0]; L_2(0, 1))$*

In order to write (1.13) as an abstract delay problem, we introduce

- $X := H_0^1(0, 1) \times L_2(0, 1) =: Z$,
- *the operator* $B := \begin{pmatrix} 0 & Id \\ \Delta & 0 \end{pmatrix}$
  with domain $D(B) := (H_0^1(0, 1) \cap H^2(0, 1)) \times H_0^1(0, 1)$
- *the function $\mathbb{R}_+ \ni t \mapsto u(t) = u(t, \cdot) \in L_2(0, 1)$, and*
- *the delay operator $\Phi : W_2^1(I; X) \to X$ defined as*

  $$\Phi \begin{pmatrix} x \\ f \end{pmatrix} := \begin{pmatrix} 0 & 0 \\ c_1 \partial_x \delta_{-h_1} & c_2 \delta_{-h_2} \end{pmatrix} \begin{pmatrix} x \\ f \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ c_1 \partial_x g(-h_1) & c_2 h(-h_2) \end{pmatrix},$$

  where δ_{-h_1} and δ_{-h_2} are the point evaluations in $-h_1$ and $-h_2$, respectively.

Then, equation (1.13) is well-posed by Theorem 1.24.

1.2. Stochastic semilinear case with Lévy noise

In this section, we introduce the main problem of this work. We build on the model presented in the previous section and extend it to a semilinear stochastic delay problem. The driving stochastic process will be a Hilbert space valued square-integrable Lévy-martingale. Our goal is to show that also in the stochastic case the transformation of the stochastic delay problem into a stochastic Cauchy problem is possible. Due to the stochastic nature of the problem we will have to use weaker solution concepts, since our setting doesn't allow for a meaningful strong solution. That is why we work with mild solutions. We start by introducing the complete setting for the stochastic case and then give definitions for the solutions of the two problems. After that we will show that they are equivalent. At the end of the section, we discuss the question of well-posedness.

1.2.1. Setting

In the stochastic case we consider the evolution on the finite time interval $[0, T]$ with $T > 0$ and we define the entire time interval of the past and the evolution time to be $\mathcal{I} := I \cup [0, T]$. Furthermore, we have to reformulate the problem in a Hilbert space setting. Therefore, we restate the standing hypotheses for the stochastic case. Assume, that

- (SH_1) H is a separable Hilber space;
- (SH_2) $B : \mathscr{D}(B) \subset H \longrightarrow H$ is a closed, densely defined, linear operator;
- (SH_3) Z is a Hilbert space such that $\mathscr{D}(B) \overset{d}{\hookrightarrow} Z \overset{d}{\hookrightarrow} H$;
- (SH_4) $f \in L_2(I; Z; d\mu)$ and $h \in H$;
- (SH_5) $\Phi : W_2^1(I; Z; d\mu) \longrightarrow H$ is a bounded linear operator, called the delay operator;
- (SH_6) $\mathscr{E}_2 := H \times L_2(I; Z; d\mu)$;
- (SH_7) $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ is the operator on $\mathscr{E}_2$ defined as

$$\mathscr{A} := \begin{pmatrix} B & \Phi \\ 0 & \frac{d}{d\sigma} \end{pmatrix},$$

with domain

$$\mathscr{D}(\mathscr{A}) := \left\{ \begin{pmatrix} h \\ f \end{pmatrix} \in \mathscr{D}(B) \times W_2^1(I; Z; d\mu) : f(0) = h \right\};$$

- (SH_8) $(\Omega, \mathcal{F}, \mathcal{F}_{t\geqslant 0}, \mathbb{P})$ is a filtered complete probability space;
- (SH_9) U is a separable Hilbert space and $(M(t))_{t\geqslant 0}$ is a U-valued square-integrable Lévy martingale equipped with its natural filtration $(\mathcal{F}_t)_{t\geqslant 0}$ and with the covariance operator Q;
- (SH_{10}) $F : \Omega \times [0, T] \times \mathscr{E}_2 \rightarrow H$ is a $\mathcal{P}_T \otimes \mathscr{B}(\mathscr{E}_2)/\mathscr{B}(H)$-measurable function satisfying Lipschitz and growth conditions, that is for all $t \in [0, T]$ and $u, v \in \mathscr{E}_2$ there exists a constant $C_F > 0$, such that for a.e. $\omega \in \Omega$

$$\|F(\omega, t, u) - F(\omega, t, v)\|_H \leqslant C_F \|u - v\|_{\mathscr{E}_2},$$

and there exists a constant k_F such that for a.e. $\omega \in \Omega$ and all $t \in [0,T],\ u \in \mathscr{E}_2$ it holds

$$\|F(\omega,t,u)\|_H \leqslant k_F(1+\|u\|_{\mathscr{E}_2}).$$

(SH_{11}) G is a map from $\Omega \times [0,T] \times \mathscr{E}_2$, which takes values in the set of all linear operators from U into H, such that

$$\Omega \times [0,T] \times \mathscr{E}_2 \ni (\omega,t,v) \mapsto G(\omega,t,v)u$$

is a $\mathcal{P}_T \otimes \mathscr{B}(\mathscr{E}_2)/\mathscr{B}(H)$-measurable function for all $u \in U$ and $GQ^{\frac{1}{2}}$ takes values in $L_{(HS)}(U,H)$, where the space $L_{(HS)}(U,H)$ is the space of all Hilbert-Schmidt operators from U to H. Furthermore, we assume G to be Lipschitz and linear bounded, that is there exists a constant $C_G > 0$ such that for a.e. $\omega \in \Omega$

$$\|(G(\omega,t,u)-G(\omega,t,v))Q^{\frac{1}{2}}\|_{L_{(HS)}(U,H)} \leqslant C_G\|u-v\|_{\mathscr{E}_2}$$

for all $t \in [0,T]$ and $u,v \in \mathscr{E}_2$, and there exists a constant k_G such that for a.e. $\omega \in \Omega$ and all $t \in [0,T],\ u \in \mathscr{E}_2$ it holds

$$\|G(\omega,t,u)Q^{\frac{1}{2}}\|_{L_{(HS)}(U,H)} \leqslant k_G(1+\|u\|_{\mathscr{E}_2}).$$

REMARK 1.30. *For the definition of $\mathcal{P}_T$ and of the stochastic integral with respect to a square-integrable Lévy process see Appendix B in particular Section B.3.*

Under these hypotheses, and for given elements $h \in H$ and $f \in L_2(I;Z;d\mu)$, the following initial value problem will be called a stochastic (abstract) delay equation

$$(S-DE)\begin{cases} dX(t) &= BX(t)dt + \Phi X_t dt + F(t,X(t),X_t)dt + G(t,X(t),X_t)dM(t),\ t \geqslant 0,\\ X(0) &= h,\\ X_0 &= f.\end{cases}$$

Our goal is to show that, just like in the deterministic case, a solution of problem $(S-DE)$ is equivalent to a solution of the following stochastic abstract Cauchy problem

$$(S_D-ACP)\begin{cases} dY(t) &= \mathscr{A}Y(t)dt + \mathscr{F}(t,Y(t))dt + \mathcal{G}(t,Y(t))d\mathcal{M}(t), \quad t \geqslant 0,\\ Y(0) &= \begin{pmatrix} h \\ f \end{pmatrix},\end{cases}$$

where

$$\mathcal{M}(t) := \begin{pmatrix} M(t) \\ 0 \end{pmatrix},\ \mathscr{F}(t,u) := \begin{pmatrix} F(t,u_1,u_2) \\ 0 \end{pmatrix},\ \text{and}\ \mathcal{G}(t,u) := \begin{pmatrix} G(t,u_1,u_2) & 0 \\ 0 & 0 \end{pmatrix}.$$

1.2.2. Equivalence of $(S-DE)$ and (S_D-ACP)

Before we can go ahead and show the claimed equivalents we have to give suitable definitions of the solution for each problem. But here we encounter the first problem. It turns out that, even for the Wiener noise case, the well-know definition of a stochastic strong solution for a stochastic abstract Cauchy problem like (S_D-ACP) of Da Prato and Zabczyk [**DPZ92**] can already imply that the solution is almost surely deterministic. To illustrate this we provide an example here which is discussed in [**CG11**].

Suppose we can show, just like we already did for the deterministic case, that if we have a strong solution Y of $(S_D - ACP)$ (see [**DPZ92**, Chapter 5, 6, and 7] for more detail), that the first component of Y solves $(S - DE)$ and the second component is the history function, that is

$$Y_1(t) = X(t) \text{ and } Y_2(t) = X_t \ \mathbb{P}\text{-a.s.},$$

then the following proposition shows that a strong solution in the sense of Da Prato and Zabczyk of $(S_D - ACP)$ is already deterministic.

PROPOSITION 1.31. *Let $H = \mathbb{R}$ $(= Z)$, $\varrho \equiv 1$, $\mathscr{F} \equiv 0$, $\mathcal{G}(\omega, t, y) = G(y)$ for all $\omega \in \Omega$, $t \in [0, T]$, $y \in \mathcal{E}_2$, and let Y be a strong solution of $(S_D - ACP)$. In particular, we have $Y(s) \in \mathscr{D}(\mathscr{A})$ for all $s \in [0, T]$ $\mathbb{P}$-a.s., then $\mathscr{T}(t)\binom{h}{f} \in \mathcal{N}(\mathcal{G})$ (the null space of $\mathcal{G}$) and $Y(s) = \mathscr{T}(s)\binom{h}{f}$ $\mathbb{P}$-a.s. for almost all $s \in [0, T]$, that is $(S_D - ACP)$ is deterministic.*

PROOF.

See [**CG11**, Proposition 4.13].

Proposition 1.31 shows that we have to ask for less regularity in the definition of a solution, in order to have meaningful stochastic objects to investigate. That is why we work with mild solutions. Thus, we define for problem $(S_D - ACP)$

$$(S_D - ACP) \begin{cases} dY(t) &= \mathscr{A}Y(t)dt + \mathscr{F}(t, Y(t))dt + \mathcal{G}(t, Y(t))d\mathcal{M}(t), \quad t \geqslant 0, \\ Y(0) &= \begin{pmatrix} h \\ f \end{pmatrix}, \end{cases}$$

the mild solution as follows.

DEFINITION 1.32 (mild solution for $(S_D - ACP)$). *A stochastic process $Y : \Omega \times [0, T] \to \mathscr{E}_2$ is called a mild solution of $(S_D - ACP)$, if Y is a predictable $\mathscr{E}_2$-valued process satisfying*

$$\sup_{t \in [0,T]} \mathbb{E}\|Y(t)\|^2_{\mathscr{E}_2} < \infty, \tag{1.14}$$

such that for every $t \in [0, T]$ we have $\mathbb{P}$-a.s.

$$Y(t) = \mathscr{T}(t)\begin{pmatrix} h \\ f \end{pmatrix} + \int_0^t \mathscr{T}(t-s)\mathscr{F}(s, Y(s))ds + \int_0^t \mathscr{T}(t-s)\mathcal{G}(s, Y(s))d\mathcal{M}(s). \tag{1.15}$$

REMARK 1.33. *Note that a mild solution of $(S_D - ACP)$ doesn't need to have a càdlàg modification.*

In order to work out a suitable definition for the solution of $(S - DE)$, we recall the equivalence of mild solutions in the deterministic setting, that is in particular Proposition A.13. This motivates the next definition, the so call integrated solution for $(S_D - ACP)$. It is our goal to show the equivalence of mild and integrated for the Lévy noise case (Theorem 1.35), which is the stochastic analog to Proposition A.13. Note that similar results for the Wiener case can be found in [**CG11**].

DEFINITION 1.34 (integrated solution of $(S_D - ACP)$). *A stochastic process $Y : \Omega \times [0,T] \to \mathscr{E}_2$ is called an integrated solution of $(S_D - ACP)$, if Y is a predictable $\mathscr{E}_2$-valued process which is $\mathbb{P}$-a.s. locally Bochner integrable, satisfying*

$$\sup_{t\in[0,T]} \mathbb{E}\|Y(t)\|_{\mathscr{E}_2}^2 < \infty,$$

and for all $t \in [0,T]$ we have

(i) $\int_0^t Y(s)ds \in \mathscr{D}(\mathscr{A})$ $\mathbb{P}$-a.s.,

(ii) $\mathcal{G}(\cdot, Y)$ *is stochastically integrable on $[0,T]$, and $\mathbb{P}$-a.s.*

$$Y(t) - \begin{pmatrix} h \\ f \end{pmatrix} = \mathscr{A} \int_0^t Y(s)ds + \int_0^t \mathscr{F}(s, Y(s))ds + \int_0^t \mathcal{G}(s, Y(s))d\mathcal{M}(s).$$

THEOREM 1.35 (equivalence of mild and integrated solution of $(S_D - ACP)$). *Consider problem $(S_D - ACP)$. Then Y is a mild solution if and only if Y is an integrated solution.*

PROOF.

Before we can go ahead and prove the equivalents of the two solution we show two helpful equalities. We start by applying the stochastic Fubini theorem (see Theorem B.43) and receive $\mathbb{P}$-a.s.

$$\begin{aligned}\int_0^t \mathscr{T}(t-s) \int_0^s \mathcal{G}(\tau, Y(\tau))d\mathcal{M}(\tau)\ ds &= \int_0^t \int_0^s \mathscr{T}(t-s)\mathcal{G}(\tau, Y(\tau))d\mathcal{M}(\tau)\ ds \\ &= \int_0^t \int_\tau^t \mathscr{T}(t-s)\mathcal{G}(\tau, Y(\tau))ds\ d\mathcal{M}(\tau). \qquad (1.16)\end{aligned}$$

Note, that for all $t \geqslant 0$ the operator $\int_0^t \mathscr{T}(s)\mathcal{G}(\tau, Y(\tau))ds \in L(\mathcal{E}_2)$ is defined by

$$\left(\int_0^t \mathscr{T}(s)\mathcal{G}(\tau, Y(\tau))ds\right) e := \int_0^t \mathscr{T}(s)\mathcal{G}(\tau, Y(\tau))e\ ds,\ \ e \in \mathcal{E}_2.$$

Observe that $\int_\tau^t \mathscr{T}(t-s)\mathcal{G}(\tau, Y(\tau))e\ ds \in \mathscr{D}(\mathscr{A})$ $\mathbb{P}$-a.s. for all $e \in H$ and $0 \leqslant \tau < t$ by Lemma A.6 *(iii)*. Then, Proposition B.42 shows that the stochastic integral in (1.16) is in $\mathscr{D}(\mathscr{A})$ and by Lemma A.6 *(iv)* we have $\mathbb{P}$-a.s.

$$\begin{aligned}\mathscr{A} \int_0^t \int_\tau^t \mathscr{T}(t-s)\mathcal{G}(\tau, Y(\tau))ds\ d\mathcal{M}(\tau) &= \int_0^t \mathscr{A}\left[\int_\tau^t \mathscr{T}(t-s)\mathcal{G}(\tau, Y(\tau))ds\right] d\mathcal{M}(\tau) \\ &= \int_0^t [\mathscr{T}(t-\tau) - I]\,\mathcal{G}(\tau, Y(\tau))d\mathcal{M}(\tau).\end{aligned}$$

Looking back at (1.16), we obtain $\mathbb{P}$-a.s.

$$\mathscr{A} \int_0^t \mathscr{T}(t-s) \int_0^s \mathcal{G}(\tau, Y(\tau))d\mathcal{M}(\tau)\ ds = \int_0^t [\mathscr{T}(t-\tau) - I]\,\mathcal{G}(\tau, Y(\tau))d\mathcal{M}(\tau). \qquad (1.17)$$

With a similar calculation we derive the second equality, that is $\mathbb{P}$-a.s.

$$\begin{aligned}\mathscr{A}\int_0^t\int_0^s \mathscr{T}(s-\tau)\mathcal{G}(\tau,Y(\tau))d\mathcal{M}(\tau)\ ds &= \mathscr{A}\int_0^t\int_\tau^t \mathscr{T}(s-\tau)\mathcal{G}(\tau,Y(\tau))ds\ d\mathcal{M}(\tau)\\ &= \int_0^t \mathscr{A}\left[\int_\tau^t \mathscr{T}(s-\tau)\mathcal{G}(\tau,Y(\tau))ds\right]d\mathcal{M}(\tau)\\ &= \int_0^t [\mathscr{T}(t-\tau)-I]\,\mathcal{G}(\tau,Y(\tau))d\mathcal{M}(\tau). \qquad (1.18)\end{aligned}$$

Note that we can perform similar calculations for the terms involving the nonlinearity $\mathscr{F}$. Now suppose that Y is a mild solution of (S_D-ACP), then $\int_0^t Y(s)ds$ exists by (1.14). Furthermore, by Lemma A.6 (iii), (1.18) and its analog for the term involving $\mathscr{F}$ show that $\int_0^t Y(s)ds$ is in $\mathscr{D}(\mathscr{A})$. By applying (1.18), its analog for the $\mathscr{F}$-term, and Lemma A.6 (iv) we find $\mathbb{P}$-a.s.

$$\begin{aligned}\mathscr{A}\int_0^t Y(s)ds &= \mathscr{A}\int_0^t \mathscr{T}(s)\begin{pmatrix} h\\ f\end{pmatrix}ds + \mathscr{A}\int_0^t\int_0^s \mathscr{T}(s-\tau)\mathscr{F}(\tau,Y(\tau))d\tau\ ds\\ &\quad + \mathscr{A}\int_0^t\int_0^s \mathscr{T}(s-\tau)\mathcal{G}(\tau,Y(\tau))d\mathcal{M}(\tau)\ ds\\ &= \mathscr{T}(t)\begin{pmatrix} h\\ f\end{pmatrix} - \begin{pmatrix} h\\ f\end{pmatrix} + \int_0^t [\mathscr{T}(t-\tau)-I]\,\mathscr{F}(\tau,Y(\tau))d\tau\\ &\quad + \int_0^t [\mathscr{T}(t-\tau)-I]\,\mathcal{G}(\tau,Y(\tau))d\mathcal{M}(\tau)\\ &= Y(t) - \begin{pmatrix} h\\ f\end{pmatrix} - \int_0^t \mathscr{F}(\tau,Y(\tau))d\tau - \int_0^t \mathcal{G}(\tau,Y(\tau))d\mathcal{M}(\tau),\end{aligned}$$

which shows that Y is also an integrated solution of (S_D-ACP). On the other hand, if Y is an integrated solution we apply (1.17) and its analog for the term involving $\mathscr{F}$ and receive $\mathbb{P}$-a.s.

$$\begin{aligned}Y(t)-&\begin{pmatrix} h\\ f\end{pmatrix} - \mathscr{A}\int_0^t Y(s)ds\\ &= \int_0^t \mathscr{F}(s,Y(s))ds + \int_0^t \mathcal{G}(s,Y(s))d\mathcal{M}(s)\\ &= \int_0^t \mathscr{T}(t-\tau)\mathscr{F}(\tau,Y(\tau))d\tau + \int_0^t \mathscr{T}(t-\tau)\mathcal{G}(\tau,Y(\tau))d\mathcal{M}(\tau)\\ &\qquad - \mathscr{A}\int_0^t \mathscr{T}(t-s)\int_0^s \mathscr{F}(\tau,Y(\tau))d\tau\ ds\\ &\qquad - \mathscr{A}\int_0^t \mathscr{T}(t-s)\int_0^s \mathcal{G}(\tau,Y(\tau))d\mathcal{M}(\tau)\ ds. \qquad (1.19)\end{aligned}$$

Let us look at the last term in the equality above without the $\mathscr{A}$. Applying the definition of the integrated solution, the closeness of $\mathscr{A}$, properties (ii) and (iv) of Lemma A.6, and

Fubinis theorem we find $\mathbb{P}$-a.s.

$$\begin{aligned}
&\int_0^t \mathscr{T}(t-s)\int_0^s \mathcal{G}(\tau,Y(\tau))d\mathcal{M}(\tau)\ ds\\
&\quad=\int_0^t \mathscr{T}(t-s)\left[Y(s)-\begin{pmatrix} h\\ f\end{pmatrix}-\mathscr{A}\int_0^s Y(\tau)d\tau-\int_0^s \mathscr{F}(\tau,Y(\tau))d\tau\right]ds\\
&\quad=\int_0^t \mathscr{T}(t-s)Y(s)ds-\int_0^t \mathscr{T}(t-s)\begin{pmatrix} h\\ f\end{pmatrix}ds\\
&\qquad-\mathscr{A}\int_0^t\int_0^s \mathscr{T}(t-s)Y(\tau)d\tau\ ds-\int_0^t \mathscr{T}(t-s)\int_0^s \mathscr{F}(\tau,Y(\tau))d\tau\ ds\\
&\quad=\int_0^t \mathscr{T}(t-s)Y(s)ds-\int_0^t \mathscr{T}(t-s)\begin{pmatrix} h\\ f\end{pmatrix}ds\\
&\qquad-\mathscr{A}\int_0^t\int_\tau^t \mathscr{T}(t-s)Y(\tau)ds\ d\tau-\int_0^t \mathscr{T}(t-s)\int_0^s \mathscr{F}(\tau,Y(\tau))d\tau\ ds\\
&\quad=\int_0^t \mathscr{T}(t-s)Y(s)ds-\int_0^t \mathscr{T}(t-s)\begin{pmatrix} h\\ f\end{pmatrix}ds\\
&\qquad-\int_0^t \left[\mathscr{T}(t-s)-I\right]Y(s)ds-\int_0^t \mathscr{T}(t-s)\int_0^s \mathscr{F}(\tau,Y(\tau))d\tau\ ds\\
&\quad=-\int_0^t \mathscr{T}(t-s)\begin{pmatrix} h\\ f\end{pmatrix}ds+\int_0^t Y(s)ds-\int_0^t \mathscr{T}(t-s)\int_0^s \mathscr{F}(\tau,Y(\tau))d\tau\ ds.
\end{aligned}$$

Returning to (1.19), using the just shown identity and property (iv) of Lemma A.6 once more we obtain $\mathbb{P}$-a.s.

$$\begin{aligned}
&Y(t)-\begin{pmatrix} h\\ f\end{pmatrix}-\mathscr{A}\int_0^t Y(s)ds\\
&=\int_0^t \mathscr{T}(t-\tau)\mathscr{F}(\tau,Y(\tau))d\tau+\int_0^t \mathscr{T}(t-\tau)\mathcal{G}(\tau,Y(\tau))d\mathcal{M}(\tau)\\
&\quad-\mathscr{A}\int_0^t \mathscr{T}(t-s)\int_0^s \mathscr{F}(\tau,Y(\tau))d\tau\ ds\\
&\quad+\mathscr{A}\left[\int_0^t \mathscr{T}(t-s)\begin{pmatrix} h\\ f\end{pmatrix}ds-\int_0^t Y(s)ds+\int_0^t \mathscr{T}(t-s)\int_0^s \mathscr{F}(\tau,Y(\tau))d\tau\ ds\right]\\
&=\int_0^t \mathscr{T}(t-\tau)\mathscr{F}(\tau,Y(\tau))d\tau+\int_0^t \mathscr{T}(t-\tau)\mathcal{G}(\tau,Y(\tau))d\mathcal{M}(\tau)\\
&\qquad\qquad+\mathscr{T}(t)\begin{pmatrix} h\\ f\end{pmatrix}-\begin{pmatrix} h\\ f\end{pmatrix}-\mathscr{A}\int_0^t Y(s)ds.
\end{aligned}$$

Hence, Y is also a mild solution of (S_D-ACP).

□

With the just shown equivalence of mild and integrated solution for $(S_D - ACP)$ and Proposition 1.20 in mind, we can motivate for the stochastic delay problem $(S - DE)$

$$(S-DE)\begin{cases} dX(t) &= BX(t)dt + \Phi X_t dt + F(t,X(t),X_t)dt + G(t,X(t),X_t)dM(t),\ t \geqslant 0,\\ X(0) &= h,\\ X_0 &= f,\end{cases}$$

the following solution definition.

Definition 1.36 (integrated solution for $(S - DE)$). *A stochastic process $X : \Omega \times \mathcal{I} \to H$ is called an integrated solution to $(S - DE)$, if X is a predictable H-valued process and one has*

(i) $\sup_{t\in[0,T]} \mathbb{E}\|X(t)\|_H^2 < \infty$,
(ii) $\mathbb{E}\int_0^T \|X(t)\|_Z^2 dt < \infty$,
(iii) $X_0 = f$,
(iv) $\int_0^t X(s)ds \in \mathscr{D}(B)$ *for all* $t \in [0,T]$ $\mathbb{P}$-*a.s. and* $\mathbb{P}$-*a.s.*

$$X(t) - h = B\int_0^t X(s)ds + \Phi\int_0^t X_s ds + \int_0^t F(s,X(s),X_s)ds + \int_0^t G(s,X(s),X_s)dM(s).$$

In order to see that all terms appearing in the definition given above are well-defined we prove the following Lemma.

Lemma 1.37. *Let $t > 0$, $\mathcal{J} := I \cup [0,t]$ and $x \in L_2(\mathcal{J};Z;d\mu)$, where we set $\varrho(\tau) = 1$ if $\tau > 0$, then the history function $h_x : [0,t] \to L_2(I;Z;d\mu)$, $h_x(s) := x_s$ is (Bochner) integrable with*

$$\int_0^t h_x(s)ds \in W_2^1(I;Z;d\mu),\ and\ \frac{d}{d\sigma}\left(\int_0^t h_x(s)ds\right) = h_x(t) - h_x(0)$$

PROOF.
First, we show the integrability. Hence, for some $s \in [0,T]$ we look at

$$\|h_x(s)\|^2_{L_2(I;Z;d\mu)} = \int_I \|h_x(s)(\sigma)\|_Z^2 d\mu(\sigma) = \int_I \|x(s+\sigma)\|_Z^2 \varrho(\sigma)d\sigma.$$

Now we substitute $\tau = s + \sigma$ and apply the properties of the function ϱ. This way we receive in the case of infinite/finite delay

$$\begin{aligned}
&\|h_x(s)\|^2_{L_2(I;Z;d\mu)}\\
&= \int_{-\infty/s-1}^s \|x(\tau)\|_Z^2\varrho(\tau-s)d\tau\\
&= \int_{-\infty/\min\{0,s-1\}}^0 \|x(\tau)\|_Z^2\varrho(\tau-s)d\tau + \int_{0/\max\{0,s-1\}}^s \|x(\tau)\|_Z^2\varrho(\tau-s)d\tau\\
&\leqslant \int_{-\infty/\min\{0,s-1\}}^0 \|x(\tau)\|_Z^2 \frac{\varrho(\tau-s)}{\varrho(\tau)}\varrho(\tau)d\tau + \big(\max_{s\in[-t,0]}\varrho(s)\big)\int_{0/\max\{0,s-1\}}^s \|x(\tau)\|_Z^2 d\tau\\
&\leqslant C_\varrho\int_{-\infty/\min\{0,s-1\}}^0 \|x(\tau)\|_Z^2 d\mu(\tau) + \big(\max_{s\in[-t,0]}\varrho(s)\big)\int_{0/\max\{0,s-1\}}^t \|x(\tau)\|_Z^2 d\tau\\
&\leqslant (C_\varrho + \max_{s\in[-t,0]}\varrho(s))\int_{\mathcal{J}} \|x(\tau)\|_Z^2 d\mu(\tau) = (C_\varrho + \max_{s\in[-t,0]}\varrho(s))\|x\|^2_{L_2(\mathcal{J};Z;d\mu)}.
\end{aligned}$$

This shows the integrability of h_x. Now we turn our attention to the derivative of the integral. First, note that by the definition of the integral and the history function we have

$$\left(\int_0^t h(s)ds\right)(\sigma) = \left(\int_0^t x_s ds\right)(\sigma) = \int_0^t x(s+\sigma)ds.$$

Thus, for the derivative we find

$$\begin{aligned}\left\|\frac{d}{d\sigma}\int_0^t h_x(s)ds\right\|^2_{L_2(I;Z;d\mu)} &= \int_I \left\|\frac{d}{d\sigma}\left(\int_0^t h_x(s)ds\right)(\sigma)\right\|^2_Z d\mu(\sigma) \\ &= \int_I \left\|\frac{d}{d\sigma}\left(\int_0^t x(s+\sigma)ds\right)\right\|^2_Z d\mu(\sigma) = \int_I \left\|\frac{d}{d\sigma}\left(\int_\sigma^{t+\sigma} x(\tau)d\tau\right)\right\|^2_Z d\mu(\sigma) \\ &= \int_I \|x(t+\sigma)-x(\sigma)\|^2_Z d\mu(\sigma) = \int_I \|h_x(t)(\sigma)-h_x(0)(\sigma)\|^2_Z d\mu(\sigma) \\ &= \|h_x(t)-h_x(0)\|^2_{L_2(I;Z;d\mu)}.\end{aligned}$$

The last term in this equality is bounded by $4(C_\varrho + \max_{s\in[-t,0]}\varrho(s))\|x\|^2_{L_2(\mathcal{J};Z;d\mu)}$. This yields

$$\int_0^t h_x(s)ds \in W_2^1(I;Z;d\mu).$$

□

Remark 1.38. *By condition (ii) in the definition of a solution for $(S-DE)$, hypothesis (SH_4), and Lemma 1.37 we see that $\mathbb{P}$-a.s. $\int_0^t X_s ds \in W_2^1(I;Z;d\mu)$, such that all terms in the definition in Definition 1.36 are well-defined. Note also that for infinite/finite delay we have*

$$\sup_{t\in[0,T]} \mathbb{E}\|X_s\|^2_{L_2(I;Z;d\mu)} \leqslant \big(C_\varrho + \max_{s\in[-T,0]}\varrho(s)\big)\big(\|f\|^2_{L_2(I;Z;d\mu)} + \int_0^T \|X(\tau)\|^2_Z d\tau\big) < \infty.$$

Remark 1.39. *Condition (ii) in Definition 1.36 looks like we require for a lot of regularity from the solution, since we take the norm in Z here and not in the state space H. But recall that the space Z is only needed if the delay operator is unbounded (with respect to the space H). In this case the operator B has to fulfill stronger assumptions in order for $\mathscr{A}$ to generate a C_0-semigroup. In particular, we require that B is a generator of an analytic semigroup. In the case of a bounded delay operator we always have $Z = H$ and condition (i) implies (ii).*

Remark 1.40. *A more straight forward definition for the solution of $(S-DE)$ would have been using the C_0-semigroup generated by the operator B and convolve the delay term as well as the nonlinear terms. But this kind of definition for a solution is not useful for the approach we want to develop here.*

Before we can show the equivalence of the solution of the two problems $(S_D - ACP)$ and $(S-DE)$ we need one auxiliary result. Therefore, recall Definition 1.17 for the index notation.

Corollary 1.41. *If Y is a mild solution of $(S_D - ACP)$, we have for all $\sigma \in I$ and $t > -\sigma$ $\mathbb{P}$-a.s.*

$$Y_2(t)(\sigma) = Y_1(t+\sigma).$$

PROOF.

Recall Proposition 1.20, there we showed

$$\left(\mathscr{T}(t)\begin{pmatrix} h \\ f \end{pmatrix}\right)_2(\sigma) = \left(\mathscr{T}(t+\sigma)\begin{pmatrix} h \\ f \end{pmatrix}\right)_1,$$

for all $(h,f)^T \in \mathcal{E}_2$, $\sigma \in I$, and $t+\sigma > 0$. For $t+\sigma \leqslant 0$ the expression above is equal to $f(t+\sigma)$, that is

$$\left(\mathscr{T}(t)\begin{pmatrix} h \\ f \end{pmatrix}\right)_2(\sigma) = f(t+\sigma),$$

for all $(h,f)^T \in \mathcal{E}_2$, $\sigma \in I$, and $t+\sigma \leqslant 0$. Applying those equalities to the convolution involving $\mathscr{F}$ in (1.15) we find for $\sigma \in I$ $\mathbb{P}$-a.s.

$$\begin{aligned}
\left(\int_0^t \mathscr{T}(t-\tau)\mathscr{F}(\tau,Y(\tau))d\tau\right)_2(\sigma) &= \int_0^{t+\sigma}\left(\mathscr{T}(t-\tau)\begin{pmatrix} F(\tau,Y_1(\tau),Y_2(\tau)) \\ 0 \end{pmatrix}\right)_2(\sigma)d\tau \\
&\quad + \int_{t+\sigma}^{t}\left(\mathscr{T}(t-\tau)\begin{pmatrix} F(\tau,Y_1(\tau),Y_2(\tau)) \\ 0 \end{pmatrix}\right)_2(\sigma)d\tau \\
&= \int_0^{t+\sigma}\left(\mathscr{T}(t-\tau)\begin{pmatrix} F(\tau,Y_1(\tau),Y_2(\tau)) \\ 0 \end{pmatrix}\right)_2(\sigma)d\tau \\
&= \int_0^{t+\sigma}\left(\mathscr{T}(t+\sigma-\tau)\begin{pmatrix} F(\tau,Y_1(\tau),Y_2(\tau)) \\ 0 \end{pmatrix}\right)_1 d\tau \\
&= \left(\int_0^{t+\sigma}\mathscr{T}(t+\sigma-\tau)\mathscr{F}(\tau,Y(\tau))d\tau\right)_1.
\end{aligned}$$

For the stochastic convolution in (1.15) we apply the approximation of the stochastic integral from Proposition B.44. This way we find for $\sigma \in I$ $\mathbb{P}$-a.s.

$$\begin{aligned}
\left(\int_0^t \mathscr{T}(t-\tau)\mathcal{G}(\tau,Y(\tau))d\mathcal{M}(\tau)\right)_2(\sigma) &= \left(\sum_{j=1}^{\infty}\int_0^t \mathscr{T}(t-\tau)\mathcal{G}(\tau,Y(\tau))\begin{pmatrix} e_j \\ 0 \end{pmatrix}dm_j(\tau)\right)_2(\sigma) \\
&= \sum_{j=1}^{\infty}\int_0^{t+\sigma}\left(\mathscr{T}(t-\tau)\mathcal{G}(\tau,Y(\tau))\begin{pmatrix} e_j \\ 0 \end{pmatrix}\right)_2(\sigma)dm_j(\tau) \\
&\quad + \int_{t+\sigma}^{t}\left(\mathscr{T}(t-\tau)\mathcal{G}(\tau,Y(\tau))\begin{pmatrix} e_j \\ 0 \end{pmatrix}\right)_2(\sigma)dm_j(\tau) \\
&= \sum_{j=1}^{\infty}\int_0^{t+\sigma}\left(\mathscr{T}(t-\tau)\mathcal{G}(\tau,Y(\tau))\begin{pmatrix} e_j \\ 0 \end{pmatrix}\right)_2(\sigma)dm_j(\tau) \\
&= \sum_{j=1}^{\infty}\int_0^{t+\sigma}\left(\mathscr{T}(t+\sigma-\tau)\mathcal{G}(\tau,Y(\tau))\begin{pmatrix} e_j \\ 0 \end{pmatrix}\right)_1 dm_j(\tau) \\
&= \left(\int_0^{t+\sigma}\mathscr{T}(t+\sigma-\tau)\mathcal{G}(\tau,Y(\tau))d\mathcal{M}(\tau)\right)_1.
\end{aligned}$$

□

The next theorem shows that the solutions of $(S_D - ACP)$ and $(S - DE)$ are equivalent in the following sense.

THEOREM 1.42 (equivalence of solutions of $(S_D - ACP)$ and $(S - DE)$).

(i) *Let X be an integrated solution of $(S - DE)$, then the stochastic process Y defined by $Y(t) := (X(t), X_t)^T$ is an integrated solution of $(S_D - ACP)$ and hence it is also a mild solution of $(S_D - ACP)$.*

(ii) *On the other hand, if Y is an integrated solution of $(S_D - ACP)$ (hence also a mild solution), then the stochastic process defined by $X_0 := f$, $X(t) := Y_1(t)$ for $t \geqslant 0$ is an integrated solution of $(S - DE)$.*

PROOF.

(i) Let X be an integrated solution to $(S - DE)$ and we define $Y(t) := (X(t), X_t)^T$. Thus,

$$\sup_{t \in [0,T]} \mathbb{E}\|Y(t)\|^2_{\mathscr{E}_2} \leqslant \sup_{t \in [0,T]} \mathbb{E}\|X(s)\|^2_H + \sup_{t \in [0,T]} \mathbb{E}\|X_s\|^2_{L_2(I;Z;d\mu)} ds < \infty,$$

by Remark 1.38. Therefore, it follows that Y is integrable and $\mathbb{P}$-a.s.

$$\int_0^t Y(s)ds = \begin{pmatrix} \int_0^t X(s)ds \\ \int_0^t X_s ds \end{pmatrix},$$

where $\int_0^t X(s)ds \in \mathscr{D}(B)$ $\mathbb{P}$-a.s. and by Lemma 1.37 $\int_0^t X_s ds \in W^1_2(I;Z;d\mu)$ $\mathbb{P}$-a.s. with $(\int_0^t X_s ds)(0) = \int_0^t X(s)ds$. Hence, $\int_0^t Y(s)ds \in \mathscr{D}(\mathscr{A})$ $\mathbb{P}$-a.s. and again by Lemma 1.37 we find $\mathbb{P}$-a.s.

$$\begin{aligned}
\mathscr{A} \int_0^t Y(s)ds &= \begin{pmatrix} B \int_0^t X(s)ds + \Phi \int_0^t X_s ds \\ X_t - f \end{pmatrix} \\
&= \begin{pmatrix} X(t) - h - \int_0^t F(s, X(s), X_s)ds - \int_0^t G(s, X(s), X_s)dM(s) \\ X_t - f \end{pmatrix} \\
&= \begin{pmatrix} X(t) \\ X_t \end{pmatrix} - \begin{pmatrix} h \\ f \end{pmatrix} - \begin{pmatrix} \int_0^t F(s, X(s), X_s)ds \\ 0 \end{pmatrix} - \begin{pmatrix} \int_0^t G(s, X(s), X_s)dM(s) \\ 0 \end{pmatrix}.
\end{aligned}$$

Combining this equality with the following identities

$$\int_0^t \mathscr{F}(s, Y(s))ds = \int_0^t \begin{pmatrix} F(s, Y_1(s), Y_2(s)) \\ 0 \end{pmatrix} ds = \begin{pmatrix} \int_0^t F(s, Y_1(s), Y_2(s))ds \\ 0 \end{pmatrix}$$

and

$$\begin{aligned}
\int_0^t \mathcal{G}(s, Y(s))d\mathcal{M}(s) &= \int_0^t \begin{pmatrix} G(s, Y_1(s), Y_2(s)) & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} dM(s) \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} \int_0^t G(s, Y_1(s), Y_2(s))dM(s) \\ 0 \end{pmatrix}
\end{aligned}$$

we see that Y satisfies Definition 1.34.

(ii) On the other hand, let Y be an integrated solution of $(S_D - ACP)$ and define $X_0 := f$, $X(t) := Y_1(t)$ for $t \geqslant 0$. From Corollary 1.41 we find $Y_2(t)(\sigma) = Y_1(t+\sigma) = X_s(\sigma)$ for $t > -\sigma$. First, we show condition (i) and (ii) of Definition 1.36. This follows easily for infinite/finite delay, since ϱ is continuous and strictly positive. Hence, we can estimate

$$\infty > \sup_{t\in[0,T]} \mathbb{E}\|Y(t)\|_{\mathcal{E}_2}^2 \geqslant \sup_{t\in[0,T]} \mathbb{E}\|Y_1(t)\|_H^2 = \sup_{t\in[0,T]} \mathbb{E}\|X(t)\|_H^2,$$

$$\infty > \sup_{t\in[0,T]} \mathbb{E}\|Y(t)\|_{\mathcal{E}_2}^2 \geqslant \sup_{t\in[0,T]} \mathbb{E}\|Y_2(t)\|_{L_2(I;Z;d\mu)}^2 = \sup_{t\in[0,T]} \mathbb{E}\int_I \|Y_2(t)(\sigma)\|_Z^2 \varrho(\sigma)d\sigma$$

$$= \sup_{t\in[0,T]} \mathbb{E}\Bigg(\int\limits_{-\infty/\min\{0,t-1\}}^{0} \|Y_2(t)(s-t)\|_Z^2 \varrho(s-t)ds + \int\limits_{0/\max\{0,t-1\}}^{t} \|Y_2(t)(s-t)\|_Z^2 \varrho(s-t)ds\Bigg)$$

$$\geqslant \sup_{t\in[0,T]} \mathbb{E}\int_{0/\max\{0,t-1\}}^t \|X(s)\|_Z^2 \varrho(s-t)ds \geqslant \epsilon_\varrho \sup_{t\in[0,T]} \mathbb{E}\int_{0/\max\{0,t-1\}}^t \|X(s)\|_Z^2 ds.$$

This implies $\mathbb{E}\int_0^T \|X(s)\|_Z^2 ds < \infty$. Furthermore, it is clear that $\int_0^t X(s)ds \in \mathscr{D}(B)$ for all $t > 0$ $\mathbb{P}$-a.s., and we find $\mathbb{P}$-a.s. by writing down the equation of the first component of Y

$$X(t) - h = B\int_0^t X(s)ds + \Phi\int_0^t X_s ds + \int_0^t F(s, X(s), X_s)ds + \int_0^t G(t, X(s), X_s)dM(s).$$

□

1.2.3. Well-posedness for the semilinear stochastic delay problem

Just like in the deterministic case, Theorem 1.42 shows that also for the stochastic case the problem of well-posedness of $(S - DE)$ is reduced to the question if the stochastic abstract Cauchy problem $(S_D - ACP)$ is well-posed. Stochastic abstract Cauchy problems with Lévy noise have been studied extensively in the literature. Therefore, we are able to apply Theorem B.53, which is a classical existence and uniqueness result for stochastic Cauchy problems.

First, note that since the nonlinearities F, G and $\mathscr{F}$, $\mathcal{G}$ respectively, fulfill (SH_{10}) and (SH_{11}), which are classical linear growth and Lipschitz conditions, they satisfy, by Remark B.51, conditions (F) and (G) from Theorem B.53. Actually, condition (GI) holds, which is stronger than (G). This means the existence of a mild solution to $(S_D - ACP)$ and hence a solution of $(S - DE)$ is guaranteed, if $\mathscr{A}$ generates a C_0-semigroup. But we have already given criteria for this in Section 1.1.3. That is why we can formulate the following Theorem.

THEOREM 1.43 (well-posedness of $(S_D - ACP)$ and $(S - DE)$). *Under the standing hypotheses $(SH_1) - (SH_{11})$ problem $(S_D - ACP)$ and hence (S-DE) is well-posed if the operator $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ generates a C_0-semigroup. In particular, this is the case if the assumptions of Theorem 1.22, Theorem 1.23, Theorem 1.24, Theorem 1.27, or Theorem 1.28 are fulfilled.*

REMARK 1.44. *Theorem 1.43 shows that we can extend the deterministic delay models by stochastic terms and remain well-posedness. However, note that the solution from Theorem 1.43 may not have a càdlàg modification.*

1.3. Notation: $(S_D - ACP)$ and $(S - ACP)$

Since we were able to transform the stochastic abstract delay equation $(S - DE)$ into the abstract stochastic Cauchy problem $(S_D - ACP)$ in the sense of Theorem 1.42, we can reduce the study of stochastic delay equations to the study of stochastic abstract Cauchy problems. This is an advantage, since for abstract Cauchy problem there is a rich deterministic and stochastic theory available. In the stochastic case the theory for the Winer noise case is far more advanced than for the Lévy noise case. Despite that our general approach to stochastic delay equation will be via the stochastic abstract Cauchy problem $(S_D - ACP)$. This means in particular that our procedure from here on will be to address all further questions in the setting of a general stochastic abstract Cauchy problem and then apply those results via Theorem 1.42 to stochastic delay equation.

In order to be able to separate the general theory for stochastic abstract Cauchy problems from stochastic delay equations we introduce the following notation.

- If we are using regular math notion, that is A, T, F, G, and M, we are dealing with a general stochastic abstract Cauchy problem in the Hilbert space H which we call $(S - ACP)$. It reads as follows

$$(S - ACP)\begin{cases} dY(t) &= AY(t)dt + F(t, Y(t))dt + G(t, Y(t))dM(t), \quad t \geqslant 0, \\ Y(0) &= y_0 \in H, \end{cases}$$

  where A generates a C_0-semigroup. However, we hold on to our standing hypotheses for M and the nonlinearities F and G, with the obvious adjustments, that is $\mathcal{E}_2$ is replaced by H in (SH_{10}) and (SH_{11}). Then $(S - ACP)$ is well-posed by Theorem B.53.

- If we use math script respectively math calligraphic notation, that is $\mathscr{A}$, $\mathscr{T}$, $\mathscr{F}$, $\mathcal{G}$, and $\mathcal{M}$, we investigate the delay problem $(S - DE)$ via its transformed version $(S_D - ACP)$ in $\mathscr{E}_2$. That is

$$(S_D - ACP)\begin{cases} dY(t) &= \mathscr{A}Y(t)dt + \mathscr{F}(t, Y(t))dt + \mathcal{G}(t, Y(t))d\mathcal{M}(t), \quad t \geqslant 0, \\ Y(0) &= \begin{pmatrix} h \\ f \end{pmatrix}, \end{cases}$$

  where, for example, the operator $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ has the special form

$$\mathscr{A} := \begin{pmatrix} B & \Phi \\ 0 & \frac{d}{d\sigma} \end{pmatrix},$$

  with domain

$$\mathscr{D}(\mathscr{A}) := \left\{ \begin{pmatrix} h \\ f \end{pmatrix} \in \mathscr{D}(B) \times W_2^1(I; Z; d\mu) : f(0) = h \right\}.$$

CHAPTER 2

Approximation of the Mild Solution of $(S - ACP)$

In the previous chapter, we have seen in Theorem 1.35 that the mild solution Y of $(S-ACP)$ has no stochastic differentials and from Remark 1.33 we know that it might not have a càdlàg modification. This means the classical assumption for the transformation formula (that is Itô's formula) are not fulfilled for the process Y. Since the transformation formula is an essential tool in stochastic analysis, our goal is to find a generalized version, which we can still apply to the mild solution Y. In order to achieve this, we introduce two approximations of Y in this chapter, which are both based on the Yosida approximation. We will see that each member of the approximating sequences has stochastic differentials. To insure that each has a càdlàg modification we look at two different situations. First, we assume that the driving linear operator generates a generalized contraction semigroup. Then, Theorem B.53 guarantees that Y itself has a càdlàg modification and the same holds true for the approximating sequence. In the second case, we consider a general C_0-semigroup, where we approximate the driving linear operator A. Then, each member of the approximating sequence has the desired property, but the solution Y itself might only be predictable. After we proved the first approximation theorem, we apply those results to our case of interest, which is the delay problem and provide criteria, when the approximation can be applied. At the end of the chapter, we prove the general approximation theorem.

2.1. Yosida approximation, the generalized contraction case

We want to use the well-known Yosida approximation to construct a sequence of càdlàg processes, which converges to the mild solution Y of $(S-ACP)$ in the space $\mathcal{X}_{T,H}$ of all predictable processes with values in H. For the definition of $\mathcal{X}_{T,H}$ see Appendix B.3. We distinguish between two situations. The first one is the case, where the operator A generates a generalized contractions C_0-semigroup $(T(t))_{t\geqslant 0}$, that is $\|T(t)\|_{L(H)} \leqslant e^{\lambda t}$, for some $\lambda \in \mathbb{R}$. In the second case, we consider a general C_0-semigroup $(T(t))_{t\geqslant 0}$, that is $\|T(t)\|_{L(H)} \leqslant Me^{\lambda t}$, where $M \geqslant 1$ and $\lambda \in \mathbb{R}$.

In this section, we deal with the first case and therefore assume that A generates a C_0-semigroup of generalized contractions $(T(t))_{t\geqslant 0}$ with $\|T(t)\|_{L(H)} \leqslant e^{\lambda t}$ for some $\lambda \in \mathbb{R}$. In this case, the idea is to smooth all the nonlinear terms and the initial condition, such that they are elements of the domain $\mathscr{D}(A)$ of the driving linear operator A. This procedure guarantees that the approximating processes have stochastic differentials. In order to see this, let Y be the mild solution of

$$(S-ACP)\begin{cases} dY(t) &= AY(t)dt + F(t,Y(t))dt + G(t,Y(t))dM(t), \quad t \geqslant 0, \\ Y(0) &= y_0 \in H, \end{cases}$$

where M, F, and G fulfill the assumption from Section 1.3. Then, we define the approximating sequence $\{Y_n\}_{n\in\mathbb{N}}$ for $n > \lambda$ as the mild solution of

$$\begin{cases} dY_n(t) &= AY_n(t)dt + R(n)F(t,Y_n(t))dt + R(n)G(t,Y_n(t))dM(t), \quad t \geqslant 0, \\ Y_n(0) &= R(n)y_0, \end{cases}$$

where $R(n) = n(n-A)^{-1}$ for all $n \in \mathbb{N}$ with $n > \lambda$ is the Yosida approximation. Note, that we always write the short form $(n-A)^{-1}$ instead of $(nI-A)^{-1}$. By the properties of the Yosida approximation from Corollary A.21 (iii) it follows that $R(n)$ is a bounded linear operator for each $n > \lambda$. Therefore, we see that each Y_n is well-defined by Theorem B.53. Furthermore, we find from Corollary A.21 (i) and (iv) that for each $n > \lambda$ the operator $R(n)$ commutes with $T(t)$ for all $t \in [0,T]$ and that $R(n)h \in \mathscr{D}(A)$ for all $h \in H$. Thus, we see by writing down the mild solution

$$\begin{aligned} Y_n(t) &= T(t)R(n)y_0 + \int_0^t T(t-s)R(n)F(s,Y_n(s))ds + \int_0^t T(t-s)R(n)G(s,Y_n(s))dM(s) \\ &= R(n)T(t)y_0 + R(n)\int_0^t T(t-s)F(s,Y_n(s))ds + R(n)\int_0^t T(t-s)G(s,Y_n(s))dM(s), \end{aligned}$$

that $Y_n(t) \in \mathscr{D}(A)$ $\mathbb{P}$-a.s. for all $t \in [0,T]$. Considering the integrated solution we find that each Y_n has stochastic differentials, since $\mathbb{P}$-a.s.

$$\begin{aligned} Y_n(t) &= R(n)y_0 + A\int_0^t Y_n(s)ds + \int_0^t R(n)F(s,Y_n(s))ds + \int_0^t R(n)G(s,Y_n(s))dM(s) \\ &= R(n)y_0 + \int_0^t AY_n(s)ds + \int_0^t R(n)F(s,Y_n(s))ds + \int_0^t R(n)G(s,Y_n(s))dM(s). \end{aligned}$$

Finally, Theorem B.53 guarantees that Y as well as Y_n for all $n > \lambda$ have a càdlàg modification, since A is the generator of a generalized contraction C_0-semigroup. Therefore, the sequence $\{Y_n\}_{n>\lambda}$ has all the desired properties.

Remark 2.1. *It might look odd, that the solution process Y should have a càdlàg modification, since we defined the mild solution to be a predictable process. In this case, the process of left limits $\tilde{Y}$, defined by $\tilde{Y}(t) := Y(t-)$ for $t \in [0,T]$, is predictable. However, the processes Y and $\tilde{Y}$ are stochastically equivalent (see* [**PZ07**, Proposition 9.10]*). Since by Theorem B.53 the mild solution is only unique up to modification the processes Y and $\tilde{Y}$ are the same process in our setting.*

Before we show that $\{Y_n\}_{n\in\mathbb{N}}$ converges to Y in the space $\mathcal{X}_{T,H}$ we prove an auxiliary result. We extract this statement from the proof of the main theorem of this section to emphasize the technique we need to apply due to the stochastic integral. It is going to be the key step in the proof of the approximation theorem later and we will use it again in the proof of the transformation formula for mild solutions in Chapter 3. Therefore, recall that the norm in the Banach space $\mathcal{X}_{T,H}$ is given by

$$\|X\|_{T,H}^2 = \sup_{t\in[0,T]} \mathbb{E}\|X(t)\|_H^2.$$

PROPOSITION 2.2. *Let $T(t) \in L(H)$ for each $t \in [0,T]$, with $\|T(t)\|_{L(H)} \leqslant c(t)$, where c is a positive, continuous function. Then,*

$$\left\| \int_0^t T(t-s)(I-R(n))G(s,Y(s))dM(s) \right\|_{T,H} \longrightarrow 0 \text{ for } n\to\infty,$$

where Y is the mild solution of $(S-ACP)$ and $R(n)$ is the Yosida approximation.

PROOF.

First, note that the Yosida approximation doesn't converge in the operator norm. This means a straight forward argument like applying the isometry of the stochastic integral and estimates using the operator norm won't work. That is why our argument has to be more complex. We start by applying the isometry of the stochastic integral and estimate

$$\begin{aligned}
&\mathbb{E}\left\| \int_0^t T(t-s)(I-R(n))G(s,Y(s))dM(s) \right\|_H^2 \\
&= \mathbb{E}\int_0^t \|T(t-s)(I-R(n))G(s,Y(s))Q^{1/2}\|_{L_{(HS)}(U,H)}^2 ds \\
&\leqslant \mathbb{E}\int_0^t \|T(t-s)\|_{L(H)}^2 \|(I-R(n))G(s,Y(s))Q^{1/2}\|_{L_{(HS)}(U,H)}^2 ds \\
&\leqslant \max_{\tau\in[0,T]} c^2(\tau)\mathbb{E}\int_0^T \|(I-R(n))G(s,Y(s))Q^{1/2}\|_{L_{(HS)}(U,H)}^2 ds \\
&= \max_{\tau\in[0,T]} c^2(\tau)\mathbb{E}\left\| \int_0^T (I-R(n))G(s,Y(s))dM(s) \right\|_H^2 \qquad (2.1)
\end{aligned}$$

From Corollary A.21 (ii) we know that for every $h \in H$

$$\|(I-R(n))h\|_H \longrightarrow 0 \text{ for } n\to\infty.$$

Since for fixed $\omega \in \Omega$ we have $\int_0^T G(\omega,s,Y(s))dM(s)(\omega) \in H$, we find that the following convergence

$$\left\| \int_0^T (I-R(n))G(s,Y(s))dM(s) \right\|_H^2 = \left\| (I-R(n))\int_0^T G(s,Y(s))dM(s) \right\|_H^2 \longrightarrow 0 \qquad (2.2)$$

holds $\mathbb{P}$-a.s. for $n\to\infty$. From Corollary A.21 (iii) we know that $\|R(n)\|_{L(H)} \leqslant C_R$ for all $n > \lambda$, thus we can estimate using the isometry of the stochastic integral

$$\begin{aligned}
&\mathbb{E}\left\| \int_0^T (I-R(n))G(s,Y(s))dM(s) \right\|_H^2 = \mathbb{E}\left\| (I-R(n))\int_0^T G(s,Y(s))dM(s) \right\|_H^2 \\
&\leqslant (1+C_R)^2\mathbb{E}\left\| \int_0^t G(s,Y(s))dM(s) \right\|_H^2 = (1+C_R)^2\int_0^t \mathbb{E}\|G(s,Y(s))Q^{1/2}\|_{L_{(HS)}(U,H)}^2 ds \\
&\leqslant (1+C_R)^2 k_G^2 \int_0^t \mathbb{E}(1+\|Y(s)\|_H)^2 ds \leqslant 2(1+C_R)^2 k_G^2 \int_0^T 1+\mathbb{E}\|Y(s)\|_H^2 ds < \infty,
\end{aligned}$$

where we applied the linear growth condition of G and the Fubini–Tonelli theorem. The expression on the right-hand side is finite, since Y is the mild solution of $(S-ACP)$. Applying Lebesgue's dominated convergence theorem yields for $t \in [0,T]$

$$\mathbb{E}\left\| \int_0^t T(t-s)(I-R(n))G(s,Y(s))dM(s) \right\|_H^2 \longrightarrow 0 \text{ for } n \to \infty.$$

From estimate (2.1) it follows that this convergence is uniform for $t \in [0,T]$, hence

$$\sup_{t\in[0,T]} \mathbb{E}\left\| \int_0^t T(t-s)(I-R(n))G(s,Y(s))dM(s) \right\|_H^2 \longrightarrow 0 \text{ for } n \to \infty.$$

□

With Proposition 2.2 in place we can proceed to the approximation theorem.

THEOREM 2.3 (approximation of mild solution of (S-ACP); generalized contraction case). *Assume that A generates a C_0-semigroup of generalized contractions and let Y be the mild solution of $(S-ACP)$. Then*

$$\|Y - Y_n\|_{T,H} = (\sup_{t\in[0,T]} \mathbb{E}\|Y(t) - Y_n(t)\|_H^2)^{1/2} \longrightarrow 0 \textit{ for } n \to \infty.$$

PROOF.

We start by splitting the difference of the mild solutions in its three parts, that is

$$\begin{aligned} Y(t) - Y_n(t) = T(t)(I-R(n))y_0 &+ \int_0^t T(t-s)(F(s,Y(s)) - R(n)F(s,Y_n(s)))ds \\ &+ \int_0^t T(t-s)(G(s,Y(s)) - R(n)G(s,Y_n(s)))dM(s) \\ =: I_1(t) + I_2(t) + I_3(t). \end{aligned}$$

A straight forward estimate yields

$$\sup_{t\in[0,T]} \mathbb{E}\|Y(t) - Y_n(t)\|_H^2 \leqslant 3(\sup_{t\in[0,T]} \mathbb{E}\|I_1(t)\|_H^2 + \sup_{t\in[0,T]} \mathbb{E}\|I_2(t)\|_H^2 + \sup_{t\in[0,T]} \mathbb{E}\|I_3(t)\|_H^2). \quad (2.3)$$

We are going to look at each summand of (2.3) individually. We start with I_1 and estimate

$$\|T(t)(I-R(n))y_0\|_H^2 \leqslant e^{2\lambda T}\|(I-R(n))y_0\|_H^2.$$

Since $y_0 \in H$ we apply Corollary A.21 (ii) and find $(I-R(n))y_0 \longrightarrow 0$ for $n \to \infty$ in H. Summing up, we have shown

$$\sup_{t\in[0,T]} \mathbb{E}\|I_1(t)\|_H^2 \leqslant e^{2\lambda T}\|(I-R(n))y_0\|_H^2 \longrightarrow 0 \text{ for } n \to \infty.$$

In order to deal with the second summand I_2 in (2.3) we start by rewriting it as follows

$$F(s,Y(s)) - R(n)F(s,Y_n(s)) = (I-R(n))F(s,Y(s)) + R(n)(F(s,Y(s)) - F(s,Y_n(s))).$$

Since $\|R(n)\| \leqslant C_R$ by Corollary A.21 (iii), we estimate $\mathbb{P}$-a.s. for all $t \in [0,T]$ and $0 \leqslant s \leqslant t$

$$\begin{aligned}
&\|T(t-s)(F(s,Y(s)) - R(n)F(s,Y_n(s)))\|_H^2 \\
&\quad \leqslant e^{2\lambda T}\|(I-R(n))F(s,Y(s)) + R(n)(F(s,Y(s)) - F(s,Y_n(s)))\|_H^2 \\
&\quad \leqslant 2e^{2\lambda T}(\|(I-R(n))F(s,Y(s))\|_H^2 + C_R^2\|F(s,Y(s)) - F(s,Y_n(s))\|_H^2).
\end{aligned}$$

By Hölder's inequality, we find

$$\begin{aligned}
\sup_{t\in[0,T]} \mathbb{E}\|I_2(t)\|_H^2 &\leqslant T \sup_{t\in[0,T]} \mathbb{E}\int_0^t \|T(t-s)(F(s,Y(s)) - R(n)F(s,Y_n(s)))\|_H^2 ds \\
&\leqslant 2Te^{2\lambda T}\big(\sup_{t\in[0,T]} \mathbb{E}\int_0^t \|(I-R(n))F(s,Y(s))\|_H^2 ds \\
&\qquad + C_R^2 \sup_{t\in[0,T]} \mathbb{E}\int_0^t \|F(s,Y(s)) - F(s,Y_n(s))\|_H^2 ds\big) \\
&=: 2Te^{2\lambda T}(I_{21}(t) + I_{22}(t)).
\end{aligned}$$

For the term I_{21}, we have for every fixed $\omega \in \Omega$ and $s \in [0,T]$ that $F(\omega,s,Y(s)) \in H$ and hence by Corollary A.21 (ii) the pointwise convergence $\|(I-R(n))F(\omega,s,Y(\omega,s))\|_H^2 \longrightarrow 0$ for $n \to \infty$. Furthermore we estimate $\mathbb{P}$-a.s. and for every $s \in [0,T]$ using the linear growth condition of F and Corollary A.21 (iii)

$$\begin{aligned}
\|(I-R(n))F(s,Y(s))\|_H^2 &\leqslant (1+C_R)^2\|F(s,Y(s))\|_H^2 \leqslant (1+C_R)^2 k_F^2(1+\|Y(s)\|_H)^2 \\
&\leqslant 2(1+C_R)^2 k_F^2(1+\|Y(s)\|_H^2).
\end{aligned}$$

Since Y is the mild solution of $(S-ACP)$ it follows that $\int_0^t \mathbb{E}\|Y(s)\|_H^2 < \infty$ for all $t \in [0,T]$. Hence, by Fubini–Tonelli theorem and Lebesgue's dominated convergence theorem we receive

$$\begin{aligned}
\int_0^t \mathbb{E}\|(I-R(n))F(s,Y(s))\|_H^2 ds &\leqslant \int_0^T \mathbb{E}\|(I-R(n))F(s,Y(s))\|_H^2 ds \\
&= \mathbb{E}\int_0^T \|(I-R(n))F(s,Y(s))\|_H^2 ds \longrightarrow 0 \text{ for } n \to \infty.
\end{aligned}$$

Obviously this convergence is uniform for $t \in [0,T]$, that is

$$I_{21}(t) = \sup_{t\in[0,T]} \mathbb{E}\int_0^t \|(I-R(n))F(s,Y(s))\|_H^2 ds \longrightarrow 0 \text{ for } n \to \infty.$$

For the term I_{22} we apply the Lipschitz continuity of F and find

$$I_{22}(t) \leqslant C_R^2 C_F^2 \int_0^T \mathbb{E}\|Y(s) - Y_n(s)\|_H^2 ds \leqslant C_R^2 C_F^2 \int_0^T \sup_{s\in[0,\tau]} \mathbb{E}\|Y(s) - Y_n(s)\|_H^2 d\tau.$$

For I_3 in (2.3) we start again by rewriting the relevant expression as follows

$$G(s,Y(s)) - R(n)G(s,Y_n(s)) = (I-R(n))G(s,Y(s)) + R(n)(G(s,Y(s)) - G(s,Y_n(s))).$$

Thus, we estimate

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\|I_3(t)\|_H^2 &\leqslant 2 \sup_{t\in[0,T]} \mathbb{E}\left\| \int_0^t T(t-s)(I-R(n))G(s,Y(s))dM(s)\right\|_H^2 \\ &\quad + 2 \sup_{t\in[0,T]} \mathbb{E}\left\| \int_0^t T(t-s)R(n)(G(s,Y(s))-G(s,Y_n(s)))dM(s)\right\|_H^2 \\ &=: 2(I_{31}(t)+I_{32}(t)).\end{aligned}$$

The term I_{31} converges to zero for $n\to\infty$ by Proposition 2.2. For the term I_{32} we apply the isometry, the Fubini–Tonelli theorem, Corollary A.21 (iii), and the Lipschitz continuity of G

$$\begin{aligned}I_{32}(t) &= \sup_{t\in[0,T]} \int_0^t \mathbb{E}\|T(t-s)R(n)(G(s,Y(s))-G(s,Y_n(s))\|^2_{L_{(HS)}(U,H)}ds \\ &\leqslant e^{2\lambda T}C_R^2 \int_0^T \mathbb{E}\|G(s,Y(s))-G(s,Y_n(s))\|^2_{L_{(HS)}(U,H)}ds \\ &\leqslant e^{2\lambda T}C_R^2C_G^2 \int_0^T \mathbb{E}\|Y(s)-Y_n(s)\|_H^2 ds \leqslant e^{2\lambda T}C_R^2C_G^2 \int_0^T \sup_{s\in[0,\tau]} \mathbb{E}\|Y(s)-Y_n(s)\|_H^2 d\tau.\end{aligned}$$

Now, we collect all terms and from (2.3) follows

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\|Y(t)-Y_n(t)\|_H^2 &\leqslant 3\tilde\varepsilon(n) + 6Te^{2\lambda T}(\tilde\varepsilon(n) + C_R^2C_F^2 \int_0^T \sup_{s\in[0,\tau]} \mathbb{E}\|Y(s)-Y_n(s)\|_H^2 d\tau) \\ &\quad + 6(\tilde\varepsilon(n) + e^{2\lambda T}C_R^2C_G^2 \int_0^T \sup_{s\in[0,\tau]} \mathbb{E}\|Y(s)-Y_n(s)\|_H^2 d\tau) \\ &\leqslant \varepsilon(n) + C(T,\lambda,C_R,C_F,C_G) \int_0^T \sup_{s\in[0,\tau]} \mathbb{E}\|Y(s)-Y_n(s)\|_H^2 d\tau,\end{aligned}$$

where $\{\tilde\varepsilon(n)\}_{n\in\mathbb{N}}$ is a suitable null sequence, $\varepsilon := (9+6Te^{2\lambda T})\tilde\varepsilon$, and C stands for a constant depending only on the values stated in the parentheses. Using Grönwall's inequality, we obtain

$$\sup_{t\in[0,T]} \mathbb{E}\|Y(t)-Y_n(t)\|_H^2 \leqslant \varepsilon(n)(1+C(T,\lambda,C_R,C_F,C_G)Te^{C(T,\lambda,C_R,C_F,C_G)T}).$$

Hence, for $n\to\infty$ we find

$$\|Y-Y_n\|_{T,H}^2 = \sup_{t\in[0,T]} \mathbb{E}\|Y(t)-Y_n(t)\|_H^2 \longrightarrow 0.$$

□

Remark 2.4. *Like we said before the most difficult convergence in the proof of Theorem 2.3 is the one of term I_{31}. That is why we treated it separately in Proposition 2.2. The crucial step in the proof of Proposition 2.2 is in line* (2.2)*. It shows the importance of the approximation being time independent, such that it can be pulled out of the stochastic integral. If we would approximate the operator A using Yosida approximation, we would end up with a sequence of uniformly continuous semigroups approximating the C_0-semigroup generated by A. However, those are time dependent and therefore our argument would no longer work. We just showed that in the case of a generalized contraction semigroup we are able to avoid the approximation* *of A. In the general case this is no longer possible and therefore a different idea is needed. We will treat this situation in Section 2.3.*

2.2. Dissipativity of the operator $\mathscr{A}$

The essential assumption for the approximation scheme from Theorem 2.3 to work is that the driving linear operator A generates a C_0-semigroup of generalized contractions. If that is the case, not only does the approximation holds true, but we know by Theorem B.53 that the solution Y itself has a càdlàg modification. We want to apply those results to stochastic delay equation in this section. Thus, we identify cases, when the operator $\mathscr{A}$ of $(S_D - ACP)$ generate a C_0-semigroup of generalized contractions. Therefore, we assume that $\mathscr{A}$ generates a C_0-semigroup $(\mathscr{T}(t))_{t\geqslant 0}$, with $\|\mathscr{T}(t)\|_{L(\mathcal{E}_2)} \leqslant Me^{\lambda t}$, where $M \geqslant 1$ and $\lambda \in \mathbb{R}$. To guarantee that $\mathscr{A}$ is a generator one can apply the results from Section 1.2.3.

The operators which generate a C_0-semigroup of generalized contractions are characterized by the Lumer-Phillips theorem (see Theorem A.16 and Remark A.17). Since $\mathscr{A}$ generates a C_0-semigroup, we know by the general Hille-Yosida theorem (see [**EN00**, Theorem 3.8 in Section II.3.a]) that $\operatorname{rg}(\lambda_0 - A) = \mathcal{E}_2$ for any $\lambda_0 > \lambda$. Hence, all what is left to do is to identify cases when the operator $\mathscr{A}$ is or is not γ-dissipative for some $\gamma \in \mathbb{R}$.

Recalling the definition of γ-dissipative operators (see Definition A.14), we must show that

$$\langle \mathscr{A}u, u\rangle_{\mathcal{E}_2} \leqslant \gamma\|u\|^2_{\mathcal{E}_2}$$

for all $u \in \mathscr{D}(\mathscr{A})$. Breaking down the expression on the left-hand side into its individual components yields for $u = \binom{u_1}{u_2}$

$$\begin{aligned}\langle \mathscr{A}u, u\rangle_{\mathcal{E}_2} &= \langle \begin{pmatrix} B & \Phi \\ 0 & \frac{d}{d\sigma}\end{pmatrix}\begin{pmatrix} u_1 \\ u_2\end{pmatrix}, \begin{pmatrix} u_1 \\ u_2\end{pmatrix}\rangle_{\mathcal{E}_2} = \langle \begin{pmatrix} Bu_1 + \Phi u_2 \\ \frac{d}{d\sigma}u_2\end{pmatrix}, \begin{pmatrix} u_1 \\ u_2\end{pmatrix}\rangle_{\mathcal{E}_2} \\ &= \langle Bu_1, u_1\rangle_H + \langle \Phi u_2, u_1\rangle_H + \langle \frac{d}{d\sigma}u_2, u_2\rangle_{L_2(I;Z;d\mu)}. \end{aligned} \tag{2.4}$$

immediately, we see from (2.4) that it is natural to assume B to be γ_0-dissipative for some $\gamma_0 \in \mathbb{R}$. We will always assume this until the end of this section. Thus (2.4) simplifies to

$$\langle \mathscr{A}u, u\rangle_{\mathcal{E}_2} \leqslant \gamma_0\|u_1\|^2_H + \langle \Phi u_2, u_1\rangle_H + \langle \frac{d}{d\sigma}u_2, u_2\rangle_{L_2(I;Z;d\mu)}. \tag{2.5}$$

Now, we will look at important examples for the delay operator Φ and check if $\mathscr{A}$ is γ-dissipative for some $\gamma \in \mathbb{R}$.

2.2.1. Single delay - bounded delay operator

We start with the simplest case. Therefore, we consider a single delay, that is

$$\Phi u := Cu(-1), \; u \in W^1_2([-1,0];H),$$

where $C \in L(H)$. Thus, we have a bounded delay operator. This implies for our setting that $Z = H$ and $\mathcal{E}_2 = H \times L_2([-1,0];H)$. Recall that in the case of a finite delay we set $\varrho \equiv 1$

without loss of generality and thus μ becomes the Lebesgue measure. Looking back at (2.5), we have to treat the last two terms on the right-hand side. We start with the last one and integrating by parts yields

$$\begin{aligned}\langle \frac{d}{d\sigma}u_2, u_2\rangle_{L_2([-1,0];H)} &= \int_{-1}^{0} \langle \frac{d}{d\sigma}u_2(\sigma), u_2(\sigma)\rangle_H d\sigma \\ &= -\int_{-1}^{0} \langle \frac{d}{d\sigma}u_2(\sigma), u_2(\sigma)\rangle_H d\sigma + \|u_2(0)\|_H^2 - \|u_2(-1)\|_H^2.\end{aligned}$$

Since $u \in \mathscr{D}(\mathscr{A})$, we have $u_2(0) = u_1$ and we conclude

$$\langle \frac{d}{d\sigma}u_2, u_2\rangle_{L_2([-1,0];H)} = \frac{1}{2}(\|u_1\|_H^2 - \|u_2(-1)\|_H^2). \tag{2.6}$$

For the term involving the delay operator we apply elementary estimates and receive

$$\begin{aligned}\langle \Phi u_2, u_1\rangle_H = \langle Cu_2(-1), u_1\rangle_H &= \langle u_2(-1), C^* u_1\rangle_H \leqslant \|u_2(-1)\|_H \|C^* u_1\|_H \\ &\leqslant \frac{1}{2}(\|u_2(-1)\|_H^2 + \|C\|_{L(H)}^2 \|u_1\|_H^2),\end{aligned} \tag{2.7}$$

where C^* is the adjoint operator of C. Summing up, we find

$$\langle \mathscr{A}u, u\rangle_{\mathcal{E}_2} \leqslant \gamma_0 \|u_1\|_H^2 + \frac{1}{2}(\|C\|_{L(H)}^2 + 1)\|u_1\|_H^2 \leqslant \frac{2\gamma_0 + \|C\|_{L(H)}^2 + 1}{2}\|u\|_{\mathcal{E}_2}^2.$$

Hence, we can formulate the following proposition.

PROPOSITION 2.5. *If B is a γ_0-dissipative operator and $\Phi u_2 = Cu_2(-1)$, where $C \in L(H)$, then $\mathscr{A}$ is $\frac{2\gamma_0 + \|C\|_{L(H)}^2 + 1}{2}$-dissipative operator.*

2.2.2. Multiple delay - bounded delay operator

The natural next step is to consider a delay operator Φ with multiple delays. It turns out that this simple change in the structure of the delay operator is already enough for the operator $\mathscr{A}$ to lose its dissipativity, like the following example demonstrates.

EXAMPLE 2.6. *We choose*

$$\Phi u = u(-1/2) + u(-1), \ u \in W_2^1([-1,0]; H),$$

to be the delay operator. Just like in the previous case, we see that Φ is a bounded operator and therefore we set $Z = H$ and $\mathcal{E}_2 = H \times L_2([-1,0], H)$. Now, we choose an $0 \neq a \in \mathscr{D}(B)$ and a sequence $u_n \in \mathscr{D}(\mathscr{A})$, such that

$$u_{n2}(0) = u_{n1} = a, \ u_{n2}(-1/2) = na, \ u_{n2}(-1) = 0,$$

and

$$\|u_{n2}\|_{L_2([-1,0],H)} = \frac{1}{n},$$

where we use the notation of Definition 1.17. By (2.4) and (2.6) we find

$$\langle \mathscr{A}u_n, u_n\rangle_{\mathcal{E}_2} = \langle Ba, a\rangle_H + n\|a\|_H^2 + \frac{1}{2}\|a\|_H^2 \to \infty \text{ for } n \to \infty.$$

On the other hand, we have

$$\|u_n\|_{\mathcal{E}_2}^2 = \|a\|_H^2 + \frac{1}{n^2} \to \|a\|_H^2 \text{ for } n \to \infty.$$

This shows, that there can't be a constant $\gamma \in \mathbb{R}$, such that

$$\langle \mathscr{A} u, u \rangle_{\mathcal{E}_2} \leqslant \gamma \|u\|_{\mathcal{E}_2}^2$$

for all $u \in \mathscr{D}(\mathscr{A})$.

REMARK 2.7. *In order to reinstate the dissipativity of the operator $\mathscr{A}$ for a delay operator with multiple delays it was suggested in* [**Kap85**] *to adjust the geometry of the history space $L_2([-1,0]; H)$, by introducing an equivalent norm in the following way. If the delay operator Φ depends on $p \in \mathbb{N}$ delays at times $h_j \in [-1,0)$, $j = 1, \ldots, p$, define the wight function*

$$g(s) := p - j + 1 \text{ for } s \in [h_j, h_{j+1})$$

and the new norm of $\mathcal{E}_2$ as

$${}_g\|u\|_{\mathcal{E}_2} := \sqrt{\|u_1\|_H^2 + \int_{-1}^{0} \|u(\tau)\|_H^2 g(\tau) d\tau}.$$

Then, we find that the operator $\mathscr{A}$ is dissipative with respect to the new norm.

2.2.3. Singel delay - unbounded delay operator

We return to the case of a single delay, but this time we consider an unbounded operator. Recall that for an unbounded delay operator Φ the operator B has to fulfill condition (1.9) and (1.10) such that $\mathscr{A}$ generates a C_0-semigroup (see Theorem 1.27). Thus, it is natural that for the dissipativity of $\mathscr{A}$ we need to impose stronger assumption on the operator B than γ_0-dissipativity. Typically B has to fulfill a generalized Gårding's inequality of the form

$$\langle Bu, u \rangle_H \leqslant -\gamma_1 \|u\|_Z^2 + \gamma_0 \|u\|_H^2, \tag{2.8}$$

where $u \in \mathscr{D}(B)$, $\gamma_1 > 0$ and $\gamma_0 \in \mathbb{R}$. We will demonstrate this in the next example.

EXAMPLE 2.8. *Consider the following delayed heat equation*

$$\begin{cases} \partial_t u(t,s) &= \Delta u(t,s) + \sum_{j=1}^n C_j \partial_j u(t-1,s), & t \geqslant 0,\ s \in G, \\ \frac{\partial u}{\partial \nu}(t,s) &= 0, & t \geqslant 0,\ s \in \partial G, \\ u(t,s) &= f(t,s), & t \in [-1,0],\ s \in G, \end{cases}$$

where $G \subset \mathbb{R}^n$ is a bounded domain with smooth boundary, $H := L_2(G)$, $Z := W_2^1(G)$, $Bu := \Delta u$, $\mathscr{D}(B) := \{g \in W_2^1(G) : \Delta g \in L_2(G) \text{ and } \frac{\delta g}{\delta \nu} = 0 \text{ on } \partial G\}$, $C_j \in L(H)$ for $j \in \{1, \ldots, n\}$, and for $u \in W_2^1([-1,0]; Z)$ we set $\Phi u := \sum_{j=1}^n C_j \frac{\partial}{\partial x_j} u(-1)$. Like in (2.7) we find for $u = \binom{u_1}{u_2} \in \mathscr{D}(\mathscr{A})$

$$\langle \Phi u_2, u_1 \rangle_H \leqslant \frac{1}{2} \sum_{j=1}^{n} (\|\frac{\partial}{\partial x_j} u_2(-1)\|_H^2 + \|C_j\|_{L(H)}^2 \|u_1\|_H^2).$$

Furthermore, we calculate using (2.6)

$$\begin{aligned}\langle \frac{d}{d\sigma} u_2, u_2 \rangle_{L_2([-1,0];Z)} &= \frac{1}{2}(\|u_1\|_Z^2 - \|u_2(-1)\|_Z^2) \\ &= \frac{1}{2}(\|u_1\|_H^2 + \sum_{i=1}^n \|\frac{\partial}{\partial x_i} u_1\|_H^2) - \frac{1}{2}(\|u_2(-1)\|_H^2 + \sum_{i=1}^n \|\frac{\partial}{\partial x_i} u_2(-1)\|_H^2).\end{aligned}$$

As for the operator B we have

$$\begin{aligned}\langle Bu_1, u_1\rangle_H &= \int_G \Delta u_1(x) u_1(x) dx = -\int_G \nabla u_1(x) \cdot \nabla u_1(x) dx + \int_{\partial G} \underbrace{\frac{\partial u_1}{\partial \nu}(y)}_{=0} u(y) dS(y) \\ &= -\sum_{i=1}^n \langle \frac{\partial}{\partial x_i} u_1, \frac{\partial}{\partial x_i} u_1\rangle_H = -\sum_{i=1}^n \|\frac{\partial}{\partial x_i} u_1\|_H^2.\end{aligned}$$

Thus, condition (2.8) *is fulfilled with* $\gamma_0 = \gamma_1 = 1$. *Inserting all terms into* (2.4) *yields*

$$\begin{aligned}\langle \mathscr{A} u, u\rangle_{\mathcal{E}_2} \leqslant &-\sum_{i=1}^n \|\frac{\partial}{\partial x_i} u_1\|_H^2 + \frac{1}{2}\sum_{j=1}^n (\|\frac{\partial}{\partial x_j} u_2(-1)\|_H^2 + \|C_j\|_{L(H)}^2 \|u_1\|_H^2) \\ &+ \frac{1}{2}(\|u_1\|_H^2 + \sum_{i=1}^n \|\frac{\partial}{\partial x_i} u_1\|_H^2) - \frac{1}{2}(\|u_2(-1)\|_H^2 + \sum_{i=1}^n \|\frac{\partial}{\partial x_i} u_2(-1)\|_H^2) \\ \leqslant &\frac{\sum_{j=1}^n \|C_j\|_{L(H)}^2 + 1}{2} \|u_1\|_H^2 \leqslant \frac{\sum_{j=1}^n \|C_j\|_{L(H)}^2 + 1}{2} \|u\|_{\mathcal{E}_2}^2.\end{aligned}$$

Therefore the operator $\mathscr{A}$ is $\frac{1}{2}(\sum_{j=1}^n \|C_j\|_{L(H)}^2 + 1)$*-dissipative.*

2.2.4. Averaging over a finite delay

If the delay operator Φ is some kind of average, the delay is often called distributed delay. In this case we can apply the following corollary.

COROLLARY 2.9. *If B is a γ_0-dissipative operator and the delay operator Φ is a bounded linear operator from $L_2([-1,0];H)$ to H, that is for $u \in L_2([-1,0];H)$*

$$\|\Phi u\|_H \leqslant C_\Phi \|u\|_{L_2([-1,0];H)},$$

for some constant $C_\Phi \geqslant 1$, then the operator $\mathscr{A}$ is a $\frac{1}{2}(2\gamma_0 + C_\Phi^2 + 1)$-dissipative operator.

PROOF.

We consider the dissipativity estimate (2.5) and for $u = \binom{u_1}{u_2} \in \mathscr{D}(\mathscr{A})$ we start the estimate with the delay term

$$\langle \Phi u_2, u_1\rangle_H \leqslant \frac{1}{2}(\|\Phi u_2\|_H^2 + \|u_1\|_H^2) \leqslant \frac{C_\Phi^2}{2}(\|u_1\|_H^2 + \|u_2\|_{L_2([-1,0];H)}^2) = \frac{C_\Phi^2}{2}\|u\|_{\mathcal{E}_2}^2.$$

Form (2.6) follows $\langle \frac{d}{d\sigma} u_2, u_2\rangle_{L_2([-1,0];H)} = \frac{1}{2}(\|u_1\|_H - \|u_2(-1)\|_H^2) \leqslant \frac{1}{2}\|u_1\|_H^2$. Collecting terms yields

$$\langle \mathscr{A} u, u\rangle_{\mathcal{E}_2} \leqslant \gamma_0 \|u_1\|_H^2 + \frac{C_\Phi^2}{2}\|u\|_{\mathcal{E}_2}^2 + \frac{1}{2}\|u_1\|_H^2 \leqslant \frac{2\gamma_0 + C_\Phi^2 + 1}{2}\|u\|_{\mathcal{E}_2}^2.$$

□

EXAMPLE 2.10. *Assume, the delay operator Φ is of the following form*

$$\Phi u = \int_{-1}^{0} h(\tau)u(\tau)d\tau,$$

for $u \in L_2([-1,0];H)$ and $h \in L_2([-1,0];L(H))$. Then, Φ is a bounded linear operator from $L_2([-1,0];H)$ to H. This is easy to see, since

$$\|\Phi u\|_H \leqslant \int_{-1}^{0} \|h(\tau)\|_{L(H)}\|u(\tau)\|_H d\tau \leqslant \|h\|_{L_2([-1,0];L(H))}\|u\|_{L_2([-1,0];H)}.$$

Thus, Corollary 2.9 can be applied.

2.2.5. Averaging over a infinite delay - constant weight function

If the delay operator Φ is an average over an infinite delay and the weight function ϱ is constant (without loss of generality assume $\varrho \equiv 1$), we can transfer the results of the finite delay case. This is due to the embedding $W_2^1(\mathbb{R}_-;H) \hookrightarrow C_0(\mathbb{R}_-;H)$, thus we have

$$\langle \frac{d}{d\sigma}u_2, u_2\rangle_{L_2(\mathbb{R}_-;H)} = \frac{1}{2}\ \|u_1\|_H^2,$$

for $u = \binom{u_1}{u_2} \in \mathscr{D}(\mathscr{A})$. Therefore, we rewrite Corollary 2.9 for an infinite delay as follows.

COROLLARY 2.11. *If B is a γ_0-dissipative operator and the delay operator Φ is a bounded linear operator from $L_2(\mathbb{R}_-;H)$ to H, that is for $u \in L_2(\mathbb{R}_-;H)$*

$$\|\Phi u\|_H \leqslant C_\Phi \|u\|_{L_2(\mathbb{R}_-;H)},$$

for some constant $C_\Phi \geqslant 1$, then the operator $\mathscr{A}$ is a $\frac{1}{2}(2\gamma_0 + C_\Phi^2 + 1)$-dissipative operator.

The examples can be transferred in just the same way. That is, if $h \in L_2(\mathbb{R}_-;L(H))$, then the operator defined by

$$\Phi u := \int_{-\infty}^{0} h(\tau)u(\tau)d\tau$$

is bounded from $L_2(\mathbb{R}_-;H)$ to H and Corollary 2.11 applies.

2.2.6. Averaging over an infinite delay - nonconstant weight function

Now we turn to the general setup for averaging over an infinite delay. Recall that the measure μ from Definition 1.5 is defined by the weight function ϱ, which fulfills $\varrho \in C^1(\mathbb{R}_-)$ with $\varrho > \epsilon_\varrho > 0$. If the weight function ϱ is not constant, we have to look at the last term of the dissipativity estimate (2.5) again, since this is there the measure μ effects our calculation. Hence, we compute

$$\langle \frac{d}{d\sigma}u_2, u_2\rangle_{L_2(\mathbb{R}_-,H,d\mu)} = \int_{-\infty}^{0} \langle \frac{d}{d\sigma}u_2(\sigma), u_2(\sigma)\rangle_H \varrho(\sigma)d\sigma = \int_{-\infty}^{0} \langle \frac{d}{d\sigma}u_2(\sigma), \varrho(\sigma)u_2(\sigma)\rangle_H d\sigma$$
$$= \big[\varrho(\sigma)\|u_2(\sigma)\|_H^2\big]_{\sigma=-\infty}^{\sigma=0} - \int_{-\infty}^{0} \langle u_2(\sigma), \frac{d}{d\sigma}u_2(\sigma)\rangle_H \varrho(\sigma)d\sigma - \int_{-\infty}^{0} \langle u_2(\sigma), u_2(\sigma)\rangle_H \dot{\varrho}(\sigma)d\sigma.$$

Thus, we have

$$\langle \frac{d}{d\sigma} u_2, u_2 \rangle_{L_2(\mathbb{R}_-,H,d\mu)} = \frac{1}{2}\varrho(0)\|u_1\|_H^2 - \frac{1}{2}\int_{-\infty}^0 \|u_2(\sigma)\|_H^2 \dot{\varrho}(\sigma) d\sigma, \tag{2.9}$$

since

$$\lim_{\sigma \to -\infty} \varrho(\sigma)\|u_2(\sigma)\|_H^2 = 0.$$

This is easy to see, if we write $u_2 = \psi v$, where ψ is a real function with $|\psi| \leqslant \varrho^{-1/2}$ and $|\dot{\psi}| \leqslant \varrho^{-1/2}$. Then, by writing down the norm for u_2, we immediately find $v \in W_2^1(\mathbb{R}_-; H)$. Thus, the limit follows from the embedding $W_2^1(\mathbb{R}_-; H) \hookrightarrow C_0(\mathbb{R}_-; H)$.

From (2.9) we see that the weight function ϱ effects the dissipativity of the operator $\mathscr{A}$. Therefore, we generalize Corollary 2.11 by formulating criteria for ϱ, such that the operator $\mathscr{A}$ is dissipative.

COROLLARY 2.12. *Assume B to be a γ_0-dissipative operator and the delay operator Φ to be a bounded linear operator from $L_2(\mathbb{R}_-; H; d\mu)$ to H, that is for $u \in L_2(\mathbb{R}_-; H; d\mu)$*

$$\|\Phi u\|_H \leqslant C_\Phi \|u\|_{L_2(\mathbb{R}_-;H;d\mu)},$$

for some constant $C_\Phi \geqslant 1$. Furthermore, let one of the following condition be fulfilled

$$a) \ \dot{\varrho} \geqslant -K\varrho,$$
$$b) \ \dot{\varrho} \geqslant -K,$$

for some $K \geqslant 0$. Then the operator $\mathscr{A}$ is a $\frac{1}{2}(2\gamma_0 + \varrho(0) + \tilde{K} + C_\Phi^2)$-dissipative operator, where $\tilde{K}$ is equal to K in case $a)$ and K/ϵ_ϱ in case $b)$.

PROOF.

Consider the dissipativity estimate (2.5) and for $u = \binom{u_1}{u_2} \in \mathscr{D}(\mathscr{A})$ we start the estimate with the last term. For the case $a)$ we find from (2.9)

$$\begin{aligned}\langle \frac{d}{d\sigma} u_2, u_2 \rangle_{L_2(\mathbb{R}_-,H,d\mu)} &\leqslant \frac{1}{2}\varrho(0)\|u_1\|_H^2 + \frac{K}{2}\int_{-\infty}^0 \|u_2(\sigma)\|_H^2 \varrho(\sigma) d\sigma \\ &= \frac{1}{2}\varrho(0)\|u_1\|_H^2 + \frac{K}{2}\|u_2\|_{L_2(\mathbb{R}_-,H,d\mu)}^2.\end{aligned}$$

For the case $b)$ we argue in the following way

$$\begin{aligned}\langle \frac{d}{d\sigma} u_2, u_2 \rangle_{L_2(\mathbb{R}_-,H,d\mu)} &\leqslant \frac{1}{2}\varrho(0)\|u_1\|_H^2 + \frac{K}{2}\int_{-\infty}^0 \|u_2(\sigma)\|_H^2 d\sigma \\ &\leqslant \frac{1}{2}\varrho(0)\|u_1\|_H^2 + \frac{K}{2}\int_{-\infty}^0 \|u_2(\sigma)\|_H^2 \frac{\varrho(\sigma)}{\varrho(\sigma)} d\sigma \\ &\leqslant \frac{1}{2}\varrho(0)\|u_1\|_H^2 + \frac{K}{2\epsilon_\varrho}\int_{-\infty}^0 \|u_2(\sigma)\|_H^2 \varrho(\sigma) d\sigma \\ &= \frac{1}{2}\varrho(0)\|u_1\|_H^2 + \frac{K}{2\epsilon_\varrho}\|u_2\|_{L_2(\mathbb{R}_-,H,d\mu)}^2.\end{aligned}$$

For the delay term we proceed similarly to the previous cases

$$\langle \Phi u_2, u_1\rangle_H \leqslant \frac{1}{2}(\|u_1\|_H^2 + \|\Phi u_2\|_H^2) \leqslant \frac{C_\Phi^2}{2}(\|u_1\|_H^2 + \|u_2\|^2_{L_2(\mathbb{R}_-,H,d\mu)}) = \frac{C_\Phi^2}{2}\|u\|^2_{\mathcal{E}_2}.$$

Collecting terms yields

$$\begin{aligned}\langle \mathscr{A}u, u\rangle_{\mathcal{E}_2} &\leqslant \gamma_0\|u_1\|_H^2 + \frac{C_\Phi^2}{2}\|u\|^2_{\mathcal{E}_2} + \frac{1}{2}\varrho(0)\|u_1\|_H^2 + \frac{\tilde{K}}{2}\|u_2\|^2_{L_2(\mathbb{R}_-,H,d\mu)}\\ &\leqslant \frac{2\gamma_0 + \varrho(0) + \tilde{K} + C_\Phi^2}{2}\|u\|^2_{\mathcal{E}_2},\end{aligned}$$

where $\tilde{K}$ is equal to K in case $a)$ and K/ϵ_ϱ in case $b)$.

□

2.3. Yosida approximation, the general case

Now, we consider the general situation, where A generates a C_0-semigroup $(T(t))_{t\geqslant 0}$ with $\|T(t)\|_{L(H)} \leqslant Me^{\lambda t}$, where $M \geqslant 1$ and $\lambda \in \mathbb{R}$. In this case we have to consider a fundamentally different approximation scheme compared to Section 2.1. There, the basic idea was to smooth all terms into the domain $\mathscr{D}(A)$ of the driving linear operator A. In the general case, we do the opposite and bring all terms of the equation down to H. This is due to the mild solution Y not necessarily having a càdlàg modification. Thus, we have to approximate the operator A using Yosida approximation. Therefore, it becomes unnecessary to smooth out all other terms of the equation. Recall that Y is the mild solution of

$$(S-ACP)\begin{cases} dY(t) &= AY(t)dt + F(t,Y(t))dt + G(t,Y(t))dM(t), \quad t\geqslant 0,\\ Y(0) &= y_0 \in H,\end{cases}$$

where M, F, and G fulfill the assumption from Section 1.3, then we define the approximating sequence $\{Z_n\}_{n\in\mathbb{N}}$ for $n > \lambda$ as the mild solution of

$$\begin{cases} dZ_n(t) &= A_nZ_n(t)dt + F(t,Z_n(t))dt + G(t,Z_n(t))dM(t), \quad t\geqslant 0,\\ Z_n(0) &= y_0,\end{cases}$$

for $n > \lambda$, where A_n is the Yosida approximation of the operator A. Note that by the properties of the Yosida approximation from Corollary A.24 it follows that A_n is a bounded linear operator for each $n > \lambda$. Furthermore, each A_n generates the uniformly continuous semigroup $T_n(t) := e^{tA_n}$. Hence, by Theorem B.53 Z_n is well-defined for each $n > \lambda$.Writing down the integrated solution we immediately see that each Z_n has stochastic differentials, since $\mathbb{P}$-a.s.

$$\begin{aligned}Z_n(t) &= y_0 + A_n\int_0^t Z_n(s)ds + \int_0^t F(s,Z_n(s))ds + \int_0^t G(s,Z_n(s))dM(s)\\ &= y_0 + \int_0^t A_nZ_n(s)ds + \int_0^t F(s,Z_n(s))ds + \int_0^t G(s,Z_n(s))dM(s).\end{aligned}$$

Remark A.25 states that A_n is trivially a generator of a generalized contraction C_0-semigroup for $n > \lambda$. Therefore, Theorem B.53 guarantees that each Z_n has as a càdlàg modification. Thus, the sequence $\{Z_n\}_{n>\lambda}$ has all the desired properties. Note that in the proof of the following theorem we will write $Z(t)$ again instead of $Z(t-)$ or $\tilde{Z}(t)$ for the same reasons as explained in Remark 2.1.

THEOREM 2.13 (approximation of mild solution of (S-ACP); general case). *Assume that A generates a C_0-semigroup and let Y be the mild solution of $(S-ACP)$. Then,*

$$\|Y - Z_n\|_{T,H} = (\sup_{t\in[0,T]} \mathbb{E}\|Y(t) - Z_n(t)\|_H^2)^{1/2} \longrightarrow 0 \text{ for } n \to \infty.$$

PROOF.

We start by splitting the difference of the mild solutions in its three parts, that is

$$\begin{aligned} Y(t) - Z_n(t) &= (T(t) - T_n(t))y_0 + \int_0^t T(t-s)F(s,Y(s)) - T_n(t-s)F(s,Z_n(s))ds \\ &\qquad + \int_0^t T(t-s)G(s,Y(s)) - T_n(t-s)G(s,Z_n(s))dM(s) \\ &=: J_1(t) + J_2(t) + J_3(t). \end{aligned}$$

A straight forward estimate yields

$$\sup_{t\in[0,T]} \mathbb{E}\|Y(t) - Z_n(t)\|_H^2 \leqslant 3(\sup_{t\in[0,T]} \mathbb{E}\|J_1(t)\|_H^2 + \sup_{t\in[0,T]} \mathbb{E}\|J_2(t)\|_H^2 + \sup_{t\in[0,T]} \mathbb{E}\|J_3(t)\|_H^2). \quad (2.10)$$

We are going to look at each summand of (2.10) individually. We start with J_1 and since y_0 is a fixed element in H, we find

$$\sup_{t\in[0,T]} \mathbb{E}\|J_1(t)\|_H^2 = \sup_{t\in[0,T]} \|(T(t) - T_n(t))y_0\|_H^2 \longrightarrow 0 \text{ for } n \to \infty,$$

by Corollary A.24 (III).

In order to deal with the second summand J_2 in (2.10), we start by rewriting it as follows

$$\begin{aligned} &T(t-s)F(s,Y(s)) - T_n(t-s)F(s,Z_n(s)) \\ &\qquad = (T(t-s) - T_n(t-s))F(s,Y(s)) + T_n(t-s)(F(s,Y(s)) - F(s,Z_n(s))). \end{aligned}$$

By Hölder's inequality and Fubini–Tonelli theorem, we find

$$\begin{aligned} \sup_{t\in[0,T]} \mathbb{E}\|J_2(t)\|_H^2 &\leqslant T \sup_{t\in[0,T]} \mathbb{E} \int_0^t \|T(t-s)F(s,Y(s)) - T_n(t-s)F(s,Z_n(s))\|_H^2 ds \\ &\leqslant 2T\big(\sup_{t\in[0,T]} \int_0^t \mathbb{E}\|(T(t-s) - T_n(t-s))F(s,Y(s))\|_H^2 ds \\ &\qquad + \sup_{t\in[0,T]} \int_0^t \mathbb{E}\|T_n(t-s)(F(s,Y(s)) - F(s,Z_n(s)))\|_H^2 ds\big) \\ &=: 2T(J_{21} + J_{22}). \end{aligned}$$

For the term J_{21} we estimate

$$J_{21} \leqslant \sup_{t\in[0,T]} \int_0^t \sup_{\tau\in[s,T]} \mathbb{E}\|(T(\tau-s)-T_n(\tau-s))F(s,Y(s))\|_H^2 ds$$
$$\leqslant \int_0^T \sup_{t\in[s,T]} \mathbb{E}\|(T(t-s)-T_n(t-s))F(s,Y(s))\|_H^2 ds$$
$$\leqslant \int_0^T \mathbb{E} \sup_{t\in[s,T]} \|(T(t-s)-T_n(t-s))F(s,Y(s))\|_H^2 ds. \tag{2.11}$$

Our aim is to show with the help of Lebesgue's dominated convergence theorem that (2.11) converges to zero. Therefore, we look at the pointwise convergence for fixed s and ω. Then, $F(s,Y(s))$ is a fixed element in H. Hence, by Corollary A.24 (III)

$$\sup_{t\in[s,T]} \|(T(t-s)-T_n(t-s))F(s,Y(s))\|_H^2 \longrightarrow 0 \text{ for } n\to\infty.$$

We gain the majorant from the following estimate

$$\sup_{t\in[s,T]} \|(T(t-s)-T_n(t-s))F(s,Y(s))\|_H^2 \leqslant 4M^2e^{2\lambda T}\|F(s,Y(s))\|_H^2.$$

In order to see that the right-hand side is integrable, we calculate

$$\int_0^T \mathbb{E}\; 4M^2e^{2\lambda T}\|F(s,Y(s))\|_H^2 ds \leqslant 4M^2e^{2\lambda T}k_F^2\int_0^T \mathbb{E}(1+\|Y(s)\|_H)^2 ds$$
$$\leqslant 8M^2e^{2\lambda T}k_F^2\int_0^T 1+\mathbb{E}\|Y(s)\|_H^2 ds \leqslant 8M^2e^{2\lambda T}k_F^2 T(1+\sup_{t\in[0,T]}\mathbb{E}\|Y(t)\|_H^2) < \infty,$$

where we used the linear growth condition of F. The right-hand side is finite, since Y is the mild solution of $(S-ACP)$. Thus, by Lebesgue's dominated convergence theorem we find

$$J_{21} = \sup_{t\in[0,T]} \mathbb{E}\int_0^t \|(T(t-s)-T_n(t-s))F(s,Y(s))\|_H^2 ds \longrightarrow 0 \text{ for } n\to\infty.$$

For the term J_{22} we apply the Lipschitz continuity of F and find

$$J_{22} \leqslant M^2e^{2\lambda T}C_F^2\int_0^T \mathbb{E}\|Y(s)-Z_n(s)\|_H^2 ds \leqslant M^2e^{2\lambda T}C_F^2\int_0^T \sup_{s\in[0,\tau]}\mathbb{E}\|Y(s)-Z_n(s)\|_H^2 d\tau.$$

For J_3 in (2.10) we start again by rewriting the relevant expression as follows

$$\int_0^t T(t-s)G(s,Y(s))-T_n(t-s)G(s,Z_n(s))dM(s)$$
$$=\int_0^t (T(t-s)-T_n(t-s))G(s,Y(s))dM(s)+\int_0^t T_n(t-s)(G(s,Y(s))-G(s,Z_n(s)))dM(s).$$

Thus, we estimate

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\|J_3(t)\|_H^2 \leqslant 2 \sup_{t\in[0,T]} \mathbb{E}\| \int_0^t (T(t-s) - T_n(t-s))G(s,Y(s))dM(s)\|_H^2 \\ + 2 \sup_{t\in[0,T]} \mathbb{E}\| \int_0^t T_n(t-s)(G(s,Y(s)) - G(s,Z_n(s)))dM(s)\|_H^2 \\ =: 2(J_{31} + J_{32}).\end{aligned}$$

To show that the term J_{31} converges to zero is the crucial step of the proof. We start by estimating using the isometry of the stochastic integral and the definition of the Hilbert-Schmidt norm

$$\begin{aligned}J_{31} &= \sup_{t\in[0,T]} \int_0^t \mathbb{E}\|(T(t-s) - T_n(t-s))G(s,Y(s))Q^{1/2}\|_{L_{(HS)}(U,H)}^2 ds \\ &\leqslant \sup_{t\in[0,T]} \int_0^t \sup_{\tau\in[s,T]} \mathbb{E}\|(T(\tau-s) - T_n(\tau-s))G(s,Y(s))Q^{1/2}\|_{L_{(HS)}(U,H)}^2 ds \\ &\leqslant \int_0^T \sup_{t\in[s,T]} \mathbb{E}\|(T(t-s) - T_n(t-s))G(s,Y(s))Q^{1/2}\|_{L_{(HS)}(U,H)}^2 ds \\ &= \int_0^T \sup_{t\in[s,T]} \mathbb{E}\sum_{k=1}^{\infty} \|(T(t-s) - T_n(t-s))G(s,Y(s))Q^{1/2}f_k\|_H^2 ds \\ &\leqslant \int_0^T \mathbb{E}\sum_{k=1}^{\infty} \sup_{t\in[s,T]} \|(T(t-s) - T_n(t-s))G(s,Y(s))Q^{1/2}f_k\|_H^2 ds,\end{aligned}$$

where $\{f_k\}_{k\in\mathbb{N}}$ is an orthonormal basis of U. In order to apply Lebesgue's dominated convergence theorem, we interpret the series as an integral with respect to the counting measure. First, we look at the pointwise convergence for fixed s, ω, and k. Then, $G(s,Y(s))Q^{1/2}f_k$ is a fixed element in H. Hence, by Corollary A.24 (III)

$$\sup_{t\in[s,T]} \|(T(t-s) - T_n(t-s))G(s,Y(s))Q^{1/2}f_k\|_H^2 \longrightarrow 0 \text{ for } n\to\infty.$$

We gain the majorant from the following estimate

$$\sup_{t\in[s,T]} \|(T(t-s) - T_n(t-s))G(s,Y(s))Q^{1/2}f_k\|_H^2 \leqslant 4M^2e^{2\lambda T}\|G(s,Y(s))Q^{1/2}f_k\|_H^2.$$

In order to see that the right-hand side is integrable we calculate

$$\begin{aligned}\int_0^T \mathbb{E}\sum_{k=1}^{\infty} 4M^2e^{2\lambda T}\|G(s,Y(s))Q^{1/2}f_k\|_H^2 ds = 4M^2e^{2\lambda T}\int_0^T \mathbb{E}\|G(s,Y(s))Q^{1/2}\|_{L_{(HS)}(U,H)}^2 ds \\ \leqslant 4M^2e^{2\lambda T}k_G^2\int_0^T \mathbb{E}(1+\|Y(s)\|_H)^2 ds \leqslant 8M^2e^{2\lambda T}k_G^2 T(1+\sup_{t\in[0,T]}\mathbb{E}\|Y(t)\|_H^2) < \infty,\end{aligned}$$

where we used the linear growth condition of G. The right-hand side is finite, since Y is the mild solution of $(S-ACP)$. Thus, by Lebesgue's dominated convergence theorem we find

$$J_{31} = \sup_{t\in[0,T]} \int_0^t \mathbb{E}\|(T(t-s) - T_n(t-s))G(s,Y(s))Q^{1/2}\|_{L_{(HS)}(U,H)}^2 ds \longrightarrow 0 \text{ for } n\to\infty.$$

For the term J_{32} we apply the isometry of the stochastic integral and the Lipschitz continuity of G, that is

$$\begin{aligned} J_{32} &= \sup_{t\in[0,T]} \| \int_0^t \mathbb{E}\|T_n(t-s)(G(s,Y(s)) - G(s,Z_n(s)))\|^2_{L_{(HS)}(U,H)} ds \\ &\leqslant M^2 e^{2\lambda T} C_G^2 \int_0^T \mathbb{E}\|Y(s) - Z_n(s)\|_H^2 ds \\ &\leqslant M^2 e^{2\lambda T} C_G^2 \int_0^T \sup_{s\in[0,\tau]} \mathbb{E}\|Y(s) - Z_n(s)\|_H^2 d\tau. \end{aligned}$$

Now, we collect all terms and from (2.10) follows

$$\begin{aligned} \sup_{t\in[0,T]} \mathbb{E}\|Y(t) - Z_n(t)\|_H^2 &\leqslant 3\tilde{\varepsilon}(n) + 6T(\tilde{\varepsilon}(n) + M^2 e^{2\lambda T} C_F^2 \int_0^T \sup_{s\in[0,\tau]} \mathbb{E}\|Y(s) - Z_n(s)\|_H^2 d\tau) \\ &\quad + 6(\tilde{\varepsilon}(n) + M^2 e^{2\lambda T} C_G^2 \int_0^T \sup_{s\in[0,\tau]} \mathbb{E}\|Y(s) - Z_n(s)\|_H^2 d\tau) \\ &\leqslant \varepsilon(n) + C(T, M, \lambda, C_F, C_G) \int_0^T \sup_{s\in[0,\tau]} \mathbb{E}\|Y(s) - Z_n(s)\|_H^2 d\tau, \end{aligned}$$

where $\{\tilde{\varepsilon}(n)\}_{n\in\mathbb{N}}$ is a suitable null sequence, $\varepsilon := (9 + 6T)\tilde{\varepsilon}$, and C stands for a constant depending only on the values stated in the parentheses. Using Grönwall's inequality, we obtain

$$\sup_{t\in[0,T]} \mathbb{E}\|Y(t) - Z_n(t)\|_H^2 \leqslant \varepsilon(n)(1 + C(T, M, \lambda, C_F, C_G) T e^{C(T,M,\lambda,C_F,C_G)T}).$$

Hence, for $n \to \infty$ we find

$$\|Y - Z_n\|_{T,H}^2 = \sup_{t\in[0,T]} \mathbb{E}\|Y(t) - Z_n(t)\|_H^2 \longrightarrow 0.$$

□

Remark 2.14. *If we compare the two approximation theorems of this Chapter, we see that the main difficulty in both proofs was to show that the term I_{31} and J_{31} respectively converges to zero. If we compare the two arguments we used to obtain the convergence, we see that the one made in Proposition 2.2 is simpler, than the one of Theorem 2.13. Another advantage of the approximation scheme from Section 2.1 is that the operator A is not affected by the approximation. This will make the proof of the transformation formula for mild solution in Chapter 3 easier. On the other hand, we are restricted to the case of a generalized contraction C_0-semigroup if we want to apply Theorem 2.3. This is the main diffence to Theorem 2.13, which can always be applied.*

Theorem 2.13 also shows that any approximation $\{A_n\}_{n\in\mathbb{N}}$ of the operator A will work, as long as the semigroups $(T_n(t))_{t\geqslant 0}$ generated by A_n fulfill the stability condition $\|T_n(t)\|_{L(H)} \leqslant Me^{\lambda t}$ for all $n \in \mathbb{N}$ and we have $T_n(t)h \to T(t)h$ for $n \to \infty$ for all $t \geqslant 0$ and all $h \in H$, where the convergence is even uniform on each interval $[0, t_0]$. This holds in particular if one wants to prove a finite dimensional approximation like the Galerkin method.

CHAPTER 3

Transformation Formula

The aim of this chapter is to prove a rigorous transformation formula, that is, Itô's formula, for the mild solution Y of the stochastic Cauchy problem $(S-ACP)$ with Lévy noise. Therefore, we first provide a transformation formula for square integrable Lévy processes with drift. Since the mild solution is not regular enough (see Theorem 1.35), we cannot apply this transformation formula directly. Another problem we face is that Y is only taking values in the state space H and not in the domain of the generator $\mathscr{D}(A)$. To overcome those problems we apply the approximation of the mild solution from Theorem 2.3 and request more regularity of the transformation function. Through this procedure, we obtain a transformation formula for the mild solution Y. Since the adjoint operator A^* of the generator A appears in this formula, we calculate the adjoint operator $\mathscr{A}^*$ of the delay equation for important examples.

3.1. Transformation formula for Lévy processes

In this section, we prove the transformation formula for a well-defined square integrable Lévy process with drift taking values in a separable Hilbert space H. Therefore, we have to change our notation to the Poisson integral for the jumps of the Lévy process. For more details on the notation we refer to Secton B.2 and Section B.4 in the Appendix. There is a rich literature for the finite dimensional case. One finds the transformation formula for example in [**GS71**, Section II.2 §6.], [**Pro05**, Section II.7], or in [**App09**, Section 4.4]. In the case of Wiener noise in infinite dimension one can turn to [**DPZ92**, Section 4.4.5] or with weaker assumptions to [**GM11a**, Section 2.3]. For the pure jump noise case in infinite dimension a transformation formula with weak conditions was proven in [**MRT13**]. An older result for Lévy processes in infinite dimension with strong assumptions on the transformation function can be found in [**Mét82**, Theorem 27.2]. Since the transformation formula is known for the Wiener noise and the pure jump noise case our goal in this section is to combine those two results to receive the transformation formula for the Lévy noise case. We mainly use results from [**App09**], [**GM11a**], and [**MRT13**] to accomplish this.

The first step is to generalize Itô's formula from Theorem B.12 for the Wiener noise case. We want to replace the deterministic times by stopping times. Thus, we have to consider stochastic integrals with random limits as introduced in Lemma B.11. For notation of derivatives we refer to Section B.1 in the Appendix.

Corollary 3.1. *Let Q be a symmetric nonnegative trace-class operator on a separable Hilbert space U. Furthermore, let $\{W(t)\}_{0\leqslant t\leqslant T}$ be a Q-Wiener process on a filtered probability space*

$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, \mathbb{P})$. *Assume that the stochastic process* $X(t)$, $0 \leqslant t \leqslant T$, *is given by*

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

where $X(0)$ *is an* $\mathcal{F}_0$*-measurable* H*-valued random variable,* Ψ *is an* H*-valued* $\mathcal{F}_s$*-adapted* $\mathbb{P}$*-a.s. Bochner-integrable process on* $[0,T]$,

$$\int_0^T \|\Psi(s)\|_H ds < \infty \quad \mathbb{P}\text{-a.s.},$$

and $\Phi \in \mathcal{P}_T^2(U,H)$.
Assume that the function $\phi : H \to \mathbb{R}$ *is such that* ϕ *is continuous and its Fréchet derivatives* ϕ', ϕ'' *are continuous and bounded on bounded subsets of* H. *Then, the following Itô formula holds for all stopping times* τ_1 *and* τ_2 *with* $\mathbb{P}(0 \leqslant \tau_1 \leqslant \tau_2 \leqslant T) = 1$

$$\begin{aligned}\phi(X(\tau_2)) = \phi(X(\tau_1)) &+ \int_{\tau_1}^{\tau_2} \langle \phi'(X(s)), \Phi(s)dW(s)\rangle_H \\ &+ \int_{\tau_1}^{\tau_2} \{\langle \phi'(X(s)), \Psi(s)\rangle_H + \frac{1}{2}\operatorname{tr}[\phi''(X(s))(\Phi(s)Q^{1/2})(\Phi(s)Q^{1/2})^*]\}ds, \qquad (3.1)\end{aligned}$$

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

PROOF.

We define the processes $Z_i(t)$, $0 \leqslant t \leqslant T$, for $i = 1, 2$ by

$$Z_i(t) = X(0) + \int_0^t \mathbb{1}_{[0,\tau_i]}(s)\Psi(s)ds + \int_0^t \mathbb{1}_{[0,\tau_i]}(s)\Phi(s)dW(s).$$

Note that $Z_i(t) = X(t)$ $\mathbb{P}$-a.s. for all $t \in [0, \tau_i]$ and by Lemma B.11 we find $\mathbb{P}$-a.s.

$$\begin{aligned}Z_i(T) &= X(0) + \int_0^T \mathbb{1}_{[0,\tau_i]}(s)\Psi(s)ds + \int_0^T \mathbb{1}_{[0,\tau_i]}(s)\Phi(s)dW(s) \\ &= X(0) + \int_0^{\tau_i} \Psi(s)ds + \int_0^{\tau_i} \Phi(s)dW(s) = X(\tau_i).\end{aligned}$$

Now, using the Itô formula for deterministic times from Theorem B.12 and Lemma B.11 again we calculate $\mathbb{P}$-a.s.

$$\begin{aligned}\phi(Z_i(T)) &= \phi(Z_i(0)) + \int_0^T \langle \phi'(Z_i(s)), \mathbb{1}_{[0,\tau_i]}(s)\Phi(s)dW(s)\rangle_H + \int_0^T \{\langle \phi'(Z_i(s)), \mathbb{1}_{[0,\tau_i]}(s)\Psi(s)\rangle_H \\ &\qquad + \frac{1}{2}\operatorname{tr}[\phi''(Z_i(s))(\mathbb{1}_{[0,\tau_i]}(s)\Phi(s)Q^{1/2})(\mathbb{1}_{[0,\tau_i]}(s)\Phi(s)Q^{1/2})^*]\}ds \\ &= \phi(X(0)) + \int_0^{\tau_i} \langle \phi'(Z_i(s)), \Phi(s)dW(s)\rangle_H \\ &\quad + \int_0^{\tau_i} \{\langle \phi'(Z_i(s)), \Psi(s)\rangle_H + \frac{1}{2}\operatorname{tr}[\phi''(Z_i(s))(\Phi(s)Q^{1/2})(\Phi(s)Q^{1/2})^*]\}ds \\ &= \phi(X(0)) + \int_0^{\tau_i} \langle \phi'(X(s)), \Phi(s)dW(s)\rangle_H \\ &\quad + \int_0^{\tau_i} \{\langle \phi'(X(s)), \Psi(s)\rangle_H + \frac{1}{2}\operatorname{tr}[\phi''(X(s))(\Phi(s)Q^{1/2})(\Phi(s)Q^{1/2})^*]\}ds.\end{aligned}$$

We obtain (3.1) by subtraction and identity (B.3)

$$\begin{aligned}\phi(X(\tau_2)) - \phi(X(\tau_1)) =& \phi(Z_2(T)) - \phi(Z_1(T))\\ =& \int_{\tau_1}^{\tau_2} \langle \phi'(X(s)), \Phi(s) dW(s)\rangle_H\\ &+ \int_{\tau_1}^{\tau_2} \{\langle \phi'(X(s)), \Psi(s)\rangle_H + \frac{1}{2}\operatorname{tr}[\phi''(X(s))(\Phi(s)Q^{1/2})(\Phi(s)Q^{1/2})^*]\} ds.\end{aligned}$$

□

Now, we turn to a square integrable Lévy process taking values in the separable Hilbert space U. First, we consider the case of big jumps.

PROPOSITION 3.2. *Let $A \in \mathscr{B}(U\backslash\{0\})$, such that $\nu(A) < \infty$. We consider the process*

$$Y(t) = Y(0) + \int_0^t \Psi(s)ds + \int_0^t \Phi(s)dW_{Q_0}(s) + \int_0^t \int_A \Upsilon(s)xN(ds,dx) + \int_0^t \int_A \Lambda(s)x\nu(dx)ds,$$

where Ψ, Φ, Υ and Λ are progressively measurable processes with

$$\int_0^t \|\Psi(s)\|_H ds + \int_0^t \int_A \|\Lambda(s)x\|_H \nu(dx)ds < \infty \quad \mathbb{P}\text{-a.s.},$$

for all $t \in [0,T]$, and $\Phi \in \mathcal{P}_T^2(U,H)$. Then, for a function $\phi \in C^2(H;\mathbb{R})$, with ϕ', ϕ'' bounded on bounded subset of H, we have $\mathbb{P}$-a.s.

$$\begin{aligned}\phi(Y(t)) = \phi(Y(0)) &+ \int_0^t \langle \phi'(Y(s-)), \Phi(s) dW_{Q_0}(s)\rangle_H\\ &+ \int_0^t \{\langle \phi'(Y(s-)), \Psi(s)\rangle_H + \frac{1}{2}\operatorname{tr}[\phi''(Y(s-))(\Phi(s)Q_0^{1/2})(\Phi(s)Q_0^{1/2})^*]\} ds\\ &+ \int_0^t \int_A \{\phi(Y(s-) + \Upsilon(s)x) - \phi(Y(s-))\} N(ds,dx)\\ &+ \int_0^t \int_A \langle \phi'(Y(s-)), \Lambda(s)x\rangle_H \nu(dx)ds,\end{aligned}$$

for all $t \in [0,T]$ and all integrals appearing above are well-defined.

PROOF.

The existence of the integrals is guaranteed by Theorem B.12 and [**MRT13**, Proposition 3.3.]. We introduce the following notation

$$\begin{aligned}Y^c(t) &:= Y(0) + \int_0^t \Psi(s)ds + \int_0^t \Phi(s)dW_{Q_0}(s) + \int_0^t \int_A \Lambda(s)x\nu(dx)ds,\\ Y^N(t) &:= \int_0^t \int_A \Upsilon(s)xN(ds,dx),\end{aligned}$$

where the c indicates the continuous part of the process Y. Obviously, we have $\mathbb{P}$-a.s.

$$Y(t) = Y^c(t) + Y^N(t).$$

We consider the stochastic process $Z(t) := \int_A xN(t,dx)$. The jump times of Z will play a crucial role for the proof. Recall that the jump times are defined recursively as $T_0^A := 0$ and, for each $n \in \mathbb{N}$, $T_n^A := \inf\{t > T_{n-1}^A : \Delta Z(t) \in A\}$. Using those stopping times we find $\mathbb{P}$-a.s.

$$\begin{aligned}\phi(Y(t)) - \phi(Y(0)) &= \sum_{j=0}^{\infty} \phi(Y(t \wedge T_{j+1}^A)) - \phi(Y(t \wedge T_j^A)) \\ &= \sum_{j=0}^{\infty} \phi(Y(t \wedge T_{j+1}^A -)) - \phi(Y(t \wedge T_j^A)) \qquad (3.2) \\ &\quad + \sum_{j=0}^{\infty} \phi(Y(t \wedge T_{j+1}^A)) - \phi(Y(t \wedge T_{j+1}^A -)). \qquad (3.3)\end{aligned}$$

We look at (3.2) first. Observe that for all t with $T_j^A < t < T_{j+1}^A$, this means t lies between two consecutive jumps, we have $\mathbb{P}$-a.s. $Y(t) = Y(T_j^A) + Y^c(t) - Y^c(T_j^A)$. Thus, each summend in (3.2) behaves $\mathbb{P}$-a.s. like a continuous stochastic process, that is for $t \in (T_j^A, T_{j+1}^A)$ the process Y does not jump and we have $Y(t) = Y(t-)$. Applying Lemma B.11 leads $\mathbb{P}$-a.s. to the following identity

$$\begin{aligned}&\int_{t\wedge T_j^A}^{t \wedge T_{j+1}^A -} \langle \phi'(Y(s)), \Phi(s) dW_{Q_0}(s)\rangle_H \\ &= \int_0^{t\wedge T_{j+1}^A -} \langle \phi'(Y(s)), \Phi(s) dW_{Q_0}(s)\rangle_H - \int_0^{t\wedge T_j^A} \langle \phi'(Y(s)), \Phi(s) dW_{Q_0}(s)\rangle_H \\ &= \int_0^T (\mathbb{1}_{[0,t\wedge T_{j+1}^A -]}(s) - \mathbb{1}_{[0,t\wedge T_j^A]}(s)) \langle \phi'(Y(s)), \Phi(s) dW_{Q_0}(s)\rangle_H \\ &= \int_0^T \mathbb{1}_{(t\wedge T_j^A, t\wedge T_{j+1}^A)}(s) \langle \phi'(Y(s)), \Phi(s) dW_{Q_0}(s)\rangle_H \\ &= \int_0^T \mathbb{1}_{(t\wedge T_j^A, t\wedge T_{j+1}^A]}(s) \langle \phi'(Y(s-)), \Phi(s) dW_{Q_0}(s)\rangle_H \\ &= \int_{t\wedge T_j^A}^{t\wedge T_{j+1}^A} \langle \phi'(Y(s-)), \Phi(s) dW_{Q_0}(s)\rangle_H.\end{aligned}$$

Obviously, the same procedure works for ds-integrals. Thus, we are able to apply Corollary 3.1 to (3.2) and receive $\mathbb{P}$-a.s.

$$\begin{aligned}&\sum_{j=0}^{\infty} \phi(Y(t \wedge T_{j+1}^A -)) - \phi(Y(t \wedge T_j^A)) \\ &= \sum_{j=0}^{\infty} \Bigg[\int_{t\wedge T_j^A}^{t\wedge T_{j+1}^A -} \langle \phi'(Y(s)), \Phi(s) dW_{Q_0}(s)\rangle_H \\ &\qquad + \int_{t\wedge T_j^A}^{t\wedge T_{j+1}^A -} \{\langle \phi'(Y(s)), \Psi(s)\rangle_H + \langle \phi'(Y(s)), \int_A \Lambda(s) x \nu(dx)\rangle_H \\ &\qquad\qquad + \frac{1}{2} \operatorname{tr}[\phi''(Y(s))(\Phi(s) Q_0^{1/2})(\Phi(s) Q_0^{1/2})^*]\} ds \Bigg]\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=0}^{\infty} \Bigg[\int_{t\wedge T_j^A}^{t\wedge T_{j+1}^A} \langle \phi'(Y(s-)), \Phi(s)dW_{Q_0}(s)\rangle_H \\
&\qquad + \int_{t\wedge T_j^A}^{t\wedge T_{j+1}^A} \{\langle \phi'(Y(s-)), \Psi(s)\rangle_H + \langle \phi'(Y(s-)), \int_A \Lambda(s)x\nu(dx)\rangle_H \\
&\qquad\qquad + \frac{1}{2}\operatorname{tr}[\phi''(Y(s-))(\Phi(s)Q_0^{1/2})(\Phi(s)Q_0^{1/2})^*]\}ds \Bigg] \\
&= \int_0^t \langle \phi'(Y(s-)), \Phi(s)dW_{Q_0}(s)\rangle_H \\
&\quad + \int_0^t \{\langle \phi'(Y(s-)), \Psi(s)\rangle_H + \frac{1}{2}\operatorname{tr}[\phi''(Y(s-))(\Phi(s)Q_0^{1/2})(\Phi(s)Q_0^{1/2})^*]\}ds \\
&\quad + \int_0^t \int_A \langle \phi'(Y(s-)), \Lambda(s)x\rangle_H \nu(dx)ds.
\end{aligned}
$$

Now we turn to (3.3). Recall that the integral with respect to a Poisson random measure is a random finite sum, that is $\int_0^t \int_A f(s,x)N(ds,dx) = \sum_{j=1}^{\infty} f(T_j^A, \Delta Z(T_j^A))\mathbb{1}_{\{T_j^A \leqslant t\}}$. Using this fact we calculate $\mathbb{P}$-a.s.

$$
\begin{aligned}
Y^N(T_i^A) &= \int_0^{T_i^A}\int_A \Upsilon(s)xN(ds,dx) = \sum_{j=0}^{\infty} \Upsilon(T_j^A)\Delta Z(T_j^A)\mathbb{1}_{\{T_j^A \leqslant T_i^A\}} \\
&= \sum_{j=0}^{\infty} \Upsilon(T_j^A)\Delta Z(T_j^A)\mathbb{1}_{\{T_j^A \leqslant T_i^A -\}} + \Upsilon(T_i^A)\Delta Z(T_i^A) \\
&= \int_0^{T_i^A -}\int_A \Upsilon(s)xN(ds,dx) + \Upsilon(T_i^A)\Delta Z(T_i^A) = Y^N(T_i^A -) + \Upsilon(T_i^A)\Delta Z(T_i^A).
\end{aligned}
$$

Note that for all jump times $T_j^A > t$ the corresponding summands in (3.3) are zero. Applying those preliminary thoughts to (3.3) yields $\mathbb{P}$-a.s.

$$
\begin{aligned}
\sum_{j=0}^{\infty} \phi(Y(t\wedge T_{j+1}^A)) - \phi(Y(t\wedge T_{j+1}^A -)) &= \sum_{j=1}^{\infty} [\phi(Y(t\wedge T_j^A)) - \phi(Y(t\wedge T_j^A -))]\mathbb{1}_{\{T_j^A \leqslant t\}} \\
&= \sum_{j=1}^{\infty} [\phi(Y^c(T_j^A) + Y^N(T_j^A)) - \phi(Y(T_j^A -))]\mathbb{1}_{\{T_j^A \leqslant t\}} \\
&= \sum_{j=1}^{\infty} [\phi(Y^c(T_j^A -) + Y^N(T_j^A -) + \Upsilon(T_j^A)\Delta Z(T_j^A)) - \phi(Y(T_j^A -))]\mathbb{1}_{\{T_j^A \leqslant t\}} \\
&= \sum_{j=1}^{\infty} [\phi(Y(T_j^A -) + \Upsilon(T_j^A)\Delta Z(T_j^A)) - \phi(Y(T_j^A -))]\mathbb{1}_{\{T_j^A \leqslant t\}} \\
&= \int_0^t \int_A \{\phi(Y(s-) + \Upsilon(s)x) - \phi(Y(s-))\}N(ds,dx).
\end{aligned}
$$

□

With the help of Proposition 3.2 we are able to prove the transformation formula for the Lévy noise case. In the formulation of the following theorem quasi-sublinear functions appear and are defined in Definition B.49 in the Appendix.

THEOREM 3.3. *Let* $\phi \in C^2(H;\mathbb{R})$ *such that*

$$\|\phi'(h)\|_H \leqslant h_1(\|h\|_H),$$
$$\|\phi''(h)\|_{L(H)} \leqslant h_2(\|h\|_H),$$

where $h \in H$ *and* $h_1, h_2 : \mathbb{R}_+ \to \mathbb{R}_+$ *are quasi-sublinear functions. We consider the following stochastic process*

$$Y(t) = Y(0) + \int_0^t \Psi(s)ds + \int_0^t \Phi(s)dW_{Q_0}(s) + \int_0^t \int_A \Upsilon(s)x\tilde{N}(ds,dx) + \int_0^t \int_{A^c} \Lambda(s)xN(ds,dx),$$

where $t \in [0,T]$, $A \in \mathscr{B}(U\backslash\{0\})$ *with* $\nu(A^c) < \infty$, *where* A^c *is the complement of* A. *Furthermore, let* Ψ, Φ, Υ *and* Λ *be progressively measurable processes, such that* $\mathbb{P}$*-a.s.*

$$\int_0^t \|\Psi(s)\|_H ds + \int_0^t \int_A \|\Upsilon(s)x\|_H^2 \nu(dx)ds + \int_0^t \int_A h_1(\|\Upsilon(s)x\|_H)^2 \|\Upsilon(s)x\|_H^2 \nu(dx)ds$$
$$+ \int_0^t \int_A h_2(\|\Upsilon(s)x\|_H)\|\Upsilon(s)x\|_H^2 \nu(dx)ds < \infty, \tag{3.4}$$

and $\Phi \in \mathcal{P}_T^2(U,H)$. *Then, we have* $\mathbb{P}$*-a.s.*

$$\begin{aligned}
\phi(Y(t)) =&\phi(Y(0)) + \int_0^t \langle \phi'(Y(s-)), \Phi(s)dW_{Q_0}(s)\rangle_H \\
&+ \int_0^t \{\langle \phi'(Y(s-)), \Psi(s)\rangle_H + \frac{1}{2}\operatorname{tr}[\phi''(Y(s-))(\Phi(s)Q_0^{1/2})(\Phi(s)Q_0^{1/2})^*]\}ds \\
&+ \int_0^t \int_A \{\phi(Y(s-) + \Upsilon(s)x) - \phi(Y(s-))\}\tilde{N}(ds,dx) \qquad (3.5)\\
&+ \int_0^t \int_A \{\phi(Y(s-) + \Upsilon(s)x) - \phi(Y(s-)) - \langle \phi'(Y(s-)), \Upsilon(s)x\rangle_H\}\nu(dx)ds \\
&+ \int_0^t \int_A \{\phi(Y(s-) + \Lambda(s)x) - \phi(Y(s-))\}N(ds,dx),
\end{aligned}$$

for all $t \in [0,T]$ *and all integrals are well-defined.*

PROOF.

For the existence of the integrals see Theorem B.12 and Theorem B.50. Let $\{B_n\}_{n\in\mathbb{N}}$ be a sequence with $B_n \in \mathscr{B}(U\backslash\{0\})$, such that $B_n \uparrow U$, where $\nu(B_n) < \infty$ for each $n \in \mathbb{N}$. For example we could choose $B_n = B_{1/n}^c(0)$. We define the following stochastic process

$$\begin{aligned}
Y_n(t) := Y(0) + \int_0^t \Psi(s)ds + \int_0^t \Phi(s)dW_{Q_0}(s) \\
+ \int_0^t \int_{A\cap B_n} \Upsilon(s)x\tilde{N}(ds,dx) + \int_0^t \int_{A^c\cap B_n} \Lambda(s)xN(ds,dx).
\end{aligned}$$

In order to see that we can apply Proposition 3.2, we rewrite the process Y_n as follows

$$\begin{aligned}Y_n(t) = Y(0) + \int_0^t \Psi(s)ds + \int_0^t \Phi(s)dW_{Q_0}(s)\\ + \int_0^t \int_{B_n} \mathbb{1}_A(x)\Upsilon(s)x + \mathbb{1}_{A^c}(x)\Lambda(s)xN(ds,dx) - \int_0^t \int_{A\cap B_n} \Upsilon(s)x\nu(dx)ds.\end{aligned}$$

Note that the stochastic process Y_n fulfills all assumptions of Proposition 3.2 for each $n \in \mathbb{N}$. In particular, we find by using assumption (3.4) and Hölder's inequality, that $\mathbb{P}$-a.s. and for all $t \in [0,T]$ $\int_0^t \int_{A\cap B_n} \|\Upsilon(s)x\|\nu(dx)ds \leqslant (t\nu(A\cap B_n))^{1/2}\Big(\int_0^t \int_{A\cap B_n} \|\Upsilon(s)x\|^2\nu(dx)ds\Big)^{1/2} < \infty$. Applying Proposition 3.2 to the process Y_n yields $\mathbb{P}$-a.s.

$$\begin{aligned}\phi(Y_n(t)) =&\phi(Y(0)) + \int_0^t \langle \phi'(Y_n(s-)), \Phi(s)dW_{Q_0}(s)\rangle_H\\ &+ \int_0^t \{\langle \phi'(Y_n(s-)), \Psi(s)\rangle_H + \frac{1}{2}\operatorname{tr}[\phi''(Y_n(s-))(\Phi(s)Q_0^{1/2})(\Phi(s)Q_0^{1/2})^*]\}ds\\ &+ \int_0^t \int_{B_n} \{\phi(Y_n(s-) + \Upsilon(s)x\mathbb{1}_A(x) + \Lambda(s)x\mathbb{1}_{A^c}(x)) - \phi(Y_n(s-))\}N(ds,dx)\\ &- \int_0^t \int_{A\cap B_n} \langle \phi'(Y_n(s-)), \Upsilon(s)x\rangle_H\nu(dx)ds,\end{aligned}$$

for all $t \in [0,T]$. We rewrite the Poisson integral in the following way and receive $\mathbb{P}$-a.s.

$$\begin{aligned}\phi(Y_n(t)) =&\phi(Y(0)) + \int_0^t \langle \phi'(Y_n(s-)), \Phi(s)dW_{Q_0}(s)\rangle_H\\ &+ \int_0^t \{\langle \phi'(Y_n(s-)), \Psi(s)\rangle_H + \frac{1}{2}\operatorname{tr}[\phi''(Y_n(s-))(\Phi(s)Q_0^{1/2})(\Phi(s)Q_0^{1/2})^*]\}ds\\ &+ \int_0^t \int_{A\cap B_n} \{\phi(Y_n(s-) + \Upsilon(s)x) - \phi(Y_n(s-))\}N(ds,dx)\\ &+ \int_0^t \int_{A^c\cap B_n} \phi(Y_n(s-) + \Lambda(s)x) - \phi(Y_n(s-))\}N(ds,dx)\\ &- \int_0^t \int_{A\cap B_n} \langle \phi'(Y_n(s-)), \Upsilon(s)x\rangle_H\nu(dx)ds,\end{aligned}$$

and for all $t \in [0,T]$. Before we pass to the limit we add the compensator. This leads to

$$\phi(Y_n(t)) = \phi(Y(0)) + \int_0^t \langle \phi'(Y_n(s-)), \Phi(s)dW_{Q_0}(s)\rangle_H \tag{3.6}$$

$$+ \int_0^t \{\langle \phi'(Y_n(s-)), \Psi(s)\rangle_H + \frac{1}{2}\operatorname{tr}[\phi''(Y_n(s-))(\Phi(s)Q_0^{1/2})(\Phi(s)Q_0^{1/2})^*]\}ds$$

$$+ \int_0^t \int_A \mathbb{1}_{B_n}(x)\{\phi(Y_n(s-) + \Upsilon(s)x) - \phi(Y_n(s-))\}\tilde{N}(ds,dx) \tag{3.7}$$

$$+ \int_0^t \int_{A^c\cap B_n} \{\phi(Y_n(s-) + \Lambda(s)x) - \phi(Y_n(s-))\}N(ds,dx)$$

$$+ \int_0^t \int_A \mathbb{1}_{B_n}(x)\{\phi(Y_n(s-) + \Upsilon(s)x) - \phi(Y_n(s-)) - \langle \phi'(Y_n(s-)), \Upsilon(s)x\rangle_H\}\nu(dx)ds, \tag{3.8}$$

$\mathbb{P}$-a.s. and for all $t \in [0,T]$. To show that the integrals converge we apply Lebesgue's dominated convergence theorem and its stochastic versions, see Theorem B.13 and Theorem B.48. In oder to find the majorants, first note that for all $h \in H$ and any $R > 0$ with $\|h\| \leqslant R$, there exists a constant $C_R > 0$, such that $\|\phi'(h)\|_H + \|\phi''(h)\|_{L(H)} \leqslant C_R$. Since Y_n converges to Y $\mathbb{P}$-a.s. there exists a $C_R(\omega) > 0$ for a.e. $\omega \in \Omega$, such that for all $s \in [0,T]$ and a.e. $\omega \in \Omega$

$$\|\phi'(Y_n(\omega,s) + \theta\Upsilon(\omega,s)x)\|_H + \|\phi''(Y_n(\omega,s) + \theta\Upsilon(\omega,s)x)\|_{L(H)} \leqslant C_R(\omega),$$

for all $0 \leqslant \theta \leqslant 1$. From the mean value theorem we deduce that for all $x, y \in H$

$$|\phi(x) - \phi(y)| \leqslant \sup_{0\leqslant\theta\leqslant1} \|\phi'(y + \theta(x-y))\|_H \|x-y\|_H$$

and

$$|\phi(x) - \phi(y) - \phi'(y)(x-y)| \leqslant \sup_{0\leqslant\theta\leqslant1} \|\phi'(y + \theta(x-y)) - \phi'(y)\|_H \|x-y\|_H$$

holds. Thus, for the compensated Poisson integral (3.7) we find for all $t \in [0,T]$ and a.e. $\omega \in \Omega$

$$\begin{aligned}&\int_0^t \int_A \mathbb{1}_{B_n}(x)|\phi(Y_n(\omega,s-) + \Upsilon(\omega,s)x) - \phi(Y_n(\omega,s-))|^2 \nu(dx)ds \\ &\qquad\leqslant \int_0^t \int_A C_R^2(\omega)\|\Upsilon(\omega,s)x\|_H^2 \nu(dx)ds < \infty.\end{aligned}$$

Hence, we obtain a $\mathbb{P}$-a.s. convergent subsequence from Theorem B.48. For (3.8), we find for all $t \in [0,T]$ and a.e. $\omega \in \Omega$

$$\begin{aligned}&\int_0^t \int_A \mathbb{1}_{B_n}(x)|\phi(Y_n(\omega,s-) + \Upsilon(\omega,s)x) - \phi(Y_n(\omega,s-)) - \langle\phi'(Y_n(\omega,s-)), \Upsilon(\omega,s)x\rangle_H|\nu(dx)ds \\ &\qquad\leqslant \int_0^t \int_A C_R(\omega)\|\Upsilon(\omega,s)x\|_H^2 \nu(dx)ds < \infty.\end{aligned}$$

Finally, recall the definition of the inner product appearing in (3.6), that is for $u \in U$

$$(\Phi^*(\omega,s)\phi'(Y_n(\omega,s-)))(u) = \langle\phi'(Y_n(\omega,s-)), \Phi(\omega,s)(u)\rangle_H,$$

for all $s \in [0,T]$ and a.e. $\omega \in \Omega$. Then, (3.6) is defined for all $t \in [0,T]$ and a.e. $\omega \in \Omega$ by

$$\int_0^t \langle\phi'(Y_n(\omega,s-)), \Phi(\omega,s)dW_{Q_0}(s)(\omega)\rangle_H = \int_0^t \Phi^*(\omega,s)\phi'(Y_n(\omega,s-))dW_{Q_0}(s)(\omega),$$

for more detail see [**GM11a**, Section 2.3.1]. Thus, we find

$$\begin{aligned}\|\Phi^*(\omega,s)\phi'(Y_n(\omega,s-))\|_{L_{(HS)}(U,H)}^2 &\leqslant \|\phi'(Y_n(\omega,s-))\|_H^2 \|\Phi(\omega)\|_{L_{(HS)}(U,H)}^2 \\ &\leqslant C_R^2(\omega)\|\Phi(\omega)\|_{L_{(HS)}(U,H)}^2 < \infty,\end{aligned}$$

for all $s \in [0,T]$ and a.e. $\omega \in \Omega$ and Theorem B.13 applies.

□

Since the transformation formula in Theorem 3.3 is composed of the Itô formula from Theorem B.12 and the transformation formula for pure jump noise from Theorem B.50 all functions ϕ, which fulfill the assumption of those two theorem, can also be use as transformation function for the transformation formula in Theorem 3.3. We provide three examples here.

EXAMPLE 3.4. *Let $\phi \in BC^2(H;\mathbb{R})$, that means ϕ and its first and second derivative are continuous and bounded. Furthermore, assume that Ψ, Φ, Υ and Λ are progressively measurable processes, such that $\mathbb{P}$-a.s.*

$$\int_0^t \|\Psi(s)\|_H ds + \int_0^t \int_A \|\Upsilon(s)x\|_H^2 \nu(dx)ds < \infty$$

for $t \in [0,T]$ and $\Phi \in \mathcal{P}_T^2(U,H)$. Then, Theorem 3.3 can be applied and yields the transformation formula (3.5).

EXAMPLE 3.5. *Assume that Ψ, Φ, Υ and Λ are progressively measurable processes, such that $\mathbb{P}$-a.s.*

$$\int_0^t \|\Psi(s)\|_H ds + \int_0^t \int_A \|\Upsilon(s)x\|_H^2 \nu(dx)ds < \infty$$

for $t \in [0,T]$ and $\Phi \in \mathcal{P}_T^2(U,H)$. If $\phi \in L(H;\mathbb{R})$ the transformation formula (3.5) *simplefies to*

$$\begin{aligned}\phi(Y(t)) =& \phi(Y(0)) + \int_0^t \langle \phi, \Phi(s)dW_{Q_0}(s)\rangle_H + \int_0^t \phi(\Psi(s))ds + \int_0^t \int_A \phi(\Upsilon(s)x)\tilde{N}(ds,dx) \\ &+ \int_0^t \int_A \phi(\Lambda(s)x)N(ds,dx).\end{aligned}$$

EXAMPLE 3.6. *One of the most important examples for the transformation function is the norm-square, that is $\phi(h) = \|h\|_H^2$. In this case we have $\phi \in C^2(H;\mathbb{R})$ with the following derivates*

$$\phi'(h)v = 2\langle h, v\rangle_H \text{ and } \phi''(h)(v,w) = 2\langle v, w\rangle_H,$$

where $v, w \in H$.Thus, we find the following estimates

$$\|\phi'(h)\|_H \leqslant 2\|h\|_H \text{ and } \|\phi''(h)\|_{L(H)} \leqslant 2.$$

Additionally, if we assume that Ψ, Φ, Υ and Λ are progressively measurable processes, such that

$$\int_0^t \|\Psi(s)\|_H ds + \int_0^t \int_A \|\Upsilon(s)x\|_H^2 \nu(dx)ds + \int_0^t \int_A \|\Upsilon(s)x\|_H^4 \nu(dx)ds < \infty$$

for $t \in [0,T]$ and $\Phi \in \mathcal{P}_T^2(U,H)$, then Theorem 3.3 yields the following transformation formula

$$\begin{aligned}\|Y(s)\|_H^2 =& \|Y(0)\|_H^2 + \int_0^t \langle Y(s-), \Phi(s)dW_{Q_0}(s)\rangle_H \\ &+ \int_0^t \{\langle Y(s-), \Psi(s)\rangle_H + \operatorname{tr}[(\Phi(s)Q_0^{1/2})(\Phi(s)Q_0^{1/2})^*]\}ds \\ &+ \int_0^t \int_A \{\|Y(s-) + \Upsilon(s)x\|_H^2 - \|Y(s-)\|_H^2\}\tilde{N}(ds,dx) \\ &+ \int_0^t \int_A \{\|Y(s-) + \Upsilon(s)x\|_H^2 - \|Y(s-)\|_H^2 - \langle Y(s-), \Upsilon(s)x\rangle_H\}\nu(dx)ds \\ &+ \int_0^t \int_A \{\|Y(s-) + \Lambda(s)x\|_H^2 - \|Y(s-)\|_H^2\}N(ds,dx).\end{aligned}$$

REMARK 3.7. *Note that if the transformation function ϕ also depends on time $t \in \mathbb{R}_+$, that is $\phi : \mathbb{R}_+ \times H \to \mathbb{R},\ (t,h) \mapsto \phi(t,h)$, then we can still formulate the transformation formula.*

In this case we have to assume in Theorem 3.3 that $\phi \in C^{1,2}(\mathbb{R}_+ \times H;\mathbb{R})$, $\dot{\phi}$ is bounded on bounded subsets, and

$$\|\phi'(t,h)\|_H \leqslant h_1(\|h\|_H), \ \|\phi''(t,h)\|_{L(H)} \leqslant h_2(\|h\|_H),$$

where $h \in H$, $t \in \mathbb{R}_+$ and $h_1, h_2 : \mathbb{R}_+ \to \mathbb{R}_+$ are quasi-sublinear functions. Then we receive the following transformation formula

$$\begin{aligned}
\phi(Y(t)) &= \phi(Y(0)) + \int_0^t \langle \phi'(Y(s-)), \Phi(s) dW_{Q_0}(s)\rangle_H + \int_0^t \dot{\phi}(s, Y(s-))ds \\
&+ \int_0^t \{\langle \phi'(Y(s-)), \Psi(s)\rangle_H + \frac{1}{2}\operatorname{tr}[\phi''(Y(s-))(\Phi(s)Q_0^{1/2})(\Phi(s)Q_0^{1/2})^*]\}ds \\
&+ \int_0^t \int_A \{\phi(Y(s-) + \Upsilon(s)x) - \phi(Y(s-))\}\tilde{N}(ds,dx) \\
&+ \int_0^t \int_A \{\phi(Y(s-) + \Upsilon(s)x) - \phi(Y(s-)) - \langle \phi'(Y(s-)), \Upsilon(s)x\rangle_H\}\nu(dx)ds \\
&+ \int_0^t \int_A \{\phi(Y(s-) + \Lambda(s)x) - \phi(Y(s-))\}N(ds,dx),
\end{aligned}$$

$\mathbb{P}$-a.s. and for all $t \in [0,T]$. For the sake of clarity and brevity of the thesis, we decided to omit the time dependency. To prove the transformation formula given above we could proceed in two ways. The first would be to include the time dependence from the beginning. This is possible since the Itô formula in Theorem B.12 and the transformation formula in Theorem B.50 both include the direct time dependency of the transformation function ϕ. Since the proofs of Theorem 3.3 and the preliminarily results are deduced from those two theorems, we could carry the time dependence through each step of the proofs. The second way to prove the transformation formula above is to apply Krylovs trick [**Kry80**, Appendix 1]. *Thus, we could define a new stochastic process Z in the Hilbert space $\mathbb{R} \times H$ as $Z(t) := (\int_0^t ds, Y(t))$ and apply Theorem 3.3.*

3.2. The space of transformation function

In this section, we provide a solution for the second problem described in the introduction of the chapter. As we have seen, we can overcome the lack of stochastic differentials of the mild solution Y by applying the Yosida approximation from Theorem 2.3. Since each member of the approximating sequence is a well-defined Lévy process, we can apply the transformation formula from Theorem 3.3. However, if we do so, we still encounter one problem when passing to the limit. This is due to the term

$$\langle AY_n(t-), \phi'(Y_n(t-))\rangle_H,$$

which appears if we apply the transformation formula to Y_n. Since the process Y in not taking values in domain of A, we cannot take the limit. The only way to overcome this problem is to find condition for ϕ such that $\phi'(h) \in D(A^*)$ for all $h \in H$. Then, we can remove the operator A from the solution process Y_n and apply A^* to $\phi'(Y(t-))$ instead. In the following we provide a sufficient criterion for that.

Such as, always let H be a separable Hilbert space and $\mathcal{M} \subset H$. By $BUC^k(\mathcal{M}, \mathbb{R}),\ k = 0, 1, \ldots$ we denote all mappings ϕ from $\mathcal{M}$ to $\mathbb{R}$, which are bounded and uniformly continuous for all Fréchet derivatives $\phi, \phi', \phi'', \ldots$ up to order k. Let the operator A generate a C_0-semigroup $(T(t))_{t \geqslant 0}$ with $\|T(t)\|_H \leqslant Me^{\lambda t}$, where $M \geqslant 1$ and $\lambda \in \mathbb{R}$. If $\lambda \geqslant 0$, we can choose a $\lambda_0 > \lambda$, such that the resolvent operator $(\lambda_0 - A)^{-1}$ is well-defined. If $\lambda < 0$ we set $\lambda_0 = 0$. For $\phi \in C(H; \mathbb{R})$, with $\phi \in BUC(\mathcal{M}; \mathbb{R})$ for each bounded subset $\mathcal{M}$ of H we define the function ϕ_A analoge to [**AFZ97**]. That is, if ϕ is such that the mapping

$$x \mapsto \phi((\lambda_0 - A)x),\ x \in \mathscr{D}(A)$$

has a continuous extension to all of H, then the extension, which is unique, will be denoted by ϕ_A.

LEMMA 3.8. *Let $\phi, \phi_A \in BUC(\mathcal{M}; \mathbb{R})$ and $\phi', \phi'_A \in BUC(\mathcal{M}; L(H, \mathbb{R}))$ for each bounded subset $\mathcal{M}$ of H. Then, $\phi'(h) \in \mathscr{D}(A^*)$ for all $h \in H$ (after the identification of $\phi'(h)$ as an element of H via the Riesz representation theorem). Furthermore $A^*\phi'(h) = \phi'_A(x) + \lambda_0\phi'(h)$ holds, where $x = (\lambda_0 - A)^{-1}h$. Therefore, the map $h \mapsto A^*\phi'(h)$ is in $BUC(\mathcal{M}; H)$ for all bounded subsets $\mathcal{M}$ of H.*

PROOF.

First, we consider the function ϕ'_A

$$\begin{aligned} \phi'_A : H &\to L(H, \mathbb{R}) \\ h &\mapsto \phi'_A(h)\bullet = \langle \bullet, z_h \rangle_H, \end{aligned}$$

where $z_h \in H$ is the unique element representing the linear functional $\phi'_A(h)$ from the Riesz representation theorem. Since the operator $(\lambda_0 - A)$ is invertible, we find

$$\forall h \in H\ \exists x \in \mathscr{D}(A) : x = (\lambda_0 - A)^{-1}h.$$

Thus, we have

$$\phi(h) = \phi((\lambda_0 - A)(\lambda_0 - A)^{-1}h) = \phi((\lambda_0 - A)x) = \phi_A(x).$$

Taking the derivative of ϕ_A with respect to x ($\in \mathscr{D}(A)$) we obtain the following identity for all $w \in \mathscr{D}(A)$

$$\phi'_A(x)w = \phi'((\lambda_0 - A)x)(\lambda_0 - A)w = \phi'(h)(\lambda_0 - A)w = \langle (\lambda_0 - A)w, v_h \rangle_H,$$

where $v_h \in H$. Again the (unique) existence of v_h follows from the Riesz representation theorem, but this time applied to the linear functional $\phi'(h)$. Summing up, we found that for all $h \in H$ and all $w \in \mathscr{D}(A)$

$$\langle (\lambda_0 - A)w, v_h \rangle_H = \phi'_A(x)w = \langle w, z_x \rangle_H.$$

Hence, we find that $v_h \in \mathscr{D}((\lambda_0 - A)^*) = \mathscr{D}(A^*)$. Since v_h is the element representing $\phi'(h)$ we have $\phi'(h) \in \mathscr{D}(A^*)$. Furthermore, we have $A^*v_h = z_x + \lambda_0 v_h$ and by substituting we find $A^*\phi'(h) = \phi'_A(x) + \lambda_0\phi'(h)$.

□

3.3. Transformation formula for mild solutions

In this section, we prove the main theorem of the chapter. However, before we do so, we provide some auxiliary results which will provide the key tools to prove the transformation formula for mild solutions. As before, we will use the a.e. ω–notation at the points we want to emphasize that these properties hold only $\mathbb{P}$-a.s. and not uniformly. Recall that the space $\mathcal{X}_{T,B}$ of all predictable process with values in a Banach space B is equipped with the norm

$$\|X\|_{T,B}^2 = \sup_{t\in[0,T]} \mathbb{E}\|X(t)\|_B^2.$$

PROPOSITION 3.9. *Let ϕ be a mapping from a Hilbert space H into a Banach space B. Furthermore, let the stochastic processes $X, X_n \in \mathcal{X}_{T,H}$ be continuous in probability for each $n \in \mathbb{N}$, such that $\|X - X_n\|_{T,H} \to 0$ for $n \to \infty$. If $\phi \in BUC(H, B)$, then we have*

$$\|\phi(X_n) - \phi(X)\|_{T,B} \to 0 \textit{ for } n \to \infty.$$

PROOF.

Before we show the claimed convergence, we prove the following auxiliary result. For an arbitrary fixed $\delta > 0$ we define the set

$$\Omega_n := \{\omega \in \Omega | \forall t \in [0,T] : \|X_n(t,\omega) - X(t,\omega)\|_H^2 \leqslant \delta\}.$$

Obviously, the complement of Ω_n is the set

$$\Omega_n^c = \{\omega \in \Omega | \exists t \in [0,T] : \|X_n(t,\omega) - X(t,\omega)\|_H^2 > \delta\}.$$

We want to show that $\mathbb{P}(\Omega_n^c) \to 0$ as $n \to \infty$. Thus, for an arbitrary $\varepsilon > 0$ we choose an $N_0 \in \mathbb{N}$, such that $\|X_n - X\|_{T,H} \leqslant \delta\varepsilon$ for all $n \geqslant N_0$ and estimate

$$\delta\varepsilon \geqslant \sup_{t\in[0,T]} \int_\Omega \|X_n(t) - X(t)\|_H^2 d\mathbb{P} \geqslant \sup_{t\in[0,T]} \int_{\Omega_n^c} \|X_n(t) - X(t)\|_H^2 d\mathbb{P} \geqslant \delta \int_{\Omega_n^c} d\mathbb{P} = \delta\ \mathbb{P}(\Omega_n^c).$$

Since $\delta > 0$, we found that $\mathbb{P}(\Omega_n^c) \leqslant \varepsilon$ for all $n \geqslant N_0$.
Now we turn to the claimed statement and choose an arbitrary $\varepsilon > 0$. Since ϕ is uniformly continuous on H, there exists an $\delta > 0$, such that

$$\forall X, Y \in H : \|X - Y\|_H^2 \leqslant \delta \Longrightarrow \|\phi(X) - \phi(Y)\|_B^2 \leqslant \frac{\varepsilon}{2}.$$

Therefore, we estimate

$$\|\phi(X_n(t)) - \phi(X(t))\|_{T,B}^2 = \sup_{t\in[0,T]} \int_\Omega \|\phi(X_n(t)) - \phi(X(t))\|_B^2 d\mathbb{P}$$

$$\leqslant \sup_{t\in[0,T]} \int_{\Omega_n} \|\phi(X_n(t)) - \phi(X(t))\|_B^2 d\mathbb{P} + \sup_{t\in[0,T]} \int_{\Omega_n^c} \|\phi(X_n(t)) - \phi(X(t))\|_B^2 d\mathbb{P} \leqslant \frac{\varepsilon}{2} + C\ \mathbb{P}(\Omega_n^c),$$

where $C > 0$ is a constant due to the boundedness of ϕ. Since $\mathbb{P}(\Omega_n^c)$ converges to 0, we choose an $N_0 \in \mathbb{N}$, such that $\mathbb{P}(\Omega_n^c) \leqslant \varepsilon/2C$ for all $n \geqslant N_0$ and receive

$$\|\phi(X_n(t)) - \phi(X(t))\|_{T,B}^2 \leqslant \varepsilon \text{ for all } n \geqslant N_0.$$

□

The next result shows that from the norm convergence follows the $\mathbb{P}$-a.s. convergence of a subsequence uniformly in $t \in [0,T]$. This will be of great importance later in the proof of the transformation formula for mild solutions.

COROLLARY 3.10. *Let H be a Hilbert space. Furthermore, assume that the stochastic processes $X, X_n \in \mathcal{X}_{T,H}$ are continuous in probability for each $n \in \mathbb{N}$ and $\|X - X_n\|_{T,H} \to 0$ for $n \to \infty$. Then, there exists a subsequence $(X_{n_k})_{k\in\mathbb{N}}$ of $(X_n)_{n\in\mathbb{N}}$, such that $(X_{n_k}(t))_{k\in\mathbb{N}}$ converges $\mathbb{P}$-a.s. and uniformly in t to $X(t)$.*

PROOF.

From the norm convergence in $\mathcal{X}_{T,H}$ we find that for all $t \in [0,T]$

$$\mathbb{E}\|X_n(t) - X(t)\|_H^2 \to 0.$$

Thus, for all $t \in [0,T]$ the sequence $(X_n(t))_{n\in\mathbb{N}}$ converges to $X(t)$ in mean square. Therefore, there exists a subsequence $(X_{n_k}(t))_{k\in\mathbb{N}}$ that converges for all $t \in [0,T]$ $\mathbb{P}$-a.s. to $X(t)$. For the sake of readability we denote this subsequence again by $(X_n(t))_{n\in\mathbb{N}}$. In the proof of Proposition 3.9 it was show that for arbitary $\delta > 0$ the probability of the set

$$\Omega_n^c = \{\omega \in \Omega | \exists t \in [0,T] : \|X_n(t,\omega) - X(t,\omega)\|_H^2 > \delta\}$$

converges to zero for $n \to \infty$. We will show that the set Ω_n^c is equal to the set

$$A_n := \{\omega \in \Omega | \sup_{t\in[0,T]} \|X_n(t,\omega) - X(t,\omega)\|_H^2 > \delta\},$$

for all $n \in \mathbb{N}$. In order to show the equality recall the formal definition of the supremum for an nonempty set $S \subset \mathbb{R}$

$$\sup S = M \iff (a)\ \forall s \in S : s \leqslant M$$
$$(b)\ \forall \varepsilon > 0\ \exists s \in S : s > M - \varepsilon.$$

Since we only have to consider the t-dependence in the argument, we introduce the function $f_\omega^n(t) := \|X_n(t,\omega) - X(t,\omega)\|_H^2$. First we show $\Omega_n^c \subset A_n$, which is straightforward

$$\exists \tilde{t} \in [0,T] : f_\omega^n(\tilde{t}) > \delta \Longrightarrow \sup_{t\in[0,T]} f_\omega^n(t) \overset{(a)}{\geqslant} f_\omega^n(\tilde{t}) > \delta.$$

For the reversed inclusion it follows from (b) that

$$\forall \varepsilon > 0\ \exists \tilde{t}_\omega^n \in [0,T] : f_\omega^n(\tilde{t}_\omega^n) > \sup_{t\in[0,T]} f_\omega^n(t) - \varepsilon.$$

Since $\sup_{t\in[0,T]} f_\omega^n(t) > \delta\ \exists \varepsilon_\omega^n > 0 : \sup_{t\in[0,T]} f_\omega^n(t) - \varepsilon_\omega^n > \delta$. Thus, we find

$$f_\omega^n(\tilde{t}_\omega^n) > \sup_{t\in[0,T]} f_\omega^n(t) - \varepsilon_\omega^n > \delta.$$

This shows that also $A_n \subset \Omega_n^c$ holds. As a final step we introduce the real nonnegative random variables Z_n, $n \in \mathbb{N}$ by

$$Z_n := \sup_{t\in[0,T]} \|X_n(t) - X(t)\|_H^2.$$

Therefore, we can rewrite A_n as

$$A_n = \{\omega \in \Omega \big| |Z_n(\omega)| > \delta\}.$$

Since $A_n = \Omega_n^c$, it follow that

$$\lim_{n\to\infty} \mathbb{P}(\{\omega \in \Omega \big| |Z_n(\omega)| > \delta\}) = 0.$$

Thus, $(Z_n)_{n\in\mathbb{N}}$ converges in probability to zero. Hence, we can extract a subsequence $(Z_{n_k})_{k\in\mathbb{N}}$, which converges $\mathbb{P}$-a.s. to zero. By the definition of Z we have shown that $(X_{n_k}(t))_{k\in\mathbb{N}}$ converges $\mathbb{P}$-a.s. and uniformly in t to $X(t)$.

□

For the sequence $(X_{n_k})_{k\in\mathbb{N}}$ from Corollary 3.10 we deduce that for all $t \in [0,T]$ and a.e. $\omega \in \Omega$ there exists a random constant $C_X(\omega) < \infty$, such that

$$\|X_{n_k}(t,\omega)\|_H, \|X(t,\omega)\|_H \leqslant C_X(\omega) \text{ for all } k \in \mathbb{N}.$$

One can formulate equivalently, that there exists a $\mathbb{P}$-a.s. bounded random subset $\mathcal{M}$ in H, such that for all $t \in [0,T]$ and a.e. $\omega \in \Omega : X_{n_k}(t,\omega), X(t,\omega) \in \mathcal{M}(\omega)$ for all $k \in \mathbb{N}$. This fact will help to prove the next convergence result, which will be used several times in the proof of the main theorem.

PROPOSITION 3.11. *Let ϕ be a mapping from a Hilbert space H into a Banach space B, which is uniformly continuous on bounded subset of H. Furthermore, let the stochastic processes $X, X_n \in \mathcal{X}_{T,H}$ be continuous in probability for each $n \in \mathbb{N}$, such that $\|X - X_n\|_{T,H} \to 0$ for $n \to \infty$. Then, there exists a subsequence $(X_{n_k})_{k\in\mathbb{N}}$ of $(X_n)_{n\in\mathbb{N}}$, such that*

$$\phi(X_{n_k}(t)) \to \phi(X(t)) \text{ } \mathbb{P}\text{-a.s. and uniformly in } t \text{ for } k \to \infty.$$

PROOF.

We apply Corollary 3.10 in order to receive a subsequence $(X_{n_k}(t))_{k\in\mathbb{N}}$ that converges $\mathbb{P}$-a.s. and uniformly in t to $X(t)$, this means

$$\forall t \in [0,T] \ \mathbb{P}(\{\omega \in \Omega | \forall \delta > 0 \ \exists N_0(\omega,\delta) \in \mathbb{N} \ \forall k \geqslant N_0 : \|X_{n_k}(t,\omega) - X(t,\omega)\|_H < \delta\}) = 1.$$

In particular there exists a rondom subset $\mathcal{M} \subset H$ with $\mathbb{P}(\{\omega \in \Omega | \ \mathcal{M}(\omega) \text{ is bounded}\}) = 1$, such that for all $t \in [0,T]$ and a.e. $\omega \in \Omega$: $X_{n_k}(t,\omega), X(t,\omega) \in \mathcal{M}(\omega)$ for all $k \in \mathbb{N}$.

Now let $\varepsilon > 0$. Since ϕ is uniformly continuous on bounded subset of H, there

$$\exists \delta(\varepsilon, \mathcal{M}(\omega)) > 0 \ \forall X, Y \in \mathcal{M}(\omega) : \|X - Y\|_H < \delta(\varepsilon, \mathcal{M}(\omega)) \Longrightarrow \|\phi(X) - \phi(Y)\|_B < \varepsilon.$$

Since the sequence $(X_{n_k}(t))_{k\in\mathbb{N}}$ converges $\mathbb{P}$-a.s. and uniformly in t, it follows that for every $\delta(\varepsilon, \mathcal{M}(\omega)) > 0$ there exists a $N_0(\omega, \delta(\varepsilon, \mathcal{M}(\omega))) \in \mathbb{N}$, such that for all $t \in [0,T]$

$$\mathbb{P}(\{\omega \in \Omega | \forall \varepsilon > 0 \ \exists N_0 \in \mathbb{N} \ \forall k \geqslant N_0 : \|\phi(X_{n_k}(t,\omega)) - \phi(X(t,\omega))\|_H < \varepsilon\}) = 1.$$

□

REMARK 3.12. *In our case of interest the stochastic process is the mild solution Y of the stochastic abstract Cauchy problem $(S-ACP)$. If Y has a càdlàg modification, we already pointed out in Remark 2.1 that the process of the left limits $\tilde{Y}$, that is $\tilde{Y}(t) := Y(t-)$ for $t \in [0,T]$, is a modification of Y. Therefore, if the process Y has a càdlàg modification, Corollary 3.10 and Proposition 3.11 hold also for the process of the left limits $\tilde{Y}$.*

The last result we require to prove the transformation formula for mild solutions, is a direct consequence of Proposition 2.2.

COROLLARY 3.13. *Let Y be the mild solution of $(S-ACP)$ and $R(n)$ the Yosida approximation from Definition A.20. If G fulfills (SH_{11}), then there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ of $(n)_{n\in\mathbb{N}}$, such that*

$$\|(I-R(n_k))G(t,Y(t))Q^{1/2}\|_{L_{(HS)}(U,H)} \longrightarrow 0 \ dt \otimes \mathbb{P}\text{-a.s. for } k \to \infty.$$

PROOF.
In Proposition 2.2 we set $T(t) = I$ for all $t \in [0,T]$ and receive

$$\sup_{t\in[0,T]} \mathbb{E}\left\| \int_0^t (I-R(n))G(s,Y(s))dM(s)\right\|_H^2 \longrightarrow 0 \text{ for } n \to \infty.$$

Applying the isometry of the stochastic integral yields

$$\sup_{t\in[0,T]} \mathbb{E}\int_0^t \|(I-R(n))G(s,Y(s))Q^{1/2}\|^2_{L_{(HS)}(U,H)}ds \longrightarrow 0 \text{ for } n \to \infty.$$

Thus, the sequence converges in L_2. This means, we can extract a subsequence with converges $dt \otimes \mathbb{P}$-a.s. to zero.

□

REMARK 3.14. *Note that by the same argument as in Remark 3.12, Corollary 3.13 holds also for the process of the left limits $\tilde{Y}$, if Y has a càdlàg modification.*

Now we move on to the main result of this Chapter. Recall that we consider the stochastic abstract Cauchy Problem $(S-ACP)$ in a separable Hilbert space H given by

$$(S-ACP)\begin{cases} dY(t) &= AY(t)dt + F(t,Y(t))dt + G(t,Y(t))dM(t), \quad t \geqslant 0, \\ Y(0) &= y, \end{cases}$$

where A generates a C_0-semigroup, $y \in H$, M fulfills (SH_9), and F fulfills (SH_{10}). In order to prove a transformation formula including stochastic jumps we switch the notation to Poisson integrals for the jumps of the Lévy process M. Therefore the stochastic term in $(S-ACP)$ takes the following form

$$G(t,Y(t))dM(t) = G_0(t,Y(t))dW_{Q_0}(t) + \int_U G_1(t,Y(t))x\tilde{N}(t,dx),$$

where G_0 fulfills (SH_{11}) with Q replaced by Q_0, which is the covariance operator of the continuous part of the Lévy process M. G_1 fulfills (SH_{11}) with covariance operator of the jump part Q_1 instead of Q. Note that by Remark B.23 it is justified to write U instead of

$U\backslash\{0\}$ in the compensated Poisson integral. Hence, the stochastic abstract Cauchy Problem $(S-ACP)$ takes the following form

$$(S-ACP)\begin{cases} dY(t) &= AY(t)dt + F(t,Y(t))dt + G_0(t,Y(t))dW_{Q_0}(t) \\ & \quad +\int_U G_1(t,Y(t))x\tilde{N}(t,dx), \quad t \geqslant 0, \\ Y(0) &= y. \end{cases}$$

We will use this to prove the following theorem, where we consider the càdlàg modification of the mild solution.

THEOREM 3.15 (transformation formula for mild solution of $(S-ACP)$). *Let Y be the mild solution of $(S-ACP)$, where the operator A is γ-dissipative and G_1 fulfills $\mathbb{P}$-a.s.*

$$\|G_1(t,h)\|_{L(U,H)} \leqslant k_{G_1}(1+\|h\|_H)$$

for a deterministic constant $k_{G_1} > 0$ and for all $t \in [0,T]$ and all $h \in H$. Furthermore, let $\phi \in C^2(H;R)$, where $\phi, \phi_A, \phi', \phi'_A, \phi''$ are bounded and uniformly continuous on bounded subsets of H. Additionally, let

$$\|\phi'(h)\|_H \leqslant h_1(\|h\|), \ \ \|\phi''(h)\|_{L(H)} \leqslant h_2(\|h\|),$$

where $h \in H$ and $h_1, h_2 : \mathbb{R}_+ \to \mathbb{R}_+$ are quasi-sublinear functions and we have

$$\int_U h_1(\|x\|)^2\|x\|^2\nu(dx) < \infty, \ \ \int_U h_2(\|x\|)\|x\|^2\nu(dx) < \infty.$$

Then, the following transformation formula holds for all $t \in [0,T]$ $\mathbb{P}$-a.s.

$$\begin{aligned}\phi(Y(t)) = \ & \phi(y) + \int_0^t \langle A^*\phi'(Y(s-)), Y(s-)\rangle_H ds + \int_0^t \langle \phi'(Y(s-)), F(s,Y(s-))\rangle_H ds \\ &+ \int_0^t \langle \phi'(Y(s-)), G_0(s,Y(s-))dW_{Q_0}(s)\rangle_H \\ &+ \frac{1}{2}\int_0^t \operatorname{tr}\left[\phi''(Y(s-))\big(G_0(s,Y(s-))Q_0^{1/2}\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)^*\right]ds \\ &+ \int_0^t\int_U \phi(Y(s-)+G_1(s,Y(s-))x) - \phi(Y(s-))\tilde{N}(ds,dx) \\ &+ \int_0^t\int_U \phi(Y(s-)+G_1(s,Y(s-))x) - \phi(Y(s-)) - \langle \phi'(Y(s-)), G_1(s,Y(s-))x\rangle_H \nu(dx)ds.\end{aligned}$$

PROOF.

Since the generator A is γ-dissipative, it follows by the Lumer–Phillips theorem that it generates a generalized contraction semigroup. Therefore, the mild solution Y of $(S-ACP)$ has a càdlàg modification and Theorem 2.3 shows, that Y can be approximated in $\mathcal{X}_{T,H}$ by a sequence of smoother processes $(Y_n)_{n>\gamma} \in \mathscr{D}(A)$ of the form

$$Y_n(t) = R(n)y + \int_0^t AY_n(s-)ds + \int_0^t R(n)F(s,Y_n(s-))ds + \int_0^t R(n)G(s,Y_n(s-))dM(s).$$

From Corollary 3.10 follows, that there exists a subsequence $(Y_{n_k}(t))_{k\in\mathbb{N}}$ of $(Y_n(t))_{n>\gamma}$ which converges $\mathbb{P}$-a.s. and uniformly in t to $Y(t)$. By Remark 3.12, the same holds true for the process of the left limits. Like before, we denote this subsequence by $(Y_n(t))_{n>\gamma}$. For each $n \in \mathbb{N}$ with $n > \gamma$ we can apply the transformation formula from Theorem 3.3 and receive for all $t \in [0,T]$ $\mathbb{P}$-a.s.

$$
\begin{aligned}
\phi(Y_n(t)) = \ & \phi(R(n)y) && (3.9)\\
& + \int_0^t \langle \phi'(Y_n(s-)), AY_n(s-)\rangle_H ds && (3.10)\\
& + \int_0^t \langle \phi'(Y_n(s-)), R(n)F(s,Y_n(s-))\rangle_H ds && (3.11)\\
& + \int_0^t \langle \phi'(Y_n(s-)), R(n)G_0(s,Y_n(s-))dW_{Q_0}(s)\rangle_H && (3.12)\\
& + \frac{1}{2}\int_0^t \operatorname{tr}\big[\phi''(Y_n(s-))\big(R(n)G_0(s,Y_n(s-))Q_0^{1/2}\big)\big(R(n)G_0(s,Y_n(s-))Q_0^{1/2}\big)^*\big]ds && (3.13)\\
& + \int_0^t\int_U \phi(Y_n(s-)+R(n)G_1(s,Y_n(s-))x) - \phi(Y_n(s-))\tilde{N}(ds,dx) && (3.14)\\
& + \int_0^t\int_U \phi(Y_n(s-)+R(n)G_1(s,Y_n(s-))x) - \phi(Y_n(s-)) && (3.15)\\
& \qquad - \langle \phi'(Y_n(s-)), R(n)G_1(s,Y_n(s-))x\rangle_H \nu(dx)ds.
\end{aligned}
$$

The goal now is to show, that the left and the right-hand side converge $\mathbb{P}$-a.s. to their counterparts as claimed in the statement of the Theorem. Recall from Corollary 3.10, that for all $t \in [0,T]$: $Y_n(t), Y_n(t-), Y(t), Y(t-) \in \mathcal{M}$, where $\mathcal{M}$ is a $\mathbb{P}$-a.s. bounded subset of H independent of t. Note that the two terms in (3.9) converge by Proposition 3.11. Therefore, we have, possibly for a subsequence, for all $t \in [0,T]$ and $\mathbb{P}$-a.s.

$$\phi(Y_n(t)) \to \phi(Y(t)) \text{ and } \phi(R(n)y) \to \phi(y).$$

From Lemma 3.8 it follows that $\phi'(h) \in \mathscr{D}(A^*)$ for all $h \in H$. Hence, we can move the unbounded operator A in (3.10) away from Y_n and onto ϕ', that is $\mathbb{P}$-a.s.

$$\langle \phi'(Y_n(s-)), AY_n(s-)\rangle_H = \langle A^*\phi'(Y_n(s-)), Y_n(s-)\rangle_H,$$

for all $n > \gamma$ and $s \in [0,T]$. Thus we can estimate $\mathbb{P}$-a.s.

$$
\begin{aligned}
&\left|\int_0^t \langle A^*\phi'(Y_n(s-)), Y_n(s-)\rangle_H - \langle A^*\phi'(Y(s-)), Y(s-)\rangle_H ds\right|\\
&= \left|\int_0^t \langle A^*\big(\phi'(Y_n(s-)) - \phi'(Y(s-))\big), Y(s-)\rangle_H + \langle A^*\phi'(Y_n(s-)), Y_n(s-) - Y(s-)\rangle_H ds\right|\\
&\leqslant \int_0^t \underbrace{\|A^*\big(\phi'(Y_n(s-)) - \phi'(Y(s-))\big)\|_H \|Y(s-)\|_H}_{=:(I_A)}\\
&\qquad\qquad + \underbrace{\|A^*\phi'(Y_n(s-))\|_H \|Y_n(s-) - Y(s-)\|_H}_{:=(II_A)} ds.
\end{aligned}
$$

First we consider (I_A). Note that for every fixed $s \in [0,T]$ we have $\mathbb{P}$-a.s.

$$\|A^*\big(\phi'(Y_n(s-)) - \phi'(Y(s-))\big)\|_H \|Y(s-)\|_H \to 0 \text{ for } n \to \infty,$$

due to Lemma 3.8 and Proposition 3.11. Lemma 3.8 also states that $A^*\phi'(\cdot) \in BUC(\mathcal{M};H)$ for all bounded subsets $\mathcal{M}$ of H. Hence, we estimate

$$\|A^*\big(\phi'(Y_n(s-)) - \phi'(Y(s-))\big)\|_H \leqslant C_{A^*} \ \mathbb{P}\text{-a.s.},$$

for some positive $\mathbb{P}$-a.s. finite constant C_{A^*}. Since Y is the mild solution of $(S-ACP)$, it fulfills

$$\int_0^t \|Y(s-)\|_H^2 ds < \infty \ \mathbb{P}\text{-a.s. for all } t > 0,$$

from which we deduce

$$\int_0^t \|Y(s-)\|_H ds \leqslant \left(T \int_0^t \|Y(s-)\|_H^2 ds\right)^{1/2} < \infty \ \mathbb{P}\text{-a.s.},$$

where we used Hölder's inequality. Thus, we find by applying Lebesgue's dominated convergence theorem

$$\int_0^t \|A^*\big(\phi'(Y_n(s-)) - \phi'(Y(s-))\big)\|_H \|Y(s-)\|_H ds \to 0 \ \mathbb{P}\text{-a.s. for } n \to \infty.$$

For the term (II_A) we argue similarly and estimate the first term by a $\mathbb{P}$-a.s. positive, finite constant C, that is $\mathbb{P}$-a.s.

$$\int_0^t \|A^*\phi'(Y_n(s-))\|_H \|Y_n(s-) - Y(s-)\|_H ds \leqslant C_{A^*} \int_0^t \|Y_n(s-) - Y(s-)\|_H ds.$$

Since $(Y_n(t-))_{n>\gamma}$ converges $\mathbb{P}$-a.s. and uniformly in t to $Y(t-)$, it follows that also the term (II_A) converges $\mathbb{P}$-a.s. to zero.

In order to show that (3.11) converges $\mathbb{P}$-a.s., we argue like we did for the pervious term. We start by estimating $\mathbb{P}$-a.s.

$$\begin{aligned}
&\left|\int_0^t \langle \phi'(Y_n(s-)), R(n)F(s, Y_n(s-))\rangle_H - \langle \phi'(Y(s-)), F(s, Y(s-))\rangle_H ds\right| \\
&\leqslant \int_0^t \underbrace{\|\phi'(Y_n(s-)) - \phi'(Y(s-))\|_H \|F(s, Y(s-))\|_H}_{=:(I_F)} \\
&\qquad\qquad + \underbrace{\|\phi'(Y_n(s-))\|_H \|R(n)F(s, Y_n(s-)) - F(s, Y(s-))\|_H}_{=:(II_F)} ds.
\end{aligned}$$

For (I_F) we find that for fixed $s \in [0,T]$ $\mathbb{P}$-a.s.

$$\|\phi'(Y_n(s-)) - \phi'(Y(s-))\|_H \|F(s, Y(s-))\|_H \to 0 \text{ for } n \to \infty,$$

due to Proposition 3.11. Since ϕ' is bounded on bounded subsets and F is linear bounded, we can estimate $\mathbb{P}$-a.s.

$$\int_0^t \|\phi'(Y_n(s-)) - \phi'(Y(s-))\|_H \|F(s, Y(s-))\|_H \leqslant C_{\phi'} k_F \int_0^t 1 + \|Y(s-)\|_H ds < \infty.$$

Thus, by Lebesgue's dominated convergence theorem it follows that (I_F) converges $\mathbb{P}$-a.s. to zero. For (II_F) we argue as we did in the proof of Theorem 2.3. The first term in (II_F) we estimate by a $\mathbb{P}$-a.s. positive, finite constant $C_{\phi'}$, that is $\mathbb{P}$-a.s.

$$\int_0^t \|\phi'(Y_n(s-))\|_H \|R(n)F(s, Y_n(s-)) - F(s, Y(s-))\|_H ds \\ \leqslant C_{\phi'} \int_0^t \|R(n)F(s, Y_n(s-)) - F(s, Y(s-))\|_H ds.$$

Therefore, it is sufficient to show that the integral on the right-hand side converges $\mathbb{P}$-a.s. to zero. In order to achieve this, we estimate $\mathbb{P}$-a.s.

$$\begin{aligned} &\int_0^t \|R(n)F(s, Y_n(s-)) - F(s, Y(s-))\|_H \\ &\leqslant \int_0^t \|(I - R(n))F(s, Y(s-))\|_H + \|R(n)(F(s, Y(s-)) - F(s, Y_n(s-)))\|_H ds \\ &\leqslant \int_0^t \underbrace{\|(I - R(n))F(s, Y(s-))\|_H}_{=:(III_F)} + C_R \underbrace{\|F(s, Y(s-)) - F(s, Y_n(s-))\|_H}_{=:(IV_F)} ds. \end{aligned}$$

For the term (III_F) we apply Lebesgue's dominated convergence theorem, since

$$\begin{aligned} \int_0^t \|(I - R(n))F(s, Y(s-))\|_H ds &\leqslant (1 + C_R) \int_0^t \|F(s, Y(s-))\|_H ds \\ &\leqslant (1 + C_R) k_F \int_0^t 1 + \|Y(s-)\|_H ds < \infty \ \mathbb{P}\text{-a.s.}, \end{aligned}$$

where we used the linear boundedness of F. Furthermore, for fixed $s \in [0, T]$ we have

$$\|(I - R(n))F(s, Y(s-))\|_H \to 0 \text{ for } n \to \infty \ \mathbb{P}\text{-a.s.},$$

by Corollary A.21 (ii). Finally, for (IV_F) we use the Lipschitz continuity of F, that is $\mathbb{P}$-a.s.

$$\int_0^t \|F(s, Y(s-)) - F(s, Y_n(s-))\|_H ds \leqslant C_F \int_0^t \|Y(s-) - Y_n(s-)\|_H.$$

Since $(Y_n(t-))_{n>\gamma}$ converges $\mathbb{P}$-a.s. and uniformly in t to $Y(t-)$, it follows that (IV_F) converges $\mathbb{P}$-a.s. to zero. Summing up, we have show that (II_F) converges $\mathbb{P}$-a.s. to zero.

We continue with the term (3.12). Due to the stochastic integral the argument is different than the previous term. We start as usual by separating the two null sequences, that is $\mathbb{P}$-a.s.

$$\begin{aligned}
&\int_0^t \langle \phi'(Y_n(s-)), R(n)G_0(s, Y_n(s-))dW_{Q_0}(s)\rangle_H - \int_0^t \langle \phi'(Y(s-)), G_0(s, Y(s-))dW_{Q_0}(s)\rangle_H \\
&\quad = \underbrace{\int_0^t \langle \phi'(Y_n(s-)) - \phi'(Y(s-)), G_0(s, Y(s-))dW_{Q_0}(s)\rangle_H}_{=:(I_W)} \\
&\qquad\qquad + \underbrace{\int_0^t \langle \phi'(Y_n(s-)), \big(R(n)G_0(s, Y_n(s-)) - G_0(s, Y(s-))\big)dW_{Q_0}(s)\rangle_H}_{:=(II_W)} .
\end{aligned}$$

In order to show that (I_W) converges $\mathbb{P}$-a.s. to zero, we want to apply Theorem B.13, which is a stochastic version of Lebesgue's dominated convergence theorem. First, note that we can rewrite (I_W) as

$$(I_W) = \int_0^t G_0(s, Y(s-))^\natural \big(\phi'(Y_n(s-)) - \phi'(Y(s-))\big)dW_{Q_0}(s),$$

where $G_0(s, Y(s-))^\natural$ is the dual mapping of $G_0(s, Y(s-))$ (see [**GM11a**, Section 2.3.1] for more detail). In order to show the $dt \otimes \mathbb{P}$-a.s. convergence, we use the linear boundedness of G_0 and Proposition 3.11, that is for all $s \in [0, T]$ and a.e. $\omega \in \Omega$

$$\begin{aligned}
&\|G_0(\omega, s, Y(s-, \omega))^\natural \big(\phi'(Y_n(s-, \omega)) - \phi'(Y(s-, \omega))\big)Q_0^{1/2}\|_{L_{(HS)}(U,\mathbb{R})} \\
&\quad \leqslant \|\phi'(Y_n(s-, \omega)) - \phi'(Y(s-, \omega))\|_H \|G_0(\omega, s, Y(s-, \omega))Q_0^{1/2}\|_{L_{(HS)}(U,H)} \\
&\quad \leqslant k_{G_0}\|\phi'(Y_n(s-, \omega)) - \phi'(Y(s-, \omega))\|_H(1 + \|Y(s-, \omega)\|_H) \longrightarrow 0 \text{ for } n \to \infty.
\end{aligned}$$

Furthermore, we estimate the following integral for a.e. $\omega \in \Omega$

$$\int_0^T \|G_0(\omega, s, Y(s-, \omega))^\natural \big(\phi'(Y_n(s-, \omega)) - \phi'(Y(s-, \omega))\big)Q_0^{1/2}\|_{L_{(HS)}(U,\mathbb{R})}ds \tag{3.16}$$

$$\begin{aligned}
&\leqslant \int_0^T \|\phi'(Y_n(s-, \omega)) - \phi'(Y(s-, \omega))\|_H \|G_0(\omega, s, Y(s-, \omega))Q_0^{1/2}\|_{L_{(HS)}(U,H)}ds \\
&\leqslant C_{\phi'}(\omega) \int_0^T \|G_0(\omega, s, Y(s-, \omega))Q_0^{1/2}\|_{L_{(HS)}(U,H)}ds
\end{aligned}$$

$$\leqslant C_{\phi'}(\omega)k_{G_0} \int_0^T 1 + \|Y(s-, \omega)\|_H ds < \infty. \tag{3.17}$$

Applying Lebesgue's dominated convergence theorem yields the $\mathbb{P}$-a.s. convergence to zero of the integral (3.16). With (3.17) we also found a majorant to (3.16) with enables us to invoke Theorem B.13, from which we find

$$(I_W) = \int_0^t \langle \phi'(Y_n(s-)) - \phi'(Y(s-)), G_0(s, Y(s-))dW_{Q_0}(s)\rangle_H \longrightarrow 0 \text{ for } n \to \infty,$$

where the limit is in probability. Thus, we can extract a subsequence which converges $\mathbb{P}$-a.s. to zero. For (II_W) we want to show the $\mathbb{P}$-a.s. convergence again via Theorem B.13. Therefore, we estimate for all $s \in [0,T]$ and a.e. $\omega \in \Omega$

$$\begin{aligned}
&\|\big(R(n)G_0(\omega,s,Y_n(s-,\omega)) - G_0(\omega,s,Y(s-,\omega))\big)^{\natural}\phi'(Y_n(s-,\omega))Q_0^{1/2}\|_{L_{(HS)}(U,\mathbb{R})}\\
&\leqslant \|\phi'(Y_n(s-,\omega))\|_H\|\big(R(n)G_0(\omega,s,Y_n(s-,\omega)) - G_0(\omega,s,Y(s-,\omega))\big)Q_0^{1/2}\|_{L_{(HS)}(U,H)}\\
&\leqslant C_{\phi'}(\omega)\big(\|(I-R(n))G_0(\omega,s,Y(s-,\omega))Q_0^{1/2}\|_{L_{(HS)}(U,H)}\\
&\qquad\qquad + \|R(n)\big(G_0(\omega,s,Y(s-,\omega)) - G_0(\omega,s,Y_n(s-,\omega))\big)Q_0^{1/2}\|_{L_{(HS)}(U,H)}\big)\\
&\leqslant C_{\phi'}(\omega)\big(\|(I-R(n))G_0(\omega,s,Y(s-,\omega))Q_0^{1/2}\|_{L_{(HS)}(U,H)} + C_R C_{G_0}\|Y(s-,\omega) - Y_n(s-,\omega)\|_H\big),
\end{aligned}$$

where we used the Lipschitz continuity of G_0. For the first term Corollary 3.13 shows, possibly for a subsequence, that it converges $dt \otimes \mathbb{P}$-a.s. to zero. The same holds for the second term, since $(Y_n(t))_{n>\gamma}$ converges $\mathbb{P}$-a.s. and uniformly in t to $Y(t)$. For the convergence in L_2 we use the estimate we just performed in order to find for a.e. $\omega \in \Omega$

$$\begin{aligned}
&\int_0^T \|\big(R(n)G_0(\omega,s,Y_n(s-,\omega)) - G_0(\omega,s,Y(s-,\omega))\big)^{\natural}\phi'(Y_n(s-,\omega))Q_0^{1/2}\|_{L_{(HS)}(U,\mathbb{R})}ds\\
&\qquad \leqslant C_{\phi'}(\omega)(1+C_R)k_{G_0}\int_0^T 1 + \|Y(s-,\omega)\|_H ds + 2C_{\phi'}(\omega)C_R C_{G_0}C_Y(\omega)T < \infty,
\end{aligned} \tag{3.18}$$

where we use the linear growth condition of G_0. Together with the pointwise convergence Lebesgue's dominated convergence theorem guarantees the $\mathbb{P}$-a.s. convergence of

$$\int_0^T \|\big(R(n)G_0(\omega,s,Y_n(s-,\omega)) - G_0(\omega,s,Y(s-,\omega))\big)^{\natural}\phi'(Y_n(s-,\omega))Q_0^{1/2}\|_{L_{(HS)}(U,\mathbb{R})}ds \longrightarrow 0,$$

for $n \to \infty$. With (3.18) we also found the necessary $\mathbb{P}$-a.s. majorant. Therefore, Theorem B.13 implies for all $t \in [0,T]$

$$(II_W) = \int_0^t \langle \phi'(Y_n(s-)), \big(R(n)G_0(s,Y_n(s-)) - G_0(s,Y(s-))\big)dW_{Q_0}(s)\rangle_H \longrightarrow 0 \text{ for } n \to \infty$$

in probability. Again, we can extract a subsequence which converges $\mathbb{P}$-a.s. to zero.

For the Itô correction term (3.13) we start by separating the two null sequences, that is $\mathbb{P}$-a.s.

$$\begin{aligned}
&\int_0^t \operatorname{tr}\big[\phi''(Y_n(s-))\big(R(n)G_0(s,Y_n(s-))Q_0^{1/2}\big)\big(R(n)G_0(s,Y_n(s-))Q_0^{1/2}\big)^*\big]ds\\
&\qquad - \int_0^t \operatorname{tr}\big[\phi''(Y(s-))\big(G_0(s,Y(s-))Q_0^{1/2}\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)^*\big]ds\\
&= \underbrace{\int_0^t \operatorname{tr}\big[\big(\phi''(Y_n(s-)) - \phi''(Y(s-))\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)^*\big]ds}_{:=(I_{\mathrm{tr}})}\\
&\qquad + \underbrace{\begin{aligned}\int_0^t \operatorname{tr}\big[\phi''(Y_n(s-))\big\{&\big(R(n)G_0(s,Y_n(s-))Q_0^{1/2}\big)\big(R(n)G_0(s,Y_n(s-))Q_0^{1/2}\big)^*\\ &-\big(G_0(s,Y(s-))Q_0^{1/2}\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)^*\big\}\big]ds.\end{aligned}}_{:=(II_{\mathrm{tr}})}
\end{aligned}$$

In order to treat (I_{tr}), we estimate for fixed $s \in [0,T]$ and $\mathbb{P}$-a.s.

$$
\begin{aligned}
&\big| \operatorname{tr}\big[\big(\phi''(Y_n(s-)) - \phi''(Y(s-))\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)^*\big]\big| \\
&\quad \leqslant \|\phi''(Y_n(s-)) - \phi''(Y(s-))\|_{L(H)} \|\big(G_0(s,Y(s-))Q_0^{1/2}\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)^*\|_{nuk} \\
&\quad \leqslant \|\phi''(Y_n(s-)) - \phi''(Y(s-))\|_{L(H)} \|G_0(s,Y(s-))Q_0^{1/2}\|^2_{L_{HS}(U,H)} \\
&\quad \leqslant k_{G_0} \|\phi''(Y_n(s-)) - \phi''(Y(s-))\|_{L(H)} (1 + \|Y(s-)\|_H)^2 \\
&\quad \leqslant k_{G_0} C_Y^2 \|\phi''(Y_n(s-)) - \phi''(Y(s-))\|_{L(H)},
\end{aligned}
$$

where we used the linear growth condition of G_0. Applying Proposition 3.11 yields

$$
\|\phi''(Y_n(s-)) - \phi''(Y(s-))\|_{L(H)} \longrightarrow 0 \text{ for } n \to \infty \ \mathbb{P}\text{-a.s.},
$$

and Corollary 3.10 shows that it is bounded by some $\mathbb{P}$-a.s. finite constant. Thus, from Lebesgue's dominated convergence theorem flows that $\mathbb{P}$-a.s.

$$
(I_{\text{tr}}) = \int_0^t \operatorname{tr}\big[\big(\phi''(Y_n(s-)) - \phi''(Y(s-))\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)^*\big] ds \longrightarrow 0,
$$

for $n \to \infty$. In order to show that (II_{tr}) converges to zero as well, we rewrite it as follows

$$
\begin{aligned}
&\operatorname{tr}\big[\phi''(Y_n(s-))\big\{\big(R(n)G_0(s,Y_n(s-))Q_0^{1/2}\big)\big(R(n)G_0(s,Y_n(s-))Q_0^{1/2}\big)^* \\
&\qquad\qquad - \big(G_0(s,Y(s-))Q_0^{1/2}\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)^*\big\}\big] \\
&= \operatorname{tr}\big[\phi''(Y_n(s-))\big\{\big(R(n)(G_0(s,Y_n(s-)) - G_0(s,Y(s-)))Q_0^{1/2}\big)\big(R(n)G_0(s,Y_n(s-))Q_0^{1/2}\big)^* \\
&\qquad\qquad - \big((I - R(n))G_0(s,Y(s-))Q_0^{1/2}\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)^* \\
&\qquad\qquad - \big(R(n)G_0(s,Y(s-))Q_0^{1/2}\big)\big(R(n)(G_0(s,Y_n(s-)) - G_0(s,Y(s-)))Q_0^{1/2}\big)^* \\
&\qquad\qquad - \big(R(n)G_0(s,Y(s-))Q_0^{1/2}\big)\big((I - R(n))G_0(s,Y(s-))Q_0^{1/2}\big)^*\big\}\big].
\end{aligned}
$$

Hence, we are able to estimate $\mathbb{P}$-a.s.

$$
\begin{aligned}
&\big| \operatorname{tr}\big[\phi''(Y_n(s-))\big\{\big(R(n)G_0(s,Y_n(s-))Q_0^{1/2}\big)\big(R(n)G_0(s,Y_n(s-))Q_0^{1/2}\big)^* \\
&\qquad\qquad - \big(G_0(s,Y(s-))Q_0^{1/2}\big)\big(G_0(s,Y(s-))Q_0^{1/2}\big)^*\big\}\big]\big| \\
&\leqslant C_{\phi''}\big[C_R^2 \big\|(G_0(s,Y_n(s-)) - G_0(s,Y(s-)))Q_0^{1/2}\big\|_{L_{(HS)}(U,H)} \big\|G_0(s,Y_n(s-))Q_0^{1/2}\big\|_{L_{(HS)}(U,H)} \\
&\qquad + \big\|(I - R(n))G_0(s,Y(s-))Q_0^{1/2}\big\|_{L_{(HS)}(U,H)} \big\|G_0(s,Y(s-))Q_0^{1/2}\big\|_{L_{(HS)}(U,H)} \\
&\qquad + C_R^2 \big\|G_0(s,Y(s-))Q_0^{1/2}\big\|_{L_{(HS)}(U,H)} \big\|(G_0(s,Y_n(s-)) - G_0(s,Y(s-)))Q_0^{1/2}\big\|_{L_{(HS)}(U,H)} \\
&\qquad + C_R \big\|G_0(s,Y(s-))Q_0^{1/2}\big\|_{L_{(HS)}(U,H)} \big\|(I - R(n))G_0(s,Y(s-))Q_0^{1/2}\big\|_{L_{(HS)}(U,H)}\big] \\
&\leqslant C_{\phi''}\big[C_R^2 k_{G_0} C_{G_0} (1 + \|Y_n(s-)\|_H) \|Y_n(s-) - Y(s-)\|_H \\
&\qquad + k_{G_0}(1 + \|Y(s-)\|_H) \big\|(I - R(n))G_0(s,Y(s-))Q_0^{1/2}\big\|_{L_{(HS)}(U,H)} \\
&\qquad + C_R^2 k_{G_0} C_{G_0} (1 + \|Y(s-)\|_H) \|Y_n(s-) - Y(s-)\|_H \\
&\qquad + C_R k_{G_0} (1 + \|Y(s-)\|_H) \big\|(I - R(n))G_0(s,Y(s-))Q_0^{1/2}\big\|_{L_{(HS)}(U,H)}\big]
\end{aligned}
$$

$$\begin{aligned}
&\leqslant C_{\phi''}\big[2C_R^2 k_{G_0} C_Y C_{G_0} \|Y_n(s-) - Y(s-)\|_H + k_{G_0} C_Y \big\|(I - R(n))G_0(s, Y(s-))Q_0^{1/2}\big\|_{L_{(HS)}(U,H)} \\
&\qquad + C_R k_{G_0} C_Y \big\|(I - R(n))G_0(s, Y(s-))Q_0^{1/2}\big\|_{L_{(HS)}(U,H)}\big] \\
&\leqslant C_{\phi''}\big[4C_R^2 k_{G_0} C_Y^2 C_{G_0} + k_{G_0}^2 C_Y^2 (1 + C_R) + C_R k_{G_0}^2 C_Y^2 (1 + C_R)\big] < \infty.
\end{aligned}$$

From Corollary 3.13 we know that the term $\|(I - R(n))G_0(s, Y(s-))Q_0^{1/2}\|_{L_{(HS)}(U,H)}$ converges, possibly for a subsequence, to zero and since $\{Y_n(s-)\}_{n>\gamma}$ converges to $Y(s-)$ $\mathbb{P}$-a.s. and uniformly in s, it follows from the penultimate line in the estimate that the integrand of (II_{tr}) coverages pointwise in s and $\mathbb{P}$-a.s. to zero. The last line provides an $\mathbb{P}$-a.s. finite majorant. Hence, Lebesgue's dominated convergence theorem yields the claimed convergence.

For the term (3.14) we want to show the convergence via Theorem B.48, which is a stochastic version of Lebesgue's dominated convergence theorem for compensated Poisson integrals. Therefore, we start by showing the pointwise convergence to zero of the following expression

$$\phi(Y_n(s-) + R(n)G_1(s, Y_n(s-))x) - \phi(Y_n(s-)) - \phi(Y(s-) + G_1(s, Y(s-))x) + \phi(Y(s-)),$$

for fixed $\omega \in \Omega$, $s \in [0, T]$, and $x \in U$. From Proposition 3.11 it follows that a subsequence of $\phi(Y(s-)) - \phi(Y_n(s-))$ converges to zero. Next, we look at the following expression

$$\begin{aligned}
&\mathbb{E}\int_U \|G_1(s, Y_n(s-))x - G_1(s, Y(s-))x\|_H^2 \nu(dx) \\
&\qquad = \mathbb{E}\big\|\big(G_1(s, Y_n(s-)) - G_1(s, Y(s-))\big)Q^{1/2}\big\|_{L_{(HS)}(U,H)}^2 \leqslant C_{G_1}^2 \mathbb{E}\|Y_n(s-) - Y(s-)\|_H^2,
\end{aligned}$$

which converges for all $s \in [0, T]$ to zero, since Y_n converges to Y in $\mathcal{X}_{T,H}$. Hence, we can extract a subsequence, which converges for all $s \in [0, T]$ $\nu\otimes\mathbb{P}$-a.s. to zero. Since ϕ is uniformly continuous on bounded subsets of H, we look at the difference of the arguments

$$\begin{aligned}
&\|Y_n(s-) + R(n)G_1(s, Y_n(s-))x - Y(s-) - G_1(s, Y(s-))x\|_H \\
&\leqslant \|Y_n(s-) - Y(s-)\|_H + \|R(n)G_1(s, Y_n(s-))x - G_1(s, Y(s-))x\|_H \\
&\leqslant \|Y_n(s-) - Y(s-)\|_H + \|(R(n) - I)G_1(s, Y(s-))x\|_H \\
&\qquad\qquad + C_R\|G_1(s, Y_n(s-))x - G_1(s, Y(s-))x\|_H.
\end{aligned}$$

We already know that the first and third term converge, possibly for a subsequence, to zero. The second term also converges to zero, since $G_1(s, Y(s-))x \in H$ for fixed $s \in [0, T]$, $\omega \in \Omega$ and $x \in U$, thus Corollary A.21 (ii) applies. Summing up, we found that

$$\phi(Y_n(s-) + R(n)G_1(s, Y_n(s-))x) - \phi(Y_n(s-)) - \phi(Y(s-) + G_1(s, Y(s-))x) + \phi(Y(s-))$$

converges pointwise to zero.

In [**MRT13**] the following $\mathbb{P}$-a.s. estimate was shown in the proof of Theorem 3.6.

$$\begin{aligned}
&\int_0^T \int_U |\phi(Y_n(s-) + R(n)G_1(s, Y_n(s-))x) - \phi(Y_n(s-))|^2 \nu(dx)ds \\
&\quad \leqslant 2C_{h_1}^2 \int_0^T \int_U h_1(\|Y_n(s-)\|_H)^2 \|R(n)G_1(s, Y_n(s-))x\|_H^2 \nu(dx)ds \\
&\qquad + 2C_{h_1}^4 h_1^2(1) \int_0^T \int_U h_1(\|R(n)G_1(s, Y_n(s-))x\|_H)^2 \|R(n)G_1(s, Y_n(s-))x\|_H^2 \nu(dx)ds.
\end{aligned}$$

Note that we already applied this to our situation. We proceed with the estimate $\mathbb{P}$-a.s.

$$\begin{aligned}
&\int_0^T \int_U |\phi(Y_n(s-) + R(n)G_1(s, Y_n(s-))x) - \phi(Y_n(s-))|^2 \nu(dx)ds \\
&\leqslant 2C_{h_1}^2 h_1(C_Y)^2 C_R^2 k_{G_1}^2 \int_0^T \int_U (1 + \|Y_n(s-)\|_H)^2 \|x\|_U^2 \nu(dx)ds \\
&+ 2C_{h_1}^6 h_1(1)^2 C_R^2 k_{G_1}^2 \int_0^T \int_U h_1(C_R k_{G_1}(1+\|Y_n(s-)\|_H))^2 h_1(\|x\|_U)^2 (1+\|Y_n(s-)\|_H)^2 \|x\|_U^2 \nu(dx)ds \\
&\leqslant 2C_{h_1}^2 h_1(C_Y)^2 C_R^2 k_{G_1}^2 C_Y^2 T \operatorname{tr}(Q_1) \\
&+ 2C_{h_1}^6 h_1(1)^2 C_R^2 k_{G_1}^2 h_1(C_R k_{G_1} C_Y)^2 C_Y^2 T \int_U h_1(\|x\|_U)^2 \|x\|_U^2 \nu(dx) < \infty.
\end{aligned}$$

Using the same procedure we find $\mathbb{P}$-a.s.

$$\begin{aligned}
&\int_0^T \int_U |\phi(Y(s-) + G_1(s, Y_n(s-))x) - \phi(Y(s-))|^2 \nu(dx)ds \\
&\quad \leqslant 2C_{h_1}^2 h_1(C)^2 k_{G_1}^2 C_Y^2 T \operatorname{tr}(Q_1) + 2C_{h_1}^6 h_1(1)^2 h_1(C_Y)^2 C_Y^2 T \int_U h_1(\|x\|_U)^2 \|x\|_U^2 \nu(dx) < \infty.
\end{aligned}$$

Applying Lebesgue's dominated convergence theorem yields $\mathbb{P}$-a.s.

$$\begin{aligned}
&\int_0^T \int_U |\phi(Y_n(s-) + R(n)G_1(s, Y_n(s-))x) - \phi(Y_n(s-)) \\
&\qquad - \phi(Y(s-) + G_1(s, Y(s-))x) + \phi(Y(s-))|\nu(dx)ds \longrightarrow 0 \text{ for } n \to \infty.
\end{aligned}$$

The two estimates above also provide the necessary majorant. Hence, Theorem B.48 yields the desired convergence in probability. Thus, we can extract a subsequence which converges $\mathbb{P}$-a.s. to zero.

For the last term (3.15) we start by showing the pointwise convergence to zero. When we dealt with the term (3.14) above, we already showed that

$$\phi(Y_n(s-) + R(n)G_1(s, Y_n(s-))x) - \phi(Y_n(s-)) - \phi(Y(s-) + G_1(s, Y(s-))x) + \phi(Y(s-)) \to 0,$$

for $n \to \infty$.

Therefore, all what is left to show is that the term involving the first derivative converges to zero as well. In order to show this, we estimate as follows

$$\begin{aligned}
&\langle \phi'(Y_n(s-)), R(n)G_1(s, Y_n(s-))x\rangle_H - \langle \phi'(Y(s-)), G_1(s, Y(s-))x\rangle_H \\
&\quad \leqslant \|\phi'(Y_n(s-)) - \phi'(Y(s-))\|_H \|G_1(s, Y(s-))x\|_H \\
&\qquad\qquad + \|\phi'(Y_n(s-))\|_H \|R(n)G_1(s, Y_n(s-))x - G_1(s, Y(s-))x\|_H.
\end{aligned}$$

For the first term note that $G_1(s, Y(s-))x$ is a fixed element in H. Hence, the first term converges to zero by Proposition 3.11. For the second term we estimate $\|\phi'(Y_n(s-))\|_H$ by a $\mathbb{P}$-a.s. finite constant, since ϕ' is bounded on bounded subsets of H. Finally, in the proof of the convergence of (3.14) above we already showned that $\|R(n)G_1(s, Y_n(s-))x - G_1(s, Y(s-))x\|_H$ converges to zero. This yields the pointwise convergence. In order to find a majorant, we apply an estimate from the proof of Theorem 3.6. in [**MRT13**]. This way we find $\mathbb{P}$-a.s.

$$\begin{aligned}
&\int_0^T \int_U |\phi(Y_n(s-) + R(n)G_1(s, Y_n(s-))x) - \phi(Y_n(s-)) \\
&\qquad\qquad - \langle \phi'(Y_n(s-)), R(n)G_1(s, Y_n(s-))x\rangle_H| \nu(dx) ds \\
&\quad \leqslant C_{h_2} \int_0^T \int_U h_2(\|Y_n(s-)\|_H) \|R(n)G_1(s, Y_n(s-))x\|_H^2 \nu(dx) ds \\
&\qquad + C_{h_2}^2 h_2(1) \int_0^T \int_U h_2(\|R(n)G_1(s, Y_n(s-))x\|_H) \|R(n)G_1(s, Y_n(s-))x\|_H^2 \nu(dx) ds.
\end{aligned}$$

We proceed with the estimate $\mathbb{P}$-a.s.

$$\begin{aligned}
&\int_0^T \int_U |\phi(Y_n(s-) + R(n)G_1(s, Y_n(s-))x) - \phi(Y_n(s-)) \\
&\qquad\qquad - \langle \phi'(Y_n(s-)), R(n)G_1(s, Y_n(s-))x\rangle_H| \nu(dx) ds \\
&\leqslant C_{h_2} h_2(C_Y) C_R^2 k_{G_1}^2 \int_0^T \int_U (1 + \|Y_n(s-)\|_H)^2 \|x\|_U^2 \nu(dx) ds \\
&+ C_{h_2}^3 h_2(1) C_R^2 k_{G_1}^2 \int_0^T \int_U h_2(C_R k_{G_1}(1 + \|Y_n(s-)\|_H)) h_2(\|x\|_U)(1 + \|Y_n(s-)\|)^2 \|x\|_U^2 \nu(dx) ds \\
&\leqslant C_{h_2} h_2(C_Y) C_R^2 k_{G_1}^2 C_Y^2 T \operatorname{tr}(Q_1) \\
&+ C_{h_2}^3 h_2(1) C_R^2 k_{G_1}^2 h_2(C_R k_{G_1} C_Y) C_Y^2 T \int_U h_2(\|x\|_U) \|x\|_U^2 \nu(dx) < \infty.
\end{aligned}$$

Using the same procedure we find $\mathbb{P}$-a.s.

$$\begin{aligned}
&\int_0^T \int_U |\phi(Y(s-) + G_1(s, Y(s-))x) - \phi(Y(s-)) - \langle \phi'(Y(s-)), G_1(s, Y(s-))x\rangle_H| \nu(dx) ds \\
&\leqslant C_{h_2} h_2(C_Y) k_{G_1}^2 C_Y^2 T \operatorname{tr}(Q_1) + C_{h_2}^3 h_2(1) k_{G_1}^2 h_2(k_{G_1} C_Y) C_Y^2 T \int_U h_2(\|x\|_U) \|x\|_U^2 \nu(dx) < \infty.
\end{aligned}$$

By Lebesgue's dominated convergence theorem we find that (3.15) converges $\mathbb{P}$-a.s. to

$$\int_0^t \int_U \phi(Y(s-) + G_1(s, Y(s-))x) - \phi(Y(s-)) - \langle \phi'(Y(s-)), G_1(s, Y(s-))x\rangle_H \nu(dx) ds.$$

In conclusion we have show that, possibly for a subsequence, the left- and right-hand sides of (3.9)-(3.15) converge $\mathbb{P}$-a.s. to the left- and right-hand sides of the transformation formula as claimed in the statement of the theorem.

□

REMARK 3.16. *If a function $\phi : H \to \mathbb{R}$ fulfills the assumptions of the transformation formula from Theorem 3.3, we can always construct a new function $\tilde{\phi} : H \to \mathbb{R}$ given by*

$$x \mapsto \phi((\gamma_0 - A)^{-1}x),$$

for a fixed $\gamma_0 > \gamma$, such that $\tilde{\phi}$ fulfills the assumptions of Theorem 3.15. For example the function φ given by $\varphi(x) = \|(\gamma_0 - A)^{-1}x\|_H^2$ fulfills the assumption of Theorem 3.15. We discuss a similar example in full detail in Section 4.1.

3.4. The adjoint operator of $\mathscr{A}$

From Theorem 3.15 it is apparent, that we need to know the adjoint operator $(A^*, \mathscr{D}(A^*))$ of the driving linear operator from $(S - SCP)$ in order to apply the transformation formula. Our case of interest is the delay equation $(S - DE)$ in its transformed version $(S_D - ACP)$. Hence, the driving linear operator is given by

$$\mathscr{A} = \begin{pmatrix} B & \Phi \\ 0 & \frac{d}{d\sigma} \end{pmatrix},$$

with domain

$$\mathscr{D}(\mathscr{A}) = \left\{ \begin{pmatrix} h \\ f \end{pmatrix} \in \mathscr{D}(B) \times W_2^1(I; Z; d\mu) : f(0) = h \right\}.$$

In this section we calculate the operator $(\mathscr{A}^*, \mathscr{D}(\mathscr{A}^*))$ for the most important examples. Therefore, we assume that we know the adjoint operator of $(B, \mathscr{D}(B))$ and we set $\varrho \equiv 1$. Furthermore, we consider the following delay operator in the finite delay case

$$\begin{aligned} \Phi_1 :\ & W_2^1([-1,0]; H) \to H \\ & f \mapsto Cf(-1) + \int_{-1}^{0} D_1(\sigma) f(\sigma) d\sigma, \end{aligned} \tag{3.19}$$

where $C \in L(H)$ and $D_1 \in L_2([-1,0]; L(H))$. In the infinite delay case we look at

$$\begin{aligned} \Phi_2 : & W_2^1(\mathbb{R}_-; H) \to H \\ & f \mapsto \int_{-\infty}^{0} D_2(\sigma) f(\sigma) d\sigma, \end{aligned} \tag{3.20}$$

where $D_2 \in L_2(\mathbb{R}_-; L(H))$.

REMARK 3.17. *From Section 2.2 we know that, if we consider the delay operator Φ_1, the operator $\mathscr{A}$ is λ-dissipative if and only if Φ_1 has a single delay (and not a multiple delay) in the discrete part. Hence, only in this case we can apply Theorem 3.15. However, one can still* *calculate the adjoint operator $(A^*, \mathscr{D}(A^*))$ for the multiple delay case. For example one can adopt the method used in* [**Kap85**].

PROPOSITION 3.18. *Consider the case of a finite delay. Assume the delay operator is given by* (3.19)*. Then, the adjoint operator of $(\mathscr{A}^*, \mathscr{D}(\mathscr{A}^*))$ is given by*

$$\mathscr{A}^* x = \begin{pmatrix} B^* x_1 + x_2(0) \\ D_1^*(\cdot) x_1 - \frac{d}{d\sigma} x_2(\cdot) \end{pmatrix},$$

for

$$x \in \mathscr{D}(\mathscr{A}^*) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathscr{D}(B^*) \times W_2^1([-1,0];H) : x_2(-1) = C^* x_1 \right\}.$$

PROOF.

Assume the operator $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ is given. Then we know from functional analysis that the operator $(\mathscr{A}^*, \mathscr{D}(\mathscr{A}^*))$ exists. Therefore, it follows for every $x \in \mathscr{D}(\mathscr{A}^*)$ that $(\mathscr{A}^* x)_1 \in H$ and $(\mathscr{A}^* x)_2 \in L_2([-1,0];H)$, where we use the notation introduced in Definition 1.17. Finally, the adjoint operator of $\mathscr{A}$ has to fulfill

$$\langle \mathscr{A}^* x, f \rangle_{\mathscr{E}_2} = \langle x, \mathscr{A} f \rangle_{\mathscr{E}_2}, \tag{3.21}$$

for all $x \in \mathscr{D}(\mathscr{A}^*)$ and all $f \in \mathscr{D}(\mathscr{A})$. We look at the right-hand side of (3.21) first

$$\begin{aligned} \langle x, \mathscr{A} f \rangle_{\mathscr{E}_2} &= \langle x_1, B f_1 \rangle_H + \langle x_1, \Phi_1 f_2 \rangle_H + \int_{-1}^0 \langle x_2(\sigma), \dot{f}_2(\sigma) \rangle_H d\sigma \\ &= \langle B f_2(0), x_1 \rangle_H + \langle x_1, C f_2(-1) \rangle_H \\ &\quad + \langle x_1, \int_{-1}^0 D_1(\sigma) f_2(\sigma) d\sigma \rangle_H + \int_{-1}^0 \langle x_2(\sigma), \dot{f}_2(\sigma) \rangle_H d\sigma \\ &= \langle B f_2(0), x_1 \rangle_H + \langle f_2(-1), C^* x_1 \rangle_H \\ &\quad + \int_{-1}^0 \langle f_2(\sigma), D_1^*(\sigma) x_1 \rangle_H d\sigma + \int_{-1}^0 \langle \dot{f}_2(\sigma), x_2(\sigma) \rangle_H d\sigma, \end{aligned}$$

where we applied $f_1 = f_2(0)$, since $f \in \mathscr{D}(\mathscr{A})$, and $\dot{f}$ denotes the (weak) derivative of f. For the right-hand side of (3.21) we find

$$\begin{aligned} \langle \mathscr{A}^* x, f \rangle_{\mathscr{E}_2} &= \langle (\mathscr{A}^* x)_1, f_1 \rangle_H + \int_{-1}^0 \langle (\mathscr{A}^* x)_2(\sigma), f_2(\sigma) \rangle_H d\sigma \\ &= \langle f_2(0), (\mathscr{A}^* x)_1 \rangle_H + \int_{-1}^0 \langle f_2(\sigma), (\mathscr{A}^* x)_2(\sigma) \rangle_H d\sigma. \end{aligned}$$

Plugging the calculated terms into (3.21), we receive

$$\begin{aligned} &\langle f_2(0), (\mathscr{A}^* x)_1 \rangle_H - \langle B f_2(0), x_1 \rangle_H - \langle f_2(-1), C^* x_1 \rangle_H \\ &\qquad = \int_{-1}^0 \langle f_2(\sigma), D_1^*(\sigma) x_1 - (\mathscr{A}^* x)_2(\sigma) \rangle_H d\sigma + \int_{-1}^0 \langle \dot{f}_2(\sigma), x_2(\sigma) \rangle_H d\sigma. \end{aligned} \tag{3.22}$$

First we will calculate $(\mathscr{A}^* x)_2$. Therefore, we are argue similar as in [**BH76**]. Thus, if we choose $f_1 = 0$ and $f_2 \in C_c^\infty((-1,0);H)$ we have $f \in \mathscr{D}(\mathscr{A})$ with $f_1 = f_2(0) = f(-1) = 0$. This way we find

$$\int_{-1}^{0} \langle f_2(\sigma), D_1^*(\sigma)x_1 - (\mathscr{A}^*x)_2(\sigma)\rangle_H d\sigma = -\int_{-1}^{0} \langle \dot{f}_2(\sigma), x_2(\sigma)\rangle_H d\sigma,$$

for all $f_2 \in C_c^\infty((-1,0);H)$. Hence, $x_2 \in W_2^1([-1,0];H)$ with

$$\dot{x}_2(\sigma) = D_1^*(\sigma)x_1 - (\mathscr{A}^*x)_2(\sigma).$$

Therefore, we found the representation for $(\mathscr{A}^*x)_2$, that is

$$(\mathscr{A}^*x)_2(\sigma) = D_1^*(\sigma)x_1 - \dot{x}_2(\sigma).$$

Furthermore, we calculate for $f \in \mathscr{D}(\mathscr{A})$ and $x \in \mathscr{D}(\mathscr{A}^*)$

$$\begin{aligned}\langle f_2(0), x_2(0)\rangle_H - \langle f_2(-1), x_2(-1)\rangle_H &= \int_{-1}^{0} \frac{d}{d\sigma}[\langle f_2(\sigma), x_2(\sigma)\rangle_H] d\sigma \\ &= \int_{-1}^{0} \langle f_2(\sigma), D_1^*(\sigma)x_1 - (\mathscr{A}^*x)_2(\sigma)\rangle_H d\sigma + \int_{-1}^{0} \langle \dot{f}_2(\sigma), x_2(\sigma)\rangle_H d\sigma.\end{aligned}$$

Thus, we can rewrite (3.22) as follows

$$\langle f_2(0), (\mathscr{A}^*x)_1 - x_2(0)\rangle_H - \langle Bf_2(0), x_1\rangle_H = \langle f_2(-1), C^*x_1 - x_2(-1)\rangle_H. \tag{3.23}$$

Note that, if we now choose $f \in \mathscr{D}(\mathscr{A})$ with $f_1 = f_2(0) = 0$, $f_2(-1)$ can take any value in H. Hence, we find

$$\langle h, C^*x_1 - x_2(-1)\rangle_H = 0,$$

for all $h \in H$. Therefore, we conclude

$$x_2(-1) = C^*x_1.$$

Thus, (3.23) reduces to

$$\langle f_2(0), (\mathscr{A}^*x)_1 - x_2(0)\rangle_H = \langle Bf_2(0), x_1\rangle_H.$$

Since we can choose $f_1 = f_2(0) \in \mathscr{D}(B)$ arbitrary, it follows that $x_1 \in \mathscr{D}(B^*)$ and

$$(\mathscr{A}^*x)_1 = B^*x_1 + x_2(0).$$

Now let us assume the operator $(\mathscr{A}^*, \mathscr{D}(\mathscr{A}^*))$ is given. Then again, we know from functional analysis that the operator $(\mathscr{A}, \mathscr{D}(\mathscr{A}))$ exists. Therefore, it follows for every $f \in \mathscr{D}(\mathscr{A})$ that $(\mathscr{A}f)_1 \in H$ and $(\mathscr{A}f)_2 \in L_2([-1,0];H)$. Finally, the operator $\mathscr{A}$ has to fulfill

$$\langle \mathscr{A}^*x, f\rangle_{\mathscr{E}_2} = \langle x, \mathscr{A}f\rangle_{\mathscr{E}_2}, \tag{3.24}$$

for all $x \in \mathscr{D}(\mathscr{A}^*)$ and all $f \in \mathscr{D}(\mathscr{A})$. We look at the left-hand side of (3.24) first

$$\begin{aligned}\langle \mathscr{A}^*x, f\rangle_{\mathscr{E}_2} &= \langle B^*x_1 + x_2(0), f_1\rangle_H + \int_{-1}^{0} \langle D_1^*(\sigma)x_1 - \dot{x}_2(\sigma), f_2(\sigma)\rangle_H d\sigma \\ &= \langle B^*x_1, f_1\rangle_H + \langle x_2(0), f_1\rangle_H + \langle x_1, \int_{-1}^{0} D_1(\sigma)f_2(\sigma)d\sigma\rangle_H - \int_{-1}^{0} \langle \dot{x}_2(\sigma), f_2(\sigma)\rangle_H d\sigma.\end{aligned}$$

For the left-hand side of (3.24) we find

$$\langle x, \mathscr{A} f\rangle_{\mathscr{E}_2} = \langle x_1, (\mathscr{A} f)_1\rangle_H + \int_{-1}^{0} \langle x_2(\sigma), (\mathscr{A} f)_2(\sigma)\rangle_H d\sigma.$$

Plugging the calculated terms into (3.24), we receive

$$\begin{aligned}\langle B^* x_1, f_1\rangle_H + \langle x_2(0), f_1\rangle_H + \langle x_1, \int_{-1}^{0} D_1(\sigma) f_2(\sigma) d\sigma - (\mathscr{A} f)_1\rangle_H \\ = \int_{-1}^{0} \langle x_2(\sigma), (\mathscr{A} f)_2(\sigma)\rangle_H d\sigma + \int_{-1}^{0} \langle \dot{x}_2(\sigma), f_2(\sigma)\rangle_H d\sigma.\end{aligned} \tag{3.25}$$

We will calculate $(Af)_2$ first. Therefore, if we choose $x_1 = 0$ and $x_2 \in C_c^\infty((-1,0); H)$, we have $x \in \mathscr{D}(\mathscr{A}^*)$ with $x_2(-1) = 0 = x_2(0)$. This way we find

$$\int_{-1}^{0} \langle x_2(\sigma), (\mathscr{A} f)_2(\sigma)\rangle_H d\sigma = -\int_{-1}^{0} \langle \dot{x}_2(\sigma), f_2(\sigma)\rangle_H d\sigma,$$

for all $x_2 \in C_c^\infty((-1,0); H)$. Hence, $f_2 \in W_2^1([-1,0]; H)$ with

$$(Af)_2(\sigma) = \dot{f}(\sigma).$$

Using this equality, we calculate for $x \in \mathscr{D}(\mathscr{A}^*)$ and $f \in \mathscr{D}(\mathscr{A})$

$$\begin{aligned}\langle x_2(0), f_2(0)\rangle_H - \langle x_2(-1), f_2(-1)\rangle_H &= \int_{-1}^{0} \frac{d}{d\sigma} \langle x_2(\sigma), f_2(\sigma)\rangle_H d\sigma \\ &= \int_{-1}^{0} \langle x_2(\sigma), (\mathscr{A} f)_2(\sigma)\rangle_H d\sigma + \int_{-1}^{0} \langle \dot{x}_2(\sigma), f_2(\sigma)\rangle_H d\sigma.\end{aligned}$$

Thus, we can rewrite (3.25) as follows

$$\begin{aligned}&\langle x_2(0), f_1 - f_2(0)\rangle_H \\ &\quad = \langle x_1, (\mathscr{A} f)_1 - \int_{-1}^{0} D_1(\sigma) f_2(\sigma) d\sigma\rangle_H - \langle x_2(-1), f_2(-1)\rangle_H - \langle B^* x_1, f_1\rangle_H \\ &\quad = \langle x_1, (\mathscr{A} f)_1 - \int_{-1}^{0} D_1(\sigma) f_2(\sigma) d\sigma\rangle_H - \langle C^* x_1, f_2(-1)\rangle_H - \langle B^* x_1, f_1\rangle_H \\ &\quad = \langle x_1, (\mathscr{A} f)_1 - C f_2(-1) - \int_{-1}^{0} D_1(\sigma) f_2(\sigma) d\sigma\rangle_H - \langle B^* x_1, f_1\rangle_H,\end{aligned} \tag{3.26}$$

where we used $x_2(-1) = C^* x_1$, since $x \in \mathscr{D}(\mathscr{A}^*)$. Note that, if we now choose $x \in \mathscr{D}(\mathscr{A}^*)$ with $x_1 = 0$, $x_2(0)$ can take any value in H. Hence, we find

$$\langle h, f_1 - f_2(0)\rangle_H = 0,$$

for all $h \in H$. Therefore, we conclude

$$f_2(0) = f_1.$$

Thus, (3.26) reduces to

$$\langle B^* x_1, f_1\rangle_H = \langle x_1, (\mathscr{A} f)_1 - C f_2(-1) - \int_{-1}^{0} D_1(\sigma) f_2(\sigma) d\sigma\rangle_H.$$

Since we can choose $x_1 \in \mathscr{D}(B^*)$ arbitrary, it follows that $f_1 \in \mathscr{D}(B^*)$ and

$$(\mathscr{A} f)_1 = Bf_1 + Cf_2(-1) + \int_{-1}^{0} D_1(\sigma) f_2(\sigma) d\sigma.$$

□

If we consider an infinite delay and the delay operator is given by (3.20), we can formulate the analog result to Proposition 3.18. Since the proof is exactly the same as in the finite delay case, we omit it.

PROPOSITION 3.19. *Consider the case of an infinite delay. Assume the delay operator is given by* (3.20)*. Then, the adjoint operator of* $(\mathscr{A}^*, \mathscr{D}(\mathscr{A}^*))$ *is given by*

$$\mathscr{A}^* x = \begin{pmatrix} B^* x_1 + x_2(0) \\ D_1^*(\cdot) x_1 - \frac{d}{d\sigma} x_2(\cdot) \end{pmatrix},$$

for

$$x \in \mathscr{D}(\mathscr{A}^*) = \mathscr{D}(B^*) \times W_2^1(\mathbb{R}_-; H).$$

REMARK 3.20. *It might seem a bit odd, that there is no extra condition for the domain of the operator* $\mathscr{A}^*$*. However, it is implicitly there, since the Sobolev space* $W_2^1(\mathbb{R}_-; H)$ *is embedded in* $C_0(\mathbb{R}; H) = \{f \in C(\mathbb{R}; H) : \lim_{x \to -\infty} f(x) = 0\}$*. For the same reason is the proof of the infinite delay case a bit simpler, since one term in the fundamental theorem of calculus is zero.*

CHAPTER 4

A Filtering Problem

In this final chapter we provide an application of the transformation formula from Theorem 3.15. First, we show a product formula for mild solutions. Then, we apply this formula in order to solve a filtering problem. Since we want to be able to calculate all the expressions in a closed analytic form, we have to choose our examples quite specifically. The advantage of this procedure is that we can see exactly how the developed tools get used.

4.1. A product formula

In this section we apply Theorem 3.15 to give an explicit example. Therefore, let H be a separable Hilbert space with inner product denoted by $\langle\cdot,\cdot\rangle_H$. Furthermore, we assume the operators $(A, \mathscr{D}(A))$ and $(B, \mathscr{D}(B))$ to be generators of C_0-semigroups of generalized contractions. Hence, there exists $\lambda_A, \lambda_B \in \mathbb{R}$, such that A is a λ_A-dissipative operator and B is a λ_B-dissipative operator. Then, we consider the operator $(C, \mathscr{D}(C))$ defined by

$$C := \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \text{ and } \mathscr{D}(C) := \mathscr{D}(A) \times \mathscr{D}(B),$$

on the Hilbert space $H \times H$. A straight forward estimate shows that the operator C is $\max\{\lambda_A, \lambda_B\}$-dissipative. Therefore, we fix a $\lambda \in \mathbb{R}$, such that $\lambda > \max\{\lambda_A, \lambda_B\}$. Furthermore, let X and Y be the mild solutions in H of the following two $(S-ACP)$

$$(S-ACP_X)\begin{cases} dX(t) &= AX(t)dt + F^A(t,X(t))dt + G^A(t,X(t))dM(t), \quad t \in [0,T], \\ X(0) &= x \in H, \end{cases}$$

and

$$(S-ACP_Y)\begin{cases} dY(t) &= BY(t)dt + F^B(t,Y(t))dt + G^B(t,Y(t))dM(t), \quad t \in [0,T], \\ Y(0) &= y \in H, \end{cases}$$

where F^A, F^B, G^A, G^B, and M fulfill the assumptions of Section 1.3 for $(S-ACP)$. Hence, both equations are well-posed. We can rewrite the two equations above equivalently as one in $H \times H$ in the following way

$$(S-ACP_C)\begin{cases} d\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} &= C\begin{pmatrix} X(t) \\ Y(t) \end{pmatrix}dt + \begin{pmatrix} F^A(t,X(t)) \\ F^B(t,Y(t)) \end{pmatrix}dt \\ &\quad + \begin{pmatrix} G^A(t,X(t)) & 0 \\ 0 & G^B(t,Y(t)) \end{pmatrix} d\begin{pmatrix} M(t) \\ M(t) \end{pmatrix}, \ t \in [0,T], \\ \begin{pmatrix} X(0) \\ Y(0) \end{pmatrix} &= \begin{pmatrix} x \\ y \end{pmatrix}. \end{cases}$$

It will be this equation we use to show the product formula. We choose the following bilinear form ϕ as transformation function

$$\phi : H \times H \to \mathbb{R}$$
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \langle (\lambda - A)^{-1} x_1, (\lambda - B)^{-1} x_2 \rangle_H.$$

The first thing we need to check, if we want to apply Theorem 3.15, is if ϕ_c exists. Recalling the definition from Section 3.2 we need to find an extension of

$$\phi \left((\lambda - C) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right),$$

for $x_1 \in \mathscr{D}(A)$ and $x_2 \in \mathscr{D}(B)$ to all of $H \times H$. Therefore, we calculate

$$\phi \left((\lambda - C) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = \phi \left(\begin{pmatrix} (\lambda - A) x_1 \\ (\lambda - B) x_2 \end{pmatrix} \right) = \langle (\lambda - A)^{-1} (\lambda - A) x_1, (\lambda - B)^{-1} (\lambda - B) x_2 \rangle_H.$$

Hence, ϕ_c is given by

$$\phi_c \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \langle x_1, x_2 \rangle_H.$$

Furthermore, we need the Fréchet derivatives of ϕ and ϕ_c. They can be calculated easily using the definition of the Fréchet derivative. Hence, we find

$$\phi' \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \langle (\lambda - A^*)^{-1} (\lambda - B)^{-1} x_2, v_1 \rangle_H + \langle (\lambda - B^*)^{-1} (\lambda - A)^{-1} x_1, v_2 \rangle_H.$$

By the Riesz representation theorem we can identify ϕ' with

$$\phi' \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} (\lambda - A^*)^{-1} (\lambda - B)^{-1} x_2 \\ (\lambda - B^*)^{-1} (\lambda - A)^{-1} x_1 \end{pmatrix}.$$

For the second derivative follows

$$\phi'' \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & (\lambda - A^*)^{-1} (\lambda - B)^{-1} \\ (\lambda - B^*)^{-1} (\lambda - A)^{-1} & 0 \end{pmatrix}.$$

For ϕ_c we have

$$\phi_c' \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \text{ and } \phi_c'' \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}.$$

Obviously, $\phi, \phi_c, \phi', \phi_c', \phi''$ are bounded and uniformly continuous on bounded subsets of $H \times H$. Furthermore, we have

$$\|\phi(h)\|_{H \times H} \leqslant K \|h\|_{H \times H} \text{ and } \|\phi''(h)\|_{L(H \times H)} \leqslant K,$$

for some constant $K > 0$. Hence, we can choose $h_1(x) = Kx$ and $h_2(x) = K$ in Theorem 3.15. Now we are able to write down the product formula

PROPOSITION 4.1 (product formula). *Let $F^A, F^B, G_0^A, G_0^B, G_1^A$, and G_1^B fulfill the assumptions of Theorem 3.15. Furthermore, let X be the mild solution of $(S - ACP_X)$ and Y be the mild solution of $(S - ACP_Y)$. Finally, assume that*

$$\int_U \|x\|^4 \nu(dx) < \infty.$$

Then, the following product formula holds

$$
\begin{aligned}
&\langle(\lambda - A)^{-1}X(t), (\lambda - B)^{-1}Y(t)\rangle_H \\
&\quad = \langle(\lambda - A)^{-1}x, (\lambda - B)^{-1}y\rangle_H \\
&\qquad + \int_0^t \langle A^*(\lambda - A^*)^{-1}(\lambda - B)^{-1}Y(s-), X(s-)\rangle_H \\
&\qquad\qquad\qquad + \langle B^*(\lambda - B^*)^{-1}(\lambda - A)^{-1}X(s-), Y(s-)\rangle_H ds \\
&\qquad + \int_0^t \langle(\lambda - A^*)^{-1}(\lambda - B)^{-1}Y(s-), F^A(s, X(s-))\rangle_H \\
&\qquad\qquad\qquad + \langle(\lambda - B^*)^{-1}(\lambda - A)^{-1}X(s-), F^B(s, Y(s-))\rangle_H ds \\
&\qquad + \int_0^t \langle(\lambda - A^*)^{-1}(\lambda - B)^{-1}Y(s-), G_0^A(s, X(s-))dW_{Q_0}(s)\rangle_H \\
&\qquad\qquad\qquad + \int_0^t \langle(\lambda - B^*)^{-1}(\lambda - A)^{-1}X(s-), G_0^B(s, Y(s-))dW_{Q_0}(s)\rangle_H \\
&\qquad + \frac{1}{2}\int_0^t \mathrm{tr}[(\lambda - A^*)^{-1}(\lambda - B)^{-1}G_0^B(s, Y(s-))Q_0 G_0^A(s, X(s-))^*] \\
&\qquad\qquad\qquad + \mathrm{tr}[(\lambda - B^*)^{-1}(\lambda - A)^{-1}G_0^A(s, X(s-))Q_0 G_0^B(s, Y(s-))^*]ds \\
&\qquad + \int_0^t \int_U \langle(\lambda - A^*)^{-1}(\lambda - B)^{-1}Y(s-), G_1^A(s, X(s-))x\rangle_H \\
&\qquad\qquad\qquad + \langle(\lambda - B^*)^{-1}(\lambda - A)^{-1}X(s-), G_1^B(s, Y(s-))x\rangle_H \\
&\qquad\qquad\qquad + \langle(\lambda - A)^{-1}G_1^A(s, X(s-))x, (\lambda - B)^{-1}G_1^B(s, Y(s-))x\rangle_H \tilde{N}(ds, dx) \\
&\qquad + \int_0^t \int_U \langle(\lambda - A)^{-1}G_1^A(s, X(s-))x, (\lambda - B)^{-1}G_1^B(s, Y(s-))x\rangle_H \nu(dx)ds.
\end{aligned}
$$

PROOF.

The product formula follows by applying the transformation formula from Theorem 3.15 to equation $(S - ACP_C)$ and the function ϕ. Observe, that we are considering the noise process

$$\bar{M} = \begin{pmatrix} M \\ M \end{pmatrix} \in U \times U \text{ which has the covariance operator } \bar{Q} = \begin{pmatrix} Q & Q \\ Q & Q \end{pmatrix}.$$

Note, for the Poisson integral in the transformation formula the following equalities hold

$$
\begin{aligned}
\bar{N}(t, A \times B)(\omega) &= \#\{0 \leqslant s \leqslant t : \Delta\bar{M}(s)(\omega) \in A \times B\} \\
&= \#\{0 \leqslant s \leqslant t : \Delta M(s)(\omega) \in A \wedge \Delta M(s)(\omega) \in B\} \\
&= \#\{0 \leqslant s \leqslant t : \Delta M(s)(\omega) \in A \cap B\} \\
&= N(t, A \cap B)(\omega), \qquad (4.1)
\end{aligned}
$$

and

$$\mathbb{1}_{A\times B}(\bar{M}(s)(\omega)) = \mathbb{1}_{A\times B}\begin{pmatrix} M(s)(\omega) \\ M(s)(\omega) \end{pmatrix} = \mathbb{1}_{A\cap B}(M(s)(\omega)),$$

for all $A, B \in \mathscr{B}(U\backslash\{0\})$ bounded below. Hence, we find

$$\int_{A\times B} f(x_1,x_2)\bar{N}(t,dx)(\omega) = \int_{A\cap B} f(y,y)N(t,dy)(\omega),$$

for all $A, B \in \mathscr{B}(U\backslash\{0\})$ bounded below. For the compensation integral we use (4.1) in order to find

$$\bar{\nu}(A\times B) = \mathbb{E}[\bar{N}(1,A\times B)] = \mathbb{E}[N(1,A\cap B)] = \nu(A\cap B),$$

for all $A, B \in \mathscr{B}(U)$. Therefore, we conclude

$$\int_{A\times B} f(x_1,x_2)\bar{\nu}(dx) = \int_{A\cap B} f(y,y)\nu(dy),$$

for all $A, B \in \mathscr{B}(U)$.

□

4.2. Filtering problem

We consider the linear filtering problem for the infinite dimensional signal process X and the finite dimensional observation process Y defined by

$$\begin{cases} dX(t) &=& AX(t-)dt + G^X(t)dM(t), \quad t\in[0,T],\\ X(0) &=& 0, \end{cases}$$

in H, and

$$\begin{cases} dY(t) &=& B(t)(\lambda-A)^{-2}X(t-)dt + G^Y(t)dW(t), \quad t\in[0,T],\\ Y(0) &=& 0, \end{cases}$$

in $\mathbb{R}^n$, where $(\lambda-A)^{-2} := (\lambda-A)^{-1}(\lambda-A)^{-1}$, and we assume the following

- H is a separable Hilbert space;
- A generates a C_0-semigroup of generalized contraction in H, with $\langle Ah,h\rangle_H \leqslant \gamma_0\|h\|_H^2$ for some fixed $\gamma_0\in\mathbb{R}$ and for all $h\in\mathscr{D}(A)$;
- M is a U-valued square integrable Lévy martingale, where U is a separable Hilbert space, with decomposition $M = W_{Q_0} + M_{Q_1}$ in the continuous part W_{Q_0} and the pure jump part M_{Q_1}, with covariance operators Q_0 and Q_1 respectively;
- $G^X(t)dM(t) = G_0^X(t)dW_{Q_0}(t) + G_1^X(t)dM_{Q_1}(t)$ for all $t\in[0,T]$, where $G_0^X : [0,T]\to L(U,H)$, such that $t\mapsto G_0^X(t)u$ is $\mathscr{B}([0,T])/\mathscr{B}(H)$-measurable for all $u\in U$ and
$$\|G_0^X(t)\|_{L(U,H)} \leqslant C < \infty \text{ for all } t\in[0,T];$$
$G_1^X : [0,T]\to L(U,H)$, such that $t\mapsto G_1^X(t)u$ is $\mathscr{B}([0,T])/\mathscr{B}(H)$-measurable for all $u\in U$ and
$$\|G_1^X(t)\|_{L(U,H)} \leqslant C < \infty \text{ for all } t\in[0,T];$$
- $B : [0,T]\to L(H,\mathbb{R}^n)$, such that $t\mapsto B(t)h$ is $\mathscr{B}([0,T])/\mathscr{B}(\mathbb{R}^n)$-measurable for all $h\in H$ and
$$\|B(t)\|_{L(H,\mathbb{R}^n)} \leqslant C < \infty \text{ for all } t\in[0,T] \text{ and } B\in W_2^1([0,T];L(H,\mathbb{R}^n));$$

- W is an m-dimensional Brownian motion independent of W_{Q_0} and M_{Q_1};
- $G^Y : [0,T] \to L(R^m, \mathbb{R}^n)$, such that $t \mapsto G^Y(t)x$ is $\mathscr{B}([0,T])/\mathscr{B}(\mathbb{R}^n)$-measurable for all $x \in \mathbb{R}^m$ and

$$\|G^Y(t)\|_{L(\mathbb{R}^m,\mathbb{R}^n)} \leqslant C < \infty \text{ and } \langle G^Y(t)G^Y(t)^* x, x\rangle_{\mathbb{R}^n} \geqslant C\|x\|^2_{\mathbb{R}^n} \text{ for all } t \in [0,T];$$

- $\lambda > \gamma_0$ is fixed.

REMARK 4.2. *From Theorem B.53 follows that the equation of the signal process X is well-posed and has a càdlàg solution. Hence, the process of left limits is well-defined and is the predictable version of the solution process. Furthermore, the linear operator $(\lambda - A)^{-1}$ is well-defined, since $\lambda > \gamma_0$. Thus, also the equation for the observation process Y is well-posed.*

Following [**GT08**] we consider the following filtering problem: For fixed $a \in \mathscr{D}(A^*)$ determine a linear unbiased estimation of the type

$$X_u(T) = \int_0^T \langle u(s), dY(s)\rangle_{\mathbb{R}^n},$$

for $\langle(\lambda - A)^{-1}X(T), (\lambda - A^*)^{-1}a\rangle_H$. That is, find a $u_0 \in L_2([0,T];\mathbb{R}^n) =: \bar{H}$, such that

$$\begin{aligned}\mathbb{E}|\langle(\lambda - A)^{-1}X(T), (\lambda - A^*)^{-1}a\rangle_H - X_{u_0}(T)|^2 \\ = \min_{u\in\bar{H}} \mathbb{E}|\langle(\lambda - A)^{-1}X(T), (\lambda - A^*)^{-1}a\rangle_H - X_u(T)|^2, \quad (4.2)\end{aligned}$$

and

$$\mathbb{E}\langle(\lambda - A)^{-1}X(T), (\lambda - A^*)^{-1}a\rangle_H = \mathbb{E}X_{u_0}(T).$$

Then, we can formulate the following Theorem.

THEOREM 4.3. *The solution of filtering problem* (4.2) *is equivalent to the solution of the deterministic control problem*

$$\min_{u\in\bar{H}} F(u),$$

where the functional F is given by

$$\begin{aligned}F(u) := &\int_0^T \|Q_0^{1/2}G_0(s)^*(\lambda - A)^{-2}z_u(s)\|^2_U ds \\ &+ \int_0^T \|Q_1^{1/2}G_1(s)^*(\lambda - A)^{-2}z_u(s)\|^2_U ds \\ &+ \int_0^T \|G^Y(s)^*u(s)\|^2_{\mathbb{R}^m} ds,\end{aligned}$$

with z_u defined by

$$z_u := T^*(T-t)a - \int_t^T T^*(s-t)B^*(s)u(s)ds.$$

Hence, the filtering problem (4.2) *has a unique solution.*

PROOF.

The proof for the equivalents of the filtering problem to the deterministic control problem is split into four parts. First, we perform an approximation of the control u. Then, we introduce some auxiliary problems and estimates. In the third step we deduce the equivalence for the approximation of u. In the final step we take the limit.

Step 1: Approximation of u

Our goal is to approximate each $u \in \bar{H} = L_2([0,T];\mathbb{R}^n)$ by a sequence of smoother functions. Therefore, consider the linear operator $(C, \mathscr{D}(C))$ defined by

$$Cv = \frac{d}{d\sigma} v \text{ for } v \in \mathscr{D}(C) := \{v \in W_2^1([0,T];\mathbb{R}^n) : v(T) = 0\}.$$

It is well-known that the operator $(C, \mathscr{D}(C))$ generates the C_0-semigroup $\{S(t)\}_{t \geqslant 0}$ of contraction in $\bar{H}$, given by

$$(S(t)v)(s) := \begin{cases} v(s+t) & s+t \leqslant T, \\ 0 & s+t > T, \end{cases}$$

for $v \in \bar{H}$. Hence, by the Hille-Yosida Theorem, it follows that

$$\|\gamma(\gamma - C)^{-1}\|_{L(\bar{H})} \leqslant 1 \text{ for all } \gamma > 0.$$

Thus, we define for $u \in \bar{H} = L_2([0,T];\mathbb{R}^n)$ and $n \in \mathbb{N}$

$$u_n := n(n - C)^{-1} u.$$

Therefore, the sequence $(u_n)_{n \in \mathbb{N}}$ has the following properties

- $u_n \in \mathscr{D}(C) \subset W_2^1([0,T];\mathbb{R}^n)$,
- $\|u_n\|_{\bar{H}} \leqslant \|n(n-C)^{-1}\|_{L(\bar{H})} \|u\|_{\bar{H}} \leqslant \|u\|_{\bar{H}}$ for all $n \in \mathbb{N}$,
- by Corollary A.21 (ii) follows $u_n \to u$ in $\bar{H} = L_2([0,T];\mathbb{R}^n)$ for $n \to \infty$.

Step 2: Auxiliary problems and estimates

In order to show the equivalence of the filtering problem and the deterministic control problem, we have to introduce some auxiliary problems. Therefore, let Z_n be the mild solution of

$$\begin{cases} \frac{dZ_n}{dt}(t) &= A^* Z_n(t) - B(T-t)^* u_n(T-t), \quad t \in [0,T], \\ Z_n(0) &= a, \end{cases}$$

for all $n \in \mathbb{N}$. Z is defined to be the mild solution of

$$\begin{cases} \frac{dZ}{dt}(t) &= A^* Z(t) - B(T-t)^* u(T-t), \quad t \in [0,T], \\ Z(0) &= a. \end{cases}$$

Since H is a Hilbert space, it follows from [**EN00**, Proposition in I.5.b and Subsection II.2.a. 2.5] that A^* is the generator of the C_0-semigroup $\{T(t)^*\}_{t \geqslant 0}$. Hence, we have

$$Z_n(t) = T(t)^* a - \int_0^t T(t-s)^* B(T-s)^* u_n(T-s) ds,$$

and

$$Z(t) = T(t)^* a - \int_0^t T(t-s)^* B(T-s)^* u(T-s) ds.$$

Recall that $u_n \in W_2^1([0,T];\mathbb{R}^n)$ for each $n \in \mathbb{N}$ and that $B \in W_2^1([0,T];L(H,\mathbb{R}^n))$. In particular, it follows that $B(T-t)^*u_n(T-t)$ is differentiable a.e. on [0,T] and its derivative is in $L_1([0,T];H)$. Thus, by [**Paz83**, Chapter 4 Corollary 2.10] we find that Z_n has a strong solution, with $Z_n(t) \in \mathscr{D}(A^*)$ a.e. on [0,T] for all $n \in \mathbb{N}$. Next, we show that Z_n approximates Z pointwise in H

$$\begin{aligned}\|Z_n(t)-Z(t)\|_H &= \left\| \int_0^t T(t-s)^*B(T-s)^*(u_n-u)(T-s)ds \right\|_H \\ &\leqslant e^{\gamma_0 T}\|B\|_{L_\infty([0,T];L(H,\mathbb{R}^n))}\int_0^T \|(u_n-u)(s)\|_{\mathbb{R}^n}ds \\ &\leqslant e^{\gamma_0 T}\|B\|_{L_\infty([0,T];L(H,\mathbb{R}^n))}\sqrt{T}\|u_n-u\|_{L_2([0,T],\mathbb{R}^n)} \\ &\longrightarrow 0,\end{aligned}$$

for $n \to \infty$ and all $t \in [0,T]$. Now we perform a time reflection. Therefore, we define

$$z_n(t) := Z_n(T-t) \text{ and } z(t) := Z(T-t).$$

Hence, we have the following representations

$$z_n(t) = T(T-t)^*a - \int_t^T T(s-t)^*B(s)^*u_n(s)ds,$$

for all $n \in \mathbb{N}$ and

$$z(t) = T(T-t)^*a - \int_t^T T(s-t)^*B(s)^*u(s)ds = z_u(t).$$

Obviously, the properties from Z_n and Z carry over to z_n and z_u. In particular, we have $z_n \in \mathscr{D}(A^*)$ for all $n \in \mathbb{N}$ and $\|z_n(t)-z_u(t)\|_H \to 0$ for $n \to \infty$ for all $t \in [0,T]$. Note that z_n fulfills the following equation

$$\begin{cases} \dfrac{dz_n}{dt}(t) &= -A^*z_n(t) - B(t)^*u_n(t), \quad t \in [0,T], \\ Z_n(T) &= a. \end{cases}$$

Before we move on to the next step, we show that $z_n, z_u \in L_2([0,T];H)$ and $z_n \to z_u$ in $L_2([0,T];H)$ for $n \to \infty$. Therefore, we estimate

$$\begin{aligned}\|z_n(t)\|_H^2 &\leqslant 2\|T(T-t)^*a\|_H^2 + 2\left\| \int_t^T T(s-t)^*B(s)^*u_n(s)ds \right\|_H^2 \\ &\leqslant 2e^{2\gamma_0 T}\|a\|_H^2 + 2\left(\int_t^T \|T(s-t)^*B(s)^*u_n(s)\|_H ds \right)^2 \\ &\leqslant 2e^{2\gamma_0 T}\|a\|_H^2 + 2e^{2\gamma_0 T}\|B\|_{L_\infty([0,T];L(H,\mathbb{R}^n))}^2 \left(\int_t^T \|u_n(s)\|_{\mathbb{R}^n} \right)^2 \\ &\leqslant 2e^{2\gamma_0 T}\|a\|_H^2 + 2e^{2\gamma_0 T}\|B\|_{L_\infty([0,T];L(H,\mathbb{R}^n))}^2 (T-t)\|u_n\|_{L_2([t,T];\mathbb{R}^n)}^2 \\ &\leqslant 2e^{2\gamma_0 T}\|a\|_H^2 + 2e^{2\gamma_0 T}\|B\|_{L_\infty([0,T];L(H,\mathbb{R}^n))}^2 T\|u\|_{L_2([0,T];\mathbb{R}^n)}^2.\end{aligned}$$

Note, that the same estimate works for z_u instead of z_n. Thus, we find

$$\|z_n\|_{L_2([0,T];H)}, \|z_u\|_{L_2([0,T];H)} \leqslant C_z < \infty.$$

Finally, we look at the convergence of z_n in $L_2([0,T];H)$. We already know that z_n converges pointwise to z_u. Hence, we are looking for a majorant

$$\begin{aligned}\|z_n(t)-z_u(t)\|_H^2 &= \left\| \int_t^T T(s-t)^* B(s)^*(u_n-u)(s)ds \right\|_H^2 \\ &\leqslant e^{2\gamma_0 T}\|B\|^2_{L_\infty([0,T];L(H,\mathbb{R}^n))} T\|u_n-u\|^2_{L_2([0,T];\mathbb{R}^n)} \\ &\leqslant 2e^{2\gamma_0 T}\|B\|^2_{L_\infty([0,T];L(H,\mathbb{R}^n))} T\|u\|^2_{L_2([0,T];\mathbb{R}^n)}.\end{aligned}$$

Obviously, the expression on the right-hand side in integrable over $[0,T]$. By Lebesgue's dominated convergence Theorem it follows

$$\|z_n - z_u\|_{L_2([0,T];H)} \to 0 \text{ for } n \to \infty.$$

Step 3: Equivalence for the approximation of u

In order to show the equivalence we apply the product formula from Proposition 4.1 to X and z_n. Note that the product formula is proven for forward equations. That is why, we apply it to

$$\langle(\lambda - A)^{-1}X(t), (\lambda - A^*)^{-1}z_n(t)\rangle_H.$$

However, we have to plug in the coefficients of z_n in the formula. Hence, we receive

$$\begin{aligned}&\langle(\lambda - A)^{-1}X(t), (\lambda - A^*)^{-1}z_n(t)\rangle_H \\ &= \int_0^t \langle A^*(\lambda - A^*)^{-2}z_n(s), X(s-)\rangle_H - \langle A(\lambda - A)^{-2}X(s-), z_n(s)\rangle_H ds \\ &\quad + \int_0^t \langle(\lambda - A)^{-2}X(s-), B(s)^*u_n(s)\rangle_H ds + \int_0^t \langle(\lambda - A^*)^{-2}z_n(s), G_0^X(s)dW_{Q_0}(s)\rangle_H \\ &\quad + \int_0^t\int_U \langle(\lambda - A)^{-1}G_1^X(s)x, (\lambda - A^*)^{-1}z_n(s)\rangle_H \tilde{N}(ds,dx). \end{aligned} \tag{4.3}$$

Since $z_n(s) \in \mathscr{D}(A^*)$ for a.e. $s \in [0,T]$, we find for all $t \in [0,T]$

$$\begin{aligned}\int_0^t \langle A(\lambda - A)^{-2}X(s-), z_n(s)\rangle_H ds &= \int_0^t \langle X(s-), (\lambda - A)^{-2}A^* z_n(s)\rangle_H ds \\ &= \int_0^t \langle X(s-), A^*(\lambda - A^*)^{-2}z_n(s)\rangle_H ds.\end{aligned}$$

The second equality holds, since the generator commutes with its resolvent. Therefore, (4.3) reduces to

$$\begin{aligned}&\langle(\lambda - A)^{-1}X(t), (\lambda - A^*)^{-1}z_n(t)\rangle_H \\ &= \int_0^t \langle(\lambda - A)^{-2}X(s-), B(s)^*u_n(s)\rangle_H ds + \int_0^t \langle(\lambda - A^*)^{-2}z_n(s), G_0^X(s)dW_{Q_0}(s)\rangle_H \\ &\quad + \int_0^t\int_U \langle(\lambda - A)^{-1}G_1^X(s)x, (\lambda - A^*)^{-1}z_n(s)\rangle_H \tilde{N}(ds,dx).\end{aligned}$$

In particular, we receive for $t = T$

$$
\begin{aligned}
&\langle(\lambda - A)^{-1}X(T), (\lambda - A^*)^{-1}a\rangle_H \\
&= \int_0^T \langle(\lambda - A)^{-2}X(s-), B(s)^* u_n(s)\rangle_H ds + \int_0^T \langle(\lambda - A^*)^{-2} z_n(s), G_0^X(s) dW_{Q_0}(s)\rangle_H \\
&\quad + \int_0^T \int_U \langle(\lambda - A)^{-1} G_1^X(s)x, (\lambda - A^*)^{-1} z_n(s)\rangle_H \tilde{N}(ds, dx).
\end{aligned}
$$

Hence, by subtracting $X_{u_n}(T)$ we find

$$
\begin{aligned}
&\langle(\lambda - A)^{-1}X(T), (\lambda - A^*)^{-1}a\rangle_H - X_{u_n}(T) \\
&= \int_0^T \langle G_0^X(s)^*(\lambda - A^*)^{-2} z_n(s), dW_{Q_0}(s)\rangle_U - \int_0^T \langle G^Y(s)^* u_n(s), dW(s)\rangle_{\mathbb{R}^m} \\
&\quad + \int_0^T \int_U \langle G_1^X(s)^*(\lambda - A^*)^{-2} z_n(s), x\rangle_U \tilde{N}(ds, dx).
\end{aligned}
$$

The unbiased property is fulfilled, since the expectations of the stochastic integrals are zero. From the isometries of the stochastic integrals we receive

$$
\begin{aligned}
&\mathbb{E}|\langle(\lambda - A)^{-1}X(T), (\lambda - A^*)^{-1}a\rangle_H - X_{u_n}(T)|^2 \\
&= \int_0^T \sum_{j=1}^{\infty} \lambda_{0j} \langle G_0^X(s)^*(\lambda - A^*)^{-2} z_n(s), e_j\rangle_U^2 ds \\
&\quad + \int_0^T \int_U \langle G_1^X(s)^*(\lambda - A^*)^{-2} z_n(s), x\rangle_U^2 \nu(dx) ds + \int_0^T \|G^Y(s)^* u_n(s)\|_{\mathbb{R}^m}^2 ds \\
&= \int_0^T \|Q_0^{1/2} G_0^X(s)^*(\lambda - A^*)^{-2} z_n(s)\|_U^2 ds \\
&\quad + \int_0^T \|Q_1^{1/2} G_1^X(s)^*(\lambda - A^*)^{-2} z_n(s)\|_U^2 ds + \int_0^T \|G^Y(s)^* u_n(s)\|_{\mathbb{R}^m}^2 ds, \qquad (4.4)
\end{aligned}
$$

where λ_{0j} and e_j are the eigenvalues and eigenvectors of Q_0. The equality for the second integral follows from Theorem B.36.

Step 4: Taking the limit $n \to \infty$

In order to show that equality (4.4) also holds for $u \in L_2([0,T]; \mathbb{R}^n)$ we take the limit $n \to \infty$ on both sides. We start with right-hand side. For the last term we assume without loss of generality that

$$
\int_0^T \|G^Y(s)^* u_n(s)\|_{\mathbb{R}^m}^2 ds \geqslant \int_0^T \|G^Y(s)^* u(s)\|_{\mathbb{R}^m}^2 ds. \qquad (4.5)
$$

In order to see that this assumption is indeed without loos of generality we first perform the estimate. After this step it will be almost obvious. Thus, we estimate

$$
\begin{aligned}
&\int_0^T \|G^Y(s)^* u_n(s)\|_{\mathbb{R}^m}^2 - \|G^Y(s)^* u(s)\|_{\mathbb{R}^m}^2 ds \\
&\quad = \int_0^T \|G^Y(s)^*(u_n - u)(s) + G^Y(s)^* u(s)\|_{\mathbb{R}^m}^2 - \|G^Y(s)^* u(s)\|_{\mathbb{R}^m}^2 ds
\end{aligned}
$$

$$
\begin{aligned}
&\leqslant \int_0^T \|G^Y(s)^*(u_n-u)(s)\|_{\mathbb{R}^m}^2 ds + 2\int_0^T \|G^Y(s)^*(u_n-u)(s)\|_{\mathbb{R}^m}\|G^Y(s)^*u(s)\|_{\mathbb{R}^m}\\
&\leqslant \|G^Y\|^2_{L_\infty([0,T];L(\mathbb{R}^n,\mathbb{R}^m))}\|u_n-u\|^2_{L_2([0,T];\mathbb{R}^n)}\\
&\qquad\qquad\qquad\qquad + 2\|G^{Y*}(u_n-u)\|_{L_2([0,T];\mathbb{R}^m)}\|G^{Y*}u\|_{L_2([0,T];\mathbb{R}^m)}\\
&\leqslant \|G^Y\|^2_{L_\infty([0,T];L(\mathbb{R}^n,\mathbb{R}^m))}\big(\|u_n-u\|^2_{L_2([0,T];\mathbb{R}^n)} + 2\|u\|_{L_2([0,T];\mathbb{R}^n)}\|u_n-u\|_{L_2([0,T];\mathbb{R}^n)}\big)\\
&\longrightarrow 0, \qquad (4.6)
\end{aligned}
$$

for $n \to \infty$. Note, that if we had assumed

$$\int_0^T \|G^Y(s)^*u_n(s)\|^2_{\mathbb{R}^m}ds \leqslant \int_0^T \|G^Y(s)^*u(s)\|^2_{\mathbb{R}^m}ds$$

instead of (4.5), the last line of the estimate above would be

$$\|G^Y\|^2_{L_\infty([0,T];L(\mathbb{R}^n,\mathbb{R}^m))}\big(\|u_n-u\|^2_{L_2([0,T];\mathbb{R}^n)} + 2\|u_n\|_{L_2([0,T];\mathbb{R}^n)}\|u_n-u\|_{L_2([0,T];\mathbb{R}^n)}\big)$$

Since $\|u_n\|_{L_2([0,T];\mathbb{R}^n)} \leqslant \|u\|_{L_2([0,T];\mathbb{R}^n)}$, we end up with the same estimate. Hence, (4.5) is indeed without loss of generality.

The first two terms of the right-hand side of (4.4) have the same structure. That is why we omit the indices and this way treat them both at the same time. Furthermore, we introduce $V : [0,T] \to L(H,U)$ defined by

$$V(t) = Q^{1/2}G^X(t)^*(\lambda - A)^{-2}.$$

It is easy to see that $V \in L_\infty([0,T];L(H,U))$, since

$$\|V(t)\|_{L(H,U)} \leqslant \|(\lambda-A)^{-1}\|^2_{L(H)}\|G^X\|_{L_\infty([0,T];L(U,H))}\|Q^{1/2}\|_{L(U)} < \infty.$$

With this notation in place we assume without loos of generality (for the same reasons as above) that

$$\int_0^T \|V(s)z_n(s)\|^2_U ds \geqslant \int_0^T \|V(s)z_u(s)\|^2_U ds.$$

Then, we estimate

$$
\begin{aligned}
&\int_0^T (\|V(s)z_n(s)\|^2_U - \|V(s)z_u(s)\|^2_U)ds\\
&\quad\leqslant \int_0^T \|V(s)(z_n-z_u)(s)\|^2_U ds + 2\int_0^T \|V(s)(z_n-z_u)(s)\|_U\|V(s)z_u(s)\|_U ds\\
&\quad\leqslant \|V\|^2_{L_\infty([0,T];L(H,U))}\big(\|z_n-z_u\|^2_{L_2([0,T];H)} + 2C_z\|z_n-z_u\|_{L_2([0,T];H)}\big)\\
&\quad\longrightarrow 0,
\end{aligned}
$$

for $n \to \infty$. For the proof of the convergence of the term on left-hand side of (4.4) we introduce the real valued random variable J, by

$$J := \langle(\lambda-A)^{-1}X(T), (\lambda-A^*)^{-1}a\rangle_H.$$

For the estimate we assume without loos of generality (for the same reasons as above) that

$$\mathbb{E}|J - X_{u_n}(T)|^2 \geqslant \mathbb{E}|J - X_u(T)|^2.$$

Then, we estimate

$$\begin{aligned}
&\mathbb{E}(|J - X_{u_n}(T)|^2 - |J - X_u(T)|^2) \\
&\quad \leqslant \mathbb{E}|X_{u_n}(T) - X_u(T)|^2 + 2\mathbb{E}(|X_{u_n}(T) - X_u(T)||J - X_u(T)|) \\
&\quad \leqslant \mathbb{E}|X_{u_n}(T) - X_u(T)|^2 + 2\sqrt{\mathbb{E}|X_{u_n}(T) - X_u(T)|^2}\sqrt{\mathbb{E}|J - X_u(T)|^2}.
\end{aligned}$$

If we can show that $\mathbb{E}|X_{u_n}(T) - X_u(T)|^2$ converges to zero and $\mathbb{E}|J - X_u(T)|^2$ is bounded, then the desired convergence follows. First, we look at the convergence of $\mathbb{E}|X_{u_n}(T) - X_u(T)|^2$ and estimate

$$\begin{aligned}
&\mathbb{E}|X_{u_n}(T) - X_u(T)|^2 \\
&\quad = \mathbb{E}\bigg|\int_0^T \langle (u_n - u)(s), B(s)(\lambda - A)^{-2}X(s-)\rangle_{\mathbb{R}^n} ds + \int_0^T \langle (u_n - u)(s), G^Y(s)dW(s)\rangle_{\mathbb{R}^n}\bigg|^2 \\
&\quad \leqslant 2\mathbb{E}\bigg|\int_0^T \langle (u_n - u)(s), B(s)(\lambda - A)^{-2}X(s-)\rangle_{\mathbb{R}^n} ds\bigg|^2 \qquad (4.7) \\
&\qquad\qquad + 2\mathbb{E}\bigg|\int_0^T \langle (u_n - u)(s), G^Y(s)dW(s)\rangle_{\mathbb{R}^n}\bigg|^2. \qquad (4.8)
\end{aligned}$$

For the term (4.7) the convergence follows form the following estimate

$$\begin{aligned}
&\mathbb{E}\bigg|\int_0^T \langle (u_n - u)(s), B(s)(\lambda - A)^{-2}X(s-)\rangle_{\mathbb{R}^n} ds\bigg|^2 \\
&\quad \leqslant \mathbb{E}\bigg(\int_0^T \|(u_n - u)(s)\|_{\mathbb{R}^n} \|B(s)(\lambda - A)^{-2}X(s-)\|_{\mathbb{R}^n} ds\bigg)^2 \\
&\quad \leqslant \mathbb{E}(\|u_n - u\|^2_{L_2([0,T];\mathbb{R}^n)} \|B(\lambda - A)^{-2}X\|^2_{L_2([0,T];\mathbb{R}^n)}) \\
&\quad \leqslant \|B\|^2_{L_\infty([0,T];L(H,\mathbb{R}^n))} \|(\lambda - A)^{-1}\|^4_{L(H)} T \|X\|_{T,H} \|u_n - u\|_{L_2([0,T];\mathbb{R}^n)} \\
&\quad \longrightarrow 0,
\end{aligned}$$

for $n \to \infty$. Furthermore, note that (4.8) is equal to $2\int_0^T \|G^Y(s)^*(u_n - u)(s)\|^2_{\mathbb{R}^m} ds$. We have already shown in (4.6), that this term converges to zero. Hence, all what is left to show is the boundedness of $\mathbb{E}|J - X_{u_n}(T)|^2$. Note that we consider X_{u_n} here instead of X_u as stated in the estimate above. We are doing this to guarantee the generality claimed at the beginning. It will turn out that the same estimate also holds for $\mathbb{E}|J - X_u(T)|^2$. We begin by estimating

$$\mathbb{E}|J - X_{u_n}(T)|^2 \leqslant 2\mathbb{E}|J|^2 + 2\mathbb{E}|X_{u_n}(T)|^2.$$

For the first term a straightforward estimate yields

$$\begin{aligned}
\mathbb{E}|J|^2 &= \mathbb{E}|\langle (\lambda - A)^{-1}X(T), (\lambda - A^*)^{-1}a\rangle_H|^2 \leqslant \mathbb{E}(\|(\lambda - A)^{-1}X(T)\|^2_H \|(\lambda - A^*)^{-1}a\|^2) \\
&\leqslant \|(\lambda - A)^{-1}\|^4_{L(H)} \|a\|^2_H \mathbb{E}\|X(T)\|^2_H \leqslant \|(\lambda - A)^{-1}\|^4_{L(H)} \|a\|^2_H \|X\|_{T,H} < \infty.
\end{aligned}$$

For the second term we find

$$\begin{aligned}\mathbb{E}|X_{u_n}(T)|^2 &= \mathbb{E}\bigg|\int_0^T \langle u_n(s), B(s)(\lambda - A)^{-2}X(s-)\rangle_{\mathbb{R}^n} ds + \int_0^T \langle u_n(s), G^Y(s)dW(s)\rangle_{\mathbb{R}^n}\bigg|^2 \\ &\leqslant 2\mathbb{E}\bigg|\int_0^T \langle u_n(s), B(s)(\lambda - A)^{-2}X(s-)\rangle_{\mathbb{R}^n} ds\bigg|^2 + 2\int_0^T \|G^Y(s)^* u_n(s)\|^2_{\mathbb{R}^m} ds.\end{aligned}$$

We consider both terms individually and start with the first one

$$\begin{aligned}&\mathbb{E}\bigg|\int_0^T \langle u_n(s), B(s)(\lambda - A)^{-2}X(s-)\rangle_{\mathbb{R}^n} ds\bigg|^2 \\ &\leqslant \mathbb{E}\bigg(\int_0^T \|u_n(s)\|_{\mathbb{R}^n} \|B(s)(\lambda - A)^{-2}X(s-)\|_{\mathbb{R}^n} ds\bigg)^2 \\ &\leqslant \mathbb{E}\|u_n\|^2_{L_2([0,T];\mathbb{R}^n)} \|B(\lambda - A)^{-2}X\|^2_{L_2([0,T];\mathbb{R}^n)} \\ &\leqslant \|B\|^2_{L_\infty([0,T];L(H,\mathbb{R}^n))} \|(\lambda - A)^{-1}\|^4_{L(H)} \|u_n\|^2_{L_2([0,T];\mathbb{R}^n)} \mathbb{E}\int_0^T \|X(s-)\|^2_H ds \\ &\leqslant \|B\|^2_{L_\infty([0,T];L(H,\mathbb{R}^n))} \|(\lambda - A)^{-1}\|^4_{L(H)} \|u\|^2_{L_2([0,T];\mathbb{R}^n)} T\|X\|_{T,H} < \infty.\end{aligned}$$

Finally, the second term can be bounded by

$$\begin{aligned}\int_0^T \|G^Y(s)^* u_n(s)\|^2_{\mathbb{R}^m} ds &\leqslant \|G^Y\|^2_{L_\infty([0,T];L(\mathbb{R}^m,\mathbb{R}^n))} \|u_n\|^2_{L_2([0,T];\mathbb{R}^n)} \\ &\leqslant \|G^Y\|^2_{L_\infty([0,T];L(\mathbb{R}^m,\mathbb{R}^n))} \|u\|^2_{L_2([0,T];\mathbb{R}^n)} < \infty.\end{aligned}$$

Hence, we have shown the convergence of the right and left-hand side of (4.4). Therefore, the filtering problem is equivalent to the deterministic control problem.

The existence and uniqueness follow from classical results for optimal control problems. Since the functional F is strictly convex, continuous, and

$$F(u) \geqslant \int_0^T \|G^Y(s)^* u(s)\|^2_{\mathbb{R}^m} ds \geqslant C\int_0^T \|u(s)\|^2_{\mathbb{R}^n} ds = C\|u\|_{L_2([0,T];\mathbb{R}^n)}.$$

Hence, we have $F(u) \to \infty$ for $\|u\|_{L_2([0,T];\mathbb{R}^n)} \to \infty$. Then, the existence of a solution follows from [**Zei85**, Proposition 38.15 (a) in Section 38.5] and the uniqueness by [**Zei85**, Theorem 38.C in Section 38.4].

□

APPENDIX A

Basics of Semigroup Theory

Here, we provide a brief overview of semigroup theory and its applications to abstract Cauchy problems. A detailed introduction can be found, for example, in [**EN00**] or [**Paz83**].

Definition A.1 (bounded linear operator). *If X and Y are Banach spaces, we denote by $L(X,Y)$ the Banach space of all bounded linear operators from X to Y. If $X = Y$ we set $L(X) := L(X,X)$.*

Definition A.2 (semigroup and infinitesimal generator). *Let X be a Banach space. A one parameter family $T(t)$, $0 \leqslant t < \infty$, of bounded linear operators from X into X is called a semigroup of bounded linear operators on X if*

(i) $T(0) = I$, where I is the identity operator on X,
(ii) $T(t+s) = T(t)T(s)$ for ever $t, s \geqslant 0$ (the semigroup property).

The linear operator defined by

$$\mathscr{D}(A) := \left\{ x \in X : \lim_{t \to 0^+} \frac{T(t)x - x}{t} \quad \text{exists} \right\}$$

and

$$Ax := \lim_{t \to 0^+} \frac{T(t)x - x}{t} \quad \text{for} \quad x \in \mathscr{D}(A)$$

is called the (infinitesimal) generator of the semigroup $(T(t))_{t \geqslant 0}$, $\mathscr{D}(A)$ is the domain of A.

Definition A.3 (C_0-semigroup). *A semigroup $T(t)$, $0 \leqslant t < \infty$, of bounded linear operators on X is called a C_0-semigroup (or strongly continuous semigroup) on X of bounded linear operators if*

$$\lim_{t \to 0^+} T(t)x = x \quad \text{for every} \quad x \in X.$$

Theorem A.4 (properties of the generator of a C_0-semigroup). *The generator of a C_0-semigroup is a closed and densely defined linear operator that determines the semigroup uniquely.*

PROOF.
See [**EN00**, Theorem 1.4 in Section II.1].

PROPOSITION A.5 (bound of C_0-semigroup). *For every C_0-semigroup $T(t)$, $0 \leqslant t < \infty$, there exist constants $w \in \mathbb{R}$ and $M \geqslant 1$, such that*

$$\|T(t)\|_{L(X)} \leqslant Me^{wt}$$

for all $0 \leqslant t < \infty$.

PROOF.

See [**EN00**, Proposition 5.5 in Section I.5.a].

LEMMA A.6 (relations of a C_0-semigroup and its generator). *For the generator $(A, \mathscr{D}(A))$ of a C_0-semigroup $(T(t))_{t\geqslant 0}$, the following properties hold.*

(i) $A : \mathscr{D}(A) \subseteq X \to X$ *is a linear opeator.*

(ii) If $x \in \mathscr{D}(A)$, then $T(t)x \in \mathscr{D}(A)$ and

$$\frac{d}{dt}T(t)x = T(t)Ax = AT(t)x \text{ for all } t \geqslant 0.$$

(iii) For every $t \geqslant 0$ and $x \in X$, one has

$$\int_0^t T(s)xds \in \mathscr{D}(A).$$

(iv) For every $t \geqslant 0$, one has

$$\begin{aligned} T(t)x - x &= A\int_0^t T(s)x\ ds \text{ if } x \in X, \\ &= \int_0^t T(s)Ax\ ds \text{ if } x \in \mathscr{D}(A). \end{aligned}$$

PROOF.

See [**EN00**, Lemma 1.3 in Section II.1].

DEFINITION A.7 (C_0-semigroup of contractions). *A C_0-semigroup $T(t)$, where $0 \leqslant t < \infty$, is called a C_0-semigroup of contractions if*

$$\|T(t)\|_{L(X)} \leqslant 1$$

holds for all $0 \leqslant t < \infty$.

DEFINITION A.8 (C_0-semigroup of generalized contractions). *A C_0-semigroup $T(t)$, where $0 \leqslant t < \infty$, is called a C_0-semigroup of generalized contractions if for some $w \in \mathbb{R}$*

$$\|T(t)\|_{L(X)} \leqslant e^{wt}$$

holds for all $0 \leqslant t < \infty$.

THEOREM A.9 (solution of abstract Cauchy problem). *Let $A : \mathscr{D}(A) \subset X \to X$ be a closed operator. For the associated abstract Cauchy problem*

$$(ACP) \qquad \begin{cases} \partial_t u - Au &= 0, \quad \text{for } t \geqslant 0, \\ \\ u(0) &= u_0, \end{cases}$$

we consider the following existence and uniqueness condition

$$(EU) \qquad \begin{cases} \text{For every } u_0 \in \mathscr{D}(A), \text{ there exists} \\ \text{a unique solution } u(\cdot, u_0) \text{ of } (ACP). \end{cases}$$

Then the following properties are equivalent

(1) *A generates a C_0-semigroup on X.*
(2) *A satisfies (EU), has dense domain, and for every sequence $(x_n)_{n\in\mathbb{N}} \subset \mathscr{D}(A)$ satisfying $\lim_{n\to\infty} x_n = 0$, one has $\lim_{n\to\infty} u(t, x_n) = 0$ uniformly in compact intervals $[0, t_0]$.*

PROOF.

See [**EN00**, Theorem 6.7 in Section II.6].

DEFINITION A.10 (well-posedness). *The abstract Cauchy problem (ACP) associated to a closed operator $A : \mathscr{D}(A) \subset X \to X$ is called well-posed if condition (2) in Theorem A.9 holds.*

COROLLARY A.11 (well-posedness of closed operator). *For a closed operator $(A, \mathscr{D}(A))$ with $A : \mathscr{D}(A) \subset X \to X$, the associated abstract Cauchy problem is well-posed if and only if A generates a C_0-semigroup on X.*

DEFINITION A.12 (mild solution of (ACP)). *If the linear operator $(A, \mathscr{D}(A))$ in (ACP) generates a C_0-semigroup and $u_0 \in X$, then the function*

$$u : t \mapsto u(t) := T(t)u_0,$$

is called the mild solution of (ACP).

PROPOSITION A.13. *Let the operator $(A, \mathscr{D}(A))$ in (ACP) generate a C_0-semigroup. The function $u : \mathbb{R}_+ \to X$ is a mild solution of (ACP) if and only if*

- *$\int_0^t u(s)ds \in \mathscr{D}(A)$ for all $t \geqslant 0$, and*
- *$u(t) = A\int_0^t u(s)ds + u_0$.*

PROOF.

See [**EN00**, Proposition 6.4 in Section II.6].

Definition A.14 (dissipative operator). *A linear operator* $(A, \mathscr{D}(A))$ *on a Banach space* X *is called*

(1) *dissipative if* $\|(\lambda - A)x\|_X \geqslant \lambda \|x\|_X$ *for all* $\lambda > 0$ *and* $x \in \mathscr{D}(A)$,
(2) ω*-dissipative if* $A - \omega$ *is dissipative,*

where we write $\lambda - A$ *instead of* $\lambda I - A$. *In a Hilbert space* H *one can define equivalently*

(1) *dissipative if* $\langle Ax, x\rangle_H \leqslant 0$ *for all* $x \in \mathscr{D}(A)$,
(2) ω*-dissipative if* $\langle Ax, x\rangle_H \leqslant \omega\|x\|^2$ *for all* $x \in \mathscr{D}(A)$.

PROOF OF THE EQUIVALENCE.
See [**Paz83**, Theorem 4.2 in Section 1.4]

Proposition A.15 (properties of dissipative operator). *For a dissipative operator* $(A, \mathscr{D}(A))$ *the following properties hold.*

(1) $\lambda - A$ *is injective for all* $\lambda > 0$ *and*
$$\|(\lambda - A)^{-1}z\|_X \leqslant \frac{1}{\lambda}\|z\|_X$$
for all z *in the range* $\operatorname{rg}(\lambda - A) := (\lambda - A)\mathscr{D}(A)$.
(2) $\lambda - A$ *is surjective for some* $\lambda > 0$ *if and only if it is surjective for each* $\lambda > 0$. *In that case, one has* $(0, \infty) \subset \rho(A)$. *Here* $\rho(A)$ *denotes the resolvent set of* A.
(3) A *is closed if and only if the range* $\operatorname{rg}(\lambda - A)$ *is closed for some (hence all)* $\lambda > 0$.
(4) *If* $\operatorname{rg}(A) \subseteq \overline{\mathscr{D}(A)}$, *e.g. if* A *is densely defined, then* A *is closeable. Its closure* $\bar{A}$ *is again dissipative and satisfies* $\operatorname{rg}(\lambda - \bar{A}) = \overline{\operatorname{rg}(\lambda - A)}$ *for all* $\lambda > 0$.

PROOF.
See [**EN00**, Proposition 3.14 in Section II.3.b].

Theorem A.16 (Lumer-Phillips). *Let* A *be a linear operator with dense domain* $\mathscr{D}(A)$ *in* X.

(1) *If* A *is dissipative and there is a* $\lambda_0 > 0$ *such that the range,* $\operatorname{rg}(\lambda_0 - A)$, *of* $\lambda_0 - A$ *is* X, *then* A *is the infinitesimal generator of a* C_0*-semigroup of contractions on* X.
(2) *If* A *is the infinitesimal generator of a* C_0*-semigroup of contraction on* X, *then* $\operatorname{rg}(\lambda - A) = X$ *for all* $\lambda > 0$ *and* A *is dissipative.*

PROOF.
See [**Paz83**, Theorem 4.3 in Section 1.4].

Remark A.17. *If the linear operator* $(A, \mathscr{D}(A))$ *generates a* C_0*-semigroup of generalized contractions* T *with* $\|T(t)\|_{L(X)} \leqslant e^{wt}$ *for some* $w \in \mathbb{R}$ *it follows by rescaling that the semigroup* $S(t) := e^{-wt}T(t)$ *is a* C_0*-semigroup of contractions and is generated by* $A - w$ *with domain* $\mathscr{D}(A)$ *(see* [**EN00**, Section II.2.a. 2.2] *for more detail). This means Theorem A.16 also applies for generalized contractions semigroups respecting the parameter shift.*

Finally we consider an approximation technique. We are interested in the Yosida approximation. Therefore, we need the following result.

Lemma A.18. *Let $(A, \mathscr{D}(A))$ be a closed, densely defined operator. Suppose there exist $w \in \mathbb{R}$ and $M > 0$, such that $[w, 0) \subset \rho(A)$ (the resolvent set of A) and $\|\lambda(\lambda - A)^{-1}\| \leqslant M$ for all $\lambda \geqslant w$. Then the following convergence statements hold for $\lambda \to \infty$.*

(1) $\lambda(\lambda - A)^{-1}x \to x$ for all $x \in X$.
(2) $\lambda A(\lambda - A)^{-1}x = \lambda(\lambda - A)^{-1}Ax \to Ax$ for all $x \in \mathscr{D}(A)$.

PROOF.
See [**EN00**, Lemma 3.4 in Section II.3.a].

Remark A.19. *Note that if $(A, \mathscr{D}(A))$ generates a C_0-semigroup we can always apply Lemma A.18. This follows from the general Hille-Yosida generation theorem (see* [**EN00**, Theorem 3.8 in II.3.a]*).*

We apply the Yosida approximation to elements of the Hilbert space H and to generators of C_0-semigroups. Therefore, we gather those results here.

Definition A.20 (Yosida approximation of an element in H). *For the generator A of a C_0-semigroup $(T(t))_{t\geqslant 0}$ with $\|T(t)\|_{L(H)} \leqslant Me^{\lambda t}$ for all $t \geqslant 0$, where $M \geqslant 1$ and $\lambda \in \mathbb{R}$, we define the Yosida approximation $R(n)h$ of an element $h \in H$ for all $n \in \mathbb{N}$ with $n > \lambda$ by*

$$R(n)h := n(n - A)^{-1}h.$$

Corollary A.21 (properties of the Yosida approximation of an element in H). *Let $R(n)h$ be the Yosida approximation of $h \in H$, then*

(i) $R(n)h \in \mathscr{D}(A)$ for all $n > \lambda$ and all $h \in H$,
(ii) $R(n)h \to h$ for $n \to \infty$ and all $h \in H$,
(iii) $R(n) \in L(H)$ for $n > \lambda$ and $\|R(n)\|_{L(H)} \leqslant C_R < \infty$, where C_R is independent of n,
(iv) $T(t)R(n) = R(n)T(t)$ for all $n > \lambda$ and $t \geqslant 0$.

PROOF.
Property *(i)* follows from the definition of the resolvent operator. Property *(ii)* was shown in Lemma A.18 *(1)*. The first part of property *(iii)* follows directly from the definition of the resolvent operator and the second by the uniform boundedness principle. Property *(iv)* follows from the integral representation of the resolvent operator (see [**EN00**, Theorem 1.10 in II.1]), that is for $h \in H$

$$(n - A)^{-1}h = \int_0^\infty e^{-ns}T(s)hds.$$

□

Definition A.22 (Yosida approximation of the generator of a contraction semigroup). *For a generator A of a C_0-semigroup $(T(t))_{t\geqslant 0}$ with $\|T(t)\|_{L(H)} \leqslant 1$ for all $t \geqslant 0$ we define the Yosida approximation A_n of A by*

$$A_n := nA(n-A)^{-1},$$

for $n \in \mathbb{N}$.

Corollary A.23 (Yosida approximation for generator; contraction case). *Let A_n be the Yosida approximation of the generator A of a contraction semigroup, then*

(i) $A_n = n^2(n-A)^{-1} - n$ *and* $A_n \in L(H)$ *for all* $n \in \mathbb{N}$,

(ii) $A_n h \to Ah$ *for* $n \to \infty$ *and all* $h \in \mathscr{D}(A)$,

(iii) A_n *generates the uniformly continuous semigroup* $T_n(t) := e^{tA_n}$, *such that*

$$\|T_n(t)\|_{L(H)} \leqslant 1 \text{ for all } t \geqslant 0,$$

(iv) $T_n(t)h \to T(t)h$ *for* $n \to \infty$ *for all* $t \geqslant 0$, *all* $h \in H$, *and uniformly on each interval* $[0, t_0]$.

PROOF.

Property *(i)* follows from [**Paz83**, Theorem 3.1 identity (3.4)]. Property *(ii)* was shown in Lemma A.18 *(2)*. Properties *(iii)* and *(iv)* are shown in [**EN00**, Proof of Theorem 3.5 in II.3.a].

Since Corollary A.23 only holds for generators of contraction semigroups, our goal is to generalize those results to an arbitrary C_0-semigroup. As it turns out one can reduce the general case, where A generates the C_0-semigroup $(T(t))_{t\geqslant 0}$ with $\|T(t)\|_{L(H)} \leqslant Me^{\lambda t}$, to the case of a contraction semigroup by the same techniques applied in the generalization of the Hille-Yosida generation theorem (see the proofs of Corollary 3.6 and Theorem 3.8 in [**EN00**, Section II.3.a]).

Our goal is to find a sequence of bounded linear operators $\{A_n\}_{n\in\mathbb{N}} \in L(H)$, such that

(I) $A_n h \to Ah$ for $n \to \infty$ and all $h \in \mathscr{D}(A)$,

(II) A_n generates the uniformly continuous semigroup $T_n(t) := e^{tA_n}$, such that

$$\|T_n(t)\|_{L(H)} \leqslant Me^{\lambda t} \text{ for all } t \geqslant 0,$$

(III) $T_n(t)h \to T(t)h$ for $n \to \infty$ for all $t \geqslant 0$, all $h \in H$, and uniformly on each interval $[0, t_0]$.

Corollary A.24 (Yosida approximation for generator; general case). *Let A be the generator of a C_0-semigroup with $\|T(t)\|_{L(H)} \leqslant Me^{\lambda t}$, where $M \geqslant 1$ and $\lambda \in \mathbb{R}$. Then there exists a sequence $\{A_n\}_{n\in\mathbb{N}} \in L(H)$ for $n > \lambda$, such that properties (I)-(III) are fulfilled. We call this sequence the Yosida approximation of the operator A.*

PROOF.

First, we consider the case, where the C_0-semigroup is a generalized contractions, that is $\|T(t)\|_{L(H)} \leqslant e^{\lambda t}$. In this case we define the C_0-semigroup $(S(t))_{t\geqslant 0}$ by $S(t) := e^{-\lambda t}T(t)$. It follows that $(S(t))_{t\geqslant 0}$ is generated by $B := A - \lambda$ with domain $\mathscr{D}(B) = \mathscr{D}(A)$ and $\|S(t)\|_{L(H)} \leqslant 1$ for all $t \geqslant 0$. By Corollary A.23 exists a sequence $\{B_n\}_{n\in\mathbb{N}} \in L(H)$, such that properties *(I)-(III)* hold. Note that $\|S_n(t)\|_{L(H)} \leqslant 1$. Now, we define the approximation sequence $\{A_n\}_{n\in\mathbb{N}} \in L(H)$ by

$$A_n := B_n + \lambda.$$

A straight forward calculation yields $A_n h = B_n h + \lambda h \to Bh + \lambda h = Ah$ for $n \to \infty$ and all $h \in \mathscr{D}(A)$, $T_n(t)h = e^{\lambda t} S_n(t)h \to e^{\lambda t} S(t)h = T(t)h$ for $n \to \infty$ and for all $t \geqslant 0$, and $\|T_n(t)\|_{L(H)} = e^{\lambda t}\|S_n(t)\|_{L(H)} \leqslant e^{\lambda t}$ for all $t \geqslant 0$.
Now, consider an arbitrary C_0-semigroup, that is $\|T(t)\|_{L(H)} \leqslant Me^{\lambda t}$, where $M \geqslant 1$ and $\lambda \in \mathbb{R}$. First, note that we can simplify this case using the same trick as before. Therefore, let the C_0-semigroup $(S(t))_{t\geqslant 0}$ be defined by $S(t) := e^{-\lambda t}T(t)$. If follows that $(S(t))_{t\geqslant 0}$ is generated by $B := A - \lambda$ with domain $\mathscr{D}(B) = \mathscr{D}(A)$ and $\|S(t)\|_{L(H)} \leqslant M$ for all $t \geqslant 0$. Now we define a new norm on H in two steps

(1) $\|h\|_\mu := \sup_{n\geqslant 0} \|\mu^n(\mu - B)^n h\|_H$ for $h \in H$ and all $\mu > 0$,
(2) $|||h|||_H := \sup_{\mu > 0} \|h\|_\mu$ for $h \in H$.

Then it follows that $\|h\|_H \leqslant |||h|||_H \leqslant M\|h\|_H$. This means the norms are equivalent and we have $|||\lambda(\lambda - B)^{-1}|||_{L(H)} \leqslant 1$ for all $\lambda > 0$. From the Hille-Yosida Theorem (see [**EN00**, Theorem 3.5 in II.3.a]) follows that $|||S(t)|||_{L(H)} \leqslant 1$ for all $t \geqslant 0$. Corollary A.23 implies the existence of a sequence $\{B_n\}_{n\in\mathbb{N}} \in L(H)$, such that the properties *(I)-(III)* hold with respect to the norm $|||\cdot|||_H$. But since this norm is equivalent to the original norm $\|\cdot\|_H$, all properties also hold for the $\|\cdot\|_H$-norm, where $\|S_n(t)\|_{L(H)} \leqslant M$ for all $t \geqslant 0$. Now, we can argue exactly like we do in the case of a generalized contraction to obtain the claimed statement.

□

REMARK A.25. *Note that for the semigroup $(T_n(t))_{t\geqslant 0}$ the trivial bound $\|T_n(t)\|_{L(H)} \leqslant e^{t\|A_n\|_{L(H)}}$ holds for fixed $n \in \mathbb{N}$. Hence, for every $n \in \mathbb{N}$ the uniformly continuous semigroup $(T_n(t))_{t\geqslant 0}$ is a generalized contraction.*

APPENDIX B

Stochastic Calculus

In this appendix, we give an overview of the stochastic calculus we use in our models. We start with the case of a Q-Wiener process as the noise term. Then, we turn our attention to Lévy processes and their integration theory in greater detail. This overview is included as these theories are not commonplace at present. Therefore, let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ be a filtered complete probability space and U a real separable Hilbert space. We will always assume that the filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ satisfies the usual conditions

(1) $\mathcal{F}_0$ contains all $\mathcal{N} \in \mathcal{F}$ such that $\mathbb{P}(\mathcal{N}) = 0$,

(2) $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$.

In the following definitions let H be a Hilbert space and $T > 0$.

Definition B.1 (adapted process). *An H-valued stochastic process X is adapted to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ if, for every $t \in [0,T]$, $X(t)$ is $\mathcal{F}_t$-measurable.*

Definition B.2 (progressively measurable process). *An H-valued stochastic process X is progressively measurable, if for each $t \in [0,T]$ it is a measurable mapping from $[0,t] \times \Omega$, where the product σ-algebra $\mathscr{B}([0,t]) \otimes \mathcal{F}_t$ is considered on $[0,t] \times \Omega$.*

Let $\mathcal{P}_T$ denote the σ-algebra of predictable sets, that is, the smallest σ-algebra of subsets of $[0,T] \times \Omega$ containing all sets of the form $\{0\} \times A_0$ and $(s,t] \times A$, where $0 \leqslant s < t \leqslant T$, $A_0 \in \mathcal{F}_0$, and $A \in \mathcal{F}_s$.

Definition B.3 (predictable process). *A stochastic process X taking values in a measurable space $(E, \mathcal{E})$ is called predictable, if it is a measurable mapping from $[0,T] \times \Omega$ to E, where the σ-algebra $\mathcal{P}_T$ is considered on $[0,T] \times \Omega$.*

Next, we provide the definition of a martingale in a Hilbert space. Therefore, let $\mathbb{E}$ denote the expectation.

Definition B.4. *Let U be a separable Hilbert space considered as a measurable space with its Borel σ-algebra $\mathscr{B}(U)$. We fix $T > 0$ and let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\leqslant T}, \mathbb{P})$ be a filtered complete probability space and $\{M_t\}_{t\leqslant T}$ be an U-valued process adapted to the filtration $\{\mathcal{F}_t\}_{t\leqslant T}$. Assume that M is integrable, that is $\mathbb{E}\|M_t\|_U < \infty$ for all $t \in [0,T]$. Then M is called a martingale if $\mathbb{P}$-a.s.*

$$\mathbb{E}(M_t|\mathcal{F}_s) = M_s \text{ for any } 0 \leqslant s \leqslant t \leqslant T.$$

B.1. Itô integral with respect to a Q-Wiener process

In this section, we give the most important definitions and results for the stochastic integration with respect to a Q-Wiener process. In the next section we will see that this process is a special example of a square integrable Lévy process. For a detailed introduction to the topic we refer the reader to [**GM11a**].

Definition B.5 (Q-Wiener process). *Let Q be a nonnegative definite symmetric trace-class operator on a separable Hilbert space U, $\{f_j\}_{j=1}^{\infty}$ be an orthonormal basis in U diagonalizing Q, and let the corresponding eigenvalues be $\{\lambda_j\}_{j=1}^{\infty}$. Let $\{w_j(t)\}_{t\geqslant 0}$, $j = 1, 2, \dots$, be a sequence of independent, scalar valued, standardized Brownian motions defined on $(\Omega, \mathcal{F}, \mathcal{F}_{t\geqslant 0}, \mathbb{P})$. The process*

$$W(t) = \sum_{j=1}^{\infty} \lambda_j^{1/2} w_j(t) f_j \tag{B.1}$$

is called a Q-Wiener process in U. We can assume that the Brownian motions $w_j(t)$ are continuous. Then, the series (B.1) *converges in $L_2(\Omega; C([0,T]; U))$ for every interval $[0,T]$. Therefore, the U-valued Q-Wiener process can be assumed to be continuous. We denote*

$$W_t(u) := \sum_{j=1}^{\infty} \lambda_j^{1/2} w_j(t) \langle f_j, u \rangle_U$$

for any $u \in U$, with the series converging in $L_2(\Omega; C([0,T]; \mathbb{R}))$ on every interval $[0,T]$.

Theorem B.6 (properties of a Q-Wiener process). *A U-valued Q-Wiener process $\{W(t)\}_{t\geqslant 0}$ has the following properties*

(1) $W(0) = 0$;
(2) W has continuous trajectories in U;
(3) W has independent increments;
(4) W is a Gaussian process with the covariance operator Q, that is for any $u, u' \in U$ and $s, t \geqslant 0$,

$$\mathbb{E}(W_t(u) W_s(u')) = (t \wedge s) \langle Qu, u' \rangle_U;$$

(5) For any arbitrary $u \in U$, the law $\mathscr{L}((W(t) - W(s))(u)) \sim N(0, (t-s)\langle Qu, u\rangle_U)$.

Next, we present the most important steps in the construction of the stochastic integral. We start with the following definition.

Definition B.7 (the space $\mathscr{L}^2(U,H)$). *Let H be a separable Hilbert space. By $\mathscr{L}^2(U,H)$ we denote the space of all linear operators from U into H, which are finite with respect to the following norm*

$$\|L\|_{\mathscr{L}^2(U,H)} := \|LQ^{1/2}\|_{L_{(HS)}(U,H)},$$

where $L_{(HS)}(U,H)$ is the space of all Hilbert-Schmidt operators from U to H. If $\{f_j\}_{j=1}^{\infty}$ is an orthonormal basis in U diagonalizing Q and $\{e_j\}_{j=1}^{\infty}$ in an orthonormal basis in H, then we have the following identities

$$\|L\|_{\mathscr{L}^2(U,H)} = \sum_{j,i=1}^{\infty} \langle L(\lambda_j^{1/2} f_j), e_i \rangle_H^2 = \sum_{j,i=1}^{\infty} \langle LQ^{1/2} f_j, e_i \rangle_H^2 = \mathrm{tr}((LQ^{1/2})(LQ^{1/2})^*).$$

Note that in particular all bounded, linear operators are elements of $\mathscr{L}^2(U,H)$, that is $L(U,H) \subset \mathscr{L}^2(U,H)$. For more detail on the space $\mathscr{L}^2(U,H)$ see [**GM11a**, Section2.2].

In the following definition we consider the natural filtration of W.

Definition B.8 (the class of elementary processes $\mathscr{E}(U,H)$). *Let $\mathscr{E}(U,H)$ denote the class of $L(U,H)$-valued elementary processes adapted to the filtration $\{\mathcal{F}_t\}_{t\leqslant T}$ that are of the form*

$$\Phi(\omega,t) = \phi(\omega)\mathbb{1}_{\{0\}}(t) + \sum_{j=0}^{n-1} \phi_j(\omega)\mathbb{1}_{(t_j,t_{j+1}]}(t),$$

where $0 = t_0 \leqslant t_1 \leqslant \cdots \leqslant t_n = T$, and $\phi, \phi_j,\ j = 0,1,\ldots,n-1$, are respectively $\mathcal{F}_0$-measurable and $\mathcal{F}_{t_j}$-measurable $\mathscr{L}^2(U,H)$-valued random variables such that $\phi(\omega), \phi_j(\omega) \in L(U,H)$ for $j = 0,1,\ldots,n-1$.

Definition B.9 (Itô integral for elementary processes). *For $\Phi \in \mathscr{E}(U,H)$, we define the Itô stochastic integral with respect to a Q-Wiener process W by*

$$\int_0^t \Phi(s)dW(s) = \sum_{j=0}^{n-1} \phi_j(W(t_{j+1}\wedge t) - W(t_j \wedge t)),$$

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

Proposition B.10 (Itô isometry for elementary processes). *For a bounded elementary process $\Phi \in \mathscr{E}(U,H)$*

$$\mathbb{E}\left\| \int_0^t \Phi(s)dW(s)\right\|_H^2 = \mathbb{E}\int_0^t \|\Phi(s)Q^{1/2}\|^2_{L_{(HS)}(U,H)}ds < \infty,$$

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

PROOF.

See [**GM11a**, Proposition 2.1].

With the help of the Itô isometry it is possible to extend the definition of the Itô integral to a larger class of stochastic processes. Therefore, we define $\tilde{\mathcal{L}}^2_T(U,H)$ to be the class of all $\mathscr{L}^2(U,H)$-valued processes measurable as mappings from $([0,T]\times\Omega, \mathscr{B}([0,T])\otimes\mathcal{F}_T)$ to $(\mathscr{L}^2(U,H), \mathscr{B}(\mathscr{L}^2(U,H)))$, adapted to the filtration $\{\mathcal{F}_t\}_{t\leqslant T}$, and satisfying the condition

$$\mathbb{E}\int_0^T \|\Phi(t)Q^{1/2}\|^2_{L_{(HS)}(U,H)}dt < \infty. \tag{B.2}$$

If we equip $\tilde{\mathcal{L}}^2_T(U,H)$ with the norm

$$\|\Phi\|_{\tilde{\mathcal{L}}^2_T(U,H)} = \big(\mathbb{E}\int_0^T \|\Phi(t)Q^{1/2}\|^2_{L_{(HS)}(U,H)}dt\big)^{1/2},$$

it becomes a Hilbert space. Since $\mathcal{E}(U,H)$ is dense in $\tilde{\mathcal{L}}^2_T(U,H)$ (see [**GM11a**, Proposition 2.2]), one can extend the definition of the Itô integral to $\tilde{\mathcal{L}}^2_T(U,H)$. Note that the space $\tilde{\mathcal{L}}^2_T(U,H)$ includes the space of all predictable processes fulfilling condition (B.2). For more detail see [**GM11a**, Section 2.2.2 and Section 2.2.3].

In the case of a Q-Wiener process it is possible to extend the definition of the the Itô integral even further. Therefore, let $\mathcal{P}_T^2(U,H)$ denote the class of $\mathscr{L}^2(U,H)$-valued stochastic processes adapted to the filtration $\{\mathcal{F}_t\}_{t\leqslant T}$, measurable from $([0,T]\times\Omega, \mathscr{B}([0,T])\otimes\mathcal{F}_T)$ to $(\mathscr{L}^2(U,H), \mathscr{B}(\mathscr{L}^2(U,H)))$ satisfying the condition

$$\mathbb{P}\left\{\int_0^T \|\Phi(t)Q^{1/2}\|^2_{L_{(HS)}(U,H)} < \infty\right\} = 1.$$

Obviously, $\tilde{\mathcal{L}}_T^2(U,H) \subset \mathcal{P}_T^2(U,H)$. Again one can show that $\mathcal{E}(U,H)$ is dense in $\mathcal{P}_T^2(U,H)$ (see [**GM11a**, Lemma 2.3]), which allows the extension of the Itô integral to $\mathcal{P}_T^2(U,H)$. For more detail see [**GM11a**, Section 2.2.3].

Next, we introduce the Itô integral with respect to a Q-Wiener process with a random limit.

LEMMA B.11. *Let $\Phi \in \mathcal{P}_T^2(U,H)$, and τ be a stopping time relative to $\{\mathcal{F}_t\}_{t\in[0,T]}$, such that $\mathbb{P}(\tau \leqslant T) = 1$. Define*

$$\int_0^\tau \Phi(t)dW(t) := \int_0^s \Phi(t)dW(t) \text{ on the set } \{\omega : \tau(\omega) = s\}.$$

Then,

$$\int_0^\tau \Phi(t)dW(t) = \int_0^T \Phi(t)\mathbb{1}_{t\leqslant\tau}dW(t).$$

PROOF.

See [**GM11a**, Lemma 2.7].

Following [**GWMS14**, Section I.1 §4.], we define the stochastic integral with respect to a Q-Wiener process W with random limits, by

$$\int_{\tau_1}^{\tau_2} \Phi(s)dW(s) = \int_0^{\tau_2} \Phi(s)dW(s) - \int_0^{\tau_1} \Phi(s)dW(s), \tag{B.3}$$

where τ_1 and τ_2 are two stopping times with $\mathbb{P}(0 \leqslant \tau_1 \leqslant \tau_2 \leqslant T) = 1$.

The next result is the Itô formula for the Itô integral with respect to a Q-Wiener process. Therefore, we denote by $C^k(H,\mathbb{R})$, $k = 0,1,\dots$ all mappings ϕ from H to $\mathbb{R}$, which are continuous with all the Fréchet derivatives $\phi, \phi', \phi'', \dots$ up to order k. If the function ϕ also depends on time, its time derivative is denoted by $\dot{\phi}$.

THEOREM B.12 (Itô formula). *Let Q be a symmetric nonnegative trace-class operator on a separable Hilbert space U, and let W be a Q-Wiener process on a filtered complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$. Assume that a stochastic process $X(t)$, $0 \leqslant t \leqslant T$, is given by*

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

where $X(0)$ is an $\mathcal{F}_0$-measurable H-valued random variable, $\Phi \in \mathcal{P}_T^2(U,H)$, and Ψ is an H-valued $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted $\mathbb{P}$-a.s. Bochner-integrable process on $[0,T]$, that is $\mathbb{P}$-a.s.

$$\int_0^T \|\Psi(s)\|_H ds < \infty.$$

Assume that a function $\phi : [0,T]\times H \to \mathbb{R}$ is, such that ϕ is continuous and it's Fréchet partial derivatives $\dot{\phi}, \phi', \phi''$ are continuous and bounded on bounded subsets of $[0,T] \times H$. Then, the following Itô's formula holds

$$\begin{aligned}\phi(t, X(t)) =& \phi(0, X(0)) + \int_0^t \langle \phi'(s, X(s)), \Phi(s)dW(s)\rangle_H \\ &+ \int_0^t \{\dot{\phi}(s, X(s)) + \langle \phi'(s, X(s)), \Psi(s)\rangle_H \\ &+ \frac{1}{2}\operatorname{tr}[\phi''(X(s))(\Phi(s)Q^{1/2})(\Phi(s)Q^{1/2})^*]\}ds,\end{aligned} \tag{B.4}$$

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

PROOF.

See [**GM11a**, Theorem 2.9].

The last result of this section is a stochastic version of Lebegues dominated convergence theorem.

THEOREM B.13. *Let $\Phi \in \mathcal{P}_T^2(U,H)$ and let $(\Phi_n)_{n\in\mathbb{N}}$ be a sequence such that $\Phi_n \in \mathcal{P}_T^2(U,H)$ for all $n \in \mathbb{N}$. Suppose that Φ_n converges $dt\otimes\mathbb{P}$-a.s. to Φ on $\Omega \times [0,T]$ for $n \to \infty$, and $\mathbb{P}$-a.s.*

$$\lim_{n\to\infty} \int_0^T \|(\Phi_n(s) - \Phi(s))Q^{1/2}\|^2_{L_{(HS)}(U,H)} ds = 0. \tag{B.5}$$

Assume there is a $\Psi \in \mathcal{P}_T^2(U,H)$ such that $\mathbb{P}$-a.s.

$$\int_0^T \|\Phi_n(s)Q^{1/2}\|^2_{L_{(HS)}(U,H)} ds \leqslant \int_0^T ||\Psi(s)Q^{1/2}\|^2_{L_{(HS)}(U,H)} ds.$$

Then, we have for all $t \in [0,T]$

$$\int_0^t \Phi(s)dW(s) = \lim_{n\to\infty} \int_0^t \Phi_n(s)dW(s),$$

where the limit is in probability.

PROOF.

Let $\varepsilon > 0$ be arbitrary. For $t \in [0,T]$ we define the random variables

$$h_1^{t,N} := \mathbb{1}_{\{\int_0^t \|\Psi(s)Q^{1/2}\|^2_{L_{(HS)}(U,H)} ds \leqslant N\}}$$

$$h_2^{t,N} := \mathbb{1}_{\{\int_0^t \|\Psi(s)Q^{1/2}\|^2_{L_{(HS)}(U,H)} ds > N\}},$$

where $N \in \mathbb{N}$. Note that $\mathbb{P}$-a.s. $h_1^{t,N} + h_2^{t,N} = 1$. Now let $t \in [0,T]$ and we consider $\mathbb{P}$-a.s.

$$\begin{aligned}
&\Big\| \int_0^t (\Phi_n - \Phi)(s)dW(s)\Big\|_H^2 \\
&\leqslant 4\Big\| \int_0^t h_1^{t,N}(\Phi_n - \Phi)(s)dW(s)\Big\|_H^2 && \text{(B.6)}\\
&\quad + 4\Big\| \int_0^t h_2^{t,N}\Phi_n(s)dW(s)\Big\|_H^2 + 4\Big\| \int_0^t h_2^{t,N}\Phi(s)dW(s)\Big\|_H^2 . && \text{(B.7)}
\end{aligned}$$

Note that all integrals appearing in (B.6) and (B.7) are well-defined (see [**GM11a**, Lemma 2.3 to Definition 2.12] for more detail). We start with the first term in (B.7) and find

$$\begin{aligned}
&\mathbb{P}\bigg(\Big\| \int_0^t h_2^{t,N}\Phi_n(s)dW(s)\Big\|_H^2 > 0\bigg) \\
&= \mathbb{P}\bigg(\Big\| \int_0^t \mathbb{1}_{\{\int_0^t \|\Psi(s)Q^{1/2}\|^2_{L_{(HS)}(U,H)}ds > N\}}\Phi_n(s)dW(s)\Big\|_H^2 > 0\bigg) \\
&\leqslant \mathbb{P}\bigg(\int_0^T \|\Psi(s)Q^{1/2}\|^2_{L_{(HS)}(U,H)}ds > N\bigg) \longrightarrow 0 \quad \text{for } N \to \infty,
\end{aligned}$$

since $\Psi \in \mathcal{P}(U,H)$. The second term in (B.7) can be treated in exactly the same way. Hence, we choose an $N \in \mathbb{N}$ large enough such that

$$\mathbb{P}\bigg(4\Big\| \int_0^t h_2^{t,N}\Phi_n(s)dW(s)\Big\|_H^2 + 4\Big\| \int_0^t h_2^{t,N}\Phi(s)dW(s)\Big\|_H^2 > 0\bigg) \leqslant \frac{2}{3}\varepsilon.$$

Since N is fix now, we find that $h_1^{t,N}\Phi_n, h_1^{t,N}\Phi \in \tilde{\mathcal{L}}_T^2(U,H)$ for all $n \in \mathbb{N}$. Therefore, we can apply the Itô isometry (see [**GM11a**, Theorem 2.3]) to (B.6)

$$\begin{aligned}
&\mathbb{E}\bigg[\Big\| \int_0^t h_1^{t,N}(\Phi_n - \Phi)(s)dW(s)\Big\|_H^2\bigg] \\
&= \mathbb{E}\bigg[\mathbb{1}_{\{\int_0^t \|\Psi(s)Q^{1/2}\|^2_{L_{(HS)}(U,H)}ds \leqslant N\}}\int_0^t \|(\Phi_n - \Phi)(s)Q^{1/2}\|^2_{L_{(HS)}(U,H)}ds\bigg] \\
&\longrightarrow 0 \quad \text{for } n \to \infty.
\end{aligned}$$

The convergence follows, since

$$\mathbb{1}_{\{\int_0^t \|\Psi(s)Q^{1/2}\|^2_{L_{(HS)}(U,H)}ds \leqslant N\}}\int_0^t \|(\Phi_n - \Phi)(s)Q^{1/2}\|^2_{L_{(HS)}(U,H)}ds \leqslant 4N,$$

such that Lebesgues dominated convergence theorem and (B.5) can be applied. Thus, we choose an $n \in \mathbb{N}$ large enough, such that

$$\mathbb{P}\bigg(4\Big\| \int_0^t h_1^{t,N}(\Phi_n - \Phi)(s)dW(s)\Big\|_H^2 > 0\bigg) \leqslant \frac{1}{3}\varepsilon.$$

□

B.2. Hilbert space-valued Lévy processes

Before we introduce the integration theory with respect to a Lévy process, we present some basic results on Lévy process. A detailed introduction can be found in [**PZ07**]. We start with the formal definition.

Definition B.14 (Hilbert space-valued Lévy process). *A stochastic process* $L = \{L(t)\}_{t \geqslant 0}$ *taking values in a separable Hilbert space* U *is called a Lévy process if*

- $L(0) = 0$ $\mathbb{P}$*-a.s.,*
- L *has independent and homogeneous increments, and*
- L *ist stochastically continuous, that is for all* $t \in \mathbb{R}_+$ *and* $\varepsilon > 0$

$$\lim_{s \to t} \mathbb{P}(\|L(s) - L(t)\|_U > \varepsilon) = 0.$$

Example B.15. *A* Q*-Wiener process is a Lévy process.*

Since Lévy processes need only be stochastically continuous, they allow jumps. Thus, the paths don't have to be continuous like they are for a Q-Wiener process. Therefore, we introduce the concept of càdlàg paths.

Definition B.16. *A stochastic process* L *is called càdlàg (continu à droite et limites à gauche) if*

- L *is* $\mathbb{P}$*-a.s. right-continuous, that is*

$$\mathbb{P}(\lim_{s \downarrow t} \|L(s) - L(t)\|_U = 0, \ \forall t \geqslant 0) = 1, \textit{and}$$

- L *has left limits* $L(t-)$*, that is*

$$\mathbb{P}(\lim_{s \uparrow t} \|L(s) - L(t-)\|_U = 0, \ \forall t \geqslant 0) = 1.$$

Theorem B.17. *Every Lévy process* $\tilde{L}$ *has a càdlàg modification, that is for every* $\tilde{L}$ *there exists a càdlàg Lévy process* L*, such that* $\mathbb{P}(\tilde{L}(t) = L(t)) = 1$ *for all* $t \geqslant 0$*.*

Definition B.18 (jump process). *Let* L *be a càdlàg process. The process of jumps of* L *is then defined by* $\Delta L(t) := L(t) - L(t-),\ t \geqslant 0$*.*

Definition B.19 (compound Poisson process). *Let* $\tilde{\nu}$ *be a finite measure on a Hilbert space* U*, such that* $\tilde{\nu}(\{0\}) = 0$*. A compound Poisson process with the Lévy measure (also called the jump intensity measure)* $\tilde{\nu}$ *is a càdlàg Lévy process* L *satisfying*

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

for all $t \geqslant 0$ *and* $A \in \mathscr{B}(U)$*. In the formula above, we use the convention that* $\tilde{\nu}^0$ *is equal to the unit measure concentrated at* $\{0\}$*, that is,* $\tilde{\nu}^0 = \delta_0$ *and the* $*$ *indicates that we take the convolution of the measures* $\tilde{\nu}$*.*

Proposition B.20. *Let L be a compound Poisson process with Lévy measure $\tilde{\nu}$.*

(1) *The process L is integrable if and only if*

$$\int_U \|x\|_U \tilde{\nu}(dx) < \infty. \tag{B.8}$$

Moreover, if (B.8) *holds, then*

$$\mathbb{E}L(t) = t \int_U x\tilde{\nu}(dx).$$

(2) *The process L, and hence the compensated process $\hat{L} := L(t) - \mathbb{E}L(t)$, is square integrable if and only if*

$$\int_U \|x\|_U^2 \tilde{\nu}(dx) < \infty. \tag{B.9}$$

Moreover, if (B.9) *holds, then*

$$\mathbb{E}\|L(t)\|_U^2 = t \int_U \|x\|_U^2 \tilde{\nu}(dx) + t^2 \left\| \int_U x\tilde{\nu}(dx) \right\|^2, \text{ and}$$

$$\mathbb{E}\|\hat{L}(t)\|_U^2 = t \int_U \|x\|_U^2 \tilde{\nu}(dx).$$

PROOF.

See [**PZ07**, Proposition 4.18].

Definition B.21 (Poisson random measure of a Lévy process). *The Poisson random measure corresponding to a càdlàg Lévy process L is defined for every $A \in \mathscr{B}(U\backslash\{0\})$ by*

$$N(t,A)(\omega) := \#\{0 \leqslant s \leqslant t : \Delta L(s)(\omega) \in A\} = \sum_{0 \leqslant s \leqslant t} \mathbb{1}_A(\Delta L(s)(\omega)),$$

which counts the jumps of L, which are in the set A.

Definition B.22 (Lévy measure of a Lévy process). *The Lévy measure or jump intensity measure on $U\backslash\{0\}$ of a càdlàg Lévy process is defined by $\nu(A) = \mathbb{E}[N(1,A)]$ for $A \in \mathscr{B}(U\backslash\{0\})$.*

Note that the Lévy measure doesn't need to be finite on $U\backslash\{0\}$, but it is always σ-finite.

Remark B.23. *We can always extend the measure ν to the entire space U by $\nu(\{0\}) = 0$. This allows us to remain a cleaner notion later on, when we use the Poisson integral notation.*

Definition B.24 (bounded below). *$A \in \mathscr{B}(U\backslash\{0\})$ is called bounded below if 0 is not an element of the closure of A, that is $0 \notin \bar{A}$.*

Lemma B.25. *If A is bounded below, then $N(t,A) < \infty$ $\mathbb{P}$-a.s. for all $t \geqslant 0$.*

PROOF.

See [**AR05**, Proposition 2.8].

Definition B.26 (compensated Poisson random measure of a Lévy process). *For each $t \geqslant 0$ and A bounded below the compensated Poisson random measure of a Lévy process is defined by*

$$\tilde{N}(t, A) = N(t, A) - t\nu(A).$$

Now, we are able to give the definition of the Poisson integral which will be important later on for the transformation formula for jump processes.

Definition B.27 (Poisson integral). *Let $f : U\backslash\{0\} \to U$ be a $\mathscr{B}(U\backslash\{0\})/\mathscr{B}(U)$-measurable function and A bounded below. Then for each $t \geqslant 0$, $\omega \in \Omega$ the following random finite sum is called the Poisson integral of f*

$$\int_A f(x)N(t, dx)(\omega) := \sum_{x \in A} f(x)N(t, \{x\})(\omega).$$

Remark B.28. *Note, since $N(t, \{x\}) \neq 0$ holds if and only if $\Delta L(s) = x$ for at least one $0 \leqslant s \leqslant t$, we have*

$$\int_A f(x)N(t, dx)(\omega) = \sum_{0 \leqslant s \leqslant t} f(\Delta L(s))\mathbb{1}_A(\Delta L(s)(\omega)).$$

Definition B.29 (compensated Poisson integral). *Let ν_A denote the restriction to A of the measure ν and let $f \in L_1(A, \nu_A)$. For $t \geqslant 0$ the compensated Poisson integral is defined by*

$$\int_A f(x)\tilde{N}(t, dx) := \int_A f(x)N(t, dx) - t\int_A f(x)\nu(dx).$$

Theorem B.30 (Lévy-Khinchin decomposition). *Let $(r_k)_{k \in \mathbb{N}}$ be an arbitrary sequence decreasing to 0, $A_0 := \{x \in U : \|x\|_U \geqslant r_0\}$, and $A_k := \{x \in U : r_k \leqslant \|x\|_U < r_{k-1}\}$ for $k \in \mathbb{N}$. Furthermore, let ν be the Lévy measure of a Lévy process L. Then, the following representation holds*

$$L(t) = at + W(t) + \sum_{k=1}^{\infty} \big(L_{A_k}(t) - t\int_{A_k} x\nu(dx)\big) + L_{A_0}(t), \ t \geqslant 0, \tag{B.10}$$

where $a \in U$, W is a Q-Wiener process, L_{A_k} is a compound Poisson process for $k \in \mathbb{N}$ with Lévy measure $\mathbb{1}_{\{x \in U: r_k \leqslant \|x\|_U < r_{k-1}\}}\nu$, and L_{A_0} is a compound Poisson process with Lévy measure $\mathbb{1}_{\{x \in U: \|x\|_U \geqslant r_0\}}\nu$. Additionally, all members of the representation are independent processes and the series converges $\mathbb{P}$-a.s. uniformly on each bounded subinterval $[0, t]$ of $[0, \infty)$.

PROOF.

See [**PZ07**, Theorem 4.23].

Remark B.31. *Using the notation of the Poisson integral and compensated Poisson integral we can rewrite the Lévy-Khinchin decomposition as follows*

$$L(t) = ta + W(t) + \int_{\{x\in U:\|x\|_U<r_0\}} x\tilde{N}(t,dx) + \int_{A_0} xN(t,dx),\ \ t \geqslant 0,$$

where $\int_{\{x\in U:\|x\|_U<r_0\}} x\tilde{N}(t,dx) := \lim_{n\to\infty}\int_{\{x\in U:r_n<\|x\|_U<r_0\}} x\tilde{N}(t,dx)$. *For more detail see* [**App06**].

For the next theorem, we assume that L is an $\{F_t\}_{t\in[0,T]}$-adapted and square integrable càdlàg Lévy process. Furthermore we assume that for $t, h > 0$ the increments $L(t+h) - L(t)$ are independent of $\{\mathcal{F}_t\}_{t\in[0,T]}$.

Theorem B.32. *There exists an $m \in U$ and a nonnegative definite symmetric trace-class operator Q such that, for all $t, s \geqslant 0$ and $x, y \in U$,*

$$\begin{aligned}\mathbb{E}\langle L(t), x\rangle_U &= \langle m, x\rangle_U t,\\ \mathbb{E}\langle L(t) - m(t), x\rangle_U \langle L(s) - m(s), y\rangle_U &= t \wedge s\langle Qx, y\rangle_U,\\ \mathbb{E}\|L(t) - mt\|_U^2 &= t\,\mathrm{tr}(Q).\end{aligned}$$

The vector m is called the mean and the operator Q is the covariance operator of the process L.

PROOF.

See [**PZ07**, Theorem 4.44].

Corollary B.33. *Let L be a Lévy process taking values in U. Furthermore, let ν be the Lévy measure of L. Then, L is square integrable if and only if*

$$\int_U \|x\|^2 \nu(dx) < \infty.$$

PROOF.

The statement follows easily by applying Proposition B.20 to the Lévy-Khinchin decomposition (Theorem B.30). See also [**PZ07**, Theorem 4.47] for more detail.

□

Remark B.34. *Note that we are taking the integral over the entire space U instead of $U\backslash\{0\}$, which is justified by Remark B.23.*

Corollary B.35. *Let L be a square integrable Lévy process taking values in U. Then L can be decomposed as follows*

$$L(t) = tb + W(t) + M_J(t), \tag{B.11}$$

where $b \in U$, W is a Q-Wiener process and M_J is a $\{\mathcal{F}_t\}_{t\in[0,T]}$ martingale, which contains all jumps of L. Furthermore W and M_J are independent square integrable Lévy processes and $\mathbb{E}L(t) = tb$ for $t \geqslant 0$.

PROOF.

The decomposition follows directly from the Lévy-Khinchin decomposition (Theorem B.30), since we are able to compensate also the compound Poisson process L_{A_0} by Proposition B.20.

□

THEOREM B.36. *Let L be a square integrable Lévy process, Q_0 be the covariance operator of the Wiener part of L, and let Q_1 be the covariance operator of the jump part. Then,*

$$\langle Q_1 x, z\rangle_U = \int_U \langle x, y\rangle_U \langle z, y\rangle_U \nu(dy), \tag{B.12}$$

for $x, z \in U$, and

$$Q = Q_0 + Q_1.$$

PROOF.

See [**PZ07**, Theorem 4.47].

REMARK B.37. *Just like we did for the Lévy-Khinchin decomposition we can rewrite* (B.11) *using the notation of compensated Poisson integral. But this time, since all members of the representation can be compensated, we only need one compensated Poisson integral. This way we can rewrite* (B.11) *as*

$$L(t) = tb + W(t) + \int_U x\tilde{N}(t, dx), \tag{B.13}$$

where $\int_U x\tilde{N}(t, dx) := \lim_{n\to\infty} \int_{\{x\in U: r_n < \|x\|_U\}} x\tilde{N}(t, dx)$, for a null sequence $(r_k)_{k\in\mathbb{N}}$ from the Lévy-Khinchin decomposition (Theorem B.30).

Another important representation of a Lévy process is the expansion with respect to an orthonormal basis. Therefore, let us assume that L is a square integrable càdlàg Lévy process with mean zero and covariance operator Q. We choose the sequence of eigenvectors $\{e_n\}_{n=1}^{\infty}$ of Q as the orthogonal basis of U. In particular, we have $Qe_n = \lambda_n e_n$ for all $n \in \mathbb{N}$, where $\lambda_n \geqslant 0$ and $\mathrm{tr}(Q) = \sum_{n=1}^{\infty} \lambda_n < \infty$. Furthermore, we can represent the operator Q by $Q = \sum_{n=1}^{\infty} \lambda_n e_n \otimes e_n$. We define the stochastic process $L_n(t) := \langle L(t), e_n\rangle_U$ for $n \in \mathbb{N}$. Then, the stochastic processes L_n are real-valued uncorrelated càdlàg Lévy processes (see [**PZ07**, Section 4.8]).

THEOREM B.38 (series representation of L). *The series*

$$L(t) = \sum_{n=1}^{\infty} L_n(t) e_n \tag{B.14}$$

converges in probability, uniformly in t on any compact interval $[0, T]$.

PROOF.

See [**PZ07**, Theorem 4.39].

COROLLARY B.39. *If L is a square integrable càdlàg Lévy process with mean zero, the series* (B.14) *converges $\mathbb{P}$-a.s. and in mean square.*

PROOF.
First note that the process L_n is square integrable for all $n \in \mathbb{N}$, since

$$\mathbb{E}L_n(t)^2 = \mathbb{E}\langle L(t), e_n\rangle_U^2 = t\langle Qe_n, e_n\rangle_U = t\lambda_n\langle e_n, e_n\rangle_U = t\lambda_n < \infty,$$

where we use the second property of Theorem B.32. Then, it follows that

$$\mathbb{E}\|\sum_{n=k}^{N} L_n(t)e_n\|_U^2 = \mathbb{E}\sum_{n,j=k}^{N} L_n(t)L_j(t)\langle e_n, e_j\rangle_U = \sum_{n=k}^{N}\mathbb{E}L_n(t)^2 = t\sum_{n=k}^{N}\lambda_n.$$

Since $\sum_{n=1}^{\infty}\lambda_n < \infty$, it follows that the series (B.14) converges in mean square. The $\mathbb{P}$-a.s. convergence follows by results from [**PZ07**, Section 4.8].

□

If we define $L_N(t) := \sum_{n=1}^{N} L_n(t)e_n$ and $L^N(t) := \sum_{n=N+1}^{\infty} L_n(t)e_n$ for $N \in \mathbb{N}$ it is obvious that both processes are square integrable càdlàg Lévy processes and $\mathbb{P}$-a.s. $L = L_N + L^N$. Furthermore, the covariance operator of L_N and L^N respectively are given by $Q_N = \sum_{n=1}^{N}\lambda_n e_n \otimes e_n$ and $Q^N = \sum_{n=N+1}^{\infty}\lambda_n e_n \otimes e_n$.

B.3. Stochastic integral with respect to a square integrable Lévy martingale

In this section, we introduce the stochastic integral with respect to a square integrable Lévy martingale. The construction is very similar to the construction of the Itô integral with respect to a Q-Wiener process, since the processes have many properties in common like Theorem $B.32$ shows. A detailed introduction of the here presented approach can be found in [**PZ07**, Chapter 8].

We start again with the definition of the stochastic integral for elementary processes (see Definition B.8).

DEFINITION B.40 (stochastic integral for elementary processes). *For $\Phi \in \mathscr{E}(U,H)$, we define the stochastic integral with respect to a square integrable Lévy martingale $M(t)$ for $t \in [0,T]$ by*

$$\int_0^t \Phi(s)dM(s) = \sum_{j=0}^{n-1}\phi_j(M(t_{j+1}\wedge t) - M(t_j \wedge t)).$$

PROPOSITION B.41 (isometry for elementary processes). *For a bounded elementary process $\Phi \in \mathscr{E}(U,H)$ and for $t \in [0,T]$ holds*

$$\mathbb{E}\left\|\int_0^t \Phi(s)dM(s)\right\|_H^2 = \mathbb{E}\int_0^t \|\Phi(s)Q^{1/2}\|_{L_{(HS)}(U,H)}^2 ds < \infty.$$

PROOF.
See [**PZ07**, Proposition 8.6].

With the help of the isometry we can extend the definition of the stochastic integral to a larger class of stochastic processes. Therefore, we define $\mathcal{L}^2_T(U,H)$ to be the class of all predictable mappings Φ from $[0,T]\times\Omega$ taking values in set of linear operators from U to H such that

$$\mathbb{E}\int_0^T \|\Phi(t)Q^{1/2}\|^2_{L_{(HS)}(U,H)}dt < \infty. \tag{B.15}$$

If we equip $\mathcal{L}^2_T(U,H)$ with the norm

$$\|\Phi\|_{\mathcal{L}^2_T(U,H)} = \big(\mathbb{E}\int_0^T \|\Phi(t)Q^{1/2}\|^2_{L_{(HS)}(U,H)}dt\big)^{1/2}$$

it becomes a Hilbert space. For more detail on the extension of the stochastic integral see [**PZ07**, Theorem 8.7].

The next result is a useful tool, when dealing with the stochastic integral.

PROPOSITION B.42. *Assume that $\Phi(t)Q^{\frac{1}{2}}$ is an $L_{(HS)}(U,H)$-predictable stochastic process for all $t\in[0,T]$. Furthermore, let $A:\mathscr{D}(A)\subset H\longrightarrow H$ be a closed linear operator with domain $\mathscr{D}(A)$ being a Borel subset of H. If for all $u\in U$ and $t\in[0,T]$ one has $\Phi(t)u\in\mathscr{D}(A)$ $\mathbb{P}-a.s.$ and $\Phi, A\Phi\in\mathcal{L}^2_T(U,H)$, then $\Phi\in\mathcal{L}^2_T(U,\mathscr{D}(A))$ and $\mathbb{P}$-a.s.*

$$A\int_0^T \Phi(s)dM(s) = \int_0^T A\Phi(s)dM(s). \tag{B.16}$$

PROOF.
Before we prove the statement let us note that for a linear operator $S:U\to H$ the following equivalence holds

$$S\in L_{(HS)}(U,\mathscr{D}(A)) \Longleftrightarrow S\in L_{(HS)}(U,H) \text{ and } AS\in L_{(HS)}(U,H).$$

This is easy to see, since

$$\begin{aligned}\|S\|^2_{L_{(HS)}(U,\mathscr{D}(A))} = \sum_{i=1}^\infty \|Sf_i\|^2_{\mathscr{D}(A)} &= \sum_{i=1}^\infty \|Sf_i\|^2_H + \sum_{i=1}^\infty \|ASf_i\|^2_H \\ &= \|S\|^2_{L_{(HS)}(U,H)} + \|AS\|^2_{L_{(HS)}(U,H)},\end{aligned}$$

for any normalized orthonormal basis $\{f_i\}_{i=1}^\infty$ of U. Applying this identity and the isometry of the stochastic integral (see [**PZ07**, Theorem 8.7 (i)]) yields

$$\begin{aligned}\mathbb{E}\left\|\int_0^T \Phi(s)dM(s)\right\|^2_{\mathscr{D}(A)} &= \mathbb{E}\int_0^T \|\Phi(s)Q^{\frac{1}{2}}\|^2_{L_{(HS)}(U,\mathscr{D}(A))}ds \\ &= \mathbb{E}\int_0^T \|\Phi(s)Q^{\frac{1}{2}}\|^2_{L_{(HS)}(U,H)}ds + \mathbb{E}\int_0^T \|A\Phi(s)Q^{\frac{1}{2}}\|^2_{L_{(HS)}(U,H)}ds.\end{aligned}$$

Since $\Phi, A\Phi\in\mathcal{L}^2_T(U,H)$ we have $\mathbb{E}\left\|\int_0^T \Phi(s)dM(s)\right\|^2_{\mathscr{D}(A)} < \infty$.

We proceed to the identity (B.16). Let Φ be a simple process at first. Through a straight forward calculation we find $\mathbb{P}$-a.s.

$$A\int_0^T \Phi(s)dM(s) = A\sum_{i=0}^{m} \Phi_i(M(t_{i+1}) - M(t_i)) = \sum_{i=0}^{m} A\Phi_i(M(t_{i+1}) - M(t_i))$$
$$= \int_0^T A\Phi(s)dM(s).$$

Since $\Phi \in \mathcal{L}_T^2(U, \mathscr{D}(A))$, there exists a sequence $\{\Phi_m\}$ of simple $\mathscr{D}(A)$-valued processes, which converges to Φ in $\mathcal{L}_T^2(U, \mathscr{D}(A))$. Hence, we have

$$\int_0^T \Phi_m(s)dM(s) \to \int_0^T \Phi(s)dM(s)$$

and

$$\int_0^T A\Phi_m(s)dM(s) \to \int_0^T A\Phi(s)dM(s).$$

This convergence is in $L_2(\Omega, \mathcal{F}, \mathbb{P}; H)$. But we have already shown that $\mathbb{P}$-a.s.

$$\int_0^T A\Phi_m(s)dM(s) = A\int_0^T \Phi_m(s)dM(s).$$

Since the operator A is closed, the claim follows.

□

The following result is a stochastic version of Fubinis theorem. Therefore let λ be a finite positive measure on a measurable space $(E, \mathcal{E})$. Then we can formulate the following stochastic Fubini theorem.

THEOREM B.43. *Assume that $\Phi \in L_1(E, \mathcal{E}, \lambda; \mathcal{L}_T^2(U, H))$. Then, $\mathbb{P}$-a.s.*

$$\int_E \int_0^T \Phi(t,x)dM(t)\lambda(dx) = \int_0^T \int_E \Phi(t,x)\lambda(dx)dM(t).$$

PROOF.
See [**PZ07**, Throrem 8.14].

For some calculations it is useful to use an approximation of the stochastic integral, which is based on the series (B.14). Therefore, we define the Banach space $\mathcal{X}_{T,B}$ of all predictable process with values in a Banach space B equipped with the norm

$$\|X\|_{T,B} := \sqrt{\sup_{t\in[0,T]} \mathbb{E}\|X(t)\|_B^2}.$$

PROPOSITION B.44. *Let M be a square integrable Lévy martingale and $\Phi \in \mathcal{L}_T^2(U, H)$. Then,*

$$\sum_{n=1}^{N} \int_0^t \Phi(s)e_n dm_n(s) \longrightarrow \int_0^t \Phi(s)dM(s),$$

for $N \to \infty$ in $\mathcal{X}_{T,H}$, where $(e_n)_{n\in\mathbb{N}}$ are the eigenvectors of Q and $m_n := \langle M, e_n\rangle_U$.

PROOF.

If Φ is an elementary process, then the stochastic integrals have the following representation

$$\int_0^t \Phi(s)dM_N(s) = \sum_{j=0}^{n-1} \phi_j(M_N(t_{j+1} \wedge t) - M_N(t_j \wedge t)),$$

and

$$\int_0^t \Phi(s)dM^N(s) = \sum_{j=0}^{n-1} \phi_j(M^N(t_{j+1} \wedge t) - M^N(t_j \wedge t)).$$

An easy calculation yields $\mathbb{P}$-a.s.

$$\int_0^t \Phi(s)dM(s) = \int_0^t \Phi(s)dM_N(s) + \int_0^t \Phi(s)dM^N(s), \tag{B.17}$$

for all $t \in [0,T]$. By the same procedure as in the construction of the integral we find that identity (B.17) holds for all $\Phi \in \mathcal{L}_T^2(U,H)$. Note that $\mathbb{P}$-a.s.

$$\int_0^t \Phi(s)dM_N(s) = \sum_{n=1}^{N} \int_0^t \Phi(s)e_n dm_n(s).$$

In order to show the claimed approximation we calculate for a $\Phi \in \mathcal{L}_T^2(U,H)$, using the isometry of the stochastic integral (see [**PZ07**, Theorem 8.7 (i)])

$$\begin{aligned}\mathbb{E}\Big\| \int_0^t \Phi(s)dM(s) - \int_0^t \Phi(s)dM_N(s)\Big\|_H^2 &= \mathbb{E}\Big\| \int_0^t \Phi(s)dM^N(s) \Big\|_H^2 \\ &= \mathbb{E}\int_0^t \|\Phi(s)(Q^N)^{\frac{1}{2}}\|^2_{L_{(HS)}(U,H)}ds = \mathbb{E}\int_0^t \sum_{n=1}^{\infty} \|\Phi(s)(Q^N)^{\frac{1}{2}}e_n\|_H^2 ds \\ &= \mathbb{E}\int_0^t \sum_{n=N+1}^{\infty} \lambda_n \|\Phi(s)e_n\|_H^2 ds \leqslant \mathbb{E}\int_0^T \sum_{n=N+1}^{\infty} \lambda_n \|\Phi(s)e_n\|_H^2 ds,\end{aligned} \tag{B.18}$$

which converges pointwise (for every s) to zero $N \to \infty$. Since $\Phi \in \mathcal{L}_T^2(U,H)$, we have

$$\infty > \mathbb{E}\int_0^T \|\Phi(s)Q^{\frac{1}{2}}\|^2_{L_{(HS)}(U,H)}ds = \mathbb{E}\int_0^T \sum_{n=1}^{\infty} \|\Phi(s)Q^{\frac{1}{2}}e_n\|_H^2 ds = \mathbb{E}\int_0^T \sum_{n=1}^{\infty} \lambda_n\|\Phi(s)e_n\|_H^2 ds,$$

which dominates (B.18). Therefore, by Lebesgue's dominated convergence theorem we receive

$$\mathbb{E}\Big\| \int_0^t \Phi(s)dM(s) - \int_0^t \Phi(s)dM_N(s)\Big\|_H^2 = \mathbb{E}\int_0^t \sum_{n=N+1}^{\infty} \lambda_n \|\Phi(s)e_n\|_H^2 ds \longrightarrow 0,$$

for $N \to \infty$. The convergence in $\mathcal{X}_{T,H}$ follows from (B.18).

□

REMARK B.45. *If one desires the convergence in the space of martingales equipped with the stronger norm* $\|X\| := \sqrt{\mathbb{E}(\sup_{t\in[0,T]} \|X(t)\|_H^2)}$ *one needs to apply a generalization of Doob's inequality. Such an inequality can be found for example in* [**MP80**, Theorem 2].

B.4. Stochastic integral w.r.t. a compensated Poisson random measure

Another way to introduce a stochastic integral for Lévy processes is motivated by the decomposition (B.13). Since for the Q-Wiener process the Itô integral is already defined, one only needs to define an integral for the stochastic process

$$\int_U x\tilde{N}(t,dx)$$

in order to find an stochastic integral with respect to a square integrable Lévy processes. This was done in [**Rüd04**] (it is called the strong integral of type 2). Since the idea is fundamentally the same as in the two cases above, we will only repeat the main ideas here.
First, let us define the σ-algebra $F_t := \mathscr{B}(\mathbb{R}_+ \times U\backslash\{0\}) \otimes \mathcal{F}_t$. For $T > 0$, we denote

$$\begin{aligned} M^T(U,H) := \{f : &\mathbb{R}_+ \times U\backslash\{0\} \times \Omega \to H, \text{ such that } f \text{ is } \mathcal{F}_T/\mathscr{B}(H) \text{ measurable} \\ &\text{and } f(t,x,\omega) \text{ is } \mathcal{F}_t\text{-adapted } \forall x \in U\backslash\{0\},\ \forall t \in [0,T]\} \end{aligned}$$

Now, we are able to introduce the so called natural integral as follows.

Definition B.46. *Let $t \in (0,T]$, A bounded below, $f \in M^T(U,H)$. Assume that $f(\cdot,\cdot,\omega)$ is Bochner integrable on $(0,T] \times A$ with respect to ν, for all $\omega \in \Omega$ fixed. The natural integral of f on $(0,t] \times A$ with respect to the compensated Poisson random measure $\tilde{N}(dx,dt)$ is*

$$\begin{aligned} &\int_0^t \int_A f(s,x,\omega)\tilde{N}(dx,ds) \\ &\quad := \sum_{0 \leqslant s \leqslant t} f(s,(\Delta M(s))(\omega),\omega)\mathbb{1}_A(\Delta M(s)(\omega)) - \int_0^t \int_A f(s,x,\omega)\nu(dx)ds,\ \omega \in \Omega, \end{aligned}$$

where the last term is understood as a Bochner integral (for $\omega \in \Omega$ fixed) of $f(s,x,\omega)$ with respect to the measure $\nu \otimes dt$.

Then, it is shown that the integral has the familiar form for the class of elementary processes (see [**Rüd04**, Proposition 3.5]) and is well-defined. Consequently the definition of the stochastic integral is extended to functions belonging to the following space

$$M_\nu^{T,2}(U,H) = \left\{ f \in M^T(U,H) : \int_0^T \int_U \mathbb{E}\|f(t,x)\|_H^2 \nu(dx)dt < \infty \right\}.$$

This is due, since the class of elementary processes is dense in $M_\nu^{T,2}(U,H)$ (see [**Rüd04**, Theorem 4.2]).

Remark B.47. *An easy calculation using* (B.12) *shows*

$$\|\Phi(s)Q_1^{1/2}\|_{L_{(HS)}(U,H)}^2 = \int_U \|\Phi(s)x\|_H^2 \nu(dx),$$

for $\Phi \in \mathcal{E}(U,H)$. Therefore, it follows that in the pure jump case

$$\Phi \in \mathcal{L}_T^2(U,H) \Longleftrightarrow \Phi \in M_\nu^{T,2}(U,H)$$

and

$$\int_0^t \Phi(s)dM(s) = \int_0^t \int_U \Phi(s)x\tilde{N}(dxds),$$

for all $t \in [0,T]$ *and* $\Phi \in \mathcal{L}_T^2(U,H)$. *For more detail see* [**App06**].

The final step is to weaken the condition for the function similar to the Wiener case. Hence, the stochastic integral is extended to function from the following class

$$N_\nu^{T,2}(U,H) = \left\{ f \in M^T(U,H) : \int_0^T \int_U \|f(t,x)\|_H^2 \nu(dx)dt < \infty \ \mathbb{P}\text{-a.s.} \right\}.$$

The complete construction can be found in [**Rüd04**, Section 7]. The extension to $N_\nu^{T,2}(U,H)$ is especially important later for the transformation formula, if we want to be able to use the norm square as a transformation function. But before we introduce the transformation formula we state the stochastic version of Lebesgues dominated convergence theorem for integrals with respect to a compensated Poisson random measure.

Theorem B.48 (stochastic Lebesgues theorem for integral with respect to compensated Poisson random measure). *Let* $f \in N_\nu^{T,2}(U,H)$ *be arbitrary and let* $(f_n)_{n\in\mathbb{N}}$ *be a sequence such that* $f_n \in N_\nu^{T,2}(U,H)$ *for all* $n \in \mathbb{N}$. *Suppose* $(f_n)_{n\in\mathbb{N}}$ *converges* $dt \otimes \nu \otimes \mathbb{P}$*-a.s. on* $[0,T] \times U \times \Omega$, *when* $n \to \infty$, *and* $\mathbb{P}$*-a.s.*

$$\lim_{n\to\infty} \int_0^T \int_U \|f_n - f\|^2 \nu(dx)dt = 0.$$

Assume, there is a $g \in N_\nu^{T,2}(U,H)$, *such that*

$$\int_0^T \int_U \|f_n\|^2 \nu(dx)dt \leqslant \int_0^T \int_U \|g\|^2 \nu(dx)dt.$$

Then, we have

$$\int_0^t \int_U f(s,x)\tilde{N}(ds,dx) = \lim_{n\to\infty} \int_0^t \int_U f_n(s,x)\tilde{N}(ds,dx),$$

where the limit is in probability.

PROOF.
See [**Rüd04**, Theorem 7.7 and Remark 7.8].

Before we can state the transformation formula we need the following definition.

Definition B.49 (quasi-sublinear function). *A continuous, nondecreasing function on* $\mathbb{R}_+$, *that is* $h : \mathbb{R}_+ \to \mathbb{R}_+$, *is called quasi-sublinear if there is a constant* $C > 0$, *such that*

$$h(x+y) \leqslant C(h(x) + h(y)), \ x, y \in \mathbb{R}_+,$$

$$h(xy) \leqslant Ch(x)h(y), \ x, y \in \mathbb{R}_+.$$

THEOREM B.50 (transformation formula for jump processes). *Suppose that*

- $\phi \in C^{1,2}(\mathbb{R}_+ \times H; \mathbb{R})$ *is a function, such that*
$$\|\phi'(s,y)\|_H \leqslant h_1(\|y\|_H),$$
$$\|\phi''(s,y)\|_{L(H)} \leqslant h_2(\|y\|_H),$$
for $(s,y) \in \mathbb{R}_+ \times H$ *and* h_1, h_2 *being quasi-sublinear function;*
- $A \in \mathscr{B}(U\backslash\{0\})$ *is a set with* $\nu(A^c) < \infty$, *where* A^c *denotes the compliment of* A*;*
- $f : \Omega \times \mathbb{R}_+ \times U \to H$ *is progressively measurable process, such that for all* $t \in \mathbb{R}_+$ *we have* $\mathbb{P}$*-a.s.*
$$\int_0^t \int_A \|f(s,x)\|_H^2 \nu(dx)ds + \int_0^t \int_A h_1(\|f(s,x)\|_H)^2 \|f(s,x)\|_H^2 \nu(dx)ds$$
$$+ \int_0^t \int_A h_2(\|f(s,x)\|_H)\|f(s,x)\|_H^2 \nu(dx)ds < \infty;$$
- $g : \Omega \times \mathbb{R}_+ \times U \to H$ *is a progressively measurable process;*
- Y *is a stochastic process of the from*
$$Y(t) = Y(0) + \int_0^t \int_A f(s,x)\tilde{N}(dx,ds) + \int_0^t \int_{A^c} g(s,x)N(dx,ds), \ t \geqslant 0.$$

Then the following statements hold:

- *For all* $t \in \mathbb{R}_+$ *we have* $\mathbb{P}$*-a.s.*
$$\int_0^t |\dot{\phi}(s, Y(s-))|ds < \infty,$$
$$\int_0^t \int_A |\phi(s, Y(s-) + f(s,x)) - \phi(s, Y(s-))|^2 \nu(dx)ds < \infty,$$
$$\int_0^t \int_A |\phi(s, Y(s-) + f(s,x)) - \phi(s, Y(s-)) - \phi'(s, Y(s-))f(s,x)|\nu(dx)ds < \infty,$$
$$\int_0^t \int_{A^c} |\phi(s, Y(s-) + g(s,x)) - \phi(s, Y(s-))|N(dx,ds) < \infty.$$
- *We have* $\mathbb{P}$*-a.s.*
$$\begin{aligned}\phi(t, Y(t)) = \ & \phi(0, Y(0)) + \int_0^t \dot{\phi}(s, Y(s-))ds \\ & + \int_0^t \int_A \phi(s, Y(s-) + f(s,x)) - \phi(s, Y(s-))\tilde{N}(dx,ds) \\ & + \int_0^t \int_A \phi(s, Y(s-) + f(s,x)) - \phi(s, Y(s-)) - \phi'(s, Y(s-))f(s,x)\nu(dx)ds \\ & + \int_0^t \int_{A^c} \phi(s, Y(s-) + g(s,x)) - \phi(s, Y(s-))N(dx,ds), \ t \geqslant 0.\end{aligned}$$

PROOF.

See [**MRT13**, Theorem 3.6].

B.5. Stochastic abstract Cauchy problem with Lévy noise

Let us consider the following abstract Cauchy problem in a separable Hilbert space H

$$(S-ACP) \quad \begin{cases} dX(t) &= (AX(t)+F(X(t)))dt + G(X(t))dM(t), \quad \text{for } t \geqslant 0, \\ \\ X(0) &= X_0, \end{cases}$$

where A, with domain $\mathscr{D}(A)$, is the generator of a C_0-semigroup $(S(t))_{t\geqslant 0}$ on H, M is a square integrable Lévy martingale taking values in a separable Hilbert space U, and X_0 is an $\mathcal{F}_0$-measurable random variable in H. For the nonlininearities we assume the following

(F) $F : \mathscr{D}(F) \to H$, $\mathscr{D}(F)$ is dense in H and there is a function $a : (0,\infty) \mapsto (0,\infty)$ satisfying $\int_0^T a(t)dt < \infty$ for all $T < \infty$, such that for all $t > 0$ and $x, y \in \mathscr{D}(F)$,

$$\|S(t)F(x)\|_H \leqslant a(t)(1 + \|x\|_H),$$
$$\|S(t)(F(x) - F(y))\|_H \leqslant a(t)\|x - y\|_H.$$

(G) We set $\mathcal{H} := Q^{1/2}(U)$. $G : \mathscr{D}(G) \to \tilde{L}(\mathcal{H}, H)$, where $\tilde{L}$ ist the class of all linear (not necessarily bounded) operators, $\mathscr{D}(G)$ is dense in H and assume there is a function $b : (0,\infty) \mapsto (0,\infty)$ satisfying $\int_0^T b(t)dt < \infty$ for all $T < \infty$, such that for all $t > 0$ and $x, y \in \mathscr{D}(G)$,

$$\|S(t)G(x)\|_{L_{(HS)}(\mathcal{H},H)} \leqslant b(t)(1 + \|x\|_H),$$
$$\|S(t)(G(x) - G(y))\|_{L_{(HS)}(\mathcal{H},H)} \leqslant b(t)\|x - y\|_H.$$

In order to receive a càdlàg solution of $(S-ACP)$ it will be necessary to strengthen condition (G).

(GI) Condition (G) holds, if $S(t) = I,\ t \geqslant 0$.

Remark B.51. *The function $t \mapsto \|S(t)\|_{L(H)}$ is bounded on any finite interval $[0,T]$. Thus, if $F : H \to H$ and $G : H \to L_{(HS)}(\mathcal{H}, H)$ are Lipschitz continuous, then (F) and (G) are satisfied.*

Definition B.52 (mild solution of $(S-ACP)$). *Let X_0 be a square integrable $\mathcal{F}_0$-measurable random variable in H. A predictable process $X : [0,\infty) \times \Omega \to H$ is called a mild solution to $(S-ACP)$ starting at time zero from X_0 if*

$$\sup_{t\in[0,T]} \mathbb{E}\|X(t)\|_H^2 < \infty \text{ for all } T \in (0,\infty),$$

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

$$X(t) = S(t)X_0 + \int_0^t S(t-s)F(X(s))ds + \int_0^t S(t-s)G(X(s))dM(s) \text{ for all } t \geqslant t_0.$$

THEOREM B.53 (existence and uniqueness of solution of $(S - ACP)$). *Assume that condition (F) und (G) are satisfied, then the following hold:*

(1) for every $\mathcal{F}_0$-measurable square integrable random variable X_0 in H there exists a unique (up to modification) mild solution $X(\cdot, X_0)$ of $(S - ACP)$.

(2) for all $T < \infty$ there exists a $L < \infty$, such that for all $x_1, x_2 \in H$,

$$\sup_{t \in [0,T]} \mathbb{E} \|X(t, x_1) - X(t, x_2)\|_H^2 \leqslant L \|x_1 - x_2\|_H^2.$$

If (F) and (GI) hold and S is a generalized contraction, then the solution has a càdlàg version.

PROOF.

See [**PZ07**, Theorem 9.15 and Theorem 9.29].

REMARK B.54. *In the finite dimensional theory presented for instance in* [**App09**] *or* [**Pro05**] *one looks for càdlàg solutions instead of predictable solutions. However, in infinite dimensions the solution my not have a càdlàg modification in H. An example of this can be found in* [**PZ07**, Section 9.4.4]. *That is why the condition on the solution process is weakened to be only predictable. For more detail on the concept of solution in infinite dimensions see* [**PZ07**, Section 9.2.1].

List of Figures

Figure 1 Path of Wiener process 1
Figure 2 Paths of fractional Brownian motions 2
Figure 3 Paths of Poisson process and compound Poisson process 2
Figure 4 Plot undelayed logistic growth model 3
Figure 5 Plots delayed logistic growth model 4

Bibliography

[AF00] Hyungsok Ahn and Raisa E. Feldman. Optimal filtering of a Gaussian signal in the presence of Lévy noise. *SIAM J. Appl. Math.*, 60(1):359–369 (electronic), 2000.

[AFZ97] Nasir Uddin Ahmed, Marco Fuhrman, and Jerzy Zabczyk. On filtering equations in infinite dimensions. *J. Funct. Anal.*, 143(1):180–204, 1997.

[AMRW08] Murray E. Alexander, Seyed M. Moghadas, Gergely Röst, and Jianhong Wu. A delay differential model for pandemic influenza with antiviral treatment. *Bull. Math. Biol.*, 70(2):382–397, 2008.

[AMS13] Sergio Albeverio, Elisa Mastrogiacomo, and Boubaker Smii. Small noise asymptotic expansions for stochastic PDE's driven by dissipative nonlinearity and Lévy noise. *Stochastic Process. Appl.*, 123(6):2084–2109, 2013.

[App06] David Applebaum. Martingale-valued measures, Ornstein-Uhlenbeck processes with jumps and operator self-decomposability in Hilbert space. In *In memoriam Paul-André Meyer: Séminaire de Probabilités XXXIX*, volume 1874 of *Lecture Notes in Math.*, pages 171–196. Springer, Berlin, 2006.

[App09] David Applebaum. *Lévy processes and stochastic calculus*, volume 116 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2009.

[AR05] S. Albeverio and B. Rüdiger. Stochastic integrals and the Lévy-Ito decomposition theorem on separable Banach spaces. *Stoch. Anal. Appl.*, 23(2):217–253, 2005.

[ARS13] John A. D. Appleby, Markus Riedle, and Catherine Swords. Bubbles and crashes in a Black-Scholes model with delay. *Finance Stoch.*, 17(1):1–30, 2013.

[Åst70] Karl J. Åström. *Introduction to stochastic control theory.* Mathematics in Science and Engineering, Vol. 70. Academic Press, New York-London, 1970.

[BH76] J. A. Burns and T. L. Herdman. Adjoint semigroup theory for a class of functional differential equations. *SIAM. J. Math. Anal.*, 7(5):729–745, 1976.

[BK14] Fred Espen Benth and Paul Krühner. Representation of infinite-dimensional forward price models in commodity markets. *Commun. Math. Stat.*, 2(1):47–106, 2014.

[BL12] Andrea Barth and Annika Lang. Simulation of stochastic partial differential equations using finite element methods. *Stochastics*, 84(2-3):217–231, 2012.

[BP05] András Bátkai and Susanna Piazzera. *Semigroups for delay equations*, volume 10 of *Research Notes in Mathematics*. A K Peters Ltd., Wellesley, MA, 2005.

[BvGL09] Joris Bierkens, Onno van Gaans, and Sjoerd Verduyn Lunel. Existence of an invariant measure for stochastic evolutions driven by an eventually compact semigroup. *J. Evol. Equ.*, 9(4):771–786, 2009.

[CG11] Sonja Cox and Mariusz Górajski. Vector-valued stochastic delay equations—a semigroup approach. *Semigroup Forum*, 82(3):389–411, 2011.

[CM78] Anna Chojnowska-Michalik. Representation theorem for general stochastic delay equations. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 26(7):635–642, 1978.

[DPZ92] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1992.

[DvGVLW95] Odo Diekmann, Stephan A. van Gils, Sjoerd M. Verduyn Lunel, and Hans-Otto Walther. *Delay equations*, volume 110 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1995. Functional, complex, and nonlinear analysis.

[EN00] Klaus-Jochen Engel and Rainer Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.

[Eps90] Irving R Epstein. Differential delay equations in chemical kinetics: Some simple linear model systems. *The Journal of Chemical Physics*, 92(3):1702–1712, 1990.

[FIM07] G. Fragnelli, A. Idrissi, and L. Maniar. The asymptotic behavior of a population equation with diffusion and delayed birth process. *Discrete Contin. Dyn. Syst. Ser. B*, 7(4):735–754 (electronic), 2007.

[FSX13] Rüdiger Frey, Thorsten Schmidt, and Ling Xu. On Galerkin approximations for the Zakai equation with diffusive and point process observations. *SIAM J. Numer. Anal.*, 51(4):2036–2062, 2013.

[FT04] G. Fragnelli and L. Tonetto. A population equation with diffusion. *J. Math. Anal. Appl.*, 289(1):90–99, 2004.

[GM11a] Leszek Gawarecki and Vidyadhar Mandrekar. *Stochastic differential equations in infinite dimensions with applications to stochastic partial differential equations*. Probability and its Applications (New York). Springer, Heidelberg, 2011.

[GM11b] B. Grigelionis and R. Mikulevicius. Nonlinear filtering equations for stochastic processes with jumps. In *The Oxford handbook of nonlinear filtering*, pages 95–128. Oxford Univ. Press, Oxford, 2011.

[Gór14] Mariusz Górajski. Vector-valued stochastic delay equations—a weak solution and its Markovian representation. *Nonlinear Anal.*, 103:55–71, 2014.

[GS71] I. I. Gichman and A. W. Skorochod. *Stochastische Differentialgleichungen.* Akademie-Verlag, Berlin, 1971. Deutsche Übersetzung von dem Russischen durch Jurgis Szlaza, In deutscher Sprache herausgegeben von Helga Bunke, Mathematische Lehrbücher und Monographien, II. Abteilung: Mathematische Monographien, Band XXV.

[GT08] Wilfried Grecksch and Constantin Tudor. A filtering problem for a linear stochastic evolution equation driven by a fractional Brownian motion. *Stoch. Dyn.*, 8(3):397–412, 2008.

[GWMS14] I.I. Gihman, K. Wickwire, Y.A. Mitropolski, and A.V. Skorohod. *Stochastic Differential Equations.* Ergebnisse der Mathematik und ihrer Grenzgebiete. 2. Folge. Springer Berlin Heidelberg, 2014.

[HG13] Erika Hausenblas and Ankik Kumar Giri. Stochastic Burgers equation with polynomial nonlinearity driven by Lévy process. *Commun. Stoch. Anal.*, 7(1):91–112, 2013.

[HMN91] Yoshiyuki Hino, Satoru Murakami, and Toshiki Naito. *Functional-differential equations with infinite delay*, volume 1473 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1991.

[HVL93] Jack K. Hale and Sjoerd M. Verduyn Lunel. *Introduction to functional-differential equations*, volume 99 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1993.

[Itô44] Kiyosi Itô. Stochastic integral. *Proc. Imp. Acad. Tokyo*, 20:519–524, 1944.

[Itô51] Kiyosi Itô. On a formula concerning stochastic differentials. *Nagoya Math. J.*, 3:55–65, 1951.

[Kap85] F. Kappel. Galerkin type approximation schemes for delay systems. *Ann. Differential Equations*, 1(1):57–82, 1985.

[Kra63] N. N. Krasovskiĭ. *Stability of motion. Applications of Lyapunov's second method to differential systems and equations with delay.* Translated by J. L. Brenner. Stanford University Press, Stanford, Calif., 1963.

[Kry80] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Applications of Mathematics*. Springer-Verlag, New York-Berlin, 1980. Translated from the Russian by A. B. Aries.

[Kus08] Harold J. Kushner. *Numerical methods for controlled stochastic delay systems*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 2008.

[KW67] Hiroshi Kunita and Shinzo Watanabe. On square integrable martingales. *Nagoya Math. J.*, 30:209–245, 1967.

[LD14] Chun Lu and Xiaohua Ding. Permanence and extinction of a stochastic delay logistic model with jumps. *Math. Probl. Eng.*, pages Art. ID 495275, 8, 2014.

[Liu06] Kai Liu. *Stability of infinite dimensional stochastic differential equations with applications*, volume 135 of *Chapman & Hall/CRC Monographs and Surveys in Pure and Applied Mathematics*. Chapman & Hall/CRC, Boca Raton, FL, 2006.

[Mao94] Xuerong Mao. *Exponential stability of stochastic differential equations*, volume 182 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 1994.

[MBP04] Thilo Meyer-Brandis and Frank Proske. Explicit solution of a non-linear filtering problem for Lévy processes with application to finance. *Appl. Math. Optim.*, 50(2):119–134, 2004.

[Mét82] Michel Métivier. *Semimartingales*, volume 2 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin-New York, 1982. A course on stochastic processes.

[Mey76] P. A. Meyer. Un cours sur les intégrales stochastiques. In *Séminaire de Probabilités, X (Seconde partie: Théorie des intégrales stochastiques, Univ. Strasbourg, Strasbourg, année universitaire 1974/1975)*, pages 245–400. Lecture Notes in Math., Vol. 511. Springer, Berlin, 1976.

[Mis08] Yuliya S. Mishura. *Stochastic calculus for fractional Brownian motion and related processes*, volume 1929 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2008.

[MLK06] Athena Makroglou, Jiaxu Li, and Yang Kuang. Mathematical models and software tools for the glucose-insulin regulatory system and diabetes: an overview. *Appl. Numer. Math.*, 56(3-4):559–573, 2006.

[MP80] M. Métivier and J. Pellaumail. On a stopped Doob's inequality and general stochastic equations. *Ann. Probab.*, 8(1):96–114, 1980.

[MRT13] Vidyadhar Mandrekar, Barbara Rüdiger, and Stefan Tappe. Itō's formula for Banach-space-valued jump processes driven by Poisson random measures. In *Seminar on stochastic analysis, random fields and applications VII. Centro Stefano Franscini, Ascona (Ticino), Switzerland, May 23–27, 2011*, pages 171–186. Basel: Birkhäuser/Springer, 2013.

[MVN68] Benoit B. Mandelbrot and John W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 10:422–437, 1968.

[MW11] V. Mandrekar and Li Wang. Asymptotic properties of stochastic partial differential equations in Hilbert spaces driven by non-Gaussian noise. *Commun. Stoch. Anal.*, 5(2):309–331, 2011.

[MY06] Xuerong Mao and Chenggui Yuan. *Stochastic differential equations with Markovian switching*. Imperial College Press, London, 2006.

[Øks03] Bernt Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.

[Paz83] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[Pro05] Philip E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[PS09] S. Popa and S. S. Sritharan. Nonlinear filtering of Itô-Lévy stochastic differential equations with continuous observations. *Commun. Stoch. Anal.*, 3(3):313–330, 2009.

[PZ07] S. Peszat and J. Zabczyk. *Stochastic partial differential equations with Lévy noise*, volume 113 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2007. An evolution equation approach.

[Rie14] Markus Riedle. Stochastic integration with respect to cylindrical Lévy processes in Hilbert spaces: an L^2 approach. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 17(1):1450008, 19, 2014.

[Rie15] Markus Riedle. Ornstein-Uhlenbeck Processes Driven by Cylindrical Lévy Processes. *Potential Anal.*, 42(4):809–838, 2015.

[Rou96] Marc R Roussel. The use of delay differential equations in chemical kinetics. *The journal of physical chemistry*, 100(20):8323–8330, 1996.

[Rüd04] B. Rüdiger. Stochastic integration with respect to compensated Poisson random measures on separable Banach spaces. *Stoch. Stoch. Rep.*, 76(3):213–242, 2004.

[Sch13] Michael Scheutzow. Exponential growth rate for a singular linear stochastic delay differential equation. *Discrete Contin. Dyn. Syst. Ser. B*, 18(6):1683–1696, 2013.

[SI01] Didier Sornette and Kayo Ide. The Kalman-Lévy filter. *Phys. D*, 151(2-4):142–174, 2001.

[SK01] Marek Szydłowski and Adam Krawiec. The Kaldor-Kalecki model of business cycle as a two-dimensional dynamical system. *J. Nonlinear Math. Phys.*, 8(suppl.):266–271, 2001. Nonlinear evolution equations and dynamical systems (Kolimbary, 1999).

[Sko75] A. V. Skorohod. On a generalization of the stochastic integral. *Teor. Verojatnost. i Primenen.*, 20(2):223–238, 1975.

[Str64] R. L. Stratonovič. A new form of representing stochastic integrals and equations. *Vestnik Moskov. Univ. Ser. I Mat. Meh.*, 1964(1):3–12, 1964.

[vNR07] Jan van Neerven and Markus Riedle. A semigroup approach to stochastic delay equations in spaces of continuous functions. *Semigroup Forum*, 74(2):227–239, 2007.

[Vol28] V. Volterra. Sur la théorie mathématique des phénomènes héréditaires. *Journal de Mathématiques Pures et Appliquées*, pages 249–298, 1928.

[Vol90] Vito Volterra. *Leçons sur la théorie mathématique de la lutte pour la vie*. Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics]. Éditions Jacques Gabay, Sceaux, 1990. Reprint of the 1931 original.

[VV14] Almut E. D. Veraart and Luitgard A. M. Veraart. Modelling electricity day-ahead prices by multivariate Lévy semistationary processes. In *Quantitative energy finance*, pages 157–188. Springer, New York, 2014.

[Wu96] Jianhong Wu. *Theory and applications of partial functional-differential equations*, volume 119 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1996.

[XPL14] Yong Xu, Bin Pei, and Yongge Li. An averaging principle for stochastic differential delay equations with fractional Brownian motion. *Abstr. Appl. Anal.*, pages Art. ID 479195, 10, 2014.

[Zei85] Eberhard Zeidler. *Nonlinear functional analysis and its applications. III*. Springer-Verlag, New York, 1985. Variational methods and optimization, Translated from the German by Leo F. Boron.

[ZL14] Yanchao Zang and Junping Li. Stability in distribution of neutral stochastic partial differential delay equations driven by α-stable process. *Adv. Difference Equ.*, pages 2014:13, 16, 2014.

[ZYL14] Wen Zhang, Jun Ye, and Haibo Li. Stability with general decay rates of stochastic differential delay equations with Poisson jumps and Markovian switching. *Statist. Probab. Lett.*, 92:1–11, 2014.

Index

$\mathscr{A}$, 14
$\mathscr{A}^*$, 79, 82
A_n, 100
A^*, 62
A^c, 58
(ACP_p), 13
adapted process, 102

$\mathscr{B}(U)$, 102
$BUC^k(\mathcal{M}, \mathbb{R})$, 63
bounded below, 109

C^*, 43
$C^k(H, \mathbb{R})$, 105
C_0-semigroup, 95
càdlàg process, 108
classical solution of (DE_p), 12
compensated Poisson integral, 110
compensated Poisson random measure, 110
complement of A, 58
compound Poisson process, 108

$\Delta L(t)$, 108
$\mathscr{D}(A)$, 95
(DE_p), 11
delay operator, 11
dissipative operator, 98
dual mapping, 72

ϵ_ϱ, 11
$\mathbb{E}$, 102
$\mathscr{E}_p$, 11
$\mathscr{E}(U, H)$, 104
elementary process, 104

$F(t, u)$, 24
$\mathscr{F}(t, u)$, 25
$\mathcal{F}_t$, 102
Fréchet derivative, 105

$G(t, u)$, 25
$G(t, u)^\natural$, 72
$\mathscr{G}(t, u)$, 25
generator of semigroup, 95

h_u, 11
history function, 11
history segment, 11

I, 10
$\mathcal{I}$, 24
$\int_0^t \Phi(s) dM(s)$, 113
$\int_0^t \int_A f(s, x) \tilde{N}(dx, ds)$, 117
$\int_0^t \Phi(s) dW(s)$, 104
integrated solution for $(S - DE)$, 30
integrated solution of $(S_D - ACP)$, 27
isometry for the Lévy noise case, 113
Itô formula, 105
Itô integral, 104
Itô integral a with random limit, 105
Itô isometry, 104

$L(X)$, 95
$L(X, Y)$, 95
$L_{(HS)}(U, H)$, 103
$\mathscr{L}^2(U, H)$, 103
$\mathcal{L}_T^2(U, H)$, 114
Lévy-Khinchin decomposition, 110
Lévy measure, 109
Lévy process, 108
Lumer-Phillips Theorem, 98

$M(t)$, 24
$\mathcal{M}(t)$, 25
μ, 11
$M^T(U, H)$, 117
$M_\nu^{T,2}(U, H)$, 117
mild solution for $(S_D - ACP)$, 26
mild solution of (DE_p), 17

mild solution of $(S-ACP)$, 120
mild solution of (ACP), 97

ν, 109
$N(t,A)$, 109
$\tilde{N}(t,A)$, 110
$N_\nu^{T,2}(U,H)$, 118

$(\Omega,\mathcal{F},\mathcal{F}_{t\geqslant 0},\mathbb{P})$, 102
ODE, 4

Φ, 11
$\dot{\phi}$, 105
ϕ', 105
$\mathcal{P}_T$, 102
$\mathcal{P}_T^2(U,H)$, 105
PDE, 4
Poisson integral, 110
Poisson random measure of a Lévy process, 109
predictable process, 102
process of lefts limits, 37
process of jumps, 108
product formula, 84
progressively measurable process, 102

Q, 103
Q_0, 112
Q_1, 112
quasi-sublinear function, 118
Q-Wiener process, 103

ϱ, 11
$R(n)$, 99

$(S-ACP)$, 120
(S_D-ACP), 25, 35
$(S-DE)$, 25
semigroup of contractions, 96
semigroup of generalized contractions, 96
SPDE, 5
standing hypotheses (deterministic case), 11
standing hypotheses (stochastic case), 24
stochastic Fubini theorem, 115
stochastic integral w.r.t. a cPoisson random measure, 117
stochastic integral w.r.t. a square integrable Lévy process, 113
stochastic Lebesgues theorem for integral w.r.t. cPoisson random measure, 118
stochastic Lebegues for Wiener noise, 106

$\mathscr{T}(t)$, 15
transformation formula for integral w.r.t. cPoisson random measure, 119
transformation formula for mild solution of $(S-ACP)$, 68
transformation formula for Square integrable Lévy process, 58

u_t, 11

$W(t)$, 103

$X(t-)$, 108
$\mathcal{X}_{T,B}$, 115

Yosida approximation for generator; contraction case, 100
Yosida approximation for generator; general case, 100
Yosida approximation of an element in H, 99
Yosida approximation of mild solution of (S-ACP); general case, 49
Yosida approximation of mild solution of (S-ACP); generalized contraction case, 39

Erklärung an Eides statt

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegeben Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Halle (Saale), 21. November 2015

Frank Wusterhausen

Curriculum Vitae

Personal Data

Name:	Frank Wusterhausen
Date of birth:	12.04.1983
Place of birth:	Halle (Saale)
Nationality:	German
Martial status:	Single

Education

since July 2015	consultant at d-fine GmbH, Frankfurt
2011 - 2015	Ph.D. student at the department of mathematics, Martin-Luther-University Halle-Wittenberg
Sept. 2011	Diploma in mathematics title of the diploma thesis: *Schrödinger Equation with Noise on the Boundary*
2005-2011	studies in mathematics (major) and economics (minor), Martin-Luther-University Halle-Wittenberg
July 2010	Diploma in economathematics title of the diploma thesis: *An Analysis of Path-Dependent Options*
2003-2010	studies in economathematics, Martin-Luther-University Halle-Wittenberg
2007-2008	studies in mathematics, University of Toronto
July 2003	"Abitur" (A-Levels)
1993-2003	"Thomas-Müntzer-Gymnasium" grammar school, Halle (Saale)

Scholarships & Awards

2012-2014	Friedrich Naumann Foundation for Freedom, PhD scholarship
2011-2012	graduate fund of the state Saxony-Anhalt, PhD scholarship
2011-2014	travel grant, Martin-Luther-University Halle-Wittenberg
Oct. 2011	award of the "Georg-Cantor-Vereinigung" for the diploma thesis
2007-2008	DAAD (German Academic Exchange Service), scholarship for academic year at the University of Toronto
2006-2009	German National Academic Foundation (Studienstiftung des Deutschen Volkes), scholarship
2004-2010	Friedrich Naumann Foundation for Freedom, scholarship
2000-2001	US Congress and the German Parliament, scholarship of the Congress Bundestag Youth Exchange Program

Talks and Posters at Conferences

18.03.2015	Workshop on Stochastic Models and Control, Kaiserslautern, Germany
09.07.2014	AIMS Conference on Dynamical Systems, Differential Equations and Applications, Madrid, Spain
31.03-03.04.2014	Stochastic Processes and Differential Equations in Infinite Dimensional Spaces, London, United Kingdom *(poster)*
06.03.2014	German Probability and Statistics Days, Ulm, Germany
02.08.2013	9th meeting of German-speaking PhD-students in stochastics, Göttingen, Germany
21.06.2013	three-states workshop in stochastic analysis, Jena, Germany
08.06.2013	German-Polish Joint Conference on Probability and Mathematical Statistics, Torun, Poland
20.03.2013	Workshop on Stochastic Models and Control, Berlin, Germany
02.08.2012	8th meeting of German-speaking PhD-students in stochastics, Freudenstadt, Germany
04.07.2012	AIMS Conference on Dynamical Systems, Differential Equations and Applications, Orlando, Florida, USA
09.03.2012	German Probability and Statistics Days, Mainz, Germany

Organized Conferences

06.08-08.08.2014	10th meeting of German-speaking PhD-students in stochastics, Halle, Germany

Publications

1. F. Wusterhausen. An Analysis of Path-Dependent Options. *Journal of Optimization Theory and Applications*, online first, 2013.
2. F. Wusterhausen. Schrödinger equation with noise on the boundary. *AMIS Journals Volume 2013 (special Issue)*, 791 – 796, 2013.

Halle (Saale), 21. November 2015

Frank Wusterhausen