Some problems in Abstract Stochastic Differential Equations on Banach spaces

P V Crewe

St John's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Mathematics

Michaelmas Term 2011

Acknowledgements

There are many people and groups to thank for their help and support in the course of writing this thesis, but principal among them is my supervisor Charles Batty. I remember my first Oxford tutorial in Charles' office more than seven years ago (that first problem sheet was a rather poor effort) and through the course of two degrees and countless tutorial and supervision sessions Charles has been an unstinting source of support and inspiration.

Outside of mathematics, much of my life at university has been dominated by rowing in one way or another. I arrived at St. John's in 2004 as a skinny, unfit, eighteen year old with little sporting background but thanks to St. John's College Boat Club and later the Oxford University Lightweight Rowing Club I have rowed on some of the most beautiful stretches of water in Europe and along side some of the fastest crews in Britain. I have won Blades in the top division of Summer VIIIs and raced two Lightweight Boat Races against Cambridge, the second as the winning President. To everyone who coached, coxed or rowed with me in that time, you are the people who made every pre-dawn start, every hour on the erg and every outing with ice growing on the riggers worth it. Also, thanks again to Charles, not just for tolerating my rowing habit, but for taking an active interest, for supporting from the riverbank and for trusting me to manage my work when other supervisors might have taken a dim view of the hours involved.

In the summer of 2010 I was invited to the Delft University of Technology, by Professor Jan van Neerven. As well as for being the source of much of the work that inspired this thesis, I'd like to thank Jan, together with Mark Veraar, Sonja Cox and the rest of the group at Delft for interesting and useful discussions and for being such gracious hosts.

Finally, thanks to Junaid, Matt, David and Felix for making the functional analysis office a place I wanted to work and to my family for being there when I needed them.

Abstract

This thesis studies abstract stochastic differential equations on Banach spaces. The well-posedness of abstract stochastic differential equations on such spaces is a recent result of van Neerven, Veraar and Weis, based on the theory of stochastic integration of Banach space valued processes constructed by the same authors.

We study existence and uniqueness for solutions of stochastic differential equations with (possibly infinite) delay in their inputs on UMD Banach spaces. Such problems are also known as functional differential equations or delay differential equations. We show that the methods of van Neerven *et al.* extend to such problems if the initial history of the system lies in a space of a type introduced by Hale and Kato. The results are essentially of a fixed point type, both autonomous and non-autonomous cases are discussed and an example is given.

We also study some long time properties of solutions to these stochastic differential equations on general Banach spaces. We show the existence of solutions to stochastic problems with almost periodicity in a weak or distributional sense. Results are again given for both autonomous and non-autonomous cases and depend heavily on estimates for R-bounds of operator families developed by Veraar. An example is given for a second order differential operator on a domain in $\mathbb{R}^d$.

Finally we consider the existence of invariant measures for such problems. This extends recent work of van Gaans in Hilbert spaces to Banach spaces of type 2.

Contents

Chapter 1

Introduction

1.1 Background

The subject of this thesis is abstract stochastic differential equations on Banach spaces and some natural extensions to the well posedness results of van Neerven, Veraar and Weis in [69] and related papers. Given existence and uniqueness of solutions to a new class of equations, it is natural to ask in what ways can we perturb the problem while preserving well posedness, or if solutions to this new class of problems exhibit interesting properties that have been studied in related systems.

Abstract stochastic differential equations are a powerful tool to study stochastic partial differential equations. Consider a problem on a domain $\mathcal{O} \subset \mathbb{R}^d$

$$\frac{\partial}{\partial t} u(t, \xi) = \big(L(t) u(t, \cdot)\big)(\xi) + f\big(t, \xi, u(t, \xi)\big) + \sum_{n=1}^{\infty} g_n\big(t, \xi, u(t, \xi)\big) \partial W_n(t), \tag{1.1}$$

with some initial and boundary conditions, where $L(t)$ is a family of differential operators and (W_n) is a sequence of independent Brownian motions. Examples include stochastic heat and wave equations. It is often the case that (1.1) can be put in the form of an abstract Cauchy problem

$$\mathrm{d}U(t) = \Big[A(t)u(t) + F(t, U(t))\Big] \,\mathrm{d}t + G(t, U(t)) \,\mathrm{d}W(t) \tag{1.2}$$

on some Banach space E of functions, where W is a cylindrical Brownian motion on an associated Hilbert space H (our model of "space-time white noise"). The hope is that then (1.2) can be studied using the techniques of operator theory and evolution equations. This approach can give insight into many qualitative properties of the problem (1.1), such as well-posedness (existence and uniqueness of solutions in various senses), regularity (continuity, smoothness, differentiability and more abstract notions which lie between these well known properties) and asymptotics (long term rates of growth and decay, as well as global behaviours such as almost periodicity, ergodicity and invariant measures).

Operator theoretic techniques and abstract Cauchy problems were developed in the first half of the 20th century as a tool for studying partial differential equations. A key notion is that the solution of a well-posed linear homogeneous problem

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}u(t) &= Au(t), \quad t \geq 0, \\ u(0) &= x \in E, \end{cases} \tag{1.3}$$

is of the form $u(t) = S(t)x$ for some family $S(t)$ of linear operators on the Banach space E. $S(t)$ is called a *semigroup* in view of the fact that

$$S(0)x = x \quad \text{and} \quad S(t+s)x = S(t)S(s)x$$

for all $x \in E$. In a strictly algebraic sense, any operator semigroup $S(t)$ is a representation of the same semigroup $(\mathbb{R}_+, +)$, however, such language will not feature again in this thesis. In general the map $t \mapsto S(t)$ is not continuous with respect to the norm on $\mathcal{L}(E)$, but if (1.3) has continuous solutions for any initial condition $x \in E$ then the map $t \mapsto S(t)x = u(t)$ is continuous for all $x \in E$ and $S(t)$ is said to be continuous in the strong operator topology, or a C_0-semigroup. The operator A in (1.3) is called the generator of $S(t)$ and we shall see (Theorem 2.5) that generation of a strongly continuous semigroup characterises well-posedness of (1.3).

Strongly continuous semigroups are the building blocks of solutions to more complicated problems. A problem is called semi-linear if it is composed of a linear part which is "large" and dominates the behaviour together with a non-linear part which is "small" in some relative sense. Consider the problem

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}u(t) &= Au(t) + F(t, u(t)), \quad t \geq 0 \\ u(0) &= x \in E. \end{cases} \tag{1.4}$$

Under the assumption that we can solve the linear problem (1.3) (i.e. that we have knowledge of the semigroup $S(t)$ generated by A) and that F is sufficiently well behaved, the solution of (1.4) is given by the variation of parameters formula

$$u(t) = S(t)x + \int_0^t S(t-s)F(s,u(s))\,\mathrm{d}s, \quad t \geq 0. \tag{1.5}$$

Typically the solution $u(t)$ to the integral equation (1.5) is then shown to exist by fixed point methods in some suitable space and this solution is considered to solve (1.4) in a mild sense.

1.1.1 Stochastic integration

In this thesis we are interested in problems such as (1.2) where there is an additional random perturbation by some space-time white noise term, which may or may not depend on the state $u(t)$ of the system. To make sense of such problems in an infinite dimensional setting it is necessary to understand the associated integral equation

$$u(t) = S(t)x + \int_0^t S(t-s)F(s,u(s))\,\mathrm{d}s + \int_0^t S(t-s)G(s,u(s))\,\mathrm{d}W(s), \quad t \geq 0, \tag{1.6}$$

and hence we need a theory of stochastic integration for functions taking values in infinite dimensional spaces.

Since the classical Itô integral extends in a natural way to Hilbert spaces, abstract stochastic problems such as (1.2) have been studied by semigroup methods on Hilbert spaces since the 1970s. Original papers on the subject include Curtain and Falb [26] and Dawson [35] and [36], but the field reached some maturity with the work of Da Prato *et al.* in the 1980s, [29], [30] and [33]. One of the most significant developments of this period is the factorisation method of [29], which formed the basis for most future study of regularity for stochastic equations and which generalises readily to Banach spaces.

Some of the early moves to handle equations such as (1.2) in a wider class of spaces were by Brzeźniak, [20] and [21] in the mid 1990s, where the stochastic integral is extended from a Hilbert space setting to Banach spaces of martingale type 2. In the early 2000s the theory developed further, with stochastic integrals in arbitrary Banach spaces constructed for the first time. Initially, existence and uniqueness of equations with constant noise terms $G(t,u(t)) = G \in \mathcal{L}(H,E)$ were considered on separable spaces by Brzeźniak and van Neerven [23], making sense of the stochastic convolution

$$\int_0^t S(t-s)G\,\mathrm{d}W(s)$$

by embedding E in a larger Hilbert space and using Hilbert space methods. Equations with arbitrary deterministic inhomogeneous terms $G(t)$ came next in work of van Neerven and Weis [70], who constructed a theory of integration intrinsically in the Banach space E and obtained a characterisation of functions $\phi : [0,T] \to \mathcal{L}(H,E)$, with any Banach space E, for which

$$\int_0^T \phi(s)\,\mathrm{d}W(s)$$

is well defined. It is in this paper that stochastic integration is first expressed in terms of γ-radonifying operators. The theory is detailed in Chapter 3, but the main feature is that stochastically integrable functions $\phi : [0,T] \to \mathcal{L}(H,E)$ form a Banach space, written as $\gamma(L^2(0,T;H),E)$, and the norm $\|\cdot\|_\gamma$ on this space gives a two sided Itô type isometry for the stochastic integrals

$$c^{-1}\|\phi\|_\gamma^2 \leq \mathbb{E}\Big\|\int_0^T \phi(s)\,\mathrm{d}W(s)\Big\|^2 \leq c\|\phi\|_\gamma^2. \tag{1.7}$$

Multiplicative noise terms, (i.e. equations where G depends on $U(t)$) and the associated theory of stochastic integration for integrands that are themselves random processes were a significantly greater challenge. In [68], van Neerven, Veraar and Weis complete the search for a good theory of stochastic integration on Banach spaces by building an integral for processes $\Phi : [0,T] \times \Omega \to \mathcal{L}(H,E)$. To do this they introduce a new Brownian motion on a different probability space and use a "decoupling inequality" of Garling [44] in order to reduce the problem to that of deterministic integrands. Garling's inequality requires the restriction to spaces with the UMD property, a class which includes L^p for $1 < p < \infty$, but through certain tricks, one sided estimates are still available in a slightly larger class that includes L^1. Once again the result is a Banach space of integrable processes along with a two sided Itô isometry of the form of (1.7).

In the paper [69], the same authors apply this theory to existence, uniqueness and regularity of (1.2). Fixed point methods are used in a precisely chosen space of continuous processes, the semigroup is assumed to be analytic and the factorisation method of Da Prato [29] is used heavily in the course of highly sophisticated estimates on a fixed point operator. In [86], Veraar studies the associated non-autonomous problem (that is, where $A(t)$ is a family of operators which change with time, rather than being a fixed operator A) on a scale of spaces that is slightly restricted in order to give a larger class of integrable processes. Further discussion of the methods used appears in Section 4.1.

Hence we approach this thesis with a well developed theory of stochastic integration in place and existence and uniqueness of solutions to (1.2) firmly established. It is natural then to ask whether related problems which have been studied on Hilbert spaces or in the context of deterministic evolution equations also extend to this new setting. We show in the following chapters that several such extensions are possible.

1.1.2 Delay equations

An active field of research in engineering, materials science, biology and financial mathematics is stochastic delay equations. These problems, also known as functional differential equations, are ones in which the evolution of the system depends not just on the current state, but also on the past states of the system in some way. If the system looks back only a finite time into the past then such equations can be modelled with semigroup methods by lifting the problem from E up to a new problem on a space of functions from the past into E. Such methods are detailed in [39, Section VI.6] and the book of Batkai and Piazzera [11]. The earliest systematic investigation of equations with delay was by Volterra in the early 20^{th} century, but the use of semigroup methods on appropriate function spaces dates to Krasovskiĭ [56]. The general theory was first formulated in the modern form by Hale in the 1960s and became well established with the book [49], with the non- and semi-linear version coming later by Webb [88].

Suppose we are given a problem of the form

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}u(t) &= Au(t) + F(t, u_t), \quad t \geq 0, \\ u(0) &= x \in E, \\ u_0 &= f : [-1, 0] \to E, \end{cases} \tag{1.8}$$

where A is the generator of a C_0-semigroup and $F : \mathbb{R}_+ \times L^1(-1, 0; E) \to E$ is sufficiently well behaved. Here $u_t : [-1, 0] \to E$ is the history function of u at a time $t \geq 0$, defined by $u_t(s) = u(t + s)$ for $-1 \leq s \leq 0$ (alternatively, one can think of u_t as just the restriction of u to the interval $[t - 1, t]$ and then shifted by t to the left). Essentially the idea is to introduce a new operator $\mathcal{A}$ on the space $E \times L^1(-1, 0; E)$

which generates a new "delay semigroup" $\mathcal{S}(t)$. The problem is thereby reduced to one of the form of (1.4) or (1.5) on the space $E \times L^1(-1, 0; E)$.

The hope in starting this project was to show that stochastic delay differential equations of the form

$$\begin{cases} \mathrm{d}U(t) &= \Big[Au(t) + F(t, U_t)\Big]\,\mathrm{d}t + G(t, U_t)\,\mathrm{d}W(t), \quad t \geq 0, \\ U(0) &= X, \\ U_0 &= \Phi, \end{cases} \tag{1.9}$$

can be well posed on general UMD Banach spaces, following in the footsteps of the work of van Neerven *et al.* [68], [69]. An immediate problem is that while the paper [69] requires the semigroup $S(t)$ to be analytic in order to obtain results, the delay semigroup $\mathcal{S}(t)$ is never analytic. Since we cannot apply the results of [69] directly, we instead set out to replicate the fixed point arguments therein to show well posedness of (1.9) from first principles.

There already exists a body of literature devoted to working with delay equations by fixed point methods, and that is in the study of equations with infinite delay. Infinitely delayed equations, in contrast to finite delay models such as (1.8), evolve based on the entire history of the system and hence never "forget" prior states. Because of this, the precise *phase* or *history* space of functions $(-\infty, 0] \to E$ in which the initial data lives takes on additional importance in relation to questions of well-posedness and regularity. The first general result to give an axiomatic description of admissible history spaces is by Hale and Kato [50], but most modern work such as [52] uses the terminology from the book of Hino [53].

Stochastic delay equations are studied in Hilbert spaces by Liu in [60] and [61] and in Banach spaces of type 2 by Cox and Górajski [25]. Another approach taken by Riedle [76] and van Neerven and Riedle [77] is to consider stochastic delay equations with the history space $C([-1, 0]; E)$, wherein the solutions lie not in $C([-1, 0]; E)$ but in the bi-dual of this space.

The problem studied in Chapter 4 is to extend the well-posedness results of [69] and [86] to abstract semi-linear stochastic delay equations with infinite delay on general UMD spaces with initial data living in a history space of the form of Hale and Kato ([50], [53]).

1.1.3 Almost periodicity and invariant measures

Another natural extension to the results of [69] is to ask if, and under what conditions, solutions exists with various long time properties. Two properties which we investigate in Chapters 5 and 6 are almost periodicity and invariance of measures.

Almost periodic functions were first defined and studied by Bohr in the 1920s [18] as the closure of the trigonometric polynomials with respect to the uniform norm on $\mathbb{R}$. Several equivalent definitions soon followed and the most common today is that of Bochner [16], which says, essentially, that a function $f : \mathbb{R} \to E$ is almost periodic if for any $\varepsilon > 0$, f comes within ε of being periodic if one looks at sufficiently long "almost periods". Almost periodic processes abound in nature with applications from planetary motion to signal processing, where their usefulness stems from being far more flexible than purely periodic functions. The existence of an almost periodic solution to a given problem is a well studied question in the theory of evolution equations.

Two common approaches to the problem are either to assume spectral properties of the generator A, typically some condition on the countability of $\sigma(A) \cap i\mathbb{R}$ [6, 12], or to assume some dissipativity condition on the semigroup [31]. When studying stochastic equations there is the problem of how to interpret almost periodicity for random processes. It is clear that we cannot expect almost periodicity to hold for stochastic problems in a pathwise sense, i.e. we cannot expect almost periodicity of the map $t \mapsto U(t, \omega)$ for all $\omega \in \Omega$ as the Brownian motion $W(t)$ is almost surely not almost periodic. We can however ask about almost periodicity of the distribution of $U(t)$ or, equivalently, of the associated measures $P_U(t, \mathrm{d}x)$ in the topology of convergence in measure.

The approach we take in Chapter 5 differs from previous approaches to stochastic almost periodicity such as [31] which relied heavily on Hilbert spaces techniques. We instead follow the approach of [64], where almost periodic solutions of non-autonomous deterministic problems are studied. The extension of this method to stochastic problems on Banach spaces rests heavily on the R-boundedness estimates for semigroups and evolutions families of [54].

In following the approach of [64] we inherit the assumption that the evolution families used be almost periodic jointly in both variables, i.e. that for any $\tau_n \uparrow \infty$, the sequence

$$\big(P(t + \tau_n, s + \tau_n)\big)_{n \geq 1} \tag{1.10}$$

has a subsequence which converges uniformly for bounded $(t-s)$. As an aside at the end of Chapter 5 we develop the theory of such families by finding new checkable conditions for this property on families of operators $A(t)$ which satisfy conditions given by Acquistapace and Terreni [1] (such families are a non-autonomous analogue of analytic semigroups and appear widely in applications). There is a body of literature which studies almost periodicity of evolution families (for example [7], [9], [10] and [63]), typically assuming almost periodicity of the map $t \mapsto R(\lambda, A(t))$ and showing almost periodicity of associated Green's functions by way of Yosida approximations and then of solutions to associated Cauchy problems. Our approach differs in that we assume almost periodicity of the map $t \mapsto A(t)$ directly (in a suitable topology) and show from these that the evolution family generated is almost periodic in the sense of (1.10). The result is proved by going back to the original construction of Acquistapace and Terreni and showing almost periodicity painstakingly, term by term.

Another long term property of stochastic evolution equations is the existence of invariant measures. A problem

$$\mathrm{d}U(t) = \Big[Au(t) + F(t, U(t))\Big]\,\mathrm{d}t + G(t, U(t))\,\mathrm{d}W(t), \quad t \geq 0, \tag{1.11}$$

has an invariant measure if there exists a probability measure μ on E such that if $U(0)$ has distribution μ then so does $U(t)$ for all $t \geq 0$. Invariant measures for stochastic problems are analogous to stationary solutions of deterministic problems and are strongly related to other important notions such as ergodicity. Again, invariant measures have been much studied on Hilbert spaces [28], [34], [40] and [42]. Most papers on the subject follow one of two methods considered in [28], either assuming a dissipativity condition on the Yosida approximations of the generator A or requiring compactness of the semigroup $S(t)$. Brzeźniak in [22] shows the existence of invariant measure on Banach spaces of martingale type 2 using a variant of the former method, whereas works based on the latter such as [14] and [41] do so by first showing tightness of certain families of measures using a compactness argument. By a theorem of Krylov and Bogoliubov (Theorem 6.3), under the right conditions, tightness of these families is equivalent to the existence of the desired invariant measure.

In [14], van Gaans *et al.* consider invariant measures for delay equations on Hilbert spaces. They show that invariant measures for (1.11) exist on a Hilbert space even if the semigroup $S(t)$ generated by A is only eventually compact. The result then follows using the fact (Theorem 2.19) that if $S(t)$ is an immediately compact semigroup and

$\mathcal{S}(t)$ is the associated (finite) delay semigroup, then $\mathcal{S}(t)$ is compact for $t > 1$. We show in Chapter 6 that this method extends to delay problems (with finite delay) on Banach spaces of type 2.

1.2 Overview

This thesis is structured as follows. In Chapter 2 we provide a basic introduction to evolution equations and specifically to abstract Cauchy problems on Banach spaces and their solutions in terms of C_0-semigroups. A special class, the analytic semigroups are recalled in Section 2.2.3 and some properties are listed that will be important later. Results in the sequel are largely concerned with inhomogeneous problems, and so perturbations by non-linear Lipschitz continuous terms are outlined in Section 2.2.4 together with the variation of parameters formula which becomes the main tool for their handling. In several cases in later chapters we also give versions of our results for problems driven by a non-autonomous linear part, and so in Section 2.3 we describe strongly continuous evolution families which play a role analogous to semigroups in such problems. Acquistapace and Terreni gave conditions under which operators $A(t)$ will generate such a family and these are included. Finally in Section 2.4 we introduce delayed differential equations and show the relation between delay equations and Cauchy problems. The theory of such problems is rather different depending on whether the delay in question is finite or infinite, the former admitting the notion of a delay semigroup while the latter requires more attention be placed on the history or phase space of initial data. Both are briefly outlined.

In Chapter 3 we include preliminary material related to probability theory and the stochastic integral of van Neerven, Veraar and Weis ([43] and [68]). We start with basic definitions of measurability and random variables on Banach spaces in Section 3.1. Many important geometric properties of Banach spaces turn out to be characterised by whether or not certain inequalities hold on expectations of sums of random variables. Properties such as the type and co-type of spaces and the UMD property are given in this way and play a key role in stochastic integration. Next we define the space $\gamma(H, E)$ of bounded operators from a Hilbert space to E, which will turn out to characterise stochastic integrability, before moving on in Section 3.4 to giving the main theorems on the existence and manipulation of the stochastic integral of van Neerven, Veraar and Weis. Finally we recall the notions of R- and

γ-boundedness for families of operators. These will turn out to play a key role as multipliers for stochastically integrable processes.

Our new results start in Chapter 4 where we follow the methods of [69] and [86] to give existence and uniqueness results for (infinitely) delayed evolution equations on UMD Banach spaces. In Section 4.2 we consider the non-autonomous problem on spaces of type 2 before moving on to the main result of the chapter, well-posedness of infinite delay equations on general UMD spaces. Section 4.4 is devoted to an application of this result to a stochastic heat equation in a material with memory.

Chapter 5 is concerned with existence of almost periodic solutions to stochastic differential equations on Banach spaces. We start by recalling some background results on almost periodic functions and R-boundedness for semigroups before presenting conditions under which such solutions will exist in both the autonomous (Section 5.3) and non-autonomous (Section 5.4) cases.

At this point we take an aside from stochastic differential equations briefly to consider almost periodicity of evolution families in the sense of (1.10) which satisfy the conditions of Acquistapace and Terreni. The existence of such families is assumed in the previous section, so in Section 5.5 we give new checkable conditions on the operators $A(t)$, under which the evolution family generated will be almost periodic in the required sense. Finally we provide an example of such an operator family and show that almost periodic solutions (in distribution) exist for the stochastic differential equation in question.

Our final chapter looks at the existence of invariant measures for stochastic evolution equations with finite delay. We present the main tool, the Krylov-Bogoliubov Theorem (Theorem 6.3) before working our way to invariant measures for delay equations by first showing existence for equations driven by compact semigroups, then eventually compact semigroups and finally passing from this case to delay problems using the fact that delay semigroups are eventually compact. Once again a stochastic heat equation in a material with memory provides an example in Section 6.6.

1.3 Notation

We adopt the following conventions:

- All Banach spaces are real unless otherwise specified.
- Banach spaces are usually E, F, etc. Hilbert spaces are H, K, etc.
- $(\Omega, \mathscr{F}, \mathbb{P})$ is assumed throughout to be a complete probability space equipped with a filtration $\{\mathscr{F}_t\}_{t\geq 0}$.
- Unless otherwise stated, vector and operator valued integrals should be read as Bochner integrals.
- The space $L^0(\Omega; E)$ denotes the space of strongly measurable E-valued random variables on Ω, equipped with the topology of convergence in probability.
- The space of bounded linear operators from E to F is denoted $\mathcal{L}(E, F)$, or $\mathcal{L}(E)$ in the case that $E = F$.
- We use $a \lesssim_p b$ for the situation where there exists a constant C (depending on p) such that $a \leq Cb$.
- We use $a \simeq_p b$ for $a \lesssim_p b \lesssim_p a$.
- Lower case $u, v, \dots, z$ are usually vectors in a Banach space E, whereas upper case $U, V, \dots, Z$ represent E-valued random variables.
- $\langle x, x^* \rangle$ represents the action of functional $x^* \in E^*$ on vector $x \in E$.
- Inner products in Hilbert spaces are denoted $[h, g]_H$ for $h, g \in H$.
- $\mathbb{B}_E$ is the Borel σ-algebra on E.

Chapter 2

Deterministic Preliminaries

2.1 Introduction

This chapter contains background material on Cauchy problems, C_0-semigroups and delay equations. Background on probability theory and stochastic integration will follow in Chapter 3. We introduce the classes of problems known as evolution equations and Cauchy problems and recall the notion of operator semigroups as solutions of homogeneous time-autonomous problems and an evolution families solving the non-autonomous case. Special classes of each, analytic semigroups and Acquistapace-Terreni evolution families respectively will be important in our results of chapters 4 and 5, so are defined here. We also consider briefly perturbed problems of each type and show how perturbations can be used to solve problems depending on delayed inputs, or equivalently on the history of the system. Finally we recall some theory of equations with infinite delay and show the importance of the notion of history spaces.

2.2 Evolution equations in Banach spaces

Suppose we are interested in studying a system, perhaps arising in physics, biology or finance that evolves along a continuous 'time' parameter $t \in \mathbb{R}$ in some state space E. Such systems are often described in terms of the change of the *state* $u(t)$ of the system with respect to time,

$$\frac{\mathrm{d}u}{\mathrm{d}t}(t) = f(t, u(t)). \tag{2.1}$$

We call this most general class of problems *evolution equations*. By considering a higher order system as a first order system on a product space (see [39, Section

VI.3]), it is no loss of generality to assume for the remainder of this thesis that all systems will be first order.

If we have an evolution equation such as (2.1) and an initial value $u(s) = x_s$ at time s, then the *solution* $u(t)$ is of the form

$$u(t) = F(t, s, x_s), \quad t \geq s$$

for some map $F : \mathbb{R}^2 \times E \to E$. We say such a system *determines a flow* F if it satisfies the functional equation

$$\begin{cases} F(t,s,x) &= F\big(t, r, F(r,s,x)\big), \quad s \leq r \leq t \\ F(s,s,x) &= x \end{cases} \tag{2.2}$$

for all x in the state space. This condition is essentially the idea that running the system from time s to time r, then taking the result $F(r, s, x_s)$ and running the system with this initial condition from time r until t is the same as letting the system run all the way from s to t. The system is said to be *autonomous* if $F(t, s, x)$ depends only on the value of $t - s$, and in this case we write $F(t, x) = F(t, 0, x)$.

2.2.1 Cauchy problems

Evolution equations such as (2.1) are usually studied on infinite dimensional Banach spaces. A typical example is the n-dimensional heat equation.

Example 2.1. Let $\mathcal{O}$ be an open bounded domain of $\mathbb{R}^n$ with Lipschitz boundary and consider the following heat equation on $\mathcal{O}$

$$\begin{cases} \dfrac{\partial u}{\partial t}(t) &= \Delta u(t), \quad t \geq 0 \\ u(t)\big|_{\partial\mathcal{O}} &= g \in L^p(\partial\mathcal{O}) \\ u(0) &= f \in L^p(\mathcal{O}). \end{cases} \tag{2.3}$$

Here it is natural to consider the problem not as a partial differential equation on a subset of $\mathbb{R}^n$, but as an ordinary differential equation on the Banach space $L^p(\mathcal{O})$. The Laplacian Δ is a closed and densely defined operator on a subspace $D(\Delta) \subseteq \{u \in W^{2,p}(\mathcal{O}) : u\big|_{\partial\mathcal{O}} = g\}$ satisfying the boundary condition.

We call an autonomous, homogeneous, linear evolution equation such as (2.3) an *abstract Cauchy problem*. In general, if E is a Banach space and A is a (typically) unbounded linear operator with domain $D(A) \subseteq E$, then the equation

$$\begin{cases} \frac{\mathrm{d}u}{\mathrm{d}t}(t) &= Au(t), \quad t \geq 0 \\ u(0) &= x \in E \end{cases} \tag{2.4}$$

is said to be an abstract Cauchy Problem. For a comprehensive introduction to the theory of Cauchy problems we refer to [39]. Since (2.4) is linear and autonomous, by (2.2) we would expect solutions to take the form of families $(S(t))_{t\geq 0} \subset \mathcal{L}(E)$ satisfying the functional equation

$$\begin{cases} S(t+s)x_0 &= S(t)S(s)x_0, \quad s,t \geq 0 \\ S(0)x_0 &= x_0 \end{cases} \tag{2.5}$$

for all $x_0 \in E$. A family $(S(t))_{t\geq 0}$ satisfying (2.5) is called a *semigroup*.

2.2.2 C_0-semigroups

Definition 2.2. Let E and F be Banach spaces. A function $S : \mathbb{R}_+ \to \mathcal{L}(E, F)$ is said to be *strongly continuous* (or equivalently, continuous in the *strong operator topology*) if the function $S(\cdot)x \to F$ is continuous for every $x \in E$. A strongly continuous semigroup $S : \mathbb{R}_+ \to \mathcal{L}(E)$ is called a *C_0-semigroup*.

Definition 2.3. Given a C_0-semigroup $(S(t))_{t\geq 0}$ on a Banach space E, we define an (unbounded) linear operator, the *generator* A of $(S(t))_{t\geq 0}$ by

$$Ax = \lim_{t\downarrow 0} \frac{1}{t}\big(S(t)x - x\big) \tag{2.6}$$

where this exists. The *domain* $D(A)$ of A is the subspace of all $x \in E$ for which the limit exists.

Since the operator A is (in general) unbounded, there is a question of how to interpret the problem (2.4) for $x \notin D(A)$. It is necessary to introduce the concepts of *classical* and *mild* solutions.

Definition 2.4. A function $u : \mathbb{R}_+ \to E$ is said to be a *classical solution* of (2.4) if $u(t) \in D(A)$ for all $t \in \mathbb{R}_+$ and $u(t)$ satisfies (2.4). On the other hand, $u(t)$ is said to be a *mild solution* of (2.4) if for all $t > 0$

$$\int_0^t u(s)\,\mathrm{d}s \in D(A) \text{ and } A\int_0^t u(s)\,\mathrm{d}s = u(t) - x.$$

We can now give the following theorem which connects C_0-semigroups to abstract Cauchy problems. See [5, Theorem 3.1.12].

Theorem 2.5. *Let $(A, D(A))$ be a closed operator on a Banach space E. The following are equivalent:*

(1) For all $x \in E$ there exists a unique mild solution to (2.4)*.*

(2) The operator A is the generator of a C_0-semigroup $(S(t))_{t \geq 0}$.

(3) The resolvent set $\rho(A) \neq \emptyset$ and for all $x \in D(A)$ there exists a unique classical solution of (2.4)*.*

In this case the (classical or mild) solution is given by $u(t) = S(t)x$ for $t \geq 0$.

2.2.3 Analytic C_0-semigroups

In Chapter 4 we will need restrictions on the regularity (smoothness) of the semigroup, specifically estimates on the map $t \mapsto (-A)^\eta S(t)x$ for $x \in E$. We will require that the semigroup generated by $(A, D(A))$ is *analytic.*

If A is a closed operator with $(0, \infty) \subset \rho(A)$ and $\|R(\lambda, A)\| \leq \frac{M}{1+\lambda}$ for all $\lambda \in (0, \infty)$ then we can define fractional powers $(-A)^\eta$ of A as follows. For $\eta > 0$ define $(-A)^{-\eta}$ by

$$(-A)^{-\eta} := \frac{1}{2\pi i} \int_\Gamma (-\lambda)^{-\eta} R(\lambda, A) \, \mathrm{d}\lambda$$

where Γ is a piecewise smooth path in some sector $\Sigma \subset \rho(A)$ running from $\infty e^{-i\delta}$ to $\infty e^{i\delta}$ for some $\delta > 0$. $(-A)^{-\eta}$ is injective for every $\eta > 0$, so we define $(-A)^\eta$ as the inverse of $(-A)^{-\eta}$ with domain $D((-A)^\eta) = \mathrm{range}((-A)^{-\eta})$. The space $D((-A)^\eta)$ becomes a Banach space (written E_η) when endowed with the graph norm $\|x\|_{E_\eta} := \|(-A)^\eta x\|_E + \|x\|_E$. For further details on fractional powers and proof of these results we refer to [39, Section II.5c].

Remark 2.6. Analytic semigroups are an analogue of holomorphic functions and thus live naturally on *complex* Banach spaces. We *complexify* our real space E in order to make this definition. Take $E^{\mathbb{C}} = E \oplus iE$ with scalar multiplication

$$(a + ib)(x + iy) = (ax - by) + i(bx + ay), \quad a, b \in \mathbb{R}, \quad x, y \in E.$$

We will define analytic semigroups on a complex Banach space, then if this space is the complexification $E^{\mathbb{C}}$ of a real space E we will take an analytic semigroup on the real space to be the restriction of the semigroup to the (real) subspace E of $E^{\mathbb{C}}$.

Definition 2.7. A closed linear operator $(A, D(A))$, densely defined on a *complex* Banach space E is said to be *sectorial (of angle θ)* if there exists $\theta \in (0, \frac{\pi}{2}]$ such that the open sector

$$\Sigma_{\frac{\pi}{2}+\theta} := \left\{\lambda \in \mathbb{C} : |\arg(\lambda)| < \frac{\pi}{2} + \theta\right\} \setminus \{0\}$$

is contained in the resolvent $\rho(A)$ and if for each $\varepsilon \in (0, \theta)$ there exists a constant $M_\varepsilon \geq 1$ such that

$$\|R(\lambda, A)\| \leq \frac{M_\varepsilon}{|\lambda|} \quad \text{for all } \lambda \in \overline{\Sigma}_{\frac{\pi}{2}+\theta-\varepsilon} \setminus \{0\}.$$

Definition 2.8. A family of operators $(S(z))_{z\in\Sigma_\theta\cup\{0\}}$ is called an *analytic semigroup* (of angle $\theta \in (0, \frac{\pi}{2}]$) if

(1) $S(0) = I$ and $S(w + z) = S(w)S(z)$ for all $w, z \in \Sigma_\theta$;

(2) the map $z \mapsto S(z)$ is analytic in Σ_θ;

(3) $\|S(z)x - x\| \to 0$ for all $x \in E$ as $z \to 0$ for $z \in \Sigma_{\theta'}$ with $0 < \theta' < \theta$.

If in addition $\|S(z)\|$ is bounded on $\Sigma_{\theta'}$ for every $0 < \theta' < \theta$ we say $S(z)$ is a bounded analytic semigroup. We define the generator A of $S(t)$ in a similar fashion to Definition 2.3, taking the limit instead as $z \to 0$ over $z \in \Sigma_\theta$.

Theorem 2.9 (See Theorem II.4.6 of [39])**.** *For an operator $(A, D(A))$ on a complex Banach space E the following are equivalent.*

(1) A is sectorial of angle $\theta \in (0, \frac{\pi}{2})$;

(2) A generates a bounded analytic semigroup $(S(z))_{z\in\Sigma_\theta\cup\{0\}}$ of angle θ;

(3) A generates a bounded C_0-semigroup $(S(t))_{t\geq 0}$ on E such that $S(t)E \subset D(A)$ for all $t > 0$ and for $\eta = 1$, and hence all $\eta \in [0, 1]$, there exists a constant $C_\eta \geq 0$ such that

$$\left\|S(t)\right\|_{\mathcal{L}(E,E_\eta)} \leq \frac{C_\eta}{t^\eta}, \quad t > 0. \tag{2.7}$$

In this case the semigroup $S(z)$ can be given for $z \in \Sigma_\theta$ by Cauchy's formula

$$S(z) = \frac{1}{2\pi i} \int_\Gamma e^{\mu z} R(\mu, A)\, \mathrm{d}\mu, \tag{2.8}$$

where Γ is a piecewise smooth path in some sector $\Sigma_{\frac{pi}{2}+\delta}$ running from $\infty e^{-i\delta}$ to $\infty e^{i\delta}$ for some $\delta \in (\frac{\pi}{2}, \frac{\pi}{2} + \theta)$.

We call (2.7) the *analytic semigroup property*. It is a major tool that we will use regularly to estimate various integrals.

Remark 2.10. If A generates a C_0-semigroup $(S(t))_{t\geq 0}$ on a real space E then (the natural extension of) A also generates a C_0-semigroup $(S^{\mathbb{C}}(t))_{t\geq 0}$ on the complexification $E^{\mathbb{C}}$. In this case E is $S^{\mathbb{C}}(t)$-invariant as a (real) subspace of $E^{\mathbb{C}}$ with $S^{\mathbb{C}}(t)|_E = S(t)$ and hence property (3) of Theorem 2.9 is inherited by $S(t)$. This justifies effectively taking (3) to be the definition of an analytic C_0-semigroup on a real Banach space.

2.2.4 Perturbations and inhomogeneous problems

We now consider the inhomogeneous Cauchy problem. Inhomogeneous problems will occur throughout this thesis when we consider either delay terms or noise terms in our problem. Let $f \in L^1(0,T;E)$ and consider the problem

$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t}(t) & = Au(t) + f(t), \quad t \in [0,T] \\ \\ u(0) & = x \in E. \end{cases} \tag{2.9}$$

Following Definition 2.4, a function $u \in C([0,T];E)$ is a mild solution of (2.9) if

$$\int_0^t u(s)\,\mathrm{d}s \in D(A), \quad t > 0$$

and

$$u(t) = x + A\int_0^t u(s)\,\mathrm{d}s + \int_0^t f(s)\,\mathrm{d}s$$

for all $t \in [0,T]$.

Proposition 2.11. *(See [5, Proposition 3.1.16]). Let A be the generator of a C_0-semigroup $(S(t))$ on E, then for every $f \in L^1(0,T;E)$ the problem* (2.9) *has a unique mild solution given by the so called* variation of parameters formula

$$u(t) = S(t)x + \int_0^t S(t-s)f(s)\,\mathrm{d}s, \quad t \in [0,T].$$

We are also interested in the case where the inhomogeneity depends on $u(t)$. In this case we will need the *Banach fixed point Theorem.*

Theorem 2.12 (Banach, [8])**.** *Let (M,d) be a non-empty complete metric space and let $\Phi : M \to M$ be a contraction mapping, i.e. there exists $c \in (0,1)$ such that*

$$d\big(\Phi(x),\Phi(y)\big) \leq c \cdot d(x,y)$$

for all $x, y \in M$. Then there exists a unique fixed point $x \in M$ such that

$$\Phi(x) = x.$$

Theorem 2.13. *Suppose A is the generator of a C_0-semigroup $(S(t))$ such that $\|S(t)\| \leq Me^{-\lambda t}$ for some $\lambda > 0$ and $M \geq 0$ and $F : [0,T] \times E \to E$ is continuous, Lipschitz continuous on E and of linear growth, that is, there exists $L_F, C_F \geq 0$ such that for all $x, y \in E$*

$$\big\|F(t,x) - F(t,y)\big\| \leq L_F\|x-y\| \quad \text{and} \quad \big\|F(t,x)\big\| \leq C_F\big(1+\|x\|\big).$$

Then the problem

$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t}(t) &= Au(t) + F(t,u(t)), \quad t \in [0,T] \\ \\ u(0) &= x \in E. \end{cases} \tag{2.10}$$

has a unique mild solution given by

$$u(t) = S(t)x + \int_0^t S(t-s)F(s,u(s))\,\mathrm{d}s, \quad t \in [0,T].$$

Sketch proof. Let $T_0 \leq T$ and define the fixed point operator Φ on $C([0,T_0];E)$ by

$$\Phi(u)(t) = S(t)x_0 + \int_0^t S(t-s)F(s,u(s))\,\mathrm{d}s.$$

We claim that Φ is a strict contraction for small enough T_0. Let $u, v \in C([0,T_0];E)$ and fix $t \in [0,T_0]$, then

$$\begin{aligned} \big\|\Phi(u)(t) - \Phi(v)(t)\big\| &\leq \int_0^t \big\|S(t-s)\big[F(s,u(s)) - F(s,v(s))\big]\big\|\,\mathrm{d}s \\ &\leq M\int_0^t e^{-\lambda(t-s)}\big\|F(s,u(s)) - F(s,v(s))\big\|\,\mathrm{d}s \\ &\leq \frac{ML_F}{\lambda}(1-e^{-\lambda T_0}) \sup_{t\in[0,T_0]} \|u(t)-v(t)\|. \end{aligned}$$

Now $\frac{ML_F}{\lambda}(1-e^{-\lambda T_0}) \to 0$ as $T_0 \downarrow 0$, hence for small enough $T_0 > 0$, Φ is a strict contraction. Therefore by Theorem 2.12 there exists a unique $u(t) \in C([0,T_0];E)$ such that

$$u(t) = S(t)x + \int_0^t S(t-s)F(s,u(s))\,\mathrm{d}s \text{ for all } t \in [0,T_0].$$

The existence of a fixed point on the whole interval $[0,T]$ follows inductively by considering the above problem on the intervals $[T_0, 2T_0], \ldots, [mT_0, T]$. The $u(t)$ thus constructed is the unique mild solution following the method in the proof of Proposition 2.11 [5, Proposition 3.1.16]. □

2.3 Non-autonomous Cauchy problems

We also consider non-autonomous problems, that is, problems where the unbounded operator A is replaced with a time indexed family $(A(t))$.

$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t}(t) &= A(t)u(t), \quad t \in [0,T], \\ \\ u(s) &= x \in E. \end{cases} \tag{2.11}$$

In a manner similar to Theorem 2.5, we would like to describe solutions to (2.11) in terms of a family of operators $(P(t,s))_{0\le s\le t\le T}$ such that $u(t) = P(t,s)x$ for each $x \in E$ and $0 \le s \le t \le T$.

Definition 2.14. A family $\big(P(t,s)\big)_{0\le s\le t\le T} \subset \mathcal{L}(E)$ is called a strongly continuous *evolution family* if

(1) $P(s,s) = I$ for all $0 \le s \le T$.

(2) $P(t,s) = P(t,r)P(r,s)$ for all $0 \le s \le r \le t \le T$.

(3) The mapping $(t,s) \mapsto P(t,s) : \{(t,s) \in \mathbb{R}^2 : 0 \le s \le t \le T\} \to \mathcal{L}(E)$ is strongly continuous.

We say that $\big(P(t,s)\big)_{0\le s\le t\le T}$ solves (2.11) if there exists a family $(Y_s)_{0\le s\le T}$ of dense subspaces of E such that $P(t,s)Y_s \subseteq Y_t \subseteq D(A(t))$ for every $0 \le s \le t \le T$ and the function $u(t) = P(t,s)x$ is in $C^1([s,\infty);E)$ with $u(t) \in D(A(t))$ and $u'(t) = A(t)u(t)$ for all $x \in Y_s$ and $s \le t \le T$. In this case we say that $A(t)$ generates the evolution family $\big(P(t,s)\big)_{0\le s\le t\le T}$.

We will consider a particular class of evolution families that are analogous to analytic semigroups from Section 2.2.3. As above, when on a real vector space E the following is understood to refer to the complexification of the objects concerned.

Definition 2.15 (see [3])**.** Let $(A(t), D(A(t)))$ be a family of closed, densely defined linear operators on Banach space E. $A(t)$ satisfies the Acquistapace-Terreni (AT) condition if

(AT1) There exist constants $N \ge 0$ and $\theta \in (0, \frac{\pi}{2})$ such that for all $0 \le t \le T$, $\Sigma_{\theta+\frac{\pi}{2}} \cup \{0\} \subset \rho(A(t))$ and

$$\big\|R(\lambda, A(t))\big\| \le \frac{N}{1+|\lambda|}.$$

for all $\lambda \in \Sigma_{\theta+\frac{\pi}{2}} \cup \{0\}$, as defined in Definition 2.7.

(AT2) There exist constants $L \geq 0$ and $\mu, \nu \in (0,1]$ with $\mu + \nu > 1$ such that for all $\lambda \in \Sigma_{\theta+\frac{\pi}{2}}$ and $s, t \in [0,T]$

$$\left\| A(t)R(\lambda, A(t))\left(A(t)^{-1} - A(s)^{-1}\right)\right\| \leq L|t-s|^{\mu}(|\lambda|+1)^{-\nu}.$$

The first condition (AT1) is essentially that the family $A(t)$ are uniformly sectorial, hence by Theorem 2.9, $A(t)$ generates a bounded analytic semigroup $\left(e^{sA(t)}\right)_{s\geq 0}$ for each $t \in [0,T]$. We define the real interpolation space $E^t_\eta = E^t_{\eta,2} := (E, D(A(t)))_{\eta,2}$ as follows.

Definition 2.16 (See [82] Section 1.6.5)**.** Let A be the generator of a C_0 semigroup $S(t)$. For $0 < \eta < 1$ and $1 \leq p < \infty$ we define the real interpolation space $(E, D(A(t)))_{\eta,p}$ as the set of $x \in E$ such that

$$\|x\|_{E^t_{\eta,p}} = \left(\int_0^\infty \left[s^{-\eta}\left\|S(s)x - x\right\|\right]^p \frac{\mathrm{d}s}{s}\right)^{\frac{1}{p}} \tag{2.12}$$

is finite. Similarly, we define the space $(E, D(A))_{\eta,\infty}$ as those $x \in E$ such that

$$\|x\|_{E^t_{\eta,\infty}} = \sup_{s\in(0,\infty)} s^{-\eta}\left\|S(s)x - x\right\| < \infty. \tag{2.13}$$

The following result, [86, Theorem 2.1] (or originally [3, Theorem 2.3]), is the analogue of Theorem 2.9 for non-autonomous problems satisfying (AT).

Theorem 2.17. *If condition (AT) of Definition 2.15 holds, then there exists a unique strongly continuous evolution family* $\left(P(t,s)\right)_{0\leq s\leq t\leq T}$ *which solves* (2.11)*, that is* $u(t) := P(t,s)x \in D(A(t))$ *and* $u'(t) = A(t)u(t)$ *for all* $0 \leq s \leq t \leq T$*. Moreover,* $t \mapsto P(t,s)$ *is (norm) continuous for* $0 \leq s < t \leq T$ *and there exists* $C \geq 0$ *such that*

$$\left\|P(t,s)x\right\|_{E^t_\eta} \leq C(t-s)^{\theta-\eta}\|x\|_{E^s_\theta} \tag{2.14}$$

for all $0 \leq s < t \leq T$ *and* $0 \leq \theta \leq \eta \leq 1$.

2.4 Delay equations and history spaces

2.4.1 Delay equations with finite delay

There are essentially two distinct theories of delay equations for finite delay, depending on whether the history space is taken to be of the type $C([-1,0];E)$ or $L^p(-1,0;E)$. In this thesis we will consider only the latter to avoid confusion. The continuous theory is outlined in [39, Section VI.6] whereas the L^p theory is extensively studied in [11]. We assume without loss of generality (by rescaling as necessary) that all finite delays are of duration 1.

Let E be a Banach space, then for a function $u : [-1, \infty) \to E$ we define the *history function* of u at time $t \geq 0$ by the map

$$u_t : [-1, 0] \to E; \quad u_t(s) = u(s + t).$$

Intuitively, we associate with each positive time t the function u_t which represents the action of u in the previous unit time interval. In the L^p theory of delay equations we take $L^p(-1, 0; E)$ for some $p \geq 1$ as the space in which history functions live, so for $u(t) \in L^p(-1, T; E)$ we have a map $t \mapsto u_t : [0, T] \to L^p(-1, 0; E)$.

Consider the *delayed Cauchy problem*

$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t}(t) &= Au(t) + F(t, u(t), u_t), \quad t \in [0, T], \\ u(0) &= x \in E, \\ u_0 &= f \in L^p(-1, 0; E). \end{cases} \tag{2.15}$$

Here A is the generator of a C_0-semigroup and F is continuous and Lipschitz on $L^p(-1, 0; E)$. In order to study (2.15), we can transform the delay equation (2.15) into the form of an undelayed Cauchy problem on the product space $\mathcal{E} := E \times L^p(-1, 0; E)$.

Define the delay operator $\mathcal{A}$ on $\mathcal{E}$ by

$$\mathcal{A} = \begin{pmatrix} A & 0 \\ 0 & \frac{\mathrm{d}}{\mathrm{d}s} \end{pmatrix} \text{ with domain } D(\mathcal{A}) := \left\{ \begin{pmatrix} x \\ f \end{pmatrix} \in D(A) \times W^{1,p}(-1, 0; E) : f(0) = x \right\} \tag{2.16}$$

and introduce a new inhomogeneous term $\mathcal{F} : [0, T] \times \mathcal{E} \to \mathcal{E}$ by

$$\mathcal{F}\Big(t, \begin{pmatrix} x \\ f \end{pmatrix}\Big) = \begin{pmatrix} F(t, x, f) \\ 0 \end{pmatrix}.$$

This allows us to rewrite (2.15) as a new inhomogeneous problem of the form (2.10) on the space $\mathcal{E}$

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} &= \mathcal{A} \begin{pmatrix} u(t) \\ v(t) \end{pmatrix} + \mathcal{F}\Big(t, \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}\Big), \quad t \in [0, T] \\ \begin{pmatrix} u(0) \\ v(0) \end{pmatrix} &= \begin{pmatrix} x \\ f \end{pmatrix} \in \mathcal{E}. \end{cases} \tag{2.17}$$

$\mathcal{A}$ generates the *delay semigroup* $\mathcal{S}(t)$ on $\mathcal{E}$ given by

$$\mathcal{S}(t) := \begin{pmatrix} S(t) & 0 \\ S_t & L_t \end{pmatrix} \tag{2.18}$$

where

$$(S_t x)(s) := \begin{cases} 0 & t+s<0 \\ S(t+s)x & t+s \geq 0 \end{cases} \quad \text{and} \quad (L_t f)(s) := \begin{cases} f(t+s) & t+s<0 \\ 0 & t+s \geq 0 \end{cases}.$$

Then (by e.g. [11, Corollary 3.7] or [14, Theorem 3.2]) $\mathcal{S}(t)$ is strongly continuous and hence (2.17) has a unique mild solution $(u(t), v(t))^T \in L^p(0,T;\mathcal{E})$ by Theorem 2.13. Solutions to (2.15) and (2.17) are related by the following theorem.

Theorem 2.18 (See Theorem 3.12 of [11])**.**

(1) If $u(t)$ is a solution of (2.15)*, then $t \mapsto \begin{pmatrix} u(t) \\ u_t \end{pmatrix}$ is a solution of* (2.17)*.*

(2) If $t \mapsto \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}$ is a solution of (2.17) *then $v(t) = u_t$ for all $t \in [0,T]$ and $u(t)$ is a solution of* (2.15)*.*

In the sequel (in particular in Chapter 6) we will be interested in *compact semigroups.* A semigroup $S(t)$ on E such that $S(t)$ is a compact operator for all $t > 0$ is said to be *immediately compact*, or just compact. If $S(t)$ is compact for all $t > t_0$ for some $t_0 > 0$ then $S(t)$ is said to be *eventually compact.* We have the following useful result about compactness for delay semigroups.

Theorem 2.19 (Theorem A.4 of [14], also appears in [11])**.** *Suppose A generates an immediately compact semigroup $S(t)$ on E, then the delay semigroup $\mathcal{S}(t)$ on $\mathcal{E}$ given by* (2.18) *is compact for all $t > 1$.*

2.4.2 Infinite history spaces

In systems with infinite delay, that is where evolution is determined by the entire history of the system from $-\infty$ to t, well-posedness becomes a more subtle question. In the case of finite delay, we have delay semigroups as in (2.17) and the problem is amenable to general semigroup theory, however for infinite delay problems the semigroup (2.18) is not strongly continuous in general. In fact the map $t \mapsto u_t$ is not continuous in general for u_t taken in an arbitrary space of functions $(-\infty, 0] \to E$ and continuous $u : (-\infty, T] \mapsto E$.

The principal difference between the two cases is that an infinite delay problem never "forgets" the initial data which may not be as regular as the output of the system, whereas after $t = 1$ a finite delay problem is determined entirely by the prior output of the system and hence inherits certain regularity properties from $S(t)$. The theory of well-posedness for infinitely delayed evolution equations with data from a general Banach space of functions $(-\infty, 0] \to E$ was first put on a firm axiomatic footing by Hale and Kato in [50]. They introduced a series of axioms on a space $\mathscr{B}$ of functions $\phi : (-\infty, 0] \to E$ under which the system

$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t}(t) &= Au(t) + F(t, u_t), \quad t \in [0, T], \\ \\ u_0(s) &= \phi(s), \qquad\qquad\quad s \in (-\infty, 0], \end{cases} \tag{2.19}$$

is well posed for any $\phi \in \mathscr{B}$. Here u_t is defined on all of $(-\infty, 0]$ by

$$u_t : (-\infty, 0] \to E; \quad u_t(s) = u(t+s)$$

and we note that since $u(t)$ will be continuous for $t \geq 0$, $u_t(0)$ is well defined and hence it is no loss of generality to write $F(t, u_t)$ rather than $F(t, u(t), u_t)$. We say that a function $u : (-\infty, T] \to E$ is a mild solution of (2.19) if $u(s) = \phi(s)$ for $s \leq 0$ and for all $t \in [0, T]$ we have $\int_0^t u(s)\,\mathrm{d}s \in D(A)$ and

$$u(t) = A\int_0^t u(s)\,\mathrm{d}s + \int_0^t F(s, u_s)\,\mathrm{d}s.$$

The following notation is now standard, and dates to the book [53] of Hino.

Definition 2.20. Let $\mathscr{B}$ be a space of functions $(-\infty, \sigma] \to E$ equipped with a semi-norm $\|\cdot\|_{\mathscr{B}}$. We will say that $\mathscr{B}$ is a *history space* if $\mathscr{B}$ satisfies

(A0) There exists a constant $H \geq 0$ and functions $K, M : \mathbb{R}_+ \to \mathbb{R}_+$, where K is continuous and M locally bounded, such that if $\phi : (-\infty, \sigma + T) \to E$ for $T > 0$ is continuous on $[\sigma, \sigma + T)$ and $\phi_\sigma \in \mathscr{B}$, then for every $t \in [\sigma, \sigma + T)$ the following hold:

$$\text{(i)} \qquad \phi_t \in \mathscr{B}; \tag{2.20}$$

$$\text{(ii)} \qquad \|\phi(t)\| \leq H\|\phi_t\|_{\mathscr{B}}; \tag{2.21}$$

$$\text{(iii)} \qquad \|\phi_t\|_{\mathscr{B}} \leq K(t-\sigma)\sup_{\sigma \leq s \leq t}\|\phi(s)\| + M(t-\sigma)\|\phi_\sigma\|_{\mathscr{B}}. \tag{2.22}$$

(A1) For the function ϕ in (A0), the map $t \mapsto \phi_t$ is a $\mathscr{B}$-valued continuous function on $[\sigma, \sigma + T)$.

We identify functions indistinguishable under $\| \cdot \|_{\mathscr{B}}$, and then require that

(B) The normed space $\mathscr{B}$ formed by this identification is complete.

It follows from (2.20) and (2.22) that if $\phi : (-\infty, 0] \to E$ is continuous with compact support then $\phi \in \mathscr{B}$ and

$$\|\phi\|_{\mathscr{B}} \leq K(a) \sup_{s \in [-a,0]} \|\phi(s)\| \tag{2.23}$$

for all $a > 0$ such that $\mathrm{supp}(\phi) \subseteq [-a, 0]$.

If $\phi : (-\infty, T] \to E$ is continuous on $[0, T]$ with $\phi_t \in \mathscr{B}$ for $t \in [0, T]$ then the pair $(\phi(t), \phi_t)$ is well defined. In this case and where there is no ambiguity we identify ϕ_t with $(\phi(t), \phi_t)$.

Theorem 2.21. *Suppose that A generates a C_0-semigroup $(S(t))$ with $\|S(t)\| \leq Me^{-\lambda t}$ for some $\lambda > 0$ and $M \geq 0$ and that $F : [0, T] \times \mathscr{B} \to E$ is continuous and uniformly Lipschitz in the second variable. Then (2.19) has a unique mild solution $u : (-\infty, T] \to E$ for every initial history $\phi \in \mathscr{B}$.*

Proof. The proof of Theorem 2.21 follows the same pattern as that of Theorem 2.13. We introduce a fixed point operator L on the complete metric space $\{u \in C([0, T_0]; E) : u(0) = \phi(0)\}$ and show that for small enough $T_0 \leq T$, L becomes a strict contraction. For $u : [0, T] \to E$ with $u(0) = \phi(0)$ we write $\widetilde{u}$ for the extension of u by ϕ to $(-\infty, T_0] \to E$, so

$$\widetilde{u}(t) = \begin{cases} \phi(t), & -\infty < t \leq 0, \\ u(t), & 0 < t \leq T_0. \end{cases}$$

By (2.23) and linearity $\widetilde{u}_t \in \mathscr{B}$ for all $t \in [0, T_0]$. Define

$$L(u)(t) = S(t)\phi(0) + \int_0^t S(t-s)F(s, u_s)\, \mathrm{d}s.$$

Let $u, v \in C([0, T_0]; E)$ with $u(0) = v(0) = \phi(0)$ and fix $t \in [0, T_0]$, then

$$\begin{aligned} \big\|L(u)(t) - L(v)(t)\big\| &\leq \int_0^t \big\|S(t-s)\big[F(s, \widetilde{u}_s) - F(s, \widetilde{v}_s)\big]\big\|\, \mathrm{d}s \\ &\leq M \int_0^t e^{-\lambda(t-s)} \big\|F(s, \widetilde{u}_s) - F(s, \widetilde{v}_s)\big\|\, \mathrm{d}s \\ &\leq \frac{ML_F}{\lambda}(1 - e^{-\lambda T_0}) \sup_{t \in [0,T_0]} \|\widetilde{u}_t - \widetilde{v}_t\|_{\mathscr{B}}. \end{aligned} \tag{2.24}$$

Now since $\widetilde{u}(s) = \widetilde{v}(s) = \phi(s)$ for $s \leq 0$, $\widetilde{u}_t - \widetilde{v}_t$ has compact support in $[-t, 0]$ and (2.23) gives

$$\begin{aligned} \|\widetilde{u}_t - \widetilde{v}_t\|_{\mathscr{B}} &\leq K(t) \sup_{s\in[-t,0]} \|\widetilde{u}_t(s) - \widetilde{v}_t(s)\| \\ &\leq \sup_{r\in[0,t]} K(r)\|u(r) - v(r)\| \end{aligned}$$

which together with (2.24) shows that L is a strict contraction for small enough T_0. The result now follows as in the proof of Theorem 2.13. □

Example 2.22. Let E be a Banach space and $\mathscr{B}$ be one of the following

(1) The space $BUC((-\infty, 0]; E)$ of all bounded uniformly continuous functions from $(-\infty, 0]$ to E with the sup norm

$$\|\phi\|_\infty := \sup_{s\leq 0} \|\phi(s)\|;$$

(2) The space $C_g^0\big((-\infty, 0], E\big)$ of all continuous functions $\phi : (-\infty, 0] \to E$ such that $\frac{\|\phi(s)\|}{g(s)} \to 0$ as $s \to -\infty$, with the norm

$$\|\phi\|_g := \sup_{s\leq 0} \frac{\|\phi(s)\|}{g(s)},$$

where $g : (-\infty, 0] \to [1, \infty)$ is such that

(a) $g(s) \to \infty$ as $s \to -\infty$;

(b) $\sup_{t\geq 0} \sup_{s\leq -t} \frac{g(t+s)}{g(s)} < \infty$.

(3) The space $E \times L_g^p\big((-\infty, 0], E\big)$ of pairs (x, ϕ) of equivalence classes of measurable functions $\phi : (-\infty, 0] \to E$ together with a vector $x \in E$ for which the norm

$$\|(x, \phi)\|_g := \left(\int_{-\infty}^0 g(s)\|\phi(s)\|^p \,\mathrm{d}s\right)^{\frac{1}{p}} + \|x\|$$

is finite. The weight g is a function $g : (-\infty, 0] \to (0, \infty)$ such that

(a) g is locally Lebesgue integrable, i.e. integrable on every interval $[-r, 0]$;

(b) there exists a non-negative and locally bounded function G on $(-\infty, 0]$ such that $g(s + t) \leq G(s)g(t)$ for all $s \leq 0$ and almost all $t \leq 0$.

Then $\mathscr{B}$ satisfies axioms (A0), (A1) and (B) of Definition 2.20, and Theorem 2.21 applies.

Proof. See [53, Theorems 1.3.6 and 1.3.8]. □

Remark 2.23. Let $\mathscr{B}$ be a history space of functions $\phi : (-\infty, 0] \to E$ and A be the generator of a uniformly bounded analytic semigroup on E with $0 \in \rho(A)$. For $\eta \in [0,1]$ we define the space $\mathscr{B}_\eta$ by

$$\mathscr{B}_\eta = \{(-A)^{-\eta} \otimes \phi : \phi \in \mathscr{B}\}$$

where $((-A)^{-\eta} \otimes \phi)(t) = (-A)^{-\eta}\phi(t)$. We endow $\mathscr{B}_\eta$ with the norm

$$\|\psi\|_{\mathscr{B}_\eta} := \|(-A)^{\eta} \otimes \psi\|_{\mathscr{B}}$$

then $\mathcal{B}_\eta$ is easily seen to also be a history space [52].

Chapter 3

Stochastic Preliminaries

In this chapter we recall the background material necessary to construct a theory of stochastic integration on a wide class of Banach spaces. We start in 3.1 with various notions of measurability for vector and operator valued functions and discuss Gaussian random variables and Brownian motion on Banach spaces. We discuss how the behaviour of certain families of random variables can be used to characterise geometric properties of Banach spaces in 3.2; it will turn out that this characterisation allows the formulation of many of the theorems that appear later. We round up the chapter by constructing the stochastic integral in 3.4 which powers the remainder of the work.

3.1 Measurability and random variables

We introduce here the basic probabilistic tools needed later, as well as definitions of the various notions of measurability.

3.1.1 Measurability

Let (S, Σ, μ) be a measure space and consider functions ϕ from S into a metric space (E, d). We recall several related notions of measurability of ϕ.

Definition 3.1. Let $\phi : S \to E$

(1) ϕ is said to be *Borel measurable* if the pre-image of any Borel set is measurable, that is, for any Borel set $B \subseteq E$ we have $\phi^{-1}(B) \in \Sigma$.

(2) ϕ is *μ-simple* if it is a finite linear combination of functions of the form $\mathbb{1}_B \otimes x$ for $B \in \Sigma$ with $\mu(B) < \infty$ and $x \in E$.

(3) ϕ is *μ-strongly measurable* if ϕ is the almost everywhere pointwise limit of a sequence of simple functions $\phi_n : S \to E$.

Where the measure μ is clear and there is no ambiguity we refer simply to *simple* and *strongly measurable* functions.

The following result gives a criterion for a function $\phi : S \to E$ to be strongly measurable.

Proposition 3.2 (Pettis' measurability theorem). *Let E be a metric space and suppose $\Gamma \subseteq C(E; \mathbb{R})$ is a set of continuous real valued functions which separates points of E, that is, for distinct $x, y \in E$, there exists $f \in \Gamma$ such that $f(x) \neq f(y)$. For a function $\phi : S \to E$ the following are equivalent:*

(1) ϕ is strongly measurable;

(2) ϕ is Borel measurable and takes values in a separable closed subspace of E almost surely (ϕ is essentially separably valued*);*

(3) ϕ is essentially separably valued and $f \circ \phi$ is Borel measurable for every $f \in \Gamma$;

(4) there exists a sequence of countably valued Borel measurable functions (ϕ_n) such that $d(\phi(s), \phi_n(s)) \to 0$ uniformly for $s \in S$ as $n \to \infty$.

Proof. [84, Propositions I.1.9 and I.1.10]. □

By a simple diagonalisation argument it follows that the pointwise limit of a sequence of strongly measurable functions is strongly measurable. We now assume that E is a Banach space and make the following further definitions.

Definition 3.3.

(1) A function $\phi : S \to E$ is said to be *weakly measurable* if $\langle \phi, x^* \rangle : S \to \mathbb{R}$ is measurable for all $x^* \in E^*$;

(2) if F is another Banach space then an operator valued function $\Phi : S \to \mathcal{L}(E, F)$ is said to be *E-strongly measurable* if the function $\Phi(\cdot)x : S \to F$ is strongly measurable for each $x \in E$.

By applying the Pettis measurability theorem with $\Gamma = E^*$ we see that $\phi : S \to E$ is strongly measurable if and only if ϕ is weakly measurable and essentially separably valued. If $\Phi : S \to \mathcal{L}(E, F)$ is E-strongly measurable and $f : S \to E$ is strongly measurable then $\Phi f : S \to F$ is also strongly measurable.

We will often want to consider the space of strongly measurable functions, which we define as follows.

Definition 3.4. Let (S, Σ, μ) be a finite measure space and let ϕ be a μ-equivalence class of functions from S to E. We say ϕ is *strongly measurable* if there is a version of ϕ which is strongly measurable. If μ is a finite measure we define $L^0(S, E)$ as the space of all strongly measurable equivalence classes of functions $\phi : S \to E$. $L^0(S, E)$ is a complete metric space when endowed with the metric

$$d(\phi, \psi) = \int_S d(\phi(s), \psi(s)) \wedge \mathbb{1} \, \mathrm{d}\mu(s)$$

and this topology coincides with the topology of *convergence in measure* (or convergence in probability if $\mu(S) = 1$).

Definition 3.5. If $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space, then a strongly measurable function $X : \Omega \to E$ is called a *random variable*. The Borel probability measure P_X on E defined by $P_X(B) = \mathbb{P}\{X \in B\}$ for Borel subsets $B \in \mathbb{B}_E$ is called the *distribution* of X.

Definition 3.6. For an E-valued random variable X on $(\Omega, \mathscr{F}, \mathbb{P})$, if $X : \Omega \to E$ is Bochner integrable then the *expectation* $\mathbb{E}(X)$ of X is defined by

$$\mathbb{E}(X) := \int_\Omega X(\omega) \, \mathrm{d}\mathbb{P}(\omega).$$

We will usually follow the convention that random variables are denoted by upper case Roman letters, as in Definition 3.5. The next result on E-valued random variables that will be needed later. For more details and proofs we recommend the survey [65].

Proposition 3.7. *Any E-valued random variables is* tight*, that is, if $X : \Omega \to E$ is a random variable, then for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset E$ such that $\mathbb{P}\{X \in K_\varepsilon\} > 1 - \varepsilon$.*

We say a family $X(t)$, $t \in I$ of random variables is *uniformly tight* if for any $\varepsilon > 0$ there exists a compact set $K \subset E$ such that $\mathbb{P}\big\{X(t) \in K\big\} > 1 - \varepsilon$ for all $t \in I$.

3.1.2 Gaussian random variables

Gaussian random variables are the foundation of the Brownian motion model of space-time white noise, so it is necessary to understand how they generalise to Banach spaces. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and E be a Banach space.

Definition 3.8. A real valued random variable η is said to be *Gaussian* (or *normal*) of mean m and variance σ^2 if for all Borel sets $B \subset \mathbb{R}$

$$\mathbb{P}\{\eta \in B\} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_B e^{-|t-m|/2\sigma^2} \, \mathrm{d}t$$

Definition 3.9. A strongly measurable function $\gamma : \Omega \to E$ is a *Gaussian* random variable on E if for all functionals $x^* \in E^*$, $\langle \gamma, x^* \rangle$ is a Gaussian random variable on $\mathbb{R}$. We define the *mean* $m \in E$ and *covariance operator* $Q \in \mathcal{L}(E^*, E)$ of γ by

$$\begin{aligned} m &:= \mathbb{E}\gamma \\ Qx^* &:= \mathbb{E}\left[\langle \gamma - m, x^* \rangle (\gamma - m)\right]. \end{aligned} \tag{3.1}$$

Both the mean and covariance operator exist, and in fact all p^{th} moments of γ are finite. This follows from Fernique's Theorem, [59, Corollary 2.6.1] which says that $\mathbb{E}e^{a\|\gamma\|} < \infty$ for $a > 0$.

Definition 3.10. An operator $A \in \mathcal{L}(E^*, E)$ is said to be *positive* if $\langle Ax^*, x^* \rangle \geq 0$ for all $x^* \in E^*$ and *symmetric* if $\langle Ax^*, y^* \rangle = \langle x^*, Ay^* \rangle$ for all $x^*, y^* \in E^*$.

It is clear that Q is both positive and symmetric. We say that γ has distribution $\mathcal{N}(m, Q)$. Conversely an operator $Q \in \mathcal{L}(E^*, E)$ is called a *covariance operator* if there exists an E-valued Gaussian random variable such that Q satisfies (3.1).

Let (γ_n) be an independent identically distributed sequence of Gaussian random variables on $\mathbb{R}$ (a *Gaussian sequence*) and (x_n) a sequence in E. The series

$$\sum_{n \in \mathbb{N}} \gamma_n x_n$$

converges almost surely if and only if it converges in probability and if and only if it converges in $L^p(\Omega, E)$ for some (and hence all) $p \in [1, \infty)$. The limit in each case is a Gaussian random variable [59, Theorems 2.1.1 and 2.2.1] on E.

For a *centred* (mean 0) Gaussian random variable γ, a version of the Kahane-Khinchine inequalities (see 3.15) holds [59, Corollary 3.4.1], that is, for all $p, q \in [1, \infty)$ we have

$$\left(\mathbb{E}\|\gamma\|^p\right)^{\frac{1}{p}} \simeq_{p,q} \left(\mathbb{E}\|\gamma\|^q\right)^{\frac{1}{q}}. \tag{3.2}$$

A consequence is that if (γ_n) is a sequence of E-valued Gaussian random variables and $\gamma : \Omega \to E$ is a random variable with $\lim_{n\to\infty} \gamma_n = \gamma$ in probability, then γ is Gaussian and $\lim_{n\to\infty} \gamma_n = \gamma$ in $L^p(\Omega, E)$ for all $p \in [1, \infty)$.

Finally in this section we recall the notion of a *Gaussian process.*

Definition 3.11. If $\mathcal{I}$ is an indexing set and ξ is a mapping from $\mathcal{I} \times \Omega$ to E, then ξ is a *Gaussian process* if for any $n \in \mathbb{N}$ and $i_1, i_2, \ldots, i_n \in \mathcal{I}$, the n-tuple

$$\big(\xi(i_1), \ldots, \xi(i_n)\big)$$

is an E^n-valued Gaussian random variable.

Example 3.12 (Brownian motion [38])**.** A real Brownian motion on $(\Omega, \mathscr{F}, \mathbb{P})$ is a family $B(t)$ of square integrable random variables such that for $0 \le s \le t$

1. $B(0) = 0$;
2. $B(t) - B(s)$ is a Gaussian random variable of mean 0 and variance $(t - s)$;
3. $B(t)$ has independent increments;
4. the map $t \mapsto B(t)$ is almost surely continuous on $t \in \mathbb{R}_+$.

A real Brownian motion $B(t)$ is a Gaussian process on $\mathbb{R}$.

3.2 Geometric properties of Banach spaces

There are many related but distinct notions characterising Banach spaces into various *types* (usually falling in the range $[1, 2]$) and *co-types* (usually in $[2, \infty]$) which find much use in functional analysis. Among these are concepts such as *martingale type* or *Fourier type*, but for our purposes we are interested in the properties usually referred to simply as *type* and *co-type.* In older literature this notion appears as both *Rademacher type* and *Gauss type* spaces, but in fact, as we shall note, these two properties are equivalent. For a discussion of the development of these concepts see [75, Section 6.1].

3.2.1 Type and co-type of Banach spaces

For this section, let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, E be a Banach space and $1 \leq p \leq 2 \leq q \leq \infty$.

Definition 3.13. A sequence (r_n) of independent identically distributed random variables on $(\Omega, \mathscr{F}, \mathbb{P})$ with

$$\mathbb{P}(r_1 = 1) = \mathbb{P}(r_1 = -1) = \frac{1}{2}$$

is called a *Rademacher sequence.*

By way of motivation we recall the Khintchine inequality on the real line and the Khintchine-Kahane inequality on a general Banach space.

Proposition 3.14 (Khintchine inequality, 1923)**.** *For $s \in (0, \infty)$, $n \in \mathbb{N}$, $\xi_1, \ldots, \xi_n \in \mathbb{R}$ and a Rademacher sequence (r_k) the following holds*

$$\left(\mathbb{E}\Big|\sum_{k=1}^{n} \xi_k r_k\Big|^s\right)^{\frac{1}{s}} \simeq_s \Big(\sum_{k=1}^{n} |\xi_k|^2\Big)^{\frac{1}{2}}. \tag{3.3}$$

Proposition 3.15 (Khintchine-Kahane inequality, 1964)**.** *For $s \in (0, \infty)$, $n \in \mathbb{N}$, $x_1, \ldots, x_n \in E$ and a Rademacher sequence (r_k) the following holds*

$$\left(\mathbb{E}\Big\|\sum_{k=1}^{n} x_k r_k\Big\|^s\right)^{\frac{1}{s}} \simeq_s \left(\mathbb{E}\Big\|\sum_{k=1}^{n} x_k r_k\Big\|^2\right)^{\frac{1}{2}}. \tag{3.4}$$

This second result is weaker than the first, as the expectation still occurs on the right hand side. It is desirable to have a version of (3.3) which holds when the real numbers ξ_k are replaced by vectors x_k in a Banach space E. Unfortunately one or both directions of (3.3) may fail if vectors are used instead. This motivates the following definitions of spaces in which one or other inequality of (3.3) holds.

Definition 3.16. E is said to have *type* $p \in [1, 2]$ (or *Rademacher type* p) if for some (and hence all) $s \in (0, \infty)$ there exists a constant $C_{s,p} \geq 1$ such that

$$\left(\mathbb{E}\Big\|\sum_{k=1}^{n} x_k r_k\Big\|^s\right)^{\frac{1}{s}} \leq C_{s,p}\Big(\sum_{k=1}^{n} \|x_k\|^p\Big)^{\frac{1}{p}} \tag{3.5}$$

for any $n \in \mathbb{N}$ and finite subset $\{x_1, \ldots, x_n\} \subset E$. Similarly, E is said to be of *co-type* $q \in [2, \infty)$ (or *Rademacher co-type* q) if there exists $c_{s,q} \geq 1$ such that

$$\Big(\sum_{k=1}^{n} \|x_k\|^q\Big)^{\frac{1}{q}} \leq c_{s,q} \left(\mathbb{E}\Big\|\sum_{k=1}^{n} x_k r_k\Big\|^s\right)^{\frac{1}{s}} \tag{3.6}$$

with the standard modification if $q = \infty$.

It follows from (3.4) that these inequalities do not depend on the choice of s, so we will usually take $s = 2$ in the above for simplicity.

Proposition 3.17 (Properties of type and co-type, see [75] Section 6.1.7)**.**

(1) If $1 \leq p \leq p' \leq 2$ then all type p' spaces are also type p.

(2) Similarly, for $2 \leq q' \leq q \leq \infty$, all co-type q' spaces are also co-type q.

(3) The space $L^r(\mathbb{R})$ is of type $\min\{r, 2\}$ and co-type $\max\{r, 2\}$.

(4) For a non-trivial measure space (S, Σ, μ) and $r \in [p, \infty)$, E has type p if and only if $L^r(S; E)$ has type p.

(5) Similarly for $r \in [1, q]$, E has co-type q if and only if $L^r(S; E)$ has co-type q.

(6) Closed subspaces of a Banach space preserve both type and co-type, while taking quotients preserves only type in general.

(7) Every Banach space has type 1 and co-type ∞ with constants $C_{s,1} = c_{s,\infty} = 1$.

(8) A Banach space is isomorphic to a Hilbert space if and only if it has type 2 and co-type 2 (Kwapień's theorem).

Following (7), E is said to have *non-trivial type* if $p > 1$ and *finite co-type* if $q < \infty$.

If the Rademacher sequences (r_n) in Definition 3.16 are replaced with sequences (γ_n) of Gaussian random variables then Proposition 3.17 still holds. These spaces were historically referred to as *Gauss type* and *Gauss co-type*, however it was shown by Hoffmann-Jørgensen and Maurey and Pisier that in fact Rademacher type = Gauss type and Rademacher co-type = Gauss co-type. See [75, Section 6.1.7] for details.

3.2.2 UMD spaces

We now discuss a class of spaces introduced by Burkholder [24] known as *UMD spaces*, or spaces with the *unconditional property for martingale differences*.

Definition 3.18. A sequence (D_n) of E-valued random variables on a filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_n\}, \mathbb{P})$ is said to be *adapted* (to the filtration $\{\mathscr{F}_n\}$) if D_n is $\mathscr{F}_n$-measurable for each $n \in \mathbb{N}$. D_n is called a *martingale difference sequence* if each D_n is Bochner integrable and

$$\mathbb{E}(D_{n+1}|\mathscr{F}_n) = 0 \text{ for all } n,$$

where $\mathbb{E}(\cdot|\mathscr{F}_n)$ is conditional expectation with respect to the sub-σ-algebra $\mathscr{F}_n$ as defined in [90, Chapter 9]. (D_n) is called a *Paley-Walsh martingale difference sequence* if it is a martingale difference sequence for a filtration $\{\mathcal{G}_n\}$ defined by $\mathcal{G}_0 = \{\emptyset, \Omega\}$ and $\mathcal{G}_n = \sigma(r_k : k = 1, \ldots, n)$ where (r_n) is a Rademacher sequence.

The name 'martingale difference sequence' arises from the case where $D_n = M_n - M_{n-1}$ for a martingale (M_n).

Definition 3.19. A Banach space E is said to be a *UMD space* if for some (and hence every) $p \in (1, \infty)$, there exists a constant $\beta_{E,p} \geq 1$ such that for every finite, p-integrable, E-valued, martingale difference sequence $(D_n)_{n=1}^N$ on a filtered space $(\Omega, \mathscr{F}, \{\mathscr{F}_n\}, \mathbb{P})$ and every sequence $(\varepsilon_n)_{n=1}^N \in \{-1, 1\}^N$ we have

$$\left(\mathbb{E}\Big\| \sum_{n=1}^N \varepsilon_n D_n \Big\|^p\right)^{\frac{1}{p}} \leq \beta_{E,p} \left(\mathbb{E}\Big\| \sum_{n=1}^N D_n \Big\|^p\right)^{\frac{1}{p}}. \tag{3.7}$$

It was shown by Maurey that it is enough to have (3.7) for just Paley-Walsh sequences (D_n). The independence of the UMD property from the choice of $p \in (1, \infty)$ in Definition 3.19 is shown in [24]

Proposition 3.20 (Properties of UMD spaces, see [75] Section 6.1.10)**.**

(1) If E is a UMD space then so is F for every closed subspace or quotient space F of E.

(2) E is a UMD space if and only if E^* is UMD.

(3) If E is UMD then E is super-reflexive and of non-trivial type.

(4) E is UMD if and only if $L^p(S, E)$ is UMD for every σ-finite measure space (S, Σ, μ) and $p \in (1, \infty)$.

On the other hand, Pisier constructed a super-reflexive space which is not UMD, so the converse of (3) is false.

3.3 The space of γ-radonifying operators

We now begin recalling the construction of the stochastic integral, as developed by van Neerven *et al.* in [70] and then [68]. The key tool will be characterising stochastically integrable functions in terms of γ-*radonifying* operators.

Definition 3.21. A linear operator $R : H \to E$ from a separable Hilbert space H into a Banach space E is said to be γ-*radonifying* if for some (and hence every) orthonormal basis (h_n) of H and every standard Gaussian sequence (γ_n) on $(\Omega, \mathscr{F}, \mathbb{P})$, the sum

$$\sum_{n=1}^{\infty} \gamma_n R h_n$$

converges in $L^2(\Omega; E)$. The space of all γ-radonifying operators is written $\gamma(H, E)$, and is a Banach space with respect to the norm

$$\|R\|_{\gamma(H,E)} := \left(\mathbb{E} \Big\| \sum_{n=1}^{\infty} \gamma_n R h_n \Big\|^2 \right)^{\frac{1}{2}}.$$

This norm is shown to be independent of the choice of orthonormal basis (h_n) and the sequence (γ_n) in the survey article [65, Section 3] of van Neerven.

Example 3.22.

(1) If E is also a Hilbert space, then $\gamma(H, E) = \mathcal{L}_2(H, E)$, the space of Hilbert-Schmidt operators from H to E ([32] or [86, Section 2.3]).

(2) Suppose $R, S \in \mathcal{L}(H, E)$ and

$$\|R^* x^*\|_H \le \|S^* x^*\|_H \quad \text{for all } x^* \in E^*.$$

If $S \in \gamma(H, E)$ then we also have $R \in \gamma(H, E)$ and $\|R\|_{\gamma(H,E)} \le \|S\|_{\gamma(H,E)}$ (see [65, Theorem 9.3]).

(3) If $E = L^p(S)$ for some σ-finite measure space S, then $\gamma(H, E)$ is characterised by the following Lemma.

Lemma 3.23 (Lemma 2.1 of [69]). *Let (S, Σ, μ) be a σ-finite measure space and let $1 \le p < \infty$. For an operator $R \in \mathcal{L}(H, L^p(S))$ the following are equivalent:*

(1) $R \in \gamma(H, L^p(S))$.

(2) For some (and hence all) orthonormal bases (h_n) of H, the function

$$\Big(\sum_{n=1}^{\infty} |Rh_n|^2\Big)^{\frac{1}{2}}$$

belongs to $L^p(S)$.

(3) There exists a function $g \in L^p(S)$ such that for all $h \in H$ we have $|(Rh)(s)| \leq g(s)\|h\|_H$ μ-almost everywhere.

(4) There exists a function $k \in L^p(S;H)$ such that $(Rh)(s) = [k(s),h]_H$ μ-almost everywhere.

In this case we have

$$\|R\|_{\gamma(H,L^p(S))} \simeq_p \|k\|_{L^p(S,H)} \simeq_p \left\| \Big(\sum_{n=1}^{\infty} |Rh_n|^2\Big)^{\frac{1}{2}} \right\|_{L^p(S)} \leq \|g\|_{L^p(S)}.$$

We will usually be concerned with the case $\gamma(\mathcal{H},E)$, where $\mathcal{H}$ is a Hilbert space of the form $L^2(I;H)$ for another Hilbert space H and an interval $I \subseteq \mathbb{R}$. In order to estimate γ-norms it will be useful to have a Fubini type theorem to interchange γ and L^p type norms. The next result, known as the γ-Fubini isomorphism, is a tool that will be used repeatedly.

Proposition 3.24 (Proposition 2.6 of [68])**.** *Let (S,Σ,μ) be a σ-finite measure space and let $p \in [1,\infty)$, then the map $F_\gamma : L^p\big(S;\gamma(H,E)\big) \to \mathcal{L}\big(H, L^p(S;E)\big)$ defined by*

$$\big(F_\gamma(\phi)h\big)(s) := \phi(s)h, \quad s \in S, h \in H$$

defines an isomorphism from $L^p\big(S;\gamma(H,E)\big)$ onto $\gamma\big(H,L^p(S;E)\big)$.

Note that in the case $E = \mathbb{R}$, then $\gamma(H,E)$ is just H, so we have

$$L^p(S;H) \simeq \gamma(H, L^p(S)). \tag{3.8}$$

In the delayed context, we will also need the following immediate corollary.

Corollary 3.25. *If (S,Σ,μ) is a σ-finite measure space and $p \in [1,\infty)$, then the map $G_\gamma : L^p\big(S;\gamma(H,E)\big) \times \gamma(H,E) \to \mathcal{L}\big(H, L^p(S;E) \times E\big)$ defined by*

$$G_\gamma\big(\phi,R\big)h := \big(\phi(\cdot)h, Rh\big)$$

is an isomorphism from $L^p\big(S;\gamma(H,E)\big) \times \gamma(H,E)$ onto $\gamma\big(H, L^p(S;E) \times E\big)$.

3.4 Construction of the stochastic integral

Now we define the stochastic integral, following the method of van Neerven *et al.* in [68].

Definition 3.26. Let H be a separable Hilbert space. An *H-cylindrical Brownian motion* on the filtered space $\big(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P}\big)$ is a family $\big(W(t)\big)_{t\geq 0}$ of bounded linear operators from H to $L^2(\Omega)$ which satisfies

(1) $Wh := \big(W(t)h\big)_{t\geq 0}$ is a real valued scaled Brownian motion on $\big(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P}\big)$ for each $h \in H$;

(2) $\mathbb{E}\big(W(s)g \cdot W(t)h\big) = (s \wedge t)[g, h]_H$ for all $s, t \geq 0$ and $g, h \in H$;

(3) the family $\big\{W(t)h : t \geq 0, h \in H\big\}$ is jointly Gaussian.

$W(t)h$ is a standard Brownian motion if and only if $\|h\|_H = 1$. Note that condition (3) is often not included in definitions in the literature, but is frequently assumed implicitly in arguments that follow. Since (3) does not obviously follow from (1) and (2) we include it explicitly.

Equivalently, and rather more elegantly, an H-cylindrical Brownian motion can be defined as an *isonormal process* $W : L^2(\mathbb{R}_+; H) \to L^2(\Omega)$. See the survery [65] for details.

Example 3.27.

(1) If H is a separable Hilbert space with orthonormal basis (h_n) and $\big(W^{(n)}\big)_{n=1}^{\infty}$ is a sequence of independent real Brownian motions, then

$$W(t)h := \sum_{n=1}^{\infty} W^{(n)}(t)[h, h_n]_H, \quad h \in H$$

defines an H-cylindrical Brownian motion.

(2) In the case where $H = L^2(D)$ where D is an open subset of $\mathbb{R}^d$ then an $L^2(D)$-cylindrical Brownian motion is the model for 'space-time white noise' on $\mathbb{R}_+ \times D$. This explains why H-cylindrical Brownian motion appears naturally in stochastic partial differential equations.

For $h \in H$ and $x \in E$ we write $h \otimes x \in \mathcal{L}(H, E)$ for the operator defined by

$$(h \otimes x)g = [h, g]_H x, \quad g \in H. \tag{3.9}$$

For $0 \leq a < b < \infty$ and an $\mathscr{F}_a$-measurable subset $A \subseteq \Omega$, the stochastic integral of the indicator process $\mathbb{1}_{(a,b \wedge t] \times A} \otimes (h \otimes x) : [0, \infty) \times \Omega \to \mathcal{L}(H, E)\big)$ is defined as

$$\int_0^t \mathbb{1}_{(a,b] \times A} \otimes (h \otimes x) \, \mathrm{d}W := \mathbb{1}_A \big(W(b \wedge t)h - W(a \wedge t)h\big)x. \tag{3.10}$$

Note that this integral is of Itô type and is an $\mathscr{F}_t$-adapted martingale by the definition of W (Definition 3.26). By linearity, this definition extends to adapted step processes $\Phi : [0, \infty) \times \Omega \to \mathcal{L}(H, E)$ which take values in the finite rank operators.

In order to extend this definition to a more general class of processes, recall the definition of an H-strongly measurable process (Definition 3.3). We first define the stochastic integral for deterministic functions $\Phi : [0, \infty) \to \mathcal{L}(H, E)$ and then extend this to random processes $\Phi : [0, T] \times \Omega \to \mathcal{L}(H, E)$.

Definition 3.28. An H-strongly measurable function $\phi : [0, \infty) \to \mathcal{L}(H, E)$ is said to be *stochastically integrable* with respect to $W(t)$ if there exists a sequence $\phi_n : [0, \infty) \to \mathcal{L}(H, E)$ of adapted step functions which take values in the finite rank operators and a pathwise continuous process $\zeta : [0, \infty) \times \Omega \to E$ such that the following two conditions hold:

(1) $\phi_n h$ converges to ϕh in measure on $[0, \infty) \times \Omega$ as $n \to \infty$ for all $h \in H$;

(2) $t \mapsto \displaystyle\int_0^t \phi_n \, \mathrm{d}W$ converges to ζ in measure on $C_b([0, \infty); E)$ as $n \to \infty$.

In this case ζ is well defined independently of the choice of ϕ_n as an element of $L^0\big(\Omega; C_b([0, \infty); E)\big)$ and is called the stochastic integral of ϕ with respect to W. We write

$$\zeta(t) = \int_0^t \phi(s) \, \mathrm{d}W(s).$$

Proposition 3.29 (Theorem 4.2 of [70]). *Let H be a separable Hilbert space. For an H-strongly measurable function $\phi : [0, \infty) \to \mathcal{L}(H, E)$ the following are equivalent:*

(1) ϕ is stochastically integrable;

(2) for all $x^* \in E^*$ the function $\phi^* x^*(t) := \phi(t)^* x^*$ is in $L^2(0,\infty;H)$ and there exists a pathwise continuous process $\xi : [0,\infty) \times \Omega \to E$ such that

$$\langle \xi(\cdot), x^* \rangle = \int_0^{\cdot} \phi(s)^* x^* \, \mathrm{d}W(s) \text{ almost surely, for all } x^* \in E^*;$$

(3) for all $x^* \in E^*$ the function $\phi^* x^*$ is in $L^2(0,\infty;H)$ and there exists an operator $R \in \gamma\big(L^2(0,\infty;H), E\big)$ such that for all $f \in L^2(0,\infty;H)$ and $x^* \in E^*$ we have

$$\langle Rf, x^* \rangle = \int_0^\infty [f(s), \phi(t)^* x^*]_H \, \mathrm{d}t.$$

We say that ϕ is the representation *of R.*

In this case ξ is uniquely determined and

$$\xi(t) = \int_0^t \phi(s) \, \mathrm{d}W(s) \text{ almost surely for all } t \in [0,\infty).$$

The process ξ is a martingale adapted to the filtration $\{\mathscr{F}_t\}$ with continuous p^{th} moments for all $p \in [1,\infty)$, and for all $p \in [1,\infty)$ we have

$$\Big(\mathbb{E} \sup_{t\in[0,\infty)} \Big\| \int_0^t \phi \, \mathrm{d}W \Big\|^p \Big)^{\frac{1}{p}} \simeq_p \Big(\mathbb{E} \sup_{t\in[0,\infty)} \Big\| \int_0^t \phi \, \mathrm{d}W \Big\|^2 \Big)^{\frac{1}{2}} = \|R\|_{\gamma(L^2(0,\infty;H),E)}. \quad (3.11)$$

Moreover, for each $p \in [1,\infty)$, the convergence of $\int_0^t \phi_n \, \mathrm{d}W$ to ζ in Definition 3.28 (2) is in $L^p(\Omega;E)$.

For functions $\phi_n : \mathbb{R} \to \mathcal{L}(H,E)$ we have a stochastic dominated convergence theorem. See [70, Theorem 6.2].

Theorem 3.30 (Dominated Convergence Theorem)**.** *Let $\phi_n : (0,\infty) \to \mathcal{L}(E,H)$ be a sequence of stochastically integrable functions and suppose there exists a function $\phi : (0,\infty) \to \mathcal{L}(H,E)$ with $\phi^* x^* \in L^2(0,\infty;H)$ for all $x^* \in E^*$ such that*

$$\lim_{n\to\infty} \int_0^\infty \big\|\phi_n^*(t)x^* - \phi^*(t)x^*\big\|_H^2 \, \mathrm{d}t = 0 \text{ for all } x^* \in E^*.$$

Assume also that there exists a stochastically integrable function $\psi : (0,T) \to \mathcal{L}(H,E)$ such that for all $x^* \in E^*$ and all $n \in \mathbb{N}$ we have

$$\int_0^\infty \big\|\phi_n^*(t)x^*\big\|_H^2 \, \mathrm{d}t \leq \int_0^\infty \big\|\psi^*(t)x^*\big\|_H^2 \, \mathrm{d}t.$$

Then Φ is stochastically integrable and

$$\lim_{n\to\infty} \mathbb{E} \Big\| \int_0^\infty \big(\phi_n(t) - \phi(t)\big) \, \mathrm{d}W(t) \Big\|^p = 0 \text{ for all } p \in [1,\infty).$$

Corollary 3.31 (Uniform Dominated Convergence). *Suppose $\phi_n(t,s) : \mathbb{R}\times(0,\infty) \to \mathcal{L}(H,E)$ is a sequence of stochastically integrable functions with*

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}} \int_0^\infty \left\|\phi_n^*(t,s)x^*\right\|_H^2 \mathrm{d}s = 0 \text{ for all } x^* \in E^*$$

and there exists a stochastically integrable function $\psi : (0,\infty) \to \mathcal{L}(H,E)$ with

$$\sup_{t\in\mathbb{R}} \int_0^\infty \left\|\phi_n^*(t,s)x^*\right\|_H^2 \mathrm{d}t \le \int_0^\infty \left\|\psi^*(s)x^*\right\|_H^2 \mathrm{d}s.$$

Then Theorem 3.30 holds uniformly for $t \in \mathbb{R}$, so

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}} \mathbb{E}\left\| \int_0^\infty \phi_n(t,s)\,\mathrm{d}W(s)\right\|^p = 0 \text{ for all } p \in [1,\infty).$$

Proof. This essentially follows from the proof of [70, Theorem 6.2] with uniform convergence in $t \in \mathbb{R}$ coming from Proposition 3.29 (3) and Example 3.22 (2). □

To extend the stochastic integral to adapted processes $\Phi : [0,T]\times\Omega \to \mathcal{L}(H,E)$ it is necessary to assume something about the Banach space E. The following result of Garling [44] characterises UMD spaces (as defined above in 3.2.2) as those in which the *decoupling inequality* holds.

Let $\big(\widetilde{W}(t)\big)$ be another H-cylindrical Brownian motion on a new filtered space $\big(\widetilde{\Omega}, \widetilde{\mathscr{F}}, (\widetilde{\mathscr{F}}_t), \widetilde{\mathbb{P}}\big)$, independent of $\big(\mathscr{F}_t\big)$. We define the *decoupled integral* of an elementary adapted process $\mathbb{1}_{(a,b]\times A} \otimes (h\otimes x) : [0,T]\times\Omega \to E$ for $A \in \mathscr{F}_a$ with respect to $\widetilde{W}$ as follows

$$\int_0^t \mathbb{1}_{(a,b]\times A} \otimes (h\otimes x)\,\mathrm{d}\widetilde{W} = \mathbb{1}_A\big(\widetilde{W}(b\wedge t)h - \widetilde{W}(a\wedge t)h\big)x.$$

This stochastic integral forms a random variable in $L^p\big(\Omega; L^p(\widetilde{\Omega}; E)\big)$. We write $\widetilde{\mathbb{E}}$ for expectation with respect to $\widetilde{\mathbb{P}}$.

Lemma 3.32. *Let H be a non-zero separable real Hilbert space and fix $p \in (1,\infty)$. The following are equivalent:*

(1) E is a UMD space;

(2) there exists a constant $C_{p,E} \ge 0$ such that for every elementary adapted (to $\{\mathscr{F}_t\}$) process $\Phi : [0,T]\times\Omega \to \mathcal{L}(H,E)$

$$C_{p,E}^{-1}\mathbb{E}\Big(\widetilde{\mathbb{E}} \sup_{t\in[0,T]} \Big\| \int_0^t \Phi(s)\,\mathrm{d}\widetilde{W}(s)\Big\|^p\Big) \le \mathbb{E} \sup_{t\in[0,T]} \Big\| \int_0^t \Phi(s)\,\mathrm{d}W(s)\Big\|^p$$
$$\le C_{p,E}\mathbb{E}\Big(\widetilde{\mathbb{E}} \sup_{t\in[0,T]} \Big\| \int_0^t \Phi(s)\,\mathrm{d}\widetilde{W}(s)\Big\|^p\Big).$$

We now give an analogue to Proposition 3.29 for arbitrary adapted processes $\Phi : [0,T] \times \Omega \to \mathcal{L}(H,E)$. We say a process $\Phi : [0,T] \times \Omega \to \mathcal{L}(H,E)$ is adapted (to the filtration $\{\mathscr{F}_t\}_{t\geq 0}$) if $\Phi(t)$ is H-strongly measurable with respect to $\mathscr{F}_t$ for all $t \in [0,T]$.

Theorem 3.33 (Proposition 3.6 of [68])**.** *Let E be a UMD space and fix $p \in (1,\infty)$. For an H-strongly measurable and adapted process $\Phi : [0,T] \times \Omega \to \mathcal{L}(H,E)$, the following are equivalent:*

(1) Φ is an L^p-stochastically integrable process with respect to $W(t)$. That is, there exists a sequence (Φ_n) of elementary adapted processes such that

(a) for all $h \in H$, $\lim_{n\to\infty} \Phi_n h = \Phi h$ in measure on $[0,T] \times \Omega$;

(b) there exists $\zeta \in L^p\big(\Omega; C([0,T];E)\big)$ such that

$$\zeta(\cdot) = \lim_{n\to\infty} \int_0^{\cdot} \Phi_n(s)\, \mathrm{d}W(s) \text{ in } L^p\big(\Omega, C([0,T];E)\big);$$

(2) for all $x^* \in E^*$ the process $\Phi^* x^*$ is in $L^p\big(\Omega; L^2(0,T;H)\big)$ and there exists a pathwise continuous process $\xi \in L^p\big(\Omega; C([0,T];E)\big)$ such that

$$\langle \xi(\cdot), x^* \rangle = \int_0^{\cdot} \Phi(s)^* x^* \,\mathrm{d}W(s) \text{ in } L^p(\Omega; C[0,T]);$$

(3) for all $x^* \in E^*$, the function $\Phi(\omega,\cdot)^* x^*$ is in $L^2(0,T;H)$ for almost all $\omega \in \Omega$ and Φ represents an element $R \in L^p\big(\Omega; \gamma\big(L^2(0,T;H);E\big)\big)$, that is, for all $f \in L^2(0,T;H)$ and $x^* \in E^*$ we have

$$\langle Rf, x^* \rangle = \int_0^T [f(t), \Phi(t)^* x^*]_H \,\mathrm{d}t \text{ almost surely;}$$

(4) for almost all $\omega \in \Omega$, the function $\Phi_\omega := \Phi(\omega,\cdot)$ is stochastically integrable (in the sense of Definition 3.28) with respect to an independent H-cylindrical Brownian motion $\widetilde{W}(t)$ and

$$\omega \mapsto \int_0^T \Phi(\omega,s)\,\mathrm{d}\widetilde{W}(s)$$

defines an element of $L^p\big(\Omega; L^p(\widetilde{\Omega};E)\big)$.

In this case the random variables ζ and ξ are uniquely determined and equal in $L^p(\Omega;E)$. The random variable R is in $L^p\big(\Omega;\gamma(L^2(0,T;H),E)\big)$, and moreover

$$\Big(\mathbb{E} \sup_{t\in[0,T]} \Big\| \int_0^t \Phi(s)\,\mathrm{d}W(s) \Big\|^p \Big)^{\frac{1}{p}} \simeq_p \Big(\mathbb{E} \sup_{t\in[0,T]} \Big\| \int_0^t \Phi(s)\,\mathrm{d}W(s) \Big\|^2 \Big)^{\frac{1}{2}} \simeq \mathbb{E}\|R\|_{\gamma(L^2(0,T;H),E)}. \tag{3.12}$$

We will refer to the final line (3.12) the Itô isometry.

If E is a Banach space of type 2 then there is a convenient subspace of integrable processes which greatly simplifies dealing with the stochastic integral on these spaces.

Corollary 3.34 (Corollary 3.10 of [68]). *Let E be a UMD space of type 2 and let $p \in (1, \infty)$. Then for every H-strongly measurable and adapted process $\Phi \in L^p\big(\Omega; L^2(0, T; \gamma(H, E))\big)$, Φ is L^p-stochastically integrable with respect to W, and*

$$\mathbb{E} \sup_{t \in [0,T]} \Big\| \int_0^t \Phi(s) \, \mathrm{d}W(s) \Big\|^p \lesssim_p \mathbb{E} \|\Phi\|^p_{L^2(0,T;\gamma(H,E))}.$$

Corollary 3.34 also holds if UMD space of type 2 is replaced with Martingale type 2 spcae. See for example [20].

3.5 R and γ-bounded families of operators

In generalising various types of multiplier theorem from Hilbert spaces to Banach spaces (for example, [89] or the Fourier multiplier theorems 1.6 and 1.10 on pages 74–76 of [57]) it is often necessary to replace the notion of uniform boundedness of a family of operators with a stronger condition, that of R-boundedness.

Definition 3.35. Let (r_n) be a Rademacher sequence on $(\Omega, \mathscr{F}, \mathbb{P})$ and let E and F be Banach spaces. A family of bounded linear operators $\mathscr{T} \subset \mathcal{L}(E, F)$ is said to be *R-bounded* if there exists $M \geq 0$ such that

$$\left(\mathbb{E} \Big\| \sum_{n=1}^N r_n T_n x_n \Big\|_F^2 \right)^{\frac{1}{2}} \leq M \left(\mathbb{E} \Big\| \sum_{n=1}^N r_n x_n \Big\|_E^2 \right)^{\frac{1}{2}}$$

for all $N \geq 1$ and all finite sequences $(T_n)_{n=1}^N \subset \mathscr{T}$ and $(x_n)_{n=1}^N \subset E$. The least such M is called the *R-bound* of $\mathscr{T}$ and is written $\mathcal{R}(\mathscr{T})$.

By the Kahane-Khinchine inequalities, Proposition 3.15, the exponent 2 can be replaced by any $p \in [1, \infty)$ (however, possibly with a different constant $\mathcal{R}_p(\mathscr{T})$).

Definition 3.36. If the Rademacher sequence in Definition 3.35 is replaced by a Gaussian sequence (γ_n) then we obtain the related notion of γ-boundedness. The least constant in this case is called the *γ-bound* of $\mathscr{T}$ and is denoted $\gamma(\mathscr{T})$. Again, the 2 may be replaced with any $p \in [1, \infty)$.

Every R-bounded family $\mathscr{T} \subset \mathcal{L}(E, F)$ is γ-bounded with $\gamma(\mathscr{T}) \leq \mathcal{R}(\mathscr{T})$. If E has finite co-type then R and γ-boundedness are equivalent and there exists $C \geq 0$ such that $\mathcal{R}(\mathscr{T}) \leq C\gamma(\mathscr{T})$. If E and F are Hilbert spaces then both notions are the same as uniform boundedness and $\mathcal{R}(\mathscr{T}) = \gamma(\mathscr{T}) = \sup_{T \in \mathscr{T}} \|T\|$.

The importance of γ-bounded families comes from the fact that they act as pointwise multipliers on spaces of stochastically integrable processes, as we shall see in following lemma.

Lemma 3.37 (Lemma 2.9 of [69])**.** *Let E, F be Banach spaces and suppose that E does not contain a copy of c_0. Let H be a separable Hilbert space and $T > 0$. Suppose $M : (0, T) \to \mathcal{L}(E, F)$ satisfies:*

(1) for all $x \in E$, the function $M(\cdot)x : (0, T) \to F$ is strongly measurable;

(2) the range $\mathcal{M} = \big\{M(t) \in \mathcal{L}(E, F) : t \in (0, T)\big\}$ is γ-bounded in $\mathcal{L}(E, F)$.

Then for all step functions $\Phi : (0, T) \to \mathcal{L}(H, E)$ which take values in the finite rank operators we have

$$\|M\Phi\|_{\gamma(L^2(0,T;H),F)} \leq \gamma(\mathcal{M})\|\Phi\|_{\gamma(L^2(0,T;H),E)}.$$

Here $(M\Phi)(t) = M(t)\Phi(t)$. The mapping $\Phi \mapsto M\Phi$ has a unique extension to a bounded operator from $\gamma\big(L^2(0, T; H), E\big)$ to $\gamma\big(L^2(0, T; H), F\big)$ of norm at most $\gamma(\mathcal{M})$.

We now give two examples of $\mathcal{R}$- (and hence, γ-) bounded families that crop up later.

Lemma 3.38 (Proposition 2.5 of [89])**.** *Let E and F be Banach spaces. If $S : [0, T] \to \mathcal{L}(E, F)$ is differentiable on $(0, T)$ and $S'(t)$ is uniformly Bochner integrable, then the family*

$$\mathscr{T}_S = \big\{S(t) : t \in (0, T)\big\}$$

is R-bounded in $\mathcal{L}(E, F)$, with

$$\mathcal{R}(\mathscr{T}_S) \leq \big\|S(0)\big\| + \int_0^T \big\|S'(t)\big\| \, \mathrm{d}t.$$

The next Lemma refers to the abstract Sobolev space $E_a = D((-A)^a)$ together with the graph norm $\|x\|_{E_a} = \|x\| + \|(-A)^a x\|$ for $x \in D((-A)^a)$.

Lemma 3.39 (Proposition 4.1 of [69]). *If $S(t)$ is an analytic C_0-semigroup on a Banach space E, then for all $a \in [0,1)$ and $\varepsilon > 0$ the family*

$$\mathscr{T}_S := \left\{ t^{a+\varepsilon} S(t) \in \mathcal{L}(E, E_a) : t \in [0,T] \right\}$$

is R-bounded in $\mathcal{L}(E, E_a)$, with R-bound of order $\mathcal{O}(T^\varepsilon)$ as $T \downarrow 0$.

Proof. Define $N : [0,T] \to \mathcal{L}(E, E_a)$ by $N(t) = t^{a+\varepsilon} S(t)$. $N(t)$ is continuously differentiable on $(0,T)$ and by the product rule $N'(t) = (a+\varepsilon)t^{a+\varepsilon-1}S(t) + t^{a+\varepsilon} AS(t)$ where A is the generator of $S(t)$. Now by (2.7) of Theorem 2.9,

$$\left\| N'(t) \right\|_{\mathcal{L}(E,E_a)} \leq C t^{\varepsilon - 1} \text{ for } t \in (0,T),$$

and so Lemma 3.38 gives

$$\mathcal{R}(\mathscr{T}_S) \leq \int_0^T \left\| N'(t) \right\|_{\mathcal{L}(E,E_a)} \mathrm{d}t \leq C T^\varepsilon. \qquad \square$$

Chapter 4

Stochastic delay problems

4.1 Introduction

In this section we build on the work of van Neerven *et al.*, particularly [69] and [86], by extending their primary results on the well posedness of two classes of stochastic differential equations with multiplicative noise term to equations based on a (possibly infinite) delay. We will give two theorems, the first concerning non-autonomous delay problems on UMD spaces of type 2 (for example, an L^p space for $p \geq 2$) and the second for an autonomous problem on general UMD spaces (which include L^p for $p \in (1, \infty)$). The theory in the later case is rather more sophisticated, leading to the restriction to autonomous problems.

In both theorems that follow we adopt the strategy taken by van Neerven, Veraar, *et al.* in [86] and [69] but replace the dependence on the current state $X(t)$ of the system with dependence on the entire history X_t of the system up time t, which lies in a space of the type $\mathscr{B}$ introduced in Definition 2.20. This causes the principal new difficulty, that of relating the various fixed-point space norms to those on the history space. In the type 2 spaces we allow the full range of abstract history spaces defined in Definition 2.20, however in UMD spaces we restrict to a class of weighted L^p spaces in order to achieve this.

The key difference between the two results in this section is the restriction in the first instance to UMD spaces of type 2 (see Definition 3.16). In a type 2 Banach space there is an easy subspace of integrable processes that is sufficient for our purposes. By Proposition 3.34, for an H-strongly measurable and adapted process $\Phi : [0, T] \times \Omega \to \gamma(H, E)$ which is in $L^2((0, T) \times \Omega; \gamma(H, E))$ we can define (see section 3.4) the stochastic integral

$$\int_0^T \Phi(s) \, \mathrm{d}W(s)$$

as a limit of integrals of adapted step processes and there exists a constant C, independent of Φ such that

$$\left(\mathbb{E}\Big\|\int_0^T \Phi(s)\,\mathrm{d}W(s)\Big\|^2\right)^{\frac{1}{2}} \leq C\|\Phi\|_{L^2((0,T)\times\Omega;\gamma(H,E))}.$$

However, if E is an arbitrary UMD space of type $\tau \in (1,2)$ then the picture is not so simple. Stochastic integrability in UMD spaces is characterised by Theorem 3.33. An H-strongly measurable and adapted process $\Phi : [0,T]\times\Omega \to \gamma(H,E)$ is stochastically integrable if Φ is a representation of an operator $R \in L^2(\Omega;\gamma(L^2(0,T;H),E))$ satisfying certain conditions and in this case the best we can say is that there exists a constant C such that

$$\mathbb{E}\Big\|\int_0^T \Phi(s)\,\mathrm{d}W(s)\Big\|^2 \leq C\mathbb{E}\|R\|^2_{\gamma(L^2(0,T;H),E)}.$$

The difficulty in proving existence results in general UMD spaces is in finding a correct space of continuous processes in which to apply a fixed point theorem. Any such space V would need the property that for $U \in V$, both the deterministic and stochastic convolutions,

$$\int_0^t S(t-s)F(s,U(s))\,\mathrm{d}s, \text{ and } \int_0^t S(t-s)G(s,U(s))\,\mathrm{d}W(s), \quad t\in[0,T]$$

respectively lie in V under suitable Lipschitz conditions on F and G. The problem that van Neerven *et al.* overcome in [69] is that, by [67, Theorem 3], if $f(u)$ is stochastically integrable for every stochastically integrable process u and Lipschitz function $f : E \to E$, then E will be isomorphic to a Hilbert space. They resolve this issue by introducing the stronger notion of L^2_γ-Lipschitz functions and require only that the stochastic convolution is in V for maps G with this property. Under the assumption that F and G are Lipschitz and L^2_γ-Lipschitz in the second variable respectively, the problem reduces to finding a space V such that $\phi \in V$ implies

$$t \mapsto \int_0^t S(t-s)\phi(s)\,\mathrm{d}s, \text{ and } t \mapsto \int_0^t S(t-s)\phi(s)\,\mathrm{d}W(s), \quad t\in[0,T]$$

are in V. The fixed point space proposed by van Neerven *et al.* is defined in Section 4.2.3, but is essentially the intersection of the adapted processes in $L^p(\Omega;C([0,T];E))$ with a weighted version of the space $L^p\big((0,T)\times\Omega;\gamma(L^2(0,T;H),E)\big)$.

4.2 Non autonomous delay equations

The first problem we consider is the following non-autonomous infinite delay equation on a UMD Banach space of type 2, following the approach of [86].

$$\begin{cases} \mathrm{d}U(t) &= \big(A(t)U(t) + F(t,U_t)\big)\,\mathrm{d}t + G(t,U_t)\,\mathrm{d}W(t), \quad t \in [0,T_0], \\ U(0) &= X : \Omega \to E, \\ U_0 &= \Phi : (-\infty,0] \times \Omega \to E. \end{cases} \tag{4.1}$$

4.2.1 Hypotheses

We assume that $A(t)$ satisfies the condition (AT) and make the following additional assumptions on E, A, F, G and the constants $\eta, \theta_F, \theta_G \geq 0$ and $p > 2$:

(H1) E is a UMD Banach space of type 2 and the operators $A(t)$ on E for $t \in [0,T_0]$ have constant domains $D(A(t)) = D(A(0)) := E_1$ with uniformly equivalent graph norms.

(H2) The constants $\eta \geq 0$, $\theta_F, \theta_G \in [0,\mu)$, for μ as in (AT2), and $p > 2$ are such that $\eta + \frac{1}{p} < \min\{1-\theta_F, \frac{1}{2} - \theta_G, \eta_0\}$.

(H3) The history space $\mathscr{B}$ of pairs (x,ϕ) for $x \in E$ and functions $\phi : (-\infty,0] \to E$ is of the form defined in Definition 2.20 and for $\eta \in [0,1]$ we define the space $\mathscr{B}_\eta$ by

$$\mathscr{B}_\eta := \{(-A(\cdot))^{-\eta}\phi(\cdot) : \phi \in \mathscr{B}\}$$

with norm

$$\|\psi\|_{\mathscr{B}_\eta} := \|(-A(\cdot))^{\eta}\psi(\cdot)\|_{\mathscr{B}}.$$

Then $\mathscr{B}_\eta$ is a history space by the uniform equivalence of graph norms for $A(t)$ and Remark 2.23.

(H4) The function $F : [0,T_0] \times \Omega \times \mathscr{B}_\eta \to E$ is such that $(-A(\cdot))^{-\theta_F}F$ is Lipschitz continuous of linear growth in its third variable, uniformly on $[0,T_0] \times \Omega$. That is, there exist constants L_F and C_F such that for all $t \in [0,T_0]$, $\omega \in \Omega$ and $\phi, \psi \in \mathscr{B}_\eta$

$$\begin{aligned} \big\|(-A(t))^{-\theta_F}\big(F(t,\omega,\phi) - F(t,\omega,\psi)\big)\big\|_E &\leq L_F\|\phi-\psi\|_{\mathscr{B}_\eta}, \\ \big\|(-A(t))^{-\theta_F}F(t,\omega,\phi)\big\|_E &\leq C_F(1+\|\phi\|_{\mathscr{B}_\eta}). \end{aligned}$$

Moreover, the function $(t, \omega, \phi) \mapsto (-A(t))^{-\theta_F} F(t, \omega, \phi)$ is strongly measurable and adapted.

(H5) The function $G : [0, T_0] \times \Omega \times \mathscr{B}_\eta \to \gamma(H, E)$ is such that $(-A(\cdot))^{-\theta_G} G$ is Lipschitz continuous of linear growth in its third variable, uniformly on $[0, T_0] \times \Omega$. That is, there exist constants L_G and C_G such that for all $t \in [0, T_0]$, $\omega \in \Omega$ and $\phi, \psi \in \mathscr{B}_\eta$

$$\begin{aligned} \big\|(-A(t))^{-\theta_G}\big(G(t, \omega, \phi) - G(t, \omega, \psi)\big)\big\|_{\gamma(H,E)} &\leq L_G \|\phi - \psi\|_{\mathscr{B}_\eta}, \\ \big\|(-A(t))^{-\theta_G} G(t, \omega, \phi)\big\|_{\gamma(H,E)} &\leq C_G(1 + \|\phi\|_{\mathscr{B}_\eta}). \end{aligned}$$

Moreover, the function $(t, \omega, \phi) \mapsto (-A(t))^{-\theta_G} G(t, \omega, \phi)$ is strongly measurable and adapted.

(H6) The initial history data $(U(0), U_0) = (X, \Phi) \in L^p(\Omega, \mathscr{F}_0, \mathscr{B}_\eta)$.

4.2.2 Convolution estimates

It will be necessary to use several convolution estimates from [86]. We give statements here without proof for brevity.

Suppose $\big(A(t), D(A(t))\big)_{t \geq 0}$ is a family of operators on E satisfying the condition (AT) and generating an evolution family $(P(t, s))_{0 \leq s \leq t \leq T}$. For a semigroup $(S(t))_{t \geq 0}$ or an evolution family $(P(t, s))_{0 \leq s \leq t}$ we will write the deterministic and stochastic convolutions with the functions $F : [0, T_0] \to E$ and $G : [0, T_0] \to \gamma(H, E)$ respectively as

$$P * F(t) := \int_0^t P(t, s) F(s) \, \mathrm{d}s, \text{ and } P \diamond G(t) := \int_0^t P(t, s) G(s) \, \mathrm{d}W(s). \tag{4.2}$$

and similarly for S with $P(t, s)$ replaced by $S(t - s)$.

Lemma 4.1 (Lemma 2.3 of [86]). *Assume (AT) and (H1). Let $\eta \in (0, \eta_0]$ and $\delta, \lambda > 0$ be such that $\delta + \lambda \leq \eta$, then there exists $C \geq 0$ such that for all $0 \leq r \leq s \leq t \leq T$ and all $x \in E_\eta^r$*

$$\|P(t, r)x - P(s, r)x\|_{E_\delta^r} \leq C|t - s|^\lambda \|x\|_{E_\eta^r}.$$

Moreover, if $\eta \in [0, \eta_0)$ then $P(\cdot, r)$ is strongly continuous on E_η^r, that is, $P(\cdot, r)x \in C([r, T]; E_\eta)$ for every $x \in E_\eta^r$.

Proposition 4.2 (Proposition 3.2 of [86]). *Assume (AT) and (H1). Let* $\theta \in [0, \mu)$ *(where* μ *is as in (AT)),* $p \in (1, \infty]$, $\delta \in [0, 1)$ *and* $\lambda \in (0, 1)$ *be such that* $\lambda + \delta + \frac{1}{p} < \min\{1 - \theta\}, \eta_0\}$. *Then there exists* $C_T \geq 0$ *with* $\lim_{T \downarrow 0} C_T = 0$ *such that for all* $\phi : [0, T] \to E$ *such that* $(-A)^{-\theta}\phi \in L^p(0, T; E)$,

$$\|P * \phi\|_{C^\lambda([0,T];E_\delta)} \leq C_T \|(-A)^{-\theta}\phi\|_{L^p(0,T;E)}.$$

Proposition 4.3 (Theorem 4.1 of [86]). *Assume (AT) and (H1). Let* $\theta \in [0, \mu \wedge \frac{1}{2})$, $p \in (2, \infty)$ *and* $\delta, \lambda > 0$ *be such that* $\delta + \lambda + \frac{1}{p} < \min\{\frac{1}{2} - \theta, \eta_0\}$. *If the map* $(-A)^{-\theta}\Phi : [0, T] \times \Omega \to \gamma(H, E)$ *is* H*-strongly measurable and adapted with* $(-A)^{-\theta}\Phi \in L^p(\Omega; L^\infty(0, T; \gamma(H, E)))$ *almost surely, then the stochastic convolution* $P \diamond \Phi$ *exists in* E_δ *and is* λ*-Hölder continuous such that*

$$\|P \diamond \Phi\|^p_{L^p(\Omega;C^\lambda([0,T];E_\delta))} \leq C_T \mathbb{E}\|((-A)^{-\theta}\Phi)(\cdot)\|^p_{L^\infty(0,T;\gamma(H,E))}$$

for some constant $C_T \geq 0$ *independent of* Φ *with* $C_T \to 0$ *as* $T \downarrow 0$.

4.2.3 Well posedness and regularity results

Definition 4.4. Assume (AT) and (H1) - (H6). An E_η-valued process $(U(t))_{t\in(-\infty,T_0]}$ is a *mild solution* of (4.1) if

(1) $U(t)$ for $t \in [0, T_0]$ is strongly measurable, adapted to $\mathscr{F}_t$ and in $L^p(0, T_0; E_\eta)$ a.s.,

(2) $U(t) = \Phi(t)$ a.s. for $t \in (-\infty, 0]$,

(3) for all $t \in [0, T_0]$,

$$U(t) = P(t, 0)X + P * F(\cdot, U.)(t) + P \diamond G(\cdot, U.)(t) \text{ a.s.}$$

where $P * F$ and $P \diamond G$ are as in (4.2).

For η, p as in (H2) and $T \in (0, T_0]$, we let $Z^p_{\eta,T}$ be the closed subspace of adapted processes in $L^p(\Omega; C([0, T]; E_\eta))$. For $\phi \in Z^p_{\eta,T}$ write $\widetilde{\phi}$ for the extension of ϕ to $(-\infty, 0]$ by Φ, so

$$\widetilde{\phi}(t) = \begin{cases} \Phi(t), & t \in (-\infty, 0), \\ \phi(t), & t \in [0, T]. \end{cases}$$

Define the fixed point operator L on $Z^p_{\eta,T}$ by

$$L(\phi)(t) = P(t, 0)X + P * F(\cdot, \widetilde{\phi}.)(t) + P \diamond G(\cdot, \widetilde{\phi}.)(t). \tag{4.3}$$

We claim that L is well defined and a strict contraction on $Z^p_{\eta,T}$ for some sufficiently small $T \in (0, T_0]$.

Proposition 4.5. *Assume (AT) and (H1) - (H6) on a Banach space E of type 2. Then the operator L is well defined and there exists $T \in (0, T_0]$ such that for all $\phi, \psi \in Z^p_{\eta,T}$*

$$\|L(\phi) - L(\psi)\|_{Z^p_{\eta,T}} \le \frac{1}{2}\|\phi - \psi\|_{Z^p_{\eta,T}}. \tag{4.4}$$

Moreover, there exists a constant C, independent of (X, Φ) such that for all $\phi \in Z^p_{\eta,T}$

$$\|L(\phi)\|_{Z^p_{\eta,T}} \le C\big(1 + \big(\mathbb{E}\|(X,\Phi)\|^p_{\mathscr{B}_\eta}\big)^{\frac{1}{p}}\big) + \frac{1}{2}\|\phi\|_{Z^p_{\eta,T}} \tag{4.5}$$

Before giving the proof of Lemma 4.5 we recall a technical Lemma for an estimate on deterministic convolution. This appears almost as in the proof of Lemma 6.1 in [86], but we present it separately here as the delay plays no role.

Lemma 4.6 (Deterministic estimate). *Under the assumptions of Proposition 4.5, let $\Psi : [0,T] \times \Omega \to E$ be such that $(-A(t))^{-\theta_F}\Psi(t) \in L^p(\Omega; L^\infty(0,T;E))$. Then for all $t \in [0,T]$,*

$$\begin{aligned}\|P * \Psi\|_{L^p(\Omega;C([0,T];E_\eta))} &\lesssim \mathbb{E}\sup_{t\in[0,T]}\Big(\int_0^t (t-s)^{-p(\eta+\theta_F)}\|A(s)^{-\theta_F}\Psi(s)\|^p_E\,\mathrm{d}s\Big)^{\frac{1}{p}} \\ &\lesssim C_T\|A(\cdot)^{-\theta_F}\Psi(\cdot)\|_{L^p(\Omega;L^\infty(0,T;E))} \end{aligned}\tag{4.6}$$

where $C_T \to 0$ as $T \downarrow 0$.

Proofs of the above are contained in the proof of [86, Lemma 6.1]. We can now give the proof of Proposition 4.5.

Proof of Proposition 4.5. By linearity, it is enough to show that there exists a $T \in (0, T_0]$ such that each term of (4.3) satisfies (4.4) and (4.5).

Initial value part. By (2.14) of Theorem 2.17, we have the estimate

$$\|P(t,0)X\|_{E_\eta} \le C\|X\|_{E_\eta}, \quad t \in [0,T],$$

which clearly implies

$$\|P(\cdot,0)X\|_{Z^p_{\eta,T}} = \|P(\cdot,0)X\|_{L^p(\Omega;C([0,T];E_\eta)} \lesssim \Big(\mathbb{E}\|X\|^p_{E_\eta}\Big)^{\frac{1}{p}} \lesssim \Big(\mathbb{E}\|(X,\Phi)\|^p_{\mathscr{B}_\eta}\Big)^{\frac{1}{p}}, \tag{4.7}$$

and this is path continuous by Lemma 4.1.

Deterministic convolution. Let $\phi, \psi \in Z^p_{\eta,T}$, then by (H4), $(-A(t))^{-\theta_F} F(t, \widetilde{\phi}_t)$ and $(-A(t))^{-\theta_F} F(t, \widetilde{\psi}_t)$ are adapted since $t \mapsto \phi_t, \psi_t$ are continuous and lie in the space $L^p(\Omega; L^\infty(0,T;E))$ by a simple application of condition (A0)(iii) from Definition 2.20.

By Lemma 4.6 and Proposition 4.2 we have $P * F(\cdot, \widetilde{\phi}.)$ and $P * F(\cdot, \widetilde{\psi}.) \in Z^p_{\eta,T}$ and

$$\begin{aligned}\big\|P * F(\cdot, \widetilde{\phi}.) - P * F(\cdot, \widetilde{\psi}.)\big\|^p_{L^p(\Omega; C([0,T];E_\eta))} &\\ \lesssim \mathbb{E} \sup_{s\in[0,T]} \big\|A(s)^{-\theta_F}\big[F(s, \widetilde{\phi}_s) - F(s, \widetilde{\psi}_s)\big]\big\|^p_E &\\ \overset{(H4)}{\lesssim} L^p_F \mathbb{E} \sup_{s\in[0,T]} \big\|\widetilde{\phi}_s - \widetilde{\psi}_s\big\|^p_{\mathscr{B}_\eta}&\end{aligned}$$

Now by (A0)(iii) (2.22) for each $\omega \in \Omega$ we have (since $\widetilde{\phi}_0 = \widetilde{\psi}_0 = \Phi$)

$$\|\widetilde{\phi}_s - \widetilde{\psi}_s\|_{\mathscr{B}_\eta} \le \sup_{s\in[0,T]} K(s) \sup_{0\le r\le s} \|\phi(r) - \psi(r)\|_{E_\eta} \le K_T \|\phi - \psi\|_{C([0,T];E_\eta)} \tag{4.8}$$

where $K_T = \sup_{t\in[0,T]} K(t)$, so

$$\big\|P * F(\cdot,\widetilde{\phi}.) - P * F(\cdot, \widetilde{\psi}.)\big\|^p_{L^p(\Omega;C([0,T];E_\eta))} \lesssim L^p_F K^p_T \|\phi - \psi\|_{L^p(\Omega;C([0,T];E_\eta))} \tag{4.9}$$

Stochastic convolution. Consider functions $\phi, \psi \in Z^p_{\eta,T}$, then by (H5) we have $(-A(t))^{-\theta_G} G(t, \widetilde{\phi}_t)$ and $(-A(t))^{-\theta_G} G(t, \widetilde{\psi}_t)$ adapted and in $L^p(\Omega; L^\infty(0,T;\gamma(H,E)))$.

By Proposition 4.3, $P \diamond G(\cdot, \phi.)$ and $P \diamond G(\cdot, \psi.)$ are in $Z^p_{\eta,T}$ and we have

$$\begin{aligned}\big\|P \diamond G(\cdot, \widetilde{\phi}.) - P \diamond G(\cdot, \widetilde{\psi}.)\big\|^p_{L^p(\Omega;C([0,T];E_\eta))} &\\ \lesssim \big\|(-A(t))^{-\theta_G}\big(G(t, \widetilde{\phi}_t) - G(t, \widetilde{\psi}_t)\big)\big\|^p_{L^p(\Omega;L^\infty(0,T;\gamma(H,E)))} &\\ \overset{(H5)}{\lesssim} L_G \big\|t \mapsto \widetilde{\phi}_t - \widetilde{\psi}_t\big\|^p_{L^p(\Omega;L^\infty(0,T;\mathscr{B}_\eta))} &\\ \overset{(A0)}{\lesssim} L_G K_T \big\|\phi - \psi\big\|^p_{L^p(\Omega;L^\infty(0,T;E_\eta))} &\end{aligned} \tag{4.10}$$

Collecting the estimates. It follows from (4.7), (4.9) and (4.10) that L is well defined. For $\phi, \psi \in Z^p_{\eta,T}$ we get

$$\|L(\phi) - L(\psi)\|_{Z^p_{\eta,T}} \lesssim C_T \|\phi - \psi\|_{Z^p_{\eta,T}}$$

where $C_T \to 0$ as $T \downarrow 0$. Taking sufficiently small T gives us

$$\|L(\phi) - L(\psi)\|_{Z^p_{\eta,T}} \le \frac{1}{2} \|\phi - \psi\|_{Z^p_{\eta,T}}$$

as required. Now recalling that $\widetilde{0}_t \in \mathscr{B}$ is defined by

$$\widetilde{0}_t(s) = \begin{cases} \Phi(t+s) & t+s<0 \\ 0 & t+s \geq 0 \end{cases}$$

we get (4.5) from (4.4) by observing that

$$\begin{aligned} \|L(0)\|_{Z^p_{\eta,T}} &\lesssim \|X\|_{L^p(\Omega,E^0_\eta)} + C_{T_0}\Big(1 + \|\widetilde{0}_r\|_{L^p(\Omega;C([0,T];\mathscr{B}_\eta))}\Big) \\ &\leq \|X\|_{L^p(\Omega,E^0_\eta)} + C_{T_0}\Big(1 + M(0)\|(0,\Phi)\|_{L^p(\Omega;\mathscr{B}_\eta)}\Big) \\ &\lesssim 1 + \big\|(X,\Phi)\big\|_{L^p(\Omega;\mathscr{B}_\eta)} \end{aligned}$$

□

Theorem 4.7. *Assume (AT) and (H1) - (H6). Then*

(1) there exists a unique mild solution U of (4.1) such that $U|_{[0,T_0]} \in Z^p_{\eta,T_0}$ and

$$\|U|_{[0,T_0]}\|_{Z^p_{\eta,T_0}} \leq C\Big(1 + \big(\mathbb{E}\big\|(X,\Phi)\big\|^p_{\mathscr{B}_\eta}\big)^{\frac{1}{p}}\Big). \tag{4.11}$$

(2) For every $\delta, \lambda > 0$ with $\delta + \lambda + \eta + \frac{1}{p} < \min\{\frac{1}{2} - \theta_G, 1 - \theta_F, \eta_0\}$ there exists a version of U such that $U|_{[0,T_0]} - P(\cdot,0)X$ is in $L^P(\Omega; C^\lambda([0,T_0]; E_{\delta+\eta}))$ with

$$\Big(\mathbb{E}\big\|U - P(\cdot,0)X\big\|^p_{C^\lambda([0,T];E_{\delta+\eta})}\Big)^{\frac{1}{p}} \leq C\Big(1 + \big(\mathbb{E}\|(X,\Phi)\|^p_{\mathscr{B}_\eta}\big)^{\frac{1}{p}}\Big). \tag{4.12}$$

Proof.

(1) By Proposition 4.5 and the Banach fixed point theorem 2.12 there exists a unique fixed point $V \in Z^p_{\eta,T}$ of L, then $U : (-\infty,T] \times \Omega \to E$ defined by $U = \widetilde{V}$ is the unique mild solution on $(-\infty,T]$.

To find a solution on all of $[0,T_0]$ we note that $(U(T),U_T) \in L^p(\Omega,\mathscr{F}_T,\mathscr{B}_\eta)$ and hence (H6) is satisfied for the initial condition $(U(T),U_T)$. By another application of Proposition 4.5 and the above we extend the solution U to the interval $(-\infty,2T]$, and by induction in this manner to the whole of $(-\infty,T_0]$. Equation (4.11) follows from (4.5).

(2) for U as in (1). we have, by Propositions 4.2 and 4.3 that

$$\begin{aligned} &\mathbb{E}\big\|P * F(\cdot,U_\cdot)\big\|^p_{C^\lambda([0,T_0];E_{\eta+\delta})} \\ &\qquad \lesssim \mathbb{E}\int_0^{T_0} \big\|(-A(t))^{-\theta_F}F(t,U_t)\big\|^p_E \,\mathrm{d}t \\ &\qquad \lesssim \mathbb{E}\int_0^{T_0} \big(1 + \|U_t\|_{\mathscr{B}_\eta}\big)^p \,\mathrm{d}t \\ &\qquad \lesssim \mathbb{E}\int_0^{T_0} \Big(1 + K(t)\sup_{0\leq s\leq t}\|U(s)\|_{E_\eta} + M(0)\|(X,\Phi)\|_{\mathscr{B}_\eta}\Big)^p \,\mathrm{d}t \\ &\qquad \lesssim 1 + \|U|_{[0,T_0]}\|_{Z^p_{\eta,T_0}} + \Big(\mathbb{E}\|(X,\Phi)\|^p_{\mathscr{B}_\eta}\Big)^{\frac{1}{p}} \end{aligned} \tag{4.13}$$

and similarly

$$\mathbb{E}\big\|P \diamond G(\cdot, U_\cdot)\big\|^p_{C^\lambda([0,T_0];E_{\eta+\delta})} \lesssim \mathbb{E}\big\|(-A)^{-\theta_G} G(\cdot, U_\cdot)\big\|^p_{L^p(0,T_0;\gamma(H,E))}$$
$$\lesssim 1 + \|U_{[0,T_0]}\|_{Z^p_{\eta,T_0}} + \Big(\mathbb{E}\|(X,\Phi)\|^p_{\mathscr{B}_\eta}\Big)^{\frac{1}{p}}. \qquad (4.14)$$

Define $V : [0, T_0] \times \Omega \to E_\eta$ by

$$V(t) = P(t, 0)X + P * F\big(\cdot, U_\cdot\big)(t) + P \diamond G\big(\cdot, U_\cdot\big)(t),$$

where we take the versions of the convolutions as above, then clearly $V = U|_{[0,T_0]}$ in Z^p_{η,T_0} and therefore $\widetilde{V}$ is the desired mild solution. Moreover, it follows from (4.13) and (4.14) that

$$\mathbb{E}\big\|V - P(\cdot, 0)X\|^p_{C^\lambda([0,T_0];E_{\delta+\eta})} \leq C\Big(1 + \|V\|_{Z^p_{\eta,T_0}} + \Big(\mathbb{E}\|(X,\Phi)\|^p_{\mathscr{B}_\eta}\Big)^{\frac{1}{p}}\Big)$$

and hence (4.12) follows from (4.11). □

4.3 Delay equations in UMD spaces

Next we look at the delay problem on a general UMD Banach space. In this case we restrict to autonomous problems and to history spaces of a certain class of weighted L^p spaces so as to make use of some explicit isometries on the history space.

Consider the following delayed stochastic Cauchy problem

$$\begin{cases} \mathrm{d}U(t) &= \big(AU(t) + F(t, U_t)\big)\,\mathrm{d}t + G(t, U_t)\,\mathrm{d}W_H(t),\ t \in [0, T_0] \\ U(0) &= X : \Omega \to E \\ U_0 &= \Phi : (-\infty, 0] \times \Omega \to E. \end{cases} \qquad (4.15)$$

4.3.1 L^2_γ-Lipschitz functions and convolution estimates

In order to avoid problems alluded to in Section 4.1, we define L^2_γ-Lipschitz functions as in [69, Section 5]. Let (S, Σ) be a countably generated measure space and let μ be a finite measure on (S, Σ). Define $L^2_\gamma(S, \mu; E)$ by

$$L^2_\gamma(S, \mu; E) := \gamma\big(L^2(S, \mu), E\big) \cap L^2(S, \mu; E),$$

the space of strongly μ-measurable functions $\phi : S \to E$ such that

$$\|\phi\|_{L^2_\gamma(S,\mu;E)} := \|\phi\|_{\gamma(L^2(S,\mu),E)} + \|\phi\|_{L^2(S,\mu;E)} < \infty. \tag{4.16}$$

The simple functions $S \to E$ form a dense subspace in $L^2_\gamma(S, \mu; E)$.

Suppose H is a non-zero separable Hilbert space and E and F are Banach spaces. Let $G : S \times E \to \mathcal{L}(H, F)$ be a function such that $s \mapsto G(s, x)$ is in $\gamma(L^2(S, \mu; H), F)$ for every $x \in E$. It follows easily that $s \mapsto G(s, \phi(s))$ is in $\gamma(L^2(S, \mu; H), F)$ for simple functions $\phi : S \to E$.

Definition 4.8. We say that $G : S \times E \to \mathcal{L}(H, F)$ is L^2_γ-*Lipschitz* if G is strongly continuous in the second variable and for any finite measure μ on (S, Σ) there exists a constant C_μ such that

$$\big\|G(\cdot, \phi(\cdot)) - G(\cdot, \psi(\cdot))\big\|_{\gamma(L^2(S,\mu;H),F)} \le C_\mu \|\phi - \psi\|_{L^2_\gamma(S,\mu;E)} \tag{4.17}$$

for all simple functions $\phi, \psi : S \to E$. In this case, by [69, Lemma 5.1] the map $S_{\mu,G} : \phi \mapsto G(\cdot, \phi(\cdot))$ extends uniquely to a Lipschitz map from $L^2_\gamma(S, \mu; E)$ into $\gamma(L^2(S, \mu; H), F)$ where the operator $S_{\mu,G}(\phi) \in \gamma(L^2(S, \mu; H), F)$ is represented by the function $G(\cdot, \phi(\cdot))$ for all $\phi \in L^2_\gamma(S, \mu; E)$.

The L^2_γ-Lipschitz property is strictly stronger than conventional Lipschitz continuity. From the proof of [67, Theorem 1] it follows that if $H \ne \{0\}$ then every Lipschitz function $g : E \to \gamma(H, F)$ is L^2_γ-Lipschitz if and only if F has type 2.

We now give several estimates on the deterministic and stochastic convolutions that will be needed in the sequel. These appear exactly as in [69] and so no proofs are given.

Proposition 4.9 (Proposition 3.5 of [69])**.** *Let E be a Banach space of type $\tau \in [1, 2]$ and let $0 \le \alpha \le \frac{1}{2}$. If $\eta, \theta \ge 0$ satisfy $\eta + \theta < \frac{3}{2} - \frac{1}{\tau}$ then there exists a constant $C \ge 0$ such that for all $0 \le t \le T \le T_0$ and all $\phi \in L^\infty(0, T; E)$,*

$$\big\|(t - \cdot)^{-\alpha}\big(S * \phi\big)(\cdot)\big\|_{\gamma(L^2(0,t);E_\eta)} \le CT^{\frac{1}{2}-\alpha}\|\phi\|_{L^\infty(0,T;E_{-\theta})}.$$

For $\alpha \in (0, 1]$ we define the following convolution,

$$R_\alpha\phi(t) := \frac{1}{\Gamma(\alpha)} \int_0^t (t - s)^{\alpha-1} S(t - s)\phi(s)\,\mathrm{d}s,$$

and then estimate space time regularity of R_α by with the following Lemma.

Lemma 4.10 (Lemma 3.6 of [69]). *Let* $\alpha \in (0,1]$, $p \in (1,\infty)$ *and* $\lambda, \eta, \theta \geq 0$ *be such that* $\lambda + \eta + \theta < \alpha - \frac{1}{p}$. *Then there exist constants* $C \geq 0$ *and* $\varepsilon > 0$ *such that for all* $\phi \in L^p(0,T;E)$ *and* $T \in [0,T_0]$,

$$\left\| R_\alpha \phi \right\|_{C^\lambda([0,T];E_\eta)} \leq CT^\varepsilon \|\phi\|_{L^p(0,T;E_{-\theta})}.$$

Next come estimates of the stochastic convolution $S \diamond \Phi$.

Proposition 4.11 (Proposition 4.2 of [69]). *Suppose* $\alpha \in (0, \frac{1}{2})$, $p \geq 2$ *and* $\lambda, \eta, \theta \geq 0$ *satisfy* $\lambda + \eta + \theta < \alpha - \frac{1}{p}$. *Let* A *be the generator of an analytic semigroup* S *on a UMD space* E *and let* $\Phi : (0,T) \times \Omega \to \mathcal{L}(H, E_{-\theta})$ *be measurable and adapted. Then there exists* $C \geq 0$ *and* $\varepsilon > 0$ *such that*

$$\mathbb{E}\left\| S \diamond \Phi \right\|^p_{C^\lambda([0,T];E_\eta)} \leq C^p T^{\varepsilon p} \int_0^T \mathbb{E}\left\| (t - \cdot)^{-\alpha} \Phi(\cdot) \right\|^p_{\gamma(L^2(0,t;H),E_{-\theta})} \,\mathrm{d}t.$$

Proposition 4.12 (Proposition 4.5 of [69]). *Let* E *be a UMD Banach space and let* $\alpha \in (0, \frac{1}{2})$ *and* $\eta, \theta \geq 0$ *satisfy* $\eta + \theta < \alpha$. *If* $\Phi : (0,T) \times \Omega \to \mathcal{L}(H, E_{-\theta})$ *is adapted and* H*-strongly measurable then for all* $p \in (1,\infty)$ *and* $0 \leq t \leq T \leq T_0$,

$$\mathbb{E}\left\| (t - \cdot)^{-\alpha} S \diamond \Phi(\cdot) \right\|^p_{\gamma(L^2(0,t;H),E_\eta)} \leq C^p T^{p(\frac{1}{2}-\eta-\theta)} \mathbb{E}\| (t - \cdot)^{-\alpha} \Phi(\cdot) \|^p_{\gamma(L^2(0,t;H),E_{-\theta})}.$$

4.3.2 Hypotheses

We make the following assumptions on A, F, G, X, Φ and the real constants $\eta, \theta_F, \theta_G \geq 0$, $\tau \in (1,2]$ and $p > 2$:

(D1) A generates an analytic C_0-semigroup on a UMD Banach space E of type τ and $W_H = \big(W_H(t)\big)_{t \in [0,T_0]} \subset \mathcal{L}\big(H, L^2(\Omega)\big)$ is an H-cylindrical Brownian motion on a Hilbert space H.

(D2) $0 \leq \eta + \theta_F < \frac{3}{2} - \frac{1}{\tau}$ and $0 \leq \eta + \theta_G < \frac{1}{2} - \frac{1}{p}$.

(D3) The history space $\mathscr{B}$ is of the form $E \times L^p_g(-\infty, 0; E)$ introduced in Example 2.22.*3*. For a function $\phi : (-\infty, a] \to E$ for $a > 0$ such that ϕ is continuous on $[0,a]$ and $(\phi(a), \phi_a) \in \mathscr{B}$, we associate ϕ with the pair $(\phi(\cdot), \phi_\cdot)$ and write $\phi_a \in \mathscr{B}$.

(D4) The function $F : [0,T_0] \times \Omega \times \mathscr{B}_\eta \to E_{-\theta_F}$ is continuous almost surely and Lipschitz of linear growth on the history space $\mathscr{B}$ uniformly with respect to $[0,T_0] \times \Omega$. In other words, there exist constants $L_F, C_F \geq 0$ such that for all $t \in [0,T_0], \omega \in \Omega$ and $\phi, \psi \in \mathscr{B}_\eta$ we have

$$\|F(t,\omega,\phi) - F(t,\omega,\psi)\|_{E_{-\theta_F}} \leq L_F \|\phi - \psi\|_{\mathscr{B}_\eta} \tag{4.18}$$

and

$$\|F(t,\omega,\phi)\|_{E_{-\theta_F}} \leq C_F\big(1 + \|\phi\|_{\mathscr{B}_\eta}\big) \tag{4.19}$$

and moreover, the function $(t,\omega,\phi) \mapsto F(t,\omega,\phi)$ is measurable and adapted (to $\mathscr{F}_t$) in $E_{-\theta_F}$.

(D5) The function $G : [0,T_0] \times \Omega \times \mathscr{B}_\eta \to \mathcal{L}(H, E_{-\theta_G})$ is continuous almost surely and L^2_γ-Lipschitz uniformly on Ω. In other words, there exist constants $L^\gamma_G, C^\gamma_G \geq 0$ such that for any finite measure μ on $([0,T_0], \mathbb{B}_{[0,T_0]})$, all $\omega \in \Omega$ and all $\phi, \psi : (-\infty, T_0] \to E_\eta$, continuous on $[0,T_0]$ with the property that the functions $(t \mapsto \phi_t)$ and $(t \mapsto \psi_t)$ lie in $L^2_\gamma(0,T_0,\mu;\mathscr{B}_\eta)$ we have

$$\begin{aligned}\|G(\cdot,\omega,\phi.) - G(\cdot,\omega,\psi.)\|_{\gamma(L^2(0,T_0,\mu;H),E_{-\theta_G})}& \\ \leq L^\gamma_G \|t \mapsto (\phi - \psi)_t\|_{L^2_\gamma(0,T_0,\mu;\mathscr{B}_\eta)}&\end{aligned} \tag{4.20}$$

and

$$\|G(\cdot,\omega,\phi.)\|_{\gamma(L^2(0,T_0,\mu;H),E_{-\theta_G})} \leq C^\gamma_G\big(1 + \|t \mapsto \phi_t\|_{L^2_\gamma(0,T_0,\mu;\mathscr{B}_\eta)}\big) \tag{4.21}$$

and moreover, for all $\phi \in \mathscr{B}_\eta$ the map $(t,\omega) \mapsto G(t,\omega,\phi)$ is H-strongly measurable and adapted (to $\mathscr{F}_t$) in $\mathcal{L}(H, E_{-\theta_G})$.

(D6) The initial history $(U(0), U_0) = (X, \Phi) : \Omega \to \mathscr{B}_\eta$ is strongly $\mathscr{F}_0$-measurable, $X \in L^p(\Omega; E_\eta)$ and if $\widehat{\Phi}$ represents the extension of Φ to $(-\infty, T_0]$ by

$$\widehat{\Phi}(t) := \begin{cases} 0 & t \in (0,T_0] \\ \Phi(t) & t \in (-\infty, 0]. \end{cases} \tag{4.22}$$

then the function mapping s to $(t-s)^{-\alpha}\widehat{\Phi}(\omega)_s$ lies in $L^2_\gamma\big(0,t;\mathscr{B}_\eta\big)$ for all $t \in [0,T_0]$, $\alpha \in (0, \frac{1}{2})$ and almost all $\omega \in \Omega$ with

$$\sup_{t\in[0,T_0]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_s\big\|_{L^2_\gamma(0,t;\mathscr{B}_\eta)} < \infty. \tag{4.23}$$

4.3.3 Well posedness and regularity results

Definition 4.13. A process $\big(U(t)\big)_{t\in(-\infty,T]}$ is said to be a *mild solution* of (4.15) if $U : [0,T_0]\times\Omega \to E_\eta$ is strongly measurable, adapted to $\mathscr{F}_t$ and for all $t \in [0,T_0]$

(i) U has continuous paths on $[0,T]$, $U(0) = X$ and $U|_{(-\infty,0]} = \Phi$ almost surely;

(ii) the function $s \mapsto S(t-s)F(s,U_s)$ is in $L^0\big(\Omega; L^1(0,t;E)\big)$;

(iii) the function $s \mapsto S(t-s)G(s,U_s)$ is H-strongly measurable, adapted to $\mathscr{F}_t$ and in $\gamma\big(L^2(0,t;H),E\big)$ almost surely;

(iv) $U(t) = S(t)X + S * F(\cdot,U.)(t) + S \diamond G(\cdot,U.)(t)$ almost surely.

Hence for a given solution U the deterministic convolution is pathwise well defined as a Bochner integral and the stochastic convolution is well defined by Theorem 3.33.

We will prove an existence and uniqueness result for (4.15) using a fixed point argument in a scale of Banach spaces introduced by van Neerven *et al.* in [69]. Fix $T \in (0,T_0]$, $p \in [1,\infty)$ and $\alpha \in (0,\frac{1}{2})$ and define $V^p_{\alpha,\infty}\big([0,T]\times\Omega;E\big)$ to be the space of all continuous adapted processes $\phi : [0,T]\times\Omega \to E$ for which

$$\begin{aligned} &\|\phi\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E)} \\ &\quad := \Big(\mathbb{E}\|\phi\|^p_{C([0,T];E)}\Big)^{\frac{1}{p}} + \sup_{t\in[0,T]} \Big(\mathbb{E}\big\|s\mapsto (t-s)^{-\alpha}\phi(s)\big\|^p_{\gamma(L^2(0,t),E)}\Big)^{\frac{1}{p}} \end{aligned} \tag{4.24}$$

is finite. Similarly, define $V^p_{\alpha,p}\big([0,T]\times\Omega;E\big)$ to be the space of pathwise continuous and adapted processes $\phi : [0,T]\times\Omega \to E$ for which

$$\begin{aligned} &\|\phi\|_{V^p_{\alpha,p}([0,T]\times\Omega;E)} \\ &\quad := \Big(\mathbb{E}\|\phi\|^p_{C([0,T];E)}\Big)^{\frac{1}{p}} + \left(\int_0^T \mathbb{E}\big\|s\mapsto (t-s)^{-\alpha}\phi(s)\big\|^p_{\gamma(L^2(0,t),E)}\,\mathrm{d}t\right)^{\frac{1}{p}} < \infty. \end{aligned} \tag{4.25}$$

Under identification of processes equivalent under the above norms, $V^p_{\alpha,\infty}\big([0,T]\times\Omega;E\big)$ and $V^p_{\alpha,p}\big([0,T]\times\Omega;E\big)$ become Banach spaces.

Our main result, Theorem 4.17, is to prove the existence and uniqueness of a mild solution of (4.15) in each of the spaces above with the initial history data (X,Φ).

Consider the fixed point operator defined on $V^p_{\alpha,\infty}$ or $V^p_{\alpha,p}$ by

$$L_T(\phi)(t) := S(t)X + S * F(\cdot,\widetilde{\phi}.)(t) + S \diamond G(\cdot,\widetilde{\phi}.)(t), \quad t\in[0,T] \tag{4.26}$$

where $\widetilde{\phi}$ is the extension of ϕ to $(-\infty, T]$ by Φ,

$$\widetilde{\phi}(t) := \begin{cases} \phi(t) & t \in [0,T] \\ \Phi(t) & t \in (-\infty, 0). \end{cases} \tag{4.27}$$

We will show that L_T is well defined and becomes a strict contraction for small enough T.

Proposition 4.14. *Suppose (D1) – (D6) are satisfied and choose $\alpha \in (0, \frac{1}{2})$ such that*

$$\eta + \theta_G < \alpha - \frac{1}{p}.$$

Then the operator L_T is well defined and bounded on either of the spaces $V = V^p_{\alpha,\infty}\big([0,T]\times\Omega; E_\eta\big)$ or $V^p_{\alpha,p}\big([0,T]\times\Omega; E_\eta\big)$, and moreover, there exist constants C, C_T with $\lim_{T\downarrow 0} C_T = 0$ such that for all $\phi, \psi \in V$

$$\|L_T(\phi) - L_T(\psi)\|_V \le C_T \|\phi - \psi\|_V \tag{4.28}$$

and

$$\begin{aligned} \|L_T(\phi)\|_V \le C\Big(1 + \big(\mathbb{E}\|(\Phi, X)\|^p_{\mathscr{B}_\eta}\big)^{\frac{1}{p}} \\ + \big(\sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_s\big\|^p_{L^2_\gamma(0,t;\mathscr{B})}\big)^{\frac{1}{p}}\Big) + C_T\|\phi\|_V \end{aligned} \tag{4.29}$$

First we give two more technical lemmas that will be required in the proof. These are exactly as in the proof of Proposition 6.1 in [69], but are presented separately here for clarity, as the delay plays no part in these estimates.

Lemma 4.15. *Assume (D1)-(D6). Under the conditions of Proposition 4.14, if $\Psi \in L^p\big(\Omega; C([0,T]; E_{-\theta_F})\big)$ then*

$$\big\|S * \Psi\big\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \le CT^{\min\{\frac{1}{2}-\alpha, 1-\eta-\theta_F\}}\|\Psi\|_{L^p(\Omega;C([0,T];E_{-\theta_F}))}. \tag{4.30}$$

Proof. Take $\psi \in C([0,T]; E_{-\theta_F})$ and find bounds for the $V^p_{\alpha,\infty}\big([0,T]\times\Omega; E_\eta\big)$-norm of $S * \psi$. Applying Lemma 4.10 with $\alpha = 1$ and $\lambda = 0$ we see that $S * \psi$ is continuous in E_η. The standard analytic property (2.7) of S gives

$$\begin{aligned} \big\|S * \psi\big\|_{C([0,T];E_\eta)} &\le C\int_0^t (t-s)^{-(\eta+\theta_F)}\,\mathrm{d}s\|\psi\|_{C([0,T];E_{-\theta_F})} \\ &\le CT^{1-\eta-\theta_F}\|\psi\|_{C([0,T];E_{-\theta_F})}. \end{aligned} \tag{4.31}$$

By assumption (D2) we satisfy the conditions of Proposition 4.9, so it follows that

$$\big\|s \mapsto (t-s)^{-\alpha} S * \psi(s)\big\|_{\gamma(L^2(0,t),E_\eta)} \leq T^{\frac{1}{2}-\alpha} \|\psi\|_{C([0,T];E_{-\theta_F})}. \tag{4.32}$$

Now let $\Psi \in L^p\big(\Omega; C([0,T]; E_{-\theta_F})\big)$. We apply (4.31) and (4.32) to the paths $\Psi(\cdot,\omega)$ and take expectations to see that $S * \Psi \in V^p_{\alpha,\infty}\big([0,T]\times\Omega; E\big)$ and

$$\big\|S * \Psi\big\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \leq C T^{\min\{\frac{1}{2}-\alpha, 1-\eta-\theta_F\}} \|\Psi\|_{L^p(\Omega;C([0,T];E_{-\theta_F}))}. \qquad \square$$

Lemma 4.16. *Assume (D1) - (D6). Under the conditions of Proposition 4.14 let* $\Psi : [0,T]\times\Omega \to \mathcal{L}(H, E_{-\theta_G})$ *be* H*-strongly measurable and adapted, and suppose that*

$$\sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\Psi(s)\big\|^p_{\gamma(L^2(0,t;H),E_{-\theta_G})} < \infty. \tag{4.33}$$

Then there exists $\varepsilon' > 0$ *such that*

$$\begin{aligned} &\big\|S \diamond \Psi\big\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \\ &\quad\leq\quad C T^{\varepsilon'} \Big(\sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\Psi(s)\big\|^p_{\gamma(L^2(0,t;H),E_{-\theta_G})}\Big)^{\frac{1}{p}}. \end{aligned} \tag{4.34}$$

Proof. Let $\Psi : [0,T]\times\Omega \to \mathcal{L}(H, E_{-\theta_G})$ be H-strongly measurable and adapted, and suppose that

$$\sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\Psi(s)\big\|^p_{\gamma(L^2(0,t;H),E_{-\theta_G})} < \infty. \tag{4.35}$$

We estimate the $V^p_{\alpha,\infty}\big([0,T]\times\Omega; E_\eta\big)$-norm of $S \diamond \Psi$. From Proposition 4.11 there exists an $\varepsilon > 0$ such that

$$\mathbb{E}\big\|S \diamond \Psi\big\|^p_{C([0,T];E_\eta)} \leq C^p T^{\varepsilon p} \sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\Psi(s)\big\|^p_{\gamma(L^2(0,t;H),E_{-\theta_G})}.$$

For the other part of the norm, by Proposition 4.12 we obtain that

$$\begin{aligned} \mathbb{E}\big\|s \mapsto &(t-s)^{-\alpha} S \diamond \Psi(s)\big\|^p_{\gamma(L^2(0,t;H),E_\eta)} \\ &\leq\quad C^p T^{(\frac{1}{2}-\eta-\theta_G)p} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\Psi(s)\big\|^p_{\gamma(L^2(0,t;H),E_{-\theta_G})}. \end{aligned}$$

Combining the above we conclude that for $\varepsilon' := \min\{\frac{1}{2}-\eta-\theta_G, \varepsilon\}$

$$\begin{aligned} &\big\|S \diamond \Psi\big\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \\ &\quad\leq\quad C T^{\varepsilon'} \Big(\sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\Psi(s)\big\|^p_{\gamma(L^2(0,t;H),E_{-\theta_G})}\Big)^{\frac{1}{p}}. \qquad \square \end{aligned} \tag{4.36}$$

We are now ready to proceed with the main result.

Proof (of Proposition 4.14). The proof proceeds along the same lines as the proof of Proposition 6.1 in [69], adapting as necessary for the delay. As in [69] we will give a detailed proof only in the case that $V = V^p_{\alpha,\infty}\big([0,T]\times\Omega; E\big)$, the other case being essentially the same.

Step 1 (Estimating the initial part).

It is clear that for fixed $\omega \in \Omega$

$$\big\|SX\big\|_{C([0,T];E_\eta)} \leq C\|X\|_{E_\eta}.$$

Let $\varepsilon \in (0, \frac{1}{2})$. From Lemma 3.37 and Lemma 3.39 we infer that for a fixed $\omega \in \Omega$, $t \in [0,T]$

$$\begin{aligned}
&\big\|s \mapsto (t-s)^{-\alpha} S(s)X\big\|_{\gamma(L^2(0,t),E_\eta)} \\
&\qquad \leq \quad C\big\|s \mapsto (t-s)^{-\alpha} s^{-\varepsilon} X\big\|_{\gamma(L^2(0,t),E_\eta)} \\
&\qquad \leq \quad C\big\|s \mapsto (t-s)^{-\alpha} s^{-\varepsilon}\big\|_{L^2(0,t)} \|X\|_{E_\eta} \\
&\qquad \leq \quad C\|X\|_{E_\eta}.
\end{aligned}$$

Then taking expectations, we get

$$\big\|S(\cdot)X\big\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E)} \leq C\|X\|_{L^p(\Omega;E_\eta)}. \tag{4.37}$$

Step 2 (Estimating the deterministic convolution).

Let $\phi, \psi \in V^p_{\alpha,\infty}\big([0,T]\times\Omega;E_\eta\big)$. Since F is continuous on $[0,T]\times\mathscr{B}_\eta$ and the map $t \mapsto \widetilde{\phi}_t$ is continuous by (A1) of Definition 2.20 we have, for fixed $\omega \in \Omega$, $t \in [0,T]$

$$\begin{aligned}
\|F(t,\widetilde{\phi}_t)\|_{E_{-\theta_F}} &\leq C_F\big(1 + \|\widetilde{\phi}_t\|_{\mathscr{B}_\eta}\big) \\
&\leq C_F\big(1 + K_T \sup_{s\in[0,t]} \|\phi(s)\|_{E_\eta} + M_T\|\Phi\|_{\mathscr{B}_\eta}\big) \\
&\lesssim 1 + \sup_{s\in[0,T]} \|\phi(s)\|_{E_\eta}.
\end{aligned} \tag{4.38}$$

Taking a supremum on the LHS and then expectations we get

$$\|F(\cdot,\widetilde{\phi}_\cdot)\|_{L^p(\Omega;C([0,T];E_{-\theta_F}))} \lesssim 1 + \|\phi\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)}$$

and the same for ψ, so both $F(\cdot,\widetilde{\phi}_\cdot)$ and $F(\cdot,\widetilde{\psi}_\cdot)$ belong to $L^p\big(\Omega;C([0,T];E_{-\theta_F})\big)$ and we can apply Lemma 4.15. The estimate (4.30) shows that $S * F(\cdot,\widetilde{\phi}_\cdot)$ and $S * F(\cdot,\widetilde{\psi}_\cdot) \in V^p_{\alpha,\infty}\big([0,T]\times\Omega;E_\eta\big)$. Note that by the property (2.22) of Definition 2.20

$$\|\widetilde{\phi}_t - \widetilde{\psi}_t\|_{\mathscr{B}_\eta} \leq \sup_{s\in[0,T]} K(s)\|\phi-\psi\|_{C([0,T],E_\eta)} \tag{4.39}$$

since $\widetilde{\phi}_0 = \widetilde{\psi}_0 = \Phi$. Combining (4.30) with the property that F is Lipschitz in its $\mathscr{B}_\eta$-variable (4.18) and setting $\delta = \min\{\frac{1}{2} - \alpha, 1 - \eta - \theta_F\}$ yields

$$\begin{aligned}
\big\|S * \big(F(\cdot,\widetilde{\phi}_\cdot) - F(\cdot,\widetilde{\psi}_\cdot)\big)\big\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \\
&\le CT^\delta \big\|F(\cdot,\widetilde{\phi}_\cdot) - F(\cdot,\widetilde{\psi}_\cdot)\big\|_{L^p(\Omega;C([0,T];E_{-\theta_F}))} \\
&= CT^\delta \Big(\mathbb{E} \sup_{t\in[0,T]} \big\|F(t,\widetilde{\phi}_t) - F(t,\widetilde{\psi}_t)\big\|^p_{E_{-\theta_F}}\Big)^{\frac{1}{p}} \\
&\le CT^\delta L_F \Big(\mathbb{E} \sup_{t\in[0,T]} \|(\widetilde{\phi} - \widetilde{\psi})_t\|^p_{\mathscr{B}_\eta}\Big)^{\frac{1}{p}} \\
&\le CT^\delta L_F K_{T_0} \Big(\mathbb{E}\|\phi - \psi\|^p_{C([0,T];E_\eta)}\Big)^{\frac{1}{p}} \\
&\le CT^\delta L_F K_{T_0} \|\phi - \psi\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)}. \qquad (4.40)
\end{aligned}$$

Step 3 (Estimating the stochastic convolution).

For $t \in [0,T]$, let $\mu_{t,\alpha}$ be the finite measure on $\big((0,t), \mathbb{B}_{(0,t)}\big)$ defined by

$$\mu_{t,\alpha}(B) := \int_0^t (t-s)^{-2\alpha} \mathbb{1}_B(s)\,\mathrm{d}s.$$

Notice that for a function $\phi \in C\big([0,t];E\big)$ we have

$$\phi \in \gamma\big(L^2(0,t,\mu_{t,\alpha}),E\big) \iff \big[s \mapsto (t-s)^{-\alpha}\phi(s)\big] \in \gamma\big(L^2(0,t),E\big).$$

and

$$\begin{aligned}
\|\phi\|_{L^2(0,t,\mu_{t,\alpha};E)} &= \big\|s \mapsto (t-s)^{-\alpha}\phi(s)\big\|_{L^2(0,t;E)} \\
&\le Ct^{\frac{1}{2}-\alpha}\|\phi\|_{C([0,T];E)}.
\end{aligned}$$

Now let $\phi,\psi \in V^p_{\alpha,\infty}\big([0,T]\times\Omega;E_\eta\big)$. Fix some $\omega \in \Omega$. The paths of ϕ,ψ belong to $L^2_\gamma\big(0,t,\mu_{t,\alpha};E_\eta\big)$ uniformly for $t \in [0,T]$ almost surely by the definition of the norm on $V^p_{\alpha,\infty}\big([0,T]\times\Omega;E_\eta\big)$ (4.24) and that the path of $s \mapsto \widehat{\Phi}_s$ is in $L^2_\gamma\big(0,t,\mu_{t,\alpha};\mathscr{B}_\eta\big)$ almost surely by assumption (D6).

Recall the definition of $\widetilde{\phi},\widetilde{\psi}$ from (4.27). We plan to show that $s \mapsto \widetilde{\phi}_s$ and $s \mapsto \widetilde{\psi}_s$ are contained in $L^2_\gamma\big(0,t,\mu_{t,\alpha};\mathscr{B}_\eta\big)$ almost surely.

Fix an $\omega \in \Omega$. From the definition (4.16) we have that

$$\big\|s \mapsto \widetilde{\phi}_s\big\|_{L^2_\gamma(0,t,\mu_{t,\alpha};\mathscr{B}_\eta)} = \big\|s \mapsto \widetilde{\phi}_s\big\|_{L^2(0,t,\mu_{t,\alpha};\mathscr{B}_\eta)} + \big\|s \mapsto \widetilde{\phi}_s\big\|_{\gamma(L^2(0,t,\mu_{t,\alpha}),\mathscr{B}_\eta)}.$$

For the first part, using property (2.22) of Definition 2.20

$$\begin{aligned}\big\|s\mapsto \widetilde{\phi}_s\big\|^2_{L^2(0,t,\mu_{t,\alpha};\mathscr{B}_\eta)} &= \int_0^t \|\widetilde{\phi}_s\|^2_{\mathscr{B}_\eta}\,\mathrm{d}\mu_{t,\alpha}(s)\\ &\le \int_0^t \Big(K_T^2\|\phi\|^2_{C([0,T];E_\eta)} + M_T^2\|\widetilde{\phi}_0\|^2_{\mathscr{B}_\eta}\Big)\,\mathrm{d}\mu_{t,\alpha}(s)\\ &\le C\Big(\|\phi\|^2_{C([0,T];E_\eta)} + \|(X,\Phi)\|^2_{\mathscr{B}_\eta}\Big) \end{aligned}\tag{4.41}$$

where $K_T = \sup_{s\in[0,T]} K(s)$, $M_T = \sup_{s\in[0,T]} M(s)$. By the corollary to the γ-Fubini isomorphism (Corollary 3.25), since $\mathscr{B}_\eta = L^p_g\big((-\infty,0];E_\eta\big)\times E_\eta$ for some g we have that

$$\begin{aligned}\gamma\Big(L^2\big(0,t,\mu_{t,\alpha}\big),\mathscr{B}_\eta\Big) &= \gamma\Big(L^2\big(0,t,\mu_{t,\alpha}\big),L^p\big(-\infty,0,g(r)\,\mathrm{d}r;E_\eta\big)\times E_\eta\Big)\\ &\simeq L^p_g\Big((-\infty,0];\gamma\big(L^2(0,t,\mu_{t,\alpha}),E_\eta\big)\Big)\\ &\qquad\times\gamma\big(L^2(0,t,\mu_{t,\alpha}),E_\eta\big).\end{aligned}\tag{4.42}$$

We now apply this isomorphism to the map $s\mapsto\widetilde{\phi}_s$

$$\begin{aligned}&\big\|s\mapsto\widetilde{\phi}_s\big\|^p_{\gamma(L^2(0,t,\mu_{t,\alpha}),\mathscr{B}_\eta)}\\ &\quad\le C\big\|s\mapsto\widetilde{\phi}_s\big\|^p_{L^p_g((-\infty,0];\gamma(L^2(0,t,\mu_{t,\alpha}),E_\eta))\times\gamma(L^2(0,t,\mu_{t,\alpha}),E_\eta)}\\ &\quad= C\Big(\big\|s\mapsto\widetilde{\phi}_s\big\|^p_{L^p_g((-\infty,0];\gamma(L^2(0,t,\mu_{t,\alpha}),E_\eta))} + \big\|s\mapsto\widetilde{\phi}_s(0)\big\|^p_{\gamma(L^2(0,t,\mu_{t,\alpha}),E_\eta)}\Big)\\ &\quad= C\big(I_1+I_2\big).\end{aligned}$$

To deal with the L^p_g term I_1 we partition the domain $(0,t)$ of $\widetilde{\phi}_s$ so as to estimate the initial history part Φ and the part in $V^p_{\alpha,\infty}\big([0,T]\times\Omega;E_\eta\big)$ separately.

$$\begin{aligned}I_1 &= C\int_{-\infty}^0 g(r)\big\|s\mapsto(t-s)^{-\alpha}\widetilde{\phi}_s(r)\big\|^p_{\gamma(L^2(0,t),E_\eta)}\,\mathrm{d}r\\ &= C\int_{-\infty}^0 g(r)\big\|s\mapsto(t-s)^{-\alpha}\widetilde{\phi}(s+r)\big\|^p_{\gamma(L^2(0,t),E_\eta)}\,\mathrm{d}r\\ &= C\int_{-\infty}^0 g(r)\big\|u\mapsto(t+r-u)^{-\alpha}\widetilde{\phi}(u)\big\|^p_{\gamma(L^2(r,t+r),E_\eta)}\,\mathrm{d}r\\ &= C\int_{-\infty}^0 g(r)\Big\|u\mapsto(t+r-u)^{-\alpha}\Big[\mathbb{1}_{[r,0\wedge(t+r)]}(u)\Phi(u)\\ &\qquad\qquad + \mathbb{1}_{(0\wedge(t+r),t+r]}(u)\phi(u)\Big]\Big\|^p_{\gamma(L^2(r,t+r),E_\eta)}\,\mathrm{d}r\\ &\le C\int_{-\infty}^0 g(r)\Bigg[\big\|u\mapsto(t+r-u)^{-\alpha}\Phi(u)\big\|^p_{\gamma(L^2(r,0\wedge(t+r)),E_\eta)}\\ &\qquad\qquad + \big\|u\mapsto(t+r-u)^{-\alpha}\phi(u)\big\|^p_{\gamma(L^2(0\wedge(t+r),t+r),E_\eta)}\Bigg]\,\mathrm{d}r.\end{aligned}$$

Now the second term in this integral is 0 for $r < -t$, so

$$\leq C\bigg[\int_{-\infty}^{0} g(r)\big\|u \mapsto (t+r-u)^{-\alpha}\widehat{\Phi}(u)\big\|^p_{\gamma(L^2(r,t+r),E_\eta)}\,\mathrm{d}r + \int_{-t}^{0} g(r)\big\|u \mapsto (t+r-u)^{-\alpha}\phi(u)\big\|^p_{\gamma(L^2(0,t+r),E_\eta)}\,\mathrm{d}r\bigg]$$

$$\leq C\bigg[\int_{-\infty}^{0} g(r)\big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_s(r)\big\|^p_{\gamma(L^2(0,t),E_\eta)}\,\mathrm{d}r + \sup_{\substack{t\in[0,T]\\ r\in[-t,0]}}\big\|u \mapsto (t+r-u)^{-\alpha}\phi(u)\big\|^p_{\gamma(L^2(0,t+r),E_\eta)}\int_{-t}^{0} g(r)\,\mathrm{d}r\bigg]$$

$$\lesssim C\bigg[\sup_{t\in[0,T]}\big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_s\big\|^p_{\gamma(L^2(0,t),\mathscr{B}_\eta)} + \sup_{t\in[0,T]}\big\|s \mapsto (t-s)^{-\alpha}\phi(s)\big\|^p_{\gamma(L^2(0,t),E_\eta)}\int_{-t}^{0} g(r)\,\mathrm{d}r\bigg], \tag{4.43}$$

and since $\phi \in V^p_{\alpha,\infty}\big([0,T]\times\Omega;E_\eta\big)$ and Φ satisfies (4.23), this is integrable with finite expectation. For I_2

$$\begin{aligned} I_2 = \big\|s \mapsto \widetilde{\phi}_s(0)\big\|^p_{\gamma(L^2(0,t,\mu_{t,\alpha}),E_\eta)} &= \big\|s \mapsto \phi(s)\big\|^p_{\gamma(L^2(0,t,\mu_{t,\alpha}),E_\eta)} \\ &\leq \sup_{t\in[0,T]}\big\|s \mapsto (t-s)^{-\alpha}\phi(s)\big\|^p_{\gamma(L^2(0,t),E_\eta)} \end{aligned} \tag{4.44}$$

which has finite expectation for the same reason. Hence $s \mapsto \widetilde{\phi}_s$ and $s \mapsto \widetilde{\psi}_s$ are contained in $L^2_\gamma\big(0,t,\mu_{t,\alpha};\mathscr{B}_\eta\big)$ almost surely and

$$\big\|s \mapsto \widetilde{\phi}_s\big\|_{L^2_\gamma(0,t,\mu_{t,\alpha};\mathscr{B}_\eta)} \lesssim 1 + \|\phi\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)}. \tag{4.45}$$

By (D5), for fixed $\omega \in \Omega$, $t \in [0,T]$ we have

$$\begin{aligned} \big\|s \mapsto (t-s)^{-\alpha}G(s,\widetilde{\phi}_s)\big\|_{\gamma(L^2(0,T;H),E_{-\theta_G})} &\overset{(4.21)}{\lesssim} 1 + \big\|s \mapsto (t-s)^{-\alpha}\widetilde{\phi}_s\big\|_{L^2_\gamma(0,T;\mathscr{B}_\eta)} \\ &\overset{(4.45)}{\lesssim} 1 + \|\phi\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \end{aligned}$$

and hence the function G is of linear growth from the space $L^2_\gamma\big(0,t,\mu_{t,\alpha};\mathscr{B}_\eta\big)$ into $\gamma(L^2(0,t,\mu_{t,\alpha};H),E_{-\theta_G}\big)$ for all $t \in [0,T]$. Now $\Psi(s) := G(s,\widetilde{\phi}_s)$ satisfies (4.33) and we can apply Lemma 4.16. Thus as $G(s,\widetilde{\phi}_s)$ and $G(s,\widetilde{\psi}_s)$ are H-strongly measurable and adapted, $S \diamond G(\cdot,\widetilde{\phi}.)$ and $S \diamond G(\cdot,\widetilde{\psi}.)$ are in $V^p_{\alpha,\infty}\big([0,T]\times\Omega;E_\eta\big)$ by (4.34).

Now using the L^2_γ-Lipschitz property of G (4.20)

$$
\begin{aligned}
&\big\|S \diamond \big(G(\cdot,\widetilde{\phi}.) - G(\cdot,\widetilde{\psi}.)\big)\big\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \\
&\overset{(4.34)}{\lesssim} T^{\varepsilon'} \sup_{t\in[0,T]} \Big(\mathbb{E}\big\|s\mapsto (t-s)^{-\alpha}\big[G(s,\widetilde{\phi}_s) - G(s,\widetilde{\psi}_s)\big]\big\|^p_{\gamma(L^2(0,t;H);E_{-\theta_G})}\Big)^{\frac{1}{p}} \\
&= T^{\varepsilon'} \sup_{t\in[0,T]} \Big(\mathbb{E}\big\|s\mapsto G(s,\widetilde{\phi}_s) - G(s,\widetilde{\psi}_s)\big\|^p_{\gamma(L^2(0,t,\mu_{t,\alpha};H);E_{-\theta_G})}\Big)^{\frac{1}{p}} \\
&\overset{(4.20)}{\lesssim} L^\gamma_G T^{\varepsilon'} \sup_{t\in[0,T]} \Big(\mathbb{E}\big\|s\mapsto (\widetilde{\phi}-\widetilde{\psi})_s\big\|^p_{L^2_\gamma(0,t,\mu_{t,\alpha};\mathscr{B}_\eta)}\Big)^{\frac{1}{p}} \\
&\lesssim L^\gamma_G T^{\varepsilon'}\Bigg[\sup_{t\in[0,T]} \Big(\mathbb{E}\big\|s\mapsto (t-s)^{-\alpha}(\widetilde{\phi}-\widetilde{\psi})_s\big\|^p_{\gamma(L^2(0,t),\mathscr{B}_\eta)}\Big)^{\frac{1}{p}} \qquad (4.46)\\
&\qquad\qquad + \sup_{t\in[0,T]} \Big(\mathbb{E}\big\|s\mapsto (t-s)^{-\alpha}(\widetilde{\phi}-\widetilde{\psi})_s\big\|^p_{L^2(0,t;\mathscr{B}_\eta)}\Big)^{\frac{1}{p}}\Bigg].
\end{aligned}
$$

Using property (2.22) of Definition 2.20 and the fact that $\widetilde{\phi}_0 = \widetilde{\psi}_0 = \Phi$ almost surely, we have that for almost all $\omega\in\Omega$

$$
\begin{aligned}
\big\|s\mapsto (t-s)^{-\alpha}(\widetilde{\phi}-\widetilde{\psi})_s\big\|^p_{L^2(0,t;\mathscr{B}_\eta)} &\le T^{(\frac{p}{2}-\alpha p)}\big\|t\mapsto (\widetilde{\phi}-\widetilde{\psi})_t\big\|^p_{C([0,T];\mathscr{B}_\eta)} \\
&\le K^p_T T^{(\frac{p}{2}-\alpha p)}\big\|\phi-\psi\big\|^p_{C([0,T];E_\eta)} \qquad (4.47)
\end{aligned}
$$

and by a further use of Corollary 3.25, the isomorphism

$$
\begin{aligned}
\gamma\Big(L^2(0,t,\mu_{t,\alpha}),\mathscr{B}_\eta\Big) \simeq L^p_g\Big((-\infty,0];\gamma\big(L^2(0,t,\mu_{t,\alpha}),E_\eta\big)\Big)& \\
\times\gamma\big(L^2(0,t,\mu_{t,\alpha}),E_\eta\big)& \qquad (4.48)
\end{aligned}
$$

gives us

$$
\begin{aligned}
&\sup_{t\in[0,T]}\big\|s\mapsto (t-s)^{-\alpha}(\widetilde{\phi}-\widetilde{\psi})_s\big\|^p_{\gamma(L^2(0,t),\mathscr{B}_\eta)} \\
&\quad\lesssim \sup_{t\in[0,T]}\int_{-\infty}^0 g(r)\big\|s\mapsto (t-s)^{-\alpha}(\widetilde{\phi}-\widetilde{\psi})(s+r)\big\|^p_{\gamma(L^2(0,t),E_\eta)}\,\mathrm{d}r \\
&\qquad\qquad + \sup_{t\in[0,T]}\big\|s\mapsto (t-s)^{-\alpha}(\phi-\psi)(s)\big\|^p_{\gamma(L^2(0,t),E_\eta)} \\
&\quad\le \big(1+\|g\|_{L^1(-T,0)}\big)\sup_{\substack{t\in[0,T]\\ r\in[-t,0]}}\big\|s\mapsto (t-s)^{-\alpha}(\widetilde{\phi}-\widetilde{\psi})(s+r)\big\|^p_{\gamma(L^2(0,t),E_\eta)} \\
&\quad= \big(1+\|g\|_{L^1(-T,0)}\big)\sup_{\substack{t\in[0,T]\\ r\in[-t,0]}}\big\|u\mapsto (t+r-u)^{-\alpha}(\phi-\psi)(u)\big\|^p_{\gamma(L^2(0,t+r),E_\eta)} \\
&\quad= \big(1+\|g\|_{L^1(-T,0)}\big)\sup_{t\in[0,T]}\big\|s\mapsto (t-s)^{-\alpha}(\phi-\psi)(s)\big\|^p_{\gamma(L^2(0,t),E_\eta)}. \qquad (4.49)
\end{aligned}
$$

Taking expectations and combining (4.46) with (4.47) and (4.49), there exists $C \geq 0$ such that

$$\big\| S \diamond \big(G(\cdot, \widetilde{\phi}.) - G(\cdot, \widetilde{\psi}.) \big) \big\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \leq CT^{\varepsilon'} \|\phi - \psi\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \tag{4.50}$$

as required.

Step 4 (Collecting the estimates).

From the estimates (4.37), (4.40) and (4.50) we see that L_T is well defined on $V^p_{\alpha,\infty}\big([0,T] \times \Omega; E_\eta\big)$ and that there exist constants $C \geq 0$ and $\beta > 0$ such that for all $\phi, \psi \in V^p_{\alpha,\infty}\big([0,T] \times \Omega; E_\eta\big)$ we have

$$\big\| L_T(\phi) - L_T(\psi) \|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \leq CT^{\beta} \|\phi - \psi\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)}. \tag{4.51}$$

Now recall the definition of $\widehat{\Phi}$ (4.22) as the extension of Φ by 0 to $(-\infty, T_0]$

$$\begin{aligned}
\|S(\cdot)X\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} &\overset{(4.37)}{\leq} C\big(\mathbb{E}\|X\|^p_{E_\eta}\big)^{\frac{1}{p}} \\
&\overset{(2.22)}{\leq} CH\big(\mathbb{E}\|(X,\Phi)\|^p_{\mathscr{B}_\eta}\big)^{\frac{1}{p}} \\
\big\|S * F(\cdot, \widehat{\Phi}.)\big\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} &\overset{(4.30)}{\leq} CT^{\beta}\Big(\mathbb{E}\big\|F(\cdot,\widehat{\Phi}.)\big\|^p_{C([0,T];E_\eta)}\Big)^{\frac{1}{p}} \\
&\overset{(4.19)}{\leq} C_F CT^{\beta}\Big(1 + \mathbb{E}\big\|(X,\Phi)\big\|^p_{\mathscr{B}_\eta)}\Big)^{\frac{1}{p}}.
\end{aligned}$$

By (4.23), $G(\cdot, \widehat{\Phi}.)$ satisfies (4.33), and hence by (4.34)

$$\begin{aligned}
&\big\|S \diamond G(\cdot, \widehat{\Phi}.)\big\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \\
&\qquad \leq CT^{\beta}\Big(\sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha} G(s, \widehat{\Phi}_s)\big\|^p_{\gamma(L^2(0,t;H),\mathscr{B}_\eta)}\Big)^{\frac{1}{p}} \\
&\qquad \overset{(4.21)}{\leq} CT^{\beta}C^{\gamma}_G\Big(1 + \sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_s\big\|^p_{L^2_\gamma(0,t;\mathscr{B}_\eta)}\Big)^{\frac{1}{p}}
\end{aligned}$$

Combining the above gives

$$\begin{aligned}
\big\|L_T(0)\big\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} &\leq C\Big(1 + \big(\mathbb{E}\|(X,\Phi)\|^p_{\mathscr{B}_\eta}\big)^{\frac{1}{p}} \\
&\qquad + \big(\sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_s\big\|^p_{L^2_\gamma(0,t;\mathscr{B}_\eta)}\big)^{\frac{1}{p}}\Big)
\end{aligned}$$

which together with (4.51) yields the result. □

Theorem 4.17 (Existence and Uniqueness)**.** *Suppose (D1) – (D6) are satisfied and choose* $\alpha \in (0, \frac{1}{2})$ *such that*

$$\eta + \theta_G < \alpha - \frac{1}{p}.$$

There exists a mild solution U *in* $V^p_{\alpha,\infty}\big([0,T_0] \times \Omega; E_\eta\big)$ *of* (4.15)*. As a mild solution in* $V^p_{\alpha,p}\big([0,T] \times \Omega; E_\eta\big)$*, this solution* U *is unique. Moreover there exists a constant* $C \geq 0$ *independent of* (x, Φ) *such that*

$$\begin{aligned}\|U\|_{V^p_{\alpha,\infty}([0,T_0]\times\Omega;E_\eta)} \leq & C\Big(1 + \big(\mathbb{E}\|(X,\Phi)\|^p_{\mathscr{B}_\eta}\big)^{\frac{1}{p}} \\ & + \big(\sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_s\big\|^p_{L^2_\gamma(0,t;\mathscr{B}_\eta)}\big)^{\frac{1}{p}}\Big). \end{aligned} \tag{4.52}$$

Proof. Again, we follow the method used in Theorem 6.3 of [69]. By Proposition 4.14 we can find $T \in (0, T_0]$, independent of (X, Φ), such that $C_T < \frac{1}{2}$. It follows from (4.28) and the Banach fixed point theorem that L_T has a unique fixed point $U \in V^p_{\alpha,\infty}\big([0,T]\times\Omega; E_\eta\big)$. This gives a continuous adapted process $U : [0,T]\times\Omega \to E_\eta$ such that almost surely for all $t \in [0,T]$

$$U(t) = S(t)X + S * F(\cdot, \widetilde{U}_\cdot)(t) + S \diamond G(\cdot, \widetilde{U}_\cdot)(t).$$

Noting that $U = \lim_{n\to\infty} L^n_T(0)$ in $V^p_{\alpha,\infty}\big([0,T] \times \Omega; E_\eta\big)$, (4.29) implies the inequality

$$\begin{aligned}\|U\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \leq & C\Big(1 + \big(\mathbb{E}\|(X,\Phi)\|^p_{\mathscr{B}_\eta}\big)^{\frac{1}{p}} \\ & + \big(\sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_s\big\|^p_{L^2_\gamma(0,t;\mathscr{B}_\eta)}\big)^{\frac{1}{p}}\Big) \\ & + C_T\|U\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)}, \end{aligned}$$

and then $C_T < \frac{1}{2}$ implies

$$\begin{aligned}\|U\|_{V^p_{\alpha,\infty}([0,T]\times\Omega;E_\eta)} \leq & C\Big(1 + \big(\mathbb{E}\|(X,\Phi)\|^p_{\mathscr{B}_\eta}\big)^{\frac{1}{p}} \\ & + \big(\sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_s\big\|^p_{L^2_\gamma(0,t;\mathscr{B}_\eta)}\big)^{\frac{1}{p}}\Big). \end{aligned} \tag{4.53}$$

To find a solution on all of $[0, T_0]$ we show that U_T satisfies (D6) and hence a solution exists on $[T, 2T]$. Then by induction we can construct mild solutions on each of the intervals $[T, 2T], \ldots, [nT, T_0]$ for some n. The induced solution U on $[0, T_0]$ is the desired mild solution of (4.15) and is unique. Moreover, by (4.53) and induction we deduce (4.52).

Now $U_T : \Omega \to \mathscr{B}_\eta$ is strongly $\mathscr{F}_T$-measurable, $U(T) \in L^p(\Omega, E_\eta)$ and setting

$$\widehat{U}(t) := \begin{cases} \Phi(t) & t \in (-\infty, 0] \\ U(t) & t \in (0, T] \\ 0 & t \in (T, T_0] \end{cases} \tag{4.54}$$

and $T_1 = T_0 - T$ we have

$$\begin{aligned}
&\sup_{t\in[0,T_1]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\widehat{U}_{T+s}\big\|_{\gamma(L^2(0,t;H),\mathscr{B}_\eta)} \\
&\quad\overset{(4.42)}{\lesssim} \sup_{t\in[0,T_1]} \mathbb{E}\int_{-\infty}^{0} g(r)\big\|s \mapsto (t-s)^{-\alpha}\widehat{U}(T+s+r)\big\|_{\gamma(L^2(0,t;H),E_\eta)}\,\mathrm{d}r \\
&\qquad\qquad + \sup_{t\in[0,T_1]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\widehat{U}(T+s)\big\|_{\gamma(L^2(0,t;H),E_\eta)} \\
&\quad= \sup_{t\in[0,T_1]} \mathbb{E}\big(I_1 + I_2\big)
\end{aligned}$$

Now $I_2 = 0$ by the definition of $\widehat{U}$ and by partitioning the range $(0,t)$ as above,

$$\begin{aligned}
I_1 &= \int_{-\infty}^{0} g(r)\big\|w \mapsto (t+T+r-w)^{-\alpha}\widehat{U}(w)\big\|_{\gamma(L^2(T+r,t+T+r;H),E_\eta)}\,\mathrm{d}r \\
&\le \int_{-\infty}^{0} g(r)\Big\|w \mapsto (t+T+r-w)^{-\alpha}\Big[\mathbb{1}_{(-\infty,0]}\Phi(w) \\
&\qquad\qquad + \mathbb{1}_{(0,T]}(w)U(w)\Big]\Big\|_{\gamma(L^2(T+r,t+T+r;H),E_\eta)}\,\mathrm{d}r \\
&\le \int_{-\infty}^{0} g(r)\Big[\big\|w \mapsto (t+T+r-w)^{-\alpha}\Phi(w)\big\|_{\gamma(L^2(T+r,(t+T+r)\wedge 0;H),E_\eta)} \\
&\quad + \big\|w \mapsto (t+T+r-w)^{-\alpha}U(w)\big\|_{\gamma(L^2((t+T+r)\wedge 0,(t+T+r)\wedge T;H),E_\eta)}\Big]\,\mathrm{d}r \\
&\le \int_{-\infty}^{0} g(r)\big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_{T+s}(r)\big\|_{\gamma(L^2(0,t;H),E_\eta)}\,\mathrm{d}r \\
&\quad + \int_{-(t+T)}^{0} g(r)\big\|w \mapsto (t+T+r-w)^{-\alpha}U(w)\big\|_{\gamma(L^2((t+T+r)\wedge 0,(t+T+r)\wedge T;H),E_\eta)}\,\mathrm{d}r
\end{aligned}$$

$$\begin{aligned}
&\overset{(4.42)}{\lesssim} \Big[\big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_{T+s}\big\|_{\gamma(L^2(0,t;H),\mathscr{B}_\eta)} \\
&\quad + \sup_{r\in[-(t+T),0]} \big\|w \mapsto (t+T+r-w)^{-\alpha}U(w)\big\|_{\gamma(L^2(0,(t+T+r)\wedge T;H),E_\eta)}\Big]\int_{-(t+T)}^{0} g(r)\,\mathrm{d}r \\
&\lesssim \big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_{T+s}\big\|_{\gamma(L^2(0,t;H),\mathscr{B}_\eta)} + \big\|s \mapsto (t-s)^{-\alpha}U(s)\big\|_{\gamma(L^2(0,t;h),E_\eta)}
\end{aligned}$$

so $\sup_{t\in[0,T_1]}\mathbb{E}(I_1)$ is finite as $U \in V^p_{\alpha,\infty}\big([0,T]\times\Omega;E_\eta\big)$ and U_T satisfies (D6). The result follows. $\square$

Space and time regularity of the solution follows very quickly from the above, as Hölder regularity of the convolutions was already shown in Lemma 4.10 and Proposition 4.11, which up until now have not been used to their full extent.

Theorem 4.18 (Regularity). *Let E be a UMD space with type $\tau \in [1,2]$ and suppose that (D1) - (D6) hold. Let $\lambda \geq 0$ and $\delta \geq \eta$ satisfy $\lambda + \delta < \min\{\frac{1}{2} - \frac{1}{p} - \theta_G, 1 - \theta_F\}$ then there exists $C \geq 0$ such that*

$$\Big(\mathbb{E}\big\|U - SX\|^p_{C^\lambda([0,T_0];E_\delta)}\Big)^{\frac{1}{p}} \leq C\Big(1 + \big(\mathbb{E}\|(X,\Phi)\|^p_{\mathscr{B}_\eta}\big)^{\frac{1}{p}} + \big(\sup_{t\in[0,T]} \mathbb{E}\big\|s \mapsto (t-s)^{-\alpha}\widehat{\Phi}_s\big\|^p_{L^2_\gamma(0,t;\mathscr{B}_\eta)}\big)^{\frac{1}{p}}\Big). \tag{4.55}$$

Proof. Choose $r \geq 1$ such that $\lambda + \delta < 1 - \theta_F - \frac{1}{r}$ and $\alpha \in (0, \frac{1}{2})$ with

$$\eta + \theta_G < \alpha - \frac{1}{p} \quad \text{and} \quad \lambda + \delta + \theta_G < \alpha - \frac{1}{p}.$$

Let $W \in V^p_{\alpha,\infty}([0,T]\times\Omega; E_\eta)$ be the mild solution from Theorem 4.17. It follows from Lemma 4.10 (with $\alpha = 1$) that we may take a version of $S * F(\cdot, \widetilde{W_\cdot})$ with

$$\begin{aligned}\mathbb{E}\big\|S * F(\cdot,\widetilde{W_\cdot})\big\|^p_{C^\lambda([0,T_0];E_\delta)} &\leq C\mathbb{E}\big\|F(\cdot,\widetilde{W_\cdot})\big\|^p_{L^r(0,T_0;E_{-\theta_F})}\\ &\leq C\mathbb{E}\big\|F(\cdot,\widetilde{W_\cdot})\big\|^p_{C([0,T_0];E_{-\theta_F})}\end{aligned}$$

Similarly, by Proposition 4.11 we may take a version of $S \diamond G(\cdot, \widetilde{W_\cdot})$ with

$$\mathbb{E}\big\|S \diamond G(\cdot,\widetilde{W_\cdot})\big\|^p_{C^\lambda([0,T_0];E_\delta)} \leq C \sup_{t\in[0,T_0]} \mathbb{E}\big\|(t-\cdot)^{-\alpha} G(\cdot,\widetilde{W_\cdot})\big\|^p_{\gamma(L^2(0,t;H),E_{-\theta_G})}. \tag{4.56}$$

Now we define $U : [0,T_0]\times\Omega \to E_\eta$ by

$$U(t) = S(t)X + S * F(\cdot,\widetilde{W_\cdot})(t) + S \diamond G(\cdot,\widetilde{W_\cdot}),$$

where we take the versions of the convolutions as above ☐

4.4 Example

Let $\mathcal{O}$ be an open and bounded domain of $\mathbb{R}^d$ with smooth boundary and consider the following perturbed heat equation with memory, equipped with Dirichlet boundary conditions

$$\begin{aligned}
\frac{\partial}{\partial t}u(t,\xi) &= \Delta u(t,\xi) + f\big(t,\xi,u_t(\cdot,\xi)\big) + \sum_{n=1}^{\infty} g_n\big(t,\xi,u_t(\cdot,\xi)\big)\partial W_n(t), \\
&\qquad\qquad t\in[0,T],\ \xi\in\mathcal{O}, \qquad (4.57)\\
u(t,\xi) &= 0, \quad t\in[0,T],\ \xi\in\partial\mathcal{O}, \\
u_0(t,\xi) &= \Phi(t,\xi), \quad t\in(-\infty,0],\ \xi\in\mathcal{O}, \\
u(0,\xi) &= \Psi(\xi), \quad \xi\in\mathcal{O}.
\end{aligned}$$

We take $E := L^q(\mathcal{O})$ for $1 < q < \infty$. It is well known that the Dirichlet Laplacian Δ with $D(\Delta) := \{g \in W^{2,q}(\mathcal{O}) : g|_{\partial\mathcal{O}} = 0\}$ generates a uniformly stable analytic C_0-semigroup on $E = L^q(\mathcal{O})$. $(W_n)_{n\geq 1}$ will be a sequence of independent standard real Brownian motions on Ω. We say that $u : [0,T]\times\Omega\times\mathcal{O}\to\mathbb{R}$ is a solution of (4.57) if the corresponding functional analytic model (4.15) has a mild solution U with $U(t,\omega)(\xi) = u(t,\omega,\xi)$.

Fix $p > 2$. We assume the functions $f, g_n : [0,T]\times\Omega\times\mathcal{O}\times L^p(-\infty,0;E)\to\mathbb{R}$ are jointly measurable and adapted, and that there exist constants $L_f, L_{g_n} \geq 0$ such that for all $t\in[0,T]$, $\omega\in\Omega$, $\xi\in\mathcal{O}$ and all functions $\phi,\psi : (-\infty,T]\to E$, that are continuous on $[0,T]$ with $\phi_0,\psi_0\in L^p(-\infty,0;E)$ we have

$$\|f(t,\omega,\cdot,\phi_0) - f(t,\omega,\cdot,\psi_0)\|_{L^q(\mathcal{O})} \leq L_f\|\phi_0-\psi_0\|_{L^p(-\infty,0;L^q(\mathcal{O}))} \qquad (4.58)$$

and

$$\int_{\mathcal{O}}\left(\sum_{n=1}^{\infty}|g_n(t,\omega,\xi,\phi_t) - g_n(t,\omega,\xi,\psi_t)|^2\right)^{\frac{q}{2}} \mathrm{d}\xi \leq L\big\|(\phi-\psi)_t\big\|^q_{L^p(-\infty,0;L^q(\mathcal{O}))}. \qquad (4.59)$$

Assume also that

$$\sup_{t\in[0,T]}\big\|f(t,\omega,\cdot,0)\big\|_{L^\infty(\Omega;L^q(\mathcal{O}))} < \infty \qquad (4.60)$$

and that

$$\Big\|\Big(\int_0^T\sum_{n\geq 1}\big|g_n(t,\omega,\cdot,0)\big|^2\,\mathrm{d}\mu(t)\Big)^{\frac{1}{2}}\Big\|_{L^\infty(\Omega;L^q(\mathcal{O}))} < \infty \qquad (4.61)$$

for all finite measures μ on $[0,T]$.

Theorem 4.19. *Assume that all the above holds, that $\Psi\in L^q(\mathcal{O})$ and if $\widehat{\Phi}$ is the extension of the initial data $\Phi : (-\infty,0]\to L^q(\mathcal{O})$ to $(-\infty,T]$ by 0, that the map from $[0,T]$ to $L^p(-\infty,0;L^q(\mathcal{O}))$ given by $s\mapsto\widehat{\Phi}_s$ lies in the mixed L^p-L^q-L^2 space*

$$L^p\big(-\infty,0;L^q\big(\mathcal{O};L^2(0,t,\mu_{\alpha,t})\big)\big)\cap L^2\big(0,t,\mu_{\alpha,t};L^p\big(-\infty,0;L^q(\mathcal{O})\big)\big),$$

uniformly for $t \in [0,T]$ for some $\alpha \in (\frac{1}{p}, \frac{1}{2})$, i.e. both the following are finite and uniformly bounded for $t \in [0,T]$,

$$\int_{-\infty}^{0} \left(\int_{\mathcal{O}} \left(\int_{0}^{t} |(t-s)^{-\alpha} \widehat{\Phi}(r+s,\xi)|^2 \,\mathrm{d}s \right)^{\frac{q}{2}} \mathrm{d}\xi \right)^{\frac{p}{q}} \mathrm{d}r,$$

$$\int_{0}^{t} \left(\int_{-\infty}^{0} \left(\int_{\mathcal{O}} |(t-s)^{-\alpha} \widehat{\Phi}(r+s,\xi)|^q \,\mathrm{d}\xi \right)^{\frac{p}{q}} \mathrm{d}r \right)^{\frac{2}{p}} \mathrm{d}s.$$

Then for the problem (4.57) *has a mild solution $U \in V^p_{\alpha,\infty}\big([0,T] \times \Omega; L^q(\mathcal{O})\big)$ which is unique in $V^p_{\alpha,p}\big([0,T] \times \Omega; L^q(\mathcal{O})\big)$*

Proof. We check the conditions of Theorem 4.17.

(D1) As noted above, the Dirichlet Laplacian $(\Delta, D(\Delta))$ generates a uniformly exponentially stable analytic C_0-semigroup on E. Let $H = \ell^2$ with the standard unit basis (e_n), then setting $W(t)e_n = W_n(t)$, W becomes an H-cylindrical Brownian motion.

(D2) In this example $\eta = \theta_F = \theta_G = 0$.

(D3) Here $\mathscr{B} = L^p\big(-\infty, 0; L^q(\mathcal{O})\big) \times L^q(\mathcal{O})$.

(D4) Follows from (4.58) and (4.60) by defining $F : [0,T] \times \Omega \times \mathscr{B} \to E$ by

$$F(t,\omega,\phi)(\xi) = f\big(t,\omega,\xi,\phi\big).$$

Then for all $\omega \in \Omega$, $t \in [0,T]$ and $\phi, \psi \in L^p(-\infty,0;L^q(\mathcal{O}))$

$$\|F(t,\omega,\phi) - F(t,\omega,\psi)\|^q_{L^q(\mathcal{O})} \le L_f \|\phi - \psi\|_{L^p(-\infty,0;L^q(\mathcal{O}))}$$

so (4.18) holds, whereas the linear growth condition and (4.19) follows from (4.60).

(D5) Define $G : [0,T] \times \Omega \times \mathscr{B} \to \mathcal{L}(H,E)$ by

$$\big(G(t,\omega,\phi)e_n\big)(\xi) = g_n\big(t,\omega,\xi,\phi\big).$$

Let $\phi, \psi : (-\infty, T] \times S \to E$ be continuous on $[0,T]$ and suppose $(t \mapsto \phi_t)$ and $(t \mapsto \psi_t)$ lie in

$$L^p\big(-\infty,0;L^q\big(\mathcal{O};L^2(0,T,\mu)\big)\big) \cap L^2\big(0,T,\mu;L^p\big(-\infty,0;L^q(\mathcal{O})\big)\big)$$

for all finite measures μ on $[0,T]$. By using the γ-Fubini isomorphism, Proposition 3.24, twice we have $(t \mapsto \phi_t)$, $(t \mapsto \psi_t) \in L^2_\gamma(0,T,\mu;\mathscr{B})$. Now by (4.59)

$$\begin{aligned}
&\big\|G(\cdot,\phi_\cdot)-G(\cdot,\psi_\cdot)\big\|^q_{\gamma(L^2(0,T,\mu;\ell^2),E)} \\
&\qquad \simeq_q \big\|G(\cdot,\phi_\cdot)-G(\cdot,\psi_\cdot)\big\|^q_{L^q(\mathcal{O};L^2(0,T,\mu;\ell^2))} \\
&\qquad = \int_{\mathcal{O}} \Big\|\Big(g_n(\cdot,\xi,\phi_\cdot)-g_n(\cdot,\xi,\psi_\cdot)\Big)_{n\geq 1}\Big\|^q_{L^2(0,T,\mu;\ell^2)} \,\mathrm{d}\xi \\
&\qquad = \int_{\mathcal{O}} \left(\sum_{n=1}^{\infty} \big\|g_n(\cdot,\xi,\phi_\cdot)-g_n(\cdot,\xi,\psi_\cdot)\big\|^2_{L^2(0,T,\mu)}\right)^{\frac{q}{2}} \mathrm{d}\xi \\
&\qquad \leq L\big\|t \mapsto (\phi-\psi)_t\big\|^q_{L^2_\gamma(0,T,\mu;L^p(-\infty,0;L^q(\mathcal{O})))}.
\end{aligned}$$

This gives us the Lipschitz part of (D5), and the linear growth part follows from (4.61).

(D6) $\Psi \in L^q(\mathcal{O})$, and by assumption

$$\begin{aligned}
&\sup_{t\in[0,T]} \big\|r \mapsto (t-r)^{-\alpha}\widehat{\Phi}_r(\cdot,*)\big\|_{L^2_\gamma(0,t;L^p(-\infty,0;L^q(\mathcal{O})))} \\
&\qquad \leq \sup_{t\in[0,T]} \big\|r \mapsto (t-r)^{-\alpha}\widehat{\Phi}_r(\cdot,*)\big\|_{L^2(0,t;L^p(-\infty,0;L^q(\mathcal{O})))} \\
&\qquad\qquad + \sup_{t\in[0,T]} \big\|r \mapsto (t-r)^{-\alpha}\widehat{\Phi}_r(\cdot,*)\big\|_{L^p(-\infty,0;L^q(\mathcal{O};L^2(0,t)))} \\
&\qquad < \infty.
\end{aligned}$$

The result now follows from Theorem 4.17. □

Chapter 5

Almost periodic solutions to stochastic problems

5.1 Introduction

In this chapter we turn to the problem of existence of *almost periodic* solutions to stochastic differential equations. Almost periodic functions were first considered by Bohr [18] in 1925, and loosely speaking are those functions from $\mathbb{R}$ to E which come arbitrarily close to being periodic when one looks over long enough time scales. Almost periodic solutions of differential equations have been an active area of research for much of the last century and have been studied in the context of Hilbert space valued stochastic differential equations, for example [31].

Since the solutions in question are random processes, we cannot expect any form of periodicity in the paths, that is, we cannot expect that a path $(X(\omega, t))_{t\in\mathbb{R}}$ is almost periodic for any given $\omega \in \Omega$ or even that almost periodicity will hold *in probability.* The correct notion to consider in this case turns out to be that of *almost periodicity in distribution*, see e.g. [85]. We will show the existence of solutions X where the probability $\mathbb{P}\{X(t) \in B\}$, is almost periodic for any $B \in \mathbb{B}_E$, or equivalently $\mathbb{E}(f(X(t)))$ is almost periodic for every bounded Lipschitz function f on E.

Typically in the literature, proofs of almost periodicity first show the existence of a bounded solution and then, under various conditions related either to the spectrum of the operator A or to stability of solutions under perturbations, show the existence of the almost periodic solution. The greatest difficulty is often in showing the existence of a bounded solution to the given problem, and indeed many papers simply assume this. In [64] a new technique is proposed, obtaining bounded solutions (to a deterministic evolution equation) by assuming that the semigroup decays 'sufficiently quickly' on a subspace on E. We will follow a similar approach, but we are forced to make the rather stronger assumption that the semigroup is exponentially stable in order to make use of results on R-boundedness of semigroups that are necessary to handle the stochastic convolution.

In Section 5.2.1 we make precise the informal notions above and give some background results. We then show in 5.3 the existence of almost periodic solutions to an autonomous stochastic differential equation with additive noise and then generalise this result in 5.4 to an SDE with non-autonomous unbounded term.

5.2 Further preliminaries

5.2.1 Almost periodicity in distribution

Definition 5.1. An E-valued process $X(t)$ on $(\Omega, \mathscr{F}, \mathbb{P})$ is said to be *L^p-bounded* if

$$\|X(t)\|^p_{L^p(\Omega;E)} = \mathbb{E}\|X(t)\|^p \leq M < \infty \tag{5.1}$$

for all $t \in \mathbb{R}$.

L^p-bounded processes are those with uniformly bounded p^{th} moment for $t \in \mathbb{R}$. As we are concerned with properties of the distribution of solutions over time, it will turn out to be enough to show L^p-boundedness of solutions as a stepping stone to almost periodicity.

Definition 5.2. A sequence (X_n) of random variables on a Banach space E is said to *converge in distribution* to a random variable X if $\mathbb{E}(f(X_n))$ converges to $\mathbb{E}(f(X))$ for every bounded Lipschitz continuous function $f : X \to \mathbb{R}$.

The following result, known as the Portmanteau Theorem (see [15, Section 2]), gives a number of conditions equivalent to convergence in distribution.

Lemma 5.3 (Portmanteau Theorem)**.** *Let X_n for $n \in \mathbb{N}$ and X be E-valued random variables on $(\Omega, \mathscr{F}, \mathbb{P})$. The following are equivalent:*

(1) X_n converges to X in distribution.

(2) $\mathbb{E}(f(X_n))$ converges to $\mathbb{E}(f(X))$ for all bounded, uniformly continuous $f : E \to \mathbb{R}$.

(3) if P_{X_n}, P_X *are the probability distributions of* X_n*,* X *on* E *respectively, then*

$$\int_E f(x)\,\mathrm{d}P_{X_n}(x) \to \int_E f(x)\,\mathrm{d}P_X(x) \text{ as } n \to \infty$$

for all bounded, uniformly continuous $f : E \to \mathbb{R}$.

(4) $\limsup_{n\to\infty} P_{X_n}(C) \leq P_X(C)$ *for all closed sets* $C \subset E$.

(5) $\liminf_{n\to\infty} P_{X_n}(U) \geq P_X(U)$ *for all open sets* $U \subset E$.

We can now define almost periodicity in distribution.

Definition 5.4. A random process $X(t)$ on E is said to be *almost periodic in distribution* if the set $\{X(t+\tau) : \tau \in \mathbb{R}\}$ of random variables on $C(\mathbb{R}, E)$ is *relatively compact in distribution*, that is, for any sequence $(\tau_n) \subset \mathbb{R}$ with $\tau_n \uparrow \infty$ there exists a subsequence (τ_{n_r}) such that $(X(t+\tau_{n_r}))$ converges in distribution uniformly for $t \in \mathbb{R}$.

$X(t)$ is *asymptotically* almost periodic in distribution if there exist continuous processes $Y(t)$ and $Z(t)$, $t \geq 0$ such that $Y(t)$ is almost periodic in distribution, $\|Z(t)\| \to 0$ in probability as $t \to \infty$ and $X(t) = Y(t) + Z(t)$ almost surely.

Note that if it is known that $X(t)$ is Gaussian for all $t \in \mathbb{R}$ then it is enough to check that the mean and covariance of $X(t+\tau_n)$ converge uniformly for $t \in \mathbb{R}$ in norm and strong operator topology respectively. See, for example, [73].

5.2.2 R-boundedness of semigroups

We will also need some results of Veraar and Hytönen [54] related to R-boundedness of semigroups. Recall the definition of R-boundedness from Section 3.5.

Proposition 5.5 (Corollary 5.5 of [54])**.** *Let* E, F *be Banach spaces such that* F *has type* $q \in [1,2]$*,* E *has co-type* $r \in [2,\infty]$ *and let* $I = (a,b)$ *with* $-\infty \leq a < b \leq \infty$*. Take* $p \in (1,\infty]$ *such that* $\frac{1}{p} \geq \frac{1}{q} - \frac{1}{r}$ *and let* $\alpha \in (\frac{1}{p}, 1)$*. If* $T \in L^p(I, \mathcal{L}(E,F))$ *and there exists a constant* M *such that*

$$\big\|T(s+h) - T(s)\big\| \leq M|h|^\alpha (1+|s|)^{-\alpha} \text{ for all } s, s+h \in I$$

then $\{T(t) \in \mathcal{L}(E,F) : t \in I\}$ *is* R*-bounded by a constant* $C = C(q,r,E,F)$ *times* M*. Moreover, if the interval* I *is bounded, then the factor* $(1+|s|)^{-\alpha}$ *can be omitted.*

Theorem 5.6 (Theorem 6.1 of [54]). *Let S(t) be a C_0-semigroup on a Banach space E with $\|S(t)\| \leq Me^{-\lambda t}$ for some $\lambda > 0$. If $q \in [1,2]$ and $r \in [2,\infty]$ are the type and co-type of E respectively, let*

$$\alpha > \frac{1}{q} - \frac{1}{r},$$

then

$$\big\{S(t) \in \mathcal{L}(D((-A)^\alpha), E) : t \in \mathbb{R}_+\big\}$$

is R-bounded. Moreover, there exists $C \geq 0$ such that for any $I := [a,b] \subset \mathbb{R}_+$ we have

$$\mathcal{R}\Big(\big\{S(t) \in \mathcal{L}(D((-A)^\alpha), E) : t \in I\big\}\Big) \leq CMe^{-\lambda a}.$$

Proof. For $\theta \in (0,1)$ let E_θ be the real interpolation space $(D(A), E)_{\theta,\infty}$ (see [82, Section 1.6.2]). $x \in E_\theta$ if and only if

$$\|x\|_{E_\theta} := \|x\| + \sup_{t \in \mathbb{R}_+} t^{-\theta} \big\|\big(T(t)x - x\big)\big\|$$

is finite, and this expression defines an equivalent norm on E_θ (see [62, Proposition 3.2.1]). Let $i_\alpha : D((-A)^\alpha) \to E$ be the identity map and $\|x\|_{D((-A)^\alpha)}$ be the graph norm $\|x\|_{D((-A)^\alpha)} = \|x\| + \|(-A)^\alpha x\|$, then for $\theta \in (\frac{1}{q} - \frac{1}{r}, \alpha)$ we have $D((-A)^\alpha) \hookrightarrow E_\theta$ [39, Proposition II.5.33] and so

$$\begin{aligned}
\sup_{t \in \mathbb{R}_+} t^{-\alpha} \big\|T(t) - i_\alpha\big\|_{\mathcal{L}(D((-A)^\alpha), E)} &= \sup_{\|x\|_{D((-A)^\alpha)} \leq 1} \Big\{ \sup_{t \in \mathbb{R}} t^{-\alpha} \|T(t)x - x\| \Big\} \\
&\leq \sup_{\|x\|_{D((-A)^\alpha)} \leq 1} \|x\|_{E_\theta} \\
&\leq C \sup_{\|x\|_{D((-A)^\alpha)} \leq 1} \|x\|_{D((-A)^\alpha)} \\
&= C.
\end{aligned}$$

Hence for $h > 0$

$$\begin{aligned}
\big\|T(s+h) - T(s)\big\|_{\mathcal{L}(D((-A)^\alpha), E)} &\leq \big\|T(s)\big\|_{\mathcal{L}(E)} \big\|T(h) - i_\alpha\big\|_{\mathcal{L}(D((-A)^\alpha), E)} \\
&\leq Ch^\alpha \big\|T(s)\big\|_{\mathcal{L}(E)} \\
&\leq CMe^{-\lambda s} h^\alpha.
\end{aligned}$$

Then the result follows from Proposition 5.5. □

5.2.3 Weak and strong solutions of SDEs

All solutions of stochastic differential equations up to this point have been *strong* solutions, in the sense that the Brownian motion $W(t)$ is given in advance and the solution $U(t)$ obtained is adapted to $(\mathscr{F}_t)$. If, on the other hand, the Brownian motion is not specified in advance and instead we find a pair of processes $\big((\widetilde{U}(t), \widetilde{W}(t)), \mathscr{H}_t\big)$ on a filtered space $\big(\Omega, \mathscr{H}, (\mathscr{H}_t), \mathbb{P}\big)$ such that $\widetilde{U}(t)$ is $\mathscr{H}_t$-adapted, $\widetilde{W}(t)$ is an $\mathscr{H}_t$-Brownian motion (that is, $\widetilde{W}(t)$ is a Brownian motion and an $\mathscr{H}_t$-martingale) and the equation is satisfied, then we call the pair $\big(\widetilde{U}(t), \widetilde{W}(t)\big)$ a weak solution.

Strong solutions are also weak solutions, but the converse is not true, for example the famous Tanaka equation [72, Example 5.3.2].

The existence and uniqueness results given up to now refer to *strong* existence and uniqueness respectively. A strong solution $U(t)$ is strongly unique if, given another strong solution $V(t)$, we have $U(t) = V(t)$ almost surely for all t. Weak uniqueness simply means that any two weak solutions are identical *in Law*, or distribution. If a stochastic differential equation admits a unique strong solution $U(t)$ for a given Brownian motion $W(t)$, then that solution will also be weakly unique, that is, given any weak solution $\big(\widetilde{V}(t), \widetilde{W}(T)\big)$ then $\widetilde{V}(t)$ and $U(t)$ will be identically distributed.

Weak uniqueness of solutions of SDEs on Hilbert spaces are considered in detail in [43] and in 2-smooth Banach spaces in [73]. For a multiplicative and stochastic noise term as in Chapter 4, weak uniqueness is argued as follows. Take a weak solution $(X_1(t), W_1(t))$ on a probability space $(\Omega_1, \mathscr{F}_1, \mathbb{P}_1)$, then by the existence theorem for strong solutions, here exists a strong solution $Y_1(t)$ with respect to W_1. By strong uniqueness, $Y_1(t) = X_1(t)$ almost surely. If $(X_2(t), W_2(t))$ is a different weak solution on a different probability space $(\Omega_2, \mathscr{F}_2, \mathbb{P}_2)$ then in the same way we can construct a strong solution $Y_2(t)$ with respect to W_2 and $Y_2(t) = X_2(t)$ almost surely. Now since Y_1 and Y_2 are constructed as a limit of a sequence of Picard approximations, it is enough to check that for every $n \in \mathbb{N}$, the n^{th} term of the respective Picard sequences have the same distribution. This is automatic for the first terms since the initial conditions are fixed, and then follows for subsequent terms by careful induction.

In the case considered in this chapter however, where the noise term is deterministic and additive, it is enough to note that for the indicator function $\mathbb{1}_{(a,b)}(h \otimes x) : \mathbb{R} \to \mathcal{L}_{FR}(H, E)$ (as defined in (3.9)),

$$\begin{aligned} \int \mathbb{1}_{(a,b)}(h \otimes x)\, \mathrm{d}W_i(t) &= \big[W_i(b)h - W_i(a)h\big]x \\ &= N_i^{\|h\|^2(b-a)}x \end{aligned}$$

where $N_i^{\sigma^2} \sim N(0,\sigma^2)$ for $i = 1,2$. Hence by linearity we have

$$\int \phi \, \mathrm{d}W_1 \overset{(d)}{=} \int \phi \, \mathrm{d}W_2$$

for all step functions ϕ taking values in the finite rank operators from H to E. Then if g is stochastically integrable on $\mathbb{R}$, by Definition 3.28, there exists a sequence of finite rank operator valued step functions (ϕ_n) such that

$$\int \phi_n \, \mathrm{d}W_i \to \int g \, \mathrm{d}W_i$$

in probability for each $i = 1,2$. It follows that the convergence also holds in distribution and hence

$$\int g \, \mathrm{d}W_1 \overset{(d)}{=} \int g \, \mathrm{d}W_2.$$

Note that for $g \in \gamma(L^2(\mathbb{R}_+; H); E)$, the integral

$$\int_0^t g(s) \, \mathrm{d}W_i(s)$$

is Gaussian for all $t \in \mathbb{R}_+$, and thus in this case it is enough to show equality of the mean and covariance operator.

5.3 Almost periodic solutions with additive noise

We will consider the following stochastic differential equation with additive noise

$$\mathrm{d}X(t) = \big(AX(t) + F(t)\big) \, \mathrm{d}t + G(t) \, \mathrm{d}W(t), \quad t \in \mathbb{R} \tag{5.2}$$

and aim to show the existence of almost periodic solutions as defined in 5.4.

We make the following standing assumptions

(P1) A is the generator of a C_0-semigroup $S(t)$ on a Banach space E of type $q \in (1,2]$ and co-type $r \in [2,\infty]$ with $\|S(t)\| \leq Me^{-\lambda t}$ for all $t \geq 0$ for some $M, \lambda > 0$.

(P2) $W(t)$ for $t \in \mathbb{R}$ is of the form

$$W(t) = \begin{cases} W_1(t) & t \geq 0 \\ W_2(-t) & t < 0 \end{cases}$$

where W_1 and W_2 are independent H-cylindrical Brownian motions on E with $W_1(0) = W_2(0) = 0$ almost surely.

(P3) $F : \mathbb{R} \to E$ is bounded and Lipschitz continuous.

(P4) $G : \mathbb{R} \to \gamma(H, D((-A)^\alpha))$ for some $\alpha > \frac{1}{q} - \frac{1}{r}$ satisfies the following

(i) G is $\frac{1}{2}$-Hölder continuous and T-periodic for some $T > 0$;

(ii) $G|_{[0,T]} \in \gamma(L^2(0,T;H), D((-A)^\alpha))$;

(iii) the function $\tau \mapsto G\big(\cdot + \tau \pmod T\big)$ from $\mathbb{R}$ to $\gamma(L^2(0,T;H), D((-A)^\alpha))$ is continuous.

First we give a result of Veraar and Hytönen [54] which guarantees that each solution of (5.2) has a version with continuous paths.

We say a process $X : [t_0, \infty) \times \Omega \to E$ is a (strong) solution of (5.2) with initial condition $X(t_0) = X_0$ for some E-valued random variable X_0 if for all $t \geq t_0$

1. $X(t)$ is strongly measurable and adapted to the filtration $\mathscr{F}_t$;

2. $X(t)$ satisfies the integral equation

$$X(t) = S(t-t_0)X_0 + \int_{t_0}^{t} S(t-s)F(s)\,\mathrm{d}s + \int_{t_0}^{t} S(t-s)G(s)\,\mathrm{d}W(s). \tag{5.3}$$

A process $X : \mathbb{R} \times \Omega \to E$ is then said to be a solution of (5.2) on $\mathbb{R}$ if the restriction of $X(t)$ to $[t_0, \infty)$ is a solution to (5.2) with initial condition $X(t_0)$ for every $t_0 \in \mathbb{R}$.

Proposition 5.7 (Theorem 6.3 of [54])**.** *Suppose E has type $q \in [1,2]$ and co-type $r \in [2,\infty]$ and that $\|S(t)\| \leq Me^{-\lambda t}$ for some $M \geq 0$ and $\lambda > 0$. Let $\alpha > \frac{1}{q} - \frac{1}{r}$ and suppose $G \in \gamma(L^2(\mathbb{R}_+;H), D((-A)^\alpha))$ and F is Lipschitz continuous. Then the equation*

$$\begin{cases} \mathrm{d}X(t) &= \big(AX(t) + F(t)\big)\,\mathrm{d}t + G(t)\,\mathrm{d}W(t), \quad t \in \mathbb{R}_+ \\ X(0) &= x_0 \end{cases}$$

has a unique mild solution $X(t)$ for each $x_0 \in E$, and moreover, if there exists $\varepsilon > 0$ such that for all $T_0 \in \mathbb{R}_+$

$$\sup_{t\in[0,T_0]} \big\|(t-\cdot)^{-\varepsilon} G(\cdot)\big\|_{\gamma(L^2(0,T_0;H),D((-A)^\alpha))} < \infty, \tag{5.4}$$

then $X(t)$ has a version with continuous paths.

Remark 5.8. Note that by [54, Remark 6.4] and [69, Lemma 3.3], a sufficient condition for (5.4) is that there exists a $p > 2$ such that G is in the Besov space $B^{\frac{1}{q}-\frac{1}{2}}_{p,q}(0,T_0;\gamma(H,E))$ for all $T_0 \geq 0$. In this case

$$\sup_{t\in[0,T_0]} \left\|(t-\cdot)^{-\varepsilon}G(\cdot)\right\|_{\gamma(L^2(0,T_0;H),D((-A)^\alpha))} \leq CT_0^{\frac{1}{2}-\frac{1}{p}-\varepsilon}\|G\|_{B^{\frac{1}{q}-\frac{1}{2}}_{p,q}(0,T_0;\gamma(H,D((-A)^\alpha))}.$$

Rather than go into a detailed description of Besov spaces, we simply note that $W^{\frac{1}{2},p}(0,T_0;\gamma(H,E)) \hookrightarrow B^{\frac{1}{q}-\frac{1}{2}}_{p,q}(0,T_0;\gamma(H,E))$ for all $q \in [1,2]$ [62]. It follows that (5.4) holds whenever G is $\frac{1}{2}$-Hölder continuous as a function $\mathbb{R} \to \gamma(H,E)$.

Proof. Suppose (5.4) holds for some $\varepsilon \in [0,\frac{1}{2})$. By Theorem 5.6, the set

$$\left\{S(t) \in \mathcal{L}(D((-A)^\alpha),E) : t \in [0,T]\right\}$$

is R-bounded by some constant C. Therefore, by Lemma 3.37, $S(s)$ acts as a multiplier between the spaces $\gamma(L^2(\mathbb{R}_+;H),D((-A)^\alpha))$ and $\gamma(L^2(\mathbb{R}_+;H),E)$, and we get that

$$\begin{aligned}\sup_{t\in[0,M]} \left\|(t-\cdot)^{-\varepsilon}S(t-\cdot)G(\cdot)\right\|_{\gamma(L^2(\mathbb{R}_+;H),E)} &\leq C\sup_{t\in[0,M]} \left\|(t-\cdot)^{-\varepsilon}G(\cdot)\right\|_{\gamma(L^2(\mathbb{R}_+;H),D((-A)^\alpha))} \\ &< \infty.\end{aligned}$$

Now the result follows from [87, Proposition 3.1 and Theorem 3.3], the first result giving the case $\varepsilon = 0$ and the second giving continuous paths for $\varepsilon \in (0,\frac{1}{2})$. □

We are now ready to give the main result of this section, the existence of almost periodic solutions to equation (5.2).

Theorem 5.9. *Assume (P1) - (P4) and consider the stochastic differential equation (5.2),*

$$\mathrm{d}X(t) = \big(AX(t) + F(t)\big)\,\mathrm{d}t + G(t)\,\mathrm{d}W(t), \quad t \in \mathbb{R}, \tag{5.2}$$

Then for $p \in (1,\infty)$ there exists an L^p-bounded solution $X(t)$ of (5.2) for all $t \in \mathbb{R}$. Moreover, if F is (asymptotically) almost periodic then the solution $X(t)$ is (asymptotically) almost periodic in distribution.

Proof. First, we claim that there exists a permanent (i.e. for all time $t \in \mathbb{R}$) *strong* solution $X(t)$ of (5.2) of the form

$$X(t) = \int_{-\infty}^{t} S(t-s)F(s)\,\mathrm{d}s + \int_{-\infty}^{t} S(t-s)G(s)\,\mathrm{d}W(s), \quad t \in \mathbb{R}, \tag{5.5}$$

for if $X(t)$ is given by (5.5), then by a simple calculation we have for any $t_0 < t$,

$$X(t) = S(t-t_0)X(t_0) + \int_{t_0}^{t} S(t-s)F(s)\,\mathrm{d}s + \int_{t_0}^{t} S(t-s)G(s)\,\mathrm{d}W(s).$$

So according to (5.3), $X(t)$ is a solution of (5.2) on every interval $[t_0, \infty)$ for $t_0 \in \mathbb{R}$, then by Proposition 5.7 and Remark 5.8, this solution is unique and has a version with continuous paths. From here on in $X(t)$ will refer to the version of (5.5) with continuous paths.

We write

$$\begin{aligned} Y(t) &:= \int_{-\infty}^{t} S(t-s)F(s)\,\mathrm{d}s, \\ Z(t) &:= \int_{-\infty}^{t} S(t-s)G(s)\,\mathrm{d}W(s), \end{aligned} \tag{5.6}$$

then

$$\begin{aligned} \|Y(t)\| &\leq \int_{-\infty}^{t} Me^{-\lambda(t-s)} \sup_{s\in\mathbb{R}} \|F(s)\|\,\mathrm{d}s \\ &\leq C \sup_{s\in\mathbb{R}} \|F(s)\| Me^{-\lambda t} \int_{-\infty}^{t} e^{-\lambda s}\,\mathrm{d}s \\ &= \frac{MC}{\lambda} \sup_{s\in\mathbb{R}} \|F(s)\| \\ &< \infty \end{aligned}$$

for all $t \in \mathbb{R}$, so $Y(t)$ is well defined and bounded on $\mathbb{R}$.

Now to show that $Z(t)$ is L^p-bounded, using the T-periodicity of G gives

$$\begin{aligned} \Big(\mathbb{E}\|Z(t)\|^p\Big)^{\frac{1}{p}} &= \Big(\mathbb{E}\Big\| \int_{-\infty}^{t} S(t-s)G(s)\,\mathrm{d}W(s)\Big\|^p\Big)^{\frac{1}{p}} \\ &= \Big(\mathbb{E}\Big\| \sum_{m=1}^{\infty} \int_{t-mT}^{t-(m-1)T} S(t-s)G(s)\,\mathrm{d}W(s)\Big\|^p\Big)^{\frac{1}{p}} \\ &= \Big(\mathbb{E}\Big\| \sum_{m=1}^{\infty} \int_{0}^{T} S(mT-r)G(r+t-mT)\,\mathrm{d}W(r+t-mT)\Big\|^p\Big)^{\frac{1}{p}} \\ &= \Big(\mathbb{E}\Big\| \sum_{m=1}^{\infty} \int_{0}^{T} S(mT-r)G(r+t)\,\mathrm{d}W(r+t-mT)\Big\|^p\Big)^{\frac{1}{p}} \\ &\leq \sum_{m=1}^{\infty} \Big(\mathbb{E}\Big\| \int_{0}^{T} S(mT-r)G(r+t)\,\mathrm{d}W(r+t-mT)\Big\|^p\Big)^{\frac{1}{p}} \end{aligned} \tag{5.7}$$

and then by Itô's isometry (3.12) we have

$$\Big(\mathbb{E}\|Z(t)\|^p\Big)^{\frac{1}{p}} \le C_p \sum_{m=1}^{\infty} \big\|S(mT-\cdot)G(\cdot+t)\big\|_{\gamma(L^2(0,T;H),E)}.$$

Next we appeal to Theorem 5.6. The semigroup $S(t)$ is R-bounded as a family of operators from the fractional domain $D((-A)^\alpha)$ to E on each interval $[(m-1)T, mT]$ with

$$\mathcal{R}(\mathscr{T}_m) := \mathcal{R}\Big(\big\{S(t) \in \mathcal{L}(D((-A)^\alpha), E) : (m-1)T \le t \le mT\big\}\Big) \le CMe^{-\lambda(m-1)T}.$$

By Lemma 3.37

$$\begin{aligned}\big\|S(mT-\cdot)G(\cdot+t)\big\|_{\gamma(L^2(0,T;H),E)} &\le \mathcal{R}(\mathscr{T}_m)\big\|G(\cdot+t)\big\|_{\gamma(L^2(0,T;H),D((-A)^\alpha))}\\ &\le CMe^{-\lambda(m-1)T}\big\|G(\cdot+t)\big\|_{\gamma(L^2(0,T;H),D((-A)^\alpha))}.\end{aligned}$$

and therefore

$$\begin{aligned}\Big(\mathbb{E}\|Z(t)\|^p\Big)^{\frac{1}{p}} &\le C_p \sum_{m=1}^{\infty} \big\|S(mT-\cdot)G(\cdot+t)\big\|_{\gamma(L^2(0,T;H),E)}\\ &\le C_p \sum_{m=0}^{\infty} CMe^{-\lambda mT}\big\|G(\cdot+t)\big\|_{\gamma(L^2(0,T;H),D((-A)^\alpha))}\\ &\le C_pCM\big\|G(\cdot+t)\big\|_{\gamma(L^2(0,T;H),D((-A)^\alpha))}\sum_{m=0}^{\infty} e^{-\lambda mT}\\ &\le C_pCM\frac{e^{\lambda T}}{e^{\lambda T}-1}\sup_{s\in[0,T]}\big\|G(\cdot+s)\big\|_{\gamma(L^2(0,T;H),D((-A)^\alpha))}\\ &< K < \infty\end{aligned}$$

for all $t \in \mathbb{R}$, so $Z(t)$ is indeed well defined and L^p-bounded on $\mathbb{R}$, and hence $X(t)$ exists and is an L^p-bounded solution of (5.2) for bounded F and T-periodic G as required.

Almost periodicity. To show almost periodicity of $X(t)$ in distribution, we want to show the relative compactness in distribution of the set $\{X(\cdot+\tau) : \tau \in \mathbb{R}\}$ in the space of continuous random processes on E. That is, for any sequence $(\tau_n) \subset \mathbb{R}$ with $\tau_n \uparrow \infty$ there exists a subsequence (without loss of generality, also (τ_n)) such that

$$\mathbb{E}\Psi\big(X(t+\tau_n)\big) \to \mathbb{E}\Psi\big(\widetilde{X}(t)\big)$$

uniformly for $t \in \mathbb{R}$ for any bounded Lipschitz function $\Psi : E \to \mathbb{R}$.

First we note that by the assumptions on G and Remark 5.8, (5.4) holds for any $\varepsilon \in (0, \frac{1}{2})$, and so by Proposition 5.7, $X(t)$ has a version with continuous paths.

Let $\tau_n \uparrow \infty$ be a real sequence. Since F and G are almost periodic and T-periodic respectively, passing to a subsequence as necessary, there exist functions $\widetilde{F}$ and $\widetilde{G}$ such that

$$\begin{aligned}
&\|F(t+\tau_n) - \widetilde{F}(t)\| \to 0 \text{ as } n \to \infty \text{ uniformly for } t \in \mathbb{R}, \\
&\|G(\cdot+\tau_n) - \widetilde{G}(\cdot)\|_{\gamma(L^2(0,T;H),D((-A)^\alpha))} \to 0 \text{ as } n \to \infty.
\end{aligned}$$

The latter holds because $\tau_n \to \theta \pmod T$ for some $\theta \in [0,T]$ and then by continuity of the map $\tau \mapsto G(\cdot+\tau)$. Consider the new SDEs

$$\mathrm{d}X(t) = \big(AX(t) + \widetilde{F}(t)\big)\,\mathrm{d}t + \widetilde{G}(t)\,\mathrm{d}W(t), \quad t \in \mathbb{R} \tag{$\widetilde{5.2}$}$$

with solution $\widetilde{X}(t)$ of the form given in (5.5) and

$$\mathrm{d}X(t) = \big(AX(t) + F(t+\tau_n)\big)\,\mathrm{d}t + G(t+\tau_n)\,\mathrm{d}W(t), \quad t \in \mathbb{R}. \tag{$(5.2)_n$}$$

If $X(t)$ is our *strong* solution of (5.2), then $(X(t+\tau_n), W(t+\tau_n))$ is a *weak* solution of $((5.2)_n)$, in the sense that $X(t+\tau_n)$ solves the equation

$$\mathrm{d}X(t) = \big(AX(t) + F(t+\tau_n)\big)\,\mathrm{d}t + G(t+\tau_n)\,\mathrm{d}W(t+\tau_n), \quad t \in \mathbb{R},$$

and moreover, if $X_n(t)$ is the *strong* solution of $((5.2)_n)$ and $P_X(t, \mathrm{d}x)$ is the probability distribution on E generated by $X(t)$, then by weak uniqueness, $P_{X_n}(t, \mathrm{d}x)$ coincides with $P_X(t+\tau_n, \mathrm{d}x)$, or equivalently,

$$\mathbb{E}\Psi\big(X(t+\tau_n)\big) = \mathbb{E}\Psi\big(X_n(t)\big)$$

for all bounded Lipschitz functions $\Psi : E \to \mathbb{R}$ and $t \in \mathbb{R}$. It is therefore enough to show that the sequence $(X_n(t))$ converges uniformly in distribution to $\widetilde{X}(t)$.

Write $X_n(t) := Y_n(t) + Z_n(t)$ and $\widetilde{X}(t) := \widetilde{Y}(t) + \widetilde{Z}(t)$ as in (5.6), then

$$\begin{aligned}
\big\|Y_n(t) - \widetilde{Y}(t)\big\| &= \Big\| \int_{-\infty}^t S(t-s)F(s+\tau_n)\,\mathrm{d}s - \int_{-\infty}^t S(t-s)\widetilde{F}(s)\,\mathrm{d}s \Big\| \\
&\le \int_0^\infty \big\|S(s)\big[F(t+\tau_n-s) - \widetilde{F}(t-s)\big]\big\|\,\mathrm{d}s \\
&\le \sup_{s\in\mathbb{R}} \big\|F(s+\tau_n) - \widetilde{F}(s)\big\| \int_0^\infty \|S(r)\|\,\mathrm{d}r \\
&\to 0 \text{ as } n \to \infty
\end{aligned} \tag{5.8}$$

uniformly for $t \in \mathbb{R}$.

For the stochastic term, fix a bounded Lipschitz $\Psi : E \to \mathbb{R}$ and

$$\begin{aligned}|\mathbb{E}\Psi(Z_n(t)) - \mathbb{E}\Psi(\widetilde{Z}(t))| &\leq \mathbb{E}\,|\Psi(Z_n(t)) - \Psi(\widetilde{Z}(t))| \\ &\leq L_\Psi \mathbb{E}\,\big\|Z_n(t) - \widetilde{Z}(t)\big\| \\ &= L_\Psi \mathbb{E}\,\Big\| \int_{-\infty}^t S(t-s)\big[G(s+\tau_n) - \widetilde{G}(s)\big]\,\mathrm{d}W(s)\Big\| \\ &\leq L_\Psi \sum_{m=1}^\infty \mathbb{E}\,\Big\| \int_0^T S(mT-s)\big[G(s+\tau_n+t-mT) \\ &\qquad - \widetilde{G}(s+t-mT)\big]\,\mathrm{d}W(s+t-mT)\Big\|.\end{aligned}$$

By Itô's isometry and R-boundedness of $\{S(t) : t \in \mathbb{R}_+\} \subset \mathcal{L}(D((-A)^\alpha), E)$ then

$$\begin{aligned}&\leq C_p L_\Psi \sum_{m=1}^\infty \big\|S(mT-\cdot)\big[G(t+\tau_n+\cdot) - \widetilde{G}(t+\cdot)\big]\big\|_{\gamma(L^2(0,T;H),E)} \\ &\leq C_p L_\Psi \sup_{t\in[0,T]} \big\|G(t+\tau_n+\cdot) - \widetilde{G}(t+\cdot)\big\|_{\gamma(L^2(0,T;H),E)} \sum_{m=1}^\infty \mathcal{R}(\mathscr{T}_m) \\ &\leq C_p L_\Psi \frac{e^{\lambda T}}{e^{\lambda T}-1} \sup_{t\in[0,T]} \big\|G(t+\tau_n+\cdot) - \widetilde{G}(t+\cdot)\big\|_{\gamma(L^2(0,T;H),E)} \\ &\to 0 \text{ as } n \to \infty. \end{aligned} \tag{5.9}$$

Collecting (5.8) and (5.9) we get the desired convergence in distribution of $X_n(t)$ to

$$\widetilde{X}(t) := \int_{-\infty}^t S(t-s)\widetilde{F}(s)\,\mathrm{d}s + \int_{-\infty}^t S(t-s)\widetilde{G}(s)\,\mathrm{d}W(s)$$

uniformly for $t \in \mathbb{R}$, and hence $X(t)$ is almost periodic in distribution as claimed.

Asymptotic almost periodicity. By linearity and the fact that any AAP function is the sum of an AP function and a c_0 function, it is enough to note that if $\|F(t)\| \to 0$ as $t \to \infty$ then $\|Y(t)\| \to 0$ as $t \to \infty$.

$$\begin{aligned}\big\|Y(t)\big\| &\leq \int_{-\infty}^t \|S(t-s)F(s)\|\,\mathrm{d}s = \int_0^\infty \|S(s)F(t-s)\|\,\mathrm{d}s \\ &\leq \int_0^\infty Me^{-\lambda s}\|F(t-s)\|\,\mathrm{d}s \to 0 \text{ as } t \to \infty\end{aligned}$$

by the dominated convergence theorem. □

Remark 5.10. There are several alternative ways to think about Theorem 5.9. One way is to notice that stochastic convolutions for additive and deterministic noise terms are simply gaussian processes, so the result could also be obtained by showing the convergence of $X_{\tau_n}(t)$ as a Gaussian process, i.e. showing that the mean and covariance of the finite dimensional distributions converge. A second way is to think of it as a result about continuous dependence on parameters. Here we think of $F_n(t) := F(t+\tau_n)$ and $G_n(t) := G(t+\tau_n)$ as a sequence of approximations to $\widetilde{F}$ and $\widetilde{G}$ and then ask if under these conditions the corresponding solutions converge. This situation is studied by van Neerven and Kunze in [58]

5.4 The non-autonomous problem

We now ask whether a result in the spirit of Theorem 5.9 holds for non-autonomous problems. Consider

$$\mathrm{d}X(t) = \big(A(t)X(t) + F(t)\big)\,\mathrm{d}t + G(t)\,\mathrm{d}W(t), \quad t \in \mathbb{R} \tag{5.10}$$

where $\big(A(t), D(A(t))\big)$ generates an family $\big(P(t,s)\big)_{s\le t}$.

We assume that $A(t)$ satisfies the condition (AT) (as defined in 2.15) and the following conditions.

(P1) $(P(t,s))_{s\le t}$ is exponentially stable, that is, there exist $M \ge 0$ and $\lambda > 0$ such that $\|P(t,s)\| \le Me^{-\lambda(t-s)}$ for all $s \le t$.

(P2) The map $\tau \mapsto \big(P(t+\tau, s+\tau)\big)_{s\le t}$ is almost periodic, in the sense that for any sequence $(\tau_n) \subset \mathbb{R}$ with $\tau_n \uparrow \infty$, there exist a subsequence (τ_{n_r}) and a strongly continuous map $\widetilde{P} : \{(t,s) \in \mathbb{R}^2 : s \le t\} \to \mathcal{L}(E)$ such that

$$P(t+\tau_{n_r}, s+\tau_{n_r})x \to \widetilde{P}(t,s)x \text{ as } r \to \infty$$

uniformly in $t \ge s$ for $t-s$ bounded. We write $P_{\tau_n}(t,s)$ for $P(t+\tau_n, s+\tau_n)$ and say that $P(t,s)$ is *jointly strongly almost periodic*, uniformly on bounded $t-s$.

By a simple argument by strong convergence, $\widetilde{P}(t,s)$ is an evolution family and $\|\widetilde{P}(t,s)\| \le Me^{-\lambda(t-s)}$. Note that if we can show that the strong convergence in (P2) is uniform for bounded $t-s$, then by exponential stability, (P1), it follows that the convergence is in fact uniform for all $t-s$.

In the [64], it is also required that the family $P(t,s)$ is *jointly strongly continuous*, that is, (P3) The map $\tau \mapsto \big(P(t+\tau, s+\tau)\big)_{s\le t}$ is strongly continuous, in the sense that for all $x \in E$ we have

$$P(t+\tau, s+\tau)x \to P(t,s)x \text{ as } \tau \to 0$$

uniformly in $t \ge s$ with $t-s$ bounded.

However, in the case the $A(t)$ satisfies (AT), this condition automatically holds.

Proposition 5.11 (Theorem 2.3 of [1]). *Let $A(t)$ satisfy the conditions (AT), then the evolution family $P(t,s)$ generated by $A(t)$ satisfies the following conditions for $0 \le s \le r \le t \le T$*

$$(i) \quad \big\|P(t,s) - P(r,s)\big\|_{\mathcal{L}(E)} \le c\Big[\log\Big(1+\frac{t-r}{r-s}\Big) + (t-r)^{\mu+\nu-1}\Big];$$
$$(ii) \quad \big\|P(t,r) - P(t,s)\big\|_{\mathcal{L}(E)} \le c\Big[\log\Big(1+\frac{r-s}{t-r}\Big) + (r-s)^{\mu+\nu-1}\Big].$$

Lemma 5.12. *Let $A(t)$ satisfy the conditions (AT), then the evolution family $P(t,s)$ generated by $A(t)$ is jointly strongly continuous, that is, for all $x \in E$,*

$$\big\|\big(P(t+\tau, s+\tau) - P(t,s)\big)x\big\| \to 0 \text{ as } \tau \to 0$$

uniformly for $t-s$ bounded.

Proof. Fix $\varepsilon > 0$, $T > 0$ and $x \in E$. Since $P(t,s)$ is strongly continuous, there exists $\delta > 0$ such that for $t - s < \delta$

$$\begin{aligned}\big\|\big(P(t+\tau, s+\tau) - P(t,s)\big)x\big\| &\le \big\|\big(P(t+\tau, s+\tau) - I\big)x\big\| + \big\|\big(P(t,s) - I\big)x\big\| \\ &< \varepsilon, \quad \text{for all } \tau \in [0,1].\end{aligned}$$

Now for $t-s \in [\delta, T]$, and $\tau < \frac{t-s}{2}$, by Proposition 5.11,

$$\begin{aligned}\big\|P(t+\tau, s+\tau) - P(t,s)\big\| &\le \big\|P(t+\tau, s+\tau) - P(t, s+\tau)\big\| + \big\|P(t, s+\tau) - P(t,s)\big\| \\ &\le c\Big[\log\Big(1+\frac{(t+\tau)-t}{t-(s+\tau)}\Big) + ((t+\tau)-t)^{\mu+\nu-1}\Big] \\ &\qquad + c\Big[\log\Big(1+\frac{(s+\tau)-s}{t-(s+\tau)}\Big) + ((s+\tau)-s)^{\mu+\nu-1}\Big] \\ &= 2c\Big[\log\Big(1+\frac{\tau}{(t-s)-\tau)}\Big) + \tau^{\mu+\nu-1}\Big] \\ &\le 2c\Big[\log\Big(1+\frac{2\tau}{\delta}\Big) + \tau^{\mu+\nu-1}\Big] \\ &\to 0 \text{ as } \tau \to 0,\end{aligned}$$

uniformly for $t-s \in [\delta, T]$. □

Before giving the main result, we first prove an R-bound for evolution families.

Proposition 5.13. *Let E be a Banach space of type $q \in [1,2]$ and co-type $r \in [2,\infty]$. Suppose P satisfies (AT) on E, with*

$$\mu > \frac{1}{p} := \frac{1}{q} - \frac{1}{r}$$

and suppose $\|P(t,s)\| \le Me^{-\lambda(t-s)}$ for some $M \ge 0$ and $\lambda > 0$. Then for any interval $J := (a,b)$ with $0 \le a < b < t \le T$ and any $c \in (b,t)$, there exist $\theta \in (0,1)$ and a constant $C \ge 0$ (independent of t) such that

$$\mathcal{R}\big\{P(t,s) \in \mathcal{L}(E) : s \in (a,b)\big\} \le \frac{C}{(c-b)^\theta} e^{-\lambda(t-c)}.$$

Proof. We apply Proposition 5.5 to the function $f : J \to \mathcal{L}(E)$ defined by

$$f(s) := P(t,s), \quad s \in (a,b).$$

Clearly $f \in L^p(J;\mathcal{L}(E))$. Choose $\theta \in (\frac{1}{p},\mu)$, then by [91, Theorem 2.3] we have that

$$\big\|P(t,s)(-A(t))^\theta\big\| \le C_1(t-s)^{-\theta}, \quad 0 \le s < t \le T. \tag{5.11}$$

By [54, Proposition 6.5], for $s,h \ge 0$ such that $s \le s+h \le t \le T$ we have

$$\big\|A(s+h)^{-\theta}\big(P(s+h,s) - I\big)\big\| \le C_2 h^\theta. \tag{5.12}$$

Now for s,h such that $s, s+h \in (a,b)$,

$$\begin{aligned}
\|f(s+h) - f(s)\| &= \big\|P(t,s+h) - P(t,s)\big\| \\
&\le \|P(t,c)\|\big\|P(c,s+h)A(s+h)^\theta A(s+h)^{-\theta}\big[I - P(s+h,s)\big]\big\| \\
&\le Me^{\lambda(t-c)}\big\|P(c,s+h)A(s+h)^\theta\big\|\big\|A(s+h)^{-\theta}\big[I - P(s+h,s)\big]\big\| \\
&\le Me^{\lambda(t-c)}\, C_1 h^\theta\, C_2(c-s-h)^{-\theta} \\
&\le \frac{MC_1C_2}{(c-b)^\theta}\, e^{\lambda(t-c)}\, h^\theta \\
&\le \frac{MC_1C_2}{(c-b)^\theta}\, e^{\lambda(T-c)}\, h^\theta.
\end{aligned} \tag{5.13}$$

Then by Proposition 5.5, f is R-bounded on (a,b) with

$$\mathcal{R}\big\{P(t,s) \in \mathcal{L}(E) : s \in (a,b)\big\} \le \frac{C}{(c-b)^\theta} e^{-\lambda(t-c)}.$$

□

Now we can give the main result of this section, the existence of (asymptotically) almost periodic solutions to equation (5.10).

Theorem 5.14. *Let E be a Banach space of type $q \in [1,2]$ and co-type $r \in [2,\infty]$. Assume that (P1) holds and that $\big(A(t), D(A(t))\big)$ satisfies (AT) with*

$$\mu > \frac{1}{q} - \frac{1}{r}$$

(see Definition 2.15). Suppose that $F : \mathbb{R} \to E$ is a bounded, measurable function and $G : \mathbb{R} \to \gamma(H, D((-A)^{\alpha}))$ for some $\alpha \in (\frac{1}{q} - \frac{1}{r}, \mu)$ satisfied the following

(i) G is $\frac{1}{2}$-Hölder continuous and T-periodic for some $T > 0$;

(ii) $G|_{[0,T]} \in \gamma(L^2(0,T;H), D((-A)^{\alpha}))$;

(iii) the map $\tau \mapsto G\big(\cdot + \tau \pmod T\big)$; $\mathbb{R} \to \gamma(L^2(0,T;H), D((-A)^{\alpha}))$ is continuous.

Then (5.10) has an L^p-bounded solution $X : \mathbb{R} \times \Omega \to E$ for any $p \in (1,\infty)$.

Moreover, if we assume (P2) and F is (asymptotically) almost periodic, then (5.10) has a solution that is (asymptotically) almost periodic in distribution.

Proof. As in the semigroup case (Proposition 5.7), the equation (5.10) has a unique mild solution with a continuous version for any initial condition $x \in E$. This follows from [54, Theorem 6.6] and [87, Theorem 3.3] by using γ-multipliers and Lemma 3.37.

As in Theorem 5.9 there exists a permanent solution of the form

$$X(t) = \int_{-\infty}^{t} P(t,s)F(s)\,\mathrm{d}s + \int_{-\infty}^{t} P(t,s)G(s)\,\mathrm{d}W(s), \quad t \in \mathbb{R}. \tag{5.14}$$

Once again we write

$$Y(t) := \int_{-\infty}^{t} P(t,s)F(s)\,\mathrm{d}s$$
$$Z(t) := \int_{-\infty}^{t} P(t,s)G(s)\,\mathrm{d}W(s)$$

then

$$\|Y(t)\| \le \int_{-\infty}^{t} Me^{-\lambda(t-s)} \sup_{s\in\mathbb{R}} \|F(s)\|\,\mathrm{d}s \le \sup_{s\in\mathbb{R}} \|F(s)\| \int_{0}^{\infty} Me^{-\lambda\tau}\,\mathrm{d}\tau. \tag{5.15}$$

To show that $Z(t)$ is L^p-bounded we proceed as in (5.7) and use Itô's isometry (3.12) to obtain

$$\Big(\mathbb{E}\|Z(t)\|^p\Big)^{\frac{1}{p}} \le C_p \sum_{m=1}^{\infty} \big\|P(t, t-mT+\cdot)G(\cdot+t)\big\|_{\gamma(L^2(0,T;H),E)}.$$

We now want to estimate each term of the above sum. Since $P(t,s)$ behaves differently when $(t-s)$ is small, we consider the $m=1$ case and $m \geq 2$ cases separately. For $m \geq 2$, by Proposition 5.13 we have

$$\big\|P(t, t-mT+\cdot)G(\cdot+t)\big\|_{\gamma(L^2(0,T;H),E)} \leq C_p e^{-\lambda(m-1)T} \|G(t+\cdot)\|_{\gamma(L^2(0,T;H),E)}$$

and by [54, Theorem 6.6], for $\varepsilon \in (0, \frac{1}{2})$

$$\big\|P(t, t-T+\cdot)G(\cdot+t)\big\|_{\gamma(L^2(0,T;H),E)} \leq C\|(T-\cdot)^{-\varepsilon}G(t+\cdot)\|_{\gamma(L^2(0,T;H),E)}$$

which is finite by Remark 5.8. Then

$$\begin{aligned}\Big(\mathbb{E}\|Z(t)\|^p\Big)^{\frac{1}{p}} &\leq C\|(T-\cdot)^{-\varepsilon}G(t+\cdot)\|_{\gamma(L^2(0,T;H),E)} + \frac{C_p}{e^{\lambda T}-1}\|G(t+\cdot)\|_{\gamma(L^2(0,T;H),E)} \\ &\leq C \sup_{r\in[0,T]} \Big[\|(T-\cdot)^{-\varepsilon}G(r+\cdot)\|_{\gamma(L^2(0,T;H),E)} + \|G(r+\cdot)\|_{\gamma(L^2(0,T;H),E)}\Big] \\ &< \infty \end{aligned} \tag{5.16}$$

uniformly for $t \in \mathbb{R}$, so $X(t)$ is a well defined L^p-bounded solution of (5.10) for bounded F and T-periodic G as required.

Almost periodicity. To show almost periodicity of $X(t)$ in distribution, we once again show relative compactness of the set $\{X(\cdot+\tau) : \tau \in \mathbb{R}\}$ in the space of continuous bounded processes on E.

As before, we take a real sequence (τ_n) with $\tau_n \uparrow \infty$, then passing to a subsequence as necessary there exists functions $\widetilde{F}$ and $\widetilde{G}$ such that

$$\|F(t+\tau_n) - \widetilde{F}(t)\| \to 0 \text{ as } n \to \infty \tag{5.17}$$

$$\|G(\cdot+\tau_n) - \widetilde{G}(\cdot)\|_{\gamma(L^2(0,T;H),E)} \to 0 \text{ as } n \to \infty \tag{5.18}$$

$$\|P_{\tau_n}(t,s)x - \widetilde{P}(t,s)x\| \to 0 \text{ as } n \to \infty \text{ for each } x \in E \tag{5.19}$$

uniformly for $s, t \in \mathbb{R}$ with $s \leq t$, $t-s$ bounded.

We introduce the new SDEs

$$\mathrm{d}X(t) = \big(A(t+\tau_n)X(t) + F(t+\tau_n)\big)\,\mathrm{d}t + G(t+\tau_n)\,\mathrm{d}W(t), \quad t \in \mathbb{R} \tag{5.10$_n$}$$

with solution $X_n(t)$ of the form given in (5.14) and define the process $\widetilde{X}(t)$ by

$$\widetilde{X}(t) = \int_{-\infty}^{t} \widetilde{P}(t,s)\widetilde{F}(s)\,\mathrm{d}s + \int_{-\infty}^{t} \widetilde{P}(t,s)\widetilde{G}(s)\,\mathrm{d}W(s).$$

If $X(t)$ is our *strong* solution of (5.10) then $(X(t+\tau_n), W(t+\tau_n))$ is a *weak* solution of (5.10_n), in the sense that $X(t+\tau_n)$ solves the equation

$$\mathrm{d}X(t) = \big(A(t+\tau_n)X(t) + F(t+\tau_n)\big)\,\mathrm{d}t + G(t+\tau_n)\,\mathrm{d}W(t+\tau_n), \quad t \in \mathbb{R},$$

and moreover, if $X_n(t)$ is the *strong* solution of (5.10_n) and $P_X(t,\,\mathrm{d}x)$ is the probability distribution on E generated by $X(t)$, then by weak uniqueness, $P_{X_n}(t,\,\mathrm{d}x)$ coincides with $P_X(t+\tau_n,\,\mathrm{d}x)$, or equivalently,

$$\mathbb{E}\Psi\big(X(t+\tau_n)\big) = \mathbb{E}\Psi\big(X_n(t)\big)$$

for all bounded Lipschitz functions $\Psi : E \to \mathbb{R}$ and $t \in \mathbb{R}$. It is therefore enough to show that the sequence $(X_n(t))$ converges uniformly in distribution to $\widetilde{X}(t)$.

Write $X_n(t) := Y_n(t) + Z_n(t)$ and $\widetilde{X}(t) := \widetilde{Y}(t) + \widetilde{Z}(t)$ as in (5.15), then

$$\begin{aligned} Y_n(t) - \widetilde{Y}(t) &= \int_{-\infty}^t P_{\tau_n}(t,s)F(s+\tau_n)\,\mathrm{d}s - \int_{-\infty}^t \widetilde{P}(t,s)\widetilde{F}(s)\,\mathrm{d}s \\ &= \int_{-\infty}^t P_{\tau_n}(t,s)\big[F(s+\tau_n) - \widetilde{F}(s)\big]\,\mathrm{d}s + \int_{-\infty}^t \Big[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\Big]\widetilde{F}(s)\,\mathrm{d}s \end{aligned}$$

and

$$\begin{aligned} \Big\| \int_{-\infty}^t P_{\tau_n}(t,s)\big[F(s+\tau_n) - \widetilde{F}(s)\big]\,\mathrm{d}s \Big\| &\le \int_{-\infty}^t \big\|P_{\tau_n}(t,s)\big[F(s+\tau_n) - \widetilde{F}(s)\big]\big\|\,\mathrm{d}s \\ &\le \int_{-\infty}^t \big\|P_{\tau_n}(t,s)\big\|\big\|F(s+\tau_n) - \widetilde{F}(s)\big\|\,\mathrm{d}s \\ &\le \sup_{s\in\mathbb{R}} \|F(s+\tau_n) - \widetilde{F}(s)\| \int_0^\infty Me^{-\lambda r}\,\mathrm{d}r \\ &\to 0 \text{ as } n \to \infty \end{aligned} \tag{5.20}$$

uniformly for $t \in \mathbb{R}$. Now $\widetilde{F}$ has relatively compact range by (5.17), so for all $\varepsilon > 0$ there exist $N \in \mathbb{N}$ and $x_1, \ldots, x_N \in E$ such that for any $s \in \mathbb{R}$ there exists $i \in \{1, \ldots, N\}$ with $\|\widetilde{F}(s) - x_i\| < \varepsilon$. Let $T_0 \ge 0$, then for any $s \in (t - T_0, t]$,

$$\begin{aligned} \big\|\big[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\big]\widetilde{F}(s)\big\| &\le \big\|P_{\tau_n}(t,s) - \widetilde{P}(t,s)\big\|\|\widetilde{F}(s) - x_i\| \\ &\qquad + \big\|\big[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\big]x_i\big\| \\ &\le 2Me^{-\lambda(t-s)} \cdot \varepsilon + \big\|\big[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\big]x_i\big\|, \end{aligned}$$

so

$$\sup_{t-s\le T_0} \big\|\big[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\big]\widetilde{F}(s)\big\| \le 2M \cdot \varepsilon + \sup_{t-s\le T_0} \max_{i=1}^N \big\|\big[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\big]x_i\big\|.$$

Letting $n \to \infty$, by (P2) we get

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}} \sup_{t-s\le T_0} \left\| \left[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\right]\widetilde{F}(s)\right\| \le 2M\cdot\varepsilon. \tag{5.21}$$

Since ε was arbitrary, in fact the limit is 0.

Since P and $\widetilde{P}$ satisfy $\|P(t,s)\| \le Me^{-\lambda(t-s)}$ we have

$$\left\|e^{\frac{\lambda}{2}(t-s)}\left[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\right]\right\| \le 2Me^{-\frac{\lambda}{2}(t-s)}.$$

Fix $\varepsilon > 0$ and $x \in E$, then there exists $T_\varepsilon \ge 0$ such that

$$\left\|e^{\frac{\lambda}{2}(t-s)}\left[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\right]\right\| \le \frac{\varepsilon}{2} \text{ for all } (t-s) \ge T_\varepsilon$$

and by (5.21) there exists $N_\varepsilon \ge 0$ such that

$$\sup_{t\in\mathbb{R}} \sup_{(t-s)\le T_\varepsilon} \left\|\left[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\right]\widetilde{F}(s)\right\| \le \frac{\varepsilon}{2}e^{-\frac{\lambda}{2}T_\varepsilon} \text{ for all } n \ge N_\varepsilon.$$

Together these give us that

$$\left\|e^{\frac{\lambda}{2}(t-s)}\left[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\right]\widetilde{F}(s)\right\| \to 0 \text{ as } n \to \infty,$$

uniformly in $t \ge s$ for each $x \in E$. Then

$$\begin{aligned}
&\left\|\int_{-\infty}^{t} \left[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\right]\widetilde{F}(s)\,\mathrm{d}s\right\| \\
&\qquad = \left\|\int_0^\infty e^{-\frac{\lambda}{2}r}e^{\frac{\lambda}{2}r}\left[P_{\tau_n}(t,t-r) - \widetilde{P}(t,t-r)\right]\widetilde{F}(t-r)\,\mathrm{d}r\right\| \\
&\qquad \le \sup_{t\in\mathbb{R}}\sup_{r\ge 0}\left\|e^{\frac{\lambda}{2}r}\left[P_{\tau_n}(t,t-r) - \widetilde{P}(t,t-r)\right]\widetilde{F}(t-r)\right\| \int_0^\infty e^{-\frac{\lambda}{2}r}\,\mathrm{d}r \\
&\qquad \to 0 \text{ as } n \to \infty.
\end{aligned}$$

Together with (5.20), this shows

$$\left\|Y_n(t) - \widetilde{Y}(t)\right\| \to 0 \text{ as } n \to \infty$$

uniformly for $t \in \mathbb{R}$. Now for the stochastic term:

$$\begin{aligned}
|\mathbb{E}\Psi(Z_n(t)) - \mathbb{E}\Psi(\widetilde{Z}(t))| &\le \mathbb{E}\,|\Psi(Z_n(t)) - \Psi(\widetilde{Z}(t))| \\
&\le L_\Psi \mathbb{E}\left\|Z_n(t) - \widetilde{Z}(t)\right\| \\
&= L_\Psi \mathbb{E}\left\|\int_{-\infty}^{t} P_{\tau_n}(t,s)\left[G(s+\tau_n) - \widetilde{G}(s)\right]\right. \\
&\qquad\qquad \left. + \left[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\right]\widetilde{G}(s)\,\mathrm{d}W(s)\right\|
\end{aligned}$$

Define $\Phi_n : \{(s,t) \in \mathbb{R}^2 : s \le t\} \to \mathcal{L}(H,E)$ by

$$\Phi_n(t,s) := P_{\tau_n}(t,s)\big[G(s+\tau_n) - \widetilde{G}(s)\big] + \big[P_{\tau_n}(t,s) - \widetilde{P}(t,s)\big]\widetilde{G}(s)$$

Then by (5.18) and (5.19) we have

$$\sup_{t\in\mathbb{R}}\sup_{r\ge 0} \|\Phi_n(t,t-r)\| \to 0 \text{ as } n \to \infty.$$

and so for every $x^* \in E^*$,

$$\sup_{t\in\mathbb{R}}\sup_{r\ge 0} \|\Phi_n^*(t,t-r)x^*\| \to 0 \text{ as } n \to \infty. \tag{5.22}$$

Now $\Phi_n^* x^*$ is uniformly dominated by

$$\big\|\Phi_n^*(t,t-r)x^*\big\|_H \le 4Me^{-\lambda r}\sup_{s\in\mathbb{R}}\|G(s)\|\|x^*\| \text{ for all } t\in\mathbb{R} \text{ and } n \in \mathbb{N},$$

so for $\varepsilon > 0$ there exists $R \ge 0$ such that

$$\int_R^\infty \big\|\Phi_n^*(t,t-r)x^*\big\|_H^2 \,\mathrm{d}r \le \Big(4M\sup_{s\in\mathbb{R}}\|G(s)\|\|x^*\|\Big)^2 \int_R^\infty e^{-2\lambda r}\,\mathrm{d}r < \frac{\varepsilon}{2} \tag{5.23}$$

for all $t \in \mathbb{R}$ and $n \in \mathbb{N}$. By (5.22) there exists $N \ge 0$ such that

$$\sup_{t\in\mathbb{R}}\int_0^R \big\|\Phi_n^*(t,t-r)x^*\big\|_H^2\,\mathrm{d}r \le R\sup_{\substack{t\in\mathbb{R}\\ r\ge 0}}\big\|\Phi_n^*(t,t-r)x^*\big\|_H^2 < \frac{\varepsilon}{2} \tag{5.24}$$

for all $n \ge N$. Combining (5.23) and (5.24) we get

$$\sup_{t\in\mathbb{R}}\int_0^\infty \big\|\Phi_n^*(t,t-r)x^*\big\|_H^2\,\mathrm{d}r \to 0 \text{ as } n\to\infty.$$

Then by the uniform version of the stochastic dominated convergence theorem, Corollary 3.31, we have

$$\mathbb{E}\Big\|\int_0^\infty \Phi_n(t,t-r)\,\mathrm{d}W(r)\Big\| \to 0 \text{ as } n \to \infty \tag{5.25}$$

Uniformly for $t \in \mathbb{R}$. Therefore $Z_n(t)$ converges to $\widetilde{Z}(t)$ in distribution, uniformly for $t \in \mathbb{R}$ as required.

Asymptotic almost periodicity follows exactly as in Theorem 5.9. □

Remark 5.15. By taking $A(t) = A$, a constant generator of an analytic semigroup in Theorem 5.14, we get a version of Theorem 5.9 which makes no assumption upon the Banach space E being on non-trivial type. Thus in the case of analytic semigroups, Theorem 5.9 holds on any general Banach space.

5.5 Aside: Almost periodicity of evolution families

In the next section we will give an example for Theorem 5.14 with a particular differential operator on a smooth domain in $\mathbb{R}^d$, but first we develop some properties of evolution families satisfying the (AT) conditions.

Almost periodicity of non-autonomous problems has been widely studied. An evolution family $P(t, s)$ is said to be T-periodic if

$$P(t + T, s + T) = P(t, s) \tag{5.26}$$

for all suitable $s \leq t$, and in [12] it is shown that in the case where $A(t)$ is T-periodic then so is $P(t, s)$, and in this case the Cauchy problem

$$u'(t) = A(t)u(t) + f(t), \quad t \geq 0$$

has an almost periodic solution $u(t)$ for every almost periodic function $f : \mathbb{R}_+ \to E$. See also [47], [48] and [79]. In the previous section we assumed (condition (P2)) the existence of an evolution family $P(t, s)$ with almost periodicity in a sense similar to (5.26), simultaneously in both variables. Mokhtar-Kharroubi in [64] uses such a family, but in that case the only example given is of a generating family $A(t)$ of multiplication operators on a space $L^p(\mathcal{O})$. The question is whether such a condition can hold and be checked for classes of evolution families which occur in wider applications.

The aim of this section is to give a set of easily checkable conditions on a family $A(t)$ satisfying the (AT) conditions of Acquistapace and Terreni which imply this property for the generated family $P(t, s)$.

Problems in which the family $A(t)$ is almost periodic in some sense have been studied in several works by Schnaubelt *et al.* for example [10] and [63] and by various other authors (for example [7], [9] or [37]). Typically these papers assume that the map $t \mapsto R(\lambda, A(t))$ be almost periodic and then proceed by way of Yosida approximations. Our approach in this section differs from established techniques in that we make assumptions directly about almost periodicity of the map $t \mapsto A(t)$ in a suitable topology and then proceed by going back to the original construction of Acquistapace and Terreni [1], [2] and [3] and proving almost periodicity term by term. Unfortunately this is a rather long and painstaking process.

It is shown in [64], in the section above and in the example to follow in Section 5.6 that this sense of almost periodicity of evolution families is useful and has applications.

Definition 5.16. A function $f : \mathbb{R}^2 \to E$ is said to be *jointly almost periodic* if for any sequence $\tau_n \uparrow \infty$, there exists a subsequence (τ_{n_r}) such that the sequence $f(t + \tau_{n_r}, s + \tau_{n_r})$ converges uniformly for $t, s \in \mathbb{R}$ with $|t - s| \leq T_0$ for any $T_0 > 0$.

The main result of this section is the following

Theorem 5.17. *Suppose $A(t)$ satisfy the (AT) conditions and that*

(i) each $A(t)$ generates an analytic semigroup $e^{sA(t)}$ and the domains $D(A(t)) =: E_1$ are constant for $t \in \mathbb{R}$ with uniformly equivalent graph norms;

(ii) for some $\alpha \in (0, 1)$ and $q \in [1, \infty)$, the difference $A(t) - A(s)$ is a bounded operator from E_α to E for all $t, s \in \mathbb{R}$, where $E_\alpha := E_{\alpha,\infty}$ is the real interpolation space $(E, E_1)_{\alpha,\infty}$;

(iii) for any sequence $\tau_n \uparrow \infty$ we can pass to a subsequence such that $\|A(t + \tau_n) - A(t + \tau_m)\|_{\mathcal{L}(E_\alpha,E)} \to 0$ as $m, n \to \infty$ uniformly for $t \in \mathbb{R}$.

Then the evolution family $P(t, s)$ generated by $A(t)$ is jointly almost periodic, uniformly for bounded $t - s$.

The interpolation spaces $E_\alpha := E_{\alpha,\infty} = (E, E_1)_{\alpha,\infty}$ are as in Definition 2.16. For each $t \in \mathbb{R}$ and $1 > \alpha > \beta > 0$ we have (by, for example [39, Proposition II.5.14 and II.5.33])

$$E_1 = D(A(t)) \hookrightarrow E_\alpha \hookrightarrow D((-A(t)^\beta) \hookrightarrow E_\beta \hookrightarrow E. \tag{5.27}$$

A motivating example of families $A(t)$ satisfying condition *(ii)* is higher order (≥ 2) differential operators on some domain in $\mathbb{R}^d$ which are constant in time in their highest order term.

Lemma 5.18 (A perturbation result for analytic semigroups)**.** *Suppose A and B are generators of analytic semigroups $S(t)$ and $T(t)$ on E respectively with $D(A) = D(B)$ and equivalent graph norms and that $(A - B)$ is in $\mathcal{L}(D((-A)^\alpha), E)$ for some $\alpha \in (0, 1)$.*

Let $\varepsilon > 0$, $T_0 \geq 0$ and write $M := \sup_{s\in[0,T_0]} \{\|S(s)\|, \|T(s)\|\}$ and $C_\alpha := \sup_{s>0} \{\|s^\alpha(-A)^\alpha S(s)\|, \|s^\alpha B^\alpha T(s)\|\}$. Then there exists $K > 0$, depending only on ε, T_0, α, M and C_α, such that if $\|A - B\|_{\mathcal{L}(D((-A)^\alpha),E)} < \frac{\varepsilon}{K}$ then

$$\|T(t) - S(t)\| \leq \varepsilon \text{ for all } t \in [0, T_0].$$

Proof. By (2.7) there exists $C_\alpha \geq 0$ such that

$$\|S(t)\|_{\mathcal{L}(E,D((-A)^\alpha))} \leq \frac{C_\alpha}{t^\alpha},$$

then for $x \in E$

$$\begin{aligned}
\int_0^1 \|(B-A)S(t)x\| \,\mathrm{d}t &\leq \int_0^1 \|B-A\|_{\mathcal{L}(D((-A)^\alpha),E)} \|S(t)x\|_{D((-A)^\alpha)} \,\mathrm{d}t \\
&\leq \frac{\varepsilon}{K} C_\alpha \|x\| \int_0^1 t^{-\alpha} \,\mathrm{d}t \\
&= \frac{\varepsilon}{K} \frac{C_\alpha}{1-\alpha} \|x\|.
\end{aligned}$$

Choosing K large enough so that $\varepsilon C_\alpha < K(1-\alpha)$, we have by [39, Corollary III.3.16] that $(B-A)$ is a Miyadera-Voigt perturbation of A and $S(t)$ and $T(t)$ are related by the variation of parameters formula

$$T(t)x - S(t)x = \int_0^t T(s)(B-A)S(t-s)x \,\mathrm{d}s, \qquad \text{for all } x \in D(A).$$

Then for $x \in D(A)$ and $t \leq T_0$

$$\begin{aligned}
\big\|T(t)x - S(t)x\big\| &\leq \int_0^t \|T(s)\| \|(B-A)\|_{\mathcal{L}(D((-A)^\alpha),E)} \|S(t-s)x\|_{D((-A)^\alpha)} \,\mathrm{d}s, \\
&\leq M \frac{\varepsilon}{K} C_\alpha \|x\| \int_0^t s^{-\alpha} \,\mathrm{d}s \leq M \frac{\varepsilon}{K} \frac{C_\alpha T_0^{1-\alpha}}{1-\alpha} \|x\| \\
&< \varepsilon \|x\|
\end{aligned}$$

for sufficiently large $K > M C_\alpha T_0^{1-\alpha}(1-\alpha)^{-1}$. Then the result follows from the density of $D(A)$. □

Remark 5.19. Note that in the Theorem 5.17 and Lemma 5.18, if the evolution family and semigroup respectively are exponentially stable then the condition that $(t-s)$ be bounded is unnecessary, as the convergence in question automatically holds in the tail.

Suppose $A(t)$, $t \in \mathbb{R}$ satisfies the conditions (AT). Fix $T_0 \geq 0$. Define the kernel $Q(t,s) \in \mathcal{L}(E)$ for $s \leq t$ by

$$Q(t,s) = A(t)^2 e^{(t-s)A(t)} \big[A(t)^{-1} - A(s)^{-1}\big], \quad t-s \leq T_0, \tag{5.28}$$

and then $Q_k(t,s)$ inductively by

$$Q_1(t,s) := Q(t,s), \qquad Q_k(t,s) := \int_s^t Q_{k-1}(t,r)Q(r,s) \,\mathrm{d}r. \tag{5.29}$$

Following [1], the evolution family $P(t,s)$ generated by $A(t)$ is given by

$$P(t,s) = e^{(t-s)A(s)} + \int_s^t Z(r,s)\,\mathrm{d}r$$

where

$$\begin{aligned} Z(t,s) := g_1(t,s) + \sum_{k=1}^{\infty} \int_s^t Q_k(t,r) g_1(r,s)\,\mathrm{d}r \\ + \sum_{k=1}^{\infty} \int_s^t \big[Q_k(t,r) - Q_n(t,s)\big] g_2(r,s)\,\mathrm{d}r \\ + \sum_{k=1}^{\infty} Q_k(t,s)\big[e^{(t-s)A(s)} - I\big], \end{aligned} \tag{5.30}$$

for $g_i(t,s)$, $i = 1,2$ given by

$$\begin{aligned} g_1(t,s) &= A(t)e^{(t-s)A(t)} - A(s)e^{(t-s)A(s)}, \\ g_2(t,s) &= A(s)e^{(t-s)A(s)}. \end{aligned}$$

We will show that each term is jointly almost periodic, uniformly for bounded $t-s$.

Lemma 5.20. *Suppose $A(t)$, $t \in \mathbb{R}$ is a family of operators on E satisfying the condition (AT1) of Definition 2.15, so for each $t \in \mathbb{R}$, $A(t)$ generates an analytic semigroup $\left(e^{sA(t)}\right)_{s\geq 0}$. Then for any $\alpha \in (0,1]$ there exist constants M, C_α such that*

$$\|e^{sA(t)}\| \leq M \quad \textit{and} \quad \|A(t)^\alpha e^{sA(t)}\| \leq \frac{C_\alpha}{s^\alpha} \tag{5.31}$$

for all $t \in \mathbb{R}$ and $s \in [0,T_0]$ for any $T_0 \geq 0$.

Proof. Let A be the generator of an analytic semigroup $S(s)$ on E with $\Sigma_{\theta+\frac{\pi}{2}} \cup \{0\} \subset \rho(A)$ and

$$\|R(\lambda, A)\| \leq \frac{N}{1+|\lambda|}$$

for all $\lambda \in \Sigma_{\theta+\frac{\pi}{2}} \cup \{0\}$ for some $\theta \in (0, \frac{\pi}{2})$. By the proof of [39, Theorem II.4.6], parts (e) $\Rightarrow$ (c) and (c) $\Rightarrow$ (a), it follow that there exist constants M and C_1 depending only on N and θ such that

$$\|S(s)\| \leq M \quad \text{and} \quad \|AS(s)\| \leq \frac{C_1}{s}$$

for all $s \in [0,T_0]$. Hence by condition (AT1), (5.31) holds for $\alpha = 1$.

The next argument follows the lines of [39, Theorem II.5.34]. Let $x_0 \in E$ and $\beta \in (0,1)$. The following estimates hold

$$\begin{aligned}\|s^{-\beta}R(s,A)x_0\| &\le s^{-\beta}\|AR(s,A)\|\|A^{-1}x_0\| \\ &\le s^{-\beta}\Big[\frac{sN}{1+s}+1\Big]\|A^{-1}x_0\| \\ &\le s^{-\beta}\big[N+1\big]\|A^{-1}x_0\|\end{aligned}$$

and

$$\begin{aligned}\|s^{-\beta}R(s,A)x_0\| &\le \frac{s^{-\beta}N}{1+s}\|x_0\| \\ &\le s^{-(1+\beta)}N\|x_0\|.\end{aligned}$$

Then by [39, Theorem II.5.28],

$$\begin{aligned}\|(-A)^{-\beta}x_0\| &= \Big\|\frac{1}{2\pi i}\big(1-e^{-2\pi i\beta}\big)\int_0^\infty s^{-\beta}R(s,A)x_0\,\mathrm{d}s\Big\| \\ &\le C'_\beta\Big\|\int_0^\tau s^{-\beta}R(s,A)x_0\,\mathrm{d}s+\int_\tau^\infty s^{-\beta}R(s,A)x_0\,\mathrm{d}s\Big\| \\ &\le C'_\beta\big[N+1\big]\|A^{-1}x_0\|\int_0^\tau s^{-\beta}\,\mathrm{d}s+C'_\beta N\|x_0\|\int_\tau^\infty s^{-(1+\beta)}\,\mathrm{d}s \\ &\le \frac{C'_\beta\big[N+1\big]}{1-\beta}\tau^{1-\beta}\|A^{-1}x_0\|+\frac{C'_\beta N}{\beta}\tau^{-\beta}\|x_0\|\end{aligned}$$

for any $\tau \in (0,\infty)$. So choosing $\tau := \|A^{-1}x_0\|^{-1}\cdot\|x_0\|$ we get

$$\begin{aligned}\|(-A)^{-\beta}x_0\| &\le \|A^{-1}x_0\|^\beta\cdot\|x_0\|^{1-\beta}\Big[\frac{C'_\beta(N+1)}{1-\beta}+\frac{CN}{\beta}\Big] \\ &=: C''_\beta\|A^{-1}x_0\|^\beta\cdot\|x_0\|^{1-\beta}.\end{aligned}$$

Now we set $\beta := 1-\alpha$ and $x_0 = AS(t)x$ for $x \in E$ and

$$\begin{aligned}\|(-A)^\alpha S(s)x\| &= \|(-A)^{-(1-\alpha)}AS(s)x\| \\ &\le C''_{(1-\alpha)}\|S(s)x\|^{1-\alpha}\|AS(s)x\|^\alpha \\ &\le C''_{(1-\alpha)}M^{1-\alpha}\|x\|^{1-\alpha}\left(\frac{C_1}{s}\right)^\alpha\|x\|^\alpha \\ &\le \frac{C_\alpha}{s^\alpha}\|x\|.\end{aligned}$$

Now C_α depends only on M, N, C_1 and α, so (5.31) holds for all $\alpha \in (0,1]$. $\square$

Remark 5.21. By Lemma 5.20, if $A(t)$, $t \in \mathbb{R}$ satisfy the (AT) conditions then the result of Lemma 5.18 holds uniformly for $t \in \mathbb{R}$. That is, we can choose the K in Lemma 5.18 independent of the choice of $A = A(t_1)$, $B = A(t_2)$ for $t_1, t_2 \in \mathbb{R}$.

Corollary 5.22. *Under the conditions of Theorem 5.17, the map $(t, s) \mapsto e^{(t-s)A(s)}$ is jointly almost periodic (in norm), uniformly for bounded $t - s$.*

Proof. Let $(\tau_n) \subset \mathbb{R}$, $\tau_n \uparrow \infty$ and $x \in E$. By Theorem 5.17 (iii), passing to a subsequence, for any $\varepsilon > 0$ and $K \geq 0$ there exists $N_0 \geq 0$ such that

$$\|A(s + \tau_n) - A(s + \tau_m)\|_{\mathcal{L}(E_\alpha, E)} \leq \frac{\varepsilon}{K}, \qquad \text{for all } s \in \mathbb{R},\ m, n \geq N_0.$$

Since, by (5.27), $D((-A(s))^\alpha) \hookrightarrow E_\alpha$ with uniformly equivalent norms for all $s \in \mathbb{R}$ we have some $N \geq N_0$ with

$$\|A(s + \tau_n) - A(s + \tau_m)\|_{\mathcal{L}(D(A(s+\tau_n)^\alpha), E)} \leq \frac{\varepsilon}{K}, \qquad \text{for all } s \in \mathbb{R},\ m, n \geq N,$$

then by Lemmas 5.18 and 5.20, there exists sufficiently large K (independent of t, s) such that

$$\left\| e^{(t-s)A(s+\tau_n)} - e^{(t-s)A(s+\tau_m)} \right\| < \varepsilon, \qquad \text{for all } (t - s) \leq T_0, m, n \geq N.$$

Hence the map $(t, s) \mapsto e^{(t-s)A(s)}$ is jointly almost periodic, uniformly for $t \geq s$ bounded. ☐

Lemma 5.23. *Under the conditions of Theorem 5.17 , let the function $f(t, s)$ be one of $\{Q(t, s), g_1(t, s)\}$. Then the following hold:*

(i) $\|f(t, s)\| \leq C(t - s)^{\mu+\nu-2}$ for μ, ν as in (AT2) such that $\mu + \nu - 2 > -1$, for some $C \geq 0$ independent of t, s for $t - s$ bounded;

(ii) for any $0 < \delta < T_0$, $f(t, s)$ is jointly almost periodic, uniformly for $\delta < t - s \leq T_0$;

(iii) for any $\tau_n \uparrow \infty$ and $T_0 \geq 0$ we can pass to a subsequence such that for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for $m, n \geq N$ and $t - s \leq T_0$

$$\int_s^t \|f(t + \tau_n, r + \tau_n) - f(t + \tau_m, r + \tau_m)\| \,\mathrm{d}r \leq \varepsilon,$$
$$\int_s^t \|f(r + \tau_n, s + \tau_n) - f(r + \tau_m, s + \tau_m)\| \,\mathrm{d}r \leq \varepsilon.$$

Proof. Let $(\tau_n) \subset \mathbb{R}$, $\tau_n \uparrow \infty$, $x \in E$, $T_0 > 0$ and $\varepsilon > 0$. By Theorem 5.17 (iii), passing to a subsequence, for any $K \geq 0$ there exists $N \geq 0$ such that

$$\|A(t+\tau_n) - A(t+\tau_m)\|_{\mathcal{L}(E_\alpha,E)} \leq \frac{\varepsilon}{K}, \qquad \text{for all } t \in \mathbb{R},\ m,n \geq N. \tag{5.32}$$

Let $f(t,s) = Q(t,s)$. By [3, Lemma 2.3 (i)] there exists $C \geq 0$ such that

$$\|Q(t,s)\| \leq C(t-s)^{\mu+\nu-2},$$

where μ, ν from (AT2) satisfy $\mu+\nu-2 > -1$, so (i) holds and there exists $\delta_0 > 0$ such that if $|t-s| < \delta_0$ then

$$\begin{aligned}
\int_s^t \big\|Q(t+\tau_n,&r+\tau_n) - Q(t+\tau_m, r+\tau_m)\big\| \,\mathrm{d}r \\
&\leq \int_s^t 2C(t-r)^{\mu+\nu-2}\,\mathrm{d}r \;=\; \frac{(t-s)^{\mu+\nu-1}}{\mu+\nu-1} \;<\; \frac{\delta_0^{\mu+\nu-1}}{\mu+\nu-1} \\
&< \frac{\varepsilon}{2}.
\end{aligned} \tag{5.33}$$

So fix $s \leq t \in \mathbb{R}$ with $|t-s| \geq \delta_0$. Write $\Delta_n(t,s) := \big(A(t+\tau_n)^{-1} - A(s+\tau_n)^{-1}\big)$ and $\tau_0 = 0$, so

$$Q(t,s) = A(t)^2 e^{(t-s)A(t)}\Delta_0(t,s).$$

Then by the triangle inequality

$$\begin{aligned}
\big\|Q(t+\tau_n, s+\tau_n) &- Q(t+\tau_m, s+\tau_m)\big\| \\
&\leq \big\|A(t+\tau_n)^2\Big[e^{(t-s)A(t+\tau_n)}\Delta_n(t,s) - e^{(t-s)A(t+\tau_m)}\Delta_m(t,s)\Big]\big\| \\
&\qquad + \big\|\Big[A(t+\tau_n)^2 - A(t+\tau_m)^2\Big]e^{(t-s)A(t+\tau_m)}\Delta_m(t,s)\big\| \\
&\leq \big\|A(t+\tau_n)^2 e^{(t-s)A(t+\tau_n)}\Big[\Delta_n(t,s) - \Delta_m(t,s)\Big]\big\| \\
&\qquad + \big\|A(t+\tau_n)^2\Big[e^{(t-s)A(t+\tau_n)} - e^{(t-s)A(t+\tau_m)}\Big]\Delta_m(t,s)\big\| \\
&\qquad + \big\|\Big[A(t+\tau_n)^2 - A(t+\tau_m)^2\Big]e^{(t-s)A(t+\tau_m)}\Delta_m(t,s)\big\| \\
&=: I_1 + I_2 + I_3.
\end{aligned}$$

We consider the terms I_i in turn, starting with I_1.

$$\begin{aligned}
I_1 :&= \big\|A(t+\tau_n)^2 e^{(t-s)A(t+\tau_n)}\Big[\Delta_n(t,s) - \Delta_m(t,s)\Big]\big\| \\
&\leq \big\|A(t+\tau_n)^2 e^{(t-s)A(t+\tau_n)}\big\|_{\mathcal{L}(E_1,E)}\big\|\Delta_n(t,s) - \Delta_m(t,s)\big\|_{\mathcal{L}(E,E_1)} \\
&\lesssim \big\|A(t+\tau_n)^2 e^{(t-s)A(t+\tau_n)}\big\|_{\mathcal{L}(E)}\Big[\big\|A(t+\tau_n)^{-1} - A(t+\tau_m)^{-1}\big\|_{\mathcal{L}(E,E_1)} \\
&\qquad\qquad + \big\|A(s+\tau_n)^{-1} - A(s+\tau_m)^{-1}\big\|_{\mathcal{L}(E,E_1)}\Big].
\end{aligned}$$

Since $E_1 \hookrightarrow E_\alpha$, there exists $C > 0$ such that $\|x\|_{E_\alpha} \le C\|x\|_{E_1}$ for $x \in E_1$, then

$$\|A(t+\tau_n) - A(t+\tau_m)\|_{\mathcal{L}(E_1,E)} \le C\|A(t+\tau_n) - A(t+\tau_m)\|_{\mathcal{L}(E_\alpha,E)} < C\frac{\varepsilon}{K}$$

for all $t \in \mathbb{R}$, $m, n \ge N$. Now write $A = A(t+\tau_n)$, $B = A(t+\tau_m)$ and

$$\begin{aligned}
\|A^{-1} - B^{-1}\|_{\mathcal{L}(E,E_1)} &= \|A^{-1}BB^{-1} - A^{-1}AB^{-1}\|_{\mathcal{L}(E,E_1)} \\
&= \|A^{-1}(B-A)B^{-1}\|_{\mathcal{L}(E,E_1)} \\
&\le \|A^{-1}\|_{\mathcal{L}(E,E_1)}\|(B-A)\|_{\mathcal{L}(E_1,E)}\|B^{-1}\|_{\mathcal{L}(E,E_1)} \\
&\le \frac{CM^2\varepsilon}{K}.
\end{aligned}$$

So by (2.7),

$$I_1 \lesssim 2C_1(t-s)^{-2}\frac{CM^2\varepsilon}{K}$$

and hence for large enough K and $|t-s| \ge \delta_0$ we have

$$I_1 < \frac{\varepsilon}{6T_0}. \tag{5.34}$$

Now by uniform equivalence of graph norms, there exists $C \ge 0$ such that $A(t)A(s)^{-1}$ is bounded with $\|A(t)A(s)^{-1}\| \le C$ for all $r, s, t \in \mathbb{R}$, so

$$\begin{aligned}
I_2 &= \big\|A(t+\tau_n)^2\Big[e^{(t-s)A(t+\tau_n)} - e^{(t-s)A(t+\tau_m)}\Big]\Delta_m(t,s)\big\| \\
&\le \big\|A(t+\tau_n)\Big[e^{(t-s)A(t+\tau_n)} - e^{(t-s)A(t+\tau_m)}\Big]\big\|\big\|A(t+\tau_n)\Delta_m(t,s)\big\| \\
&\le C\big\|A(t+\tau_n)\Big[e^{(t-s)A(t+\tau_n)} - e^{(t-s)A(t+\tau_m)}\Big]\big\|.
\end{aligned}$$

Then by the Cauchy's formula representation of $e^{sA(t)}$ (2.8), for $A = A(t+\tau_n)$, $B = A(t+\tau_m)$ as above

$$\begin{aligned}
\big\|A\big[e^{(t-s)A} - e^{(t-s)B}\big]\big\| &\le \big\|Ae^{(t-s)A} - Be^{(t-s)B}\big\| + \big\|(B-A)e^{(t-s)B}\big\| \\
&\le \Big\|\frac{1}{2\pi i}\int_\Gamma \lambda e^{\lambda(t-s)}\Big[R(\lambda,A) - R(\lambda,B)\Big]\,\mathrm{d}\lambda\Big\| \\
&\qquad + \|B-A\|_{\mathcal{L}(E_\alpha,E)}\|e^{(t-s)B}\|_{\mathcal{L}(E,E_\alpha)} \\
&\le \Big\|\frac{1}{2\pi i}\int_\Gamma \lambda e^{\lambda(t-s)}R(\lambda,A)\big(B-A\big)R(\lambda,B)\,\mathrm{d}\lambda\Big\| \\
&\qquad + C_\alpha(t-s)^{-\alpha}\|B-A\|_{\mathcal{L}(E_\alpha,E)}
\end{aligned}$$

$$
\begin{aligned}
&\le C\|B-A\|_{\mathcal{L}(E_\alpha,E)}\left[\int_\Gamma \frac{|\lambda e^{\lambda(t-s)}|}{1+|\lambda|}\,\mathrm{d}\lambda + (t-s)^{-\alpha}\right] \\
&\le C\|B-A\|_{\mathcal{L}(E_\alpha,E)}\left[2\int_1^\infty e^{r(t-s)\cos\theta}\,\mathrm{d}r + 2\theta e^{(t-s)} + (t-s)^{-\alpha}\right] \\
&\le C\|B-A\|_{\mathcal{L}(E_\alpha,E)}\Big[\frac{2e^{(t-s)\cos\theta}}{(t-s)\cos\theta} + 2\theta e^{(t-s)} + (t-s)^{-\alpha}\Big]. \qquad (5.35)
\end{aligned}
$$

So for $|t-s| \in (\delta_0, T_0)$ and sufficiently large K we have

$$
I_2 < \frac{\varepsilon}{6T_0}. \qquad (5.36)
$$

$$
\begin{aligned}
I_3 &= \left\|\Big[A(t+\tau_n)^2 - A(t+\tau_m)^2\Big] e^{(t-s)A(t+\tau_m)}\Delta_m(t,s)\right\| \\
&\le \left\|A(t+\tau_n)^2 - A(t+\tau_m)^2\right\|_{\mathcal{L}(E_2,E)} \left\|e^{(t-s)A(t+\tau_m)}\Delta_m(t,s)\right\|_{\mathcal{L}(E,E_2)} \\
&\le \left\|A(t+\tau_n)^2 - A(t+\tau_m)^2\right\|_{\mathcal{L}(E_2,E)} C(t-s)^{-1},
\end{aligned}
$$

and since $E_\beta \hookrightarrow E_\alpha$ for all $\beta \ge \alpha$ we have for $A = A(t+\tau_n)$, $B = A(t+\tau_m)$ as above,

$$
\begin{aligned}
\|A^2 - B^2\|_{\mathcal{L}(E_2,E)} &\le \|BA - A^2\|_{\mathcal{L}(E_2,E)} + \|B^2 - BA\|_{\mathcal{L}(E_2,E)} \\
&\le \|B-A\|_{\mathcal{L}(E_1,E)}\|A\|_{\mathcal{L}(E_2,E_1)} + \|B-A\|_{\mathcal{L}(E_1,E)}\|B\|_{\mathcal{L}(E_2,E_1)} \\
&\le \frac{C\varepsilon}{K}.
\end{aligned}
$$

So for $|t-s| > \delta_0$ and sufficiently large K we have

$$
I_3 < \frac{\varepsilon}{6T_0}. \qquad (5.37)
$$

So (ii) holds, that is, there exists $\delta > 0$ such that $Q(t,s)$ is jointly almost periodic, uniformly for $(t-s) \in (\delta 6, T_0]$.

Now collecting (5.33) with (5.34), (5.36) and (5.37) we see that for $|t-s| \le T_0$

$$
\begin{aligned}
&\int_s^t \left\|Q(t+\tau_n, r+\tau_n) - Q(t+\tau_m, r+\tau_m)\right\| \mathrm{d}r \\
&\qquad = \int_s^{t-\delta_0} \left\|Q(t+\tau_n, r+\tau_n) - Q(t+\tau_m, r+\tau_m)\right\| \mathrm{d}r \\
&\qquad\qquad + \int_{t-\delta_0}^t \left\|Q(t+\tau_n, r+\tau_n) - Q(t+\tau_m, r+\tau_m)\right\| \mathrm{d}r \\
&\qquad < \int_s^{t-\delta_0} \frac{1}{T_0}\Big(\frac{\varepsilon}{6} + \frac{\varepsilon}{6} + \frac{\varepsilon}{6}\Big)\,\mathrm{d}r \quad + \quad \frac{\varepsilon}{2} \\
&\qquad < \varepsilon. \qquad (5.38)
\end{aligned}
$$

The same calculation shows that

$$\int_s^t \big\|Q(r+\tau_n, s+\tau_n) - Q(r+\tau_m, s+\tau_m)\big\| \,\mathrm{d}r < \varepsilon$$

for $|t-s| \le T_0$, so (iii) holds.

Now let $f(t,s) = g_1(t,s) = A(t)e^{(t-s)A(t)} - A(s)e^{(t-s)A(s)}$. From [3, Lemma 1.10 (i)] it follows that

$$\|g_1(t,s)\| \le C(t-s)^{\mu+\nu-2},$$

and since $\mu+\nu-2 > -1$, (i) holds and there exists $\delta_0 > 0$ such that for $|t-s| < \delta_0$

$$\int_s^t \|g_1(r,s)\| \,\mathrm{d}r \le \int_s^t C(r-s)^{\mu+\nu-2} \,\mathrm{d}r < \frac{\delta_0^{\mu+\nu-1}}{\mu+\nu-1} < \frac{\varepsilon}{2}. \tag{5.39}$$

By the resolvent calculation (5.35) we have

$$\|g_1(t+\tau_n, s+\tau_n) - g_1(t+\tau_m, s+\tau_m)\| \le \frac{C}{t-s}\Big[\|A(s+\tau_n) - A(s+\tau_m)\|_{\mathcal{L}(E_\alpha,E)} + \|A(t+\tau_n) - A(t+\tau_m)\|_{\mathcal{L}(E_\alpha,E)}\Big]$$

and so for large enough K and $|t-s| \ge \delta_0$ we have

$$\|g_1(t+\tau_n, s+\tau_n) - g_1(t+\tau_m, s+\tau_m)\| \le \frac{C\varepsilon}{K(t-s)} \le \frac{C\varepsilon}{K\delta_0} < \frac{\varepsilon}{2T_0}. \tag{5.40}$$

So (ii) holds, and combining (5.39) and (5.40) in the manner of (5.38), we see that

$$\int_s^t \big\|g_1(t+\tau_n, r+\tau_n)g_1(t+\tau_m, r+\tau_m)\big\| \,\mathrm{d}r < \varepsilon,$$

uniformly for $|t-s| \le T_0$ and sufficiently large m,n and y the same calculation, the same holds for (t,r) replaces by (r,s), so (iii) holds for g_1. □

Remark 5.24. it follows from (iii) of Lemma 5.23 that for $f \in \{G, g_1\}$, the maps $(t,s) \mapsto \int_s^t f(t,r)\,\mathrm{d}r$ and $(t,s) \mapsto \int_s^t f(r,s)\,\mathrm{d}r$ are jointly almost periodic in norm, uniformly for bounded $t-s$.

Lemma 5.25. *Suppose $f(t,s)$ and $g(t,s) : \{(t,s) \in \mathbb{R}^2 : s \le t\} \to \mathcal{L}(E)$ satisfy the following*

(a) there exists $k \in \mathbb{N}$ such that for any $T_0 \ge 0$ there exists a constant $C \ge 0$ such that $\|f(t,s)\| \le C(t-s)^{k(\mu+\nu-1)-1}$, and $\|g(t,s)\| \le C(t-s)^{\mu+\nu-2}$ (where μ,ν are as in (AT), so $\mu+\nu > 1$);

(b) for any $\tau_n \uparrow \infty$*,* $\varepsilon > 0$ *and* $T_0 > 0$ *we can pass to a subsequence such that*

(i) for any $\delta_0 > 0$*, there exists* $N_{\delta_0} \geq 0$ *such that for all* $m, n \geq N_{\delta_0}$ *and* $(t-s) \in (\delta_0, T_0]$ *both*

$$\|f(t+\tau_n, s+\tau_n) - f(t+\tau_m, s+\tau_m)\| \leq \varepsilon,$$
$$\|g(t+\tau_n, s+\tau_n) - g(t+\tau_m, s+\tau_m)\| \leq \varepsilon;$$

(ii) for the same subsequence there exists $N \geq 0$ *such that for* $m, n \geq N$ *and* $t - s \leq T_0$

$$\int_s^t \|f(t+\tau_n, r+\tau_n) - f(t+\tau_m, r+\tau_m)\| \,\mathrm{d}r \leq \varepsilon,$$
$$\int_s^t \|g(r+\tau_n, s+\tau_n) - g(r+\tau_m, s+\tau_m)\| \,\mathrm{d}r \leq \varepsilon.$$

Then the functions

$$G(t,s) = \int_s^t f(t,r)g(r,s)\,\mathrm{d}r \qquad \textit{and} \qquad H(t,s) = \int_s^t f(t,r)\,\mathrm{d}r$$

are bounded and satisfy

$$\|G(t+\tau_n, s+\tau_n) - G(t+\tau_m, s+\tau_m)\| < \varepsilon,$$
$$\|H(t+\tau_n, s+\tau_n) - H(t+\tau_m, s+\tau_m)\| < \varepsilon \tag{5.41}$$

for $m, n \geq N$ *uniformly for* $t - s \leq T_0$*.*

Proof. Let $(\tau_n) \subset \mathbb{R}$, $\tau_n \uparrow \infty$, $x \in E$, $T_0 > 0$ and $\varepsilon > 0$. By (a), there exists $\delta_0 > 0$ such that for $t - s < \delta_0$ we have

$$\int_s^t \|f(t,r)\|\,\mathrm{d}r < \frac{\varepsilon}{4} \quad \text{and} \quad \int_s^t \|g(r,s)\|\,\mathrm{d}r < \frac{\varepsilon}{4},$$

and there exists $C \geq 0$ such that $\|f(t,s)\|$, $\|g(t,s)\| \leq C$ for all $(t-s) \in (\delta_0, T_0)$.

By (b) there exists $N \geq 0$ such that for $m, n \geq N_1$ we have both

$$\|f(t+\tau_n, s+\tau_n) - f(t+\tau_m, s+\tau_m)\| \leq 1$$
$$\|g(t+\tau_n, s+\tau_n) - g(t+\tau_m, s+\tau_m)\| \leq 1,$$

and by (c) there exists $N \geq N_1$ such that for $m, n \geq N$ and $(t-s) > \delta_0$ we have both

$$\int_s^t \big\|f(t+\tau_n, r+\tau_n) - f(t+\tau_m, r+\tau_m)\big\|\,\mathrm{d}r < \frac{\varepsilon}{4CT_0}$$
$$\int_s^t \big\|g(r+\tau_n, s+\tau_n) - g(r+\tau_m, s+\tau_m)\big\|\,\mathrm{d}r < \frac{\varepsilon}{4CT_0}.$$

Then

$$
\begin{aligned}
&\|G(t+\tau_n, s+\tau_n) - G(t+\tau_m, s+\tau_m)\| \\
&\quad = \Big\| \int_s^t f(t+\tau_n, r+\tau_n) g(r+\tau_n, s+\tau_n) - f(t+\tau_m, r+\tau_m) g(r+\tau_m, s+\tau_m)\, \mathrm{d}r \Big\| \\
&\quad \le \Big\| \int_s^t f(t+\tau_n, r+\tau_n) \Big[g(r+\tau_n, s+\tau_n) - g(r+\tau_m, s+\tau_m) \Big]\, \mathrm{d}r \Big\| \\
&\qquad\qquad + \Big\| \int_s^t \Big[f(t+\tau_n, r+\tau_n) - f(t+\tau_m, r+\tau_m) \Big] g(r+\tau_m, s+\tau_m)\, \mathrm{d}r \Big\| \\
&\quad =: I_1 + I_2.
\end{aligned}
$$

We will deal with I_1 and I_2 separately. For $(t-s) \le T_0$

$$
\begin{aligned}
I_1 &\le \Big\| \int_s^{(t-\delta_0)\vee s} f(t+\tau_n, r+\tau_n) \Big[g(r+\tau_n, s+\tau_n) - g(r+\tau_m, s+\tau_m) \Big]\, \mathrm{d}r \Big\| \\
&\qquad + \Big\| \int_{(t-\delta_0)\vee s}^t f(t+\tau_n, r+\tau_n) \Big[g(r+\tau_n, s+\tau_n) - g(r+\tau_m, s+\tau_m) \Big]\, \mathrm{d}r \Big\| \\
&\le C \int_s^{(t-\delta_0)\vee s} \big\| g(r+\tau_n, s+\tau_n) - g(r+\tau_m, s+\tau_m) \big\|\, \mathrm{d}r + \int_{(t-\delta_0)\vee s}^t \big\| f(t+\tau_n, r+\tau_n) \big\| \cdot 1\, \mathrm{d}r \\
&< C \cdot \frac{\varepsilon(t-s)}{4CT_0} + \frac{\varepsilon}{4} \\
&\le \frac{\varepsilon}{2}.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
I_2 &\le \Big\| \int_s^{(t-\delta_0)\vee s} \big[f(t+\tau_n, r+\tau_n) - f(t+\tau_m, r+\tau_m) \big] g(r+\tau_m, s+\tau_m)\, \mathrm{d}r \Big\| \\
&\qquad + \Big\| \int_{(t-\delta_0)\vee s}^t \big[f(t+\tau_n, r+\tau_n) - f(t+\tau_m, r+\tau_m) \big] g(r+\tau_m, s+\tau_m)\, \mathrm{d}r \Big\| \\
&\le C \int_s^{(t-\delta_0)\vee s} \big\| f(t+\tau_n, r+\tau_n) - f(t+\tau_m, r+\tau_m) \big\|\, \mathrm{d}r + \int_{(t-\delta_0)\vee s}^t 1 \cdot \big\| g(r+\tau_m, s+\tau_m) \big\|\, \mathrm{d}r \\
&< C \cdot \frac{\varepsilon(t-s)}{4CT_0} + \frac{\varepsilon}{4} \\
&\le \frac{\varepsilon}{2}.
\end{aligned}
$$

Hence $G(t,s)$ is satisfies (5.41), and the same for $H(t,s)$ follows from a calculation identical to (5.38). □

For brevity, we write $\eta := \mu + \nu - 1 \in (0,1]$ for μ, ν as in (AT2).

Lemma 5.26 (Lemma 1.2 of [2])**.** *For $Q_k(t,s)$, $k \in \mathbb{N}$ as defined in* (5.29)*, there exists $C_Q \geq 0$ such that*

$$\|Q_k(t,s)\| \leq \frac{C_Q^k \Gamma(\eta)^k}{\Gamma(k\eta)}(t-s)^{k\eta-1} \tag{5.42}$$

for $(t-s) \leq T_0$.

Hence for $k \geq \frac{1}{\eta}$, $\|Q_k(t,s)\|$ is bounded for $(t-s) \leq T_0$ with

$$\|Q_k(t,s)\| \leq \frac{C_Q^k \Gamma(\eta)^k T_0^{k\eta-1}}{\Gamma(k\eta)}.$$

Proof. By [3, Lemma 2.3 (i)], there exists $C_Q \geq 0$ such that

$$\|Q_1(t,s)\| = \|Q(t,s)\| \leq C_Q(t-s)^{\eta-1}.$$

Suppose that (5.42) holds for some k, then

$$\begin{aligned}
\|Q_{k+1}(t,s)\| &= \left\| \int_s^t Q_k(t,r)Q(r,s)\,\mathrm{d}r \right\| \\
&\leq \int_s^t \frac{C_Q^k \Gamma(\eta)^k}{\Gamma(k\eta)}(t-r)^{k\eta-1} \cdot C_Q(r-s)^{\eta-1}\,\mathrm{d}r \\
&= \frac{C_Q^{k+1}\Gamma(\eta)^k}{\Gamma(k\eta)} \int_0^{t-s} (t-s-r)^{k\eta-1} r^{\eta-1}\,\mathrm{d}r \\
&= \frac{C_Q^{k+1}\Gamma(\eta)^k}{\Gamma(k\eta)} \int_0^1 (t-s)^{k\eta-1}(1-u)^{k\eta-1} \cdot (t-s)^{\eta-1}u^{\eta-1} \cdot (t-s)\,\mathrm{d}u \\
&= \frac{C_Q^{k+1}\Gamma(\eta)^k}{\Gamma(k\eta)}(t-s)^{(k+1)\eta-1} \int_0^1 (1-u)^{k\eta-1}u^{\eta-1}\,\mathrm{d}u \\
&= \frac{C_Q^{k+1}\Gamma(\eta)^k}{\Gamma(k\eta)}(t-s)^{(k+1)\eta-1} \cdot \frac{\Gamma(k\eta)\Gamma(\eta)}{\Gamma((k+1)\eta)} \\
&= \frac{C_Q^{k+1}\Gamma(\eta)^{k+1}}{\Gamma((k+1)\eta)}(t-s)^{(k+1)\eta-1}.
\end{aligned}$$

Hence (5.42) holds for $k+1$, and the result follows by induction. □

Write $M_k = \frac{C_Q^k \Gamma(\eta)^k}{\Gamma(k\eta)}$ and note that M_k converges to zero faster than any exponential sequence.

Corollary 5.27. *For the functions $Q_k(t,s)$, as defined in* (5.29)*, given a sequence $\tau_n \uparrow \infty$ and $\varepsilon > 0$ we can pass to a subsequence such that there exists $N \geq 0$ such that for all $k \geq 2$*

$$\|Q_k(t+\tau_n, s+\tau_n) - Q_k(t+\tau_m, s+\tau_m)\| < \frac{\varepsilon}{2^k} \tag{5.43}$$

for all $m, n \geq N$ and $(t-s) \leq T_0$.

Proof. Fix an $\varepsilon > 0$. By Lemma 5.26, for $k \geq \frac{1}{\eta}$, $Q_k(t,s)$ is bounded on $(t-s) \leq T_0$ with

$$\|Q_k(t,s)\| \leq M_k(t-s)^{k\eta-1} \leq M_k T_0^{k\eta-1}.$$

Now by the super-exponential growth of $\Gamma(k\eta)$, the sequence $2^{k+1}T_0^{k\eta-1}M_k$ converges to 0 for any $T_0 > 0$, so there exists $K_0 \geq \frac{1}{\eta}$ such that for $k \geq K_0$

$$T_0^{k\eta-1}M_k < \frac{\varepsilon}{2^{k+1}}.$$

Therefore

$$\|Q_k(t,s)\| < \frac{\varepsilon}{2^{k+1}}$$

for all $k \geq K_0$ and $(t-s) \leq T_0$, and so

$$\|Q_k(t+\tau_n, s+\tau_n) - Q_k(t+\tau_m, s+\tau_m)\| \leq \frac{\varepsilon}{2^k}$$

for any $m, n \in \mathbb{N}$.

Now let $2 \leq k < K_0$. By Lemma 5.23, $Q(t,s)$ satisfies conditions (a) and (b) of Lemma 5.25 so $Q_2(t,s)$ is jointly almost periodic, uniformly for $t-s$ bounded. It is easy to see that a function which is jointly almost periodic, uniformly for $t-s$ bounded satisfies (a) and (b) of Lemma 5.25 so by induction we have that $Q_k(t,s)$ is jointly almost periodic for all $2 \leq k < K_0$. Hence for any sequence $\tau_n \uparrow \infty$, we can pass (repeatedly, K_0-times) to a subsequence so that there exists $N \in \mathbb{N}$ such that

$$\|Q_k(t+\tau_n, s+\tau_n) - Q_k(t+\tau_m, s+\tau_m)\| \leq \frac{\varepsilon}{2^k}$$

for all $m, n \geq N$ and each $2 \leq k < K_0$. □

Remark 5.28. In the above, we have shown that for any sequence $\tau_n \uparrow \infty$ and any $\varepsilon > 0$, there exists a single subsequence and a single $N \geq 0$ such that the convergence in (5.43) holds for all $k \geq 2$.

We now begin to pull the above results together to construct $Z(t,s)$. Recall from (5.30) that

$$\begin{aligned} Z(t,s) := g_1(t,s) &+ \sum_{k=1}^{\infty} \int_s^t Q_k(t,r) g_1(r,s)\,\mathrm{d}r \\ &+ \sum_{k=1}^{\infty} \int_s^t \big[Q_k(t,r) - Q_n(t,s)\big] g_2(r,s)\,\mathrm{d}r \\ &+ \sum_{k=1}^{\infty} Q_k(t,s)\big[e^{(t-s)A(s)} - I\big] \end{aligned}$$

for $g_i(t,s)$, $i=1,2$ given by

$$g_1(t,s) = A(t)e^{(t-s)A(t)} - A(s)e^{(t-s)A(s)},$$
$$g_2(t,s) = A(s)e^{(t-s)A(s)}.$$

We will show that each term of $Z(t,s)$ satisfies conditions (a) and (b) of Lemma 5.25. The first term $g_1(s,t)$ already satisfies conditions (a) and (b) of Lemma 5.25 by Lemma 5.23.

Lemma 5.29. *The function*

$$G_1(t,s) := \sum_{k=1}^{\infty} \int_s^t Q_k(t,r) g_1(r,s)\,\mathrm{d}r$$

is such that for any $\tau_n \uparrow \infty$, $\varepsilon > 0$ and $T_0 \geq 0$ we can pass to a subsequence such that there exists $N \geq 0$ such that

$$\|G_1(t+\tau_n, s+\tau_n) - G_1(t+\tau_m, s+\tau_m)\| < \varepsilon$$

for $m,n \geq N$, uniformly for $(t-s) \leq T_0$.

Proof. Let $\tau_n \uparrow \infty$, $\varepsilon > 0$ and $T_0 > 0$. By Lemmas 5.23 and 5.26 we have

$$\|Q_k(t,s)\| \leq M_k (t-s)^{k\eta - 1} \quad \text{and} \quad \|g_1(t,s)\| \leq C_{g_1}(t-s)^{\eta-1}$$

and we can pass to a subsequence such that for any $\delta_0 > 0$ there exists $N_0 \in \mathbb{N}$ such that

$$\|g_1(t+\tau_n, s+\tau_n) - g_1(t+\tau_m, s+\tau_m)\| \leq \varepsilon$$

for all $(t-s) \in (\delta_0, T_0]$. By Corollary 5.27 we can pass to a further subsequence such that for some $N \geq N_0$

$$\|Q_k(t+\tau_n, s+\tau_n) - Q_k(t+\tau_m, s+\tau_m)\| \leq \frac{\varepsilon}{2^k} \tag{5.44}$$

for all $m,n \geq N$, $k \geq 2$ and $(t-s) \leq T_0$. Now

$$\begin{aligned}
&\left\| \int_s^t Q_k(t+\tau_n, s+\tau_n) g_1(r+\tau_n, s+\tau_n) - Q_k(t+\tau_m, s+\tau_m) g_1(r+\tau_m, s+\tau_m)\,\mathrm{d}r \right\| \\
&\quad \leq \int_s^t \|Q_k(t+\tau_n, r+\tau_n)\big[g_1(r+\tau_n, s+\tau_n) - g_1(r+\tau_m, s+\tau_m)\big]\|\,\mathrm{d}r \\
&\qquad + \int_s^t \|\big[Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_m, r+\tau_m)\big] g_1(r+\tau_m, s+\tau_m)\|\,\mathrm{d}r \\
&\quad =: I_1 + I_2.
\end{aligned}$$

We consider I_1 and I_2 separately. For $k \geq \frac{1}{\eta}$

$$\begin{aligned} I_1 &\leq \int_s^t M_k(t-r)^{k\eta-1} \cdot 2C_{g_1}(r-s)^{\eta-1}\,\mathrm{d}r \\ &= 2C_{g_1}M_k(t-s)^{\eta(k+1)-1}\int_0^1 (1-u)^{k\eta-1}u^{\eta-1}\,\mathrm{d}u \\ &= 2C_{g_1}M_k(t-s)^{\eta(k+1)-1}\frac{\Gamma(k\eta)\Gamma(\eta)}{\Gamma((k+1)\eta)} \\ &\leq 2C_{g_1}C_Q\frac{\Gamma(\eta)^{k+1}}{\Gamma((k+1)\eta)}T_0^{(k+1)\eta-1}. \end{aligned}$$

So by the super exponential growth of $\Gamma(k\eta)$, there exists $K \in \mathbb{N}$ such that

$$I_1 < \frac{\varepsilon}{2^{k+1}}$$

for all $k \geq K$. Now for $1 \leq k < K$, choose $\delta_0 > 0$ small enough so that

$$\int_s^{s+\delta_0} \|Q_k(t+\tau_n, r+\tau_n)\big[g_1(r+\tau_n, s+\tau_n) - g_1(r+\tau_m, s+\tau_m)\big]\|\,\mathrm{d}r < \frac{\varepsilon}{2^{K+2}}.$$

Then there exists $N_0 \in \mathbb{N}$ such that for all $(t-s) \in (\delta_0, T_0]$,

$$\|g_1(t+\tau_n, s+\tau_n) - g_1(t+\tau_m, s+\tau_m)\| < \frac{\varepsilon}{C_K 2^{K+2}}$$

for all $m, n \geq N_0$, where

$$C_K := \max_{k=1}^{K}\left\{\frac{M_k T_0^{k\eta}}{k\eta}\right\}.$$

Then

$$\begin{aligned} &\int_{s+\delta_0}^t \|Q_k(t+\tau_n, r+\tau_n)\big[g_1(r+\tau_n, s+\tau_n) - g_1(r+\tau_m, s+\tau_m)\big]\|\,\mathrm{d}r \\ &\qquad < \int_{s+\delta_0}^t M_k(t-r)^{k\eta-1}\cdot\frac{\varepsilon}{C_K 2^{K+2}}\,\mathrm{d}r \\ &\qquad \leq \frac{\varepsilon}{2^{K+2}} \end{aligned}$$

for all $m, n \geq N_0$. Hence we have

$$I_1 < \frac{\varepsilon}{2^{k+1}}$$

for all $k \in \mathbb{N}$, $m, n \geq N_0$.

Recall that by Lemma 5.26 we have

$$\|Q_k(t,s)\| \leq M_k(t-s)^{k\eta-1}$$

for $t-s \le T_0$ where $M_k \to 0$ faster than any exponential sequence. This means there exists $K_1 \ge 2$ such that

$$\|Q_k(t,s)\| \le \frac{1}{2} \cdot \frac{\varepsilon}{2^{k+1}} \cdot \frac{\eta}{C_{g_1} T_0^\eta}$$

for all $k \ge K_1$. Now choose $N_1 \ge N_0$ such that

$$\|Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_m, r+\tau_m)\| < \frac{\varepsilon}{2^{K+1}} \cdot \frac{\eta}{C_{g_1} T_0^\eta}$$

for all $m, n \ge N_1$ and $2 \le k < K_1$. Hence we have an N_1 independent of k such that

$$\|Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_m, r+\tau_m)\| < \frac{\varepsilon}{2^{k+1}} \cdot \frac{\eta}{C_{g_1} T_0^\eta} \tag{5.45}$$

for all $m, n \ge N_1$ and $k \ge 2$. Then

$$\begin{aligned} I_2 &= \int_s^t \|\big[Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_m, r+\tau_m)\big] g_1(r+\tau_m, s+\tau_m)\| \,\mathrm{d}r \\ &< \frac{\varepsilon}{2^{k+1}} \cdot \frac{\eta}{C_{g_1} T_0^\eta} \int_s^t C_{g_1}(r-s)^{\eta-1} \,\mathrm{d}r \\ &\le \frac{\varepsilon}{2^{k+1}} \end{aligned} \tag{5.46}$$

for all $k \ge 2$. For the $k = 1$ case, choose $\delta_1 > 0$ such that

$$\int_{t-\delta_1}^t \|\big[Q(t+\tau_n, r+\tau_n) - Q(t+\tau_m, r+\tau_m)\big] g_1(r+\tau_m, s+\tau_m)\| \,\mathrm{d}r < \frac{\varepsilon}{8}$$

then by Lemma 5.23, choose $N \ge N_1$ such that

$$\|Q(t+\tau_n, s+\tau_n) - Q(t+\tau_m, s+\tau_m)\| \le \frac{\varepsilon}{8} \frac{\eta}{C_{g_1} T_0^\eta}$$

for all $n, m \ge N$ and $(t-s) \in (\delta_1, T_0]$. So

$$\begin{aligned} I_2 &\le \int_s^{t-\delta_1} \|\big[Q(t+\tau_n, r+\tau_n) - Q(t+\tau_m, r+\tau_m)\big] g_1(r+\tau_m, s+\tau_m)\| \,\mathrm{d}r \\ &\qquad + \int_{t-\delta_1}^t \|\big[Q(t+\tau_n, r+\tau_n) - Q(t+\tau_m, r+\tau_m)\big] g_1(r+\tau_m, s+\tau_m)\| \,\mathrm{d}r \\ &< \int_s^{t-\delta_1} \frac{\varepsilon}{8} \frac{\eta}{C_{g_1} T_0^\eta} \cdot C_{g_1}(r-s)^{\eta-1} \,\mathrm{d}r + \frac{\varepsilon}{8} \\ &\le \frac{\varepsilon}{4}. \end{aligned}$$

So we have now that for all $k \ge 1$

$$I_2 < \frac{\varepsilon}{2^{k+1}}. \tag{5.47}$$

Then by (5.45) and (5.47) we have

$$\Big\| \int_s^t Q_k(t+\tau_n, s+\tau_n) g_1(r+\tau_n, s+\tau_n) - Q_k(t+\tau_m, s+\tau_m) g_1(r+\tau_m, s+\tau_m)\,\mathrm{d}r \Big\| < \frac{\varepsilon}{2^k}$$

for all $k \in \mathbb{N}$, and so

$$\|G_1(t+\tau_n, s+\tau_n) - G_1(t+\tau_m, s+\tau_m)\| \le \varepsilon \sum_{k=1}^{\infty} 2^{-k} = \varepsilon. \qquad \square$$

Lemma 5.30. *The function*

$$G_2(t,s) := \sum_{k=0}^{\infty} \int_s^t \big[Q_k(t,r) - Q_k(t,s)\big] g_2(r,s)\,\mathrm{d}r$$

is locally integrable in each variable and for $\tau_n \uparrow \infty$ and $\varepsilon > 0$ we can pass to a subsequence such that there exists $N \in \mathbb{N}$ such that

$$\Big\| \int_s^t G_2(r+\tau_n, s+\tau_n) - G_2(r+\tau_m, s+\tau_m)\,\mathrm{d}r \Big\| < \varepsilon$$

for all $m, n \ge N$ and $(t-s) \le T_0$.

Proof. Let $\tau_n \uparrow \infty$, $\varepsilon > 0$ and $T_0 > 0$. By [2, Lemma 1.2 (iii)] we have that for some $C_0 \ge 1$,

$$\int_s^t \frac{\|Q_k(t,r) - Q_k(t,s)\|}{r-s}\,\mathrm{d}r \le C_0 \frac{C_Q^{k-1}\Gamma(\eta)^{k-1}}{\Gamma((k-1)\eta)} (t-s)^{k\eta-1} = C_0 M_{k-1} (t-s)^{k\eta-1} \quad (5.48)$$

for all $(t-s) \le T_0$ and $k \ge 2$. By (2.7),

$$\|g_2(t,s)\| \le C_{g_2}(t-s)^{-1}$$

so for $k \ge 2$

$$\int_s^t \|\big[Q_k(t,r) - Q_k(t,s)\big] g_2(r,s)\|\,\mathrm{d}r \le C_0 C_{g_2} M_{k-1} (t-s)^{k\eta-1}. \quad (5.49)$$

The next estimate comes from [1, Lemma 2.1 (iii)]; for any $\beta \in (0,\eta)$ there exists $C_\beta \ge 0$ such that

$$\|Q_k(t,r) - Q_k(t,s)\| \le C_\beta (r-s)^\beta (t-r)^{\eta-\beta-1} \quad (5.50)$$

for all $k \in \mathbb{N}$, $s \le r \le t$ with $t - s \le T_0$. So fix $\beta \in (0, \eta)$, then for $k = 1$

$$\begin{aligned}
\int_s^t \left\| \big[Q(t,r) - Q(t,s)\big] g_2(r,s) \right\| \mathrm{d}r &\le \int_s^t C_\beta (r-s)^\beta (t-r)^{\eta-\beta-1} \cdot C_{g_2}(r-s)^{-1} \,\mathrm{d}r \\
&= C_\beta C_{g_2} (t-s)^{\eta-1} \int_0^1 (1-r)^{\eta-\beta-1} r^{\beta-1} \,\mathrm{d}r \\
&= C_\beta C_{g_2} (t-s)^{\eta-1} \frac{\Gamma(\eta-\beta)\Gamma(\beta)}{\Gamma(\eta)} \\
&=: C_0 C_{g_2} M_0 (t-s)^{\eta-1}
\end{aligned} \tag{5.51}$$

for $M_0 := \frac{C_\beta}{C_0} \cdot \frac{\Gamma(\eta-\beta)\Gamma(\beta)}{\Gamma(\eta)}$. Now if

$$F_j(r) := \sum_{k=1}^{j} C_0 C_{g_2} M_{k-1} r^{k\eta-1},$$

then F_j is integrable on $[0, T_0]$ for all $j \ge 1$ and $\big(F_j\big)$ is non-decreasing in j, so by the monotone convergence theorem, using the fact that

$$\sum_{k=1}^{\infty} C_0 C_{g_2} M_{k-1} \int_0^{T_0} r^{k\eta-1} \,\mathrm{d}r$$

is convergent, we have that $F(r) := \lim_{j\to\infty} F_j(r)$ is integrable and then by dominated convergence $G_2(t,s)$ is integrable in each variable for $(t-s) \le T_0$.

Now

$$\begin{aligned}
\int_s^t \Big\| &\big[Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_n, s+\tau_n)\big] g_2(r+\tau_n, s+\tau_n) \\
&\quad - \big[Q_k(t+\tau_m, r+\tau_m) - Q_k(t+\tau_m, s+\tau_m)\big] g_2(r+\tau_m, s+\tau_m) \Big\| \,\mathrm{d}r \\
&\le \int_s^t \Big\| \big[Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_n, s+\tau_n)\big] \\
&\qquad \cdot \big[g_2(r+\tau_n, s+\tau_n) - g_2(r+\tau_m, s+\tau_m)\big] \Big\| \,\mathrm{d}r \\
&\quad + \int_s^t \Big\| Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_n, s+\tau_n) - Q_k(t+\tau_m, r+\tau_m) \\
&\qquad + Q_k(t+\tau_m, s+\tau_m) \Big\| \left\| g_2(r+\tau_m, s+\tau_m) \right\| \,\mathrm{d}r \\
&=: I_1 + I_2.
\end{aligned} \tag{5.52}$$

We consider I_1 and I_2 separately. By (5.48) and (5.51),

$$I_1(t,s) \le 2C_0 M_{k-1} (t-s)^{k\eta-1},$$

for all $k \in \mathbb{N}$, so since $2^{k+2} T_0^{k\eta-1} M_{k-1} \to 0$, there exists $K \ge \frac{1}{\eta}$ such that

$$I_1(t,s) < \frac{\varepsilon}{2^{k+1} T_0} \tag{5.53}$$

for all $k \geq K$. Now fix $1 \leq k < K$, by (5.53) there exists $\delta_k > 0$ such that if $(t-s) < \delta_k$ then

$$\int_s^t I_1(r,s)\,\mathrm{d}r < \frac{\varepsilon}{2^{k+2}}. \tag{5.54}$$

Let $\delta_0 := \min_{1\leq k<K}\big\{\delta_k\big\}$. By the calculation (5.35), we have that passing to a subsequence, there exists $N_0 \in \mathbb{N}$ such that for $m, n \geq N_0$ and $(t-s) \leq T_0$

$$\|g_2(r+\tau_n, s+\tau_n) - g_2(r+\tau_m, s+\tau_m)\| \leq C_2\Big(1 + \frac{1}{r-s}\Big)\frac{\varepsilon}{C_K},$$

where

$$C_K := 2^{K+2}T_0 C_\beta C_2\Big[\delta_0^{\eta-1}M_0 + T_0^\eta\frac{\Gamma(\eta-\beta)\Gamma(\beta+1)}{\Gamma(\eta+1)}\Big]$$

By (5.50), if $(t-s) \in (\delta_0, T_0]$ then

$$\begin{aligned}
I_1 &< \int_s^t \|Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_n, s+\tau_n)\| \cdot C_2\Big(1+\frac{1}{r-s}\Big)\frac{\varepsilon}{C_K}\,\mathrm{d}r \\
&\leq \int_s^t C_\beta(r-s)^\beta(t-r)^{\eta-\beta-1}\cdot C_2\Big(1+\frac{1}{r-s}\Big)\frac{\varepsilon}{C_K}\,\mathrm{d}r \\
&= C_\beta C_2\frac{\varepsilon}{C_K}\left[(t-s)^{\eta-1}\int_0^1(1-r)^{\eta-\beta-1}r^{\beta-1}\,\mathrm{d}r + (t-s)^\eta\int_0^1(1-r)^{\eta-\beta-1}r^\beta\,\mathrm{d}r\right] \\
&= C_\beta C_2\frac{\varepsilon}{C_K}\Big[(t-s)^{\eta-1}\frac{\Gamma(\eta-\beta)\Gamma(\beta)}{\Gamma(\eta)} + (t-s)^\eta\frac{\Gamma(\eta-\beta)\Gamma(\beta+1)}{\Gamma(\eta+1)}\Big] \\
&\leq C_\beta C_2\frac{\varepsilon}{C_K}\Big[\delta_0^{\eta-1}M_0 + T_0^\eta\frac{\Gamma(\eta-\beta)\Gamma(\beta+1)}{\Gamma(\eta+1)}\Big] \\
&\leq \frac{\varepsilon}{2^{K+2}T_0}.
\end{aligned}$$

for all $1 \leq k < K$. So together with (5.54) we have that

$$\begin{aligned}
\int_s^t I_1(r,s)\,\mathrm{d}r &= \int_s^{s+\delta_0} I_1(r,s)\,\mathrm{d}r + \int_{s+\delta_0}^t I_1(r,s)\,\mathrm{d}r \\
&< \frac{\varepsilon}{2^{k+2}} + \int_{s+\delta_0}^t \frac{\varepsilon}{2^{k+2}T_0}\,\mathrm{d}r \\
&\leq \frac{\varepsilon}{2^{k+1}}
\end{aligned} \tag{5.55}$$

for all $k \in \mathbb{N}$. Now for I_2. Recall from (5.49) and (5.51)

$$\int_s^t \|\big[Q_k(t,r) - Q_k(t,s)\big]g_2(r,s)\|\,\mathrm{d}r \leq C_0 C_{g_2} M_{k-1}(t-s)^{k\eta-1}$$

for all $k \geq 1$, so since $2^k T_0^{k\eta} M_k \to 0$ as $k \to \infty$ there exists $K \geq \frac{1}{\eta}$ such that

$$\begin{aligned} I_2 = \int_s^t \big\| & Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_n, s+\tau_n) - Q_k(t+\tau_m, r+\tau_m) \\ & + Q_k(t+\tau_m, s+\tau_m)\big\| \|g_2(r+\tau_m, s+\tau_m)\| \,\mathrm{d}r \\ < \frac{\varepsilon}{2^{k+1}T_0} & \end{aligned} \tag{5.56}$$

for all $k \geq K$, $t - s \leq T_0$. Now for $1 \leq k < K$, recall from (5.50) that for $\beta \in (0, \eta)$ there exists $C_\beta \geq 0$ such that

$$\|Q_k(t,r) - Q_k(t,s)\| \leq C_\beta (r-s)^\beta (t-r)^{\eta-\beta-1}$$

for all $s \leq r \leq t$ and $(t-s) \leq T_0$. Hence for $0 < \delta < (t-s)$

$$\begin{aligned} \int_s^{s+\delta} \frac{\|Q_k(t,r) - Q_k(t,s)\|}{r-s} \,\mathrm{d}r &\leq \int_s^{s+\delta} C_3 (r-s)^{\beta-1}(t-r)^{\eta-\beta-1} \,\mathrm{d}r \\ &\leq C_3 (t-s)^{\eta-1} \int_0^{\frac{\delta}{t-s}} r^{\beta-1}(1-r)^{\eta-\beta-1} \,\mathrm{d}r. \end{aligned}$$

For $r \leq \frac{1}{2}$, we have $(1-r)^{\eta-\beta-1} \leq \left(\frac{1}{2}\right)^{\eta-\beta-1}$, and so if $\frac{\delta}{t-s} \leq \frac{1}{2}$ then

$$\int_0^{\frac{\delta}{t-s}} r^{\beta-1}(1-r)^{\eta-\beta-1} \,\mathrm{d}r \leq \int_0^{\frac{\delta}{t-s}} \frac{r^{\beta-1}}{2^{\eta-\beta-1}} \,\mathrm{d}r = \frac{(t-s)^{-\beta}}{\beta \cdot 2^{\eta-\beta-1}} \cdot \delta^\beta$$

and therefore

$$\int_s^{s+\delta} \frac{\|Q_k(t,r) - Q_k(t,s)\|}{r-s} \,\mathrm{d}r \leq C_6 \delta^\beta (t-s)^{\eta-\beta-1}$$

for all $(t-s) \in (2\delta, T_0)$. Consider

$$\int_s^t I_2(r,s) \,\mathrm{d}r = \int_s^{s+2\delta} I_2(r,s) \,\mathrm{d}r + \int_{s+2\delta}^t I_2(r,s) \,\mathrm{d}r. \tag{5.57}$$

By (5.51) we have

$$\int_s^{s+2\delta} I_2(r,s) \,\mathrm{d}r \leq \int_s^{s+2\delta} C_4 (r-s)^{\eta-1} \,\mathrm{d}r = \frac{C_4}{\eta}(2\delta)^\eta, \tag{5.58}$$

for some $C_4 \geq 0$, so choose $\delta_0 > 0$ such that

$$C_6 \delta_0^\beta < \frac{\varepsilon(\eta-\beta)}{2^{K+2} C_{g_2} T_0^{\eta-\beta}} \quad \text{and} \quad \frac{C_4}{\eta}(2\delta_0)^\eta < \frac{\varepsilon}{2^{K+1}}.$$

Then for $(t-s) \in (2\delta_0, T_0)$

$$\int_s^{s+\delta_0} \big\|Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_n, s+\tau_n) - Q_k(t+\tau_m, r+\tau_m) + Q_k(t+\tau_m, s+\tau_m)\big\| \|g_2(r+\tau_m, s+\tau_m)\| \,\mathrm{d}r$$

$$\begin{aligned}
&\leq \int_s^{s+\delta_0} C_{g_2} \frac{\|Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_n, s+\tau_n)\|}{r-s} + C_{g_2} \frac{\|Q_k(t+\tau_m, r+\tau_m) - Q_k(t+\tau_m, s+\tau_m)\|}{r-s} \,\mathrm{d}r \\
&< \frac{\varepsilon(\eta-\beta)}{2^{K+2} T_0^{\eta-\beta}} (t-s)^{\eta-\beta-1}. \qquad (5.59)
\end{aligned}$$

Now by Corollary 5.27, we can pass to a further subsequence of (τ_n) such that there exists $N \in \mathbb{N}$ such that for all $1 \leq k < K$, $t - s \leq T_0$ and $m, n \geq N$

$$\|Q_k(t+\tau_n, s+\tau_n) - Q_k(t+\tau_m, s+\tau_m)\| < \frac{\varepsilon \delta_0}{2^{K+2} C_{g_2} T_0^2}. \qquad (5.60)$$

Then for $m, n \geq N$ and $(t-s) \in (2\delta_0, T_0)$

$$\begin{aligned}
I_2 &= \int_s^t \big\|Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_n, s+\tau_n) - Q_k(t+\tau_m, r+\tau_m) + Q_k(t+\tau_m, s+\tau_m)\big\| \|g_2(r+\tau_m, s+\tau_m)\| \,\mathrm{d}r \\
&\leq C_{g_2} \Bigg[\int_s^{s+\delta_0} \frac{\|Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_n, s+\tau_n)\|}{r-s} + \frac{\|Q_k(t+\tau_m, r+\tau_m) - Q_k(t+\tau_m, s+\tau_m)\|}{r-s} \,\mathrm{d}r \\
&\qquad + \int_{s+\delta_0}^t \frac{\|Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_m, r+\tau_m)\|}{r-s} + \frac{\|Q_k(t+\tau_n, s+\tau_n) - Q_k(t+\tau_m, s+\tau_m)\|}{r-s} \,\mathrm{d}r \Bigg] \\
&< \frac{\varepsilon(\eta-\beta)}{2^{K+2} T_0^{\eta-\beta}} (t-s)^{\eta-\beta-1} + 2 \int_{s+\delta_0}^t \frac{1}{\delta_0} \cdot \frac{\varepsilon \delta_0}{2^{K+2} T_0^2} \,\mathrm{d}r \\
&\leq \frac{\varepsilon}{2^{K+2}} \left(\frac{(\eta-\beta)(t-s)^{\eta-\beta-1}}{T_0^{\eta-\beta}} + \frac{1}{T_0} \right) \qquad (5.61)
\end{aligned}$$

where the second inequality follows from (5.59) and (5.60). Putting (5.58) and (5.61) into (5.57) gives us that

$$\begin{aligned}\int_s^t I_2(r,s)\,\mathrm{d}r &< \int_s^{s+2\delta_0} C_4(r-s)^{\eta-1}\,\mathrm{d}r + \int_{s+2\delta_0}^t \frac{\varepsilon}{2^{K+1}}\left(\frac{(\eta-\beta)(r-s)^{\eta-\beta-1}}{T_0^{\eta-\beta}} + \frac{1}{T_0}\right)\mathrm{d}r \\ &< \frac{\varepsilon}{2^{K+1}} + \frac{\varepsilon}{2^{K+2}}\left[\frac{(t-s)^{\eta-\beta}}{T_0^{\eta-\beta}} + \frac{(t-s)}{T_0}\right] \\ &= \frac{\varepsilon}{2^{K+1}}. \end{aligned} \tag{5.62}$$

Putting (5.56) and (5.62) together we get that

$$\int_s^t I_2(r,s)\,\mathrm{d}r < \frac{\varepsilon}{2^{k+1}}$$

for all $k \in \mathbb{N}$. Then using (5.52) and (5.55) we have

$$\begin{aligned}\int_s^t &\|\big[Q_k(t+\tau_n, r+\tau_n) - Q_k(t+\tau_n, s+\tau_n)\big]g_2(r+\tau_n, s+\tau_n) \\ &\qquad - \big[Q_k(t+\tau_m, r+\tau_m) - Q_k(t+\tau_m, s+\tau_m)\big]g_2(r+\tau_m, s+\tau_m)\|\,\mathrm{d}r \\ &< \frac{\varepsilon}{2^k}\end{aligned}$$

for all $k \in \mathbb{N}$. Finally, by absolute convergence we can exchange the sum and integral to get

$$\begin{aligned}&\left\|\int_s^t G_2(r+\tau_n, s+\tau_n) - G_2(r+\tau_m, s+\tau_m)\,\mathrm{d}r\right\| \\ &\le \left\|\int_s^t \sum_{k=1}^{\infty}\int_s^r \big[Q_k(r+\tau_n, u+\tau_n) - Q_k(r+\tau_n, s+\tau_n)\big]g_2(u+\tau_n, s+\tau_n)\right. \\ &\qquad \left. - \big[Q_k(r+\tau_m, u+\tau_m) - Q_k(r+\tau_m, s+\tau_m)\big]g_2(u+\tau_m, s+\tau_m)\,\mathrm{d}u\,\mathrm{d}r\right\| \\ &\le \sum_{k=1}^{\infty}\int_s^t I_1(r,s) + I_2(r,s)\,\mathrm{d}r \\ &< \varepsilon\sum_{k=1}^{\infty} 2^{-k} = \varepsilon\end{aligned}$$

for all sufficiently large m, n. □

Now for the final term of $Z(t,s)$.

Lemma 5.31. *The function*

$$G_3(t,s) := \sum_{k=0}^{\infty} Q_k(t,s)\big[e^{(t-s)A(s)} - I\big]$$

is locally integrable in each variable and for $\tau_n \uparrow \infty$ and $\varepsilon > 0$ we can pass to a subsequence such that there exists $N \in \mathbb{N}$ such that

$$\Big\| \int_s^t G_3(r+\tau_n, s+\tau_n) - G_3(r+\tau_m, s+\tau_m)\,\mathrm{d}r \Big\| < \varepsilon$$

for all $m, n \geq N$ and $(t-s) \leq T_0$.

Proof. Let $\tau_n \uparrow \infty$, $\varepsilon > 0$ and $T_0 > 0$.

$$\begin{aligned}
\int_s^t &\|Q_k(r+\tau_n, s+\tau_n)e^{(r-s)A(s+\tau_n)} - Q_k(r+\tau_m, s+\tau_m)e^{(r-s)A(s+\tau_m)}\|\,\mathrm{d}r \\
&\leq \int_s^t \|\big[Q_k(r+\tau_n, s+\tau_n) - Q_k(r+\tau_m, s+\tau_m)\big]e^{(r-s)A(s+\tau_n)}\|\,\mathrm{d}r \\
&\qquad + \int_s^t \|Q_k(r+\tau_m, s+\tau_m)\big[e^{(r-s)A(s+\tau_n)} - e^{(r-s)A(s+\tau_m)}\big]\|\,\mathrm{d}r \\
&=: I_1 + I_2.
\end{aligned}$$

As usual, we consider I_1 and I_2 separately. Let $M := \sup_{s\in\mathbb{R}} \sup_{t\in[0,T_0]} \|e^{tA(s)}\|$, which exists by (AT1), then

$$I_1 \leq M \int_s^t \|Q_k(r+\tau_n, s+\tau_n) - Q_k(r+\tau_m, s+\tau_m)\|\,\mathrm{d}r$$

Now for $k = 1$ by Lemma 5.23, we can pass to a subsequence such that there exists $N_0 \in \mathbb{N}$ with

$$\int_s^t \|Q(r+\tau_n, s+\tau_n) - Q(r+\tau_m, s+\tau_m)\|\,\mathrm{d}r < \frac{\varepsilon}{4M}$$

for all $m, n \geq N_0$ and $(t-s) \leq T_0$. Then by Corollary 5.27, we can pass to a further subsequence and find an $N_1 \geq N_0$ such that

$$\|Q_k(r+\tau_n, s+\tau_n) - Q_k(r+\tau_m, s+\tau_m)\| \leq \frac{\varepsilon}{2^{k+1}MT_0}$$

for all $k \geq 2$, $m, n \geq N_1$ and $(t-s) \leq T_0$. Hence we have

$$I_1 < \frac{\varepsilon}{2^{k+1}} \tag{5.63}$$

for all $k \in \mathbb{N}$, $m, n \geq N_1$ and $(t-s) \leq T_0$. Now for I_2, recall from Lemma 5.26 that

$$\|Q_k(t,s)\| \leq M_k(t-s)^{k\eta-1}.$$

Define

$$M_\infty := \sup_{k\in\mathbb{N}} \left\{ \frac{2^{k+1}M_k T_0^{k\eta}}{k\eta} \right\},$$

then $M_\infty < \infty$ by the super-exponential decay of M_k. By Corollary 5.22, we have that for a further subsequence and some $N_2 \geq N_1$

$$\|e^{(t-s)A(s+\tau_n)} - e^{(t-s)A(s+\tau_m)}\| < \frac{\varepsilon}{M_\infty}$$

for all $m, n \geq N_2$. Then

$$\begin{aligned} I_2 &\leq \int_s^t M_k (r-s)^{k\eta-1} \|e^{(r-s)A(s+\tau_n)} - e^{(r-s)A(s+\tau_m)}\| \,\mathrm{d}r \\ &< \int_s^t M_k (r-s)^{k\eta-1} \frac{\varepsilon}{M_\infty} \,\mathrm{d}r \\ &\leq M_k \frac{(t-s)^{k\eta}}{k\eta} \cdot \varepsilon \frac{k\eta}{2^{k+1} M_k T_0^{k\eta}} \\ &\leq \frac{\varepsilon}{2^{k+1}} \end{aligned} \tag{5.64}$$

for all $k \in \mathbb{N}$ and $(t-s) \leq T_0$. Collecting (5.63) and (5.64) we get

$$\begin{aligned} &\left\| \int_s^t G_3(r+\tau_n, s+\tau_n) - G_3(r+\tau_m, s+\tau_m) \,\mathrm{d}r \right\| \\ &\quad \leq \int_s^t \sum_{k=1}^\infty \left\| Q_k(r+\tau_n, s+\tau_n) e^{(r-s)A(s+\tau_n)} - Q_k(r+\tau_m, s+\tau_m) e^{(r-s)A(s+\tau_m)} \right\| \mathrm{d}r \\ &\quad = \sum_{k=1}^\infty \int_s^t \left\| Q_k(r+\tau_n, s+\tau_n) e^{(r-s)A(s+\tau_n)} - Q_k(r+\tau_m, s+\tau_m) e^{(r-s)A(s+\tau_m)} \right\| \mathrm{d}r \\ &\quad < \varepsilon \sum_{k=1}^\infty 2^{-k} = \varepsilon. \end{aligned}$$

□

Now we are ready to prove Theorem 5.17.

Proof of Theorem 5.17. Recall that from [1],

$$\begin{aligned} P(t,s) &= e^{(t-s)A(s)} + \int_s^t Z(r,s) \,\mathrm{d}r \\ &= e^{(t-s)A(s)} + \int_s^t \Big[g_1(r,s) + G_1(r,s) + G_2(r,s) + G_3(r,s) \Big] \,\mathrm{d}r. \end{aligned}$$

Choose a sequence $\tau_n \uparrow \infty$ and $\varepsilon > 0$, $T_0 \geq 0$. By the following results for each term we can pass to a subsequence such that there exists $N \geq 0$ such that

$$\|P(t+\tau_n, s+\tau_n) - P(t+\tau_m, s+\tau_m)\| \leq \varepsilon$$

for all $m, n \geq N$, $t - s \leq T_0$:

- $e^{(t-s)A(s)}$ – Lemma 5.22;
- $\int_s^t g_1(r,s)\,\mathrm{d}r$ – Lemma 5.23 (iii);
- $\int_s^t G_1(r,s)\,\mathrm{d}r$ – Lemma 5.29;
- $\int_s^t G_2(r,s)\,\mathrm{d}r$ – Lemma 5.30;
- $\int_s^t G_3(r,s)\,\mathrm{d}r$ – Lemma 5.31.

We now construct (by diagonalisation) a single subsequence $(\sigma_n) \subseteq (\tau_n)$ independent of ε such that $\big(P(t+\sigma_n, s+\sigma_n)\big)$ converges uniformly for $t-s \leq T_0$.

By the above (possibly discarding the first N terms), we can pass to a subsequence $\big(\tau_n^1\big) \subseteq \big(\tau_n\big)$ such that

$$\|P(t+\tau_n^1, s+\tau_n^1) - P(t+\tau_m^1, s+\tau_m^1)\| \leq \frac{1}{2^{k+1}}$$

for all $n, m \geq 1$, $t-s \leq T_0$. Write $\sigma_1 := \tau_1^1$. Then by applying the above again to the sequence $\big(\tau_n^1\big)$ we can find a subsequence $\big(\tau_n^2\big) \subseteq \big(\tau_n^1\big)$ such that

$$\|P(t+\tau_n^2, s+\tau_n^2) - P(t+\tau_m^2, s+\tau_m^2)\| \leq \frac{1}{2^{k+2}}$$

for all $n, m \geq 2$, $t-s \leq T_0$. Write $\sigma_2 := \tau_2^2 \geq \tau_1^2 \geq \tau_1^1 = \sigma_1$. Continue in this way by induction, at each step finding a subsequence $\big(\tau_n^{\ell+1}\big) \subseteq \big(\tau_n^{\ell}\big)$ such that

$$\|P(t+\tau_n^{\ell+1}, s+\tau_n^{\ell+1}) - P(t+\tau_m^{\ell+1}, s+\tau_m^{\ell+1})\| \leq \frac{1}{2^{k+\ell+1}}$$

for all $n, m \geq \ell+1$, $t-s \leq T_0$ and set $\sigma_{\ell+1} := \tau_{\ell+1}^{\ell+1} \geq \sigma_\ell$.

Then we have constructed a subsequence $(\sigma_n) \subseteq (\tau_n)$ such that for any $\varepsilon > 0$, we have

$$\|P(t+\sigma_n, s+\sigma_n) - P(t+\sigma_m, s+\sigma_m)\| < \frac{\varepsilon}{2^k}$$

for all $m, n \geq N$, $t-s \leq T_0$ where N is such that $2^{-N} < \varepsilon$.

□

5.6 Example

Let $p \in (1,\infty)$, $\mathcal{O}$ be an open and bounded domain in $\mathbb{R}^d$ with C^2-boundary, and consider the following stochastic differential equation:

$$\begin{cases} \dfrac{\partial}{\partial t}u(t,\xi) &= A(t,\xi)u(t,\xi) + f(t,\xi) + \displaystyle\sum_{n=1}^{\infty} g_n(t,\xi)\partial W_n(t), \\ & \qquad\qquad\qquad\qquad\qquad t \in \mathbb{R}, \xi \in \mathcal{O}, \\ u(t,\xi) &= 0, \quad t \in \mathbb{R}, \xi \in \partial\mathcal{O}, \end{cases} \tag{5.65}$$

where (W_n) is a sequence of independent standard real Brownian motions, $f : \mathbb{R}\times\mathcal{O} \to L^p(\mathcal{O})$ is almost periodic and $g_n : \mathbb{R}\times\mathcal{O} \to L^p(\mathcal{O})$ are all 2π-periodic and $\frac{1}{2}$-Hölder continuous with Hölder constants H_n, such that

$$\int_{\mathcal{O}} \Big\| \Big(\sum_{n=1}^{\infty} g_n(\cdot,\xi)^2 \Big)^{\frac{1}{2}} \Big\|^p_{L^2(0,2\pi)} \mathrm{d}\xi < \infty \qquad \text{and} \qquad \sum_{n=1}^{\infty} H_n^2 < \infty. \tag{5.66}$$

The operator $A(t,\xi)$ is given by

$$A(t,\xi) = \sum_{i,j=1}^{d} a_{ij}(\xi)\partial_i\partial_j + \sum_{k=1}^{d} a_k(t,\xi)\partial_k + a_0(t,\xi). \tag{5.67}$$

We assume that the functions a_{ij}, a_k and a_0 are in $C_b^{\mu}(\mathbb{R}, C(\overline{\mathcal{O}}))$, the space of bounded μ-Hölder continuous functions from $\mathbb{R}$ to $C(\overline{\mathcal{O}})$ for some $\mu \in (\frac{1}{2},1)$, that $a_{ij} = a_{ji}$, i.e. that the matrix $\big(a_{ij}\big)_{i,j=1}^{d}$ is symmetric and that

$$\sum_{i,j=1}^{d} a_{ij}(\xi)\xi'_i\xi'_j \geq \beta|\xi'|^2 \tag{5.68}$$

for some constant $\beta > 0$ and all $\xi \in \overline{\mathcal{O}}$, $\xi' \in \mathbb{R}^d$. Set $E := L^p(\mathcal{O})$ and define the family $A(t)$ of operators on E by

$$\big(A(t)\phi\big)(\xi) := A(t,\xi)\phi(\xi), \quad \xi \in \mathcal{O},\ t \in \mathbb{R},$$

with domain

$$D(A(t)) := \big\{\phi \in W^{2,p}(\mathcal{O}) : \phi(\xi) = 0 \text{ for } \xi \in \partial\mathcal{O}\big\} =: E_1.$$

The family $A(t)$ then satisfies the conditions (AT), see [1, Section 6], [10, Example 5.6], [80, Example 2.9] or [4], and so there exists a evolution family $P(t,s)$ on E which solves the problem

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}u(t) &= A(t)u(t) \qquad t \geq s, \\ u(s) &= x_s \in E, \end{cases}$$

with solution $u(t) = P(t,s)x_s$. It is also necessary to assume that $P(t,s)$ is uniformly stable. Unfortunately a simple characterisation of stability of evolution families in terms of the generators is not available. However, note that if $a_0(t,\xi) < -\lambda$ for sufficiently large $\lambda > 0$ then the analytic semigroup $S_{0,\lambda}(t)$ generated by the second order term minus λI will be uniformly stable. Now $P(t,s)$ is a lower order perturbation of the evolution family $P_{0,\lambda}(t,s) := S_{0,\lambda}(t-s)$ and by [39, Theorem VI.9.19], $P(t,s)$ is exponentially bounded, and hence for sufficiently large λ, exponentially stable.

From (5.67) and (5.68), the graph norm of $A(t)$ is uniformly equivalent to the norm on $W^{2,p}(\mathcal{O})$ for all $t \in \mathbb{R}$.

By [62, Example 1.10] or [83, Section 2.3, 2.4], if $k \in \mathbb{N}$ and $\theta \in (0,1)$ is such that $k\theta \notin \mathbb{N}$, then the following interpolation space and fractional Sobolev space coincide

$$\big(L^p(\mathcal{O}), W^{k,p}(\mathcal{O})\big)_{\theta,p} = W^{k\theta,p}(\mathcal{O}).$$

We also use the following embedding.

Proposition 5.32 (Proposition 1.4 of [62]). *If $0 < \beta < \alpha < 1$ then*

$$(E, E_1)_{\alpha,q_1} \hookrightarrow (E, E_1)_{\beta,q_2}$$

for any $q_1, q_2 \in [1, \infty]$.

Hence for $\frac{1}{2} < \alpha - \varepsilon < \alpha < 1$ and $1 \le q \le p$ we have

$$\begin{aligned} W^{2,p}(\mathcal{O}) &\hookrightarrow \big(L^p(\mathcal{O}), W^{2,p}(\mathcal{O})\big)_{\alpha,\infty} \\ &\hookrightarrow \big(L^p(\mathcal{O}), W^{2,p}(\mathcal{O})\big)_{\alpha-\varepsilon,p} = W^{2(\alpha-\varepsilon),p}(\mathcal{O}) \\ &\hookrightarrow W^{1,p}(\mathcal{O}). \end{aligned} \tag{5.69}$$

Now the difference $A(t,\xi) - A(s,\xi)$ is just the first order differential operator

$$A(t,\xi) - A(s,\xi) = \sum_{k=1}^{d} \big[a_k(t,\xi) - a_k(s,\xi)\big]\partial_k + \big[a_0(t,\xi) - a_0(s,\xi)\big],$$

and so $A(t) - A(s)$ is bounded from $W^{1,p}(\mathcal{O})$ to $L^p(\mathcal{O})$.

If F_1 and F_2 are Banach spaces with $F_2 \hookrightarrow F_1$, then there exists a constant $C \ge 0$ such that for $x \in F_2$

$$\|x\|_{F_1} \le C\|x\|_{F_2}.$$

So if $B \in \mathcal{L}(F_1, E)$, $x \in F_2$ we have

$$\|Bx\| \le \|B\|_{\mathcal{L}(F_1,E)}\|x\|_{F_1} \le C\|B\|_{\mathcal{L}(F_1,E)}\|x\|_{F_2},$$

then $B \in \mathcal{L}(F_2, E)$ and $\|B\|_{\mathcal{L}(F_2,E)} \leq C\|B\|_{\mathcal{L}(F_1,E)}$. Thus by (5.69), $A(t) - A(s)$ is bounded from E_α to $L^p(\mathcal{O})$ for every $\alpha \in (\frac{1}{2}, 1)$ with

$$\big\|A(t) - A(s)\big\|_{\mathcal{L}(E_\alpha,E)} \leq C_\alpha \big\|A(t) - A(s)\big\|_{\mathcal{L}(W^{1,p}(\mathcal{O}),E)} \tag{5.70}$$

for all $s, t \in \mathbb{R}$.

Now, if $a_k(t,\xi)$ is in $AP(\mathbb{R}; C(\overline{\mathcal{O}}))$ for all $k = 0, 1, \dots, d$ then for any $\tau_n \uparrow \infty$ we can pass to a subsequence such that $a_k(t + \tau_n, \cdot)$ converges in $C(\overline{\mathcal{O}})$ uniformly for $t \in \mathbb{R}$ for $k = 0, 1, \dots, d$. Then for $\phi \in W^{1,p}(\mathcal{O})$

$$\begin{aligned}
&\|A(t + \tau_n)\phi - A(t+\tau_m)\phi\|_{L^p(\mathcal{O})} \\
&\qquad \leq \sum_{k=1}^{d} \sup_{\xi\in\mathcal{O}} |a_k(t + \tau_n, \xi) - a_k(t + \tau_m, \xi)| \|\partial_k\phi\|_{L^p(\mathcal{O})} \\
&\qquad\qquad + \sup_{\xi\in\mathcal{O}} |a_0(t + \tau_n, \xi) - a_0(t + \tau_m, \xi)| \|\phi\|_{L^p(\mathcal{O})} \\
&\qquad \leq \max_{k=0}^{d} \sup_{\xi\in\mathcal{O}} |a_k(t + \tau_n, \xi) - a_k(t + \tau_m, \xi)| \|\phi\|_{W^{1,p}(\mathcal{O})} \\
&\qquad \to 0 \text{ as } m, n \to \infty
\end{aligned}$$

uniformly for $t \in \mathbb{R}$. Hence by (5.70), for any sequence $\tau_n \uparrow \infty$ we can pass to a subsequence such that $\|A(t + \tau_n) - A(t + \tau_m)\|_{\mathcal{L}(E_\alpha,E)} \to 0$ as $m, n \to \infty$ uniformly for $t \in \mathbb{R}$.

Hence we have that $A(t)$ satisfies all three condition of Theorem 5.17, and therefore the evolution family $P(t, s)$ generated by $A(t)$ will be jointly almost periodic.

Now define the ℓ^2-cylindrical Brownian motion $W(t)$ as in Theorem 4.19 by

$$W(t)e_n = W_n(t), \quad n \in \mathbb{N}$$

and $G : \mathbb{R} \to \mathcal{L}(\ell^2, L^p(\mathcal{O}))$ by

$$\big(G(t)e_n\big)(\xi) = g_n(t, \xi) \quad n \in \mathbb{N},$$

where (e_n) is the usual basis of ℓ^2. We check conditions (i) to (iii) of Theorem 5.14.

(i) For $t, s \in \mathbb{R}$

$$\begin{aligned}
\|G(t) - G(s)\|_{\gamma(\ell^2,L^p(\mathcal{O}))} &\simeq \|G(t) - G(s)\|_{L^p(\mathcal{O},\ell^2)} \\
&\leq \left(\int_{\mathcal{O}} \left(\sum_{n=1}^{\infty} |g_n(t, \xi) - g_n(s, \xi)|^2\right)^{\frac{p}{2}} \mathrm{d}\xi\right)^{\frac{1}{p}} \\
&\leq |t - s|^{\frac{1}{2}} \left(\sum_{n=1}^{\infty} H_n^2\right)^{\frac{1}{2}} m(\mathcal{O})^{\frac{1}{p}}.
\end{aligned}$$

(ii) By the γ-Fubini isomorphism (3.8) and (5.66),

$$
\begin{aligned}
\|G\|_{\gamma(L^2(0,2\pi;\ell^2),L^p(\mathcal{O}))} &\lesssim \|G\|_{L^p(\mathcal{O};L^2(0,2\pi;\ell^2))} \\
&\leq \left(\int_{\mathcal{O}} \left(\int_0^{2\pi} \sum_{n=1}^{\infty} g_n(t,\xi)^2 \, \mathrm{d}t \right)^{\frac{p}{2}} \mathrm{d}\xi \right)^{\frac{1}{p}} \\
&< \infty.
\end{aligned}
$$

(iii) For $\tau \in \mathbb{R}$

$$
\begin{aligned}
&\|G(\cdot) - G(\cdot + \tau)\|_{\gamma(L^2(0,2\pi;\ell^2),L^p(\mathcal{O}))} \\
&\qquad \lesssim \|G(\cdot) - G(\cdot + \tau)\|_{L^p(\mathcal{O};L^2(0,2\pi;\ell^2))} \\
&\qquad \leq \left(\int_{\mathcal{O}} \left(\int_0^{2\pi} \sum_{n=1}^{\infty} |g_n(t,\xi) - g_n(t+\tau,\xi)|^2 \, \mathrm{d}t \right)^{\frac{p}{2}} \mathrm{d}\xi \right)^{\frac{1}{p}} \\
&\qquad \leq \tau^{\frac{1}{2}} \sqrt{2\pi} \left(\sum_{n=1}^{\infty} H_n^2 \right)^{\frac{1}{2}} m(\mathcal{O})^{\frac{1}{p}}.
\end{aligned}
$$

Hence all the conditions of Theorem 5.14 are satisfied, and so there exists a solution $X : \mathbb{R} \times \Omega \to L^p(\mathcal{O})$ of (5.65) which is L^q-bounded (i.e. $\sup_{t\in\mathbb{R}} \mathbb{E}\|X(t)\|^q < \infty$) for all $q \in [1, \infty)$ and $X(t)$ is almost periodic in distribution. $\square$

Chapter 6

Invariant measures and delay problems

6.1 Introduction

In this Chapter we investigate the existence of invariant measures for stochastic differential equations and stochastic delay equations on Banach spaces. We consider problems of the form

$$\mathrm{d}X(t) = \big[AX(t) + F\big(X(t), X_t\big)\big]\,\mathrm{d}t + G\big(X(t), X_t\big)\,\mathrm{d}W(t), \quad t \geq 0 \tag{6.1}$$

and ask whether there exists a measure μ on the space $\mathcal{E}_p := E \times L^p(-1, 0; E)$ with the property that if $(X(0), X_0)$ has distribution μ then $(X(t), X_t)$ has distribution μ for all $t \geq 0$.

The fundamental tool for showing existence of invariant measures is the Krylov-Bogoliubov Theorem (see Theorem 6.3 below), which relates tightness of solutions to invariance of measures. The basic pattern of many results in this area is to use compactness of the semigroup $S(t)$ and an assumption on the existence of a bounded solution (in some sense) to show tightness of the family of distributions induced by $X(t)$ and then to use Krylov-Bogoliubov. Existence of invariant measures for the non-delayed problem

$$\mathrm{d}X(t) = \big[AX(t) + F(X(t))\big]\,\mathrm{d}t + G(X(t))\,\mathrm{d}W(t), \quad t \geq 0 \tag{6.2}$$

was studied initially by Da Prato *et al.* [28] on Hilbert spaces, using both a compactness argument as outlined above and also an alternative approach based on a dissipativity condition on the Yosida approximations of A. Both approaches have been well studied in recent years, for example Brzeźniak *et al.* consider dissipativity assumptions on Banach spaces of martingale type 2 [22], whereas van Gaans *et al.* consider compactness in Hilbert spaces. Alternative methods also exist, for instance van Neerven and Weis [71] study the case of constant additive noise $G \in \mathcal{L}(E)$ and obtain invariant measures by way of spectral theory and γ-boundedness of resolvents $R(\lambda, A)$ on the half-plane $\{\mathrm{Re}\lambda > 0\}$.

Invariant measures for delay problems are also studied, van Neerven and Riedle [66] show existence of such measures for certain delay problems with additive noise on spaces of continuous functions. Our principal reference for the results that follow regarding delay problems though will be the paper of van Gaans *et al.* [14] who find invariant measures for problems driven by *eventually* compact semigroups on Hilbert spaces, which therefore implies the same for delay problems with immediately compact linear part. For well posedness results and equivalence of solutions of (6.1) with the delay Cauchy problem on the delay space $\mathcal{E}$ of the form on (2.17), we refer to Cox and Górajski [25].

6.2 Preliminaries

We consider first the undelayed problem (6.2) and make the following, by now standard, assumptions

(M1) A generates a C_0-semigroup $S(t)$ on a separable UMD Banach space E of type 2, $W(t)$ is an H-cylindrical Brownian motion and there exists $M \geq 0$ and $\lambda > 0$ such that $\|S(t)\| \leq Me^{-\lambda t}$ for all $t \geq 0$.

(M2) $F : E \to E$ is Lipschitz, that is, there exists $L_F \geq 0$ such that for all $x, y \in E$

$$\|F(x) - F(y)\| \quad \leq \quad L_F\|x - y\|.$$

(M3) $G : E \to \gamma(H, E)$ is Lipschitz, that is, there exists $L_G \geq 0$ such that for all $x, y \in E$

$$\|G(x) - G(y)\|_{\gamma(H,E)} \quad \leq \quad L_G\|x - y\|.$$

By the paper of Cox and Górajski [25], (M1) - (M3) are enough to guarantee a unique mild solution $X(t, x)$ to (6.2) with continuous paths for any initial condition $x \in E$. To show invariant measures we also require that solutions are bounded, uniformly in probability:

(M4) For any $x \in E$ and $\varepsilon > 0$, there exists $N \geq 0$ such that

$$\mathbb{P}\big(\|X(t,x)\| \geq N\big) \leq \varepsilon$$

for any $t \geq 0$.

Note that by Markov's inequality, (M4) holds whenever there exists L^p-bounded solutions (see Definition 5.1). It is typically not possible to establish (M4) at this level of generality, i.e. to give an easily checkable criterion for A, F, G and E such that (M4) will hold. Assumptions of this type are very often required in the literature and rely on being checked on a case by case basis where applied.

6.2.1 The Krylov-Bogoliubov Theorem

The Krylov-Bogoliubov Theorem essentially says that the existence of invariant measures for certain stochastic problems is equivalent to the tightness of families of measures associated to bounded solutions of the problem. The main idea underlying this chapter is to find conditions on equations (6.1) and (6.2) such that solution families are *tight* in some sense and then apply the Krylov-Bogoliubov Theorem to get existence of invariant measures.

In order to state the theorem we first recall the notion of the *Markov semigroup* associated with a process $X(t)$.

Definition 6.1 (Definition 5.1 of [27]). A *Markov semigroup* (P_t) on the space $\mathbb{B}_b(E)$ of bounded Borel functions on E is a map $\mathbb{R}_+ \to \mathcal{L}(\mathbb{B}_b(E))$ such that

(i) $P_0 = I$ and $P_{t+s} = P_t P_s$ for all $t, s \geq 0$.

(ii) For any $t \geq 0$ and $x \in E$ there exists a probability measure $\pi_{t,x}$ on E such that

$$P_t\varphi(x) = \int_E \varphi(y)\pi_{t,x}(\,\mathrm{d}y) \quad \text{for all } \varphi \in \mathbb{B}_b(E). \tag{6.3}$$

(iii) For any $x \in E$, the map $t \mapsto P_t\varphi(x)$ is continuous for all $\varphi \in C_b(E)$ (the continuous bounded functions on E) and Borel measurable for any $\mathbb{B}_b(E)$ (the bounded Borel measurable functions on E).

We say P_t is *Feller* (or has the *Feller property*) if $C_b(E)$ is an invariant subspace for all $t \geq 0$, i.e. if $P_t\varphi \in C_b(E)$ for any $\varphi \in C_b(E), t \geq 0$.

Example 6.2 (See [27], Section 5.1)**.** Let $X(t,x)$ be the solution of (6.2) with initial condition $x \in E$, then $X(t,x)$ is a continuous Markov process on E and the family $P_t \in \mathcal{L}(\mathbb{B}_b(E))$ defined by

$$P_t\varphi(x) := \mathbb{E}\big(\varphi(X(t,x)\big), \quad t \geq 0, x \in E. \tag{6.4}$$

is a Markov semigroup, with associated measure $\pi_{t,x}$ on E given by $\pi_{t,x}(B) := \mathbb{P}(X(t,x) \in B)$ for $B \in \mathbb{B}(E)$. Moreover, P_t is Feller continuous.

Proof. That $X(t,x)$ is a Markov process is shown in [72, Theorem 7.1.2] (the proof works exactly the same for Banach spaces). Define P_t as in (6.4), then

(i) For $\varphi \in \mathbb{B}_b(E)$

$$P_0\varphi(x) = \mathbb{E}\big(\varphi(X(0,x)\big) = \mathbb{E}\varphi(x) = \varphi(x) \text{ for all } x \in E$$

so $P_0 = I$ and for $t, s \geq 0$

$$\begin{aligned} P_tP_s\varphi(x) &= \mathbb{E}\big(\big(P_s\varphi\big)(X(t,x)\big) = \mathbb{E}\big(\varphi(X(s,X(t,x))\big) \\ &= \mathbb{E}\big(\varphi(X(s+t,x)\big) = P_{t+s}\varphi(x). \end{aligned}$$

(ii) $\pi_{t,x}$ as above clearly satisfies

$$\int_E \varphi(y)\pi_{t,x}(\,\mathrm{d}y) = \mathbb{E}\big(\varphi(X(t,x)\big). \tag{6.5}$$

(iii) Since $t \mapsto X(t,x)$ is continuous, $t \mapsto P_t\varphi(x)$ is continuous (resp. Borel) for every $\varphi \in C_b(E)$ (resp. $\mathbb{B}_b(E)$) and $x \in E$.

To show the Feller property we proceed as in [72, Lemma 8.1.4]. Fix $0 \leq t \leq T$ for some $T \geq 0$ and let $x, y \in E$, then by Lipschitz continuity of F and G and Corollary 3.34

$$\begin{aligned} \mathbb{E}\|X(t,x) - X(t,y)\|^2 \leq M\|x-y\|^2 &+ \mathbb{E}\Big\|\int_0^t S(t-s)\big[F(X(s,x)) - F(X(s,y))\big]\,\mathrm{d}s\Big\|^2 \\ &+ \mathbb{E}\Big\|\int_0^t S(t-s)\big[G(X(s,x)) - G(X(s,y))\big]\,\mathrm{d}W(s)\Big\|^2 \\ \leq M\|x-y\|^2 &+ TC_1\int_0^t \mathbb{E}\big\|F(X(s,x)) - F(X(s,y))\big\|^2\,\mathrm{d}s \\ &+ TC_2\int_0^t \mathbb{E}\big\|G(X(s,x)) - G(X(s,y))\big\|^2_{\gamma(H,E)}\,\mathrm{d}s \\ \leq M\|x-y\|^2 &+ C_T\int_0^t \mathbb{E}\big\|X(s,x) - X(s,y)\big\|^2\,\mathrm{d}s \end{aligned}$$

and hence by Gronwall's Lemma, [13, Lemma 1],

$$\mathbb{E}\|X(t,x) - X(t,y)\|^2 \leq M\|x-y\|^2 e^{C_T t}.$$

We have $x \mapsto X(t,x)$ is continuous from E to $L^2(\Omega; E)$ for each $t \geq 0$, so if $(x_n) \subset E$ is a sequence converging to $x \in E$, we can pass to a subsequence (also (x_n)) such that $X(t,x_n) \to X(t,x)$ almost surely. By the Fatou Lemma and the continuity of φ

$$\begin{aligned} P_t\varphi(x) = \mathbb{E}\varphi(X(t,x)) &= \mathbb{E}\big[\liminf_{n\to\infty} \varphi(X(t,x_n))\big] \\ &\leq \liminf_{n\to\infty} \mathbb{E}\big[\varphi(X(t,x_n))\big] = \liminf_{n\to\infty} P_t\varphi(x_n). \end{aligned} \tag{6.6}$$

Now (see [74, Proposition 1.5.11]), a function $f : E \to \mathbb{R}$ is *lower semicontinuous* if and only if for every convergent sequence $(y_n) \subset E$ we have

$$f\big(\lim_{n\to\infty} y_n\big) \leq \liminf_{n\to\infty} f(y_n).$$

Hence by (6.6), $P_t\varphi$ is lower semicontinuous and applying the same to $-\varphi$, $P_t\varphi$ is also upper semicontinuous. Hence we conclude the Feller property for P_t. □

A measure μ on E is an invariant measure for $X(t)$ if and only if

$$\int_E P_t\varphi\mu(\,\mathrm{d}x) = \int_E \varphi\mu(\,\mathrm{d}x) \quad \text{for all } \varphi \in \mathbb{B}_b(E) \text{ and } t \geq 0.$$

Theorem 6.3 (Krylov-Bogoliubov, Theorem 7.1 of [27])**.** *Suppose P_t is a Markov semigroup on $\mathbb{B}_b(E)$, that P_t is Feller continuous and that the associated family of measures $\{\pi_{t,x_0}\}_{t\geq 0}$ is uniformly tight for some $x_0 \in E$. Then there exists an invariant measure for P_t*

In order to give the proof, we first recall the following theorem of Prokhorov.

Theorem 6.4 (Prokhorov, see [55] Theorem 8.9)**.** *Let E be a separable metric space, then a family of probability measures $\{\mu_i : i \in I\}$ is tight if and only if it is weakly compact (in the sense of Definition 5.2).*

Proof of Theorem 6.3. The family $\{\pi_{t,x_0}\}_{t\geq 0}$ is uniformly tight for some $x_0 \in E$. Define a new family $\{\mu_T\}_{T\geq 0}$ by

$$\mu_T(B) = \frac{1}{T}\int_0^T \pi_{t,x_0}(B)\,\mathrm{d}t \quad B \in \mathbb{B}(E), \tag{6.7}$$

and note that if $\pi_{t,x_0}(K) \geq 1-\varepsilon$ for all $t \geq 0$ then $\mu_T(K) \geq 1-\varepsilon$, and so $\{\mu_T\}_{T\geq 0}$ is tight. By Prokhorov, $\{\mu_T\}_{T\geq 0}$ is weakly compact so there exists a sequence $T_n \uparrow \infty$ and a probability measure μ on E such that

$$\lim_{n\to\infty} \int_E \varphi(x)\mu_{T_n}(\,\mathrm{d}x) = \int_E \varphi(x)\mu(\,\mathrm{d}x) \quad \text{for all } \varphi \in C_b(E).$$

By Fubini and (6.3), this is equivalent to

$$\lim_{n\to\infty} \frac{1}{T_n}\int_0^{T_n} P_t\varphi(x_0)\,\mathrm{d}t = \int_E \varphi(x)\mu(\,\mathrm{d}x) \quad \text{for all } \varphi \in C_b(E). \tag{6.8}$$

Now for $\psi \in C_b(E)$, let $\varphi = P_s\psi$, then by the Feller property, $\varphi \in C_b(E)$ and

$$\lim_{n\to\infty} \frac{1}{T_n}\int_0^{T_n} P_{t+s}\psi(x_0)\,\mathrm{d}t = \int_E P_s\psi(x)\mu(\,\mathrm{d}x) = \int_E \varphi(x)\mu(\,\mathrm{d}x) \quad \text{for all } \psi \in C_b(E). \tag{6.9}$$

It remains to show that the left-hand side of (6.9) is equal to

$$\int_E \psi\mu(\,\mathrm{d}x),$$

and this will prove that μ is invariant. By (6.8) we have

$$\begin{aligned}
&\frac{1}{T_n}\int_0^{T_n} P_{t+s}\psi(x_0)\,\mathrm{d}t \\
&\quad= \frac{1}{T_n}\int_s^{T_n+s} P_t\psi(x_0)\,\mathrm{d}t \\
&\quad= \frac{1}{T_n}\int_0^{T_n} P_t\psi(x_0)\,\mathrm{d}t + \frac{1}{T_n}\int_{T_n}^{T_n+s} P_{t+s}\psi(x_0)\,\mathrm{d}t - \frac{1}{T_n}\int_0^{s} P_t\psi(x_0)\,\mathrm{d}t \\
&\quad\to \int_E \psi(x)\mu(\,\mathrm{d}x) \text{ as } n\to\infty.
\end{aligned}$$

□

Note that by (6.7), we can replace (M4) by the following slightly weaker assumption with only very minor adjustments.

(M5) For all $x \in E$ and $\varepsilon > 0$, there exists $M \geq 0$ such that for all $T > 1$,

$$\frac{1}{T}\int_0^T \mathbb{P}(\|X(t,x)\| \geq M)\,\mathrm{d}t < \varepsilon.$$

6.3 Invariant measures for SDEs in Banach space

To establish the base case, the existence of invariant measures for non-delayed stochastic differential equations of the form of (6.2), we follow the method of Da Prato [28], making the adjustments necessitated by the Banach space setting.

We start by recalling a lemma of Da Prato [29, Proposition 1] known as the *factorisation method* for stochastic convolutions. The factorisation method was heavily used in the papers [69] and [86] on which Chapter 4 was based (see for example Lemma 4.10), however the details were largely suppressed in that section as the results in question from those papers were referenced without change. Here we give more details. The two results that follow were originally presented for Hilbert spaces, but the proofs work more generally on Banach spaces without change.

For $p > 1$, $g \in L^p(0,1;E)$ and $\alpha \in \left(\frac{1}{p}, 1\right]$, define the operator $R_\alpha : L^p(0,1;E) \to L^p(0,1;E)$ by

$$R_\alpha g(t) := \int_0^t (t-s)^{\alpha-1} S(t-s) g(s) \, \mathrm{d}s.$$

Proposition 6.5 (Lemma 1 in [29] and Proposition 1 in [46])**.** *If $S(t)$ is a C_0-semigroup then for any $p > 1$ and $\alpha \in (\frac{1}{p}, 1]$, R_α is a continuous operator from $L^p(0,1;E)$ to $C([0,1];E)$. Moreover, if $S(t)$ is a compact semigroup then R_α is compact.*

Proposition 6.6 (Proposition 1 of [29])**.** *Let S be a C_0-semigroup, $p > 2$ and $\alpha \in \left(\frac{1}{p}, \frac{1}{2}\right)$. Then for any $t \in [0,1]$ and any adapted $\phi \in L^p(\Omega \times [0,1]; \gamma(H,E))$*

$$\int_0^t S(t-s)\phi(s) \, \mathrm{d}W(s) = \frac{\sin(\pi\alpha)}{\pi} R_\alpha Y(t) \ \text{a.s.}$$

where

$$Y(t) = \int_0^t (t-s)^{-\alpha} S(t-s) \phi(s) \, \mathrm{d}W(s).$$

By Example 6.2, $X(t,x)$, the solution of (6.2) with initial condition $x \in E$, is a continuous Markov process on E. Denote by $P(t,x,\Gamma)$ the transition probabilities for $X(t,x)$,

$$P(t,x;\Gamma) := \mathbb{P}(X(t,x) \in \Gamma) \tag{6.10}$$

for $\Gamma \subseteq E$ and $t \in [0,1]$. Now for $M \geq 0$ we write

$$K(M,\alpha) := \left\{ y \in E : \begin{array}{c} y = S(1)x + (R_1 f)(1) + (R_\alpha g)(1) \\ \text{for all } \|x\|^p, \|f\|_p^p, \|g\|_p^p \leq M \end{array} \right\} \tag{6.11}$$

Lemma 6.7. *Assume (M1) - (M3) and chose* $\alpha \in (\frac{1}{p}, \frac{1}{2})$. *Then there exists a constant* $C \geq 0$ *such that*

$$P\big(1, x; K(M,\alpha)\big) \geq 1 - \frac{C}{M}\big(1 + \|x\|^p\big) \tag{6.12}$$

for every $\|x\|^p \leq M$.

Proof. Fix $x \in E$ and write $Z(t,x) = \mathbb{E}\|X(t,x)\|^p$ for $0 \leq t \leq 1$. Then by Jensen's inequality and Corollary 3.34

$$\begin{aligned} Z(t,x) &\leq C\|x\|^p + \mathbb{E}\Big\| \int_0^t S(t-s)F(X(s))\,\mathrm{d}s\Big\|^p + \mathbb{E}\Big\|\int_0^t S(t-s)G(X(s))\,\mathrm{d}W(s)\Big\|^p \\ &\leq C\|x\|^p + C_1\Big(1 + \int_0^t Z(s,x)\,\mathrm{d}s\Big) + C_2\mathbb{E}\Big(\int_0^t \|B(X(s,x))\|^2_{\gamma(H,E)}\,\mathrm{d}s\Big)^{\frac{p}{2}} \\ &\leq C\|x\|^p + C_3\Big(1 + \int_0^t Z(s,x)\,\mathrm{d}s\Big). \end{aligned} \tag{6.13}$$

Hence by Gronwall's inequality, [13, Lemma 1], we have

$$Z(t,x) \leq C_4\big(1 + \|x\|^p\big). \tag{6.14}$$

Write

$$Y^x(t) := \int_0^t (t-s)^{-\alpha} S(t-s)G(X(s,x))\,\mathrm{d}W(s) \text{ and } F^x(t) = F(X(t,x)),$$

then by corollary 3.34

$$\begin{aligned} \mathbb{E}\int_0^1 \|Y^x(s)\|^p\,\mathrm{d}s &= \mathbb{E}\int_0^1 \Big\|\int_0^s (s-r)^{-\alpha}S(s-r)G(X(r,x))\,\mathrm{d}W(r)\Big\|^p\,\mathrm{d}s \\ &\leq C\mathbb{E}\int_0^1 \Big(\int_0^s (s-r)^{-2\alpha}\|G(X(r,x))\|^2_{\gamma(H,E)}\,\mathrm{d}r\Big)^{\frac{p}{2}}\,\mathrm{d}s. \end{aligned} \tag{6.15}$$

Now we apply Young's inequality, [45, Theorem 9.4.1], which says that if $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$ and $\phi \in L^p(0,1)$, $\psi \in L^q(0,1)$ then $\phi * \psi \in L^r(0,1)$ and

$$\|\phi * \psi\|_{L^r(0,1)} \leq \|\phi\|_{L^p(0,1)}\|\psi\|_{L^q(0,1)}.$$

We get

$$\begin{aligned} \mathbb{E}\int_0^1 \|Y^x(s)\|^p\,\mathrm{d}s &\leq C\Big(1 + \int_0^1 \mathbb{E}\|X(s,x)\|\,\mathrm{d}s\Big) \\ &\leq C\big(1 + \|x\|^p\big). \end{aligned} \tag{6.16}$$

By (6.14) and (6.16)

$$\mathbb{E}\big(\|Y^x\|^p_{L^p(0,1;E)} + \|F^x\|^p_{L^p(0,1;E)}\big) \le C\big(1 + \|x\|^p\big). \tag{6.17}$$

By Proposition 6.6 we can write

$$X(1,x) = S(1)x + \big(R_1 F^x\big)(1) + \frac{\sin(\pi\alpha)}{\pi}\big(R_\alpha Y^x\big)(1),$$

hence if $\|x\|^p, \|F^x\|^p_{L^p(0,1;E)} \le M$ and $\|Y^x\|^p_{L^p(0,1;E)} \le \frac{\pi^p M}{\sin(\pi\alpha)^p}$ then $X(1,x) \in K(M,\alpha)$. Then by Markov's inequality, (6.11) and (6.17) we have

$$\begin{aligned}
\mathbb{P}\big(X(1,x) \notin K(M,\alpha)\big) &\le \mathbb{P}\big(\|F^x\|^p_{L^p(0,1;E)} > M\big) + \mathbb{P}\Big(\|Y^x\|^p_{L^p(0,1;E)} > \frac{\pi^p M}{\sin(\pi\alpha)^p}\Big) \\
&\le \frac{1}{M}\Big(\mathbb{E}\|F^x\|^p_{L^p(0,1;E)} + \frac{\sin(\pi\alpha)^p}{\pi^p}\mathbb{E}\|Y^x\|^p_{L^p(0,1;E)}\Big) \\
&\le \frac{C}{M}\big(1 + \|x\|^p\big).
\end{aligned} \tag{6.18}$$

Thus (6.12) holds. □

Proposition 6.8. *Assume (M1) - (M3) and (M4). If the semigroup $S(t)$ is compact then the family $\{X(t,x)\}_{t\ge 1}$ is tight for any $x \in E$.*

Proof. Fix $p > 2$ and $\alpha \in (\frac{1}{p}, \frac{1}{2})$. Since the operators $S(1), R_1$ and R_α are compact and the map $C([0,1];E) \to E$ which evaluates functions at 1 is continuous, we see that the set $K(M,\alpha)$ is relatively compact. By (M4), (6.12) and the Markov property, we have for any $t \ge 1$

$$\begin{aligned}
\mathbb{P}\big(X(t,x) \in K(M,\alpha)\big) &= \mathbb{E}P\big(1, X(t-1,x); K(M,\alpha)\big) \\
&\ge \mathbb{E}\Big[P\big(1, X(t-1,x); K(M,\alpha)\big)\Big|\|X(t-1,x)\|^p \le N\Big] \\
&\ge \Big(1 - \frac{C}{M}(1+N)\Big)\mathbb{P}\big(\|X(t-1,x)\| \le N\big) \\
&\ge 1 - \varepsilon
\end{aligned}$$

for sufficiently large M for any $\varepsilon > 0$, where $P(t,x,\Gamma)$ is defined as in (6.10). Thus $\{X(t,x)\}_{t\ge 1}$ is tight. □

We can now prove the existence of invariant measures for equation (6.2).

Theorem 6.9. *Assume (M1) - (M4) and that A generates a compact semigroup. Then there exists an invariant measure for*

$$\mathrm{d}X(t) = \big[AX(t) + F(X(t))\big]\,\mathrm{d}t + G(X(t))\,\mathrm{d}W(t), \quad t \ge 0. \tag{6.2}$$

Proof. By Proposition 6.8, the family $\{X(t,x)\}_{t\ge 1}$ is tight for any $x \in E$ and by Example 6.2, the associated Markov Semigroup is Feller continuous. Hence by the Krylov-Bogoliubov Theorem 6.3, there exists such an invariant measure. □

6.4 Invariant measures for eventually compact semigroups

Having established existence of invariant measures for our 'base', undelayed problem, we hope to establish the same for delayed stochastic differential equations. As an intermediate step we follow the paper of van Gaans [14] and show conditions under which invariant measures exist in problems driven by eventually compact semigroups. Then in section 6.5, using the fact (Theorem 2.19) that if $S(t)$ is compact for $t > 0$ then the delay semigroup $\mathcal{S}(t)$ defined in (2.18) is compact for $t > 1$, we reach our goal.

The following result is given for Hilbert spaces by van Gaans in [14, Lemma 2.3] and Bogachev [17, Example 3.8.13(ii)], but we give here a version for a Banach space with basis.

Proposition 6.10. *Let E be a Banach space with Schauder basis $(e_i)_{i\in\mathbb{N}}$ and let $K \subset E$, $K \notin \{\emptyset, \{0\}\}$ be a compact set. There exists an injective compact operator $T \in \mathcal{L}(E)$ such that*

$$K \subseteq T(B(0,1)).$$

Proof. As $(e_i)_{i\in\mathbb{N}}$ is a Schauder basis of E, for each $x \in E$ there exists a sequence $(\alpha_i^x)_{i\in\mathbb{N}}$ such that

$$x = \sum_{i=1}^{\infty} \alpha_i^x e_i.$$

Write S_n for the projection onto $\operatorname{span}\{e_i : i = 1, \ldots, n\}$

$$S_n x := \sum_{i=1}^{n} \alpha_i^x e_i,$$

then by [51, Theorem 4.13] S_n are bounded operators and $\sup_{n\in\mathbb{N}} \|S_n\| < \infty$. Consequently for any $n \in \mathbb{N}$,

$$\Big\| \sum_{i=n}^{\infty} \alpha_i^x e_i \Big\| = \|(I - S_{n-1})x\| \leq \Big(1 + \sup_{n\in\mathbb{N}} \|S_n\|\Big) \|x\| =: C\|x\|.$$

First we show that

$$\lim_{n\to\infty} \sup_{x\in K} \Big\| \sum_{i=n}^{\infty} \alpha_i^x e_i \Big\| = 0. \tag{6.19}$$

Suppose not, so there exists $\delta > 0$ such that there exists a sequence $(x_n) \subset K$ such that
$$\Big\| \sum_{i=n}^{\infty} \alpha_i^{x_n} e_i \Big\| \geq \delta$$
for all $n \in \mathbb{N}$.

Fix $n \in \mathbb{N}$. Since the sum representing x_n is convergent, there exists $M > n$ such that for all $m \geq M$
$$\Big\| \sum_{i=m}^{\infty} \alpha_i^{x_n} e_i \Big\| < \frac{\delta}{2}.$$
But then
$$\|x_n - x_m\| \geq \frac{1}{C} \big\| (I - S_m)(x_n - x_m) \big\| = \frac{1}{C} \Big\| \sum_{i=m}^{\infty} (\alpha_i^{x_n} - \alpha_i^{x_m}) e_i \Big\| > \frac{\delta}{2C}$$
and hence $(x_n)_{n\in\mathbb{N}}$ cannot contain a Cauchy subsequence, which contradicts the compactness of K.

By (6.19) then, we can choose an increasing sequence $(N_n)_{n\in\mathbb{N}}$ such that
$$\sup_{x\in K} \Big\| \sum_{i=N_n+1}^{\infty} \alpha_i^x e_i \Big\| \leq \frac{2^{-2n}}{C} \quad \text{for all } n \in \mathbb{N}. \tag{6.20}$$
We define the operator $T \in \mathcal{L}(E)$ as follows. For $i \in \mathbb{N}$, define $t_i > 0$ by
$$t_i = \begin{cases} 2C \sup_{x\in K} \|x\| & 1 \leq i \leq N_1 \\ 2^{-n+1} & N_n < i \leq N_{n+1} \end{cases}$$
then let
$$T\Big(\sum_{i=1}^{\infty} \alpha_i^x e_i \Big) := \sum_{i=1}^{\infty} t_i \alpha_i^x e_i.$$
Then T is clearly injective since $t_i > 0$ for all $i \in \mathbb{N}$ and
$$T = \big(2C \sup_{x\in K} \|x\|\big) S_{N_1} + \sum_{n=1}^{\infty} 2^{-n+1} \big(S_{N_{n+1}} - S_{N_n}\big),$$
where the series is absolutely convergent in norm. Hence T is the norm limit of finite rank operators, and therefore T is compact.

It remains to show that $K \subseteq T(B(0,1))$. Let $x \in K$ and define $y \in E$ by
$$y := \sum_{i=1}^{\infty} \frac{\alpha_i^x}{t_i} e_i.$$

Then $Ty = x$ and

$$\begin{aligned}
\|y\| &= \Big\| \sum_{i=1}^{\infty} \frac{\alpha_i^x}{t_i} e_i \Big\| \\
&\le \Big\| \sum_{i=1}^{N_1} \frac{\alpha_i^x}{t_i} e_i \Big\| + \sum_{n=1}^{\infty} \Big\| \sum_{i=N_n+1}^{N_{n+1}} \frac{\alpha_i^x}{t_i} e_i \Big\| \\
&\le \frac{1}{2C \sup_{x' \in K} \|x'\|} \|S_{N_1} x\| + \sum_{n=1}^{\infty} 2^{n-1} \Big\| \sum_{i=N_n+1}^{N_{n+1}} \alpha_i^x e_i \Big\|
\end{aligned}$$

Now by (6.20),

$$\Big\| \sum_{i=N_n+1}^{N_{n+1}} \alpha_i^x e_i \Big\| \le C \Big\| \sum_{i=N_n+1}^{\infty} \alpha_i^x e_i \Big\| \le 2^{-2n}$$

so the series for y converges absolutely in E and

$$\|y\| \;\le\; \frac{C\|x\|}{2C \sup_{x' \in K} \|x'\|} + \sum_{n=1}^{\infty} 2^{n-1} \cdot 2^{-2n} \;\le\; \frac{1}{2} + \sum_{n=1}^{\infty} 2^{-(n+1)} \;=\; 1. \tag{6.21}$$

Hence for all $x \in K$, $x = Ty$ for some $y \in B(0,1)$, i.e. $K \subseteq T(B(0,1))$. □

The method for eventually compact semigroups follows that in Section 6.3. We state an analogue of Lemma 6.5 which does not assume that the semigroup $S(t)$ is compact.

Lemma 6.11 (Lemma 2.2 of [14]). *Let $S(t)$ be a C_0-semigroup on E and suppose $C \in \mathcal{L}(E)$ is compact. Define the operator D by*

$$Df := \int_0^1 S(1-s) C f(s) \, \mathrm{d}s, \quad f \in L^p(0,1;E).$$

Then $D \in \mathcal{L}(L^p(0,1;E), E)$ is compact.

Lemma 6.11 suggests the following modification of the assumptions (M2) and (M3) of Section 6.2.

(M2′) $F : E \to E$ is Lipschitz and admits a factorisation $F = C_1 \circ F_1$ where $F_1 : E \to E_1$ for some Banach space E_1 and $C_1 \in \mathcal{L}(E_1, E)$ is compact.

(M3′) $G : E \to \gamma(H, E)$ is Lipschitz and admits a factorisation $G = C_2 \circ G_1$ where $G_1 : E \to \gamma(H, E_2)$ for some Banach space E_2 and $C_2 \in \mathcal{L}(E_2, E)$ is compact

We now give a series of results which appear in [14], culminating in an invariant measure for an eventually compact semigroup, but which we give here in type 2 Banach spaces with basis, rather than Hilbert spaces as originally presented. Very similar results are also contained in [34, Sections 5 and 6].

Assume (M1) and (M3′) and write

$$Y_x := \int_0^1 S(1-s)G(X(s,x))\,\mathrm{d}W(s),$$

where once again $X(t,x)$ represents the solution of (6.2) at time $t \geq 0$ and initial condition $X(0) = x \in E$.

Lemma 6.12 (Lemma 2.4 of [14]). *Suppose E is of type 2 and has a Schauder basis. For all $\varepsilon > 0$ and $M \geq 0$ there exists a compact set $K(M,\varepsilon)$ such that*

$$\mathbb{P}\big(Y_x \in K(M,\varepsilon)\big) > 1-\varepsilon$$

for all $\|x\| \leq M$.

Proof. Recall from (M3′) that $G = C_2 \circ G_1$ where $C_2 \in \mathcal{L}(E_2, E)$ is compact. Let

$$V := \big\{S(t)C_2 y : t \in [0,1], y \in E_2, \|y\|_{E_2} \leq 1\big\},$$

then we claim that V is relatively compact. Let $(v_n) \subset V$, then there exists $(t_n) \subset [0,1]$ and $(y_n) \in E_2$ with $\|y_n\|_{E_2} \leq 1$ such that

$$v_n = S(t_n)C_2 y_n.$$

Then by compactness of C_2, we can pass to a subsequence such that $C_2 y_n \to z \in E$, $\|z\| \leq \|C_2\|$. Since $[0,1]$ is compact, we can pass to a further subsequence such that (t_n) converges, and then by strong continuity of $S(t)$ we have $S(t_n)z \to x_0 \in S([0,1])z$. Then

$$\begin{aligned}\big\|S(t_n)C_2 y_n - x_0\big\| &\leq \big\|S(t_n)C_2 y_n - S(t_n)z\big\| + \big\|T(t_n)z - x_0\big\| \\ &\leq \big\|S(t_n)\big\|\big\|C_2 y_n - z\big\| + \big\|T(t_n)z - x_0\big\| \\ &\to 0 \text{ as } n \to \infty.\end{aligned}$$

So V is relatively compact, and hence if K is the closed convex hull of V, then K is compact [78, Theorem 3.25]. By Proposition 6.10 there exists an injective, compact operator $T \in \mathbb{E}$ such that $K \subseteq T(B(0,1))$. Since T is injective we define T^{-1} on $D(T^{-1}) = \text{range}(T)$, so $V \subseteq K \subseteq D(T^{-1})$. Let $K(\lambda) := \lambda T\big(\overline{B(0;1)}\big)$ for $\lambda > 0$, then

$$\begin{aligned} Y_x &= \int_0^1 S(1-s)C_2G_1(X(s,x))\,\mathrm{d}W(s) \\ &= \int_0^1 TT^{-1}S(1-s)C_2G_1(X(s,x))\,\mathrm{d}W(s) \\ &= T\int_0^1 T^{-1}S(1-s)C_2G_1(X(s,x))\,\mathrm{d}W(s), \end{aligned}$$

so

$$\begin{aligned} Y_x \in K(\lambda) \quad &\Longleftrightarrow \quad \int_0^1 T^{-1}S(1-s)C_2G_1(X(s,x))\,\mathrm{d}W(s) \in \lambda B(0,1) \\ &\Longleftrightarrow \quad \Big\| \int_0^1 T^{-1}S(1-s)C_2G_1(X(s,x))\,\mathrm{d}W(s)\Big\| \le \lambda. \end{aligned}$$

Now since $S(1-s)C_2\big[\,\overline{B(0,1)}\,\big] \subseteq K \subseteq T\big(\overline{B(0;1)}\big)$ we see that $T^{-1}S(1-s)C_2$ is an operator of norm less than 1, and so by Markov's inequality and the ideal property of $\gamma(H,E)$ ([65, Theorem 6.2])

$$\begin{aligned} \mathbb{P}\big(Y_x \notin K(\lambda)\big) &\le \mathbb{P}\Big(\int_0^1 T^{-1}S(1-s)C_2G_1(X(s,x))\,\mathrm{d}W(s) \notin \lambda B(0,1)\Big) \\ &\le \frac{1}{\lambda^2}\mathbb{E}\Big\|\int_0^1 T^{-1}S(1-s)C_2G_1(X(s,x))\,\mathrm{d}W(s)\Big\|^2 \\ &\simeq \frac{1}{\lambda^2}\mathbb{E}\int_0^1 \big\|T^{-1}S(1-s)C_2G_1(X(s,x))\big\|^2_{\gamma(H,E)}\,\mathrm{d}s \\ &\lesssim \frac{1}{\lambda^2}\mathbb{E}\int_0^1 \big\|G_1(X(s,x))\big\|^2_{\gamma(H,E)}\,\mathrm{d}s \\ &\le \frac{c}{\lambda^2}\big(1+\|x\|^2\big), \end{aligned}$$

where the last line follows as in (6.14). Finally choose λ large enough so that

$$\frac{c^2}{\lambda^2}(1+M^2) < \varepsilon$$

and the result follows. $\square$

Lemma 6.13 (Lemma 2.5 of [14]). *Suppose that E has a Schauder basis, that $S(t)$ is compact for $t \ge t_1 > 0$ and that (M1), (M2′), (M3′) and (M5) are satisfied. For any $\varepsilon > 0$ and $M \ge 0$ there exists a compact set $K(M,\varepsilon)$ such that*

$$\mathbb{P}\big(X(t_1,x) \in K(M,\varepsilon)\big) \ge 1-\varepsilon$$

for all $x \in E$ with $\|x\| \le M$.

Proof. We recall that

$$X(t_1,x) = S(t_1)x + \int_0^{t_1} S(t_1-s)F(X(s,x))\,\mathrm{d}s + \int_0^{t_1} S(t_1-s)G(X(s,x))\,\mathrm{d}W(s),$$

and treat each term separately. Since $S(t_1)$ is compact then there exists a compact set $K_1(M)$ such that $S(t_1)x \in K(M)$ for every $\|x\| \le M$.

Let $p \ge 1$, then it follows from (6.14) that

$$\mathbb{E}\int_0^{t_1} \|F_1(X(s,x))\|^p\,\mathrm{d}s \le C\big(1+\|x\|^p\big),$$

then for $\lambda > 0$

$$\begin{aligned}\mathbb{P}\big(\|F_1(X(\cdot,x))\|_{L^p(0,t_1;E_1)} > \lambda\big) &\le \frac{1}{\lambda^p}\mathbb{E}\|F_1(X(\cdot,x))\|_{L^p(0,t_1;E_1)} \le \frac{C}{\lambda^p}\big(1+\|x\|^p\big)\\ &\le \frac{C}{\lambda^p}\big(1+M^p\big).\end{aligned}$$

So for sufficiently large λ

$$\mathbb{P}\big(\|F_1(X(\cdot,x))\|_{L^p(0,t_1;E_1)} > \lambda\big) \le \frac{\varepsilon}{2}$$

and so by Lemma 6.11 there exists a compact set $K_2(M,\varepsilon) := K_2(\lambda)$ such that

$$\mathbb{P}\Big(\int_0^{t_1} S(t_1-s)F(X(s,x))\,\mathrm{d}s \in K_2(M,\varepsilon)\Big) > 1-\frac{\varepsilon}{2}.$$

By Lemma 6.12, there exists a compact set $K_3(M,\varepsilon)$ such that

$$\mathbb{P}\Big(\int_0^{t_1} S(t_1-s)G(X(s,x))\,\mathrm{d}W(s) \in K_3(M,\varepsilon)\Big) > 1-\frac{\varepsilon}{2}.$$

Hence we conclude that

$$\mathbb{P}\big(X(t_1,x) \in K_1(M) + K_2(M,\varepsilon) + K_3(M,\varepsilon)\big) \ge 1-\varepsilon$$

for all $\|x\| \le M$. □

Theorem 6.14 (Theorem 2.6 of [14])**.** *Suppose that E has a Schauder basis, that $S(t)$ is compact for $t \ge t_1 > 0$ and that (M1), (M2′), (M3′) and (M5) are satisfied. Then there exists an invariant measure for* (6.2).

Proof. Let $K(M,\varepsilon)$ be as in Lemma 6.13. For $t > t_1$, we write $P(t,x,\Gamma)$ for the Markov transition probabilities

$$P(t,x,\Gamma) = \mathbb{P}\big(X(t,x) \in \Gamma\big),$$

for $x \in E$ and $\Gamma \in \mathbb{B}(E)$. Then

$$\begin{aligned}\mathbb{P}\big(X(t,x) \in K(M,\varepsilon)\big) &= \mathbb{E}\big[P\big(t_1, X(t-t_1,x), K(M,\varepsilon)\big)\big] \\ &\geq \mathbb{E}\big[P\big(t_1, X(t-t_1,x), K(M,\varepsilon)\big) \cdot \mathbb{1}_{\{\|X(t-t_1,x)\| \leq M\}}\big].\end{aligned}$$

By Lemma 6.13,

$$\mathbb{P}\big(X(t,x) \in K(M,\varepsilon) \mid \|X(t-t_1,x)\| \leq M\big) \geq (1-\varepsilon)\,\mathbb{P}\big(\|X(t-t_1,x)\| \leq M\big),$$

so

$$\frac{1}{T}\int_{t_1}^{T+t_1} \mathbb{P}\big(X(t,x) \in K(M,\varepsilon)\big)\,\mathrm{d}t \;\geq\; \frac{1-\varepsilon}{T}\int_0^T \mathbb{P}\big(\|X(t,x)\| \leq M\big)\,\mathrm{d}t.$$

Now by $(M5)$, taking M large enough we see that the family

$$\frac{1}{T}\int_{t_1}^{T+t_1} P(t,x,\cdot)\,\mathrm{d}t, \quad T \geq t_1$$

is tight. Hence by Krylov-Bogoliubov (Theorem 6.3) there exists an invariant measure. □

6.5 Delay problems

We are now in a position to consider the delayed problem (6.1)

$$\mathrm{d}X(t) = \big[AX(t) + F\big(X(t), X_t\big)\big]\,\mathrm{d}t + G\big(X(t), X_t\big)\,\mathrm{d}W(t), \quad t \geq 0. \tag{6.1}$$

Let $p \in [2,\infty)$ and recall the definition of the delayed Cauchy problem on the space $\mathcal{E} = E \times L^p(-1,0;E)$ from Section 2.4.1.

The delay operator $\mathcal{A}$ on $\mathcal{E}$ is defined as

$$\mathcal{A} = \begin{pmatrix} A & 0 \\ 0 & \frac{\mathrm{d}}{\mathrm{d}s}\end{pmatrix} \text{ with domain } D(\mathcal{A}) := \left\{\begin{pmatrix} x \\ \phi\end{pmatrix} \in D(A) \times W^{1,p}(-1,0;E) : \phi(0) = x\right\}$$

and we introduce new functions $\mathcal{F} : \mathcal{E} \to \mathcal{E}$ and $\mathcal{G} : \mathcal{E} \to \mathcal{L}(H,E) \times \mathcal{L}(H, L^p(-1,0;E))$ given by

$$\mathcal{F}\begin{pmatrix} x \\ \phi\end{pmatrix} = \begin{pmatrix} F(x,\phi) \\ 0\end{pmatrix} \qquad \text{and} \qquad \mathcal{G}\begin{pmatrix} x \\ \phi\end{pmatrix} = \begin{pmatrix} G(x,\phi) \\ 0\end{pmatrix}.$$

This allows us to rewrite (6.1) as a new problem on $\mathcal{E}$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathfrak{X}(t) = \big[\mathcal{A}\,\mathfrak{X}(t) + \mathcal{F}(\mathfrak{X}(t))\big]\,\mathrm{d}t + \mathcal{G}(\mathfrak{X}(t))\,\mathrm{d}W(t) \quad t \geq 0. \tag{6.22}$$

By [25, Theorem 4.3], for any $(X, \Phi)^T \in L^p(\Omega; \mathcal{E})$ there exists a unique solution $\mathfrak{X}(t, X, \Phi)$ of (6.22). Moreover, by [25, Theorem 4.7] (the stochastic analogue of Theorem 2.18), (6.1) has a solution $X(t)$ if and only if (6.22) has a solution $\mathfrak{X}(t)$, and in this case

$$\mathfrak{X}(t) = \begin{pmatrix} X(t) \\ X_t \end{pmatrix} \text{ a.s..}$$

As in Section 2.4.1, $\mathcal{A}$ generates the *delay semigroup* $\mathcal{S}(t)$ on $\mathcal{E}$ given by

$$\mathcal{S}(t) := \begin{pmatrix} S(t) & 0 \\ S_t & L_t \end{pmatrix} \tag{6.23}$$

where

$$(S_t x)(s) := \begin{cases} 0 & t+s<0 \\ S(t+s)x & t+s \geq 0 \end{cases} \quad \text{and} \quad (L_t f)(s) := \begin{cases} f(t+s) & t+s<0 \\ 0 & t+s \geq 0. \end{cases}$$

By Theorem 2.19, $\mathcal{S}(t)$ is compact for all $t > 1$ if and only if $S(t)$ is compact for all $t > 0$. Hence we can appeal to Theorem 6.14 with any $t_1 > 1$.

We will need the following assumptions on F, G and the unique solution $X(t, x, \phi)$ for initial conditions $(x, \phi) \in \mathcal{E}$.

(DM2′) $F : \mathcal{E} \to E$ is Lipschitz and admits a factorisation $F = C_1 \circ F_1$ where $F_1 : \mathcal{E} \to E_1$ for some Banach space E_1 and $C_1 \in \mathcal{L}(E_1, E)$ is compact.

(DM3′) $G : \mathcal{E} \to \gamma(H, E)$ is Lipschitz and admits a factorisation $G = C_2 \circ G_1$ where $G_1 : \mathcal{E} \to \gamma(H, E_2)$ for some Banach space E_2 and $C_2 \in \mathcal{L}(E_2, E)$ is compact.

(DM5) For all $(x, \phi) \in \mathcal{E}$ and $\varepsilon > 0$, there exists $M \geq 0$ such that for all $T > 1$,

$$\frac{1}{T}\int_0^T \mathbb{P}\big(\|X(t, x, \phi)\| \geq M\big)\,\mathrm{d}t < \varepsilon,$$

and

$$\frac{1}{T}\int_0^T \mathbb{P}\Big(\int_{-1}^0 \|\widehat{X}(t+s, x, \phi)\|^p\,\mathrm{d}s \geq M\Big)\,\mathrm{d}t < \varepsilon,$$

where $\widehat{X}(\cdot, x, \phi) : [-1, \infty) \to E$ is the extension of $X(t, x, \phi)$ to $[-1, \infty)$ by ϕ

$$\widehat{X}(t, x, \phi) = \begin{cases} \phi(t) & -1 \leq t < 0, \\ X(t, x, \phi) & 0 \leq t < \infty. \end{cases}$$

Theorem 6.15. *Suppose that (M1), (DM2′), (DM3′) and (DM5) hold, that E has a Schauder basis and that A generates an immediately compact C_0-semigroup $S(t)$. Then there exists an invariant measure for* (6.1)*.*

Proof. We will show that $\mathcal{E}$ and $\mathcal{S}(t)$ satisfy (M1) and $\mathcal{F}$, $\mathcal{G}$ and $\mathfrak{X}$ satisfy (M2′), (M3′) and (M5) respectively, then apply Theorem 6.14 to equation (6.22).

First however we note that by [19] if E is a UMD space with a Schauder basis then so is $L^p(-1,0;E)$ and hence so is $\mathcal{E}$.

(M1) If E is a UMD space of type 2 then so is $L^p(-1,0;E)$ for any $2 \leq p < \infty$ and hence so is $\mathcal{E}$. As $S(t)$ is compact for all $t > 0$, by Theorem 2.19, $\mathcal{S}(t)$ is compact for all $t \geq t_1$ for any $t_1 > 1$. By [11, Proposition 3.19] we have $\sigma(\mathcal{A}) = \sigma(A)$ and $\|S(t)\| \leq Me^{-\lambda t}$ for some $\lambda > 0$ so $\sigma(\mathcal{A}) = \sigma(A) \subset \{z \in \mathbb{C} : \mathrm{Re} z \leq -\lambda\}$, therefore by [39, Theorem V.1.10], $\mathcal{S}(t)$ is uniformly exponentially stable.

(M2′) Write $\mathcal{E}_1 := E_1 \times L^p(-1,0;E_1)$ and

$$\mathcal{F}_1 : \begin{pmatrix} x \\ \phi \end{pmatrix} \mapsto \begin{pmatrix} F_1(x,\phi) \\ 0 \end{pmatrix} : \mathcal{E} \to \mathcal{E}_1,$$

$$\mathcal{C}_1 = \begin{pmatrix} C_1 & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{L}(\mathcal{E}_1, \mathcal{E}),$$

then $\mathcal{C}_1$ is compact by (DM2′) and $\mathcal{F} = \mathcal{C}_1 \circ \mathcal{F}_1$. $\mathcal{F}$ is easily seen to be Lipschitz so (M2′) holds for (6.22).

(M3′) Similarly, write $\mathcal{E}_2 := E_1 \times L^p(-1,0;E_2)$ and

$$\mathcal{G}_1 : \begin{pmatrix} x \\ \phi \end{pmatrix} \mapsto \begin{pmatrix} G_1(x,\phi) \\ 0 \end{pmatrix} : \mathcal{E} \to \gamma(H, \mathcal{E}_2),$$

$$\mathcal{C}_2 = \begin{pmatrix} C_2 & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{L}(\mathcal{E}_2, \mathcal{E}),$$

then $\mathcal{C}_2$ is compact by (DM3′) and $\mathcal{G} = \mathcal{C}_2 \circ \mathcal{G}_1$. $\mathcal{G}$ is easily seen to be Lipschitz so (M3′) holds for (6.22).

(M5) Let $(x,\phi) \in \mathcal{E}$ and $\varepsilon > 0$. By [25, Theorem 4.7] we have $\|\mathfrak{X}(t)\|_{\mathcal{E}} = \|X(t)\|_E + \|X(t+\cdot)\|_{L^p(-1,0;E)}$, and hence for $M \geq 0$

$$\|\mathfrak{X}(t)\|_{\mathcal{E}} \geq M \quad \Longrightarrow \quad \left[\|X(t)\|_E \geq \frac{M}{2} \quad \text{or} \quad \|X(t+\cdot)\|_{L^p(-1,0;E)} \geq \frac{M}{2} \right]$$

so

$$\mathbb{P}\big(\|\mathfrak{X}(t)\|_{\mathcal{E}} \geq M\big) \leq \mathbb{P}\big(\|X(t)\|_E \geq \tfrac{M}{2}\big) + \mathbb{P}\big(\|X(t+\cdot)\|_{L^p(-1,0;E)} \geq \tfrac{M}{2}\big).$$

Then

$$\begin{aligned}\frac{1}{T}\int_0^T \mathbb{P}\big(\|\mathfrak{X}(t)\|_{\mathcal{E}} \geq M\big)\,\mathrm{d}t &\leq \frac{1}{T}\int_0^T \mathbb{P}\big(\|X(t)\|_E \geq \tfrac{M}{2}\big)\,\mathrm{d}t \\ &\qquad + \frac{1}{T}\int_0^T \mathbb{P}\big(\|X(t+\cdot)\|_{L^p(-1,0;E)} \geq \tfrac{M}{2}\big)\,\mathrm{d}t \\ &< \varepsilon\end{aligned}$$

for sufficiently large M.

Hence by Theorem 6.14, there exists an invariant measure for $\mathfrak{X}(t)$, which is therefore also an invariant measure for $X(t)$. □

6.6 Example

We turn now to a perturbed stochastic functional differential equation on $\mathbb{R}^d$. Consider the problem

$$\frac{\partial}{\partial t}u(t) \quad = \quad Bu(t) + f(u(t), u_t) + \sum_{n=1}^{d} g_n(u(t), u_t)\frac{\partial W_n(t)}{\partial t}, \qquad t \geq 0. \tag{6.24}$$

Where $B \in \mathcal{L}(\mathbb{R}^d)$, $2 \leq p < \infty$, $E := \mathbb{R}^d$ and $\mathcal{E} := \mathbb{R}^d \times L^p(-1,0;\mathbb{R}^d)$, then E and $\mathcal{E}$ are of type 2 and by [81, Section 2.5.5], both have Schauder bases. $W_n(t)$ is a standard real Brownian motion for $n = 1, \ldots, d$ and $f, g_n : \mathcal{E} \to \mathbb{R}^d$ are Lipschitz functions with

$$\big|g_n(x,\phi) - g_n(y,\psi)\big|_{\mathbb{R}^d} \leq L_{g_n}\big\|(x-y, \phi-\psi)\big\|_{\mathcal{E}}$$

for $(x,\phi), (y,\psi) \in \mathcal{E}$. We define an ℓ^2-cylindrical Brownian motion by

$$W(t)e_n = W_n(t).$$

Assume that $\sigma(B) \subset \{z : \mathrm{Re}(z) < 0\}$, then $\exp(Bt)$ is a norm continuous, stable and compact semigroup on $\mathbb{R}^d$.

Define the process $U : \mathbb{R}_+ \times \Omega \to E$ by $U(t) = u(t)$ and the functions $F : \mathcal{E} \to E$, $G : \mathcal{E} \to \mathcal{L}(\ell^2, E)$ by

$$F(x,\phi) = f(x,\phi) \quad \text{and} \quad \big(G(x,\phi)e_n\big) = g_n(x,\phi)$$

for $x \in E$, $\phi \in L^p(-1,0;E)$ and e_n the standard basis of ℓ^2. Then F and G are Lipschitz and since each has finite dimensional range, there exists compact decompositions as in (DM2′) and (DM3′).

By all of the above and [25, Theorem 4.3], the delayed stochastic Cauchy problem

$$\mathrm{d}U(t) = \big[BU(t) + F(U(t), U_t)\big]\,\mathrm{d}t + G(U(t), U_t)\,\mathrm{d}W(t), \quad t \geq 0 \tag{6.25}$$

has a unique solution for any initial conditions $(x, \phi) \in \mathcal{E}$.

Then if (DM5) holds, i.e. if for all $x \in E$, $\phi \in L^p(-1,0;E)$ and $\varepsilon > 0$, there exists $M \geq 0$ such that for all $T > 1$,

$$\frac{1}{T}\int_0^T \mathbb{P}\big(\|U(t,x,\phi)\| \geq M\big)\,\mathrm{d}t < \varepsilon,$$

and

$$\frac{1}{T}\int_0^T \mathbb{P}\Big(\int_{-1}^0 \|\widehat{U}(t+s,x,\phi)\|^p\,\mathrm{d}s \geq M\Big)\,\mathrm{d}t < \varepsilon,$$

then (6.25) satisfies all conditions of Theorem 6.15, and therefore there exists an invariant measure for (6.24). □

Bibliography

[1] P Acquistapace. Evolution operators and strong solutions of abstract linear parabolic equations. *Differential Integral Equations*, 1:433–457, 1988.

[2] P Acquistapace and B Terreni. On fundamental solutions for abstract parabolic equations. In *Differential equations in Banach spaces, Proceedings of a conference held at the University of Bologna, 1985*, Lecture Notes in Math., 1223, pages 1–11, Berlin, 1986. Springer.

[3] P Acquistapace and B Terreni. A unified approach to abstract linear nonautonomous parabolic equations. *Rend. Sem. Mat. Univ. Padova*, 78:47–107, 1987.

[4] H Amann. *Linear and quasilinear parabolic problems*, volume 1, Abstract linear theory of *Monographs in Mathematics, 89.* Birkhäuser Boston, Inc., Boston, MA, 1995.

[5] W Arendt, C Batty, M Hieber, and F Neubrander. *Vector-valued Laplace transforms and Cauchy problems.* Monographs in Mathematics, 96. Birkhäuser Verlag, Basel, 2001.

[6] W Arendt and C J K Batty. Asymptotically almost periodic solutions of inhomogeneous cauchy problems on the half-line. *Bull. London Math. Soc.*, 31:291–304, 1999.

[7] D Bahuguna and S Abbas. Pseudo almost periodic mild solutions of retarded functional differential equations. *Glob. J. Pure Appl. Math.*, 3:27–36, 2007.

[8] S Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fund. Math.*, 3:133–181, 1922.

[9] M Baroun, T Boulite, S and; Diagana, and L Maniar. Almost periodic solutions to some semilinear non-autonomous thermoelastic plate equations. *J. Math. Anal. Appl.*, 349:47–84, 2009.

[10] M Baroun, L Maniar, and R Schnaubelt. Almost periodicity of parabolic evolution equations with inhomogeneous boundary values. *Integral Equations Operator Theory*, 65(2):169–193, 2009.

[11] A Bátkai and S Piazzera. *Semigroups for delay equations.* Research Notes in Mathematics, 10. A K Peters, Ltd., Wellesley, MA, 2005.

[12] C J K Batty, W Hutter, and F Räbiger. Almost periodicity of mild solutions of inhomogeneous periodic cauchy problems. *J. Differential Equations*, 156:309–327, 1999.

[13] R Bellman. The stability of solutions of linear differential equations. *Duke Math. J.*, 10:643–647, 1943.

[14] J Bierkens, O van Gaans, and S Lunel. Existence of an invariant measure for stochastic evolutions driven by an eventually compact semigroup. *J. Evol. Equ.*, 9(4):771–786, 2009.

[15] P Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., New York, 1999.

[16] S Bochner. Beiträge zur theorie der fastperiodischen funktionen. *Math. Ann.*, 96(1):119–147, 1927.

[17] V Bogachev. *Gaussian measures.* Mathematical Surveys and Monographs, 62. American Mathematical Society, Providence, RI, 1998.

[18] H Bohr. Zur theorie der fastperiodischen funktionen. *Acta Math.*, 46(1-2):101–214, 1925.

[19] J Bourgain. Some remarks on Banach spaces in which martingale difference sequences are unconditional. *Ark. Mat.*, 21(2):163–168, 1983.

[20] Z Brzeźniak. Stochastic partial differential equations in m-type 2 Banach spaces. *Potential Anal.*, 4(1):1–45, 1995.

[21] Z Brzeźniak. On stochastic convolution in Banach spaces and applications. *Stoch. Stoch. Rep.*, 61(3-4):245–295, 1997.

[22] Z Brzeźniak, L Hongwei, and I Simão. Invariant measures for stochastic evolution equations in M-type 2 Banach spaces. *J. Evol. Equ.*, 10(4):785–810, 2010.

[23] Z Brzeźniak and J van Neerven. Stochastic convolution in separable Banach spaces and the stochastic linear Cauchy problem. *Studia Math.*, 143(1):43–74, 2000.

[24] D Burkholder. A geometrical characterization of Banach spaces in which martingale difference sequences are unconditional. *Ann. Probab.*, 9(6):997–1011, 1981.

[25] S Cox and M Górajski. Vector-valued stochastic delay equations - A semigroup approach. *Semigroup Forum*, 82(3):389–411, 2010.

[26] R Curtain and P Falb. Stochastic differential equations in Hilbert space. *J. Differential Equations*, 10:412–430, 1971.

[27] G Da Prato. *An introduction to infinite-dimensional analysis.* Universitext. Springer-Verlag, Berlin, 2006.

[28] G Da Prato, D Gątarek, and J Zabczyk. Invariant measures for semilinear stochastic equations. *Stochastic Anal. Appl.*, 10(4):387–408, 1992.

[29] G Da Prato, S Kwapień, and J Zabczyk. Regularity of solutions of linear stochastic equations in Hilbert spaces. *Stochastics*, 23(1):1–23, 1987.

[30] G Da Prato and L Tubaro. Some results on semilinear stochastic differential equations in Hilbert spaces. *Stochastics*, 15(4):271–281, 1985.

[31] G Da Prato and C Tudor. Periodic and almost periodic solutions for semilinear stochastic equations. *Stochastic Anal. Appl.*, 13(1):13–33, 1995.

[32] G Da Prato and J Zabczyk. A note on stochastic convolution. *Stochastic Anal. Appl.*, 10(2):143–153, 1992.

[33] G Da Prato and J Zabczyk. *Stochastic equations in infinite dimensions.* Encyclopedia of Mathematics and its Applications, 44. Cambridge University Press, Cambridge, 1992.

[34] G Da Prato and J Zabczyk. *Ergodicity for infinite-dimensional systems.* London Mathematical Society Lecture Note Series, 229. Cambridge University Press, Cambridge, 1996.

[35] D Dawson. Stochastic evolution equations. *Math. Biosci.*, 15:287–316, 1972.

[36] D Dawson. Stochastic evolution equations and related measure processes. *J. Multivariate Anal.*, 5:1–52, 1975.

[37] T Diagana. Existence of weighted pseudo-almost periodic solutions to some classes of nonautonomous partial evolution equations. *Nonlinear Anal.*, 74:600–615, 2011.

[38] A Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 322(8):549–560, 1905.

[39] K Engel and R Nagel. *One-parameter semigroups for linear evolution equations.* Graduate Texts in Mathematics, 194. Springer-Verlag, New York, 2000.

[40] A Es-Sarhir and B Farkas. Invariant measures and regularity properties of perturbed Ornstein-Uhlenbeck semigroups. *J. Differential Equations*, 233(1):87–104, 2007.

[41] A Es-Sarhir, M Scheutzow, and O van Gaans. Invariant measures for stochastic functional differential equations with superlinear drift term. *Differential Integral Equations*, 23(1-2):189–200, 2010.

[42] A Es-Sarhir and W Stannat. Invariant measures for semilinear SPDE's with local Lipschitz drift coefficients and applications. *J. Evol. Equ.*, 8:129–154, 2008.

[43] O van Gaans. A series approach to stochastic differential equations with infinite dimensional noise. *Integral Equations Operator Theory*, 51(3):435–458, 2005.

[44] D Garling. Brownian motion and UMD-spaces. In *Probability and Banach spaces (Zaragoza, 1985)*, Lecture Notes in Mathematics 1221, pages 36–49. Springer, Berlin, 1986.

[45] D Garling. *Inequalities: a journey into linear analysis.* Cambridge University Press, Cambridge, 2007.

[46] D Gątarek and B Gołdys. On weak solutions of stochastic equations in Hilbert spaces. *Stochastics Stochastics Rep.*, 46(1-2):41–51, 1994.

[47] G Gühring and F Räbiger. Asymptotic properties of mild solutions of nonautonomous evolution equations with applications to retarded differential equations. *Abstr. Appl. Anal.*, 4(3):169–194, 1999.

[48] G Gühring, F Räbiger, and R Schnaubelt. A characteristic equation for non-autonomous partial functional differential equations. *Abstr. Appl. Anal.*, 4(3):169–194, 1999.

[49] J Hale. *Functional differential equations*. Applied Mathematical Sciences, Vol. 3. Springer-Verlag, New York-Heidelberg, 1978.

[50] J Hale and J Kato. Phase space for retarded equations with infinite delay. *Funkcial. Ekvac.*, 21(1):11–41, 1978.

[51] C Heil. *A basis theory primer*. Applied and Numerical Harmonic Analysis. Birkhäuser, New York, 2011.

[52] H Henríquez. Regularity of solutions of abstract retarded functional-differential equations with unbounded delay. *Nonlinear Anal.*, 28(3):513–531, 1997.

[53] Y Hino, S Murakami, and T Naito. *Functional-differential equations with infinite delay*. Springer-Verlag, Berlin, 1991.

[54] T Hytönen and M Veraar. R-boundedness of smooth operator-valued functions. *Integral Equations Operator Theory*, 63(3):373–402, 2009.

[55] L Koralov and Y Sinai. *Theory of probability and random processes*. Springer, Berlin, 2007.

[56] N Krasovskiĭ. *Stability of motion. Applications of Lyapunov's second method to differential systems and equations with delay*. Stanford University Press, Stanford, Calif., 1963.

[57] P Kunstmann and L Weis. *Maximal L_p-regularity for Parabolic Equations, Fourier Multiplier Theorems and H^∞-functional Calculus*, pages 65–311. Lecture Notes in Mathematics, 1855. Springer-Verlag, Berlin, 2004.

[58] M Kunze and J van Neerven. Continuous dependence on the coefficients and global existence for stochastic reaction diffusion equations. *ArXiv e-prints:1104.4258*, 2011.

[59] S Kwapień and W Woyczyński. *Random series and stochastic integrals: single and multiple.* Probability and its Applications. Birkhäuser, Boston, 1992.

[60] K Liu. Existence, uniqueness, and asymptotic behaviour of mild solutions to stochastic functional differential equations in Hilbert spaces. *J. Differential Equations*, 181(1):72–91, 2002.

[61] K Liu. Stochastic retarded evolution equations: Green operators, convolutions, and solutions. *Stoch. Anal. Appl.*, 26(3):624–650, 2008.

[62] A Lunardi. *Interpolation Theory.* Appunti. Scuola Normale Superiore, Pisa, 1999.

[63] L Maniar and R Schnaubelt. Almost periodicity of inhomogeneous parabolic evolution equations. In *Evolution equations*, Lecture Notes in Pure and Appl. Math., 234, pages 299–318. Dekker, New York, 2003.

[64] M Mokhtar-Kharroubi. On permanent regimes for non-autonomous linear evolution equations in Banach spaces with applications to transport theory. *Kinet. Relat. Models*, 3(3):473–499, 2010.

[65] J van Neerven. γ-radonifying operators - a survey. In *Proceedings of the CMA*, volume 44, pages 1–62, 2010.

[66] J van Neerven and M Riedle. A Semigroup Approach to Stochastic Delay Equations in Spaces of Continuous Functions. *Semigroup Forum*, 74:227–239, 2007.

[67] J van Neerven and M Veraar. On the action of Lipschitz functions on vector-valued random sums. *Arch. Math.*, 85(6):544–553, 2005.

[68] J van Neerven, M Veraar, and L Weis. Stochastic integration in UMD Banach spaces. *Ann. Probab.*, 35(4):1438–1478, 2007.

[69] J van Neerven, M Veraar, and L Weis. Stochastic evolution equations in UMD Banach spaces. *J. Funct. Anal.*, 255(4):940–993, 2008.

[70] J van Neerven and L Weis. Stochastic integration of functions with values in a Banach space. *Studia Math.*, 166(2):131–170, 2005.

[71] J van Neerven and L Weis. Invariant measures for the linear stochastic Cauchy problem and R-boundedness of the resolvent. *J. Evol. Eqn.*, 6(2):205–228, 2006.

[72] B Øksendal. *Stochastic Differential Equations: An Introduction with Applications (Sixth edition).* Springer, Berlin, 2003.

[73] M Ondreját. Uniqueness for stochastic evolution equations in Banach spaces. *Dissert. Math.*, 426, 2004.

[74] G Pedersen. *Analysis now.* Graduate Texts in Mathematics, 118. Springer-Verlag, New York, 1989.

[75] A Pietsch. *History of Banach spaces and linear operators.* Birkhäuser, Boston, 2007.

[76] M Riedle. Solutions of affine stochastic functional differential equations in the state space. *J. Evol. Equ.*, 8(1):71–97, 2008.

[77] M Riedle and J van Neerven. A semigroup approach to stochastic delay equations in spaces of continuous functions. *Semigroup Forum*, 74(2):227–239, 2007.

[78] W Rudin. *Functional analysis.* International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York, second edition, 1991.

[79] W Ruess and W Summers. Weak almost periodicity and the strong ergodic limit theorem for periodic evolution systems. *J. Funct. Anal.*, 94(1):177–195, 1990.

[80] R Schnaubelt. Asymptotic behaviour of parabolic nonautonomous evolution equations. In *unctional analytic methods for evolution equations*, Lecture Notes in Math., 1855, pages 401–472. Springer, Berlin, 2004.

[81] H Triebel. *Theory of function spaces.* Monographs in Mathematics, 78. Birkhäuser Verlag, Basel, 1983.

[82] H Triebel. *Theory of function spaces. II.* Monographs in Mathematics, 84. Birkhäuser Verlag, Basel, 1992.

[83] H Triebel. *Interpolation theory, function spaces, differential operators.* Johann Ambrosius Barth, Heidelberg, second edition, 1995.

[84] N Vakhania, V Tarieladze, and S Chobanyan. *Probability Distributions on Banach Spaces.* Mathematics and its Applicatons (Soviet Series), vol. 14. D. Reidel Publishing Co., Dordrecht, 1987.

[85] C Vârsan. Stochastic differential equations and asymptotic almost periodic solutions. *Rev. Roumaine Math. Pures Appl.*, 35(5):485–493, 1990.

[86] M Veraar. Non-autonomous stochastic evolution equations and applications to stochastic partial differential equations. *J. Evol. Equ.*, 10(1):85–127, 2010.

[87] M Veraar and J Zimmerschied. Non-autonomous stochastic Cauchy problems in Banach spaces. *Studia Math*, 185(1):1–34, 2008.

[88] G Webb. Autonomous nonlinear functional differential equations and nonlinear semigroups. *J. Math. Anal. Appl.*, 46:1–12, 1974.

[89] L Weis. Operator-valued Fourier multiplier theorems and maximal L_p-regularity. *Math. Ann.*, 319(4):735–758, 2001.

[90] D Williams. *Probability with martingales.* Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

[91] A Yagi. Abstract quasilinear evolution equations of parabolic type in Banach spaces. *Boll. Un. Mat. Ital. B (7)*, 5(2):341–368, 1991.