Long term behavior of dichotomous stochastic differential equations in Hilbert spaces

Onno van Gaans and Sjoerd Verduyn Lunel*

Abstract

We study existence of invariant measures for semilinear stochastic differential equations in Hilbert spaces. The noise is infinite dimensional, white in time, and colored in space. We show that if the equation is exponentially dichotomous in the sense that the semigroup generated by the linear part is hyperbolic and the Lipschitz constants of the nonlinearities are not too large, then existence of a solution with bounded mean squares implies existence of an invariant measure. Moreover, we show that every bounded solution satisfies a certain 'Cauchy condition', which implies that its distributions converge weakly to a limit distribution.

1 Introduction

One of the main interests in the study of long term behavior of stochastic evolution equations is existence of invariant measures. In this paper we study existence of invariant measures for a class of dichotomous semilinear stochastic evolution equations in Hilbert spaces. Such equations arise, for instance, in the study of stochastic delay differential equations, neutral functional differential equations, and classes of partial differential equations.

Let H be a separable real Hilbert space and consider an equation of the form

$$dX(t) = \big(AX(t) + F(X(t))\big)dt + \sum_{i=1}^{\infty} \Phi_i(X(t))dW_i(t), \quad t \geq 0, \tag{1}$$

where A is the generator of a C_0-semigroup $(S(t))_{t\geq 0}$ on H, the maps F and $\Phi_1, \Phi_2, \ldots$ are Lipschitz continuous from H to H, and $(W_i)_{i=0}^{\infty}$ is a sequence of independent scalar Wiener processes (Brownian motions). We will consider solutions of (1) that are continuous functions from $[0,\infty)$ to the Bochner space $L^2(\Omega; H)$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is the underlying probability space.

*the research of both authors is supported by an N.W.O. PIONIER-grant under 600-61-410.

If X is a solution of (1), then for each $t \geq 0$, $X(t)$ is a random variable with values in H and its distribution is a Borel probability measure on H. If $\mu_{X(t)} = \mu$ for all $t \geq 0$, then μ is called a *stationary distribution* for (1). This definition can be extended to the notion of *invariant measure*, which is a Borel probability measure on H that is invariant under the dual of the transition semigroup associated to (1). See Section 8 for further information.

A necessary condition for existence of an invariant measure for (1) is existence of a solution that is bounded in probability. If H is finite dimensional, then existence of such a bounded solution is sufficient as well. For the case that H is infinite dimensional, several instances have been established where existence of a bounded solution implies existence of an invariant measure. The existing results are mainly tailored for parabolic partial differential equations. Da Prato, Gatarek, and Zabczyk showed in [3] (also in [4, Theorem 11.29, p.338]) that existence of a bounded solution of (1) implies existence of an invariant measure when $S(t)$ is compact for every $t > 0$. Similar results are known for dissipative systems (see [3, 12, 4]) and linear equations ([3, 4]). Delay equations and several other interesting applications, however, do in general not fit in these frameworks.

It is our aim to replace compactness or dissipativity assumptions on the underlying deterministic semiflow by a condition of an exponential dichotomy. We will show that if the semigroup $(S(t))_{t\geq 0}$ is hyperbolic and the Lipschitz constants of F and Φ_i, $i \geq 1$, are sufficiently small, and if $X : [0, \infty) \to L^2(\Omega; X)$ is a solution of (1) such that $t \mapsto \mathbb{E} \|X(t)\|^2$ is bounded on $[0, \infty)$, then (i) for every $\varepsilon > 0$ there is a compact $K_\varepsilon \subset H$ such that $\mathbb{P}(X(t) \in K_\varepsilon) \geq 1 - \varepsilon$ for all $t \geq 0$, (ii) $\mu_{X(t)}$ converges to a Borel probability measure μ on H as $t \to \infty$, and (iii) the measure μ from (ii) is a stationary distribution for (1). Our method resembles methods for linear equations, and we believe that some of our steps are of independent interest.

The paper consists of three parts. In Sections 2–3 we complete the introductory part by a description of the precise setting, the statement of the main result, and an outline of its proof. The second part of the paper is concerned with the proof of the main result, which is divided over Sections 5–8. Finally, we illustrate the main result in Section 9 by means of examples of stochastic delay differential equations.

2 Setting and preliminaries

Throughout the paper we will consider the following situation. For sake of reference we call it

Situation (*). Let H be a separable real Hilbert space, let $(S(t))_{t\geq 0}$ be a strongly continuous semigroup of bounded linear operators on H with generator A, and let $F : H \to H$ be a Lipschitz continuous function with Lipschitz constant L_F. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration in $\mathcal{F}$, and let $(W_i)_{i=1}^{\infty}$ be a sequence of independent normalized scalar Wiener processes on Ω with respect to $(\mathcal{F}_t)_{t\geq 0}$. Let $\Phi = (\Phi_i)_{i=1}^{\infty}$ be a sequence of Lipschitz continuous functions $\Phi_i : H \to H$ with Lipschitz constants L_{Φ_i} such that

$$L_\Phi := \Big(\sum_{i=1}^{\infty} L_{\Phi_i}^2 \Big)^{1/2} < \infty \quad \text{and} \quad \sum_{i=1}^{\infty} \|\Phi_i(0)\|^2 < \infty.$$

Let $X_0 \in L^2(\Omega; H)$ be $\mathcal{F}_0$-measurable.

We interpret equation (1) with initial value $X(0) = X_0$ as the stochastic integral equation

$$X(t) = S(t)X_0 + \int_0^t S(t-s)F(X(s))ds + \int_0^t S(t-s)\Phi(X(s))dW(s), \quad t \geq 0. \tag{2}$$

The stochastic integral in (2) is defined by

$$\int_0^t S(t-s)\Phi(X(s))dW(s) = \sum_{i=1}^{\infty} \int_0^t S(t-s)\Phi_i(X(s))dW_i(s), \quad t \geq 0, \tag{3}$$

where the latter integrals are H-valued Itô integrals. Note that

$$\begin{aligned}
&\int_0^t \mathbb{E} \sum_i \|S(t-s)\Phi_i(X(s))\|^2 \, ds \\
&\quad \leq \sup_{0\leq s\leq t} \|S(t)\|^2 \int_0^t \sum_{i=1}^{\infty} \mathbb{E}\, [2\|\Phi_i(0)\|^2 + 2L_{\Phi_i}^2 \|X(s)\|^2] ds,
\end{aligned}$$

so that the series is indeed convergent in $L^2(\Omega; H)$ if $t \mapsto \|X(t)\|^2$ is bounded on each bounded interval. The series definition (3) of the stochastic integral is discussed in [7]. It actually turns out that the integral defined by (3) is equivalent to the integral for Hilbert-Schmidt operator valued processes described in [4]. We summarize the relevant properties from [7] below.

Lemma 2.1. *Let H be a separable real Hilbert space and let $(W_i)_{i=1}^{\infty}$ be a sequence of independent normalized scalar Wiener processes on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$ in $\mathcal{F}$. Let $(\Psi_i)_{i=1}^{\infty}$ be a sequence of adapted piecewise uniformly continuous functions from $[0, \infty)$ to $L^2(\Omega; H)$ such that $\int_0^t \sum_{i=1}^{\infty} \mathbb{E}\, \|\Psi_i(s)\|^2 ds < \infty$ for every $t \geq 0$. Then:*

(i) $t \mapsto \int_0^t \Psi(s)dW(s) := \sum_{i=1}^{\infty} \int_0^t \Psi_i(s)dW_i(s) : [0, \infty) \to L^2(\Omega; H)$ is continuous and adapted,

(ii) (Itô isometry)

$$\mathbb{E}\,\|\int_0^t \Psi(s)dW(s)\|^2 = \int_0^t \mathbb{E}\sum_{i=1}^{\infty}\|\Psi_i(s)\|^2 ds, \quad t \geq 0,$$

(iii) if $T : H \to H$ is a bounded linear operator, then for each $t \geq 0$ one has that

$$\sum_{i=1}^{\infty}\int_0^t T \circ (\Psi_i(s))dW_i(s) = T \circ \Big(\sum_{i=1}^{\infty}\int_0^t \Psi_i(s)dW_i(s)\Big).$$

By a solution of (1) or (2) we always mean a *mild solution*, which is a continuous function $t \mapsto X(t)$ from $[0,\infty)$ into the Bochner space $L^2(\Omega; H)$ which is adapted to $(\mathcal{F}_t)_{t\geq 0}$ and satisfies the integral equation (2). Note that the assumptions on X indeed imply that the integrals in (2) are well-defined. In Situation (*) equation (2) has a unique solution. The solution is also weakly unique, which will be important in the sequel. The definition of weak uniqueness is as follows. Assume Situation (*) and let X be the solution of (2). Let $(\Omega', \mathcal{F}', \mathbb{P}')$ be another probability space and let $(W_i')_{i=1}^{\infty}$ be a sequence of independent normalized scalar Wiener processes on Ω' with respect to a filtration $(\mathcal{F}_t')_{t\geq 0}$ in $\mathcal{F}'$. Let $X_0' \in L^2(\Omega'; H)$ be $\mathcal{F}_0'$-measurable with the same distribution as X_0, and let $X' : [0,\infty) \to L^2(\Omega'; H))$ be the continuous, adapted solution of

$$X'(t) = S(t)X_0' + \int_0^t S(t-s)F(X'(s))ds + \int_0^t S(t-s)\Phi(X'(s))dW'(s),$$

$t \geq 0$. If for any such an X' the random variables $X(t)$ and $X'(t)$ have the same distribution for all $t \geq 0$, then X is called *weakly unique* or *unique in law.* It turns out that in the setting of Situation (*) even a stronger conclusion holds true, namely that the random vectors $(X(t_1),\dots,X(t_n))$ and $(X'(t_1),\dots,X'(t_n))$ have the same distribution for any $0 \leq t_1 \leq \cdots \leq t_n$.

Theorem 2.2 (existence and uniqueness). *Assume Situation (*). There exists one and only one continuous adapted function $X : [0,\infty) \to L^2(\Omega; H)$ that satisfies*

$$X(t) = S(t)X_0 + \int_0^t S(t-s)F(X(s))ds + \int_0^t S(t-s)\Phi(X(s))dW(s), \quad t \geq 0.$$

Moreover, the solution X is also weakly unique.

A detailed proof of Theorem 2.2 can be found in [7].

The main condition on the class of equations that we will study is that the equations have an exponential dichotomy, which is defined by means of the following property of the semigroup $(S(t))_{t\geq 0}$.

Definition 2.3. The semigroup $(S(t))_{t\geq 0}$ is called *hyperbolic* with exponents $\alpha_1, \alpha_2 > 0$ if there exist closed linear subspaces H_1 and H_2 of H and a constant $M \geq 0$ such that $H = H_1 \oplus H_2$, H_1 and H_2 are invariant under $S(t)$ for all $t \geq 0$,

$$\|S(t)x\| \leq Me^{-\alpha_1 t}\|x\| \quad \text{for all } t \geq 0,\ x \in H_1, \tag{4}$$

and $(S(t)|_{H_2})_{t\geq 0}$ extends to a strongly continuous group on H_2 with

$$\|S(-t)x\| \leq Me^{-\alpha_2 t}\|x\| \quad \text{for all } t \geq 0,\ x \in H_2. \tag{5}$$

We will denote by P_1 the projection on H_1 along H_2 and by P_2 the projection on H_2 along H_1.

Hyperbolicity of the semigroup $(S(t))_{t\geq 0}$ can be formulated in terms of its generator A. Gearhart [8] and Herbst [10] showed that $(S(t))_{t\geq 0}$ is hyperbolic if and only if there is an open strip U in the complex plane containing the imaginary axis such that A has no spectrum in U and its resolvent is uniformly bounded on U (see also [13]). If the semigroup is eventually compact, that is, there is a $t_0 > 0$ such that $S(t)$ is compact for all $t \geq t_0$, then it is known that the spectrum of A consists only of isolated points (see [5, Exercise IV.2.8] or [18, Corollary A-III.6.7]). So in this case $(S(t))_{t\geq 0}$ is hyperbolic if and only if A has no spectrum on the imaginary axis. Moreover, H_2 is then finite dimensional.

We say that equation (2) has an *exponential dichotomy* if one can choose $(S(t))_{t\geq 0}$ and F in such a way that $(S(t))_{t\geq 0}$ is hyperbolic and the Lipschitz constant of F is smaller than the exponents of hyperbolicity.

Hyperbolicity and dichotomies are important and well-studied notions in the theory of evolution equations, especially in relation with long term behavior and stability issues (see [1, 9, 22]).

3 Main result and outline of the proof

In our study of long term behavior of equation (2) we will focus on three properties, namely tightness of the set of distributions $\{\mu_{X(t)} : t \geq 0\}$ of the solution $t \mapsto X(t)$, convergence of $\mu_{X(t)}$ to a limit distribution, and existence of an invariant measure. We will also encounter more quantitative properties such as in the conclusion of Theorem 5.1. The notion of tightness is discussed in Section 6. The set of distributions $\{\mu_{X(t)} : t \geq 0\}$ is *tight* if and only if for every $\varepsilon > 0$ there is a compact set $K_\varepsilon \subset H$ with $\mu_{X(t)}(K_\varepsilon) \geq 1-\varepsilon$, or, in other words, $\mathbb{P}(X(t) \in K_\varepsilon) \geq 1 - \varepsilon$ for all $t \geq 0$. An *invariant measure* for equation (2) is a Borel probability measure on the Hilbert space H which is invariant under the action of the dual of the transition semigroup induced by (2). A complete definition is given in Section 8. The main result in this paper is as follows.

Theorem 3.1. *Assume Situation (*) and assume that the semigroup* $(S(t))_{t\geq 0}$ *is hyperbolic with* H_1, H_2*,* α_1, α_2, M*,* P_1*, and* P_2 *as in Definition 2.3. Let*

$$L_1 = 3\left(\frac{L_F^2}{\alpha_1} + L_\Phi^2\right) M^2 \|P_1\|^2, \quad L_2 = 3\left(\frac{L_F^2}{\alpha_2} + L_\Phi^2\right) M^2 \|P_2\|^2,$$

and assume that the Lipschitz constants L_F *and* L_Φ *are so small that* $2L_1 < \alpha_1$*,* $2L_2 < \alpha_2$*, and* $2\alpha_1 L_1 + 2\alpha_2 L_1 < \alpha_1\alpha_2$*. Let* $X : [0,\infty) \to L^2(\Omega; H)$ *be the mild solution of (2) and let* $\mu_{X(t)}$ *denote the distribution of* $X(t)$*,* $t \geq 0$*. If* $t \mapsto \mathbb{E}\,\|X(t)\|^2$ *is bounded on* $[0,\infty)$*, then*

(i) the set of distributions $\{\mu_{X(t)} : t \geq 0\}$ *is tight,*

(ii) there exists a Borel probability measure μ *on* H *with* $\int_H \|x\|^2 d\mu(x) < \infty$ *such that*

$$\mu_{X(t)} \to \mu \quad \textit{as } t \to \infty$$

in the sense of weak convergence of measures, and

(iii) the measure μ *from (ii) is an invariant measure for equation (2).*

Let us explain the idea of the proof. As a first approach to studying long term limit behavior of the solution X of (2) one could consider

$$\mathbb{E}\,\|X(t+\tau) - X(t)\|^2$$

for some $\tau > 0$ and let $t \to \infty$. However, $\mathbb{E}\,\|X(t+\tau) - X(t)\|^2$ will not tend to zero due to the time shift in the Wiener processes occurring in the right hand side of the equation (2) for $t+\tau$ and t, and the independence of its increments. Clearly, it is too much to expect that the random variables $X(t)$ converge themselves as $t \to \infty$, as the Wiener process keeps adding new stochastically independent information.

The crucial idea of our method is to involve another continuous function $Y : [0,\infty) \to L^2(\Omega; H)$ such that $Y(t)$ has the same distribution as $X(t)$ for every $t \geq 0$ and such that

$$\mathbb{E}\,\|X(t+\tau) - Y(t)\|^2 \to 0 \quad \text{as } t \to \infty$$

for some $\tau > 0$. We consider existence of such a function Y as a 'Cauchy condition' for the function $t \mapsto X(t)$. The first step of the proof of Theorem 3.1 is then to show that such a function Y exists, the second step that it follows that $t \mapsto X(t)$ has a tight family of distributions, and the third step to conclude existence of a limit distribution and an invariant measure.

In order to find such a function Y we will exploit the weak uniqueness of (2). Indeed, we will let $\tau > 0$ be arbitrary and choose Y to be the solution of the equation with the Wiener process shifted over a time $\tau > 0$:

$$Y(t) = S(t)X_0 + \int_0^t S(t-s)F(Y(s))ds + \int_0^t S(t-s)\Phi(Y(s))dW(\tau+s), \quad (6)$$

$t \geq 0$. Then $Y(t)$ and $X(t)$ have the same distribution for all $t \geq 0$ and the dichotomy yields that $\mathbb{E}\,\|X(t+\tau) - Y(t)\|^2 \to 0$ as $t \to \infty$. Our main tools here are the Itô isometry and Gronwall type lemmas.

For the second step it turns out that it does not matter that X is the solution of a stochastic differential equation; it suffices that X is a continuous function from $[0,\infty)$ to $L^2(\Omega; H)$. We do need to know, however, that

$$t \mapsto \mathbb{E}\,\|X(t+\tau) - Y(t)\|^2$$

decays with at least a certain speed. To be precise, we need that

$$\sum_{n=1}^{\infty} \left(\max_{(n-1)\tau \leq t \leq n\tau} \mathbb{E}\,\|X(t+\tau) - Y(t)\|^2 \right)^{1/3} < \infty. \quad (7)$$

For the function Y of the first step the decay is exponentially, which is amply sufficient. Our arguments in the second step involve standard elements of the theory of probability measures on metric spaces, without reference to differential equations.

To complete the proof of Theorem 3.1, we will show that the tightness of the set $\{\mu_{X(t)} : t \geq 0\}$ provides existence of a limit distribution and existence of an invariant measure. One could do so by applying standard results of stochastic differential equations, but we will show that existence of a limit distribution follows as well in the more general setting from (7) if $\mathbb{E}\,\|X(t+\tau) - Y(t)\|^2 \to 0$ as $t \to \infty$ uniformly in $\tau \geq 0$, where $Y = Y_\tau$ may depend on τ. In fact, our results in the second and third step amount to the following theorem.

Theorem 3.2. *Let H be a separable real Hilbert space and let $(\Omega; \mathcal{F}, \mathbb{P})$ be a probability space. Let $X : [0,\infty) \to L^2(\Omega; H)$ be a bounded continuous function and let $\mu_{X(t)}$ denote the distribution of $X(t)$ for $t \geq 0$. Assume that we have for every $\tau \geq 0$ a bounded continuous function $Y_\tau : [0,\infty) \to L^2(\Omega; H)$ such that $Y_\tau(t)$ has the same distribution as $X(t)$ for every $t \geq 0$.*

(i) If for some $\tau > 0$ *we have that $\mathbb{E}\,\|X(t+\tau) - Y_\tau(t)\|^2$ tends to zero so fast that*

$$\sum_{n=1}^{\infty} \left(\max_{(n-1)\tau \leq t \leq n\tau} \mathbb{E}\,\|X(t+\tau) - Y_\tau(t)\|^2 \right)^{1/3} < \infty,$$

then $\{\mu_{X(t)} : t \geq 0\}$ is tight.

(ii) *If* $\{\mu_{X(t)} : t \geq 0\}$ *is tight, then for every sequence* $(t_n)_n$ *in* $[0,\infty)$ *with* $t_n \to \infty$ *there is a subsequence* $(t_{n_k})_k$ *and a Borel probability measure* μ *on* H *with* $\int_H \|x\|^2 d\mu(x) < \infty$ *such that* $\mu_{X(t_{n_k})} \to \mu$ *in the sense of weak convergence of measures.*

(iii) *If* $\mathbb{E}\,\|X(t+\tau) - Y_\tau(t)\|^2 \to 0$ *as* $t \to \infty$ *uniformly in* $\tau \geq 0$*, then there is* at most one *Borel probability measure* μ *on* H *such that* $\mu_{X(t_n)} \to \mu$ *as* $t_n \to \infty$*.*

Corollary 3.3. *In the setting of Theorem 3.2 we have that if there exist* $C, \alpha > 0$ *such that*

$$\mathbb{E}\,\|X(t+\tau) - Y_\tau(t)\|^2 \leq C(t+1)^{-(3+\alpha)} \quad \text{for all } t, \tau \geq 0,$$

then $\{\mu_{X(t)} : t \geq 0\}$ *is tight and* $\mu_{X(t)} \to \mu$ *as* $t \to \infty$ *for some Borel probability measure* μ *on* H *with* $\int_H \|x\|^2 d\mu(x) < \infty$*.*

Finally, we will return to our equation and conclude that the obtained limit distribution is an invariant measure.

A special class of hyperbolic semigroups are the semigroups that are exponentially stable, that is, that satisfy the conditions of Definition 2.3 with $H_2 = \{0\}$. If the semigroup $(S(t))_{t\geq 0}$ in Theorem 3.1 is exponentially stable, then it turns out that for any initial condition X_0 the solution X is bounded, and existence of an invariant measure follows. We will thus obtain the following result.

Theorem 3.4. *Assume Situation (*) and assume that* $(S(t))_{t\geq 0}$ *is exponentially stable, that is,*

$$\|S(t)x\| \leq Me^{-\alpha_1 t}\|x\| \quad \text{for all } t \geq 0,\ x \in H,$$

for certain constants $M, \alpha_1 > 0$*. Assume also that* $2L_1 < \alpha_1$*, where*

$$L_1 = 3\left(\frac{L_F^2}{\alpha_1} + L_\Phi^2\right)M^2.$$

Then there exists an invariant measure for equation (2). More than that, for every $\mathcal{F}_0$*-measurable initial condition* $X_0 \in L^2(\Omega; H)$ *the solution* $t \mapsto X(t)$ *of (2) is bounded, and we have the conclusions (i)–(iii) of Theorem 3.1.*

4 An extension of Gronwall's lemma

In the first step of the proof of Theorem 3.1 we need Gronwall's lemma and an extension of it. The next Gronwall lemma is standard (cf. [11, Theorem 1.5.7, p.19]).

Lemma 4.1 (Gronwall). *Let $T > 0$ and let $x, y \in C[0,T]$, $x, y \geq 0$. Let $C, K \in [0,\infty)$. If*

$$x(t) \leq C + K\int_0^t (x(s) + y(s))ds \quad \text{for all } t \in [0,T],$$

then

$$x(t) \leq Ce^{Kt} + K\int_0^t e^{K(t-s)}y(s)ds \quad \text{for all } t \in [0,T].$$

Our extension of Gronwall's lemma is based on the following observation.

Lemma 4.2. *Let $(E, \|.\|)$ be a partially ordered Banach space with $E^+ := \{x \in E : x \geq 0\}$ closed. Let $T : E \to E$ be a linear strict contraction that is positive (i.e. $T(E^+) \subset E^+$). Then $I - T$ is a bijection and $(I - T)^{-1}$ is positive.*

Proof. It is well-known that $I-T$ is a bijection (and that $(I-T)^{-1}$ is linear and bounded). Let $y \in E^+$. Define $G(x) := y + Tx$, $x \in E^+$. Then $G(E^+) \subset E^+$ and

$$\|G(x_1) - G(x_2)\| = \|Tx_1 - Tx_2\| \leq \|T\| \, \|x_1 - x_2\|, \quad x_1, x_2 \in E^+,$$

so G is a strict contraction on E^+. By Banach's fixed point theorem, there is an $x \in E^+$ with $G(x) = x$. This means that $y + Tx = x$, or, in other words, $y = (I - T)x$, so that $(I - T)^{-1}y = x \geq 0$. Hence $(I - T)^{-1}$ is positive. □

We are now ready to extend Gronwall's lemma to inequalities including an integral running from t to ∞.

Lemma 4.3. *Let $x \in C[0,\infty)$ be bounded and such that $x \geq 0$ and*

$$x(t) \leq Ae^{-at} + B\int_0^t e^{-b(t-s)}x(s)ds + C\int_t^\infty e^{-c(s-t)}x(s)ds \text{ for all } t \geq 0,$$

where $A, B, C, a, b, c \in [0,\infty)$ are such that $C < c$ and $\frac{Bc}{c-C} < b$ (equivalently, $\frac{B}{b} + \frac{C}{c} < 1$). If $C > 0$, then for every $\gamma \geq 0$ with $\gamma \leq b - \frac{Bc}{c-C}$ and $\gamma \leq a$ we have

$$x(t) \leq M_\gamma e^{-\gamma t} \quad \text{for all } t \geq 0,$$

where

$$M_\gamma = \frac{A}{1 - \left(\frac{B}{b-\gamma} + \frac{C}{c+\gamma}\right)} \text{ if } \gamma < b \quad \text{and} \quad M_\gamma = \frac{A}{1 - \frac{C}{c+\gamma}} \text{ if } \gamma = b. \tag{8}$$

If $C = 0$, then the estimate holds true for every $\gamma \geq 0$ with $\gamma < b - B$ and $\gamma \leq a$.

Proof. Define

$$h(t,s) := \begin{cases} Be^{-b(t-s)}, & s \le t, \\ Ce^{-c(s-t)}, & s > t, \end{cases}$$

and

$$(Tx)(t) := \int_0^\infty h(t,s)x(s)ds, \quad t \ge 0,$$

for every bounded and continuous function $x : [0,\infty) \to \mathbb{R}$. For $\gamma \in [0,\infty)$, let

$$E_\gamma := \{x \in C[0,\infty) : t \mapsto e^{\gamma t}x(t) \text{ is bounded}\},$$
$$\|x\|_\gamma := \sup_{t\ge 0} |e^{\gamma t}x(t)|, \quad x \in E_\gamma.$$

The space $(E_\gamma, \|.\|_\gamma)$ endowed with the natural ordering defined by

$$x \le y \quad \Longleftrightarrow \quad x(t) \le y(t) \text{ for all } t \ge 0$$

is a partially ordered Banach space and the set of nonnegative functions is closed. For every $\gamma > 0$, E_γ may be viewed as a linear subspace of E_0. We complete the proof in three steps. Assume first that $C > 0$.

(a) We have for any $\gamma \in [0, b - \frac{Bc}{c-C}]$ and any $x \in E_\gamma$ that

$$\begin{aligned} \sup_{t\ge 0} |e^{\gamma t}(Tx)(t)| &= \sup_{t\ge 0} \left| B\int_0^t e^{-b(t-s)}e^{\gamma t}x(s)ds + C\int_t^\infty e^{-c(s-t)}e^{\gamma t}x(s)ds \right| \\ &\le \sup_{t\ge 0} \left\{ B\int_0^t e^{-(b-\gamma)(t-s)}ds\|x\|_\gamma + C\int_t^\infty e^{-(c+\gamma)(s-t)}ds\|x\|_\gamma \right\} \\ &\le \left\{ \frac{B}{b-\gamma} + \frac{C}{c+\gamma} \right\} \|x\|_\gamma \quad \text{if } \gamma < b \end{aligned} \tag{9}$$

and if $\gamma = b$, then $B = 0$ and

$$\sup_{t\ge 0} |e^{\gamma t}(Tx)(t)| \le \frac{C}{c+\gamma}\|x\|_\gamma.$$

Hence $T(E_\gamma) \subset E_\gamma$ and $\|T\|_\gamma \le \frac{B}{b} + \frac{C}{c} < 1$ if $\gamma = 0$ and $\|T\|_\gamma \le \frac{c-C}{c} + \frac{C}{c+\gamma} < 1$ if $\gamma > 0$. By the previous lemma it follows that $(I-T)$ is a bijection from E_γ to E_γ and that $(I-T)^{-1}$ is positive on E_γ.

(b) We are given that $0 \le x \le y + Tx$, where $y(t) = Ae^{-at}$, $t \ge 0$. Let $\gamma \ge 0$ with $\gamma \le b - \frac{Bc}{c-C}$ and $\gamma \le a$. We have that $y \in E_\gamma \subset E_0$ and $x \in E_0$, hence $(I-T)x \le y$ holds in E_0. By using (a) for $\gamma = 0$ we know that $(I-T)^{-1}$ is positive on E_0, so that $0 \le x \le (I-T)^{-1}y$ in E_0. As $y \in E_\gamma$ we have $(I-T)^{-1}y \in E_\gamma$ and it follows that $x \in E_\gamma$. More quantitatively,

$$\|x\|_\gamma \le \|(I-T)^{-1}y\|_\gamma \le \|(I-T)^{-1}\|_\gamma\|y\|_\gamma \le \frac{1}{1-\|T\|_\gamma}A$$

and

$$\frac{1}{1-\|T\|_\gamma} \le \frac{1}{1-\left(\frac{B}{b-\gamma}+\frac{C}{c+\gamma}\right)} \text{ if } \gamma < b \quad \text{and} \quad \frac{1}{1-\|T\|_\gamma} \le \frac{1}{1-\frac{C}{c+\gamma}} \text{ if } \gamma = b.$$

(c) If $C = 0$, we see from (9) that $\|T\|_\gamma \le \frac{B}{b-\gamma} < 1$ for every $\gamma \ge 0$ with $\gamma < b - B$ and the proof can be completed along the same lines as above. □

Remark 4.4. If $C > 0$ in the above lemma, then we have

$$M_\gamma \le \frac{Ac(c+\gamma)}{C\gamma}, \quad \text{for any} \quad 0 < \gamma \le b - \frac{Bc}{c-C}.$$

Indeed, if $B = 0$ then the inequality follows from $C < c$, and if $B > 0$ then

$$1 - \left(\frac{B}{b-\gamma} + \frac{C}{c+\gamma}\right) \ge 1 - \frac{B}{Bc/(c-C)} - \frac{C}{c+\gamma} = \frac{C\gamma}{c(c+\gamma)}.$$

5 Exploiting the exponential dichotomy

We will establish in this section that if all the assumptions of Theorem 3.1 are satisfied, then the solution $t \mapsto X(t)$ has the property that for any $\tau > 0$ there exists a bounded continuous function $Y_\tau : [0,\infty) \to L^2(\Omega; H)$ such that $Y_\tau(t)$ has the same distribution as $X(t)$ for every $t \ge 0$ and $\mathbb{E}\,\|X(t+\tau) - Y(t)\|^2 \to 0$ exponentially fast and uniformly in $\tau \ge 0$ as $t \to \infty$. As we pointed out in Section 3, we find such a function Y_τ as the solution of the equation (6) with shifted Wiener processes.

Theorem 5.1. *Assume Situation (*). Assume that the semigroup $(S(t))_{t\ge 0}$ is hyperbolic with H_1, H_2, α_1, α_2, M, P_1, and P_2 as in Definition 2.3 and assume that*

$$2L_1 < \alpha_1, \quad 2L_2 < \alpha_2, \quad \text{and} \quad 2\alpha_1 L_2 + 2\alpha_2 L_1 < \alpha_1\alpha_2,$$

where

$$L_1 = 3\left(\frac{L_F^2}{\alpha_1} + L_\Phi^2\right) M^2 \|P_1\|^2, \quad L_2 = 3\left(\frac{L_F^2}{\alpha_2} + L_\Phi^2\right) M^2 \|P_2\|^2.$$

Let $X : [0,\infty) \to L^2(\Omega; H)$ be the continuous adapted solution of

$$X(t) = S(t)X_0 + \int_0^t S(t-s)F(X(s))ds + \int_0^t S(t-s)\Phi(X(s))dW(s),$$

$t \ge 0$, and $Y : [0,\infty) \to L^2(\Omega; H)$ the continuous adapted solution of

$$Y(t) = S(t)X_0 + \int_0^t S(t-s)F(Y(s))ds + \int_0^t S(t-s)\Phi(Y(s))dW(\tau+s),$$

$t \geq 0$, *for some fixed* $\tau > 0$. *If* $t \mapsto \mathbb{E}\,\|X(t)\|^2$ *is bounded on* $[0, \infty)$, *then*

$$\mathbb{E}\,\|X(t+\tau) - Y(t)\|^2 \leq De^{-\gamma t} \text{ for all } t \geq 0,$$

where

$$D = 24M^2 \left\{ 1 + \frac{(\alpha_1 - \gamma)\alpha_2(\alpha_2 + \gamma)}{(\alpha_1 - 2L_1 - \gamma)(\alpha_1 + \alpha_2)\gamma} \right\} \|P_1\|^2 \left(\sup_{s \geq 0} \mathbb{E}\,\|X(s)\|^2 \right).$$

Here γ *may be any number in* $(0, \gamma_0]$ *with*

$$\begin{aligned} \gamma_0 &= \frac{(\alpha_1 - 2L_1 + \alpha_2)(\alpha_1\alpha_2 - 2\alpha_1 L_2 - 2\alpha_2 L_1)}{\alpha_1\alpha_2 - 2\alpha_1 L_2 - 2\alpha_2 L_1 + \alpha_2(\alpha_2 - 2L_2)} \\ &= \alpha_1 - 2L_1 - \frac{4L_1 L_2 \alpha_2}{\alpha_1\alpha_2 - 2\alpha_1 L_2 - 2\alpha_2 L_1 + \alpha_2(\alpha_2 - 2L_2)}. \end{aligned}$$

Proof. Define for $t \geq 0$

$$z_1(t) := \mathbb{E}\,\|P_1(X(t+\tau) - Y(t))\|^2 \text{ and } z_2(t) := \mathbb{E}\,\|P_2(X(t+\tau) - Y(t))\|^2. \quad (10)$$

We will first derive that

$$\begin{aligned} z_1(t) &\leq 3M^2 e^{-2\alpha_1 t} z_1(0) + 2L_1 \int_0^t e^{-\alpha_1(t-s)}(z_1(s) + z_2(z))ds, && (11) \\ z_2(t) &\leq 2L_2 \int_t^\infty e^{-\alpha_2(s-t)}(z_1(s) + z_2(s))ds, \quad t \geq 0, && (12) \end{aligned}$$

and then apply the previous lemmas to obtain the conclusion of the theorem.

For $t \geq 0$ we have, with aid of Lemma 2.1,

$$\begin{aligned} X(t+\tau) &= S(t+\tau)X_0 + \int_0^{t+\tau} S(t+\tau-s)F(X(s))ds \\ &\quad + \int_0^{t+\tau} S(t+\tau-s)\Phi(X(s))dW(s) \\ &= S(t)S(\tau)X_0 + S(t)\int_0^\tau S(\tau-s)F(X(s))ds \\ &\quad + S(t)\int_0^\tau S(\tau-s)\Phi(X(s))dW(s) + \int_\tau^{t+\tau} S(t+\tau-s)F(X(s))ds \\ &\quad + \int_\tau^{t+\tau} S(t+\tau-s)\Phi(X(s))dW(s) \\ &= S(t)X(\tau) + \int_0^t S(t-s)F(X(\tau+s))ds && (13) \\ &\quad + \int_0^t S(t-s)\Phi(X(\tau+s))dW(\tau+s), \end{aligned}$$

so that for all $t \geq 0$

$$\begin{aligned} X(t+\tau) - Y(t) &= S(t)(X(\tau) - X_0) \\ &\quad + \int_0^t S(t-s)\Big\{F(X(\tau+s)) - F(Y(s))\Big\}ds \\ &\quad + \int_0^t S(t-s)\Big\{\Phi(X(\tau+s)) - \Phi(Y(s))\Big\}dW(\tau+s). \end{aligned}$$

Because P_1 commutes with $S(s)$ for all s and because of Lemma 2.1, it follows for all $t \geq 0$ that

$$\begin{aligned} P_1(X(t+\tau) - Y(t)) &= S(t)P_1(X(\tau) - X_0) \\ &\quad + \int_0^t S(t-s)P_1\Big\{F(X(\tau+s)) - F(Y(s))\Big\}ds \\ &\quad + \int_0^t S(t-s)P_1\Big\{\Phi(X(\tau+s)) - \Phi(Y(s))\Big\}dW(\tau+s). \end{aligned} \tag{14}$$

By taking norms at both sides, using the triangle inequality, taking squares, using that $(a+b+c)^2 \leq 3a^2 + 3b^2 + 3c^2$ for any real numbers a, b, c, and taking expectations, we find for all $t \geq 0$

$$\begin{aligned} \mathbb{E}\,\|P_1(X(t+\tau) - Y(t))\|^2 &\leq 3\mathbb{E}\,\|S(t)P_1(X(\tau) - X_0)\|^2 \\ &\quad +3\mathbb{E}\,\|\int_0^t S(t-s)P_1\Big\{F(X(\tau+s)) - F(Y(s))\Big\}ds\|^2 \\ &\quad +3\mathbb{E}\,\|\int_0^t S(t-s)P_1\Big\{\Phi(X(\tau+s)) - \Phi(Y(s))\Big\}dW(\tau+s)\|^2. \end{aligned} \tag{15}$$

With aid of the Itô isometry, the latter term becomes

$$\begin{aligned} &3\mathbb{E}\,\|\int_0^t S(t-s)P_1\Big\{\Phi(X(\tau+s)) - \Phi(Y(s))\Big\}ds\|^2 \\ &\leq 3\int_0^t \|S(t-s)P_1\|^2 \mathbb{E} \sum_{i=1}^{\infty} \|\Phi_i(X(\tau+s)) - \Phi_i(Y(s))\|^2 ds. \end{aligned}$$

To deal with the second term at the right hand side of (15), note that

$$\begin{aligned} &\|\int_0^t S(t-s)P_1\Big\{F(X(\tau+s)) - F(Y(s))\Big\}ds\|^2 \\ &\leq \left(\int_0^t \|S(t-s)P_1\|^{1/2}\|S(t-s)P_1\|^{1/2}\|F(X(\tau+s)) - F(Y(s))\|ds\right)^2 \\ &\leq \int_0^t \|S(s)P_1\|ds \int_0^t \|S(t-s)P_1\|\,\|F(X(\tau+s)) - F(Y(s))\|^2 ds. \end{aligned}$$

It follows that for all $t \geq 0$ we have

$$\begin{aligned} &\mathbb{E}\,\|P_1(X(t+\tau) - Y(t))\|^2 \leq 3\mathbb{E}\,\|S(t)P_1(X(\tau) - X_0)\|^2 \\ &\quad +3\int_0^t \|S(s)P_1\|ds \int_0^t \|S(t-s)P_1\|L_F^2\mathbb{E}\,\|X(\tau+s) - Y(s)\|^2 ds \\ &\quad +3\int_0^t \|S(t-s)P_1\|^2 L_\Phi^2 \mathbb{E}\,\|X(\tau+s) - Y(s)\|^2 ds. \end{aligned}$$

Since $\|S(s)P_1x\| \leq Me^{-\alpha_1 s}\|P_1x\|$ for all $x \in H$ and $s \geq 0$, a direct computation leads to

$$\begin{aligned}\mathbb{E}\,\|P_1(X(t+\tau)-Y(t))\|^2 &\leq 3M^2e^{-2\alpha_1 t}\mathbb{E}\,\|P_1(X(\tau)-X_0)\|^2\\ &\quad+\frac{3}{\alpha_1}M^2\|P_1\|^2L_F^2\int_0^t e^{-\alpha_1(t-s)}\mathbb{E}\,\|X(\tau+s)-Y(s)\|^2ds\\ &\quad+3M^2\|P_1\|^2L_\Phi^2\int_0^t e^{-\alpha_1(t-s)}\mathbb{E}\,\|X(\tau+s)-Y(s)\|^2ds,\end{aligned}$$

$t \geq 0$. With L_1 as defined in the theorem, we obtain that for all $t \geq 0$

$$\begin{aligned}\mathbb{E}\,\|P_1(X(t+\tau)-Y(t))\|^2 &\leq 3M^2e^{-2\alpha_1 t}\mathbb{E}\,\|P_1(X(\tau)-X_0)\|^2\\ &\quad+L_1\int_0^t e^{-\alpha_1(t-s)}\mathbb{E}\,\|X(\tau+s)-Y(s)\|^2ds.\end{aligned}\tag{16}$$

With z_1 and z_2 as defined in (10), it is clear that

$$\begin{aligned}\mathbb{E}\,\|X(s+\tau)-Y(s)\|^2 &= \mathbb{E}\,\|(P_1+P_2)(X(s+\tau)-Y(s))\|^2\\ &\leq 2z_1(s)+2z_2(s), \quad s \geq 0,\end{aligned}$$

so that we find for all $t \geq 0$ that

$$z_1(t) \leq 3M^2e^{-2\alpha_1 t}z_1(0)+2L_1\int_0^t e^{-\alpha_1(t-s)}(z_1(s)+z_2(s))ds,$$

which is the announced estimate (11).

To find a suitable estimate for z_2, we will consider the equation for X in backward time. Thus, let $t_0 \geq 0$ and observe that similar to (13) we have for all $t \in [0,t_0]$

$$\begin{aligned}X(t_0) &= S(t)X(t_0-t)+\int_0^t S(s)F(X(t_0-s))ds\\ &\quad+\int_0^t S(s)\Phi(X(t_0-s))dW(t_0-s).\end{aligned}$$

Applying $S(-t)P_2$ and rearranging terms yields

$$\begin{aligned}P_2X(t_0-t) &= S(-t)P_2X(t_0)-\int_0^t S(-(t-s))P_2F(X(t_0-s))ds\\ &\quad-\int_0^t S(-(t-s))P_2\Phi(X(t_0-s))dW(t_0-s).\end{aligned}$$

Now substitute t_0-t for t and replace t by $t+\tau$ and t_0 by $t_0+\tau$, then we find for all $t \in [0,t_0]$

$$\begin{aligned}P_2X(t+\tau) &= S(-(t_0-t))P_2X(t_0+\tau)\\ &\quad-\int_0^{t_0-t} S(-(t_0-t-s))P_2F(X(t_0+\tau-s))ds\\ &\quad-\int_0^{t_0-t} S(-(t_0-t-s))P_2\Phi(X(t_0+\tau-s))dW(t_0+\tau-s).\end{aligned}$$

If we perform similar steps for Y as well, then we obtain

$$\begin{aligned}P_2Y(t) &= S(-(t_0-t))P_2Y(t_0)\\ &\quad -\int_0^{t_0-t} S(-(t_0-t-s))P_2F(Y(t_0-s))ds\\ &\quad -\int_0^{t_0-t} S(-(t_0-t-s))P_2\Phi(Y(t_0-s))dW(\tau+t_0-s),\end{aligned}$$

for $t\in[0,t_0]$, and hence

$$\begin{aligned}&P_2(X(t+\tau)-Y(t)) = S(-(t_0-t))P_2(X(t_0+\tau)-Y(t_0)) \qquad (17)\\ &\quad -\int_0^{t_0-t} S(-(t_0-t-s))P_2\Big\{F(X(t_0+\tau-s))-F(Y(t_0-s))\Big\}ds\\ &\quad -\int_0^{t_0-t} S(-(t_0-t-s))P_2\Big\{\Phi(X(t_0+\tau-s))-\Phi(Y(t_0-s))\Big\}\\ &\qquad\qquad dW(t_0+\tau-s).\end{aligned}$$

Along the same lines as (16) was derived from (14), we can derive from (17) that for all $t\in[0,t_0]$

$$\begin{aligned}&\mathbb{E}\,\|P_2(X(t+\tau)-Y(t))\|^2 \le 3M^2e^{-2\alpha_2(t_0-t)}\mathbb{E}\,\|P_2(X(t_0+\tau)-Y(t_0))\|^2\\ &\quad +L_2\int_0^{t_0-t} e^{-\alpha_2(t_0-t-s)}\mathbb{E}\,\|X(t_0+\tau-s)-Y(t_0-s)\|^2ds\\ &\quad = 3M^2e^{-2\alpha_2(t_0-t)}\mathbb{E}\,\|P_2(X(t_0+\tau)-Y(t_0))\|^2 \qquad (18)\\ &\quad +L_2\int_t^{t_0} e^{-\alpha_2(s-t)}\mathbb{E}\,\|X(s+\tau)-Y(s)\|^2ds,\end{aligned}$$

where L_2 is as in the theorem. Because of the weak uniqueness (Theorem 2.2), $Y(t)$ and $X(t)$ have the same distribution and therefore $\mathbb{E}\,\|Y(t)\|^2=\mathbb{E}\,\|X(t)\|^2$ for all t. Consequenly, the assumptions yield that $t\mapsto \mathbb{E}\,\|X(t+\tau)-Y(t)\|^2\le 2\mathbb{E}\,\|X(t+\tau)\|^2+2\mathbb{E}\,\|Y(t)\|^2$ is bounded. Let now in (18) $t\ge 0$ be fixed and $t_0\to\infty$, then

$$\mathbb{E}\,\|P_2(X(t+\tau)-Y(t))\|^2 \le L_2\int_t^{\infty} e^{-\alpha_2(s-t)}\mathbb{E}\,\|X(s+\tau)-Y(s)\|^2ds,$$

so that

$$z_2(t)\le 2L_2\int_t^{\infty} e^{-\alpha_2(s-t)}(z_1(s)+z_2(s))ds,$$

which is (12).

We will now derive explicit estimates from the implicit ones in (11) and (12). Multiply (11) by $e^{\alpha_1 t}$ and apply Lemma 4.1 to find that

$$z_1(t)\le 3M^2z_1(0)e^{-\tilde{\alpha}_1 t}+2L_1\int_0^t e^{-\tilde{\alpha}_1(t-s)}z_2(s)ds,\quad t\ge 0, \qquad (19)$$

where

$$\tilde{\alpha}_1 := \alpha_1 - 2L_1.$$

By assumption, $\tilde{\alpha}_1 > 0$. Inserting (19) in (12) yields that

$$\begin{aligned} z_2(t) &\leq 2L_2 \int_t^\infty e^{-\alpha_2(s-t)} z_2(s) ds + 6L_2 M^2 z_1(0) \int_t^\infty e^{-\alpha_2(s-t)} e^{-\tilde{\alpha}_1 s} ds \\ &\quad + 4L_1 L_2 \int_t^\infty \int_0^s e^{-\alpha_2(s-t)} e^{-\tilde{\alpha}_1(s-t)} z_2(\sigma) d\sigma ds \\ &= \frac{6L_2 M^2}{\tilde{\alpha}_1 + \alpha_2} z_1(0) e^{-\tilde{\alpha}_1 t} + \frac{4L_1 L_2}{\tilde{\alpha}_1 + \alpha_2} \int_0^t e^{-\tilde{\alpha}_1(t-s)} z_2(s) ds \\ &\quad + \left(2L_2 + \frac{4L_1 L_2}{\tilde{\alpha}_1 + \alpha_2}\right) \int_t^\infty e^{-\alpha_2(s-t)} z_2(s) ds, \end{aligned}$$

for all $t \geq 0$. We now want to apply Lemma 4.3 with

$$A = \frac{6L_2^2 M^2}{\tilde{\alpha}_1 + \alpha_2} z_1(0), \quad B = \frac{4L_1 L_2}{\tilde{\alpha}_1 + \alpha_2}, \quad C = 2L_2 + \frac{4L_1 L_2}{\tilde{\alpha}_1 + \alpha_2} = \frac{2L_2(\alpha_1 + \alpha_2)}{\alpha_1 - 2L_1 + \alpha},$$

$a = b = \tilde{\alpha}_1$, and $c = \alpha_2$. As the assumption $2\alpha_1 L_2 + 2\alpha_2 L_1 < \alpha_1 \alpha_2$ implies that $4L_1 L_2 < \tilde{\alpha}_1(\alpha_2 - 2L_2)$, and $2L_1 < \alpha_1$ and $2L_2 < \alpha_2$, we find that

$$C < 2L_2 + \frac{\tilde{\alpha}_1}{\tilde{\alpha}_1 + \alpha_2}(\alpha_2 - 2L_2) < c$$

and

$$\frac{Bc}{c - C} = \frac{\tilde{\alpha}_1 \frac{4L_1 L_2}{\tilde{\alpha}_1 + \alpha_2} \alpha_2}{\tilde{\alpha}_1(\alpha_2 - 2L_2) - \tilde{\alpha}_1 \frac{4L_1 L_2}{\tilde{\alpha}_1 + \alpha_2}} < \frac{\frac{\tilde{\alpha}_1 \alpha_2}{\tilde{\alpha}_1 + \alpha_2}}{1 - \frac{\tilde{\alpha}_1}{\tilde{\alpha}_1 + \alpha_2}} = \tilde{\alpha}_1,$$

so that the conditions of Lemma 4.3 are satisfied. Application leads to

$$z_2(t) \leq M_\gamma e^{-\gamma t}, \quad t \geq 0, \tag{20}$$

for any number $\gamma \in [0, \gamma_0]$ with

$$\gamma_0 = \alpha_1 - 2L_1 - \frac{4L_1 L_2 \alpha_2}{\alpha_1 \alpha_2 - 2\alpha_1 L_2 - 2\alpha_2 L_1 + \alpha_2(\alpha_2 - 2L_2)}$$

and M_γ given by (8). With Remark 4.4 we find for $\gamma \in (0, \gamma_0]$ that

$$M_\gamma \leq \frac{\alpha_2(\alpha_2 + \gamma)}{(\alpha_1 + \alpha_2)\gamma} 3M^2 z_1(0) \tag{21}$$

(note that $M_\gamma = 0$ if $C = 0$). Inserting (20) in (19) yields

$$\begin{aligned} z_1(t) &\leq 3M^2 z_1(0) e^{-\tilde{\alpha}_1 t} + \frac{2L_1 M_\gamma}{\tilde{\alpha}_1 - \gamma}(e^{-\gamma t} - e^{-\tilde{\alpha}_1 t}) \\ &\leq \left(3M^2 z_1(0) + \frac{2L_1 M_\gamma}{\tilde{\alpha}_1 - \gamma}\right) e^{-\gamma t}, \quad t \geq 0. \end{aligned} \tag{22}$$

As $\mathbb{E}\|X(t+\tau)-Y(t)\|^2 \le 2z_1(t)+2z_2(t)$ for all $t\ge 0$, (20), (22), and (21) together yield for any $\gamma\in(0,\gamma_0]$ that

$$\begin{aligned}\mathbb{E}\|X(t+\tau)-Y(t)\|^2 &\le 2\left\{3M^2 z_1(0)+\frac{2L_1M_\gamma}{\tilde{\alpha}_1-\gamma}+M_\gamma\right\}e^{-\gamma t}\\ &\le 6\left\{1+\frac{(\alpha_1-\gamma)\alpha_2(\alpha_2+\gamma)}{(\alpha_1-2L_1-\gamma)(\alpha_1+\alpha_2)\gamma}\right\}M^2\mathbb{E}\|P_1(X(\tau)-Y(0))\|^2 e^{-\gamma t}\end{aligned}$$

for all $t\ge 0$, and we obtain the inequality that had to be proved. □

With some extra attention we may conclude more from the above proof in case that the unstable subspace H_2 is trivial.

Corollary 5.2. *Assume the notations of Theorem 5.1. If the semigroup $(S(t))_{t\ge 0}$ is exponentially stable, that is, $P_2=0$, then the condition in Theorem 5.1 that $t\mapsto \mathbb{E}\|X(t)\|^2$ be bounded on $[0,\infty)$ follows from the other assumptions.*

Proof. Observe that the boundedness condition is not needed for the proof of the estimates (11) and (19). If $P_2=0$, then $z_2=0$, and we obtain from (19) that

$$\mathbb{E}\|X(t+\tau)-Y(t)\|^2 = z_1(t)\le 3M^2 z_1(0)e^{-(\alpha_1-2L_1)t},\quad t\ge 0.$$

In this estimate only $Y=Y_\tau$ and $z_1(0)$ depend on τ. Since

$$z_1(0)\le 2\mathbb{E}\|X(\tau)\|^2+2\mathbb{E}\|Y(0)\|^2 = 2\mathbb{E}\|X(\tau)\|^2+2\mathbb{E}\|X(0)\|^2$$

and $t\mapsto X(t)$ is continuous, we can define

$$c_n = 3M^2\sup_{0\le\tau\le 1}4\mathbb{E}\|X(\tau)\|^2 e^{-(\alpha_1-2L_1)n},\quad n\in\mathbb{N},$$

and obtain

$$\mathbb{E}\|X(n+\tau)-Y_\tau(n)\|^2\le c_n,\ \text{for all } 0\le\tau\le 1 \text{ and } n\in\mathbb{N}.$$

As $\mathbb{E}\|Y_\tau(t)\|^2=\mathbb{E}\|X(t)\|^2$ for all $t,\tau\ge 0$, the triangle inequality yields that

$$\left|\left(\mathbb{E}\|X(n+\tau)\|^2\right)^{1/2}-\left(\mathbb{E}\|X(n)\|^2\right)^{1/2}\right|\le c_n^{1/2}$$

for all $0\le\tau\le 1$ and $n\in\mathbb{N}$. If we let

$$x_n := \sup_{n\le t\le n+1}\left(\mathbb{E}\|X(t)\|^2\right)^{1/2},$$

then it follows that $x_n\le x_{n-1}+c_n^{1/2}$ for all $n\ge 1$. Because $(c_n^{1/2})_n$ is summable, the sequence $(x_n)_n$ is bounded and therefore $t\mapsto\mathbb{E}\|X(t)\|^2$ is bounded on $[0,\infty)$. □

6 Tight set of distributions

In this section we demonstrate that the estimate

$$\mathbb{E}\,\|X(t+\tau) - Y(t)\|^2 \le De^{-\gamma t}, \quad t \ge 0,$$

in Theorem 5.1 implies that the distributions $\mu_{X(t)}$, $t \ge 0$, of the solution $t \mapsto X(t)$ form a tight set. As announced in Section 3 this result holds true for any bounded continuous function $X : [0,\infty) \to L^2(\Omega; H)$ satisfying the weaker estimate (7) and therefore we study it independently of the stochastic differential equation. We begin with some notions from the theory of probability measures on metric spaces.

Let (M,d) be a separable metric space and let $\mathcal{P}(M)$ denote the set of all Borel probability measures on M. A sequence of measures $(\mu_i)_{i=1}^{\infty}$ in $\mathcal{P}(M)$ is said to *converge weakly* to a $\mu \in \mathcal{P}(M)$ if

$$\int f\, d\mu_i \to \int f\, d\mu \quad \text{as } i \to \infty$$

for every bounded continuous function $f : M \to \mathbb{R}$. The *weak topology* of $\mathcal{P}(M)$ is the topology corresponding to weak convergence in $\mathcal{P}(M)$ and it is known to be metrizable ([20, Theorem 6.2, p.43], [24, Theorem 1.1.2, p.9]).

A subset Γ of $\mathcal{P}(M)$ is called *tight* if for every $\varepsilon > 0$ there exists a compact $K_\varepsilon \subset M$ such that $\mu(K_\varepsilon) \ge 1 - \varepsilon$ for all $\mu \in \Gamma$. The following proposition describes the compact subsets of $\mathcal{P}(M)$ if M is complete (see [20, Theorem 6.7, p.47], [24, Theorems 1.1.3–4, p.9–10]).

Proposition 6.1. *Let M be a complete separable metric space and let Γ be a subset of $\mathcal{P}(M)$. Then $\overline{\Gamma}$ is compact (in the weak topology of $\mathcal{P}(M)$) if and only if Γ is tight.*

We will also need a purely topological result on the pre-compact sets in a complete separable metric space M. Recall that a subset $A \subset M$ is *pre-compact*, that is, its closure $\overline{A}$ is compact, if and only if it is *totally bounded*, which means that for every $\varepsilon > 0$ there is a finite collection of ε-balls that covers A.

For a set $A \subset M$ and an $r \in [0,\infty)$ we denote

$$B(A,r) := \{x \in M : \exists a \in A \text{ with } d(x,a) < r\} = \cup_{a \in A} B_r(a),$$

where $B_r(a)$ denotes the open ball around a with radius r. The following lemma says that the union of a sequence of pre-compact sets is pre-compact if the sets are sufficiently near to one another.

Lemma 6.2. *Let (M, d) be a complete metric space. Let $A_1, A_2, \ldots$ be a sequence of pre-compact subsets of M such that there is a summable sequence $r_1, r_2, \ldots$ in $[0, \infty)$ with $A_{n+1} \subset B(A_n, r_n)$ for all n. Then $\cup_n A_n$ is pre-compact.*

Proof. Let $\varepsilon > 0$. Take $N \in \mathbb{N}$ such that $\sum_{n=N}^{\infty} < \varepsilon/2$. Since $A_1 \cup \cdots \cup A_N$ is pre-compact, it is covered by finitely many $\varepsilon/2$-balls. Inflate these balls to ε-balls. They cover $\cup_n A_n$, because $A_m \subset B(A_N, \sum_{n=N}^{\infty} r_n) \subset B(A_N, \varepsilon/2)$ for all $m \geq N$. Hence $\cup_n A_n$ is pre-compact. □

Consider now a continuous function $X : [0, \infty) \to L^2(\Omega; H)$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and H a separable real Hilbert space. Denote the distribution of $X(t)$ by $\mu_{X(t)}$ for each $t \geq 0$. We will show that the set $\{\mu_{X(t)} : t \geq 0\}$ is tight if the function $t \mapsto X(t)$ satisfies the 'Cauchy condition' of Theorem 3.2(i). Indeed, if there exist a $\tau > 0$ and a continuous function $Y : [0, \infty) \to L^2(\Omega; H)$ such that $Y(t)$ has the same distribution as $X(t)$ for every $t \geq 0$ and if $\mathbb{E}\,\|X(t + \tau) - Y(t)\|^2 \to 0$ sufficiently fast as $t \to \infty$, then $\{\mu_{X(t)} : t \geq 0\}$ is tight. First we establish that the continuity of X yields that $\{\mu_{X(t)} : t \in [0, T]\}$ is tight for every $T \geq 0$.

Lemma 6.3. *Let E be a separable real Banach space and let $a < b$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $X : [a, b] \to L^2(\Omega; H)$ be continuous, and let $\mu_{X(t)}$ denote the distribution of $X(t)$, $t \in [a, b]$. Then the subset $\{\mu_{X(t)} : t \in [a, b]\}$ of $\mathcal{P}(E)$ is compact.*

Proof. Denote $\Gamma := \{\mu_{X(t)} : t \in [a, b]\}$. Let $(\mu_n)_n$ be a sequence in Γ. We have to prove that $(\mu_n)_n$ has a convergent subsequence. There are $(t_n)_n$ in $[a, b]$ with $\mu_n = \mu_{X(t_n)}$, $n \in \mathbb{N}$. Choose a subsequence $(t_{n_k})_k$ such that $t_{n_k} \to t$ for a certain $t \in [a, b]$. Then $X(t_{n_k}) \to X(t)$ in $L^2(\Omega; H)$, so there is a subsubsequence $(t_{n_{k_l}})_l$ and a null set $N \subset \Omega$ such that $X(t_{n_{k_l}}) \to X(t)$ pointwise on $\Omega \setminus N$. Then $f(X(t_{n_{k_l}})) \to f(X(t))$ on $\Omega \setminus N$ for every bounded continuous function $f : E \to \mathbb{R}$. By Lebesgue's dominated convergence theorem, we obtain that $\mathbb{E}\, f(X(t_{n_{k_l}})) \to \mathbb{E}\, f(X(t))$ for every bounded continuous f, which means that $\mu_{X(t_{n_{k_l}})} \to \mu_{X(t)}$ in $\mathcal{P}(E)$. Hence $(\mu_{n_{k_l}})_l$ is convergent in Γ, and Γ is compact. □

The idea of the proof of Theorem 3.2(i) is now as follows. Let $\varepsilon > 0$. As $X : [0, \infty) \to L^2(\Omega; H)$ is continuous, it follows from Lemma 6.3 that we can take for each $n \geq 1$ a compact set $C_n \subset H$ such that $\mathbb{P}(X(t) \in C_n) \geq 1 - \varepsilon$ for all $t \in [(n - 1)\tau, n\tau]$. Of course, $\cup_n C_n$ need not be compact, unless we are able to choose the C_n closer and closer together as n increases and apply Lemma 6.2. It actually turns out that the estimate on the speed of convergence of $\mathbb{E}\,\|X(t + \tau) - Y(t)\|^2$ as $t \to \infty$ allows us to do so.

Proof of Theorem 3.2(i). Denote

$$c_n := \max_{n\tau \le t \le (n+1)\tau} \mathbb{E}\,\|X(t) - Y(t-\tau)\|^2, \quad n = 1, 2, \dots.$$

By assumption we have $\sum_{n=1}^{\infty} c_n^{1/3} < \infty$. Let $\varepsilon > 0$ and let

$$\begin{aligned} r_n &:= \rho c_n^{1/3}, \\ \varepsilon_n &:= c_n^{1/3}/\rho^2, \quad n = 1, 2, \dots, \end{aligned}$$

where we choose $\rho > 0$ so large that $\sum_{n=1}^{\infty} \varepsilon_n < \varepsilon/2$. By Markov's inequality we have for every $n \ge 1$ that

$$\begin{aligned} \mathbb{P}(\|X(t) - Y(t-\tau)\| \ge r_n) &\le \frac{1}{r_n^2} \mathbb{E}\,\|X(t) - Y(t-\tau)\|^2 \\ &\le \varepsilon_n, \qquad t \in [n\tau, (n+1)\tau]. \end{aligned} \tag{23}$$

For each $n \ge 1$, the set $\{\mu_{X(t)} : t \in [(n-1)\tau, n\tau]\}$ is compact according to Lemma 6.3 and therefore tight. Thus, we can take a compact set $C_n \subset H$ such that

$$\mathbb{P}(X(t) \in C_n) \ge 1 - \varepsilon_n \text{ for all } t \in [(n-1)\tau, n\tau]. \tag{24}$$

We now define inductively

$$A_1 := C_1, \quad A_{n+1} := C_{n+1} \cap (A_n + B_{r_n}(0)), \; n = 1, 2, \dots.$$

Clearly, every A_n is pre-compact, and for $n \ge 1$ and $t \in [n\tau, (n+1)\tau]$, we obtain from (23) and (24) that

$$\begin{aligned} \mathbb{P}(X(t) \in A_{n+1}) &\ge 1 - \mathbb{P}(X(t) \notin C_{n+1}) - \mathbb{P}(X(t) \notin A_n + B_{r_n}(0)) \\ &= \mathbb{P}(X(t) \in C_{n+1}) - 1 + \mathbb{P}(X(t) \in A_n + B_{r_n}(0)) \\ &\ge -\varepsilon_{n+1} + \mathbb{P}\big(Y(t-\tau) \in A_n \text{ and } \|X(t) - Y(t-\tau)\| < r_n\big) \\ &= -\varepsilon_{n+1} + \mathbb{P}(Y(t-\tau) \in A_n) \\ &\qquad - \mathbb{P}\big(Y(t-\tau) \in A_n \text{ and } \|X(t) - Y(t-\tau)\| \ge r_n\big) \\ &\ge -\varepsilon_{n+1} + \mathbb{P}(Y(t-\tau) \in A_n) - \varepsilon_n \\ &= \mathbb{P}(X(t-\tau) \in A_n) - \varepsilon_n - \varepsilon_{n+1}. \end{aligned}$$

Hence, for each $t \ge 0$, we have that

$$\mu_{X(t)}(\cup_{n=1}^{\infty} A_n) = \mathbb{P}(X(t) \in \cup_{n=1}^{\infty} A_n) \ge 1 - \varepsilon.$$

Furthermore, since the sequence $r_1, r_2, \dots$ is summable, Lemma 6.2 yields that $\cup_{n=1}^{\infty} A_n$ is pre-compact. Thus, the set $\{\mu_{X(t)} : t \ge 0\}$ is tight. $\square$

Remark 6.4. Theorem 3.2(i) is in spirit the same as Theorems 1.3.1 and 1.3.2 in [24]. In [24], however, $\Omega = C([0,\infty);\mathbb{R}^d)$ and since the proofs involve the the Arzela-Ascoli Theorem, the proofs do not extend to the present setting considered here.

7 Limit distribution

We will now prove (ii) and (iii) of Theorem 3.2. With part (i) of Theorem 5.1 we then obtain that the distributions $\mu_{X(t)}$ of the solution $t \mapsto X(t)$ of our equation (2) converge to a Borel probability measure μ on H. This limit distribution is an invariant measure for (2), as we will discuss in Section 8.

Proof of Theorem 3.2(ii) and (iii). (ii): The set $\{\mu_{X(t)} : t \geq 0\}$ is tight and therefore pre-compact in the separable metric space $\mathcal{P}(H)$ of Borel probability measures on H. Hence for any sequence $(t_n)_n$ in $[0,\infty)$ with $t_n \to \infty$ there is a subsequence $(t_{n_k})_k$ and a $\mu \in \mathcal{P}(H)$ such that $\mu_{X(t_{n_k})} \to \mu$ in the weak topology, i.e.

$$\mathbb{E}\, f(X(t_{n_k})) \to \int f\, d\mu$$

for every bounded continuous function $f : H \to \mathbb{R}$.

If $\mathbb{E}\,\|X(t)\|^2 \leq C$ for all $t \geq 0$, then $\int_H \|x\|^2\, d\mu_{X(t)} = \int_\Omega \|X(t)\|^2\, d\mathbb{P} \leq C$ for all $t \geq 0$. For each $N \geq 0$ the function $x \mapsto \min\{\|x\|^2, N\}$ is bounded and continuous on H and therefore

$$\int_H \min\{\|x\|^2, N\}\, d\mu(x) = \lim_{k\to\infty} \int \min\{\|x\|^2, N\}\, d\mu_{X(t_{n_k})}(x) \leq C.$$

By monotone convergence, $\int_H \|x\|^2\, d\mu(x) \leq C$.

(iii): We claim that because $\mathbb{E}\,\|X(t+\tau) - Y_\tau(t)\|^2 \to 0$ uniformly in $\tau \geq 0$, we have that

$$\mathbb{E}\, f(X(t+\tau)) - \mathbb{E}\, f(X(t)) \to 0 \quad \text{as } t \to \infty \text{ uniformly in } \tau \geq 0,$$

for every bounded *uniformly* continuous function $f : H \to \mathbb{R}$. Indeed, let f be such a function. Then for every $\varepsilon > 0$ there is a $\delta_\varepsilon > 0$ such that $\|x - y\| < \delta_\varepsilon$ implies that $|f(x) - f(y)| < \varepsilon$. For every $t \geq 0$ and $\tau \geq 0$ we have then

$$\begin{aligned}
\mathbb{P}(|f(X(t+\tau)) - f(Y_\tau(t))| \geq \varepsilon) &\leq \mathbb{P}(\|X(t+\tau) - Y_\tau(t)\| \geq \delta_\varepsilon) \\
&\leq \delta_\varepsilon^{-1}\mathbb{E}\,\|X(t+\tau) - Y_\tau(t)\| \\
&\leq \delta_\varepsilon^{-1}\mathbb{E}\,\|X(t+\tau) - Y_\tau(t)\|^2.
\end{aligned}$$

As f is bounded,

$$\begin{aligned}
\mathbb{E}\,|f(X(t+\tau)) - f(Y_\tau(t))| &\leq 2\|f\|_\infty \mathbb{P}(|f(X(t+\tau)) - f(Y_\tau(t))| \geq \varepsilon) + \varepsilon \\
&\leq 2\|f\|_\infty \delta_\varepsilon^{-1}\mathbb{E}\,\|X(t+\tau) - Y_\tau(t)\|^2 + \varepsilon.
\end{aligned}$$

Hence for every $\varepsilon > 0$ there is a $t_0 \geq 0$ such that

$$\mathbb{E}\,|f(X(t+\tau)) - f(Y_\tau(t))| < \varepsilon \quad \text{for all } t \geq t_0,\ \tau \geq 0,$$

so that

$$\begin{aligned} |\mathbb{E}\,f(X(t+\tau)) - \mathbb{E}\,f(X(t))| &= |\mathbb{E}\,f(X(t+\tau)) - \mathbb{E}\,f(Y_\tau(t))| \\ &\leq \mathbb{E}\,|f(X(t+\tau)) - f(Y_\tau(t))| \to 0 \quad \text{as } t \to \infty, \end{aligned}$$

uniformly in $\tau \geq 0$.

Consequently, if $(s_n)_n$ and $(t_n)_n$ are sequences in $[0,\infty)$ with $s_n \to \infty$, $t_n \to \infty$, and $\mu_{X(s_n)} \to \mu$, $\mu_{X(t_n)} \to \nu$, for certain $\mu, \nu \in \mathcal{P}(H)$, then

$$\int f\,d\mu = \lim_{n\to\infty} \mathbb{E}\,f(X(s_n)) = \lim_{n\to\infty} \mathbb{E}\,f(X(t_n)) = \int f\,d\nu$$

for all bounded uniformly continuous $f : H \to \mathbb{R}$, and therefore $\mu = \nu$. $\square$

8 Invariant Measure

Let us now return to the equation (2) in the setting of Situation (*). For any initial value X_0, the solution $t \mapsto X(t, X_0)$ of (2) may be viewed as the evolution of the state X_0 in time. Often the real interest is more in the distributions of the solution than in the function itself. Because of the weak uniqueness (see Theorem 2.2), equation (2) also describes evolutions of measures. Indeed, if μ is a measure on H that occurs as the distribution of some initial value $X_0 \in L^2(\Omega; H)$, then $\mu_t = \mu_{X(t,X_0)}$, $t \geq 0$, gives the evolution of μ under (2) and it is well-defined as it does only depend on μ. The class of measures μ that can be dealt with this way is quite large, because the distributions do not depend on the underlying probability space and we may enlarge it. If μ is any Borel probability measure on H such that $\int_H \|x\|^2\,d\mu(x) < \infty$, we can take as probability space the product of $(H, \mathcal{B}(H), \mu)$ and $(\Omega, \mathcal{F}, \mathbb{P})$ and set $X_0(x,\omega) := x$, $x \in H$, $\omega \in \Omega$. (Here $\mathcal{B}(H)$ denotes the Borel σ-algebra of H.)

The measure μ is called a *stationary distribution* if $\mu_{X(t,X_0)} = \mu_{X_0} = \mu$ for all $t \geq 0$. The usual way to describe this notion of invariance is by means of the adjoints of the so called 'transition operators'. We present the definitions and some relevant properties below. We include their proofs, though all results can be collected from the existing literature (see [4, Chapter 11]).

Assume Situation (*). Denote the solution of (2) by $t \mapsto X(t, X_0)$. Accordingly, denote for $x \in H$ by $t \mapsto X(t,x)$ the solution of

$$X(t) = S(t)x + \int_0^t S(t-s)F(X(s))ds + \int_0^t S(t-s)\Phi(X(s))dW(s), \quad t \geq 0.$$

Let

$$C_b(H) := \{f : H \to \mathbb{R} : f \text{ is continuous and bounded}\},$$

which is a Banach space relative to the supremum norm $\|.\|_\infty$. The operators $T(t)$, $t \geq 0$, defined by

$$(T(t)f)(x) := \mathbb{E}\, f(X(t,x)), \quad x \in H,\ f \in C_b(H),$$

are called the *transition operators* of (2). For every $t \geq 0$ the operator $T(t)$ is a bounded linear operator from $C_b(H)$ to $C_b(H)$. Indeed, to see that $T(t)$ maps into $C_b(H)$, consider first two initial values $X_0, Y_0 \in L^2(\Omega; H)$ that are $\mathcal{F}_0$-measurable and consider the corresponding solutions $t \mapsto X(t, X_0)$ and $t \mapsto X(t, Y_0)$. By similar estimates as in the proof of Theorem 5.1 we find for $t \geq 0$ that

$$\begin{aligned}
&\mathbb{E}\,\|X(t,X_0) - X(t,Y_0)\|^2 \\
&\quad\leq\ 3\mathbb{E}\,\|S(t)X_0 - S(t)Y_0\|^2 \\
&\qquad +3\mathbb{E}\left(\int_0^t \|S(t-s)L_F\|X(s,X_0) - X(s,Y_0)\|\,ds\right)^2 \\
&\qquad +3\int_0^t \mathbb{E}\,\|S(t-s)\|^2 L_\Phi^2 \|X(s,X_0) - X(s,Y_0)\|^2\,ds \\
&\quad\leq\ 3\mathbb{E}\,\|S(t)\|^2\|X_0 - Y_0\|^2 \\
&\qquad +3\int_0^t \|S(t-s)\|^2 L_F^2 ds \int_0^t \mathbb{E}\,\|X(s,X_0) - X(s,Y_0)\|^2 ds \\
&\qquad +3\int_0^t \|S(t-s)\|^2 L_\Phi^2 \mathbb{E}\,\|X(s,X_0) - X(s,Y_0)\|^2 ds \\
&\quad\leq\ 3M_t^2\|X_0 - Y_0\|^2 \\
&\qquad +3(L_F^2 t + L_\Phi^2)M_t^2 \int_0^t \mathbb{E}\,\|X(s,X_0) - X(s,Y_0)\|^2 ds,
\end{aligned}$$

where $M_t = \max_{0\leq s\leq t}\|S(s)\|$. Gronwall's Lemma (Lemma 4.1) then yields

$$\mathbb{E}\,\|X(t,X_0) - X(t,Y_0)\|^2 \leq 3M_t^2\mathbb{E}\,\|X_0 - Y_0\|^2 e^{3(L_F^2 t + L_\Phi^2)M_t^2 t}, \qquad t \geq 0.$$

Hence if $t \geq 0$ and $(x_n)_n$ in H and $x \in H$ are such that $x_n \to x$, then $X(t,x_n) \to X(t,x)$ in $L^2(\Omega; H)$, so that for every subsequence $(x_{n_k})_k$ there is a subsubsequence such that $X(t, x_{n_{k_l}}) \to X(t,x)$ in H a.e. on Ω. For any $f \in C_b(H)$ we then have $f(X(t,x_{n_{k_l}})) \to f(X(t,x))$ a.e. on Ω and by dominated convergence $\mathbb{E}\,f(X(t,x_{n_{k_l}})) \to \mathbb{E}\,f(X(t,x))$. It follows that $\mathbb{E}\,f(X(t,x_n)) \to \mathbb{E}\,f(X(t,x))$ and thus continuity of $x \to (T(t)f)(x)$ is established. It is clear that $T(t) : C_b(H) \to C_b(H)$ is linear and bounded and that $\|T(t)\| = 1$ for every $t \geq 0$. Furthermore, for each $t \geq 0$ the operator $T(t)$ is positive, that is, $T(t)f$ is nonnegative for every nonnegative f.

To the operators $(T(t))_{t\geq 0}$ we can associate dual operators acting on the vector space $M(H)$ of finite signed Borel measures on H. Endowed with the norm given by $\|\mu\| = |\mu|(H)$, $\mu \in M(H)$, $M(H)$ is a Banach space. Denote by $C_b(H)'$ the norm dual of $(C_b(H), \|.\|_\infty)$ and let for $t \geq 0$ the operator $T(t)'$ be the adjoint of $T(t)$, that is,

$$(T(t)'\varphi)(f) = \varphi(T(t)f) \quad \text{for all } f \in C_b(H),\ \varphi \in C_b(H)'.$$

As for any measure $\mu \in M(H)$ the map $f \mapsto \int f d\mu$ is a member of $C_b(H)'$, we can identify $M(H)$ with a subspace of $C_b(H)'$, and the isomorphism is isometric as well as order preserving. For each $t \geq 0$ the restriction of $T(t)'$ to this subspace can then be viewed as an operator on $M(H)$ and we denote it by $T(t)^*$. Thus, $T(t)^*$ is the unique bounded linear operator on $M(H)$ such that

$$\int f\, d(T(t)^*\mu) = \int T(t)f\, d\mu \quad \text{for all } f \in C_b(H),\ \mu \in M(H).$$

For each $t \geq 0$, $\|T(t)^*\| = 1$, $T(t)^*$ is nonnegative, and for any Borel probability measure μ on H the measure $T(t)^*\mu$ is a Borel probability measure. The operators $T(t)^*$, $t \geq 0$, describe the evolution of the distributions:

Proposition 8.1. *Assume Situation (*). Let $(T(t))_{t\geq 0}$ be the transition operators associated to (2). Let $t \mapsto X(t, X_0)$ denote the solution of (2) and $\mu_{X(t,X_0)}$ the distribution of $X(t, X_0)$, $t \geq 0$. Then:*

(i) $\mu_{X(t,X_0)} = T(t)^*\mu_{X_0}$ *for every* $t \geq 0$,

(ii) $T(t+\tau)^*\mu_{X_0} = T(t)^*T(\tau)^*\mu_{X_0}$ *for every* $t, \tau \geq 0$,

(iii) $(T(t))_{t\geq 0}$ *is a semigroup on* $C_b(H)$, *and*

(iv) $(T(t)^*)_{t\geq 0}$ *is a semigroup on* $M(H)$.

Proof. (i): (See [4, Proposition 11.1, p.303].) As X_0 and $(W(t))_{t\geq 0}$ are independent, we may assume that $\Omega = \Omega_1 \times \Omega_2$, $\mathbb{P} = \mathbb{P}_1 \otimes \mathbb{P}_2$, and that X_0 is constant in its second argument and $W(t)$ in its first for all $t \geq 0$. For $t \geq 0$

and $f \in C_b(H)$ we have

$$\begin{aligned}
\int_H f\, d\mu_{X(t,X_0)} &= \int_\Omega f(X(t,X_0)(\omega_1,\omega_2))\, d\mathbb{P}(\omega_1,\omega_2) \\
&= \int_{\Omega_2}\int_{\Omega_1} f(X(t,X_0)(\omega_1,\omega_2))\, d\mathbb{P}_1(\omega_1) d\mathbb{P}_2(\omega_2) \\
&= \int_{\Omega_2}\int_{\Omega_1} f(X(t,X_0(\omega_1))(\omega_2))\, d\mathbb{P}_1(\omega_1) d\mathbb{P}_2(\omega_2) \\
&= \int_{\Omega_2}\int_H f(X(t,x))\, d\mu_{X_0} d\mathbb{P}_2(\omega_2) \\
&= \int_H \mathbb{E}\, f(X(t,x))\, d\mu_{X_0} \\
&= \int_H T(t)f\, d\mu_{X_0} = \int_H f\, d(T(t)^*\mu_{X_0}).
\end{aligned}$$

(ii): We have $\mu_{X(t,X_0)} = T(t)^*\mu_{X_0}$ for every $t \geq 0$. By rewriting (2) we get for any $t, \tau > 0$ that

$$X(t+\tau) = S(t)X(\tau) + \int_0^t S(t-s)F(X(s+\tau))ds + \int_0^t S(t-s)\Phi(X(s+\tau))\, dW(s+\tau) \tag{25}$$

and because of the weak uniqueness we obtain by applying (i) to (25) instead of (2) that

$$\mu_{X(t+\tau)} = T(t)^*\mu_{X(\tau)}.$$

Hence

$$T(t+\tau)^*\mu_{X_0} = \mu_{X(t+\tau)} = T(t)^*T(\tau)^*\mu_{X_0}, \qquad t, \tau \geq 0.$$

(iii): Let $t, \tau \geq 0$ and let $x \in H$. Applying (ii) to the initial value $X_0 = x$ yields that $T(t)^*\delta_x = T(t)^*T(\tau)^*\delta_x$, where δ_x denotes the Dirac measure at x. Hence for $f \in C_b(H)$ we have

$$\int T(t+\tau)f\, d\delta_x = \int T(\tau)T(t)f\, d\delta_x,$$

so that

$$(T(t+\tau)f)(x) = (T(\tau)T(t)f)(x).$$

Thus, $T(t+\tau) = T(\tau)T(t)$.

(iv): From (iii). □

A Borel probability measure μ on H is called an *invariant measure* for (2) (and also for (1)) if $T(t)^*\mu = \mu$ for all $t \geq 0$.

Proposition 8.2. *Assume Situation (*). Let $t \mapsto X(t, X_0)$ be the solution of (2) and denote by $\mu_{X(t,X_0)}$ the distribution of $X(t, X_0)$, $t \geq 0$. If $\mu_{X(t,X_0)}$ converges weakly as $t \to \infty$ to some Borel probability measure μ on H, then μ is an invariant measure for (2).*

Proof. For every $t, \tau \geq 0$ we have

$$\mu_{X(t+\tau)} = T(t+\tau)^* \mu_{X_0} = T(t)^* T(\tau)^* \mu_{X_0} = T(t)^* \mu_{X(\tau)},$$

so that

$$\int f \, d\mu_{X(t+\tau)} = \int T(t) f \, d\mu_{X(\tau)} \tag{26}$$

for all $f \in C_b(H)$. As $\int f \, d\mu_{X(r)} \to \int f \, d\mu$ as $r \to \infty$ for each $f \in C_b(H)$, we obtain by letting $\tau \to \infty$ in (26) that

$$\int f \, d\mu = \int T(t) f \, d\mu$$

for all $f \in C_b(H)$. Hence $\mu = T(t)^* \mu$ for all $t \geq 0$ and μ is an invariant measure. □

We have now established all ingredients for the proof of Theorem 3.1.

Proof of Theorem 3.1. (i) and (ii) follow from Theorem 5.1 and Corollary 3.3, and (iii) follows from (ii) and Proposition 8.2. □

Proof of Theorem 3.4. Let $P_1 = I$, $P_2 = 0$, $L_2 = 0$, and take $\alpha_2 > 0$ arbitrary. According to Corollary 5.2, the function $t \mapsto \mathbb{E} \|X(t)\|^2$ is bounded on $[0, \infty)$. Therefore we are in a position to apply Theorem 5.1 and the assertions then follow with aid of Theorem 3.2. □

9 Examples

In this final section we present some examples of stochastic delay differential equations to illustrate Theorem 3.1. They are not meant to show the full scope of possible applications, but serve merely as illustration and show at least one class of problems that could not be dealt with by earlier results. For general theory on (deterministic) delay differential equations we refer to [2] and [5]. Other approaches to stochastic delay differential equations can be found in the work by X. Mao, S. Mohammed, and M. Scheutzow, see [14, 15, 16, 17, 23].

The first example shows that delay differential equations do in general not fit in the framework of compact semigroups and dissipative generators, as we claimed in the introduction. In Example 9.3 we will add stochastic noise to the presented system and show that it does satisfy the conditions of Theorem 3.1.

Example 9.1. Consider the following system of delay differential equations

$$\left\{ \begin{array}{l} \frac{d}{dt} x_1(t) = -x_2(t-1) \\ \frac{d}{dt} x_2(t) = -x_1(t-1) \end{array} \right., \quad t \geq 0. \tag{27}$$

Let $W^{1,2} = W^{1,2}([-1,0];\mathbb{R}^2)$ denote the Sobolev space of functions with distributional derivative in $L^2([-1,0];\mathbb{R}^2)$ and denote for $\varphi \in W^{1,2}$ its derivative by φ'. It is well-known that $W^{1,2} \subset C([-1,0];\mathbb{R}^2)$. Define the spaces

$$\begin{aligned} H &= \mathbb{R}^2 \times L^2([-1,0];\mathbb{R}^2), \\ D(A) &= \{(c,\varphi) \in H : \varphi \in W^{1,2},\ c = \varphi(0)\}. \end{aligned}$$

If we let $L : C([-1,0];\mathbb{R}^2) \to \mathbb{R}^2$ be the bounded linear map defined by

$$L(\varphi_1,\varphi_2) = (-\varphi_2(-1), -\varphi_1(-1)), \quad (\varphi_1,\varphi_2) \in C([-1,0];\mathbb{R}^2),$$

then (27) can be formulated as the evolution equation

$$\frac{d}{dt} y(t) = Ay(t), \quad t \geq 0,$$

where y takes values in H and

$$A(c,\varphi) = (L\varphi, \varphi'), \quad (c,\varphi) \in D(A).$$

The operator A is the generator of a C_0-semigroup $(S(t))_{t\geq 0}$ in H and if $(c(t),u(t)) = S(t)(c_0,\varphi)$, $t \geq 0$, then $t \mapsto c(t)$ solves (27) with initial conditions $c(s) = \varphi(s)$ for a.e. $s \in [-1,0]$ and $c(0) = c_0$. Moreover, $u(t)(s) = c(t+s)$ for a.e. $s \in [-1,0]$.

We claim that A is not dissipative and that $S(t)$ is not compact for $0 \leq t < 1$. For a proof of the former, use the characterization of [21, Theorem 4.2, p.14]:

$$A \text{ is dissipative} \iff \|(\lambda I - A)x\| \geq \lambda \|x\| \text{ for all } x \in D(A) \text{ and } \lambda > 0. \quad (28)$$

Let $\varphi_{z,u}(t) = e^{zt}u$, $t \in [-1,0]$, with $z \in \mathbb{C}$ and $u \in \mathbb{R}^2$, and observe that $(u, \varphi_{z,u}) \in D(A)$ and

$$(zI - A)(u, \varphi_{z,u}) = (\Delta(z)u, 0),$$

where

$$\Delta(z) = \begin{pmatrix} z & e^{-z} \\ e^{-z} & z \end{pmatrix}, \quad z \in \mathbb{C}.$$

With $u = (1,0)$ we have for $\lambda > 0$ that

$$\begin{aligned} \|(\lambda I - A)(u, \varphi_{\lambda,u})\|^2 &= \lambda^2 + e^{-2\lambda}, \\ \|(u, \varphi_{\lambda,u})\|^2 &= 1 + (1 - e^{-2\lambda})/2\lambda. \end{aligned}$$

Hence for λ real and large the condition at the right hand side of (28) is not satisfied and it follows that A is not dissipative.

In order to establish that $S(t)$ is not compact for $0 \le t < 1$, let $(c, \varphi) \in H$ and let $x = (x_1, x_2)$ be the solution of (27) with initial condition $x(s) = \varphi(s)$ for a.e. $s \in [-1, 0]$ and $x(0) = c$. If we denote $x_t := s \mapsto x(t+s)$ $(-1 \le s \le 0)$ and $\varphi_t := s \mapsto \varphi(t+s)$ $(-1 \le s \le t-1)$, then for $0 \le t < 1$ we have $x_t(s) = \varphi(s)$ for a.e. $s \in [-1, t-1]$, so that $S(t)(c, \varphi) = (x(t), \varphi_t \mathbb{1}_{[-1,t-1]} + x_t \mathbb{1}_{[t-1,0]})$. It follows that the operator $S(t)$ on H is not compact.

The general form of the infinite dimensional equation that we considered in the preceding sections allows analysis of delay systems that have noise 'in the memory' of the system as well. We present an instance next.

Example 9.2. Consider the stochastic delay differential equation

$$dx(t) = \big(-x(t-1) + f(x(t), x_t)\big)dt + g(x(t), x_t)dW(t), \quad t \ge 0, \tag{29}$$

where

$$x_t(s) = x(t+s), \quad -1 \le s \le 0,\ t \ge 0,$$

$f, g : \mathbb{R} \times L^2[-1, 0] \to \mathbb{R}$ are Lipschitz continuous functions, and $(W(t))_{t\ge 0}$ is a normalized scalar Wiener process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In the same way as in Example 9.1, equation (29) can be formulated as an evolution equation

$$dy(t) = \big(Ay(t) + F(y(t))\big)dt + G(y(t))dW(t), \quad t \ge 0, \tag{30}$$

on the Hilbert space $H = \mathbb{R} \times L^2[-1, 0]$ with

$$F(c, \varphi) = (f(c, \varphi), 0), \quad G(c, \varphi) = (g(c, \varphi), 0), \quad (c, \varphi) \in H.$$

It can be shown that the linear operator A generates a strongly continuous semigroup on H that is exponentially stable. Therefore, if the Lipschitz constants of f and g are sufficiently small, Corollary 5.2 yields that equation (30) has an invariant measure. More than that, there is a $(c, \varphi) \in L^2(\Omega; H)$ such that equation (30) with initial condition $y(0) = (c, \varphi)$ is stationary. It follows that the solution x of (29) with initial conditions $x(0) = c$ and $x_0 = \varphi$ is such that the distribution of $x(t)$ equals the distributin of c for all $t \ge 0$ and the distribution of x_t $(\in L^2(\Omega; L^2[-1, 0]))$ equals the distribution of φ for all $t \ge 0$.

Finally we apply Theorem 3.1 to the system presented in Example 9.1 with additonal noise terms.

Example 9.3. Consider, again, the system of delay equations (27). We will first show that (27) is dichotomous. With the same notations as in Example 9.1, we have that the spectrum of A consists of the $z \in \mathbb{C}$ with $\det \Delta(z) = 0$. As

$$\det \Delta(z) = (z - e^{-z})(z + e^{-z}),$$

we find that there is no spectrum in a strip U around the imaginary axis and that $\lambda \mapsto \|(\lambda I - A)^{-1}\|$ is bounded on U. Hence the Gearhart-Herbst Theorem implies that $(S(t))_{t\geq 0}$ is hyperbolic and (27) is dichotomous.

Let $f, g : \mathbb{R}^2 \times L^2([-1,0];\mathbb{R}^2) \to \mathbb{R}$ be Lipschitz continuous functions and let $(W(t))_{t\geq 0}$ be a normalized scalar Wiener process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and consider the system

$$\left\{ \begin{array}{l} dx_1(t) = \big(-x_2(t-1) + f(x(t), x_t)\big)dt + g(x(t), x_t)dW(t) \\ dx_2(t) = \big(-x_1(t-1) + f(x(t), x_t)\big)dt + g(x(t), x_t)dW(t) \end{array} \right. , \qquad t \geq 0, \qquad (31)$$

where

$$x_t(s) := x(t+s) = (x_1(t+s), x_2(t+s)), \quad -1 \leq s \leq 0, \ t \geq 0.$$

By adding and subtracting we obtain

$$\left\{ \begin{array}{l} d(x_1 + x_2)(t) = \big(-(x_1 + x_2)(t-1) + 2f(x(t), x_t)\big)dt + 2g(x(t), x_t)dW(t) \\ d(x_1 - x_2)(t) = (x_1 - x_2)(t-1)dt \end{array} \right. , \qquad (32)$$

$t \geq 0$. This system can be written in the form of (1) or (2) with $H = \mathbb{R}^2 \times L^2([-1,0];\mathbb{R}^2)$. We see that if we choose an initial condition $(c_1, c_2, \varphi_1, \varphi_2) \in L^2(\Omega; H)$ such that $c_1 - c_2 = 0$ and $\varphi_1 - \varphi_2 = 0$, then $(x_1 - x_2)(t) = 0$ for all $t \geq 0$. Since the equation for $x_1 + x_2$ is stable (see also Example 9.2) any such initial condition yields a bounded solution if the Lipschitz constants of f and g are sufficiently small. It follows then from Theorem 3.1 that there exists an initial condition $(c, \varphi) \in L^2(\Omega; H))$ such that for the solution x of equation (31) with $x(0) = c$ and $x_0 = \varphi$ one has that the distribution of $(x(t), x_t)$ equals the distribution of (c, φ) for all $t \geq 0$.

References

[1] Coppel, W.A., *Dichotomies in stability theory* [LNM 629], Springer-Verlag, Berlin, New York, 1978.

[2] Curtain, Ruth F. and Hans Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory* [Texts in Applied Mathematics 21], Springer-Verlag, New York, 1995.

[3] Da Prato, Giuseppe, D. Gatarek, and Jerzy Zabczyk, Invariant measures for semilinear stochastic equations, *Stochastic Anal. Appl.* **10** (1992) no. 4, 387–408.

[4] Da Prato, Giuseppe and Jerzy Zabczyk, *Stochastic equations in infinite dimensions* [Encyclopedia of mathematics and its applications 45], Cambridge University Press, 1992.

[5] Diekmann, O., S.A. van Gils, S.M. Verduyn Lunel, and H.-O. Walther, *Delay Equations, Functional-, Complex-, and Nonlinear Analysis* [Applied Mathematical Sciences 110], Springer-Verlag, New York, 1995.

[6] Eckmann, J.-P. and M. Hairer, Invariant measures for stochastic partial differential equations in unbounded domains, *Nonlinearity* **14** (2001), 133–151.

[7] van Gaans, Onno, A series approach to stochastic differential equations with infinite dimensional noise, report MI 2002-07, Universiteit Leiden, 2002.

[8] Gearhart, L., Spectral theory for contraction semigroups on Hilbert space, *Trans. Amer. Math. Soc.* **236** (1978), 385–394.

[9] Henry, Daniel, *Geometric theory of semilinear parabolic equations* [LNM 840], Springer-Verlag, Berlin, New York, 1981.

[10] Herbst, I., The spectrum of Hilbert space semigroups, *J. Operator Theory* **10** (1983) no. 1, 87–94.

[11] Hille, Einar, *Lectures on Ordinary Differential Equations*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1969.

[12] Ichikawa, Akira, Semilinear Stochastic Evolution Equations: Boundedness, Stability and Invariant Measures, *Stochastics* **12** (1984), 1–39.

[13] Kaashoek, M.A. and S.M. Verduyn Lunel, An integrability condition on the resolvent for hyperbolicity of the semigroup, *J. Differential Equations* **112** (1994) no. 2, 374–406.

[14] Mao, Xuerong, Stability of stochastic integro-differential equations, *Stochastic Anal. Appl.* **18** (2000) no. 6, 1005–1017.

[15] Mao, Xuerong, Almost sure exponential stability of delay equations with damped stochastic perturbation. *Stochastic Anal. Appl.* **19** (2001) no. 1, 67–84.

[16] Mohammed, S.E.A., *Stochastic functional differential equations* [Research Notes in Mathematics 99], Pitman, Boston, 1984.

[17] Mohammed, Salah-Eldin A. and Michael K.R. Scheutzow, Lyapunov exponents of linear stochastic functional differential equations. II. Examples and case studies, *Ann. Probab.* **25** (1997) no. 3, 1210–1240.

[18] Nagel, R. (ed.), *One-Parameter semigroups of Positive Operators* [LNM 1184], Springer-Verlag, Berlin, 1980.

[19] Øksendal, Bernt, *Stochastic Differential Equations*, 5th edition, Springer-Verlag, Berlin - Heidelberg, 1998.

[20] Parthasarathy, K.R., *Probability Measures on Metric Spaces*, Academic Press, New York and London, 1967.

[21] Pazy, A., *Semigroups of linear operators and applications to partial differential equations* [Applied Mathematical Sciences 44], Springer-Verlag, New York, 1983.

[22] Sacker, R.J. and G.R. Sell, Dichotomies for linear evolution equations in Banach spaces, *J. Diff. Eq.* **113** (1994), 17–67.

[23] Scheutzow, M., Qualitative behaviour of stochastic delay equations with a bounded memory, *Stochastics* **12** (1984) no. 1, 42–80.

[24] Stroock, Daniel W. and S.R.S. Varadhan, *Multidimensional Diffusion Processes,* Springer-Verlag, Berlin, 1979.

O.W. van Gaans (`o.w.vangaans@its.tudelft.nl`)
Mathematical Institute, Leiden University, The Netherlands
Current address:
Department of Applied Mathematical Analysis, Faculty ITS
Delft University of Technology
P.O. Box 5031
2600 GA Delft
The Netherlands

S.M. Verduyn Lunel (`verduyn@math.leidenuniv.nl`)
Mathematical Institute
Leiden University
P.O. Box 9512
2300 RA Leiden
The Netherlands