CrossMark

Ergodicity and Kolmogorov Equations for Dissipative SPDEs with Singular Drift: a Variational Approach

Carlo Marinelli[1] · Luca Scarpa[1]

Abstract

We prove existence of invariant measures for the Markovian semigroup generated by the solution to a parabolic semilinear stochastic PDE whose nonlinear drift term satisfies only a kind of symmetry condition on its behavior at infinity, but no restriction on its growth rate is imposed. Thanks to strong integrability properties of invariant measures μ, solvability of the associated Kolmogorov equation in $L^1(\mu)$ is then established, and the infinitesimal generator of the transition semigroup is identified as the closure of the Kolmogorov operator. A key role is played by a generalized variational setting.

Keywords Stochastic PDEs · Invariant measures · Ergodicity · Monotone operators

Mathematics Subject Classification (2010) 60H15 · 47D07 · 47H06 · 37A25

1 Introduction

Our goal is to study the asymptotic behavior of solutions to semilinear stochastic partial differential equations on a smooth bounded domain $D \subseteq \mathbb{R}^n$ of the form

$$dX_t + AX_t \, dt + \beta(X_t) \, dt \ni B(X_t) \, dW_t, \qquad X(0) = X_0. \qquad (1.1)$$

Here $A : V \to V'$ is a linear maximal monotone operator from a Hilbert space V to its dual V', and $V \subset H := L^2(D) \subset V'$ is a so-called Gelfand triple; β is a maximal monotone graph everywhere defined on $\mathbb{R}$; W is a cylindrical Wiener process on a separable Hilbert space U, and B takes values in the space of Hilbert-Schmidt operators from U to $L^2(D)$. Precise assumptions on the data of the problem are given in Section 2 below. The most salient point is that β is *not* assumed to satisfy any growth assumption, but just a kind of symmetry on its rate of growth at plus and minus infinity – see assumption (vi) in Section 2 below. Well-posedness of Eq. 1.1 in the strong (variational) sense has recently been obtained in [17] by a combination of classical results by Pardoux and Krylov-Rozovskiĭ (see [12,

✉ Luca Scarpa
luca.scarpa.15@ucl.ac.uk

[1] Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK

19]) with pathwise estimates and weak compactness arguments. The minimal assumptions on the drift term β imply that, in general, the operator $A + \beta$ does not satisfy the coercivity and boundedness assumptions required by the variational approach of [12, 19]. For this reason, questions such as ergodicity and existence of invariant measures for Eq. 1.1 cannot be addressed using the results by Barbu and Da Prato in [4], which appear to be the only ones available for equations in the variational setting (cf. also [18]). On the other hand, there is a very vast literature on these problems for equations cast in the mild setting, references to which can be found, for instance, in [7, 8, 20]. Even in this case, however, we are not aware of results on equations with a drift term as general as in Eq. 1.1. Our results thus considerably extend, or at least complement, those on reaction-diffusion equations in [6–8], for instance, where polynomial growth assumptions are essential. More recent existence and integrability results for invariant measures of semilinear equations have been obtained, e.g., in [10, 11], but still under local Lipschitz-continuity or other suitable growth assumptions on the drift. Another possible advantage of our results is that we use only standard monotonicity assumptions, whereas in a large part of the cited literature one encounters assumptions of the type

$$\langle Ax + \beta(x + y)z \rangle \leq f(\|y\|) - k\|x\|$$

for some (or all) z belonging to the subdifferential of $\|x\|$, where f is a function and k a constant. Here A actually stands for the part of A in a Banach space E continuously embedded in $L^2(D)$, $\langle \cdot \cdot \rangle$ stands for the duality between E and its dual, and the condition is assumed to hold for those x, y for which all terms are well defined. Often E is chosen as a space of continuous functions such as $C(\overline{D})$. This monotonicity-type condition on A and β is precisely what one needs in order to obtain a priori estimates by reducing the original equation to a deterministic one with random coefficients, under the assumption of additive noise. Using a figurative but rather accurate expression, this methods amounts to "subtracting the stochastic convolution". Our estimates are obtained mostly by stochastic calculus, for which the standard notion of monotonicity suffices. Among such estimates we obtain the integrability of (the potential of) the nonlinear drift term β with respect to the invariant measure μ, which is known to be a delicate issue, especially for non-gradient systems (cf. the discussion in [10]). These results allow us to show that the Kolmogorov operator associated to the stochastic Eq. 1.1 with additive noise is essentially m-dissipative in $L^1(H, \mu)$. This implies that the closure of the Kolmogorov operator in $L^1(H, \mu)$ generates a Markovian semigroup of contractions, which is a μ-version of the transition semigroup generated by the solution to the stochastic equation. It is worth mentioning that the variational-type setting, while allowing for a very general drift term β, gives raise to quite many technical issues in the study of Kolmogorov equations, for instance because test functions in function spaces on V and V' naturally appear.

We conclude this introductory section with a brief description of the structure of the paper and of the main results. In Section 2 we state the basic assumptions which are in force throughout the paper, and recall the well-posedness result for Eq. 1.1 obtained in [17]. For the reader's convenience we collect in Section 3 some tools needed in the sequel, such as Prokhorov's theorem on compactness of sets of probability measures, and the Krylov–Bogoliubov criterion for the existence of an invariant measure for a Markovian transition semigroup. Section 4 is devoted to auxiliary results, most of which should be interesting in their own right, that underpin our subsequent arguments. In particular, we prove two generalized versions of the classical Itô formula in the variational setting for Eq. 1.1: one for the square of the norm, and another one extending a very useful but not-so-well known version for more general smooth functions, originally obtained by

Pardoux (see [19, p. 62–ff]). Furthermore, we establish results on the first and second-order differentiability with respect to the initial datum, both in the Gâteaux and Fréchet sense, of (variational) solutions to semilinear equations with regular drift and additive noise. In Section 5 we prove that the transition semigroup P generated by the solution to Eq. 1.1 admits an ergodic invariant measure μ, which in also shown to be unique and strongly mixing if β is superlinear. These results follow mainly by a priori estimates (which, in turn, are obtained by stochastic calculus) and compactness. Finally, Section 6 deals with the Kolmogorov equation associated to (1.1) with additive noise. In particular, we characterize the infinitesimal generator $-L$ of the transition semigroup P on $L^1(H, \mu)$ as the closure of the Kolmogorov operator $-L_0$. After showing that L_0 is dissipative and coincides with L on a suitably chosen dense subset of $L^1(H, \mu)$, we prove that the image of $I + L_0$ is dense in $L^1(H, \mu)$, so that the Lumer-Phillips theorem can be applied. Due to the variational formulation of the problem, the latter point turns out to be rather delicate, even though the general approach follows a typical scheme: we first introduce appropriate regularizations of L_0, for which the Kolmogorov equation can be solved by established techniques, then we pass to the limit in the regularization's parameters. Here the generalized Itô formulas and the differentiability results proved in Section 4 play a key role.

2 General Assumptions and Well-Posedness

Before stating the hypotheses on the coefficients and on the initial datum of Eq. 1.1 that will be in force throughout the paper, let us fix some notation.

2.1 Notation

Given two Banach (real) spaces E and F, the space of bounded linear operators from E to F will be denoted by $\mathscr{L}(E, F)$. When $F = \mathbb{R}$, we shall just write E'. If E and F are Hilbert spaces, $\mathscr{L}^2(E, F)$ stands for the ideal of $\mathscr{L}(E, F)$ of Hilbert-Schmidt operators. The Hilbert space $L^2(D)$ will be denoted by H, and its norm and scalar product by $\| \cdot \|$ and $\langle \cdot \cdot \rangle$, respectively. For any topological space E, the Borel σ-algebra on E will be denoted by $\mathscr{B}(E)$. All measures on E are intended to be defined on its Borel σ-algebra, unless otherwise stated. The spaces of bounded Borel-measurable and bounded continuous functions on E will be denoted by $B_b(E)$ and $C_b(E)$, respectively.

2.2 Assumptions

Let V be a separable Hilbert space densely, continuously and compactly embedded in $H = L^2(D)$. The duality form between V and V' is also denoted by $\langle \cdot \cdot \rangle$, as customary. We assume that $A \in \mathscr{L}(V, V')$ satisfies the following properties:

(i)   there exists $C > 0$ such that $\langle Avv \rangle \geq C\|v\|_V^2$ for every $v \in V$;
(ii) the part of A in H can be uniquely extended to an m-accretive operator A_1 on $L^1(D)$;
(iii) for every $\delta > 0$, the resolvent $(I + \delta A_1)^{-1}$ is sub-Markovian, i.e. for every $f \in L^1(D)$ such that $0 \leq f \leq 1$ a.e. on D, we have $0 \leq (I + \delta A)^{-1} f \leq 1$ a.e. on D;
(iv) there exists $m \in \mathbb{N}$ such that $(I + \delta A_1)^{-m} \in \mathscr{L}(L^1(D), L^\infty(D))$.

Let us now consider the non-linear term in the drift. We assume that

(v) $\beta \subset \mathbb{R} \times \mathbb{R}$ is a maximal monotone graph such that $0 \in \beta(0)$ and $\mathsf{D}(\beta) = \mathbb{R}$.

Let $j : \mathbb{R} \to \mathbb{R}_+$ be the unique convex lower semicontinuous function such that $j(0) = 0$ and $\beta = \partial j$, i.e. β is the subdifferential of j in the sense of convex analysis.[1] We assume that

(vi) $\displaystyle \limsup_{|r| \to \infty} \frac{j(r)}{j(-r)} < \infty.$

This hypothesis is obviously satisfied if j is even (or, equivalently, if β is odd). For example, the function $j : \mathbb{R} \ni r \mapsto |r|^p$, with $p \geq 1$, satisfies (vi). More generally, (vi) is satisfied by any convex polynomial function with even leading coefficient. The absence of growth conditions on j allows us to deal also with functions whose rate of growth at infinity is faster than polynomial, such as $j(r) = e^{r^2}$. Examples of superpolynomial functions j satisfying (vi) without being even can be easily constructed, for instance

$$j(x) := \begin{cases} e^{|x|} - 1, & x \leq 0, \\ 0, & x \in [0, 1], \\ e^{|x|} - e, & x \geq 1. \end{cases}$$

Denoting the convex conjugate of j by j^*, it is well known that the hypothesis $\mathsf{D}(\beta) = \mathbb{R}$ is equivalent to the superlinearity of j^* at infinity (see, e.g., [3, Proposition 1.8]), i.e.

$$\lim_{|r| \to \infty} \frac{j^*(r)}{|r|} = \infty.$$

We are going to need the following property implied by assumption (vi): there exists a strictly positive number η such that, for every measurable function $y : D \to \mathbb{R}$, $j^*(y) \in L^1(D)$ implies $j^*(\eta|y|) \in L^1(D)$. In fact, from (vi) we deduce that there exist $R > 0$ and $M_1 = M_1(R) > 0$ such that $j(r) \leq M_1 j(-r)$ for $|r| \geq R$. Since $j \geq 0$, one can choose $M_1 > 1$ without loss of generality. Setting $M_2 := \max\{j(r) : |r| \leq R\}$, which is finite by continuity of j, we deduce that

$$j(r) \leq M_1 j(-r) + M_2 \qquad \forall r \in \mathbb{R}.$$

Taking convex conjugates on both sides we infer that

$$j^*(r) \geq M_1 j^*(-r/M_1) - M_2 \qquad \forall r \in \mathsf{D}(j^*).$$

Setting $\eta := 1/M_1 < 1$ and recalling that $j^*(0) = 0$, hence j^* is positive on $\mathbb{R}$ and increasing on $\mathbb{R}_+$, one has

$$\begin{aligned} j^*(\eta|y|) &= j^*(\eta y) 1_{\{y \geq 0\}} + j^*(-\eta y) 1_{\{y < 0\}} \\ &\leq j^*(y) 1_{\{y \geq 0\}} + \eta j^*(y) 1_{\{y \geq 0\}} + \eta M_2 \\ &\leq j^*(y) + M_2 \in L^1(D). \end{aligned}$$

The assumptions on the Wiener process W and the diffusion coefficient B are standard: let U be a separable Hilbert space and W a cylindrical Wiener process on U, defined on a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \in [0,T]}, \mathbb{P})$ satisfying the so-called usual conditions.[2] We assume that

(vii) $B : H \to \mathscr{L}^2(U, H)$ is Lipschitz-continuous, i.e. that there exists a positive constants L_B such that

$$\|B(x) - B(y)\|_{\mathscr{L}^2(U,H)} \leq L_B \|x - y\| \qquad \forall x, y \in H.$$

---

[1] All notions of convex analysis used throughout the paper can be found, e.g., in [3, Chapters 1–2].
[2] Expressions involving random elements are always meant to hold $\mathbb{P}$-a.s. unless otherwise stated.

Note that this readily implies that B has linear growth, as

$$\begin{aligned}
\|B(x)\|_{\mathscr{L}^2(U,H)} &\leq \|B(x) - B(0)\|_{\mathscr{L}^2(U,H)} + \|B(0)\|_{\mathscr{L}^2(U,H)} \\
&\leq \|B(0)\|_{\mathscr{L}^2(U,H)} + L_B\|x\|.
\end{aligned}$$

Finally, the initial datum X_0 is assumed to be $\mathscr{F}_0$-measurable and such that $\mathbb{E}\|X_0\|^2$ is finite. All hypotheses just stated will be tacitly assumed to hold throughout.

The following well-posedness result for Eq. 1.1 is a refined version, proved in [15], of the main result of [17], where the diffusion coefficient B can also be random and time-dependent.

Theorem 2.1 *There is a unique pair* (X, ξ)*, with* X *a* V*-valued adapted process and* ξ *an* $L^1(D)$*-valued predictable process, such that*

$$X \in L^2(\Omega; C([0, T]; H)) \cap L^2(\Omega; L^2(0, T; V)), \qquad \xi \in L^1(\Omega \times (0, T) \times D),$$
$$j(X) + j^*(\xi) \in L^1(\Omega \times (0, T) \times D), \qquad \xi \in \beta(X) \quad \text{a.e. in } \Omega \times (0, T) \times D,$$

and

$$X(t) + \int_0^t AX(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(X(s))\,dW(s) \quad \forall t \in [0, T], \quad \mathbb{P}\text{-a.s.}$$

in $V' \cap L^1(D)$*. Moreover, the solution map*

$$L^2(\Omega; H) \longrightarrow L^2(\Omega; C([0, T]; H)) \cap L^2(\Omega; L^2(0, T; V))$$
$$X_0 \longmapsto X$$

is Lipschitz-continuous.

Remark 2.2 Theorem 2.1 was proved in [17] under the stronger assumption that j is even, which was later removed in [15] (see also [16] for a more general setting). In particular, the symmetry of j was only used in [17, p. 1483, p. 1492] to infer that if a measurable function $v : D \to \mathbb{R}$ is such that $j(v) \in L^1(D)$, then one also has $j(-v) \in L^1(D)$. As is easily checked, the argument continues to hold assuming only that $j(-r) \lesssim 1 + j(r)$ for all $r \in \mathbb{R}$. By continuity of j, this is the case if j satisfies condition (vi) above. Similarly, in [17] the solution was only shown to be pathwise weakly continuous, while the strong continuity of trajectories was obtained in [15]. In the latter work it is also shown that well-posedness continues to hold under local Lipschitz-continuity and linear growth assumptions on B.

3 Preliminaries

3.1 Compactness in Spaces of Probability Measures

The set of probability measures on E is denoted by $\mathscr{M}_1(E)$ and endowed with the topology $\sigma(\mathscr{M}_1(E), C_b(E))$, which we shall call the narrow topology. We recall that a subset $\mathscr{N}$ of $\mathscr{M}_1(E)$ is called (uniformly) *tight* if for every $\varepsilon > 0$ there exists a compact set K_ε such that $\mu(E \setminus K_\varepsilon) < \varepsilon$ for all $\mu \in \mathscr{N}$. The following characterization of relative compactness of sets of probability measures is classical (see, e.g., [5, §5.5]).

Theorem 3.1 (Prokhorov) *Let* E *be a complete separable metric space. A subset of* $\mathscr{M}_1(E)$ *is relatively compact in the narrow topology if and only if it is tight.*

3.2 Markovian Semigroups and Ergodicity

A family $P = (P_t)_{t \geq 0}$ of Markovian kernels on a measure space $(E, \mathcal{E})$ such that $P_{t+s} = P_t P_s$ for all $t, s \geq 0$ is called a Markovian semigroup. We recall that a Markovian kernel on $(E, \mathcal{E})$ is a map $K : E \times \mathcal{E} \to [0, 1]$ such that (i) $x \mapsto K(x, A)$ is $\mathcal{E}$-measurable for each $A \in \mathcal{E}$, (ii) $A \mapsto K(x, A)$ is a measure on $\mathcal{E}$ for each $x \in E$, and (iii) $K(x, E) = 1$ for each $x \in E$. A Markovian kernel K on $(E, \mathcal{E})$ can naturally be extended to the space $b\mathcal{E}$ of $\mathcal{E}$-measurable bounded functions by the prescription

$$f \longmapsto Kf := \int_E f(y) \, K(\cdot, dy).$$

Then $K : b\mathcal{E} \to b\mathcal{E}$ is a linear, bounded, positive, σ-order continuous map. Similarly, K can be extended to positive measures on $\mathcal{E}$ setting

$$\mu \longmapsto \mu K(\cdot) := \int_E K(x, \cdot) \, \mu(dx).$$

The notations $P_t f$ and μP_t, with f $\mathcal{E}$-measurable bounded or positive function and μ positive measure on $\mathcal{E}$, are hence to be understood in this sense. We shall also assume that $P_0 = I$ and that $(t, x) \mapsto P_t f(x)$ is $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$-measurable.

A probability measure μ on $\mathcal{E}$ is said to be an invariant measure for the Markovian semigroup P if

$$\int_E P_t f \, d\mu = \int_E f \, d\mu \qquad \forall f \in b\mathcal{E}, \quad \forall t \geq 0,$$

or, equivalently, if $\mu P_t = \mu$ for all $t \geq 0$. If P admits an invariant measure μ, then it can be extended to a Markovian semigroup on $L^p(E, \mu)$, for every $p \geq 1$.

The invariant measure μ is said to be ergodic for P if

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t P_s f \, ds = \int_E f \, d\mu \qquad \text{in } L^2(E, \mu) \quad \forall f \in L^2(E, \mu),$$

and strongly mixing if

$$\lim_{t \to +\infty} P_t f = \int_E f \, d\mu \qquad \text{in } L^2(E, \mu) \quad \forall f \in L^2(E, \mu).$$

We recall the following classical fact on the structure of the set of ergodic measures: the ergodic invariant measures for P are the extremal points of the set of its invariant measures. In particular, if P admits a unique invariant measure μ, then μ is ergodic.

In order to state a criterion for the existence of invariant measures, let us introduce, for any probability measure $\nu \in \mathcal{M}_1(E)$, the family of averaged measures $(\mu_t^\nu)_{t \geq 0}$ defined as

$$\mu_t^\nu := \frac{1}{t} \int_0^t \nu P_s \, ds.$$

Theorem 3.2 (Krylov and Bogoliubov) *Let $(P_t)_{t \geq 0}$ be a (time-homogeneous) Markovian transition semigroup on a complete separable metric space E. Assume that*

(a) *$(P_t)_{t \geq 0}$ has the Feller property, i.e. that it maps $C_b(E)$ into $C_b(E)$;*
(b) *there exists $\nu \in \mathcal{M}_1(E)$ such that the $(\mu_t^\nu)_{t \geq 0} \subset \mathcal{M}_1(E)$ is tight.*

Then the set of invariant measures for $(P_t)_{t \geq 0}$ is non-empty.

Note that if $x \in E$ and ν is the Dirac measure at x, then $\nu P_s = P_s(x, \cdot)$. Then condition (b) is satisfied if there exists $x \in E$ such that the family of measures

$$\left(\frac{1}{t} \int_0^t P_s(x, \cdot) \, ds \right)_{t \geq 0}$$

is tight. It is easily seen that this latter condition is in turn satisfied if $(P_t(x, \cdot))_{t \geq 0} \subset \mathcal{M}_1(E)$ is tight.

4 Auxiliary Results

To prove the main results we shall need some auxiliary results that are interesting in their own right, and that are collected in this section. In particular, we recall or prove some Itô-type formulas and provide conditions for the differentiability of solutions to equations in variational form with respect to the initial datum.

4.1 Itô Formulas

The proof of the following version of Itô's formula for the square of the H-norm in a generalized variational setting follows from the proof of [17, Proposition 6.2].

Proposition 4.1 *Assume that an adapted process*

$$Y \in L^0(\Omega; L^\infty(0, T; H)) \cap L^0(\Omega; L^2(0, T; V))$$

is such that

$$Y(t) + \int_0^t AY(s) \, ds + \int_0^t g(s) \, ds = Y_0 + \int_0^t G(s) \, dW(s)$$

in $L^1(D)$ for all $t \in [0, T]$, where $Y_0 \in L^0(\Omega, \mathscr{F}_0; H)$, G is a progressive $\mathscr{L}^2(U, H)$-valued process such that

$$G \in L^2(\Omega \times (0, T); \mathscr{L}^2(U, H)),$$

g is an adapted $L^1(D)$-valued process such that

$$g \in L^0(\Omega; L^1(0, T; L^1(D))),$$

and there exists $\alpha > 0$ for which

$$j(\alpha Y) + j^*(\alpha g) \in L^1(\Omega \times (0, T) \times D).$$

Then

$$\frac{1}{2} \|Y(t)\|^2 + \int_0^t \langle AY(s), Y(s) \rangle \, ds + \int_0^t \int_D g(s, x) Y(s, x) \, dx \, ds$$
$$= \frac{1}{2} \|Y_0\|^2 + \frac{1}{2} \int_0^t \|G(s)\|_{\mathscr{L}^2(U, H)}^2 \, ds + \int_0^t Y(s) G(s) \, dW(s) \qquad \forall t \in [0, T].$$

We shall also need a simplified version of an Itô formula in the variational setting, due to Pardoux, for functions other than the square of the H-norm. For its proof (in a more general context) we refer to [19, p. 62-ff.].

Proposition 4.2 *Let $Y \in L^0(\Omega; L^2(0, T; V))$ be such that*

$$Y(t) = Y_0 + \int_0^t v(s) \, ds + \int_0^t G(s) \, dW(s)$$

for all $t \in [0, T]$, where $Y_0 \in L^0(\Omega, \mathscr{F}_0, \mathbb{P}; H)$ and

$$v \in L^0(\Omega; L^1(0, T; H)) \oplus L^0(\Omega; L^2(0, T; V'))$$

is adapted and $G \in L^2(\Omega \times (0, T); \mathscr{L}^2(U, H))$ is progressively measurable. Then, for any $F \in C_b^2(H) \cap C_b^1(V)$, one has

$$F(Y(t)) = F(Y_0) + \int_0^t DF(Y(s))v(s) \, ds + \int_0^t DF(Y(s))G(s) \, dW(s)$$

$$+ \frac{1}{2} \int_0^t \text{Tr}\big(G^*(s)D^2F(Y(s))G(s)\big) \, ds$$

for every $t \in [0, T]$, $\mathbb{P}$-almost surely.

The previous Itô formula can be extended to processes satisfying weaker integrability conditions, in analogy to Proposition 4.1.

Proposition 4.3 *Let $Y \in L^0(\Omega; L^2(0, T; V)) \cap L^0(\Omega; L^\infty(0, T; H))$ be such that*

$$Y(t) = Y_0 + \int_0^t Av(s) \, ds + \int_0^t g(s) \, ds + \int_0^t G(s) \, dW(s)$$

for all $t \in [0, T]$, where $Y_0 \in L^0(\Omega, \mathscr{F}_0, \mathbb{P}; H)$ and

$$v \in L^0(\Omega; L^2(0, T; V)), \qquad g \in L^0(\Omega; L^1(0, T; L^1(D)))$$

are adapted and $G \in L^2(\Omega \times (0, T); \mathscr{L}^2(U, H))$ is progressively measurable. Then, for any $F \in C_b^2(H) \cap C_b^1(V') \cap C_b^1(L^1(D))$, one has

$$F(Y(t)) = F(Y_0) + \int_0^t \langle Av(s), DF(Y(s)) \rangle \, ds + \int_0^t \int_D g(s)DF(Y(s)) \, ds$$

$$+ \int_0^t DF(Y(s))G(s) \, dW(s) + \frac{1}{2} \int_0^t \text{Tr}\big(G^*(s)D^2F(Y(s))G(s)\big) \, ds$$

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

Proof Since the resolvent of A_1 is ultracontractive by assumption, there exists $m \in \mathbb{N}$ such that

$$(I + \delta A_1)^{-m} : L^1(D) \to H \qquad \forall \delta > 0.$$

Using a superscript δ to denote the action of $(I + \delta A_1)^{-m}$, we have

$$Y^\delta(t) = Y_0^\delta + \int_0^t Av^\delta(s) \, ds + \int_0^t g^\delta(s) \, ds + \int_0^t G^\delta(s) \, dW(s) \qquad \forall t \in [0, T],$$

where $Av^\delta + g^\delta \in L^0(\Omega; L^1(0, T; H)) \oplus L^0(\Omega; L^2(0, T; V'))$. Hence, by Proposition 4.2, for every $\delta > 0$ we have

$$F(Y^\delta(t)) = F(Y_0^\delta) + \int_0^t \langle Av^\delta(s), DF(Y^\delta(s)) \rangle \, ds + \int_0^t \int_D g^\delta(s)DF(Y^\delta(s)) \, ds$$

$$+ \int_0^t DF(Y^\delta(s))G^\delta(s) \, dW(s) + \frac{1}{2} \int_0^t \text{Tr}\big((G^\delta)^*(s)D^2F(Y^\delta(s))G^\delta(s)\big) \, ds$$

for every $t \in [0, T]$, $\mathbb{P}$-almost surely. Let us pass to the limit as $\delta \to 0$ in the previous equation. It is clear from the fact that $Y(t), Y_0 \in H$ and the continuity of F that

$$F(Y^\delta(t)) \to F(Y(t)), \qquad F(Y_0^\delta) \to F(Y_0).$$

Moreover, since $v^\delta + \delta A v^\delta = v$ in V, taking the duality pairing with $Av^\delta \in V'$, we have

$$\langle Av^\delta, v^\delta \rangle + \delta \|Av^\delta\|^2 = \langle Av^\delta, v \rangle \leq \|A\|_{\mathscr{L}(V,V')} \|v^\delta\|_V \|v\|_V,$$

from which, by coercivity of A,

$$\|v^\delta\|_V \leq \frac{\|A\|_{\mathscr{L}(V,V')}}{C} \|v\|_V \qquad \forall \delta > 0.$$

Taking into account that $v \in L^2(0, T; V)$, we deduce that $v^\delta \to v$ weakly in $L^2(0, T; V)$. Since $Y^\delta \to Y$ in $L^2(0, T; H)$, by continuity of A and the fact that $DF \in C_b(H, V)$, we have $Av^\delta \to Av$ weakly in $L^2(0, T; V')$ and $DF(Y^\delta) \to DF(Y)$ in $L^2(0, T; V)$, hence

$$\int_0^t \langle Av^\delta(s), DF(Y^\delta(s)) \rangle \, ds \longrightarrow \int_0^t \langle Av(s), DF(Y(s)) \rangle \, ds.$$

Furthermore, since $Y^\delta(t) \to Y(t)$ in H for every $t \in [0, T]$, recalling that $DF \in C_b(H, L^\infty(D))$ and $g^\delta \to g$ in $L^1(0, T; L^1(D))$, we have (possibly along a subsequence)

$$\int_D g^\delta(s) DF(Y^\delta(s)) \longrightarrow \int_D g(s) DF(Y(s)) \qquad \text{for a.e. } s \in (0, T).$$

Taking into account that $\int_D g^\delta DF(Y^\delta) \leq \|DF\|_{C_b(H, L^\infty(D))} \|g\|_{L^1(D)} \in L^1(0, T)$, by the dominated convergence theorem we then have

$$\int_0^t \int_D g^\delta(s) DF(Y^\delta(s)) \, ds \longrightarrow \int_0^t \int_D g(s) DF(Y^\delta(s)) \, ds.$$

Moreover, since $Y^\delta(t) \to Y(t)$ in H for every $t \in [0, T]$, recalling that $D^2 F \in C(H, \mathscr{L}(H))$ and $G^\delta \to G$ in $L^2(\Omega; L^2(0, T; \mathscr{L}^2(U, H)))$, we have (possibly along a subsequence)

$$\text{Tr}\big((G^\delta)^*(s) D^2 F(Y^\delta(s)) G^\delta(s)\big) \to \text{Tr}\big(G^*(s) D^2 F(Y(s)) G(s)\big) \qquad \text{for a.e. } s \in (0, T).$$

Since $\text{Tr}\big((G^\delta)^* D^2 F(Y^\delta) G^\delta\big) \leq \|D^2 F\|_{C(H, \mathscr{L}(H))} \|G\|_{\mathscr{L}^2(U, H)}^2 \in L^1(0, T)$, the dominated convergence theorem yields

$$\int_0^t \text{Tr}\big((G^\delta)^*(s) D^2 F(Y^\delta(s)) G^\delta(s)\big) \, ds \longrightarrow \int_0^t \text{Tr}\big(G^*(s) D^2 F(Y(s)) G(s)\big) \, ds.$$

Finally, by the Burkholder-Davis-Gundy inequality (with exponent equal to one) and by the ideal property of Hilbert-Schmidt operators, we have

$$\mathbb{E} \sup_{t \in [0,T]} \left| \int_0^t DF(Y^\delta(s)) G^\delta(s) \, dW(s) - \int_0^t DF(Y(s)) G(s) \, dW(s) \right|$$

$$\lesssim \mathbb{E} \left(\int_0^T \| DF(Y^\delta(s)) G^\delta(s) - DF(Y(s)) G(s) \|^2_{\mathscr{L}(U,\mathbb{R})} \, ds \right)^{1/2}$$

$$\lesssim \mathbb{E} \left(\int_0^T \| DF(Y^\delta(s)) \|^2 \| G^\delta(s) - G(s) \|^2_{\mathscr{L}^2(U,H)} \, ds \right)^{1/2}$$

$$+ \mathbb{E} \left(\int_0^T \| DF(Y^\delta(s)) - DF(Y(s)) \|^2 \| G^\delta(s) \|^2_{\mathscr{L}^2(U,H)} \, ds \right)^{1/2}$$

$$\leq \| DF \|_{C(H,H)} \| G^\delta - G \|_{L^2(\Omega; L^2(0,T; \mathscr{L}^2(U,H)))}$$

$$+ \mathbb{E} \left(\int_0^T \| G(s) \|^2_{\mathscr{L}^2(U,H)} \| DF(Y^\delta(s)) - DF(Y(s)) \|^2 \, ds \right)^{1/2},$$

where the first term on the right-hand side converges to 0 because

$$G^\delta \to G \qquad \text{in } L^2(\Omega; L^2(0,T; \mathscr{L}^2(U,H))).$$

Similarly, since $DF(Y^\delta) \to DF(Y)$ a.e., it follows by the dominated convergence theorem that the second term on the right-hand side converges to zero as well. Therefore, passing to subsequence if necessary, one has

$$\int_0^t DF(Y^\delta(s)) G^\delta(s) \, dW(s) \longrightarrow \int_0^t DF(Y(s)) G(s) \, dW(s). \qquad \square$$

4.2 Differentiability with Respect to the Initial Datum for Solutions to Equations in Variational Form

Let $g \in C_b^2(\mathbb{R})$ and consider the equation

$$dX + AX \, dt = g(X) \, dt + G \, dW, \qquad X(0) = x,$$

in the variational sense, where A satisfies the hypotheses of Section 2, $G \in \mathscr{L}^2(U, H)$, and $x \in H$. Note that throughout this subsection the diffusion coefficient G is independent of X, so that we are considering only the case of equations with additive noise. For compactness of notation we shall write E in place of $C([0, T]; H) \cap L^2(0, T; V)$. The above equation admits a unique variational solution $X^x \in L^2(\Omega; E)$. Here and in the following we often use superscripts to denote the dependence on the initial datum. We are going to provide sufficient conditions ensuring that the solution map $x \mapsto X^x$ belongs to $C_b^2(H; L^2(\Omega; E))$. The problem of regular dependence on the initial datum for equations in the variational setting does not seem to be addressed in the literature. On the other hand, several results are available for mild solutions (see, e.g., [6, 8, 14]), where an approach via the implicit function theorem depending on a parameter is adopted. Here we proceed in a more direct and, we believe, clearer way. The results are non-trivial (and probably not easily accessible via the implicit function theorem) in the sense that the solution map is Fréchet differentiable even though, as is well known, the superposition operator associated to g is *never* Fréchet differentiable unless g is affine. The first and second Fréchet derivative of the solution map shall be denoted by DX and D^2X, respectively. These are maps with domain H and codomain $\mathscr{L}(H, L^2(\Omega; E))$ and $\mathscr{L}_2(H; L^2(\Omega; E))$, respectively. Here and in the following

we denote the space of continuous bilinear mappings from $H \times H$ to a Banach space F by $\mathscr{L}_2(H; F)$.

We begin with first-order differentiability.

Theorem 4.4 *The solution map $x \mapsto X^x : H \to L^2(\Omega; E)$ is continuously (Fréchet) differentiable with bounded derivative. Moreover, for any $h \in H$, setting $Y_h := (DX)h$, one has*

$$Y_h' + AY_h = g'(X^x)Y_h, \qquad Y_h(0) = h, \tag{4.1}$$

in the variational sense.

Proof Classical (deterministic) results imply that Eq. 4.1 admits a unique solution $Y_h \in E$ for $\mathbb{P}$-a.e. $\omega \in \Omega$. Since X^x is an adapted process and h is non-random, it follows that Y_h is itself adapted. Alternatively, and more directly, one can apply the stochastic variational theory to Eq. 4.1, deducing that $Y_h \in L^2(\Omega; E)$ is adapted.

Let us set, for compactness of notation,

$$X_\varepsilon := X^{x+\varepsilon h}, \qquad z_\varepsilon := \frac{1}{\varepsilon}(X_\varepsilon - X) - Y_h,$$

where ε is an arbitrary real number. Elementary calculations show that

$$z_\varepsilon(t) + \int_0^t Az_\varepsilon(s)\, ds = \int_0^t \left(\frac{1}{\varepsilon}\big(g(X_\varepsilon(s)) - g(X(s))\big) - g'(X(s))Y_h(s) \right) ds.$$

Writing

$$g(X_\varepsilon) - g(X) = g(X + \varepsilon Y_h) - g(X) + g(X_\varepsilon) - g(X + \varepsilon Y_h)$$

yields

$$\begin{aligned}
\frac{1}{\varepsilon}\big(g(X_\varepsilon) - g(X)\big) - g'(X)Y_h &= \frac{1}{\varepsilon}\big(g(X + \varepsilon Y_h) - g(X)\big) - g'(X)Y_h \\
&\quad + \frac{1}{\varepsilon}\big(g(X_\varepsilon) - g(X + \varepsilon Y_h)\big) \\
&=: R_\varepsilon + S_\varepsilon.
\end{aligned}$$

By the integration-by-parts formula applied to the equation for z_ε we get

$$\frac{1}{2}\|z_\varepsilon(t)\|^2 + \int_0^t \langle Az_\varepsilon(s), z_\varepsilon(s)\rangle\, ds = \int_0^t \langle R_\varepsilon(s), z_\varepsilon(s)\rangle\, ds + \int_0^t \langle S_\varepsilon(s), z_\varepsilon(s)\rangle\, ds,$$

where $\langle S_\varepsilon, z_\varepsilon\rangle \leq \|S_\varepsilon\|\|z_\varepsilon\|$ and, by the Lipschitz continuity of g,

$$\|S_\varepsilon\| \leq \|g\|_{\dot{C}^{0,1}} \frac{1}{\varepsilon}\|X_\varepsilon - X - \varepsilon Y_h\| = \|g\|_{\dot{C}^{0,1}}\|z_\varepsilon\|,$$

so that $\langle S_\varepsilon, z_\varepsilon\rangle \leq \|g\|_{\dot{C}^{0,1}}\|z_\varepsilon\|^2$. Since $\langle R_\varepsilon, z_\varepsilon\rangle \leq \big(\|R_\varepsilon\|^2 + \|z_\varepsilon\|^2\big)/2$, we are left with

$$\frac{1}{2}\|z_\varepsilon(t)\|^2 + \int_0^t \langle Az_\varepsilon(s), z_\varepsilon(s)\rangle\, ds \leq \big(1/2 + \|g\|_{\dot{C}^{0,1}}\big)\int_0^t \|z_\varepsilon(s)\|^2\, ds + \frac{1}{2}\int_0^t \|R_\varepsilon(s)\|^2\, ds.$$

For an arbitrary $t > 0$ one has, by the coercivity of A,

$$\frac{1}{2}\|z_\varepsilon\|_{C([0,t];H)}^2 + C\int_0^t \|z_\varepsilon(s)\|_V^2\, ds \leq \big(1 + 2\|g\|_{\dot{C}^{0,1}}\big)\int_0^t \|z_\varepsilon\|_{C([0,s];H)}^2\, ds + \int_0^t \|R_\varepsilon(s)\|^2\, ds,$$

hence also, by Fubini's theorem and Gronwall's inequality,

$$\mathbb{E}\|z_\varepsilon\|_{C([0,T];H)}^2 \leq e^{(2+4\|g\|_{\dot{C}^{0,1}})T}\mathbb{E}\int_0^T \|R_\varepsilon(s)\|^2\, ds.$$

It is clear from the hypotheses on g and the definition of R_ε that $R_\varepsilon \to 0$ in $L^0(\Omega \times [0, T] \times D)$ as $\varepsilon \to 0$ for every $s \in [0, T]$. Moreover, it follows by the Lipschitz continuity of g and elementary estimates that $|R_\varepsilon| \lesssim \|g\|_{\dot{C}^{0,1}} |Y_h|\|$, where $Y_h \in L^2(\Omega \times [0, T] \times D)$. The dominated convergence theorem thus yields

$$\lim_{\varepsilon \to 0} \mathbb{E} \int_0^T \|R_\varepsilon(s)\|^2 \, ds = 0.$$

Since

$$C\mathbb{E} \int_0^T \|z_\varepsilon(s)\|_V^2 \, ds \le \left(1 + 2\|g\|_{\dot{C}^{0,1}}\right) T \, \mathbb{E} \|z_\varepsilon\|_{C([0,T];H)}^2 + \mathbb{E} \int_0^T \|R_\varepsilon(s)\|^2 \, ds,$$

we conclude that

$$\lim_{\varepsilon \to 0} \|z_\varepsilon\|_{L^2(\Omega; E)} = 0.$$

This proves that the solution map is differentiable in every direction of H, and that its directional derivative in the direction $h \in H$ is given by the (unique) solution Y_h to Eq. 4.1. It is then clear that the map $h \mapsto Y_h$ is linear. Let us prove that it is also continuous: in analogy to computations already carried out above, the integration-by-parts formula yields

$$\frac{1}{2}\|Y_h(t)\|^2 + \int_0^t \langle AY_h(s), Y_h(s)\rangle \, ds = \|h\|^2 + \int_0^t \langle g'(X^x(s))Y_h(s)\|Y_h(s) \, ds,$$

from which one infers

$$\|Y_h\|_{C([0,t];H)}^2 + \|Y_h\|_{L^2(0,t;V)}^2 \lesssim \|h\|^2 + \int_0^t \|Y_h\|_{C([0,s];H)}^2 \, ds,$$

hence also, by Gronwall's inequality and elementary estimates,

$$\|Y_h\|_E \lesssim \|h\|.$$

It is important to note that this inequality holds $\mathbb{P}$-a.s. with a non-random implicit constant that depends only on T and on the Lipschitz constant of g, but not on the initial datum x. From this it follows that

$$\|Y_h\|_{L^p(\Omega; E)} \lesssim_T \|h\| \qquad \forall p \ge 0,$$

hence, in particular, that $h \mapsto Y_h$ is the Gâteaux derivative of $x \mapsto X^x$. Setting $Y^x := h \mapsto Y_h$, we are going to prove that the map

$$H \longrightarrow \mathscr{L}(H, L^2(\Omega; E))$$
$$x \longmapsto Y^x$$

is continuous. This implies, by a well-known criterion (see, e.g., [2, Theorem 1.9]), that $x \mapsto X^x$ is Fréchet differentiable with Fréchet derivative (necessarily) equal to Y^x. Let $(x_n) \subset H$ be a sequence converging to x in H, and write for simplicity $X^n := X^{x_n}$, $Y^n := Y^{x_n}$, $X := X^x$, and $Y := Y^x$, with a subscript h to denote their action on a fixed element $h \in H$. One has

$$Y_h^n(t) - Y_h(t) + \int_0^t A(Y_h^n(s) - Y_h(s)) \, ds = x_n - x + \int_0^t \left(g'(X^n)Y_h^n - g'(X)Y_h\right)(s) \, ds,$$

for which the integration-by-parts formula yields

$$\frac{1}{2}\|Y_h^n(t) - Y_h(t)\|^2 + C\int_0^t \|Y_h^n(s) - Y_h(s)\|_V^2 \, ds$$
$$\le \frac{1}{2}\|x^n - x\|^2 + \int_0^t \langle g'(X^n)Y_h^n - g'(X)Y_hY_h^n - Y_h\rangle(s) \, ds,$$

where

$$\langle g'(X^n)Y_h^n - g'(X)Y_h, Y_h^n - Y_h \rangle = \langle g'(X_n)(Y_h^n - Y_h), Y_h^n - Y_h \rangle$$
$$+ \langle (g'(X^n) - g'(X))Y_h, Y_h^n - Y_h \rangle,$$

so that, by elementary estimates,

$$\|Y_h^n(t) - Y_h(t)\|^2 + 2C \int_0^t \|Y_h^n(s) - Y_h(s)\|_V^2 \, ds$$

$$\leq \|x_n - x\|^2 + \left(2\|g\|_{\dot{C}^{0,1}} + 1\right) \int_0^t \|Y_h^n(s) - Y_h(s)\|^2 \, ds$$

$$+ \int_0^t \|\left(g'(X^n(s)) - g'(X(s))\right)Y_h(s)\|^2 \, ds.$$

Taking the supremum in time, Gronwall's inequality implies

$$\|Y_h^n - Y_h\|_E \lesssim \|x_n - x\| + \|(g'(X^n) - g'(X))Y_h\|_{L^2(0,T;H)},$$

where the implicit constant depends on C, T and on the Lipschitz constant of g. Furthermore, since, as observed above, $h \mapsto Y_h$ is a linear bounded map from H to $C([0, T]; H)$ $\mathbb{P}$-a.s. with non-random operator norm, i.e.

$$\sup_{\|h\| \leq 1} \|Y_h\|_{C(0,T];H)} \lesssim_{T,g} 1,$$

one has

$$\mathbb{E} \sup_{\|h\| \leq 1} \|(g'(X^n) - g'(X))Y_h\|_{L^2(0,T;H)} \lesssim \mathbb{E} \|g'(X^n) - g'(X)\|_{C(0,T];H)},$$

and the last term converges to zero as $n \to \infty$ by the dominated convergence theorem, because $X^n \to X$ in $L^2(\Omega; C([0, T]; H))$ and $g \in C_b^2$ (in particular, g' is Lipschitz-continuous). It immediately follows that $x \mapsto Y^x$ is a continuous map on H with values in $\mathscr{L}(H, L^2(\Omega; E))$. Furthermore, since we have shown that $\|Y_h^x\|_{L^p(\Omega;E)} \lesssim \|h\|$ for all $p \geq 0$ with a constant independent of x, we conclude that $x \mapsto X^x$ is of class C_b^1 from H to $L^2(\Omega; E)$. □

To establish the second-order Fréchet differentiability of $x \mapsto X^x$, it is convenient to consider the equation

$$Z'_{hk} + AZ_{hk} = g'(X)Z_{hk} + g''(X)Y_h Y_k, \qquad Z_{hk}(0) = 0, \tag{4.2}$$

where $h, k \in H$ and Y_h, Y_k are the solutions to Eq. 4.1 with initial conditions h and k, respectively. This is manifestly the equation formally satisfied by the second-order Fréchet derivative of $x \mapsto X^x$ evaluated at (h, k).

In order to prove that Eq. 4.2 is well-posed, we need the following lemma, which is probably well known, but for which we could not find a reference, except for the classical case where $f \in L^2(0, T; V')$ (see, e.g., [13]).

Lemma 4.5 *Let $y_0 \in H$, $f \in L^1(0, T; H)$, and $\ell \in L^\infty((0, T) \times D)$. Then there exists a unique*

$$y \in C([0, T]; H) \cap L^2(0, T; V)$$

such that

$$y(t) + \int_0^t Ay(s) \, ds = y_0 + \int_0^t \ell(s)y(s) \, ds + \int_0^t f(s) \, ds \qquad \forall t \in [0, T].$$

Moreover, one has

$$\frac{1}{2}\|y(t)\|^2 + \int_0^t \langle Ay(s), y(s)\rangle\, ds = \frac{1}{2}\|y_0\|^2 + \int_0^t \int_D \ell(s)|y(s)|^2\, ds$$

$$+ \int_0^t \langle f(s), y(s)\rangle\, ds \qquad \forall t \in [0, T].$$

Proof Let (f_n) be a sequence in $L^2(0, T; H)$ such that $f_n \to f$ in $L^1(0, T; H)$ as $n \to \infty$. By the variational theory of deterministic equations, for every $n \in \mathbb{N}$ there exists a unique

$$y_n \in H^1(0, T; V') \cap L^2(0, T; V) \hookrightarrow C([0, T]; H)$$

such that

$$y_n'(t) + Ay_n(t) = \ell(t)y_n(t) + f_n(t) \quad \text{in } V' \text{ for a.e. } t \in (0, T), \qquad y_n(0) = y_0.$$

Therefore, for every $n, m \in \mathbb{N}$, the integration-by-parts formula and an easy computation show that

$$\|y_n(t) - y_m(t)\|^2 + 2C\int_0^t \|y_n(s) - y_m(s)\|_V^2\, ds$$

$$\leq 2\|\ell\|_{L^\infty((0,T)\times D)} \int_0^t \|y_n(s) - y_m(s)\|^2\, ds + 2\int_0^t \langle f_n(s)$$

$$- f_m(s), y_n(s) - y_m(s)\rangle\, ds$$

$$\leq 2\|\ell\|_{L^\infty((0,T)\times D)} \int_0^t \|y_n(s) - y_m(s)\|^2\, ds + 2\|y_n$$

$$- y_m\|_{C([0,t];H)} \|f_n - f_m\|_{L^1(0,T;H)}$$

for every $t \in [0, T]$. By the Young inequality we infer then that, for every $\varepsilon \geq 0$,

$$\|y_n - y_m\|_{C([0,t];H)}^2 + \|y_n - y_m\|_{L^2(0,t;V)}^2$$

$$\lesssim \varepsilon\|y_n - y_m\|_{C([0,t];H)}^2 + \frac{1}{4\varepsilon}\|f_n - f_m\|_{L^1(0,T;H)}^2 + \int_0^t \|y_n - y_m\|_{C([0,s];H)}^2\, ds$$

for every $t \in [0, T]$, from which, thanks to Gronwall's inequality,

$$\|y_n - y_m\|_{C([0,T];H)\cap L^2(0,T;V)} \lesssim \|f_n - f_m\|_{L^1(0,T;H)}.$$

We deduce that there exists $y \in C([0, T]; H) \cap L^2(0, T; V)$ such that

$$y_n \to y \qquad \text{in } C([0, T]; H) \cap L^2(0, T; V).$$

It clear follows from $y \in L^2(0, T; V)$ and $A \in \mathscr{L}(V, V')$ that $Ay \in L^2(0, T; V')$ and $Ay_n \to Ay$ in $L^2(0, T; V')$ as $n \to \infty$. Moreover, we also have that

$$\frac{1}{2}\|y_n(t)\|^2 + \int_0^t \langle Ay_n(s), y_n(s)\rangle\, ds = \frac{1}{2}\|y_0\|^2 + \int_0^t \int_D \ell(s)|y_n(s)|^2\, ds + \int_0^t \langle f_n(s), y_n(s)\rangle\, ds$$

for all $t \in [0, T]$. Hence the last assertion follows letting $n \to \infty$. The uniqueness of y is a consequence of the monotonicity of A. $\qquad\square$

In order to prove second-order Fréchet differentiability of the solution map $x \mapsto X^x$ we need to make the further assumption that V is continuously embedded in $L^4(D)$. This is

satisfied, for instance, if $V = H_0^1$ and $d \leq 4$. In fact, by the Sobolev embedding theorem, $H_0^1 \hookrightarrow L^{2^*}$, where

$$\frac{1}{2^*} = \frac{1}{2} - \frac{1}{d}$$

for $d \geq 3$ and $2^* = +\infty$ otherwise.

We proceed as follows: first we establish well-posedness for equation Eq. 4.2, and then we show that its unique solution identifies $D^2 X$.

Proposition 4.6 *Assume that V is continuously embedded in $L^4(D)$. Then* Eq. 4.2 *admits a unique variational solution Z_{hk} for any $h, h \in H$. Moreover, the map*

$$Z^x : H \times H \to L^2(\Omega, E), \qquad (h, k) \mapsto Z_{hk}^x$$

is bilinear and continuous for any $x \in H$, and there exists a positive constant $M > 0$ such that

$$\|Z^x\|_{\mathscr{L}_2(H; L^2(\Omega; E))} \leq M \qquad \forall x \in H.$$

Proof Hölder's inequality and the boundedness of g'' yield

$$\|g''(X)Y_h Y_k\| \leq \|g''\|_{\dot{C}^{0,1}} \|Y_h\|_{L^4(D)} \|Y_k\|_{L^4(D)} \lesssim \|Y_h\|_V \|Y_k\|_V,$$

so that $g''(X)Y_h Y_k \in L^1(0, T; H)$ since $Y_h, Y_k \in L^2(0, T; V)$. Hence, by Lemma 4.5 there is a unique

$$Z_{hk} \in C([0, T]; H) \cap L^2(0, T; V)$$

such that

$$Z_{hk}(t) + \int_0^t A Z_{hk}(s)\, ds = \int_0^t g'(X(s)) Z_{hk}(s)\, ds + \int_0^t g''(X(s)) Y_h(s) Y_k(s)\, ds$$
$$\forall t \in [0, T].$$

Let us show that $(h, k) \mapsto Z_{hk}$ is a continuous bilinear map. The bilinearity is clear from Eq. 4.2. Moreover, testing by Z_{hk} and using the coercivity of A we have that

$$\|Z_{hk}(t)\|^2 + \int_0^t \|Z_{hk}(s)\|_V^2\, ds \lesssim \|g\|_{C_b^1} \int_0^t \|Z_{hk}(s)\|^2\, ds$$

$$+ \|g\|_{C_b^2} \int_0^t \|Y_h(s)\|_V \|Y_k(s)\|_V\, ds$$

$$\leq \|g\|_{C_b^1} \int_0^t \|Z_{hk}(s)\|^2\, ds + \|g\|_{C_b^2} \|Y_h\|_{L^2(0,T;V)} \|Y_k\|_{L^2(0,T;V)}$$

$$\lesssim_T \|g\|_{C_b^1} \int_0^t \|Z_{hk}(s)\|^2\, ds + \|g\|_{C_b^2} \|h\| \|k\|$$

and Gronwall's inequality yields

$$\|Z_{hk}^x\|_{L^2(\Omega; C([0,T];H)) \cap L^2(\Omega; L^2(0,T;V))} \lesssim \|h\| \|k\| \qquad \forall h, k, x \in H,$$

from which the last assertion follows. □

Theorem 4.7 *Assume that V is continuously embedded in $L^4(D)$. Then the solution map $x \mapsto X^x$ is of class C_b^2 from H to $L^2(\Omega; E)$.*

Proof We are going to prove first that the Fréchet derivative of the solution map is Gâteaux-differentiable, with Gâteaux derivative equal to $Z^x := (h, k) \mapsto Z_{hk}^x$, then we shall then show that $x \mapsto Z^x$ is continuous and bounded as a map from H to $\mathscr{L}_2(H; L^2(\Omega; E))$.

Step 1. Let $x \in H$ be arbitrary but fixed, and consider the family of maps $z^\varepsilon \in \mathscr{L}_2(H; L^2(\Omega; E))$, indexed by $\varepsilon \in \mathbb{R}$, defined as

$$z^\varepsilon : (h, k) \longmapsto z^\varepsilon_{hk} := \frac{1}{\varepsilon} \left(Y^{x+\varepsilon k}_h - Y^x_h \right) - Z^x_{hk}.$$

Elementary manipulations based on the equations satisfied by Y^x and Z^x show that

$$z^\varepsilon_{hk}(t) + \int_0^t A z^\varepsilon_{hk}(s)\, ds = \int_0^t \left(\frac{g'(X^\varepsilon)Y^\varepsilon_h - g'(X)Y_h}{\varepsilon} - g'(X)Z_{hk} - g''(X)Y_h Y_k \right)(s)\, ds,$$

where the integrand on the right-hand side can be written as $R_\varepsilon + S_\varepsilon$, with

$$R_\varepsilon = \left(\frac{g'(X^\varepsilon) - g'(X)}{\varepsilon} - g''(X)Y_k \right) Y_h,$$

$$S_\varepsilon = \left(g'(X^\varepsilon) \frac{Y^\varepsilon_h - Y_h}{\varepsilon} - g'(X)Z_{hk} \right).$$

Further algebraic manipulations show that $R_\varepsilon = R'_\varepsilon + R''_\varepsilon$ and $S_\varepsilon = S'_\varepsilon + S''_\varepsilon$, where

$$R'_\varepsilon := \left(\frac{g'(X + \varepsilon Y_k) - g'(X)}{\varepsilon} - g''(X)Y_k \right) Y_h,$$

$$R''_\varepsilon := \frac{g'(X^\varepsilon) - g'(X + \varepsilon Y_k)}{\varepsilon} Y_h,$$

$$S'_\varepsilon := g'(X) z^\varepsilon_{hk},$$

$$S''_\varepsilon := \left(g'(X^\varepsilon) - g'(X) \right) \frac{Y^\varepsilon_h - Y_h}{\varepsilon}.$$

The integration-by-parts formula and obvious estimates yield

$$\frac{1}{2} \| z^\varepsilon_{hk}(t) \|^2 + C \int_0^t \| z^\varepsilon_{hk}(s) \|^2_V\, ds \leq \| g \|_{\dot{C}^{0,1}} \int_0^t \| z^\varepsilon_{hk}(s) \|^2_V\, ds + \int_0^t \langle R_\varepsilon + S''_\varepsilon, z^\varepsilon_{hk} \rangle(s)\, ds,$$

Taking the supremum on both sides, one is left with, thanks to Young's inequality,

$$\| z^\varepsilon_{hk} \|^2_{C([0,t];H)} + \| z^\varepsilon_{hk} \|^2_{L^2(0,t;V)} \lesssim \delta \| z^\varepsilon_{hk} \|^2_{C([0,t];H)} + \int_0^t \| z^\varepsilon_{hk} \|^2_{C([0,s];H)}\, ds$$
$$+ \frac{1}{\delta} \| R_\varepsilon + S''_\varepsilon \|^2_{L^1(0,T;H)}$$

for all $\delta > 0$, from which it follows, taking δ sufficiently small and applying Gronwall's inequality,

$$\mathbb{E} \| z^\varepsilon_{hk} \|^2_E \lesssim \mathbb{E} \| R_\varepsilon \|^2_{L^1(0,T;H)} + \mathbb{E} \| S''_\varepsilon \|^2_{L^1(0,T;H)}.$$

We are going to show that the right-hand side tends to zero as $\varepsilon \to 0$. Since $g \in C^2_b$, it is evident that $R'_\varepsilon \to 0$ almost everywhere as $\varepsilon \to 0$ as well as that

$$|R'_\varepsilon| \leq 2 \| g'' \|_\infty |Y_k Y_h|.$$

Since

$$\sup_{\|h\| \leq 1} \| Y_h Y_k \|_{L^1(0,T;H)} \lesssim \sup_{\|h\| \leq 1} \| Y_h \|_{L^2(0,T;V)} \| Y_k \|_{L^2(0,T;V)} \lesssim \sup_{\|h\| \leq 1} \|h\| \|k\| \leq \|k\|,$$

the dominated convergence theorem yields

$$\lim_{\varepsilon \to 0} \sup_{\|h\| \leq 1} \mathbb{E} \| R'_\varepsilon \|^2_{L^1(0,T;H)} = 0.$$

Moreover, we have

$$|R_\varepsilon''| \leq \|g''\|_\infty \left| \frac{X^{x+\varepsilon k} - X^x}{\varepsilon} - Y_k^x \right| |Y_h^x|,$$

so that

$$\|R_\varepsilon''\|_{L^1(0,T;H)} \lesssim \left\| \frac{X^{x+\varepsilon k} - X^x}{\varepsilon} - Y_k^x \right\|_{L^2(0,T;V)} \|Y_h^x\|_{L^2(0,T;V)},$$

where $\|Y_h^x\|_{L^2(0,T;V)} \lesssim \|h\|$, hence, by Theorem 4.4,

$$\sup_{\|h\| \leq 1} \mathbb{E} \|R_\varepsilon''\|_{L^1(0,T;H)}^2 \lesssim \mathbb{E} \left\| \frac{X^{x+\varepsilon k} - X^x}{\varepsilon} - Y_k^x \right\|_{L^2(0,T;V)}^2 \to 0$$

Finally, from

$$|S_\varepsilon''| \leq \|g''\|_\infty \left| \frac{X^{x+\varepsilon k} - X^x}{\varepsilon} \right| |Y_h^{x+\varepsilon k} - Y_h^x|$$

we deduce

$$\|S_\varepsilon'\|_{L^1(0,T;H)} \lesssim \left\| \frac{X^{x+\varepsilon k} - X^x}{\varepsilon} \right\|_{L^2(0,T;V)} \|Y_h^{x+\varepsilon k} - Y_h^x\|_{L^2(0,T;V)}.$$

Since $(X^{x+\varepsilon k} - X^x)/\varepsilon \to Y_k^x$ in E as $\varepsilon \to 0$ and $x \mapsto Y_h^x$ is continuous from H to E, we infer that $\|S_\varepsilon'\|_{L^1(0,T;H)} \to 0$. Moreover, it follows from

$$\|Y_h^{x+\varepsilon k} - Y_h^x\|_{L^2(0,T;V)} \leq 2\|h\|$$

that

$$\sup_{\|h\| \leq 1} \|S_\varepsilon'\|_{L^1(0,T;H)} \lesssim \left\| \frac{X^{x+\varepsilon k} - X^x}{\varepsilon} \right\|_{L^2(0,T;V)}.$$

Recalling that, by Theorem 4.4, $(X^{x+\varepsilon k} - X^x)/\varepsilon \to Y_k^x$ in $L^2(\Omega; E)$ as $\varepsilon \to 0$, this implies

$$\lim_{\varepsilon \to 0} \sup_{\|h\| \leq 1} \mathbb{E} \|S_\varepsilon''\|_{L^1(0,T;H)}^2 = 0.$$

We thus conclude that

$$\lim_{\varepsilon \to 0} \sup_{\|h\| \leq 1} \|z_{hk}^\varepsilon\|_{L^2(\Omega; E)} = 0 \qquad \forall k \in H,$$

i.e. the directional derivative of $x \mapsto Y^x : H \mapsto \mathcal{L}(H, L^2(\Omega; E))$ exists for all directions and is given by the map $x \mapsto Z^x : H \to \mathcal{L}_2(H; L^2(\Omega; E))$. Since we have already proved that $(h, k) \mapsto Z_{hk}^x$ is bilinear and continuous, we infer that $x \mapsto Y^x$ is Gâteaux differentiable with derivative Z^x.

Step 2. In order to conclude that $x \mapsto Y^x$ is Fréchet differentiable (with derivative necessarily equal to Z) it is enough to show, in view of a criterion already mentioned, that the map

$$x \longmapsto Z^x$$
$$H \longrightarrow \mathcal{L}_2(H; L^2(\Omega; E))$$

is continuous. Let $(x_n)_n \subseteq H$ be a sequence converging to x in H. We have, writing Z^n in place of Z^{x_n} for simplicity,

$$(Z_{hk}^n - Z_{hk})' + A(Z_{hk}^n - Z_{hk}) = g'(X^n)Z_{hk}^n - g'(X)Z_{hk} + g''(X^n)Y_h^n Y_k^n - g''(X)Y_h Y_k,$$

with initial condition $Z_{hk}^n(0) - Z_{hk}(0) = 0$. The right-hand side of the equation can be written as $R = \sum_{i \leq 4} R_i$, with

$$R_1 := g'(X^n)(Z_{hk}^n - Z_{hk}), \qquad\qquad R_2 := (g'(X^n) - g'(X))Z_{hk},$$
$$R_3 := g''(X^n)(Y_h^n Y_k^n - Y_h Y_k), \qquad R_4 := (g''(X^n) - g''(X))Y_h Y_k,$$

so that, by the integration-by-parts formula,

$$\frac{1}{2}\|Z_{hk}^n(t) - Z_{hk}(t)\|^2 + C\int_0^t \|Z_{hk}^n(s) - Z_{hk}(s)\|_V^2\, ds \leq \int_0^t \langle R, Z_{hk}^n - Z_{hk}\rangle(s)\, ds,$$

where

$$\int_0^t \langle R_1, Z_{hk}^n - Z_{hk}\rangle(s)\, ds \leq \|g'\|_\infty \int_0^t \|Z_{hk}^n(s) - Z_{hk}(s)\|^2\, ds,$$

and, for $i \neq 1$, by Young's inequality,

$$\int_0^t \langle R_i, Z_{hk}^n - Z_{hk}\rangle(s)\, ds \leq \|Z_{hk}^n - Z_{hk}\|_{C([0,t];H)}\|R_i\|_{L^1(0,t;H)}$$

$$\leq \delta\|Z_{hk}^n - Z_{hk}\|_{C([0,t];H)}^2 + \frac{1}{\delta}\|R_i\|_{L^1(0,t;H)}^2.$$

By an argument based on the Gronwall's inequality already used several times we obtain

$$\|Z_{hk}^n - Z_{hk}\|_E^2 \lesssim \|R_2 + R_3 + R_4\|_{L^1(0,T;H)}^2,$$

where $\|R_2\| \leq \|g''\|_\infty\|(X^n - X)Z_{hk}\|$ and, by the bilinearity of Z,

$$\|(X^n - X)Z_{hk}\|_{L^1(0,T;H)} \lesssim \|X^n - X\|_{L^2(0,T;V)}\|Z_{hk}\|_{L^2(0,T;V)} \lesssim \|X^n - X\|_{L^2(0,T;V)}\|h\|\|k\|,$$

from which it follows

$$\sup_{\|h\|,\|k\|\leq 1} \mathbb{E}\|(X^n - X)Z_{hk}\|_{L^1(0,T;H)}^2 \lesssim \|X^n - X\|_{L^2(\Omega;L^2(0,T;V))} \to 0$$

because $x \mapsto X^x$ is continuous from H to $L^2(\Omega; E)$. Moreover, since $\|R_3\| \leq \|g''\|_\infty\|Y_h^n Y_k^n - Y_h Y_k\|$, we have, recalling that $V \hookrightarrow L^4$,

$$\|R_3\|_{L^1(0,T;H)} \leq \|Y_h^n - Y_h\|_{L^2(0,T;V)}\|Y_k\|_{L^2(0,T;V)} + \|Y_k^n - Y_k\|_{L^2(0,T;V)}\|Y_h^n\|_{L^2(0,T;V)},$$

where both terms on the right-hand side tend to zero because $Y_h^n \to Y_h$ in $L^2(0,T;V)$ for all $h \in H$. The estimate

$$\|Y_h^n Y_k^n - Y_h Y_k\|_{L^1(0,T;H)} \lesssim \|h\|\|k\|$$

then implies, by the dominated convergence theorem,

$$\sup_{\|h\|,\|k\|\leq 1} \mathbb{E}\|R_3\|_{L^1(0,T;H)}^2 \lesssim \sup_{\|h\|,\|k\|\leq 1} \mathbb{E}\|Y_h^n Y_k^n - Y_h Y_k\|_{L^1(0,T;H)}^2 \to 0.$$

It remains to consider R_4: it is clear that $(g''(X^n) - g''(X))Y_h Y_k \to 0$ almost everywhere by the continuity of g'', and, as before,

$$\|(g''(X^n) - g''(X))Y_h Y_k\|_{L^1(0,T;H)} \lesssim \|g''\|_\infty\|h\|\|k\|,$$

hence the dominated convergence theorem yields

$$\sup_{\|h\|,\|k\|\leq 1} \mathbb{E}\|R_4\|_{L^1(0,T;H)}^2 \to 0.$$

We have thus proved that, as $n \to \infty$,

$$\|Z^n - Z\|_{\mathscr{L}_2(H;L^2(\Omega;E))} = \sup_{\|h\|,\|k\|\leq 1} \|Z_{hk}^n - Z_{hk}\|_{L^2(\Omega;E)} \to 0.$$

Recalling that $x \mapsto Z^x$ is bounded on H, we conclude that $x \mapsto X^x$ is twice Fréchet-differentiable with continuous and bounded derivatives. $\qquad\square$

5 Invariant measures

Throughout this section, we consider Eq. 1.1 with $X_0 \in H$. Since all coefficients do not depend explicitly on $\omega \in \Omega$, it follows by a standard argument that the solution X to Eq. 1.1 is Markovian. Let $P = (P_t)_{t \geq 0}$ be the transition semigroup defined by

$$(P_t \varphi)(x) := \mathbb{E}\, \varphi(X^x(t)) \qquad \forall x \in H, \quad \varphi \in C_b(H).$$

Lemma 5.1 *The transition semigroup P is Feller, i.e. $C_b(H)$ is invariant under the action of P_t for every $t \geq 0$.*

Proof Let $\varphi \in C_b(H)$, $t \geq 0$, $x \in H$, and $(x_n) \subset H$ be a sequence converging to x in H as $n \to \infty$. It is enough to show that $(P_t \varphi)(x_n) \to (P_t \varphi)(x)$ as $n \to \infty$, as the boundedness of $P_t \varphi$ is immediate by definition of P. By the continuous dependence property with respect to the initial datum established by Theorem 2.1, we have that

$$\| X^x - X^{x_n} \|_{L^2(\Omega; C([0,T]; H))} \lesssim \| x - x_n \|,$$

hence, in particular, $X^{x_n}(t) \to X^x(t)$ in $L^2(\Omega; H)$. This in turn implies that $X^{x_n}(t) \to X^x(t)$ in probability, thus also, by continuity of φ, that $\varphi(X^{x_n}(t)) \to \varphi(X^x(t))$ in probability. Vitali's theorem then readily yields that $\varphi(X^{x_n}(t)) \to \varphi(X^x(t))$ in $L^1(\Omega)$, from which it immediately follows that

$$|(P_t \varphi)(x) - (P_t \varphi)(x_n)| \leq \mathbb{E}\, |\varphi(X^x(t)) - \varphi(X^{x_n}(t))| \longrightarrow 0$$

as $n \to \infty$. $\qquad\square$

We shall assume from now on that the pair (A, B) satisfies the coercivity condition

$$\langle Ax, x \rangle \geq \frac{1}{2} \| B(x) \|^2_{\mathscr{L}^2(U, H)} + C \| x \|^2_V - C_0 \qquad \forall x \in V, \tag{5.1}$$

with $C_0 > 0$ a constant.

Theorem 5.2 *The set of invariant measures for the transition semigroup $(P_t)_{t \geq 0}$ is not empty.*

Proof Let (X, ξ) be the unique strong solution to Eq. 1.1. For every $t \geq 0$ one has, by Proposition 4.1,

$$\frac{1}{2} \| X(t) \|^2 + \int_0^t \langle AX(s), X(s) \rangle\, ds + \int_0^t \int_D \xi(s) X(s)\, ds$$
$$= \frac{1}{2} \| x \|^2 + \frac{1}{2} \int_0^t \| B(X(s)) \|^2_{\mathscr{L}^2(U, H)}\, ds + \int_0^t X(s) B(X(s))\, dW(s).$$

Let us show that the stochastic integral $M := XB(X) \cdot W$ on the right-hand side of this identity is a martingale. For this it suffices to show that $\mathbb{E}[M, M]_T^{1/2}$ is finite:

one has, by the ideal property of Hilbert-Schmidt operators and the Cauchy-Schwarz inequality,

$$\mathbb{E}[M, M]_T^{1/2} = \mathbb{E}\left(\int_0^T \|XB(X)\|_{\mathscr{L}^2(U,\mathbb{R})}^2 \, ds\right)^{1/2}$$

$$\leq \mathbb{E}\, \|X\|_{L^\infty(0,T;H)}\left(\int_0^T \|B(X)\|_{\mathscr{L}^2(U,H)}^2 \, ds\right)^{1/2}$$

$$\leq \left(\mathbb{E}\, \|X\|_{L^\infty(0,T;H)}^2\right)^{1/2}\left(\mathbb{E}\int_0^T \|B(X)\|_{\mathscr{L}^2(U,H)}^2 \, ds\right)^{1/2},$$

where the last term is finite thanks to Theorem 2.1 and the assumption of linear growth on B. Therefore, recalling that, for any $r, s \in \mathbb{R}$, $j(r) + j^*(s) = rs$ if and only if $s \in \beta(r)$, one has, taking the coercivity condition Eq. 5.1 into account,

$$C\mathbb{E}\int_0^t \|X(s)\|_V^2 \, ds + \mathbb{E}\int_0^t\int_D j(X(s)) \, ds + \mathbb{E}\int_0^t\int_D j^*(\xi(s)) \, ds \leq \frac{1}{2}\|x\|^2 + C_0 t \quad (5.2)$$

for all $t \geq 0$. Let $x = 0$. For any $t \geq 0$ the law of the random variable $X(t)$ is a probability measure on H, which we shall denote by π_t. We are now going to show that the family of measures $(\mu_t)_{t>0}$ on H defined by

$$\mu_t : E \longmapsto \frac{1}{t}\int_0^t \pi_s(E) \, ds$$

is tight. The ball B_n in V of radius $n \in \mathbb{N}$ is a compact subset of H, because the embedding $V \hookrightarrow H$ is compact. Moreover, Markov's inequality and Eq. 5.2 yield

$$\mu_t(B_n^c) = \frac{1}{t}\int_0^t \pi_s(B_n^c) \, ds = \frac{1}{t}\int_0^t \mathbb{P}\left(\|X(s)\|_V^2 > n^2\right) ds$$

$$\leq \frac{1}{tn^2}\int_0^t \mathbb{E}\, \|X(s)\|_V^2 \, ds \leq \frac{1}{Ctn^2} C_0 t = \frac{C_0}{Cn^2},$$

hence also

$$\sup_{t>0} \mu_t(B_n^c) \leq \frac{C_0}{Cn^2} \to 0 \quad \text{as } n \to \infty.$$

It follows by Prokhorov's theorem that there exists a probability measure μ on H and a sequence $(t_k)_{k\in\mathbb{N}}$ increasing to infinity such that μ_{t_k} converges to μ in the topology $\sigma(\mathscr{M}_1(H), C_b(H))$ as $k \to \infty$. Furthermore, μ is an invariant measure for the transition semigroup P, thanks to the Krylov-Bogoliubov theorem. $\qquad\square$

We are now going to prove integrability properties of *all* invariant measures, which in turn provide information on their support. We start with a (relatively) simple yet crucial estimate.

Proposition 5.3 *Let μ be an invariant measure for the transition semigroup (P_t). Then one has*

$$\int_H \|x\|^2 \, \mu(dx) \leq \frac{K^2 C_0}{C},$$

where K is the norm of the embedding $V \hookrightarrow H$.

Proof We are going to apply the Itô formula of Proposition 4.2 to the process X and the function $G_\delta : x \mapsto g_\delta(\|x\|^2)$, where $g_\delta \in C_b^2(\mathbb{R}_+)$ is defined as

$$g_\delta(r) = \frac{r}{1 + \delta r}, \qquad \delta > 0,$$

so that

$$g_\delta'(r) = \frac{1}{(1 + \delta r)^2}, \qquad g_\delta''(r) = -\frac{2\delta}{(1 + \delta r)^3},$$

whence

$$g_\delta(\|X(t)\|^2) + 2 \int_0^t g_\delta'(\|X(s)\|^2)(\langle AX(s), X(s)\rangle + \langle \xi(s), X(s)\rangle)\, ds$$

$$-2 \int_0^t g_\delta''(\|X(s)\|^2) \|X(s)B(X(s))\|_{\mathscr{L}^2(U,\mathbb{R})}^2\, ds$$

$$= g_\delta(\|x\|^2) + 2 \int_0^t g_\delta'(\|X(s)\|^2) X(s)B(X(s))\, dW(s)$$

$$+ \int_0^t g_\delta'(\|X(s)\|^2) \|B(X(s))\|_{\mathscr{L}^2(U,\mathbb{R})}^2\, ds.$$

Since $g_\delta' > 0$ and $g_\delta'' < 0$, the coercivity condition Eq. 5.1 and the monotonicity of β imply

$$g_\delta(\|X(t)\|^2) + 2 \int_0^t g_\delta'(\|X(s)\|^2)(C\|X(s)\|_V^2 - C_0)\, ds$$

$$\leq g_\delta(\|x\|^2) + 2 \int_0^t g_\delta'(\|X(s)\|^2) X(s)B(X(s))\, dW(s).$$

Taking into account that $|g_\delta'| \leq 1$, the stochastic integral is a martingale, exactly as in the proof of Theorem 5.2, hence has expectation zero, so that

$$\mathbb{E}\, G_\delta(X(t)) + 2C\, \mathbb{E} \int_0^t g_\delta'(\|X(s)\|^2)\|X(s)\|_V^2\, ds \leq G_\delta(x) + 2Ct.$$

By definition of (P_t) we have $P_t G_\delta(x) = \mathbb{E}\, G_\delta(X(t))$, from which it follows, by the boundedness of G_δ and by definition of invariant measure,

$$C \int_H \mathbb{E} \int_0^t g_\delta'(\|X(s)\|^2)\|X(s)\|_V^2\, ds\, \mu(dx) \leq C_0 t.$$

Denoting the norm of the embedding $V \hookrightarrow H$ by K, we get

$$\int_H \int_0^t \mathbb{E} \frac{\|X(s)\|^2}{(1 + \delta\|X(s)\|^2)^2}\, ds\, d\mu \leq \frac{K^2 C_0}{C} t.$$

Let $f_\delta : r \mapsto r/(1 + \delta r)^2$, $\delta > 0$, and $F_\delta := f_\delta \circ \|\cdot\|^2$. Then

$$\mathbb{E} \frac{\|X(s)\|^2}{(1 + \delta\|X(s)\|^2)^2} = P_s F_\delta,$$

hence, by Tonelli's theorem and invariance of μ,

$$\int_H \int_0^t \mathbb{E} \frac{\|X(s)\|^2}{(1 + \delta\|X(s)\|^2)^2}\, ds\, d\mu = \int_0^t \int_H P_s F_\delta\, d\mu\, ds = t \int_H F_\delta\, d\mu \leq \frac{K^2 C_0}{C} t.$$

Taking the limit as $\delta \to 0$, the monotone convergence theorem yields

$$\int_H \|x\|^2 \, \mu(dx) \le \frac{K^2 C_0}{C}.$$

$\square$

In order to state the next integrability results for invariant measures, we need to define the following subsets of H:

$$J := \left\{ u \in H : j(u) \in L^1(D) \right\},$$
$$J^* := \left\{ u \in H : \exists\, v \in L^1(D) : v \in \beta(u) \text{ a.e. in } D \text{ and } j^*(v) \in L^1(D) \right\},$$

whose Borel measurability will be proved in Lemma 5.5 below.

Theorem 5.4 *Let μ be an invariant measure for the transition semigroup P. Then one has*

$$\int_H \|u\|_V^2 \, \mu(du) + \int_H \int_D j(u) \, \mu(du) + \int_H \int_D j^*(\beta^0(u)) \, \mu(du) \le \frac{K^2 C_0}{2C} + C_0,$$

where K is the norm of the embedding $V \hookrightarrow H$. In particular, μ is concentrated on $V \cap J \cap J^*$.

Proof Let us introduce the functions $\Phi, \Psi, \Psi_* : H \to \mathbb{R}_+ \cup \{+\infty\}$ defined as

$$\Phi : u \longmapsto \|u\|_V^2 \, 1_V(u) + \infty \cdot 1_{H \setminus V}(u),$$
$$\Psi : u \longmapsto \left(\int_D j(u) \right) 1_J(u) + \infty \cdot 1_{H \setminus J}(u),$$
$$\Psi_* : u \longmapsto \left(\int_D j^*(\beta^0(u)) \right) 1_{J^*}(u) + \infty \cdot 1_{H \setminus J^*}(u),$$

as well as their approximations $\Phi_n, \Psi_n, \Psi_{*n} : H \to \mathbb{R}_+ \cup \{+\infty\}$, $n \in \mathbb{N}$, defined as (here $B_n(V)$ denotes the ball of radius n in V)

$$\Phi_n : u \longmapsto \begin{cases} \|u\|_V^2 & \text{if } u \in B_n(V), \\ n^2 & \text{if } u \in H \setminus B_n(V), \end{cases}$$

$$\Psi_n : u \longmapsto \begin{cases} \int_D j(u) & \text{if } \int_D j(u) \le n, \\ n & \text{otherwise,} \end{cases}$$

and

$$\Psi_{*n} : u \longmapsto \begin{cases} \int_D j^*(\beta_{1/n}(u)) & \text{if } \int_D j^*(\beta_{1/n}(u)) \le n, \\ n & \text{otherwise.} \end{cases}$$

One obviously has

$$\int_H \Phi_n \, d\mu = \int_0^1 \int_H \Phi_n \, d\mu \, ds,$$

as well as, by invariance of μ and boundedness of Φ_n,

$$\int_H \Phi_n \, d\mu = \int_H P_s \Phi_n \, d\mu,$$

thus also, by Tonelli's theorem ($\Phi_n \ge 0$ and P is positivity preserving, being Markovian)

$$\int_H \Phi_n \, d\mu = \int_0^1 \int_H P_s \Phi_n \, d\mu \, ds = \int_H \int_0^1 \mathbb{E}\, \Phi_n(X(s)) \, ds \, d\mu.$$

The same reasoning also yields

$$\int_H \Psi_n \, d\mu = \int_H \int_0^1 \mathbb{E} \, \Psi_n(X(s)) \, ds \, d\mu, \qquad \int_H \Psi_{*n} \, d\mu = \int_H \int_0^1 \mathbb{E} \, \Psi_{*n}(X(s)) \, ds \, d\mu,$$

with

$$\mathbb{E} \, \Phi_n(X(s)) = 7 \, \mathbb{E}\big(\|X(s)\|_V^2 \wedge n^2\big) \le \mathbb{E} \, \|X(s)\|_V^2,$$

$$\mathbb{E} \, \Psi_n(X(s)) = \mathbb{E}\Big(n \wedge \int_D j(X(s))\Big) \le \mathbb{E} \int_D j(X(s)),$$

$$\mathbb{E} \, \Psi_{*n}(X(s)) = \mathbb{E}\Big(n \wedge \int_D j^*(\beta_{1/n}(X(s)))\Big) \le \mathbb{E} \int_D j^*(\xi(s)),$$

where, in the last inequality, we have used the fact that for every $r \in D(\beta) = \mathbb{R}$ the sequence $\{\beta_\lambda(r)\}_\lambda$ converges from below to $\beta^0(r)$, where $\beta^0(r)$ is the unique element in $\beta(r)$ such that $|\beta^0(r)| \le |y|$ for every $y \in \beta(r)$ (note that the uniqueness of $\beta^0(r)$ follows from the maximal monotonicity of β). Thanks to estimate (5.2) we have, by Tonelli's theorem,

$$C \int_0^1 \big(\mathbb{E} \, \Phi_n(X(s)) + \mathbb{E} \, \Psi_n(X(s)) + \mathbb{E} \, \Psi_{*n}(X(s))\big) \, ds$$

$$\le C\mathbb{E} \int_0^1 \|X(s)\|_V^2 \, ds + \mathbb{E} \int_0^1 \int_D j(X(s)) \, ds + \mathbb{E} \int_0^1 \int_D j^*(\xi(s)) \, ds$$

$$\le \frac{1}{2}\|x\|^2 + C_0,$$

therefore, integrating with respect to μ and taking the previous proposition into account,

$$\int_H \big(C\Phi_n + \Psi_n + \Psi_{*n}\big) \, d\mu \le \frac{1}{2} \int_H \|x\|^2 \, \mu(dx) + C_0 \le \frac{K^2 C_0}{2C} + C_0$$

uniformly with respect to n. Since Φ_n and Ψ_n converge pointwise and monotonically from below to Φ and Ψ, respectively, the monotone convergence theorem yields

$$\int_H \Phi \, d\mu \le \frac{C_0(K^2 + 2C)}{2C^2}, \qquad \int_H \Psi \, d\mu \le \frac{C_0(K^2 + 2C)}{2C},$$

hence, in particular, $\mu(V) = \mu(J) = 1$. Similarly, note that $\beta_{1/n} \in \beta((I + (1/n)\beta)^{-1})$ and $0 \in \beta(0)$ imply that $|\beta_{1/n}|$ converges pointwise to $|\beta^0|$ monotonically from below as $n \to \infty$, hence the same holds for the convergence of $j^*(\beta_{1/n})$ to $j^*(\beta^0)$ because j^* is convex and continuous with $j^*(0) = 0$. Therefore Ψ_{*n} converges to Ψ pointwise monotonically from below as $n \to \infty$. We conclude, again by the monotone convergence theorem, that $\Psi_* \in L^1(H, \mu)$, thus also that $\mu(J^*) = 1$. $\qquad \square$

As mentioned above, the sets J and J^* are Borel measurable.

Lemma 5.5 *The sets*

$$J := \big\{u \in H : j(u) \in L^1(D)\big\},$$

$$J^* := \big\{u \in H : \exists \, v \in L^1(D) : v \in \beta(u) \text{ a.e. in } D \text{ and } j^*(v) \in L^1(D)\big\},$$

belong to the Borel σ-algebra of H.

Proof Setting, for every $n \in \mathbb{N}$,

$$J_n := \Big\{ u \in H : \int_D j(u) \leq n \Big\},$$

$$J_n^* := \Big\{ u \in H : \exists\, v \in L^1(D) : v \in \beta(u) \text{ a.e. in } D \text{ and } \int_D j^*(v) \leq n \Big\},$$

it is immediately seen that

$$J = \bigcup_{n=1}^{\infty} J_n \quad \text{and} \quad J^* = \bigcup_{n=1}^{\infty} J_n^*.$$

Moreover, the lower semicontinuity of convex integrals implies that J_n is closed in H for every n, hence Borel-measurable, so that $J \in \mathscr{B}(H)$. Let us show that, similarly, J_n^* is also closed in H for every $n \in \mathbb{N}$: if $(u_k)_k \subset J_n^*$ and $u_k \to u$ in H, then for every k there exists $v_k \in L^1(D)$ with $v_k \in \beta(u_k)$ and

$$\int_D j^*(v_k) \leq n \qquad \forall k \in \mathbb{N}.$$

Since j^* is superlinear at infinity, this implies that the family $(v_k)_k$ is uniformly integrable in D, hence by the Dunford-Pettis theorem also weakly relatively compact in $L^1(D)$. Consequently, there is a subsequence $(v_{k_i})_i$ and $v \in L^1(D)$ such that $v_{k_i} \to v$ weakly in $L^1(D)$. The weak lower semicontinuity of convex integrals easily implies that

$$\int_D j^*(v) \leq \liminf_{i \to \infty} \int_D j^*(v_{k_i}) \leq n.$$

Let us show that $v \in \beta(u)$ almost everywhere in D: by definition of subdifferential, for every $k \in \mathbb{N}$ and for every measurable set $E \subseteq D$ we have

$$\int_E j(u_k) + \int_E v_k(z - u_k) \leq \int_E j(z) \qquad \forall z \in L^{\infty}(D).$$

By Egorov's theorem, for any $\varepsilon > 0$ there exists a measurable set $E_\varepsilon \subseteq D$ with $|E_\varepsilon^c| \leq \varepsilon$ and $u_k \to u$ uniformly in E_ε. Taking $E = E_\varepsilon$ in the last inequality, letting $k \to \infty$ we get

$$\int_{E_\varepsilon} j(u) + \int_{E_\varepsilon} v(z - u) \leq \int_{E_\varepsilon} j(z) \quad \forall z \in L^{\infty}(D),$$

which in turn implies by a classical localization argument that

$$j(u) + v(z - u) \leq j(z) \quad \text{a.e. in } E_\varepsilon, \qquad \forall z \in \mathbb{R}.$$

Hence, by the arbitrariness of ε, $v \in \beta(u)$ almost everywhere in D, thus also $u \in J_n^*$. This implies that J_n^* is closed in H for every n, therefore also that $J^* \in \mathscr{B}(H)$. $\qquad \square$

The estimates proved above implies that the set of ergodic invariant measures is not empty.

Theorem 5.6 *There exists an ergodic invariant measure for the transition semigroup (P_t).*

Proof Recall that, as it follows by the Krein-Milman theorem, for a Markovian transition semigroup the set of ergodic invariant measures coincides with the extreme points of the set of all invariant measures (see, e.g., [1, Thm. 19.25]). Let $\mathscr{I}$ be the set of all invariant measures for P: by Theorem 5.2, we know that $\mathscr{I}$ is not empty and we need to show that

$\mathscr{I}$ admits at least an extreme point. Let us prove that $\mathscr{I}$ is tight. By Theorem 5.4, we know that there exists a constant N such that

$$\int_H \|x\|_V^2 \, \mu(dx) \leq N \qquad \forall \mu \in \mathscr{I}.$$

Therefore, using the notation of the proof of Theorem 5.2, by Markov inequality

$$\sup_{\mu \in \mathscr{I}} \mu(B_n^c) = \sup_{\mu \in \mathscr{I}} \mu\big(\{x \in H : \|x\|_V > n\}\big) \leq \frac{1}{n^2} \sup_{\mu \in \mathscr{I}} \int_H \|x\|_V^2 \mu(dx) \leq \frac{N}{n^2} \to 0$$

as $n \to \infty$. Hence $\mathscr{I}$ is tight, and thus admits extreme points. $\qquad\square$

Under a mild superlinearity condition one can also obtain uniqueness.

Theorem 5.7 *If there exist $c > 0$ and $\delta > 0$ such that*

$$\langle A(v_1 - v_2), v_1 - v_2 \rangle + \int_D (\eta_1 - \eta_2)(v_1 - v_2) \, dx - \frac{1}{2}\|B(v_1) - B(v_2)\|_{\mathscr{L}^2(U,H)}^2$$
$$\geq \frac{c}{2}\|v_1 - v_2\|^{2+\delta}$$

for all $v_1, v_2 \in V$ with $(v_i, \eta_i) \in \beta$ a.e. in D, then there exists a unique invariant measure μ for the transition semigroup P. Moreover, μ is strongly mixing.

Proof By Proposition 4.1 applied to the difference $X^x - X^y$ (see the proof of [17, Prop 6.2] for further details) one has, with obvious meaning of the notation,

$$\|X^x(t) - X^y(t)\|^2 + 2\int_0^t \langle A(X^x - X^y), X^x - X^y \rangle(s) \, ds$$
$$+2\int_0^t \int_D (\xi^x - \xi^y)(X^x - X^y)(x,s) \, dx \, ds$$
$$= \|x - y\|^2 + \int_0^t \|B(X^x(s)) - B(X^y(s))\|_{\mathscr{L}^2(U,H)}^2 \, ds + M_t,$$

where M is the local martingale defined as

$$M_t := \int_0^t \big(X^x(s) - X^y(s)\big)\big(B(X^x(s)) - B(X^y(s))\big) \, dW(s).$$

By an argument entirely analogous to the one used in the proof of Theorem 5.2 it follows that M is in fact a martingale. Taking expectation on the above identity, the superlinearity assumption and Jensen's inequality yield

$$\mathbb{E}\|X^x(t) - X^y(t)\|^2 + c\int_0^t \big(\mathbb{E}\|X^x(s) - X^y(s)\|^2\big)^{1+\frac{\delta}{2}} \, ds \leq \|x - y\|^2$$

Denoting the solution to the Cauchy problem

$$\phi' + \phi^{1+\frac{\delta}{2}} = 0, \qquad \phi(0) = \phi_0 \geq 0,$$

by ϕ^{ϕ_0}, it is easily seen that

$$c(t) := \sup_{\phi_0 \geq 0} \phi^{\phi_0}(t)$$

is positive for every $t \geq 0$ and tends to zero as $t \to \infty$. By an elementary comparison argument it follows that

$$\mathbb{E}\|X^x(t) - X^y(t)\|^2 \leq c(t) \qquad \forall t \geq 0.$$

Let μ be an invariant measure for P. For any $\varphi \in C_b^1(H)$ and $x \in H$ we have

$$\left| P_t\varphi(x) - \int_H \varphi(y)\,\mu(dy) \right|^2 \leq \|D\varphi\|_\infty^2 \int_H \mathbb{E}\,\|X^x(t) - X^y(t)\|^2\,\mu(dy) \leq \|D\varphi\|_\infty^2 c(t)$$

uniformly in x, and since $C_b^1(H)$ is dense in $L^2(H, \mu)$, we deduce that

$$\left| P_t\varphi(x) - \int_H \varphi(y)\,\mu(dy) \right| \longrightarrow 0$$

as $t \to \infty$ for every $\varphi \in L^2(H, \mu)$. We have thus shown that P admits a unique invariant measure, which is strongly mixing as well. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 5.8 Let κ denote the norm in $\mathscr{L}(H)$ of the embedding $V \hookrightarrow H$. The superlinearity assumption of Theorem 5.7 is satisfied if $C/\kappa^2 \geq L_B/2$ and β satisfies the superlinearity condition

$$(y_i - y_2)(x_1 - x_2) \geq \frac{c}{2}|x_1 - x_2|^{2+\delta}$$

for all $(x_i, y_i) \in \beta$, $i = 1, 2$. In fact, the latter inequality obviously implies

$$\int_D (\eta_1 - \eta_2)(v_1 - v_2)\,dx \geq \frac{c}{2}\|v_1 - v_2\|^{2+\delta}$$

for every $(v_i, \eta_i) \in \beta$ a.e. in D. Moreover, by assumptions (i) and (vii), one has

$$\langle A(v_1 - v_2), v_1 - v_2 \rangle - \frac{1}{2}\|B(v_1) - B(v_2)\|_{\mathscr{L}^2(U,H)}^2$$

$$\geq C\|v_1 - v_2\|_V^2 - \frac{L_B}{2}\|v_1 - v_2\|^2 \geq (C/\kappa^2 - L_B/2)\|v_1 - v_2\|^2 \geq 0$$

for every $v_1, v_2 \in V$.

6 The Kolmogorov Equation

Throughout this section we shall assume that β is a function, rather than just a graph.

Let $P = (P_t)_{t\geq 0}$ be the Markovian semigroup on $B_b(H)$ generated by the unique solution to Eq. 1.1, as in the previous section, and μ be an invariant measure for P. Then P extends to a strongly continuous linear semigroup of contractions on $L^p(H, \mu)$ for every $p \geq 1$. These extensions will all be denoted by the same symbol. Let $-L$ be the infinitesimal generator in $L^1(H, \mu)$ of P, and $-L_0$ be Kolmogorov operator formally associated to Eq. 1.1, i.e.

$$[L_0 f](x) = -\frac{1}{2}\mathrm{Tr}\big(D^2 f(x)B(x)B^*(x)\big) + \langle Ax, Df(x)\rangle + \langle \beta(x), Df(x)\rangle, \qquad x \in V \cap J^*,$$

where f belongs to a class of sufficiently regular functions introduced below. Our aim is to characterize the "abstract" operator L as the closure of the "concrete" operator L_0. Even though this will be achieved only in the case of additive noise, some intermediate results will be proved in the more general case of multiplicative noise.

Let us first show that L_0 is a proper linear (unbounded) operator on $L^1(H, \mu)$ with domain

$$\mathsf{D}(L_0) := \big\{ f : V \to \mathbb{R} : f \in C_b^1(V') \cap C_b^2(H) \cap C_b^1(L^1(D)) \big\}.$$

Here $f \in C_b^1(V')$ means that, for any $x \in V$ and $v' \in V'$, $|Df(x)v'| \leq N\|v'\|_{V'}$, with the constant N independent of x and v', and that $x \mapsto Df(x) \in C(V', V)$. Analogously, $f \in$

$C_b^1(L^1(D))$ means that, for any $x \in V$ and $k \in L^1(D)$, there is a constant N independent of x and k such that $|Df(x)k| \le N\|k\|_{L^1(D)}$ and $x \mapsto Df(x) \in C(L^1(D), L^\infty(D))$. For any $f \in C_b^2(H)$, one has, recalling the linear growth condition on B,

$$\mathrm{Tr}\big(D^2 f(x) B(x) B^*(x)\big) \lesssim \|B(x)\|_{\mathscr{L}^2(U,H)}^2 \lesssim 1 + \|x\|^2,$$

and $\|\cdot\|^2 \in L^1(H, \mu)$. Moreover, since $A \in \mathscr{L}(V, V')$, one has $\|Ax\|_{V'} \lesssim \|x\|_V$, so that, for any $f \in C_b^1(V')$,

$$|\langle Ax, Df(x) \rangle| \le \|Ax\|_{V'} \sup_{x \in V} \|Df(x)\|_V \lesssim \|x\|_V,$$

hence $x \mapsto \langle Ax, Df(x) \rangle \in L^1(H, \mu)$ because $\|\cdot\|_V^2 \in L^1(H, \mu)$. Similarly, writing

$$|\langle \beta(x), Df(x) \rangle| \le \|j^*(\beta(x))\|_{L^1(D)} + \|j(Df(x))\|_{L^1(D)}$$

and recalling that $x \mapsto \|j^*(\beta(x))\|_{L^1(D)} \in L^1(H, \mu)$ by Theorem 5.4, it is enough to consider the second term on the right-hand side: for any $f \in C_b^1(L^1(D))$, $\sup_{x \in V} \|Df(x)\|_{L^\infty(D)}$ is finite, hence, recalling that $j \in C(\mathbb{R})$, $j(Df(x))$ is bounded pointwise in D, thus also in $L^1(D)$, uniformly over $x \in V$. In particular, $x \mapsto \|j(Df(x))\|_{L^1(D)} \in L^1(H, \mu)$.

Let us now show that the infinitesimal generator $-L$ restricted to $\mathsf{D}(L_0)$ coincides with the operator $-L_0$ defined above. Indeed, by Proposition 4.3, for every $g \in \mathsf{D}(L_0)$ we have

$$g(X^x(t)) + \int_0^t \langle AX^x(s), Dg(X^x(s)) \rangle \, ds + \int_0^t \langle \beta(X^x(s)), Dg(X^x(s)) \rangle \, ds$$

$$= g(x) + \frac{1}{2} \int_0^t \mathrm{Tr}[B^*(X^x(s)) D^2 g(X^x(s)) B(X^x(s))] \, ds$$

$$+ \int_0^t Dg(X^x(s)) B(X^x(s)) \, dW(s),$$

from which we infer, taking expectations and applying Fubini's theorem,

$$\frac{P_t g(x) - g(x)}{t} = -\frac{1}{t} \int_0^t P_s L_0 g(x) \, ds \qquad \forall x \in V \cap J^*.$$

Since $g \in \mathsf{D}(L_0)$, we have that $L_0 g \in L^1(H, \mu)$, as proved above. Therefore, recalling that P is strongly continuous on $L^1(H, \mu)$, we have that $t \mapsto P_t L_0 g$ is continuous from $[0, T]$ to $L^1(H, \mu)$. Hence, letting $t \to 0$, we have

$$\frac{P_t g - g}{t} \to -L_0 g \qquad \text{in } L^1(H, \mu),$$

which implies that $L = L_0$ on $\mathsf{D}(L_0)$.

We are now going to construct a regularization of the operator L_0. For any $\lambda \in (0, 1)$, let

$$\beta_\lambda : \mathbb{R} \to \mathbb{R}, \qquad \beta_\lambda := \frac{1}{\lambda}\big(I - (I + \lambda\beta)^{-1}\big),$$

be the Yosida approximation of β. Denoting a sequence of mollifiers on $\mathbb{R}$ by (ρ_n), the function $\beta_{\lambda n} := \beta_\lambda * \rho_n$ is monotone and infinitely differentiable with all derivatives bounded. Let us consider the regularized equation

$$dX_{\lambda n} + AX_{\lambda n} \, dt + \beta_{\lambda n}(X_{\lambda n}) \, dt = B(X_{\lambda n}) \, dW(t), \qquad X_{\lambda n}(0) = x. \qquad (6.1)$$

Since $\beta_{\lambda n}$ is Lipschitz-continuous, Eq. 6.1 admits a unique strong (variational) solution $X_{\lambda n}^x \in L^2(\Omega; E)$, where, as before, $E := C([0, T]; H) \cap L^2(0, T; V)$. Furthermore, the

generator of the Markovian transition semigroup $P^{\lambda n} = (P_t^{\lambda n})_{t \geq 0}$ on $B_b(H)$ defined by $P_t^{\lambda n} f(x) := \mathbb{E} f(X_{\lambda n}^x(t))$, restricted to $C_b^1(V') \cap C_b^2(H)$, is given by $-L_0^{\lambda n}$, where

$$[L_0^{\lambda n} f](x) = -\frac{1}{2} \operatorname{Tr}\big(D^2 f(x)B(x)B^*(x)\big) + \langle AxDf(x)\rangle + \langle \beta_{\lambda n}(x)Df(x)\rangle, \qquad x \in V.$$

This follows arguing as in the case of L_0 (even using the simpler Itô formula of Proposition 4.2, rather than the one of Proposition 4.3).

Let us now consider the stationary Kolmogorov equation

$$\alpha v + L_0^{\lambda n} v = g, \qquad g \in \mathsf{D}(L_0), \quad \alpha > 0. \tag{6.2}$$

In view of the well-known relation between (Markovian) resolvents and transition semigroups, one is led to considering the function

$$v_{\lambda n}(x) := \mathbb{E} \int_0^\infty e^{-\alpha t} g(X_{\lambda n}^x(t))\, dt,$$

which is the natural candidate to solve (6.2). If we show that $v_{\lambda n} \in C_b^1(V') \cap C_b^2(H)$, then an application of Itô's formula (in the version of Proposition 4.2) shows that indeed $v_{\lambda n}$ solves (6.2).

We are going to obtain regularity properties of $v_{\lambda n}$ via pathwise differentiability of the solution map $x \mapsto X_{\lambda n}$ of the regularized stochastic Eq. 6.1.

From now on we shall work under the following further assumptions:

(viii)   the noise is of additive type, i.e. $B \in \mathscr{L}^2(U, H)$ is non-random and time-independent;

(ix) V is continuously embedded in $L^4(D)$

These are needed in order to apply the second-order differentiability results of Section 4.2. In particular, assumptions (viii) and (ix) imply, thanks to Theorems 4.4 and 4.7, that the solution map $x \mapsto X_{\lambda n} : H \to L^2(\Omega; E)$ is Lipschitz continuous and twice Fréchet differentiable. Moreover, denoting its first order Fréchet differential by

$$DX_{\lambda n} : H \to \mathscr{L}(H, L^2(\Omega; E)),$$

for any $h \in H$ the process $Y_h := (DX_{\lambda n})h \in L^2(\Omega; E)$ satisfies the linear deterministic equation with random coefficients

$$Y_h'(t) + AY_h(t) + \beta_{\lambda n}'(X_{\lambda n}(t))Y_h(t) = 0, \qquad Y_h(0) = h. \tag{6.3}$$

Similarly, denoting the second order Fréchet differential of $x \mapsto X_{\lambda n}$ by

$$D^2 X_{\lambda n} : H \to \mathscr{L}_2(H; L^2(\Omega; E)),$$

for any $h, k \in H$ the process $Z_{hk} := D^2 X_{\lambda n}(h, k) \in L^2(\Omega; E))$ satisfies the linear deterministic equation with random coefficients

$$Z_{hk}'(t) + AZ_{hk}(t) + \beta_{\lambda n}'(X_{\lambda n}(t))Z_{hk}(t) + \beta_{\lambda n}''(X_{\lambda n}(t))Y_h(t)Y_k(t) = 0, \qquad Z_{hk}(0) = 0. \tag{6.4}$$

We shall need the following result on the connection between variational and mild solutions in the deterministic setting. We recall that A_2 denotes the part of A on H.

Lemma 6.1 *Let $F : [0, T] \times H \to H$ be Lipschitz continuous in the second variable, uniformly with respect to the first, with $F(\cdot, 0) = 0$, and $u_0 \in H$. If $u \in C([0, T]; H) \cap$*

$L^2(0, T; V)$ and $v \in C([0, T]; H)$ are the (unique) variational and mild solution to the problems

$$u' + Au = F(\cdot, u), \quad u(0) = u_0, \qquad \text{and} \qquad v' + A_2 v = F(\cdot, v), \quad v(0) = u_0,$$

respectively, then $u = v$.

Proof Let us first assume that $u' + Au = f$ and $v' + A_2 v = f$, where $f \in L^2(0, T; H)$. Then we have

$$u(t) + \int_0^t Au(s)\, ds = u_0 + \int_0^t f(s)\, ds,$$

$$v(t) = S(t)u_0 + \int_0^t S(t-s) f(s)\, ds$$

for all $t \in [0, T]$, where S is the the semigroup generated on H by $-A_2$. Let us show that $u = v$. For $m \in \mathbb{N}$, applying $(I + \varepsilon A_2)^{-m}$ to the second equation we have (with obvious meaning of notation)

$$v'_\varepsilon + A_2 v_\varepsilon = f_\varepsilon, \qquad v_\varepsilon(0) = u_0^\varepsilon$$

in the strong sense, since $v_\varepsilon \in C([0, T]; D(A_2^m))$. In particular, v_ε is also a variational solution of the equation

$$v'_\varepsilon + A v_\varepsilon = f_\varepsilon, \qquad v_\varepsilon(0) = u_0^\varepsilon.$$

By construction we have that $v_\varepsilon \to v$ in $C([0, T]; H)$; moreover, since $f_\varepsilon \to f$ in $L^2(0, T; H)$ and $u_0^\varepsilon \to u_0$ in H, arguing as in the proof of Lemma 4.5 we have also that $v_\varepsilon \to u$ in $C([0, T]; H) \cap L^2(0, T; V)$. Since mild and variational solutions are unique, we conclude that $u = v$. We shall now extend this argument to the case where u and v are the unique variational and mild solutions to the equations

$$u' + Au = F(\cdot, u), \qquad v' + A_2 v = F(\cdot, v), \qquad u(0) = v(0) = u_0,$$

respectively. Setting $f := F(\cdot, v)$, the assumptions on F imply that $f \in L^2(0, T; H)$, hence v is a mild solution to $v' + A_2 v = f$, $v(0) = u_0$. It then follows by the previous argument that v is also the unique variational solution to $v' + Av = f$, $v(0) = u_0$. Therefore

$$u' + Au = F(\cdot, u), \qquad v' + Av = F(\cdot, v), \qquad u(0) = v(0) = u_0$$

in the variational sense. Using the integration-by-parts formula, the Lipschitz continuity of F, and Gronwall's inequality, it is then a standard matter to show that $u = v$. $\quad\square$

The following estimates are crucial. Recall that hypotheses (viii) and (ix) are in force.

Proposition 6.2 *One has, for every $x, h, k \in H$ and $t > 0$,*

$$\|Y_h^x\|_{C([0,t]; H) \cap L^2(0,t; V)} \lesssim \|h\|,$$

$$\|Z_{hk}^x\|_{C([0,t]; H) \cap L^2(0,t; V)} \lesssim_{\lambda, n} \|h\| \|k\|,$$

$$\|Y_h^x\|_{C([0,t]; L^1(D))} \leq \|h\|_{L^1(D)}.$$

Regarding A as an unbounded operator on V', assume that there exists $\delta \in (0, 1)$ and $\eta > 0$ such that $H = D((\eta I + A)^\delta)$. Then

$$\|Y_h^x(t)\|_H \lesssim (1 \vee t^{-\delta}) \|h\|_{V'}.$$

Proof Let $\Omega' \subseteq \Omega$ with $\mathbb{P}(\Omega') = 1$ be such that Eq. 6.3 holds true for all $t \in [0, T]$ and all $\omega \in \Omega'$. Let $\omega \in \Omega'$ be fixed. Recalling that A is coercive and that $\beta'_{\lambda n}$ is positive because $\beta_{\lambda n}$ is increasing, taking the scalar product with $Y_h(t)$ in Eq. 6.3 and integrating in time yields

$$\frac{1}{2}\|Y_h^x(t)\|^2 + C\int_0^t \|Y_h^x(s)\|_V^2 \, ds \le \frac{1}{2}\|h\|^2$$

for all $t \in [0, T]$, and the first estimate is thus proved. The second estimate follows directly from Proposition 4.7. Furthermore, denoting the Yosida approximation of the part of A in H by A_ε, let $Y_{h\varepsilon}^x \in C([0, T]; H)$ be the unique strong solution to the equation

$$Y'_{h\varepsilon}(t) + A_\varepsilon Y_{h\varepsilon}(t) + \beta'_{\lambda n}(X_{\lambda n}(t))Y_{h\varepsilon}(t) = 0, \qquad Y_{h\varepsilon}(0) = h.$$

Let (σ_k) be a sequence of smooth increasing functions approximating pointwise the (maximal monotone) signum graph, and $\widehat{\sigma}_k$ be the primitive of σ_k with $\widehat{\sigma}_k(0) = 0$. Taking the scalar product of the previous equation with $\sigma_k(Y_{h\varepsilon}^x)$ and integrating in time we get, for every $t > 0$,

$$\int_D \widehat{\sigma}_k(Y_{h\varepsilon}^x(t)) + \int_0^t \langle A_\varepsilon Y_{h\varepsilon}^x(s)\sigma_k(Y_{h\varepsilon}^x(s))\rangle \, ds + \int_0^t \int_D \beta'_{\lambda n}(X_{\lambda n}(s))\sigma_k(Y_{h\varepsilon}^x(s))Y_{h\varepsilon}^x(s) \, ds$$
$$\le \int_D \widehat{\sigma}_k(h).$$

Since, as $k \to \infty$, $\sigma_k(Y_{h\varepsilon}^x)$ converges a.e. to a measurable function $w_\varepsilon \in \text{sgn}(Y_{h\varepsilon}^x)$ and $\widehat{\sigma} \to |\cdot|$, letting $k \to \infty$ we get, for every $t \ge 0$,

$$\|Y_{h\varepsilon}^x(t)\|_{L^1(D)} + \int_0^t \langle A_\varepsilon Y_{h\varepsilon}^x(s)w_\varepsilon(s))\rangle \, ds \le \|h\|_{L^1(D)} \qquad \forall t \in [0, T].$$

Recalling that A_2 extends to an m-accretive operator on $L^1(D)$, the second term on the left-hand side is non-negative, and taking into account that $Y_{h\varepsilon}^x \to Y_h^x$ in $C([0, T]; H)$ as $\varepsilon \to 0$, the third inequality follows. Finally, since Y_h is the unique variational solution to Eq. 6.3, by Lemma 6.1 we have that Y_h is also mild solution to the same equation, i.e.

$$Y_h^x(t) = S(t)h - \int_0^t S(t-s)\beta'_{\lambda n}(X^x(s))Y_h^x(s) \, ds \qquad \forall t \in [0, T], \quad \mathbb{P}\text{-a.s.}$$

Recall that $-A$ generates an analytic semigroup on V' extending S, denoted by the same symbol. Since $H = \mathsf{D}((\eta I + A)^\delta)$, we have $\|S(t)h\| \lesssim t^{-\delta}\|h\|_{V'}$ for every $t > 0$. By the contractivity of S in H we also have, for every $t > 0$,

$$\|Y_h^x(t)\| \lesssim t^{-\delta}\|h\|_{V'} + \|\beta'_{\lambda n}\|_\infty \int_0^t \|Y_h^x(s)\| \, ds$$

from which Gronwall's inequality implies

$$\|Y_h^x(t)\| \lesssim t^{-\delta}\|h\|_{V'} + \|\beta'_{\lambda n}\|_\infty \int_0^t s^{-\delta}e^{\|\beta'_{\lambda n}\|_\infty(t-s)}\|h\|_{V'} \, ds.$$

Therefore we have, for every $t \in [0, 1]$,

$$\|Y_h^x(t)\| \lesssim t^{-\delta}\|h\|_{V'} + \|\beta'_{\lambda n}\|_\infty e^{\|\beta'_{\lambda n}\|_\infty}\|h\|_{V'}\int_0^1 s^{-\delta} \, ds = \left(t^{-\delta} + \frac{1+\delta}{1+\delta}\right)\|h\|_{V'}$$
$$\lesssim (1 + t^{-\delta})\|h\|_{V'}$$

as well as, for every $t \geq 1$,

$$\|Y_h^x(t)\| \leq \|Y_h^x(1)\| \lesssim 1^{-\delta} \|h\|_{V'} = \|h\|_{V'},$$

which implies the last estimate. □

Lemma 6.3 *Let $\alpha > 0$ and $g \in C_b^1(V') \cap C_b^2(H) \cap C_b^1(L^1(D))$. For every $n \in \mathbb{N}$ and $\lambda \in (0, 1)$, the function $v_{\lambda n} : H \to \mathbb{R}$ defined as*

$$v_{\lambda n}(x) := \mathbb{E} \int_0^{+\infty} e^{-\alpha t} g(X_{\lambda n}^x(t)) \, dt$$

belongs to $\mathsf{D}(L_0)$ and solves (6.2). Moreover, there exists a positive constant M, independent of λ and n, such that

$$\|v_{\lambda n}\|_{C_b^1(H) \cap C_b^1(L^1(D))} \leq M \tag{6.5}$$

for all $n \in \mathbb{N}$ and $\lambda \in (0, 1)$.

Proof Since $g \in C_b^1(H)$, for any $h \in H$ we have, by the first estimate of Proposition 6.2,

$$D\left(g(X_{\lambda n}^x(t))\right) h = Dg(X_{\lambda n}^x(t)) DX_{\lambda n}^x(t) h = Dg(X_{\lambda n}^x(t)) Y_h^x(t)$$
$$\leq \|Dg\|_{C(H;H)} \|Y_h^x\|_{C([0,T];H)} \leq \|Dg\|_{C(H;H)} \|h\|,$$

hence, by the dominated convergence theorem, $v_{\lambda n} \in C_b^1(H)$ and

$$Dv_{\lambda n}(x) h = \mathbb{E} \int_0^{+\infty} e^{-\alpha t} Dg(X_{\lambda n}^x(t)) Y_h^x(t) \, dt. \tag{6.6}$$

The uniform boundedness of $\|v_{\lambda n}\|_{C_b^1(H)}$ in λ and n follows directly from these computations. Similarly, using the fact that $g \in C_b^2(H)$ and the second estimate of Proposition 6.2, we have, for every $k \in H$,

$$D(D(g(X_{\lambda n}^x(t)) h) k = D^2 g(X_{\lambda n}^x(t))(Y_h^x(t), Y_k^x(t)) + Dg(X_{\lambda n}^x(t)) Z_{hk}^x(t)$$
$$\leq \|D^2 g\|_{C(H;\mathscr{L}_2(H;\mathbb{R}))} \|Y_h^x\|_{C([0,T];H)} \|Y_k^x\|_{C([0,T];H)}$$
$$+ \|Dg\|_{C(H,H)} \|Z_{hk}^x\|_{C([0,T];H)}$$
$$\lesssim_{\lambda,n} \|g\|_{C_b^2} \|h\| \|k\|,$$

hence, by the dominated convergence theorem, $v_{\lambda n} \in C_b^2(H)$ and

$$D^2 v_{\lambda n}(x)(h, k) = \mathbb{E} \int_0^{+\infty} e^{-\alpha t} \left(D^2 g(X_{\lambda n}^x(t)) Y_h^x(t) Y_k^x(t) + Dg(X_{\lambda n}^x(t)) Z_{hk}^x(t)\right) dt. \tag{6.7}$$

Moreover, using the third estimate of Proposition 6.2 and the fact that $g \in C_b^1(L^1(D))$, it follows by Hölder's inequality and Eq. 6.6 that

$$Dv_{\lambda n}(x) h \leq \mathbb{E} \int_0^{+\infty} e^{-\alpha t} \|Dg\|_{C(H;L^\infty(D))} \|Y_h^x(t)\|_{L^1(D)} \, dt \leq \frac{1}{\alpha} \|Dg\|_{C(H;L^\infty(D))} \|h\|_{L^1(D)},$$

which implies that $v_{\lambda n} \in C_b^1(L^1(D))$ and the estimate (6.5). Finally, by the last estimate of Proposition 6.2 and the fact that $g \in C_b^1(V')$, we have

$$Dv_{\lambda n}(x) h \leq \mathbb{E} \int_0^{+\infty} e^{-\alpha t} \|Dg\|_{C(H;V)} \|Y_h^x(t)\|_{V'} \, dt$$
$$\lesssim \|Dg\|_{C(H;V)} \|h\|_{V'} \int_0^{+\infty} (1 \vee t^{-\delta}) e^{-\alpha t} \, dt.$$

Since $t \mapsto (1 \vee t^{-\delta})e^{-\alpha t}$ belongs to $L^1(0, +\infty)$, we have

$$Dv_{\lambda n}(x)h \lesssim_{\lambda,n} \|h\|_{V'},$$

thus also $v_{\lambda n} \in C_b^1(V')$.

Let us show now that $v_{\lambda n}$ solves (6.2). Indeed, since $g \in C_b^2(H) \cap C_b^1(V')$, by Itô's formula in the version of Proposition 4.2 we get

$$g(X_{\lambda n}^x(t)) + \int_0^t \langle AX_{\lambda n}^x(s)Dg(X_{\lambda n}^x(s))\rangle \, ds + \int_0^t \langle \beta_{\lambda n}(X_{\lambda n}^x(s))Dg(X_{\lambda n}^x(s))\rangle \, ds$$

$$= g(x) + \frac{1}{2}\int_0^t \mathrm{Tr}[B^*(X_{\lambda n}^x(s))D^2g(X_{\lambda n}^x(s))B(X_{\lambda n}^x(s))] \, ds$$

$$+ \int_0^t Dg(X_{\lambda n}^x(s))B(X_{\lambda n}^x(s)) \, dW(s)$$

for every $t > 0$. Thanks to the boundedness of Dg, taking expectations and using Fubini's theorem we deduce that, for every $\alpha > 0$ and $x \in V$,

$$e^{-\alpha t}\,\mathbb{E}\,g(X_{\lambda n}^x(t)) + \alpha\,\mathbb{E}\int_0^t e^{-\alpha s}g(X_{\lambda n}^x(s)) \, ds - \int_0^t P_s^{\lambda n}L_0^{\lambda n}g(x) \, ds = g(x).$$

Since $g \in C_b(H)$, it is clear that, as $t \to +\infty$, the first and second term on the left-hand side converge to zero and $\alpha v_{\lambda n}(x)$, respectively, hence, by difference, we deduce that

$$\int_0^t P_s^{\lambda n}L_0^{\lambda n}g(x) \to \int_0^{+\infty} P_s^{\lambda n}L_0^{\lambda n}g(x) \, ds.$$

Letting then $t \to +\infty$ we infer that

$$\alpha v_{\lambda n}(x) - \int_0^{+\infty} e^{-\alpha t}P_t^{\lambda n}L_0^{\lambda n}g(x) \, dt = g(x),$$

hence

$$\alpha v_{\lambda n}(x) - L_0^{\lambda n}v_{\lambda n}(x) = g(x) \qquad \forall x \in V. \qquad \square$$

Lemma 6.4 *One has*

$$\lim_{\lambda \to 0}\lim_{n \to \infty}\left\|L_0 v_{\lambda n} - L_0^{\lambda n}v_{\lambda n}\right\|_{L^1(H,\mu)} = 0.$$

Proof By definition of L_0 and $L_0^{\lambda n}$, the claim amounts to showing that

$$\lim_{\lambda \to 0}\lim_{n \to \infty}\int_H |\langle \beta_{\lambda n}(x) - \beta(x)Dv(x)\rangle| \, \mu(dx) \to 0.$$

Since $\beta_{\lambda n}$ is Lipschitz-continuous with Lipschitz constant bounded by $1/\lambda$ for every $n \in \mathbb{N}$, we have

$$|\langle \beta_\lambda(x) - \beta_{\lambda n}(x)Dv(x)\rangle| \lesssim \frac{1}{\lambda}\|x\|,$$

so that, recalling that $\|\cdot\| \in L^2(H, \mu)$ and $\beta_{\lambda n} \to \beta_\lambda$ pointwise as $n \to \infty$, the dominated convergence theorem yields

$$\lim_{n \to \infty}\int_H |\langle \beta_\lambda(x) - \beta_{\lambda n}(x)Dv(x)\rangle| \, \mu(dx) = 0.$$

Since $Dv_{\lambda n}(x)$ is bounded in $L^\infty(D)$ uniformly over λ, n and x by estimate (6.5), one has

$$|(\beta(x) - \beta_\lambda(x))\,Dv(x)| \lesssim |\beta(x) - \beta_\lambda(x)|,$$

hence

$$(\beta(x) - \beta_\lambda(x)) \, Dv_{\lambda n}(x) \to 0$$

in $L^0(D)$ as $\lambda \to 0$ for every $x \in V$.

Recalling the definition of η in Section 2, we deduce that $j^*(\eta|\beta(x)|) \in L^1(D)$. Appealing to Young's inequality in the form

$$a|b| \le j(a) + j^*(|b|) \qquad \forall a, b \in \mathbb{R},$$

we have

$$\eta|\beta(x)| + \eta|\beta_\lambda(x)| \le 2j(1) + j^*(\eta|\beta(x)|) + j^*(\eta|\beta_\lambda(x)|)$$

hence also, since j^* is increasing on $\mathbb{R}_+$ and $|\beta_\lambda| \le |\beta|$,

$$|(\beta(x) - \beta_\lambda(x)) \, Dv_{\lambda n}(x)| \lesssim j(1) + j^*(\eta|\beta(x)|).$$

which belongs to $L^1(D)$ for every $x \in J^*$. Therefore, by the dominated convergence theorem,

$$\lim_{\lambda \to 0} \langle \beta(x) - \beta_\lambda(x) \rangle \, Dv(x) = 0$$

for every $x \in H \cap J^*$. Using again the uniform boundedness in $L^\infty(D)$ of $v_{\lambda n}(x)$ we also have

$$|\langle \beta(x) - \beta_\lambda(x) Dv_{\lambda n}(x) \rangle| \lesssim 1 + \int_D j^*(\delta|\beta(x)|),$$

where the right-hand side belongs to $L^1(H, \mu)$ by Theorem 5.4. A further application of the dominated convergence theorem thus yields

$$\lim_{\lambda \to 0} \int_H |\langle \beta(x) - \beta_\lambda(x) Dv_{\lambda n}(x) \rangle| \, \mu(dx) = 0. \qquad \square$$

We are now in the position to state and prove the main result of this section, that gives a positive answer to the problem of L^1-uniqueness for the Kolmogorov operator L_0. The question is whether the extension to $L^1(H, \mu)$ of the transition semigroup P, generated by the solution to the stochastic Eq. 1.1, is the only strongly continuous semigroup on $L^1(H, \mu)$ whose infinitesimal generator is an extension of the Kolmogorov operator L_0. Recall that, apart of the standing assumptions of Section 2, we are also assuming that β is a function, B is non-random and does not depend on the unknown, V is continuously embedded in $L^4(D)$, and H is the domain of a fractional power of (a shift of) A, seen as the negative generator of an analytic semigroup in V'.

Theorem 6.5 *The generator L of the extension to $L^1(H, \mu)$ of the transition semigroup P is the closure of L_0 in $L^1(H, \mu)$.*

Proof Since the extension of the transition semigroup P to $L^1(H, \mu)$ is contractive, it follows by the Lumer-Phillips theorem that L is m-accretive. As L coincides with L_0 on $\mathsf{D}(L_0)$, this implies that L_0 is accretive in $L^1(H, \mu)$, hence, in particular, closable. We are going to show that the image of $\alpha I + L_0$ is dense in $L^1(H, \mu)$ for all $\alpha > 0$. Let $f \in L^1(H, \mu)$ and $\varepsilon > 0$. Since $\mathsf{D}(L_0)$ is dense in $L^1(H, \mu)$, there exists $g \in \mathsf{D}(L_0)$ such that $\|f - g\|_{L^1(H,\mu)} < \varepsilon/2$. Setting, for any $n \in \mathbb{N}$ and $\lambda \in (0, 1)$,

$$v_{\lambda n}(x) := \int_0^\infty e^{-\alpha t} \, \mathbb{E} \, g(X_{\lambda n}^x(t)) \, dt,$$

if follows by Lemma 6.3 that $v_{\lambda n} \in \mathsf{D}(L_0)$ and that

$$\alpha v_{\lambda n}(x) + L_0^{\lambda n} v_{\lambda n}(x) = g(x)$$

for every $x \in V \cap J \cap J^*$, hence also

$$\alpha v_{\lambda n}(x) + L_0 v_{\lambda n}(x) - g(x) = L_0 v_{\lambda n}(x) - L_0^{\lambda n} v_{\lambda n}(x).$$

Thanks to Lemma 6.4, there exist $\lambda_0 > 0$ and $n_0 \in \mathbb{N}$ such that

$$\left\| L_0 v_{\lambda_0 n_0} - L_0^{\lambda_0 n_0} v_{\lambda_0 n_0} \right\|_{L^1(H,\mu)} < \varepsilon/2,$$

hence, setting $\varphi := v_{\lambda_0 n_0}$,

$$\|\alpha \varphi + L_0 \varphi - f\|_{L^1(H,\mu)} \le \|\alpha \varphi + L_0 \varphi - g\|_{L^1(H,\mu)} + \|f - g\|_{L^1(H,\mu)} < \varepsilon.$$

As $\varepsilon > 0$ was arbitrary, it follows that the image of $\alpha I + L_0$ is dense in $L^1(H, \mu)$. Since L_0 is closable, the Lumer-Phillips theorem implies that $-\overline{L_0}$, the closure of $-L_0$ in $L^1(H, \mu)$, generates a strongly continuous semigroup of contractions in $L^1(H, \mu)$. Recalling that L is an extension of L_0, it follows again by the Lumer-Phillips theorem that $L = \overline{L_0}$ (see, for instance, [9, Theorem 1.12]).   □

References

1. Aliprantis, C.D., Border, K.C. Infinite dimensional analysis, 3rd edn. Springer, Berlin (2006). MR 2378491
2. Ambrosetti, A., Prodi, G.: A primer of nonlinear analysis. Cambridge University Press, Cambridge (1995). MR 1336591 (96a:58019)
3. Barbu, V.: Nonlinear differential equations of monotone types [sic] in Banach spaces. Springer, New York (2010). MR 2582280
4. Barbu, V., Da Prato, G.: Ergodicity for nonlinear stochastic equations in variational formulation. Appl. Math. Optim. **53**(2), 121–139 (2006). MR 2172782 (2007d:60030)
5. Bourbaki, N.: Intégration. Chapitre IX: Intégration sur les espaces topologiques séparés. Hermann, Paris (1969). MR 0276436
6. Cerrai, S.: Second order PDE's in finite and infinite dimension, Lecture Notes in Mathematics, vol. 1762. Springer-Verlag, Berlin (2001). MR 2002j:35327
7. Da Prato, G.: Kolmogorov equations for stochastic PDEs. Birkhäuser Verlag, Basel (2004). MR 2111320 (2005m:60002)
8. Da Prato, G., Zabczyk, J.: Ergodicity for infinite-dimensional systems. Cambridge University Press, Cambridge (1996). MR 1417491 (97k:60165)
9. Eberle, A.: Uniqueness and non-uniqueness of semigroups generated by singular diffusion operators, Lecture Notes in Mathematics, vol. 1718. Springer-Verlag, Berlin (1999). MR 1734956 (2001c:60122)
10. Es-Sarhir, A., Stannat, W.: Invariant measures for semilinear SPDE's with local Lipschitz drift coefficients and applications. J. Evol. Equ. **8**(1), 129–154 (2008). MR 2383485
11. Es-Sarhir, A., Stannat, W.: Improved moment estimates for invariant measures of semilinear diffusions in Hilbert spaces and applications. J. Funct. Anal. **259**(5), 1248–1272 (2010). MR 2652188
12. Krylov, N.V., Rozovskiǐ, B.L.: Stochastic evolution equations, Current problems in mathematics, Vol. 14 (Russian), Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow, 1979, pp. 71–147, 256. MR 570795 (81m:60116)
13. Lions, J.-L., Magenes, E.: Problèmes aux limites non homogènes et applications, vol. 1. Dunod, Paris (1968). MR 0247243

14. Marinelli, C., Prévôt, C., Röckner, M.: Regular dependence on initial data for stochastic evolution equations with multiplicative Poisson noise. J. Funct. Anal. **258**(2), 616–649 (2010). MR 2557949
15. Marinelli, C., Scarpa, L.: Refined existence and regularity results for a class of semilinear dissipative SPDEs. arXiv:1711.11091
16. Marinelli, C., Scarpa, L.: Well-posedness of monotone semilinear SPDEs with semimartingale noise. arXiv:1805.07562
17. Marinelli, C., Scarpa, L.: A variational approach to dissipative SPDEs with singular drift. Ann. Probab. **46**(3), 1455–1497 (2018). MR 3785593
18. Marinelli, C., Ziglio, G.: Ergodicity for nonlinear stochastic evolution equations with multiplicative Poisson noise. Dyn. Partial Differ. Equ. **7**(1), 1–23 (2010). MR 2656416
19. Pardoux, E.: Equations aux derivées partielles stochastiques nonlinéaires monotones. Ph.D. thesis, Université Paris XI (1975)
20. Stannat, W.: Stochastic partial differential equations: Kolmogorov operators and invariant measures. Jahresber. Dtsch. Math.-Ver. **113**(2), 81–109 (2011). MR 2768734