University College London

DEPARTMENT OF MATHEMATICS
PhD degree in Mathematics

A variational approach to some classes of singular stochastic PDEs

Candidate:
Mr. Luca Scarpa

Supervisor:
Prof. Carlo Marinelli

Secondary supervisors:
Prof. Leonid Parnovski
Dr. Nadia Sidorova

Thesis submitted in 2018

I, Luca Scarpa, confirm that
the work presented in this thesis is my own.
Where information has been derived from other sources,
I confirm that this has been indicated in the thesis.

Abstract

This thesis contains an analysis of certain classes of parabolic stochastic partial differential equations with singular drift and multiplicative Wiener noise. Equations of this type have been studied so far only under rather restrictive hypotheses on the growth and smoothness of the drift. By contrast, we give here a self-contained treatment for such equations under minimal assumptions.

The first part of the thesis is focused on semilinear SPDEs with singular drift. In particular, the nonlinearity in the drift is the superposition operator associated to a maximal monotone graph everywhere defined on the real line, on which neither continuity nor growth assumptions are imposed. The hypotheses on the diffusion coefficient are also very general. First of all, well-posedness is established for the equation through a combination of variational techniques and a priori estimates. Secondly, several refined well-posedness results are provided, allowing the initial datum to be only measurable and the diffusion coefficient to be locally Lipschitz-continuous. Moreover, existence, uniqueness and integrability properties of invariant measures for the Markovian semigroup generated by the solution are proved. Furthermore, the associated Kolmogorov equation is studied in L^p spaces with respect to the invariant measure and the infinitesimal generator of the transition semigroup is characterized as the closure of the corresponding Kolmogorov operator.

The second part of the thesis focuses on equations with monotone singular drift in divergence form. Due to rather general assumptions on the growth of the nonlinearity in the drift, which, in particular, is allowed to grow faster than polynomially, existing techniques are not applicable. Equations of this type are typically doubly nonlinear, making their treatment more challenging in comparison to the semilinear case. Well-posedness for such equations is established in several cases, suitably generalizing the techniques for semilinear equations to an abstract generalized variational setting.

Impact statement

The research results presented in this thesis deal with a general variational approach to some classes of singular stochastic partial differential equations. The greatest impact of this analysis concerns the abstract research field of the study of stochastic evolution equations with singular terms. The main interest and relevance of the entire thesis are thus of academic type, both from a mathematical perspective and in terms of applications to other fields where SPDEs play an important role.

The available literature on stochastic evolution equations is very developed, and addresses problems such as well-posedness, long-time behaviour and regularity of solutions. However, several of the existing techniques dealing with stochastic evolution equations strongly rely on growth and smoothness assumptions on the drift, which is particular cannot exceed prescribed polynomial growth rates, and hence are not applicable in many concrete situations. By contrast, in this thesis we introduce a generalized variational approach and give a self-contained analysis of certain classes of equations, where the drift is allowed to grow faster than polynomially. This is obtained through a combination of tools from montone and convex analysis, and compactness results in spaces of vector-valued functions.

Such problems are clearly interesting on their own in a mathematical perspective, and represent a first step towards a generalization of the existing variational techniques also to singular stochastic evolution equations of monotone type. The relevance of the results presented in this work could be strongly beneficial to the academic research field of stochastic evolution equations. Indeed, one of the main topics that are currently being investigated is the study of stochastic equations with possibly degenerate terms. Even if this thesis focuses in detail on semilinear and divergence-form equations, the abstract variational setting that we introduce here can be applied also to several other types of stochastic equations, and provides an effective tool to analyse stochastic equations with singular terms in a variational framework.

Furthermore, the techniques presented in this thesis can be adapted in order to deal with other types of stochastic PDEs, which are much more relevant for applications to physics. Let us mention, above all, the stochastic Allen-Cahn and Cahn-Hilliard equations, which play an important role in phase transition modelling for example. While these have received much attention in the deterministic setting in the last years, only few works are focused the corresponding stochastic equations. This is due in large part to the the fact that the high singularity of the drift gives rise to several difficulties in the stochastic case. For these reasons, the contribution of this thesis could be beneficial also to the applicative aspect of the study of stochastic PDEs, since it provides an effective way of handling parabolic stochastic PDEs with possibly degenerate terms.

Contents

Introduction

This PhD thesis is devoted to the analysis of some classes of parabolic stochastic partial differential equations with singular drift and multiplicative Wiener noise. The main motivation of the entire work is that equations of this type have been studied in the available literature on stochastic evolution equations only under rather restrictive hypotheses on the growth and smoothness of the drift.

It is well known that in the deterministic setting a complete and satisfactory theory is available for evolution equations of monotone type in the form

$$\frac{d}{dt}u + Au \ni f, \qquad u(0) = u_0,$$

where A is a (multivalued) maximal monotone operator on a Hilbert space H or, more generally, an m-accretive operator on a Banach space E, $f \in L^1(0,T;E)$ and u_0 is a given initial datum. The mere assumption of maximal monotonicity on the operator A is very broad and allows to include in this general treatment also several degenerate equations arising from applications. The monotonicity of the equation provides a very powerful tool to obtain a complete well-posedness theory also in highly singular settings.

The corresponding stochastic evolution equation in this abstract setting reads

$$du + Au\,dt \ni B(u)\,dW, \qquad u(0) = u_0,$$

where W is a cylindrical Wiener process on a certain separable Hilbert space and B is a suitable stochastically integrable operator with respect to W. However, a general theory of existence of solutions and continuous dependence on the data for such equations still seems out of reach, even in some simplified setting where B is nonrandom and independent of both u and t.

Among the current literature on stochastic evolution equations, significant results have been obtained only in some specific cases. In this direction, there are two main available approaches to the study of stochastic PDEs.

First of all, for semilinear equations, a powerful tool is offered by the semigroup approach: the concept of solution is formulated in analogy with the corresponding theory for deterministic equations, using the semigroup generated by the linear component of the drift. Here, one of the main difficulties consists in giving appropriate sense to the so-called "stochastic convolution". The approach has been widely studied and is very effective in terms of both well-posedness issues, regularity and long-time behaviour: for example, the general theory of semigroups on Banach spaces allows to assume very broad requirements for the initial datum, and to investigate several interesting qualitative properties such as maximal regularity and optimal estimates. On the other side, the theory presents some drawbacks: due to the formulation of solutions, the equations must necessarily be semilinear and, most notably, the nonlinearity in the drift cannot grow faster than polynomially at infinity.

The second main approach to nonlinear stochastic PDEs is the variational approach. This is a very natural and powerful generalization to the stochastic case of the classical variational theory for deterministic evolution equations by Lions and Magenes (see for example [51–53]). It was introduced in the classical works [46, 72] by Pardoux, Krylov and Rozovskiĭ, and it has been extensively developed in the recent years as well in terms of regularity, invariant measures and long-time behaviour of solutions. One of the main advantages of the variational approach is the possibility to deal with fully nonlinear equations, where the operator A is defined from a Banach space to its dual, and can be random and time-dependent. On the other side, the operator A must satisfy some classical monotonicity, coercivity and boundedness conditions. Even if these can be weakened, this forces the nonlinearity in the equation not to exceed a prescribed polynomial growth.

In this PhD thesis, we study some classes of stochastic evolution equations which do not fall in the classical variational approach because of the high singularity in the drift, and we give a self-contained treatment in terms of well-posedness, regularity and long-time behaviour in a generalized variational setting. In particular, the results obtained here considerably extend the classical ones of the variational theory, and are a very powerful tool to deal with stochastic equations with singular drift exceeding polynomial growth.

The first part of the thesis is focused on singular semilinear equations on a smooth bounded domain D of $\mathbb{R}^d$ in the form

$$dX + AX\,dt + \beta(X)\,dt \ni B(X)\,dW, \qquad X(0) = X_0, \tag{0.0.1}$$

where A is a linear coercive maximal monotone operator on $L^2(D)$ and β is a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$. The singularity of the equation is contained in the graph β. The analysis of multivalued operators β is crucial: indeed, since any increasing function of $\mathbb{R}$, possibly with infinitely many discontinuities (jumps), can be canonically embedded into a maximal monotone graph, we include in our analysis also reaction-diffusion equations with discontinuous reaction terms. Secondly, and more importantly for us, no growth assumption will be in order for β, which is allowed to grow at any rate at infinity. The main requirement on β in our study is that its effective domain is the whole real line. Even if this assumption is not needed in the deterministic theory, in the stochastic case it seems to be essential. Nevertheless, this setting clearly does not fall in the classical variational framework, and we are able to give appropriate sense also to semilinear stochastic equations with rapidly growing drift.

In Chapter 2 we present a natural well-posedness result for such equations, which is part of the joint work [65] with Carlo Marinelli, to appear on *Annals of Probability*. The proof is based on a double approximation of the problem: the diffusion coefficient B is firstly regularized through a suitable smoothing elliptic operator, and secondly the graph β is approximated using its Yosida approximation. The corresponding regularized equations can be solved invoking the classical variational theory. Then, we prove several uniform estimates on the approximated solutions, both pathwise and in expectation, which allow us to pass to the limit and obtain a solution for the original problem. The proof is strongly based on arguments from monotone and convex analysis, as well as compactness and lower semicontinuity results in spaces of Bochner-integrable functions.

In Chapter 3 we show some refined well-posedness results for the equation (0.0.1), which have been object of the article [63] with Carlo Marinelli. First of all, we prove that the solution has strongly continuous trajectories in $H := L^2(D)$. Secondly, we use such continuity property to extend the well-posedness result also to the case when X_0 is only measurable and B is locally-

Lipschitz continuous in its last argument. More specifically, we show that the the well-posedness of the equation (0.0.1) can be extrapolated to the whole range $p \in [0, +\infty[$, meaning that if the initial datum X_0 belongs to $L^p(\Omega; H)$, then the solution X belongs to $L^p(\Omega; E)$, where E is the natural space of the trajectories, and the map $X_0 \mapsto X$ is continuous. In the particular case $p = 0$, the convergence in probability of the initial data yields the uniform convergence of the solutions in $[0, T]$ in probability. The proofs are based on the introduction of suitable stopping times, depending on the initial datum and the coefficient B, as well as a generalized Itô's formula in an abstract variational setting.

In Chapter 4 we investigate the ergodicity properties of solutions to equation (0.0.1): these results have been collected in the joint paper [62] with Carlo Marinelli. Since the equation cannot be treated within the classical variational setting, we cannot rely on established tools to study ergodicity. First of all, we show that the transition semigroup induced by the solution X admits an ergodic invariant measure μ, which is also unique and strongly mixing if β is superlinear: this result follows mainly by a priori estimates and compactness results obtained by a suitable version of Itô's formula. Secondly, we study the Kolmogorov equation associated to (0.0.1), and we characterize the infinitesimal generator of the transition semigroup on $L^1(H, \mu)$ as the closure of the Kolmogorov operator associated to (0.0.1). This is done regularizing first the Yosida approximation of β through convolutions with mollifiers and solving the Kolmogorov equations of the corresponding approximated problem through existing techniques. Then, several regular dependence results of the corresponding approximated solutions with respect to the initial datum are proved, allowing us to verify that the Kolmogorov operator coincides with the infinitesimal generator of the transition semigroup on a dense subset of $L^1(H, \mu)$. Finally, a careful application of the Lumer-Phillips theorem together with the m-dissipativity of the Kolmogorov operator yields the desired result.

In Chapter 5 we complement the analysis of semilinear singular stochastic equations with a regularity result, which is part of the above-mentioned work [63]. We prove that if X_0 and B are more regular, then the regularity of the solution X also improves, irrespectively of the singularity in β. For example, if A (better said, the part of A in H) is self-adjoint, the solution has paths belonging to the domain of A in H if X_0 and B, roughly speaking, take values in the domain of $A^{1/2}$. This implies that X is a strong solution in the classical analytical sense, not just in the variational sense.

The second part of the thesis is devoted to the study of singular stochastic equations with drift in divergence form of the type

$$dX - \operatorname{div} \gamma(\nabla X)\, dt + \beta(X)\, dt \ni B(X)\, dW, \qquad X(0) = X_0, \tag{0.0.2}$$

where γ and β are maximal monotone graphs on $\mathbb{R}^d$ and $\mathbb{R}$, respectively. Such equations are doubly-nonlinear, hence more difficult to handle than the semilinear case. The available literature on divergence-form equations as (0.0.2) is very limited and entirely focused on the classical case where γ and β satisfy suitable coercivity assumptions and are polynomially bounded (the so-called Leray-Lions conditions). In this setting, some qualitative results have been obtained in the specific case $\gamma(x) = |x|^{p-2}x$, which corresponds to the p-Laplace equation. We study instead equations of the form (0.0.2) under no growth hypotheses on γ and β, thus widely improving the results in the existing literature. In analogy with semilinear equations, we only need to assume that γ and β are everywhere defined: the final well-posedness result is much more difficult to achieve in this case, due to the fact that the double nonlinearity gives rise to several nontrivial measurability problems, and it will follow after some intermediate steps.

In Chapter 6 we study equation (0.0.2) in the case where γ satisfies the above-mentioned Leray-Lions polynomial conditions and β satisfies the same assumptions considered in the case of semilinear equations. The results presented in this chapter have been collected in the article [75], published on *Journal of Differential Equations*. Through a double approximation analogous to the one used in Chapter 2, well-posedness is proved through monotonicity and compactness arguments. However, when γ is multivalued, the uniqueness of the solution components $-\operatorname{div}\gamma(\nabla X)$ and $\beta(X)$ may not hold necessarily, and this causes in turn nontrivial measurability problems. For these reasons, in this chapter γ is assumed to be single-valued.

In Chapter 7 we focus our attention entirely on the divergence term and we study equation (0.0.2) in the case $\beta = 0$ and without any polynomial growth condition on γ. The results presented here have been object of the article [67] with Carlo Marinelli, published on *Stochastics & Partial Differential Equations: Analysis and Computations*. Due to the lack of coercivity and growth conditions, the first step consists in the regularization of the problem, replacing γ with its Yosida approximation and adding a "small" elliptic term, thus obtaining a family of equations for which well-posedness is known to hold. As a second step, we prove that the family of solutions to the regularized equations is compact in suitable topologies, so that, by passage to the limit in the regularization parameters (roughly speaking), a process can be constructed that, in a final step, is shown to actually be the unique solution to the original problem. The fact that γ is the subdifferential of a certain convex function will be used to recover a suitable integrability condition on ∇X. As in the previous chapter, if γ is multivalued several measurability problems arise, as the uniqueness of $-\operatorname{div}\gamma(\nabla X)$ does not imply uniqueness of $\gamma(\nabla X)$, hence we still assume that γ is single-valued.

In Chapter 8 we give an alternative self-contained treatment to the case $\beta = 0$, including in our analysis also multivalued operators γ. These results are collected in the short contribution [64] with Carlo Marinelli, published on *Springer Proceedings in Mathematics & Statistics*. The main idea is to work only using estimates in expectation, so that the measurability of the limit processes is ensured by the weak compactness itself. However, in order to do so, the price to pay is a loss of regularity of the solutions in comparison with the corresponding result in Chapter 7.

Finally, in Chapter 9 we consider equation (0.0.2) in its most general form, with no growth conditions on the operators γ and β, also allowing γ to be multivalued. We thus obtain a general well-posedness result that unifies and extends those proved in the previous chapters. Such final result has been object of the joint paper [66] with Carlo Marinelli, to appear in *Atti Accademia Nazionale Lincei. Rendiconti Lincei. Matematica e Applicazioni*. In this case, we can prove that $-\operatorname{div}\gamma(\nabla X) + \beta(X)$ is unique, hence that it is measurable, but showing that each one of them is unique (hence measurable) seems difficult, if not impossible. This is the reason why γ was assumed to be single-valued in the previous chapters. However, we show here that it is possible to construct two suitable limiting processes which are "sections" of $\gamma(\nabla X)$ and $\beta(X)$, and that are indeed measurable. In other words, we prove that the uniqueness of the limit processes (which is still not present here) is not necessary to have measurability: the intuitive idea is that we restore uniqueness working on a suitable quotient space.

The main goal of this thesis is to give a rigorous presentation of a new general variational approach that allows to handle stochastic evolution equations with possibly singular and rapidly growing drift. Let us conclude this introduction with some remarks on the advantages and the possibilities that the generalized variational approach described in this thesis offers also in other contexts. Even if we have studied in detail the case of semilinear and divergence-form equations, it is worth emphasizing that these techniques can be adapted also to several other types of stochastic PDEs with singular drift terms.

First of all, these ideas can be adapted for example to handle singular Allen-Cahn equations with dynamic boundary conditions: in this case, the presence of a nonlinearity also on the boundary makes the problem more complicated. Thanks to a compatibility conditions between the nonlinearities in the interior of the domain and on the boundary, well-posedness is shown to hold irrespectively of the growth rate of the drifts. We refer for further detail to the joint work [69] with Carlo Orrieri.

Moreover, a similar approach has been developed to give a general treatment to the stochastic "pure" Cahn-Hilliard equation with a singular double-well potential. Here, the structure of the equation is completely different since, as it is well-known, it is of order four in space and monotone in a suitable dual space. Nevertheless, through a combination of *ad hoc* a priori estimates and the ideas exposed in this thesis, well-posedness is proved for any everywhere-defined potential. We point out in this direction the contribution [76], to appear in *Nonlinear analysis*.

Finally, in a work in preparation, the well-posedness result for semilinear equations is extended also to the case where β is random and time-dependent, and the noise is given by a general Hilbert-space-valued semimartingale.

List of Symbols

$\mathbb{N}$	The set of natural numbers
$\mathbb{R}$	The real line
$\mathbb{R}^d$	The d-dimensional Euclidean space, $d \in \mathbb{N} \setminus \{0\}$
D	Smooth bounded domain in $\mathbb{R}^d$
$\lVert\cdot\rVert_E$	The norm in the Banach space E
E^*, E'	The dual space of the Banach space E
$\to$, $\rightharpoonup$, $\overset{*}{\rightharpoonup}$	Strong, weak and weak* convergence in Banach spaces, respectively
$\mathscr{B}(E)$	The Borel σ-algebra of the space E
$B_b(E)$	The space of $\mathscr{B}(E)$-measurable bounded functions on E
$C_b(E)$	The subspace of $B_b(E)$ of continuous functions on E
$E \hookrightarrow F$	$E \subseteq F$ and the inclusion map $i : E \to F$ is continuous
$\mathscr{L}(E, F)$	The space of linear bounded operators from E to F
$\mathscr{L}_s(E, F)$	The space $\mathscr{L}(E, F)$ endowed with the strong operator topology
$\mathscr{L}^1(E, F)$	The space of trace class operators from E to F
$\mathscr{L}^2(E, F)$	The space of Hilbert-Schmidt operators from E to F
$\overline{K}$	The closure of the set K
$\operatorname{Int} K$	The interior of the set K
$\mathsf{D}(G)$	The domain of a generic map G
$\mathsf{R}(G)$	The range of a generic map G
$C^k(D)$	The space of continuously differentiable functions up to order $k \in \mathbb{N}$
$C^k_c(D)$	The space of functions in $C^k(D)$ with compact support in D
$L^p(D)$	The space of Lebesgue-p-integrable functions on D, $p \in [1, +\infty]$
$W^{m,p}(D)$	The Sobolev space of index $p \in [1, +\infty]$ and order $m \in \mathbb{N}$
$W^{m,p}_0(D)$	The closure of $C^\infty_c(D)$ in $W^{m,p}(D)$
$H^m(D)$	The space $W^{m,2}(D)$

$H_0^m(D)$	The space $W_0^{m,2}(D)$
$L^p(Y;E)$	The space p-Bochner-integrable functions from a measure space $(Y,\mathscr{A},\mu)$ to E
$L^0(Y;E)$	The space of strongly measurable functions from $(Y,\mathscr{A},\mu)$ to E
$C([a,b];E)$	The space of continuous functions from $[a,b]$ to E
$C_w([a,b];E)$	The space of weakly continuous functions from $[a,b]$ to E
$(\Omega,\mathscr{F},\mathbb{P})$	Probability space
$(\mathscr{F}_t)_{t\in[0,T]}$	Filtration on $(\Omega,\mathscr{F},\mathbb{P})$, $T>0$
$\mathscr{M}_1(E)$	The space of probability measures on the measurable space $(E,\mathscr{E})$
W	Cylindrical Wiener process
$G\cdot W$	Stochastic integral of G with respect to W
X_t^*	The quantity $\Vert X\Vert_{L^\infty(0,t;E)}$, with X an E-valued process
$a\lesssim b$	There exists $N>0$ such that $a\leq Nb$, with $a,b\in\mathbb{R}$

Chapter 1

Prerequisites

This first chapter contains some theoretical prerequisites that may be useful for the reader throughout the thesis. For convenience, we recall some basic notions of convex and monotone analysis, the classical variational approach to SPDEs and standard Markovianity and ergodicity properties.

We stress that the classical results of functional analysis, probability theory and stochastic integration in infinite dimensions are taken for granted. The reader may refer to the classical textbooks [10, 21–23, 28, 30, 51–53, 56].

1.1 Functional analysis

Let $(E, \|\cdot\|_E)$ be a real Banach space. The dual space is denoted by E^* and the duality pairing between E^* and E is indicated with the symbol $\langle \cdot, \cdot \rangle_E$. The dual norm $\|\cdot\|_{E^*}$ on E^* is defined as

$$\|\cdot\|_{E^*} := \sup_{\|x\|_E \leq 1} \langle y, x \rangle_E \,, \qquad y \in E^* \,.$$

The duality mapping of E is the set-valued function $J : E \to 2^{E^*}$ given by

$$J(x) := \left\{ y \in E^* : \quad \langle y, x \rangle_E = \|x\|_E^2 = \|y\|_{E^*}^2 \right\} \,, \qquad x \in E \,.$$

It is well known that the set $J(x)$ is not empty for every $x \in E$ as a consequence of the Hahn-Banach theorem. Let us collect some some useful properties in the following proposition: these are well-known results and the proof can be found in the textbooks quoted above.

Proposition 1.1.1. *If E is reflexive, then the duality mapping of E^* is the inverse map $J^{-1} : E^* \to E$, given by $J^{-1}(y) := \{x \in E : y \in J(x)\}$ for $y \in E^*$. If E^* is strictly convex, then the duality mapping J is single-valued and demicontinuous i.e. continuous from E to E^*_w (the space E^* endowed with the weak-star topology). If E^* is uniformly convex, then J is uniformly continuous on every bounded subset of E.*

Example 1.1.2. If $E = H$ is a Hilbert space, the duality mapping J is the canonical Riesz isomorphism from H to H^*. If $E = L^p(\Omega, \mathscr{F}, \mu)$, where $(\Omega, \mathscr{F}, \mu)$ is a measure space and $1 < p < +\infty$, the duality mapping is given by

$$J : L^p(\Omega) \to L^{\frac{p}{p-1}}(\Omega) \,, \qquad J(f) := |f|^{p-2}_{L^p(\Omega)} f \|f\|^{2-p}_{L^p(\Omega)} \,, \quad f \in L^p(\Omega) \,,$$

while if $p = 1$, then J is generally multivalued and given by

$$J(f) := \{g \in L^\infty(\Omega) : g \in \mathrm{sign}(f) \text{ a.e. in } \Omega\} , \quad f \in L^1(\Omega) ,$$

where

$$\mathrm{sign} : \mathbb{R} \to 2^{\mathbb{R}} , \qquad \mathrm{sign}(r) := \begin{cases} -1 & \text{if } r < 0 , \\ [-1, 1] & \text{if } r = 0 , \\ 1 & \text{if } r > 0 . \end{cases}$$

1.2 Convex analysis

We recall here some basic concepts of convex analysis: the reader may refer to [10] for further details.

Let $(E, \|\cdot\|_E)$ be a real Banach space and $\phi : E \to (-\infty, +\infty]$. We say that ϕ is convex, proper or lower semicontinuous (l.s.c.) when, respectively,

- $\phi(\vartheta x + (1 - \vartheta)y) \leq \vartheta\phi(x) + (1 - \vartheta)\phi(y) \qquad \forall\, \vartheta \in [0, 1] , \quad \forall\, x, y \in E ,$
- $\exists\, x \in E : \quad \phi(x) < +\infty ,$
- $\phi(x) \leq \liminf_{y \to x} \phi(y) \qquad \forall\, x \in E .$

The domain and the epigraph of ϕ are defined as

$$\mathsf{D}(\phi) := \{x \in E : \ \phi(x) < +\infty\} , \qquad \mathrm{epi}\, \phi := \{(x, s) \in E \times \mathbb{R} : \ \phi(x) \leq s\} .$$

The notions of domain and epigraph play an important role in convex analysis: in particular, we have the following well-known properties.

Proposition 1.2.1. *The function ϕ is convex if and only if $\mathrm{epi}\, \phi$ is convex in $E \times \mathbb{R}$; ϕ is proper if and only if $\mathsf{D}(\phi) \neq \emptyset$; ϕ is l.s.c. if and only if the set $\{\phi \leq s\} := \{x \in E : \ \phi(x) \leq s\}$ is closed in E for every $s \in \mathbb{R}$. Moreover, if ϕ is convex, proper and l.s.c. then ϕ is continuous in $\mathrm{Int}\, \mathsf{D}(\phi)$, and there exist $\bar{y} \in E^*$ and $b \in \mathbb{R}$ such that*

$$\phi(x) \geq \langle \bar{y}, x \rangle_E + b \quad \forall\, x \in E .$$

An important role in convex analysis is played by the concept of conjugate functions and their properties.

Definition 1.2.2. *If ϕ is proper, the conjugate function of ϕ (or Legendre transform fo ϕ) is defined as*

$$\phi^* : E^* \to (-\infty, +\infty] , \qquad \phi^*(y) := \sup_{x \in E} \{\langle y, x \rangle - \phi(x)\} , \quad y \in E^* .$$

By the previous definition, the generalized Young inequality follows directly:

$$\langle y, x \rangle_E \leq \phi(x) + \phi^*(y) \qquad \forall\, x \in E , \ \forall\, y \in E^* .$$

It is well-known that ϕ^* is always convex and l.s.c. Furthermore, if ϕ is also proper and l.s.c. then ϕ^* is proper as well.

The class of convex, proper and l.s.c. functions on Banach spaces is widely studied as it represents a useful tool in the analysis of PDEs and SPDEs of monotone type. A fundamental notion that we must introduce is the subdifferential.

Definition 1.2.3. *Let ϕ be convex, proper and l.s.c. The subdifferential of ϕ is the multivalued function*

$$\partial\phi : E \to 2^{E^*}\,, \qquad \partial\phi(x) := \{y \in E^* :\ \phi(x) + \langle y, z - x\rangle \le \phi(z)\ \ \forall\, z \in E\}\ .$$

The domain and the range of the subdifferential $\partial\phi$ are defined as

$$\mathsf{D}(\partial\phi) := \{x \in E :\ \partial\phi(x) \neq \emptyset\}\ , \qquad \mathsf{R}(\partial\phi) := \bigcup_{x\in E} \partial\phi(x)\,.$$

The following properties are well-known.

Proposition 1.2.4. *If ϕ is convex, proper and l.s.c. then*

$$\mathsf{D}(\partial\phi) \subseteq \mathsf{D}(\phi) \quad \textit{densely}\,, \qquad \mathrm{Int}\,\mathsf{D}(\phi) \subseteq \mathsf{D}(\partial\phi)\,.$$

Proposition 1.2.5. *If E is reflexive and ϕ is convex, proper and l.s.c. then for every $x \in E$ and $y \in E^*$ the following three conditions are equivalent:*

$$y \in \partial\phi(x)\,, \qquad x \in \partial\phi^*(y)\,, \qquad \langle y, x\rangle_E = \phi(x) + \phi^*(y)\,.$$

In particular, $\partial\phi^* = (\partial\phi)^{-1}$.

Proposition 1.2.6. *If ϕ is weakly l.s.c. and every level set $\{x \in E :\ \phi(x) \le s\}$ is weakly compact for every $s \in \mathbb{R}$, there exists $\bar{x} \in E$ such that $\phi(\bar{x}) = \inf_{x\in E} \phi(x)$. In particular, this is true if E is reflexive and ϕ is convex, proper, l.s.c. such that*

$$\lim_{\|x\|_E \to +\infty} \phi(x) = +\infty\,.$$

The next result is very well known, but we prefer to stress it and give a possible proof as it plays a fundamental role in the key idea of the main results presented in this thesis.

Proposition 1.2.7. *If E is reflexive, then the following conditions are equivalent:*

a) $\mathsf{R}(\partial\phi) = E^*$ *and $\partial\phi^*$ is bounded on bounded sets,*

b) $\lim_{\|x\|_E \to +\infty} \frac{\phi(x)}{\|x\|_E} = +\infty\,.$

Proof. $a) \Rightarrow b)$. By the generalized Young inequality we have

$$\phi(x) \ge \langle y, x\rangle_E - \phi^*(y)$$

for every $x \in E$ and $y \in E$. Denoting by J the duality mapping of E and choosing $y = \eta z \|x\|_E^{-1}$ with $z \in J(x)$ (for $x \neq 0$ and $\eta > 0$) we have

$$\phi(x) \ge \eta\|x\|_E - \phi^*(\eta z\|x\|_E^{-1}) \qquad \forall\, \eta > 0\,, \quad \forall\, x \in E \setminus \{0\}\,.$$

Let now $M > 0$ be arbitrary. By definition of J and the choice of z we have

$$\|\eta z\|x\|_E^{-1}\|_{E^*} = \eta\|x\|_E^{-1}\|z\|_{E^*} = \eta\,.$$

Hence, since the boundedness hypothesis on $\partial\phi^*$ in $a)$ implies the same for ϕ^* by definition of subdifferential, we deduce that there exists $C_\eta > 0$ such that $|\phi^*(\eta z\|x\|_E^{-1})| \leq C_\eta$, so that

$$\phi(x) \geq \eta\|x\|_E - C_\eta \qquad \forall\, x \in E \setminus \{0\}\,, \quad \forall\, \eta > 0\,.$$

Choosing for example $\eta = 2M$, this implies that for $\|x\|_E \geq \frac{C_{2M}}{M}$ we have

$$\frac{\phi(x)}{\|x\|_E} \geq \eta - \frac{C_\eta}{\|x\|_E} \geq 2M - M = M\,,$$

from which $b)$ follows by arbitrariness of M.
$b) \Rightarrow a)$. Let $y \in E^*$. The function $\psi_y : E \to (-\infty, +\infty]$ defined as

$$\psi_y(x) := \phi(x) - \langle y, x\rangle_E\,, \qquad x \in E\,,$$

is convex, proper and l.s.c. in E. Moreover, we have that $\partial\psi_y = \partial\phi - y$: indeed, for every $x \in E$,

$$\begin{aligned}
\partial\psi_y(x) &= \{w \in E^* : \ \psi_y(x) + \langle w, \tilde{x} - x\rangle_E \leq \psi_y(\tilde{x}) \quad \forall\, \tilde{x} \in E\}\\
&= \{w \in E^* : \ \phi(x) + \langle w + y, \tilde{x} - x\rangle_E \leq \phi(\tilde{x}) \quad \forall\, \tilde{x} \in E\}\\
&= \{z - y \in E^* : \ \phi(x) + \langle z, \tilde{x} - x\rangle_E \leq \phi(\tilde{x}) \quad \forall\, \tilde{x} \in E\} = \partial\phi(x) - y\,.
\end{aligned}$$

Now, the superlinearity assumption on ϕ contained in $b)$ implies that

$$\lim_{\|x\| \to +\infty} \psi_y(x) = +\infty\,.$$

Hence, since E is reflexive and ψ_y is convex, proper and l.s.c. we deduce that ψ_y attains its minimum on E. Consequently, by definition of subdifferential, there is $x \in E$ such that

$$0 \in \partial\psi_y(x) = \partial\phi(x) - y\,, \qquad \text{i.e.} \quad y \in \partial\phi(x)\,.$$

We infer that $\mathsf{R}(\partial\phi) = E^*$ by arbitrariness of y. Finally, let us show that $\partial\phi^* = (\partial\phi)^{-1}$ is bounded on every bounded subset of E^*. By contradiction, assume that there are two sequences $(x_n)_n \subseteq E$, $(y_n)_n \subseteq E^*$ and a constant $C_0 > 0$ such that $y_n \in \partial\phi(x_n)$ and $\|y_n\|_{E^*} \leq C_0$ for every n, but $\|x_n\|_E \to +\infty$ as $n \to \infty$. By definition of subdifferential we have

$$\phi(x_n) - C_0\|x_n\|_E \leq \phi(x_n) - \langle y_n, x_n\rangle = \phi(x_n) + \langle y_n, 0 - x_n\rangle \leq \phi(0) \quad \forall\, n \in \mathbb{N}\,.$$

Letting $n \to \infty$, by $b)$ the left hand side diverges to $+\infty$, and this is a contradiction. □

1.3 Monotone analysis

We recall here some fundamental results of monotone analysis: the reader may refer to [10] for further details.

Let $(E, \|\cdot\|_E)$ be a real Banach space. We shall denote by the symbol $E \times E^*$ the usual cartesian product between E and E^*, and by the bracket $(\cdot, \cdot)$ the generic element in $E \times E^*$.

A (multivalued, or set-valued) operator from E to E^* is a function

$$A : E \to 2^{E^*}\,.$$

First of all, let us point out that there is a bijection between the set of operators $A : E \to 2^{E^*}$ and the subsets $G \subseteq E \times E^*$. Indeed, given an operator $A : E \to 2^{E^*}$, one can define the graph of A as

$$\mathsf{G}(A) := \{(x, y) \in E \times E^* : \ y \in A(x)\} \subseteq E \times E^* ,$$

and, conversely, given a subset $G \subseteq E \times E^*$, it is a standard matter to check that the operator $A : E \to 2^{E^*}$ defined as

$$A(x) := \{y \in E^* : \ (x, y) \in G\} , \qquad x \in E ,$$

satisfies $\mathsf{G}(A) = G$. Consequently, as it is usually done in literature, we shall identity any operator with its graph, and we will use the terminology *graph* or *operator* with no distinction when convenient. Analogously, we shall use the notation $y \in A(x)$ or $(x, y) \in A$ equivalently for any $x \in E$ and $y \in E^*$.

The domain and the range of $A : E \to 2^{E^*}$ are defined as

$$\mathsf{D}(A) := \{x \in E : \ A(x) \neq \emptyset\} , \qquad \mathsf{R}(A) := \bigcup_{x \in E} A(x) .$$

Throughout this section, $A : E \to 2^{E^*}$ is a given operator. Let us recall some well known concepts that we will use in the sequel.

Definition 1.3.1. *We say that A is monotone when*

$$\langle y_1 - y_2, x_1 - x_2 \rangle_E \geq 0 \qquad \forall\, (x_i, y_i) \in A , \quad i = 1, 2 .$$

We say that A is maximal montone if it is montone and is maximal in the ordered set $(E \times E^*, \subseteq)$, or, equivalently, if A is not properly contained in any monotone subset of $E \times E^*$.

Definition 1.3.2. *Let $A : E \to E^*$ be a single-valued operator with $\mathsf{D}(A) = E$. We say that A is hemicontinuous if*

$$A(x_1 + t x_2) \overset{*}{\rightharpoonup} A(x_1) \quad \text{in } E^* \quad \text{as } t \to 0 \qquad \forall\, x_1, x_2 \in E .$$

We say that A is demicontinuous if it is continuous from $(E, \|\cdot\|_E)$ to the space E^* endowed with the weak topology, or, in other words, if for every $x \in E$ and $(x_n)_n \subseteq E$*

$$x_n \to x \quad \text{in } E \qquad \Rightarrow \qquad A(x_n) \overset{*}{\rightharpoonup} A(x) \quad \text{in } E^* .$$

We say that A is coercive if there is $x_0 \in E$ such that

$$\lim_{n \to \infty} \frac{\langle y_n, x_n - x_0 \rangle}{\|x_n\|_E} = +\infty$$

for every sequence $(x_n, y_n)_n \subseteq A$ with $\|x_n\|_E \to +\infty$ as $n \to \infty$.

The following properties are well-known.

Proposition 1.3.3. *Let A be maximal monotone in $E \times E^*$. Then,*

- *A is weakly-strongly closed, i.e for every $x, y \in E$ and $(x_n, y_n)_n \subseteq A$, if $x_n \to x$ in E and $y_n \overset{*}{\rightharpoonup} y$ in E^*, then $y \in A(x)$;*
- *A^{-1} is maximal monotone in $E^* \times E$;*

- *$A(x)$ is closed and convex in E^* for every $x \in \mathsf{D}(A)$.*

The fundamental characterization of maximal monotone operators is recalled in the following proposition.

Proposition 1.3.4. *Assume that E and E^* are reflexive and strictly convex. Then, A is maximal montone if and only if for every $\lambda > 0$ (or, equivalently, for some $\lambda > 0$) $\mathsf{R}(A+\lambda J) = E^*$, where $J : E \to E^*$ is the duality mapping of E.*

Remark 1.3.5. Note that the hypothesis that E and E^* are strictly convex can be omitted. Indeed, by Asplund's theorem, one may always choose a norm $|\cdot|_E$ in E, equivalent to the usual one $\|\cdot\|_E$, such that E is strictly convex with respect to $|\cdot|_E$ and E^* is strictly convex with respect to the corresponding dual norm. For this reason, we can only assume that E is reflexive (hence so is E^*).

The previous result allows to prove several properties on maximal monotone operators in a very direct way. We recall the most important ones.

Proposition 1.3.6. *Assume that E is reflexive. If A is maximal monotone and $B : E \to E^*$ is hemicontinuous, monotone and bounded, then $A + B$ is maximal monotone.*

Proposition 1.3.7. *Assume that E is reflexive. Then,*

- *any monotone and hemicontinuous operator in $E \times E^*$ is maximal monotone;*
- *any coercive and maximal monotone operator in $E \times E^*$ is surjective (i.e. its range is E^*);*
- *any monotone, hemicontinuous and coercive operator in $E \times E^*$ is surjective.*

One of the major issues in the study of nonlinear PDEs is to approximate maximal monotone operators in a reasonable way, possibly preserving monotonicity and maximality. To this end, there is canonical way to smooth out any generic maximal monotone operator, that we present now. Let us assume that E is reflexive, strictly convex with strictly convex dual E^*, and that $A : E \to 2^{E^*}$ is maximal monotone.

For any $\lambda > 0$ and $x \in E$, the operator $\lambda A + J(\cdot - x)$ is surjective thanks to the last two propositions, so that there is $x_\lambda \in E$ such that

$$\lambda A(x_\lambda) + J(x_\lambda - x) \ni 0\,.$$

Moreover, such x_λ is unique in E: indeed, if x_λ^1 and x_λ^2 satisfy the equation above, testing by $x_\lambda^1 - x_\lambda^2$ and using the monotonicity of A it is readily seen that

$$\langle J(x_\lambda^1 - x) - J(x_\lambda^2 - x), x_\lambda^1 - x_\lambda^2\rangle \leq 0\,,$$

so that, recalling the definition of J,

$$\left(\|x_\lambda^1 - x\|_E - \|x_\lambda^2 - x\|_E\right)^2 \leq \langle J(x_\lambda^1 - x) - J(x_\lambda^2 - x), x_\lambda^1 - x_\lambda^2\rangle \leq 0\,,$$

and $x_\lambda^1 = x_\lambda^2$. Consequently, with this notation, the following definitions make sense.

Definition 1.3.8. *Let E be reflexive and strictly convex with its dual E^*. The resolvent and the Yosida approximation of A are defined, respectively, as*

$$J_\lambda : E \to E\,, \quad J_\lambda(x) := x_\lambda\,, \qquad A_\lambda : E \to E^*\,, \quad A_\lambda(x) := \frac{1}{\lambda} J(x_\lambda - x)\,,$$

for every $x \in E$ and $\lambda > 0$.

The following properties are crucial.

Proposition 1.3.9. *Assume that E and E^* are reflexive and strictly convex, and A is maximal monotone. Then,*

- *$A_\lambda : E \to E^*$ is montone, demicontinuous and bounded;*
- *$A_\lambda(x) \in A(J_\lambda(x))$ for every $x \in E$;*
- *$\|A_\lambda(x)\|_E \leq \|A^0(x)\|_E$, where $A^0(x)$ is the minimum-norm element in $A(x)$;*
- *$J_\lambda : E \to E$ is bounded and*
$$\lim_{\lambda \to 0^+} J_\lambda(x) = x \qquad \forall\, x \in \overline{\operatorname{conv} \mathsf{D}(A)}\,;$$
- *$A_\lambda(x) \overset{*}{\rightharpoonup} A^0(x)$ for every $x \in \mathsf{D}(A)$ as $\lambda \to 0^+$. If E^* is uniformly convex, then the convergence is strong in E^*;*
- *for every $x \in E$, $y \in E^*$ and $(x_n, y_n)_n \subseteq A$ such that $x_n \rightharpoonup x$ in E and $y_n \overset{*}{\rightharpoonup} y$ in E^* as $n \to \infty$, if*
$$\limsup_{n\to\infty} \langle y_n, x_n \rangle_E \leq \langle y, x \rangle_E\,,$$
then $y \in A(x)$.

Proposition 1.3.10. *If $E = H$ is a Hilbert space identified with its dual H^*, then*

- *$J_\lambda = (I + \lambda A)^{-1} : H \to H$ is nonexpansive;*
- *$A_\lambda : H \to H$ is $\frac{1}{\lambda}$-Lipschitz continuous.*

A very large and important class of maximal montone operators is represented by the subdifferentials of proper, convex and l.s.c. functions on Banach spaces. We have indeed the following well-known result.

Proposition 1.3.11. *If $\phi : E \to (-\infty, +\infty]$ is proper, convex and l.s.c., then $\partial\phi : E \to 2^{E^*}$ is maximal montone.*

As the Yosida approximations provide a good regularization of an arbitrary maximal monotone operator, preserving monotonicity for example, in a similar fashion there is a canonical way of regularizing any arbitrary proper, convex and l.s.c. function, preserving fundamental properties such as convexity. In this direction, we have the following definition.

Definition 1.3.12. *Let $\phi : E \to (-\infty, +\infty]$ be proper, convex and l.s.c. For any $\lambda > 0$ the Moreau-Yosida regularization of ϕ is the proper, convex and l.s.c. function defined as*

$$\phi_\lambda : E \to \mathbb{R}\,, \qquad \phi_\lambda(x) := \inf_{w \in E} \left\{ \frac{\|x - w\|_E^2}{2\lambda} + \phi(w) \right\}, \quad x \in E\,.$$

Note that ϕ_λ is actually well-defined and $\mathsf{D}(\phi_\lambda) = E$, so that in particular ϕ_λ is continuous. Moreover, we recall the following properties.

Proposition 1.3.13. *Let E be reflexive and strictly convex with its dual E^* and let $\phi : E \to (-\infty, +\infty]$ be proper, convex and l.s.c. and set $A := \partial\phi$. Then, for any $\lambda > 0$ $\phi_\lambda : E \to \mathbb{R}$ is convex, continuous and Gateaux differentiable with $D\phi_\lambda = A_\lambda : E \to E^*$. Moreover,*

$$\begin{aligned} \phi_\lambda(x) = \frac{\|x - J_\lambda(x)\|_E^2}{2\lambda} + \phi(J_\lambda(x)) \qquad & \forall\, x \in E\,, \quad \forall\, \lambda > 0\,, \\ \phi(J_\lambda(x)) \le \phi_\lambda(x) \le \phi(x) \qquad & \forall\, x \in E\,, \quad \forall\, \lambda > 0\,, \\ \lim_{\lambda \to 0^+} \phi_\lambda(x) = \phi(x) \qquad & \forall\, x \in E\,. \end{aligned}$$

We shall also need a result about passing to the limit "within" maximal monotone graphs due to Brézis, see [21, Theorem 18, p. 126].

Lemma 1.3.14. *Let γ be a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ with $\mathsf{D}(\gamma) = \mathbb{R}$ and $0 \in \gamma(0)$. Assume that the sequences $(y_n)_{n\in\mathbb{N}}$, $(g_n)_{n\in\mathbb{N}}$ of real-valued measurable functions on a finite measure space $(Y, \mathscr{A}, \mu)$ are such that $y_n \to y$ μ-a.e. as $n \to \infty$, $g_n \in \gamma(y_n)$ μ-a.e. for all $n \in \mathbb{N}$, and $(g_n y_n)$ is a bounded subset of $L^1(Y, \mathscr{A}, \mu)$. Then there exists $g \in L^1(Y, \mathscr{A}, \mu)$ and a subsequence n' such that $g_{n'} \to g$ weakly in $L^1(Y, \mathscr{A}, \mu)$ as $n' \to \infty$ and $g \in \gamma(y)$ μ-almost everywhere.*

Finally, we recall a simplified version of an "abstract" Jensen's inequality, due to Haase (see [41, Theorem 3.4]), that will be used to prove a priori estimates for convex functionals of stochastic processes.

Lemma 1.3.15. *Let $(Y, \mathscr{A}, \mu)$, $(Z, \mathscr{B}, \nu)$ be measure spaces, $E \subset L^0(Y, \mathscr{A}, \mu)$ a Banach function space, and*

$$T : E \longrightarrow L^0(Z, \mathscr{B}, \nu)$$

a linear continuous sub-Markovian operator. Moreover, let $\varphi : \mathbb{R} \to [0, \infty[$ be a convex lower semicontinuous function with $\varphi(0) = 0$. Then

$$\varphi(Tf) \le T\varphi(f)$$

for all $f \in E$ such that $\varphi(f) \in E$.

1.4 Continuity and compactness for spaces of vector-valued functions

The following result by Strauss, see [79, Theorem 2.1], provides sufficient conditions for a vector-valued function to be weakly continuous. It will be used to establish the pathwise weak continuity of solutions to several stochastic equations. We recall that, given a Banach space E and an interval $I \subseteq \mathbb{R}$, the space of weakly continuous functions from I to E is denoted by $C_w(I; E)$.

Lemma 1.4.1. *Let E and F be Banach spaces such that E is dense in F, $E \hookrightarrow F$, and E is reflexive. Then*

$$L^\infty(0, T; E) \cap C_w([0, T]; F) = C_w([0, T]; E).$$

The next result is a classical integration-by-parts formula, whose proof can be found, for instance, in [8, §1.3]. Let $\mathcal{V}$ and $\mathcal{H}$ be Hilbert spaces such that $\mathcal{V} \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{V}^*$, and denote by $W(a, b; \mathcal{V})$ the set of functions $u \in L^2(a, b; \mathcal{V})$ such that $u' \in L^2(a, b; \mathcal{V}^*)$, where the derivative

u' is meant in the sense of $\mathcal{V}^*$-valued distributions. The duality of $\mathcal{V}$ and $\mathcal{V}^*$ as well as the scalar product of $\mathcal{H}$ will be denoted by $\langle \cdot, \cdot \rangle$.

Lemma 1.4.2. *Let $u \in W(a,b;\mathcal{V})$. Then there exists $\tilde{u} \in C([a,b];\mathcal{H})$ such that $u(t) = \tilde{u}(t)$ for almost all $t \in [a,b]$. Moreover, for any $v \in W(a,b;\mathcal{V})$, $\langle u, v\rangle$ is absolutely continuous on $[a,b]$ and*

$$\frac{d}{dt}\big\langle u(t), v(t)\big\rangle = \big\langle u'(t), v(t)\big\rangle + \big\langle u(t), v'(t)\big\rangle.$$

The following compactness criterion is due to Simon, see [77, Corollary 4, p. 85].

Lemma 1.4.3. *Let E_1, E_2, E_3 be three Banach spaces such that $E_1 \hookrightarrow E_2$ and $E_2 \hookrightarrow E_3$ compactly. Assume that F is a bounded subset of $L^p(0,T;E_1) \cap W^{1,1}(0,T;E_3)$ for some $p \geq 1$. Then F is relatively compact in $L^p(0,T;E_2)$.*

1.5 Hilbert-Schmidt operators

Let us recall now some standard facts about linear maps. For general definitions and properties of Hilbert-Schmidt operators we refer to [56]. We recall that the space of continuous linear operators from a Banach space E to another one F, equipped with the strong operator topology, is denoted by $\mathscr{L}_s(E,F)$. If E and F are Hilbert spaces, the space of Hilbert-Schmidt operators $\mathscr{L}^2(E,F)$ is an operator ideal, in particular it is stable with respect to pre-composition as well as post-composition with continuous linear operators: if H and K are also Hilbert spaces, and

$$H \xrightarrow{R} E \xrightarrow{T} F \xrightarrow{L} K,$$

with R and L continuous linear operators, then $LTR \in \mathscr{L}^2(H,K)$, with

$$\big\|LTR\big\|_{\mathscr{L}^2(H,K)} \leq \big\|L\big\|_{\mathscr{L}(F,K)}\big\|T\big\|_{\mathscr{L}^2(E,F)}\big\|R\big\|_{\mathscr{L}(H,E)}$$

(see, e.g., [20, p. V.52]). It follows from these properties that, for any $T \in \mathscr{L}^2(E,F)$, the mapping

$$\begin{aligned} \Phi_T : \mathscr{L}_s(F,K) &\longrightarrow \mathscr{L}^2(E,K) \\ L &\longmapsto LT \end{aligned}$$

is continuous: $L_n \to L$ in $\mathscr{L}_s(F,K)$ implies that $L_n T \to LT$ in $\mathscr{L}^2(E,K)$. If E and F are separable, then $\mathscr{L}^2(E,F)$ is itself a separable Hilbert space.

Lemma 1.5.1. *Given two Hilbert spaces U and H, if G is a progressively measurable process with values in $\mathscr{L}^2(U,H)$ such that*

$$\mathbb{E}\int_0^T \big\|G(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds < \infty$$

and F is a progressively measurable H-valued process such that $\mathbb{E}(F_T^*)^2 < \infty$, then, for any $\varepsilon > 0$,

$$\mathbb{E}\big((FG)\cdot W\big)_T^* \leq \varepsilon\,\mathbb{E}\big(F_T^*\big)^2 + N(\varepsilon)\,\mathbb{E}\int_0^T \big\|G(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds.$$

Proof. By the ideal property of Hilbert-Schmidt operators, one has

$$\begin{aligned}\|F(s)G(s)\|_{\mathscr{L}^2(U,\mathbb{R})} &\leq \|F(s)\|_H \|G(s)\|_{\mathscr{L}^2(U,H)} \\ &\leq (F_T^*)\|G(s)\|_{\mathscr{L}^2(U,H)}\end{aligned}$$

for all $s \in [0,T]$, hence

$$\int_0^T \|F(s)G(s)\|^2_{\mathscr{L}^2(U,\mathbb{R})}\,ds \leq (F_T^*)^2 \int_0^T \|G(s)\|^2_{\mathscr{L}^2(U,H)}\,ds,$$

where the right-hand side is finite $\mathbb{P}$-a.s. thanks to the assumptions on F and G. Then $(FG)\cdot W$ is a local martingale, for which Davis' inequality yields

$$\begin{aligned}\mathbb{E}\big((FG)\cdot W\big)^*_T &\lesssim \mathbb{E}\big[(FG)\cdot W, (FG)\cdot W\big]^{1/2}_T \\ &= \mathbb{E}\left(\int_0^T \|F(s)G(s)\|^2_{\mathscr{L}^2(U,\mathbb{R})}\,ds\right)^{1/2} \\ &\leq \mathbb{E}(F_T^*)\left(\int_0^T \|G(s)\|^2_{\mathscr{L}^2(U,H)}\,ds\right)^{1/2}.\end{aligned}$$

The proof is finished invoking the elementary inequality

$$ab \leq \frac{1}{2}\big(\varepsilon a^2 + \frac{1}{\varepsilon} b^2\big) \qquad \forall a, b \in \mathbb{R}.$$

□

1.6 Classical variational approach to SPDEs

Let us recall the classical variational approach to stochastic evolution equations: the reader may refer to [56].

Given a positive real number T, let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space endowed with a filtration $(\mathscr{F}_t)_{t\in[0,T]}$ which is saturated and right-continuous. Let us consider also a separable Hilbert space H and a separable Banach space V which is continuously and densely embedded in H. Then, if we identify H with its dual H^* through the Riesz isomorphism, we have the following dense and continuous inclusions:

$$V \hookrightarrow H \hookrightarrow V^*.$$

Let $W = (W(t))_{t\in[0,T]}$ be a cylindrical Wiener process on a separable Hilbert space U, and consider some operators

$$A : \Omega \times [0,T] \times V \to V^*, \qquad B : \Omega \times [0,T] \times V \to \mathscr{L}^2(U,H)$$

which are progressively measurable, i.e. the restrictions of A and B to $\Omega \times [0,t] \times V$ are $\mathscr{F}_t \otimes \mathscr{B}([0,t]) \otimes \mathscr{B}(V)$-measurable for every $t \in [0,T]$.

In this setting, we are interested in solving stochastic evolution equations in the following variational form:

$$dX(t) + A(t, X(t))\,dt = B(t, X(t))\,dW(t), \qquad X(0) = X_0.$$

The classical variational approach to this type of stochastic evolution equations requires the following assumption on A and B.

(i) (Hemicontinuity). For every $x, y, z \in V$ and $(\omega, t) \in \Omega \times [0, T]$, the map

$$s \mapsto \langle A(\omega, t, x + sy), w \rangle, \qquad s \in \mathbb{R},$$

is continuous.

(ii) (Weak monotonicity). There exists a constant $k > 0$ such that, for every $x, y \in V$ and $(\omega, t) \in \Omega \times [0, T]$,

$$\begin{aligned}\langle A(\omega, t, x) - A(\omega, t, y), x - y \rangle &- \frac{1}{2}\|B(\omega, t, x) - B(\omega, t, y)\|^2_{\mathscr{L}^2(U,H)} \\ &\geq -k\|x - y\|^2.\end{aligned}$$

(iii) (Weak coercivity). There exist constants $k_1, k_2 > 0$, $p > 1$ and an adapted process $f \in L^1(\Omega \times (0, T))$ such that, for every $x \in V$ and $(\omega, t) \in \Omega \times [0, T]$,

$$\langle A(\omega, t, x), x \rangle - \frac{1}{2}\|B(\omega, t, x)\|^2_{\mathscr{L}^2(U,H)} \geq k_1\|x\|^p_V - k_2\|x\|^2 - f(\omega, t).$$

(iv) (Weak boundedness). There exists a constant $k_3 > 0$ and an adapted process $g \in L^1(\Omega \times (0, T))$ such that, for every $x \in V$ and $(\omega, t) \in \Omega \times [0, T]$,

$$\|A(\omega, t, x)\|^{\frac{p}{p-1}}_{V^*} \leq k_3\|x\|^p_V + g(\omega, t).$$

In this setting, we have the following main results.

Proposition 1.6.1. *Let A and B satisfy conditions (i)–(iv) and let $X_0 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P}; H)$. Then, there exists a unique adapted process*

$$X \in L^2(\Omega; C([0, T]; H)) \cap L^p(\Omega \times (0, T); V)$$

such that

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

Proposition 1.6.2. *Let $Y_0 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P}; H)$, $p > 1$ and Z, G be progressively measurable processes such that*

$$Z \in L^{\frac{p}{p-1}}(\Omega \times (0, T); V^*), \qquad G \in L^2(\Omega; L^2(0, T); \mathscr{L}^2(U, H)).$$

If a process Y satysfies

$$Y \in L^p(\Omega \times (0, T); V), \qquad Y(t) \in L^2(\Omega; H) \quad \textit{for a.e. } t \in (0, T)$$

and

$$Y(t) + \int_0^t Z(s)\, ds = Y_0 + \int_0^t G(s)\, dW(s) \quad \forall\, t \in [0, T], \quad \mathbb{P}\text{-}a.s.,$$

then $Y \in L^2(\Omega; C([0,T]; H))$ *and the following Itô's formula holds:*

$$\frac{1}{2}\|Y(t)\|^2 + \int_0^t \langle Z(s), Y(s)\rangle\, ds = \frac{1}{2}\|Y_0\|^2 + \frac{1}{2}\int_0^t \|G(s)\|^2_{\mathscr{L}^2(U,H)}\, ds + \int_0^t Y(s)G(s)\, dW(s) \quad \forall\, t \in [0,T], \quad \mathbb{P}\text{-}a.s.$$

1.7 Tightness, Markovian semigroups, ergodicity

Let $(E, \mathscr{E})$ be a measurable space. The set of probability measures on $(E, \mathscr{E})$ 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., [19, §5.5]).

Theorem 1.7.1 (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.*

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

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

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

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

The notations $P_t f$ and μP_t, with f $\mathscr{E}$-measurable bounded or positive function and μ positive measure on $\mathscr{E}$, 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 $\mathscr{B}(\mathbb{R}_+) \otimes \mathscr{E}$-measurable.

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

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

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 $\nu \in \mathscr{M}_1(E)$, 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 1.7.2 (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 $\nu \in \mathscr{M}_1(E)$ such that the $(\mu_t^\nu)_{t\geq 0} \subset \mathscr{M}_1(E)$ 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. This latter condition is satisfied, for example, if $(P_t(x,\cdot))_{t\geq 0} \subset \mathscr{M}_1(E)$ is tight.

Chapter 2

Singular semilinear equations: global well-posedness

In this chapter, we prove global well-posedness for a class of dissipative semilinear stochastic evolution equations with singular drift and multiplicative Wiener noise. In particular, the nonlinear term in the drift is the superposition operator associated to a maximal monotone graph everywhere defined on the real line, on which neither continuity nor growth assumptions are imposed. The hypotheses on the diffusion coefficient are also very general, in the sense that the noise does not need to take values in spaces of continuous, or bounded, functions in space and time. Our approach combines variational techniques with a priori estimates, both pathwise and in expectation, on solutions to regularized equations.

The results presented in this chapter are part of a joint work with Carlo Marinelli, to apper on *Annals of Probability*: see [65].

2.1 The problem: literature and main goals

Our aim is to establish existence and uniqueness of solutions, and their continuous dependence on the initial datum, to the following semilinear stochastic evolution equation on $L^2(D)$, with $D \subset \mathbb{R}^d$ a bounded domain:

$$dX(t) + AX(t)\,dt + \beta(X(t))\,dt \ni B(t, X(t))\,dW(t), \qquad X(0) = X_0, \tag{2.1.1}$$

where A is a linear maximal monotone operator on $L^2(D)$ associated to a coercive Markovian bilinear form, β is a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ defined everywhere, 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 $L^2(D)$ and satisfies suitable Lipschitz continuity assumptions. Precise assumptions on the data of the problem and on the definition of solution are given below. Since any increasing function $\beta_0 : \mathbb{R} \to \mathbb{R}$ can be extended in a canonical way to a maximal monotone graph of $\mathbb{R} \times \mathbb{R}$ by "filling the gaps" (i.e., setting $\beta(x) := [\beta_0(x^-), \beta_0(x^+)]$ for all $x \in \mathbb{R}$, where $\beta(x^-)$ and $\beta(x^+)$ denote the limit from the left and from the right of β_0 at x, respectively), Equation (2.1.1) can be interpreted as a formulation of the stochastic evolution equation

$$dX(t) + AX(t)\,dt + \beta_0(X(t))\,dt = B(t, X(t))\,dW(t), \qquad X(0) = X_0.$$

Semilinear equations with singular and rapidly growing drift appear, for instance, in mathematical models of Euclidean quantum field theory (see, e.g., [1] for an equation with exponentially growing drift), and, most importantly for us, cannot be directly treated with the existing methods, hence are interesting from a purely mathematical perspective as well. In particular, the variational approach (see [46, 72]) works only assuming that β satisfies suitable polynomial growth conditions depending on the dimension n of the underlying Euclidean space (see also [56, pp. 137-ff.] for improved sufficient conditions, still dependent on the dimension), whereas most available results relying on the semigroup approach require just polynomial growth, although usually compensated by rather stringent hypotheses on the noise (see, e.g., [27, 28]). Under natural assumptions on the noise, well-posedness in L^p spaces is proven, with different methods, in [47], under the further assumption that β is locally Lipschitz continuous, and in [58]. A common basis for both works is the semigroup approach on UMD Banach spaces. A special mention deserves the short note [9], where the author considers problem (2.1.1) with $A = -\Delta$ and B independent of X, and proves existence of a pathwise solution* assuming that the solution Z to the equation with $\beta \equiv 0$ (i.e., the stochastic convolution) is jointly continuous in space and time. Furthermore, assuming that

$$\mathbb{E}\int_0^T\int_D j(Z) < \infty,$$

where j is a primitive of β, he obtains that the pathwise solution may admit a version that can be considered as a generalized mild solution to (2.1.1). This is the only result we are aware of about existence of solutions to stochastic semilinear parabolic equations *without* growth assumptions on the drift in any dimension.

It is well known that a well-posedness theory for stochastic evolution equations on a Hilbert space H of the type

$$du + Au\,dt \ni B(u)\,dW, \qquad u(0) = u_0,$$

with A an arbitrary (nonlinear) maximal monotone operator, is, in full generality, not yet available, even if B does not depend on u and is a fixed non-random operator. However, a satisfactory treatment in the finite-dimensional case has been given by Pardoux and Răşcanu in [73, §4.2], where the authors consider stochastic differential equations in $\mathbb{R}^n$ of the type

$$dX_t + A(X_t)\,dt + F(t, X_t)\,dt \ni G(t, X_t)\,dB_t,$$

where A is a (multivalued) maximal monotone operator whose domain has non-emtpy interior, B is a k-dimensional Wiener process, G satisfies standard Lipschitz continuity assumptions, and $F(t,\cdot)$ is continuous and monotone (not necessarily Lipschitz continuous). While the assumptions on A are not restrictive in finite dimensions, unbounded linear operators generating contraction semigroups in infinite-dimensional spaces, as in our case, have dense domain, whose interior is hence empty.

On the other hand, in the deterministic setting complete results have long been known for equations of the type

$$\frac{du}{dt} + Au \ni f, \qquad u(0) = u_0,$$

even in the much more general setting where A is a (multivalued) m-accretive operator on a Banach space E and $f \in L^1(0, T; E)$ (see, e.g., [8, 22]). Although a solution to the general

*To avoid misunderstandings, we should clarify once and for all that with this expression we do *not* refer to a solution in the sense of rough paths, but simply "with ω fixed".

stochastic problem does not currently seem within reach, significant results have been obtained in special cases: apart of the above-mentioned works on semilinear equations, well-posedness for the stochastic porous media equation under fairly general assumptions is known (see [12], where the same hypotheses on β imposed here are used and the noise is assumed to satisfy suitable boundedness conditions, and [13] for an extension to jump noise). Moreover, the variational theory by Pardoux, Krylov and Rozovskiĭ is essentially as complete as the corresponding deterministic theory. As mentioned above, however, large classes of maximal monotone operators on $H = L^2(D)$ cannot be cast in the variational framework.

The main contribution of this chapter is a well-posedness result for (2.1.1) under the most general conditions known so far, to the best of our knowledge. These conditions are quite sharp for A, but not for β. In particular, the conditions on A are close to those needed to show that $A + \beta(\cdot)$ is maximal monotone on $L^2(D)$, but the hypothesis that β is finite on the whole real line is not needed in the deterministic theory. Finally, the conditions on B are the natural ones to have function-valued noise, and are in this sense as general as possible. Equations with white noise in space and time, that have received much attention lately, are not within the scope of our approach (nor of others, most likely, under such general conditions on β).

Let us now briefly outline the structure of the chapter and the main ideas of the proof. Section 2.2 contains the statement of the main well-posedness result, and in Section 2.3 we discuss the hypotheses on the drift and diffusion coefficients, providing corresponding examples. In Section 2.4 we consider a version of equation (2.1.1) with additive noise satisfying a strong boundedness assumption. Using the Yosida regularization of β, we obtain a family of approximating equations with Lipschitz coefficients, which can be treated by the standard variational theory. The solutions to such equations are shown to satisfy suitable uniform estimates, both pathwise and in expectation. Such estimates allow us to obtain key regularity and integrability properties for the solution to the equation with additive bounded noise. A crucial role is played by Simon's compactness criterion, which is applied pathwise, and by compactness criteria in L^1 spaces, applied both pathwise and in expectation. It is, in essence, precisely this interplay between pathwise and "averaged" arguments that permits to avoid many restrictive hypotheses of the existing literature. An abstract version of Jensen's inequality for positive operators, combined with the lower semicontinuity of convex integrals, is also an essential tool. In Section 2.5 we prove well-posedness for equations with additive noise removing the boundedness assumption of the previous section. This is accomplished by a further regularization scheme, this time on the diffusion operator B, and by a priori estimates for solutions to the regularized equations. A key role is played again by a combination of estimates and passages to the limit both pathwise and in expectation. We also prove continuity of the solution map with respect to the initial datum and the diffusion coefficient, by means of Itô's formula and regularizations, for which smoothing properties of the resolvent of A are essential. Finally, in Section 2.6 we obtain well-posedness in the general case by a fixed-point argument, using the Lipschitz continuity of B only. Introducing weighted spaces of stochastic processes, we obtain directly global well-posedness, thus avoiding a tedious construction by "patching" local solutions.

Some tools and reasonings used in this chapter are obviously not new: weak compactness arguments in L^1, for instance, are extensively used in the literature on partial differential equations (see, e.g., [17, 21] and references therein), as well as, to a lesser extent, in the stochastic setting (cf. [9,12,60]). However, even where similarities are present, our arguments are considerably streamlined and more general. The pathwise application of Simon's compactness criterion, made possible by a construction based on the variational framework, seems to be new, at least in the context of stochastic evolution equations. It is in fact somewhat surprising that the variational setting, which notoriously fails when dealing with semilinear equations, is at a basis of an approach that leads to well-posedness of those same equations, even with singular and rapidly increasing drift.

2.2 Main results

In this section, after fixing notation and conventions used throughout the chapter, we state our main result.

2.2.1 Notation and assumptions

All functional spaces will be defined on a smooth bounded domain $D \subset \mathbb{R}^d$. We shall denote $L^2(D)$ by H and its inner product by $\langle \cdot, \cdot \rangle$.

All random quantities will be defined on a fixed probability space $(\Omega, \mathscr{F}, \mathbb{P})$ endowed with a right-continuous and saturated filtration $\mathbb{F} := (\mathscr{F}_t)_{t\in[0,T]}$, where T is a positive number. All expressions involving random quantities are meant to hold $\mathbb{P}$-almost surely, unless otherwise stated. With W we shall denote a cylindrical Wiener process on a separable Hilbert space U, that may coincide with H, but does not have to.

The following assumptions on the data of the problem are assumed to be in force throughout and will not always be recalled explicitly.

Assumption A. Let V be Hilbert space that is densely, continuously, and compactly embedded in H. The linear operator A belongs to $\mathscr{L}(V, V^*)$ and satisfies the following properties:

(i) there exists $C > 0$ such that

$$\langle Av, v\rangle \geq C\|v\|_V^2 \qquad \forall v \in V;$$

(ii) the part of A in H admits a unique m-accretive extension A_1 in $L^1(D)$;

(iii) the resolvent $\big((I + \lambda A_1)^{-1}\big)_{\lambda>0}$ is sub-Markovian;

(iv) there exists $m \in \mathbb{N}$ such that

$$\big\|(I + A_1)^{-m}\big\|_{\mathscr{L}(L^1(D),L^\infty(D))} < \infty.$$

Here we have used $\langle \cdot, \cdot \rangle$ also to denote the duality pairing of V and V^*, which is compatible with the scalar product in H. In fact, identifying H with its dual, one has the so-called Gel'fand triple

$$V \hookrightarrow H \hookrightarrow V^*,$$

where both embeddings are dense (see, e.g., [50, §2.9]). Moreover, we recall that the part of A in H is the operator A_2 on H defined as $\mathsf{D}(A_2) := \{x \in V : Au \in H\}$ and $A_2x := Ax$ for all $x \in \mathsf{D}(A_2)$. If one identifies the operators with their graphs, this is equivalent to setting $A_2 := A \cap (V \times H)$. We shall often refer to condition (i) as the coercivity of A. The sub-Markovianity condition (iii) amounts to saying that, for all functions $f \in L^1(D)$ such that $0 \leq f \leq 1$, one has

$$0 \leq (I + A_1)^{-1} f \leq 1.$$

In other words, $(I + A_1)^{-1}$ is positivity preserving and contracting in $L^\infty(D)$.

From Section 2.4 onwards, we shall often use the symbol A to denote also A_1 and A_2.

Let us observe that if A is the negative Laplacian with Dirichlet boundary conditions, all hypotheses are met. Much wider classes of operators satisfying hypotheses (i)-(iv) will be given below.

Assumption B. β is a maximal monotone graph of $\mathbb{R} \times \mathbb{R}$ such that $\mathsf{D}(\beta) = \mathbb{R}$, $0 \in \beta(0)$, and its potential j is even.

We recall that the potential j of β is the convex, proper, lower semicontinuous function $j : \mathbb{R} \to \mathbb{R}_+$, with $j(0) = 0$, such that $\partial j = \beta$, where ∂ stands for the subdifferential in the sense of convex analysis.

Assumption C. The diffusion coefficient

$$B : \Omega \times [0,T] \times H \to \mathscr{L}^2(U,H)$$

is Lipschitz continuous and grows linearly in its third argument, uniformly over $\Omega \times [0,T]$, i.e., there exist constants L_B, N_B such that

$$\begin{aligned} \big\|B(\omega,t,x) - B(\omega,t,y)\big\|_{\mathscr{L}^2(U,H)} &\leq L_B \|x-y\|_H, \\ \big\|B(\omega,t,x)\big\|_{\mathscr{L}^2(U,H)} &\leq N_B\big(1 + \|x\|_H\big) \end{aligned}$$

for all $\omega \in \Omega$, $t \in [0,T]$, and x, $y \in H$. Moreover, $B(\cdot,\cdot,x)$ is progressively measurable for all $x \in H$, i.e., for all $t \in [0,T]$, the map $(\omega,s) \mapsto B(\omega,s,x)$ from $\Omega \times [0,t]$, endowed with the σ-algebra $\mathscr{F}_t \otimes \mathscr{B}([0,t])$, to $\mathscr{L}^2(U,H)$, endowed with its Borel σ-algebra, is strongly measurable. We recall that, since U and H are separable, the space of Hilbert-Schmidt operators $\mathscr{L}^2(U,H)$ is itself a separable Hilbert space, hence strong and weak measurability coincide. Whenever we deal with maps with values in separable Banach spaces, since strong and weak measurability coincide, we shall drop the qualifier "strong".

2.2.2 The well-posedness result

Definition 2.2.1. *Let X_0 be an H-valued $\mathscr{F}_0$-measurable random variable. A* strong solution *to the stochastic equation* (2.1.1) *is a pair (X,ξ) satisfying the following properties:*

(i) *X is a measurable adapted V-valued process such that $AX \in L^1(0,T;V^*)$ and $B(\cdot,X) \in L^2(0,T;\mathscr{L}^2(U,H))$;*

(ii) *ξ is a measurable adapted $L^1(D)$-valued process such that $\xi \in L^1(0,T;L^1(D))$ and $\xi \in \beta(X)$ almost everywhere in $(0,T) \times D$;*

(iii) *one has, as an equality in $L^1(D) \cap V^*$,*

$$X(t) + \int_0^t AX(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s,X(s))\,dW(s)$$

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

Note that $L^1(D) \cap V^*$ is not empty because D has finite Lebesgue measure, hence, for instance, H is contained in both spaces.

Let us denote by $\mathscr{J}$ the set of pairs (ϕ,ζ), where ϕ and ζ are measurable adapted processes with values in H and $L^1(D)$, respectively, such that

$$\begin{aligned} \phi &\in L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V)), \\ \zeta &\in L^1(\Omega \times [0,T] \times D), \\ j(\phi) + j^*(\zeta) &\in L^1(\Omega \times [0,T] \times D). \end{aligned}$$

We shall say that (2.1.1) is well posed in $\mathscr{J}$ if there exists a unique process in $\mathscr{J}$ which is a strong solution and such that the solution map $X_0 \mapsto X$ is continuous from $L^2(\Omega; H)$ to $L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V))$.

The central result of this chapter is the following.

Theorem 2.2.2. *Let $X_0 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P}; H)$. Then* (2.1.1) *is well-posed in $\mathscr{J}$. Moreover, the solution map $X_0 \mapsto X$ is Lipschitz continuous and the paths of X are weakly continuous with values in H.*

Let us stress the fact that the more general problem of unconditional well-posedness (i.e. without the extra condition that strong solutions belong to $\mathscr{J}$) remains open and is beyond the scope of the techniques used in this chapter. In particular, we can only prove uniqueness of solutions within $\mathscr{J}$.

2.3 Examples and remarks

Some comments and examples on the assumptions on the data of the problem are in order. In particular, the hypotheses on A deserve special attention. The coercivity condition $\langle Av, v\rangle \geq C\|v\|_V^2$ for all $v \in V$ is equivalent to $A \in \mathscr{L}(V, V^*)$ being determined by a bounded V-elliptic[†] bilinear form $\mathscr{E} : V \times V \to \mathbb{R}$, i.e. such that

$$|\mathscr{E}(u,v)| \lesssim \|u\|_V \|v\|_V, \qquad \mathscr{E}(v,v) \geq C\|v\|_V^2 \qquad \forall u, v \in V.$$

This is an immediate consequence of the Lax-Milgram theorem, which also implies that A is an isomorphism between V and V^* (see, e.g., [7, §5.2] or [70, Lemma 1.3]).

The bilinear form $\mathscr{E}$ can also be seen as a closed unbounded form on H with domain V. This defines a (unique) linear m-accretive operator A_2 on H, that is nothing else than the part of A in H (see, e.g., [7, §5.3] or [70, p. 34]). Conversely, given a positive closed bilinear form $\mathscr{E}$ on H with dense domain $\mathsf{D}(\mathscr{E})$ satisfying the *strong sector condition*[‡]

$$|\mathscr{E}(u,v)| \lesssim \mathscr{E}(u,u)^{1/2} \mathscr{E}(v,v)^{1/2} \qquad \forall u, v \in \mathsf{D}(\mathscr{E}),$$

and such that $\mathscr{E}(u,u) > 0$ for all $u \in \mathsf{D}(\mathscr{E})$, $u \neq 0$, setting $V := \mathsf{D}(\mathscr{E})$ with inner product given by the symmetric part $\mathscr{E}^s$ of $\mathscr{E}$, that is

$$\mathscr{E}^s(u,v) := \frac{1}{2}\big(\mathscr{E}(u,v) + \mathscr{E}(v,u)\big), \qquad u, v \in \mathsf{D}(\mathscr{E}),$$

there is a unique linear operator $A \in \mathscr{L}(V, V^*)$ such that $\mathscr{E}(u,v) = \langle Au, v\rangle$ for all $u, v \in V$. This amounts to trivial verifications, since, obviously, $\mathscr{E}(u,u) = \mathscr{E}^s(u,u)$ for all $u \in \mathsf{D}(\mathscr{E})$. As a particular case, let A' be a linear positive self-adjoint (unbounded) operator H such that

[†] We prefer this terminology, taken from [50], over the currently more common "V-coercive", to avoid possible confusion with related terminology used in the theory of Dirichlet forms, where coercivity is meant in a somewhat different sense (cf. [57, Definition 2.4, p. 16]).

[‡] Throughout this section we shall follow the terminology on Dirichlet forms of [57].

$\langle A'u, u\rangle > 0$ for all $u \in \mathsf{D}(A)$, $u \neq 0$. Then A' admits a square root $\sqrt{A'}$, which is in turn a linear positive self-adjoint operator on H. One can then define the Hilbert space $V := \mathsf{D}(\sqrt{A'})$, endowed with the inner product

$$\langle u, v\rangle_V := \big\langle \sqrt{A'}u, \sqrt{A'}v \big\rangle,$$

and the symmetric bounded bilinear form $\mathscr{E} : V \times V \to \mathbb{R}$,

$$\mathscr{E}(u, v) := \big\langle \sqrt{A'}u, \sqrt{A'}v \big\rangle, \qquad u, v \in V,$$

which is obviously V-elliptic. By a theorem of Kato ([44, Theorem 2.23, p. 331]), there is in fact a bijective correspondence between linear positive self-adjoint operators on H and positive densely-defined closed symmetric bilinear forms. More generally, if A' is a linear (unbounded) m-accretive operator on H such that

$$\big|\langle A'u, v\rangle\big| \lesssim \langle A'u, u\rangle^{1/2}\langle A'v, v\rangle^{1/2} \qquad \forall u, v \in \mathsf{D}(A'),$$

and $\langle A'u, u\rangle > 0$ for all $u \in \mathsf{D}(A')$, $u \neq 0$, then there exists a (unique) closed V-elliptic bilinear form $\mathscr{E}$ that determines an operator $A \in \mathscr{L}(V, V^*)$, with $V := \mathsf{D}(\mathscr{E})$ and $\langle \cdot, \cdot\rangle_V := \mathscr{E}^s$, such that A' is the part on H of A. This follows, for instance, by [57, p. 27].

Note, however, that in the previous examples V may *not* be continuously embedded in H, unless $\mathscr{E}$ satisfies a Poincaré inequality, i.e. $\|u\|_H^2 \lesssim \mathscr{E}(u, u)$ for all $u \in \mathsf{D}(\mathscr{E})$ (as is the case, for instance, for the Dirichlet Laplacian). This limitation is resolved by the following important observation: our well-posedness result continues to hold if we assume, in place of hypothesis (i), the following weaker one:

(i') there exist constants $C_1 > 0$, $C_2 \in \mathbb{R}$ such that

$$\langle Av, v\rangle \geq C_1\|v\|_V^2 - C_2\|v\|_H^2 \qquad \forall v \in V,$$

which is clearly equivalent to assuming that $\tilde{A} := A + C_2 I$ is V-elliptic. Under this assumption, equation (2.1.1) can equivalently be written as

$$dX(t) + \tilde{A}X(t)\,dt + \beta(X(t))\,dt = C_2 X(t)\,dt + B(t, X(t))\,dW(t).$$

The only added complication in the proofs to follow would be the appearance of functional spaces with an exponential weight in time, very much as in the proof of Proposition 2.5.2 below. An analogous argument, in a slightly different context, is developed in detail in [58]. This seemingly trivial observation allows to considerably extend the class of operators A that can be treated. For instance, one has the following criterion.

Lemma 2.3.1. *A coercive closed form $\mathscr{E}$ on H uniquely determines an operator A satisfying* (i').

Proof. The hypothesis of the Lemma means that $\mathscr{E}$ is a densely defined bilinear form such that its symmetric part $\mathscr{E}^s$ is closed and $\mathscr{E}$ satisfies the *weak sector condition*

$$\big|\mathscr{E}_1(u, v)\big| \lesssim \mathscr{E}_1(u, u)^{1/2}\mathscr{E}_1(v, v)^{1/2} \qquad \forall u, v \in \mathsf{D}(\mathscr{E}),$$

where $\mathscr{E}_1 := \mathscr{E} + I$. In other words, $\mathscr{E}$ satisfies the weak sector condition if the shifted form $\mathscr{E} + I$ satisfies the strong sector condition. Therefore, adapting in the obvious way an argument used above, it is enough to take $V := \mathsf{D}(\mathscr{E})$ with inner product $\langle\cdot,\cdot\rangle_V := \langle\cdot,\cdot\rangle_H + \mathscr{E}^s$ to obtain that the generator A_2 of $\mathscr{E}$ can be (uniquely) extended to an operator $A \in \mathscr{L}(V, V^*)$ satisfying (i') with $C_1 = C_2 = 1$. □

Note that in all the above constructions one has $V \hookrightarrow H$ densely and continuously (under appropriate assumptions), but the embedding is *not* necessarily compact. The latter condition has to be proved depending on the situation at hand. For a general compactness criterion in terms of ultracontractivity properties, see Proposition 2.3.3 below.

As regards condition (ii), the simplest sufficient condition ensuring that A_2 admits an m-accretive extension A_1 in $L^1(D)$ is that $-A_2$ is the generator of a symmetric Markovian semigroup of contractions S_2 on H, or, equivalently, that A_2 is positive self-adjoint with a Markovian resolvent. In fact, this implies that, for any $p \in [1,\infty[$, there exists a (unique) symmetric Markovian semigroup of contractions S_p on $L^p(D)$ such that all S_p, $1 \le p < \infty$, are consistent, hence the corresponding negative generators A_p coincide on the intersections of their domains (see, e.g., [32, Theorem 1.4.1]). In the general case, i.e. if A_2 is not self-adjoint, the same conclusion remains true if the semigroup S_2 and its adjoint S_2^* are both sub-Markovian, or, equivalently, if S_2 is sub-Markovian and L^1-contracting (cf. [7, Lemma 10.13 and Theorem 10.15] or [70, Corollary 2.16]). In particular, if A_2 is the generator of a Dirichlet form on H, these conclusions hold. Moreover, since the resolvent of A_1 is sub-Markovian if and only if the resolvent of A_2 is sub-Markovian, we obtain the following complement to the previous Lemma.

Lemma 2.3.2. *A Dirichlet form $\mathscr{E}$ on H uniquely determines an operator A satisfying* (i'), (ii), *and* (iii).

Without assuming that S_2^* is sub-Markovian (which is the case, for instance, if A is determined by a semi-Dirichlet form on H, so that (i') and (iii) only are satisfied), we note that $D(A_2)$ is dense in $L^1(D)$, and the image of $I + A_2$ is dense in $L^1(D)$: the former assertion follows by $D(A_2) \subset L^2(D)$ densely and $L^2(D) \subset L^1(D)$ densely and continuously. Moreover, since A_2 generates a contraction semigroup in $L^2(D)$, the Lumer-Phillips theorem (see, e.g., [36, p. 83]) implies that $\mathsf{R}(I + A_2) = L^2(D)$, hence $\mathsf{R}(I + A_2)$ is dense in $L^1(D)$. The Lumer-Phillips theorem again guarantees that the closure of A_2 in $L^1(D)$ is m-accretive if A_2 is accretive in $L^1(D)$. The latter property is often not difficult to verify in concrete examples.

The most delicate condition is (iv), i.e. the ultracontractivity of suitable powers of the resolvent of A_1. If A_2 is self-adjoint, a simple duality arguments shows that, for any $t \ge 0$,

$$\big\|S_2(t)\big\|_{\mathscr{L}(L^1,L^\infty)} \le \big\|S_2(t/2)\big\|^2_{\mathscr{L}(L^2,L^\infty)}.$$

Sufficient conditions for $S_2(t)$ to be bounded from $L^2(D)$ to $L^\infty(D)$ are known in terms, for instance, of logarithmic Sobolev inequalities, Sobolev inequalities, and Nash inequalities (see, e.g., [32, Chapter 2] and [70, Chapter 6]). The non-symmetric case is more difficult, but ultracontractivity estimates are known in many special cases, such as in the examples that we are going to discuss next. Ultracontractivity estimates for powers of the resolvent can then be obtained from estimates for the semigroup, as explained below. The following result (probably known, but for which we could not find a reference) shows that hypothesis (iv) guarantees that the embedding $\mathsf{D}(\mathscr{E}) \hookrightarrow H$ is compact, thus answering a question left open above.

Proposition 2.3.3. *Let A_2 be the generator of a closed coercive form $\mathscr{E}$ in H. If there exists $m \in \mathbb{N}$ such that the m-th power of the resolvent of A_2 is bounded from $L^2(D)$ to $L^\infty(D)$, then $\mathsf{D}(\mathscr{E})$ is compactly embedded in H.*

Proof. Let $(u_k)_k$ be a bounded sequence in $\mathsf{D}(\mathscr{E})$, i.e., there exists a constant N such that

$$\|u_k\|_H^2 + \mathscr{E}^s(u_k, u_k) < N \qquad \forall k \in \mathbb{N}.$$

In particular, there exists a subsequence of k, denoted by the same symbol, such that u_k converges weakly to u in H as $k \to \infty$. The goal is to show that the convergence is in fact strong. Since $\mathsf{D}(A_2^m) \subset L^\infty(D)$ by assumption, it follows by a result of Arendt and Bukhvalov, see [6, Theorem 4.16(b)], that the resolvent $J_\lambda := (I + \lambda A_2)^{-1}$ is a compact operator on H for all $\lambda > 0$. The triangle inequality yields

$$\|u_k - u\| \le \|u_k - J_\lambda u_k\| + \|J_\lambda u_k - J_\lambda u\| + \|J_\lambda u - u\|,$$

where the second term on the right-hand side converges to zero as $k \to \infty$ by compactness of J_λ. Moreover, since $J_\lambda \to I$ in $\mathscr{L}_s(H,H)$ as $\lambda \to 0$, the third term on the right-hand side can be made arbitrarily small. Therefore we only have to bound the first term on the right-hand side: note that $I - J_\lambda = \lambda A_\lambda$, where A_λ, $\lambda > 0$, stands for the Yosida approximation of A_2, hence $\|u_k - J_\lambda u_k\| = \lambda \|A_\lambda u_k\|$, and

$$\begin{aligned}\langle A_\lambda u_k, u_k\rangle &= \langle A_\lambda u_k, u_k - J_\lambda u_k + J_\lambda u_k\rangle = \lambda \|A_\lambda u_k\|^2 + \langle A_\lambda u_k, J_\lambda u_k\rangle \\ &\ge \lambda \|A_\lambda u_k\|^2,\end{aligned}$$

where we have used, in the last step, the identity $A_\lambda = A_2 J_\lambda$ and the monotonicity of A_2. Since, by [57, Lemma 2.11(iii), p. 20], one has

$$\big|\mathscr{E}_1^{(\lambda)}(u,v)\big| \lesssim \mathscr{E}_1(u,u)^{1/2} \mathscr{E}_1^{(\lambda)}(v,v)^{1/2} \qquad \forall u \in \mathsf{D}(\mathscr{E}),\ v \in H,$$

where $\mathscr{E}^{(\lambda)}(u,v) := \langle A_\lambda u, v\rangle$, $u, v \in H$, and the implicit constant depends only on $\mathscr{E}$, it follows that

$$\mathscr{E}_1^{(\lambda)}(u,u) \lesssim \mathscr{E}_1(u,u) \qquad \forall u \in \mathsf{D}(\mathscr{E}),$$

hence

$$\|u_k - J_\lambda u_k\|^2 = \lambda^2 \|A_\lambda u_k\|^2 \le \lambda \langle A_\lambda u_k, u_k\rangle = \lambda \mathscr{E}_1^{(\lambda)}(u_k, u_k) \lesssim \lambda \mathscr{E}_1(u_k, u_k).$$

By the assumptions on the sequence (u_k),

$$\mathscr{E}_1(u_k, u_k) = \|u_k\|^2 + \mathscr{E}(u_k, u_k) = \|u_k\|^2 + \mathscr{E}^s(u_k, u_k)$$

is bounded uniformely over k, hence $\|u_k - J_\lambda u_k\|^2$ can be made arbitrarily small as well, thus proving the claim. □

Let us now consider some concrete examples: we first consider the case of A being a suitable “realization” of a second-order differential operator, and then of a nonlocal operator.

Example 2.3.4 (Symmetric divergence-form operators)**.** Consider the bilinear form $\mathscr{E}$ on $V := H_0^1(D)$ defined by

$$\mathscr{E}(u,v) := \big\langle a\nabla u, \nabla v\big\rangle = \sum_{j,k=1}^n a_{jk} \partial_j u \partial_k v,$$

where $a = (a_{jk})$ with $a_{jk} \in L^\infty(D)$ for all j, k, and $a_{jk} = a_{kj}$. The (formal) differential operator associated to $\mathscr{E}$ is

$$A_0 u := -\operatorname{div}\big(a\nabla u\big), \qquad u \in C_c^\infty(D),$$

where $C_c^\infty(D)$ stands for the set of infinitely differentiable functions with compact support contained in D. The form $\mathscr{E}$ is V-elliptic if there exists $C > 0$ such that $\langle a\xi, \xi\rangle \geq C|\xi|^2$ for all $\xi \in \mathbb{R}^d$. Moreover, if there exists a positive function $\mu \in C(D)$ such that $\langle a\xi, \xi\rangle \leq \mu(\xi)|\xi|^2$ for all $\xi \in \mathbb{R}^d$, then A_2 has sub-Markovian resolvent (details can be found, e.g., in [32, Chapter 1] and, in much more generality, in [57, Chapter II]). Ultracontractivity estimates follow as a special case of the corresponding estimates for non-symmetric forms treated next.

Example 2.3.5 (Non-symmetric divergence operators with lower-order terms)**.** Consider the differential operator on smooth functions

$$\begin{aligned} A_0 u &:= -\operatorname{div}(a\nabla u) + b\cdot\nabla u - \operatorname{div}(cu) + a_0 u \\ &= -\sum_{j,k=1}^n \partial_j(a_{jk}\partial_k u) + \sum_{j=1}^n \big(b_j\partial_j u - \partial_j(c_j u)\big) + a_0 u, \end{aligned}$$

where a_{jk}, b_j, c_j, $a_0 \in L^\infty(D)$, and the associated (non-symmetric) bilinear form $\mathscr{E}$ on $V := H_0^1(D)$ is defined as

$$\begin{aligned} \mathscr{E}(u,v) &= \langle a\nabla u, \nabla v\rangle + \langle b\cdot\nabla u, v\rangle + \langle u, c\cdot\nabla v\rangle + \langle a_0 u, v\rangle \\ &= \int_D \Big(\sum_{jk} a_{jk}\partial_j u\,\partial_k v + \sum_j \big(b_j\partial_j u\, v + c_j u\partial_j v\big) + a_0 uv\Big). \end{aligned}$$

The bilinear form $\mathscr{E}$ is continuous, as it easily follows from the boundedness of its coefficients. If there exists a constant $C > 0$ such that $\langle a\xi, \xi\rangle \geq C|\xi|^2$, then $\mathscr{E}$ is not V-elliptic, but satisfies the weaker estimate

$$\mathscr{E}(u,u) \geq C_1\|u\|_V^2 - C_2\|u\|_H^2 \qquad \forall u \in V,$$

where $C_1 > 0$ and $C_2 \in \mathbb{R}$ (see, e.g., [7, §11.2] or [70, p. 100]), i.e. the corresponding operator A satisfies (i'), but not (i). Using the Poincaré inequality, it is not difficult to show that $\mathscr{E}$ is V-elliptic if the diameter of D is small enough (see [31, pp. 385–387]). If we furthermore assume that $a_0 - \operatorname{div} c \geq 0$ (in the sense of distributions), then the semigroup S_2 is sub-Markovian, and so is also the resolvent of A_2. Similarly, if $a_0 - \operatorname{div} b \geq 0$,[§] then the semigroup S_2 is L^1-contracting (these results can be found, for instance, in [7, Proposition 11.14], or deduced from [70, §4.3]). As already mentioned above, this implies that S_2 can be extended to a consistent family of semigroups S_p for all $p \in [1, \infty[$. Finally, let us discuss ultracontractivity: if $\mathscr{E}$ is V-elliptic, and S_2 as well as S_2^* are sub-Markovian, then a reasoning based on the Nash inequality

$$\big\|u\big\|_{L^2}^{2+4/n} \leq N\big\|u\big\|_{H_0^1}^2\big\|u\big\|_{L^1}^{4/n} \qquad \forall u \in H_0^1,$$

implies the estimate

$$\big\|S_2(t)\big\|_{\mathscr{L}(L^1,L^\infty)} \leq N_1 t^{-n/2},$$

where $N_1 := \big(Nn/(2\alpha)\big)^{n/2}$. For a proof, see, e.g., [5, Theorem 12.3.2] or [70, p. 159]. The Laplace transform representation of the resolvent yields

$$(I + \lambda A_1)^{-m} = \frac{\lambda^m}{(m-1)!}\int_0^\infty t^{m-1}e^{-\lambda t}S(t)\,dt$$

[§]These two conditions involving a_0 and the divergence of b, c, are not restrictive, as they are close to necessary to ensure that the bilinear form $\mathscr{E}$ is positive. This can be seen by a simple computation based on integration by parts, cf. [57, p. 48].

(see, e.g., [7, p. 17] or [74, p. 21]), hence

$$\big\|(I+\lambda A_1)^{-m}\big\|_{\mathscr{L}(L^1,L^\infty)} \lesssim \frac{\lambda^m}{(m-1)!}\int_0^\infty t^{m-1-n/2}e^{-\lambda t}\,dt.$$

Thus it suffices to choose m large enough to infer the ultracontractivity of the m-th power of the resolvent.

Example 2.3.6 (Fractional Laplacian)**.** Let Δ be the Dirichlet Laplacian on H. Since it is a positive self-adjoint operator, it follows that, for any $\alpha \in \,]0,1[$, $(-\Delta)^\alpha$ is itself a positive self-adjoint (densely defined) operator on H. Furthermore, the bilinear form

$$\mathscr{E}(u,v) := \big\langle(-\Delta)^\alpha u, v\big\rangle = \big\langle(-\Delta)^{\alpha/2}u,(-\Delta)^{\alpha/2}v\big\rangle, \qquad u,v\in\mathsf{D}\big((-\Delta)^{\alpha/2}\big),$$

is a symmetric Dirichlet form on H, which, as already seen, uniquely determines an operator A satisfying conditions (i'), (ii), and (iii): in particular, $V=\mathsf{D}\big((-\Delta)^{\alpha/2}\big)$, equipped with the scalar product $\langle\cdot,\cdot\rangle_V := \langle\cdot,\cdot\rangle+\mathscr{E}$, and A is just the extension of $(-\Delta)^\alpha$, generator of $\mathscr{E}$, to V. In order to prove (iv), we are going to use again an argument based on the Nash inequality, which is however more involved as before. In particular, since $-\Delta$ satisfies the Nash inequality

$$\big\|u\big\|_{L^2}^{2+4/n} \lesssim \big\langle -\Delta u,u\big\rangle\big\|u\big\|_{L^1}^{4/n} \qquad \forall u\in H^1_0,$$

a result by Bendikov and Maheux, see [14, Theorem 1.3], implies that the fractional power $(-\Delta)^\alpha$ satisfies the Nash inequality

$$\big\|u\big\|_{L^2}^{2+4\alpha/n} \lesssim \big\langle (-\Delta)^\alpha u,u\big\rangle\big\|u\big\|_{L^1}^{4\alpha/n} \qquad \forall u\in\mathsf{D}(\mathscr{E}).$$

It follows by a general criterion of Varopoulos, Saloff-Coste and Coulhon (attributed to Ph. Bénilan), see [81, Theorem II.5.2], that the semigroup S_α on H generated by $(-\Delta)^\alpha$ satisfies the ultracontractivity estimate

$$\big\|S_\alpha(t)\big\|_{\mathscr{L}(L^1,L^\infty)} \lesssim t^{-n/2\alpha},$$

from which corresponding estimates for suitable powers of the resolvent can be deduced, as in the previous example.

Related results on ultracontractivity and smoothing properties of semigroups generated by non-local operators, arising as generators of Markov processes, can be found, e.g., in [39, 48].

We proceed with a brief discussion about the relation between our hypotheses on A and those needed in the deterministic setting, where it is enough to prove that $A+\beta$ is maximal monotone in H to get well-posedness of the nonlinear equation, for any right-hand side belonging to $L^1(0,T;H)$. Probably the most widely used criterion for the maximal monotonicity of the sum of two maximal monotone operators on H, at least with applications to PDE in mind, is the following: let F be a maximal monotone operator on H and φ a lower semi-continuous proper convex function on H. If

$$\varphi\big((I+\lambda F)^{-1}u\big) \le \varphi(u)+C\lambda \qquad \forall\lambda>0,\ \forall u\in\mathsf{D}(\varphi), \tag{2.3.2}$$

then $F+\partial\varphi$ is maximal monotone (see [21, Theorem 9, p. 108]). In the case of semilinear perturbations of the Laplacian of the type $-\Delta+\beta$, this result is used as follows: let φ be such that $-\Delta = \partial\varphi$, and

$$\psi : u \mapsto \begin{cases} \displaystyle\int_D j(u)\,dx, & \text{if } j(u) \in L^1(D), \\ +\infty, & \text{if } j(u) \notin L^1(D). \end{cases}$$

Then $\psi : H \to \mathbb{R} \cup \{+\infty\}$ is proper convex lower semicontinuous, and $F := \partial\psi$ is maximal monotone, with $F(u) = \beta(u)$ a.e. for all $u \in H$ such that $j(u) \in L^1(D)$. Then one has, recalling that $(I + \lambda\beta)^{-1}$ is a contraction on $\mathbb{R}$,

$$\begin{aligned} \varphi\big((I + \lambda F)^{-1}u\big) &= \int_D \big|\nabla(I + \lambda\beta)^{-1}u\big|^2\,dx \\ &\le \int_D |\nabla u|^2\,dx = \varphi(u), \end{aligned}$$

so that (2.3.2) is satisfied, and $-\Delta+\beta$ is maximal monotone. If one replaces $-\Delta$ with a general positive self-adjoint operator A on H, it is not clear how to adapt such reasoning. However, if we assume that A is the generator of a symmetric Dirichlet form $\mathscr{E}$ on H, then (2.3.2) is satisfied, with $C = 0$ and $\varphi = \mathscr{E}$. This follows from the fact that $(I + \lambda\beta)^{-1}$ is a normal contraction on $\mathbb{R}$ and that, for any normal contraction T on $\mathbb{R}$, $u \in \mathsf{D}(\mathscr{E})$ implies $Tu \in \mathsf{D}(\mathscr{E})$ and $\mathscr{E}(Tu, Tu) \le \mathscr{E}(u, u)$, a proof of which can be found, e.g., in [57, Theorem 4.12, p. 36].

On the other hand, if A is maximal monotone but not self-adjoint, we cannot express it as the subdifferential of a convex function on H. Hence we are led to "dualize" the previous argument, i.e. we can try to show that

$$\psi\big((I + \lambda A)^{-1}u\big) \le \psi(u) + C\lambda \qquad \forall\lambda > 0,\ \forall u \in \mathsf{D}(\varphi).$$

Knowing only that the resolvent is a contraction does not seem enough to proceed. However, if we assume that the resolvent is sub-Markovian, we can apply Jensen's inequality (see Lemma 1.3.15 below), so that

$$j\big((I + \lambda A)^{-1}u\big) \le (I + \lambda A)^{-1}j(u),$$

hence, integrating,

$$\psi\big((I + \lambda A)^{-1}u\big) = \int_D j\big((I + \lambda A)^{-1}u\big)\,dx \le \int_D (I + \lambda A)^{-1}j(u)\,dx.$$

Assuming also that the resolvent is contracting in L^1, we obtain $\psi\big((I+\lambda A)^{-1}u\big) \le \psi(u)$, hence that $A + \beta$ is maximal monotone in H. Recall that A is contracting in L^1 if it is the generator of a (nonsymmetric) Dirichlet form. It results from this discussion that our conditions (ii) and (iii) on A are not restrictive and are probably close to optimal, while the ultracontractivity condition (iv) is completely superfluous in the deterministic setting. Moreover, while condition (i') is always satisfied if A is self-adjoint, it is equally superfluous in the deterministic case if A is non-symmetric.

Let us now comment on the Lipschitz continuity assumption on B. It is natural to ask whether a well-posedness result analogous to Theorem 2.2.2 holds under the weaker assumption that B is progressively measurable, linearly growing, and just locally Lipschitz continuous, i.e. assuming that there exists a sequence $(L_B^n)_n$ of positive real numbers such that

$$\big\|B(\omega, t, x) - B(\omega, t, y)\big\|_{\mathscr{L}^2(U,H)} \le L_B^n\|x - y\|_H$$

for every $(\omega, t) \in \Omega \times [0,T]$ and $x, y \in H$ with $\|x\|_H, \|y\|_H \leq n$, for every $n \in \mathbb{N}$. In this case, introducing the globally Lipschitz continuous truncated operators

$$B_n : \Omega \times [0,T] \times H \to \mathscr{L}^2(U,H), \qquad B_n(\omega,t,x) := B(\omega,t,nPx),$$

for all $n \in \mathbb{N}$, where $P : H \to H$ is the projection on the closed unit ball in H, the stochastic evolution equation

$$dX_n + AX_n\,dt + \beta(X_n)\,dt \ni B_n(t,X_n)\,dW, \qquad X_n(0) = X_0,$$

is well-posed in $\mathscr{J}$ for all $n \in \mathbb{N}$. One would now expect to be able to construct a global solution by suitably "gluing" the solutions (X_n, ξ_n). In fact, this technique has been successfully applied in several situations (cf., e.g., [25, 47, 80]): the key argument is to introduce the sequence of stopping times $(\tau_n)_n$ defined as

$$\tau_n := \inf\bigl\{t \in [0,T] : \|X_n(t)\| \geq n\bigr\} \wedge T,$$

and to show that, for any $m > n$, one has $X_m = X_n$ on

$$[\![0, \tau_n]\!] := \bigl\{(\omega,t) \in \Omega \times [0,T] :\, 0 \leq t \leq \tau_n(\omega)\bigr\}.$$

For this construction to work, it seems essential to assume that X_n has continuous trajectories for all $n \in \mathbb{N}$ (as is the case in *op. cit.*). However, in our case, we only know that the trajectories of X_n are weakly continuous in H, hence the above construction does not seem to work. On the other hand, we conjecture that strong solutions in $\mathscr{J}$ to (2.1.1) are indeed pathwise continuous under suitable polynomial boundedness assumption on β, and that, in this case, equations with locally Lipschitz diffusion coefficient can be shown to be well-posed. This will be treated in the forthcoming Chapter 3. We conclude remarking that such a well-posedness result for semilinear equations with polynomially growing drift does not follow from the classical variational approach (see, e.g., [56, Example 5.1.8]).

2.4 Well-posedness for a regularized equation

Let V_0 be a separable Hilbert space such that V_0 is a dense subset of V, $V_0 \hookrightarrow V$, and $V_0 \hookrightarrow L^\infty(D)$. The goal of this section is to establish existence and uniqueness of solutions to the stochastic evolution equation

$$dX(t) + AX(t)\,dt + \beta(X(t))\,dt \ni B(t)\,dW(t), \qquad X(0) = X_0, \tag{2.4.3}$$

where B is an $\mathscr{L}^2(U,V_0)$-valued process. In particular, this stochastic equation can be interpreted as a version of (2.1.1) with additive and more regular noise.

Proposition 2.4.1. *Assume that $X_0 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P}; H)$ and that*

$$B \in L^2(\Omega; L^2(0,T; \mathscr{L}^2(U,V_0)))$$

is measurable and adapted. Then equation (2.4.3) admits a unique strong solution (X, ξ) such *that*

$$X \in L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V)),$$
$$j(X) + j^*(\xi) \in L^1((0,T) \times D) \qquad \mathbb{P}\textit{-almost surely.}$$

Moreover, $X(\omega,\cdot) \in C_w([0,T];H)$ *for* $\mathbb{P}$*-almost all* $\omega \in \Omega$*.*

The rest of this section is devoted to the proof of Proposition 2.4.1, which is structured as a follows: we consider a regularized version of (2.4.3), where the nonlinear term β is replaced by its Yosida approximation, and obtain suitable a priori estimates, both pathwise and in expectation. Taking limits in appropriate topologies of the solutions to these regularized equations, we construct solutions to (2.4.3), that are finally shown to be unique.

Let

$$\beta_\lambda := \frac{1}{\lambda}\big(I - (I+\lambda\beta)^{-1}\big), \qquad \lambda > 0,$$

be the Yosida approximation of β, and consider the regularized equation

$$dX_\lambda(t) + AX_\lambda(t)\,dt + \beta_\lambda(X_\lambda(t))\,dt = B(t)\,dW(t), \qquad X_\lambda(0) = X_0. \tag{2.4.4}$$

Since β_λ is monotone and Lipschitz continuous, it is easy to check that the operator $A + \beta_\lambda$ satisfies, for any $\lambda > 0$, the classical conditions of Pardoux, Krylov and Rozovskiĭ [46, 72]. For completeness, a proof is given next.

Lemma 2.4.2. *Let* $\lambda > 0$*. The operator* $A_\lambda := A + \beta_\lambda : V \to V^*$ *satisfies the following conditions:*

(i) A_λ *is hemicontinuous, i.e. the map* $\mathbb{R} \ni \eta \mapsto \langle A_\lambda(u+\eta v), x\rangle$ *is continuous for all* u*,* v*,* $x \in V$*;*

(ii) A_λ *is monotone, i.e.* $\langle A_\lambda u - A_\lambda v, u - v\rangle \geq 0$ *for all* u*,* $v \in V$*;*

(iii) A_λ *is coercive, i.e. there exists a constant* $C_1 > 0$ *such that* $\langle A_\lambda v, v\rangle \geq C_1\|v\|_V^2$ *for all* $v \in V$*;*

(iv) A_λ *is bounded, i.e. there exists a constant* $C_2 > 0$ *such that* $\|A_\lambda v\|_{V^*} \leq C_2\|v\|_V$ *for all* $v \in V$*.*

Proof. (i) For any $u, v, x \in V$, one has

$$\langle A_\lambda(u+\eta v), x\rangle = \langle Au, x\rangle + \eta\langle Av, x\rangle + \int_D \beta_\lambda(u+\eta v)x.$$

It clearly suffices to check that the last term depends continuously on η, which follows immediately by the Lipschitz continuity of β_λ. (ii) Since both A and β_λ are monotone, one has

$$\langle A_\lambda u - A_\lambda v, u - v\rangle = \langle Au - Av, u - v\rangle + \int_D (\beta_\lambda(u) - \beta_\lambda(v)(u - v) \geq 0.$$

(iii) Similarly, since $0 \in \beta(0)$ implies $\beta_\lambda(0) = 0$, coercivity of A and monotonicity of β_λ imply

$$\langle A_\lambda v, v\rangle = \langle Av, v\rangle + \int_D \beta_\lambda(v)v \geq \langle Av, v\rangle \geq C\|v\|_V^2$$

(in particular, C_1 can be chosen equal to C, the coercivity constant of A itself). (iv) Using again the fact that $\beta_\lambda(0) = 0$, and recalling that β_λ is Lipschitz continuous with Lipschitz constant bounded by $1/\lambda$, one has

$$\begin{aligned}\langle A_\lambda v, u\rangle = \langle Av, u\rangle + \int_D \beta_\lambda(v)u &\leq \|Av\|_{V^*}\|u\|_V + \frac{1}{\lambda}\|v\|_H\|u\|_H \\ &\leq \bigl(\|A\|_{\mathscr{L}(V,V^*)} + k/\lambda\bigr)\|v\|_V\|u\|_V,\end{aligned}$$

where k is the norm of the continuous embedding $\iota : V \to H$. □

Hence (2.4.4) admits a unique variational solution, that is, there exists a unique adapted process

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

such that, in V^*,

$$X_\lambda(t) + \int_0^t AX_\lambda(s)\,ds + \int_0^t \beta_\lambda(X_\lambda(s))\,ds = X_0 + \int_0^t B(s)\,dW(s) \tag{2.4.5}$$

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

In the next lemmata we establish a priori estimates for X_λ and $\beta_\lambda(X_\lambda)$. We begin with a pathwise estimate.

Lemma 2.4.3. *There exists $\Omega' \subseteq \Omega$ with $\mathbb{P}(\Omega') = 1$ and $M : \Omega' \to \mathbb{R}$ such that*

$$\bigl\|X_\lambda(\omega)\bigr\|^2_{C([0,T];H)\cap L^2(0,T;V)} + \bigl\|j_\lambda(X_\lambda(\omega))\bigr\|_{L^1(0,T;L^1(D))} < M(\omega)$$

for all $\omega \in \Omega'$.

Proof. Setting $Y_\lambda := X_\lambda - B \cdot W$, Itô's formula¶ yields

$$\bigl\|Y_\lambda(t)\bigr\|^2_H + 2\int_0^t \bigl\langle AX_\lambda(s), Y_\lambda(s)\bigr\rangle\,ds + 2\int_0^t \bigl\langle \beta_\lambda(X_\lambda), Y_\lambda(s)\bigr\rangle\,ds = \bigl\|X_0\bigr\|^2_H,$$

where $\|X_\lambda\|_H \leq \|Y_\lambda\|_H + \|B\cdot W\|_H$ by the triangle inequality, hence

$$\|Y_\lambda(t)\|^2_H \geq \frac{1}{2}\|X_\lambda(t)\|^2_H - \|B\cdot W(t)\|^2_H.$$

Moreover, writing $\langle AX_\lambda, Y_\lambda\rangle = \langle AX_\lambda, X_\lambda\rangle - \langle AX_\lambda, B\cdot W\rangle$, one has

$$\langle AX_\lambda, X_\lambda\rangle \geq C\|X_\lambda\|^2_V$$

by the coercivity of A, and

$$\begin{aligned}\langle AX_\lambda, B\cdot W\rangle &\leq \|A\|_{\mathscr{L}(V,V^*)}\|X_\lambda\|_V\|B\cdot W\|_V \\ &\leq \frac{1}{2}C\|X_\lambda\|^2_V + \frac{1}{2\varepsilon}\|B\cdot W\|^2_V,\end{aligned}$$

where we have used the elementary inequality $ab \leq \frac{1}{2}(\varepsilon a^2 + b^2/\varepsilon)$ for all $a, b \in \mathbb{R}$, with $\varepsilon := C\|A\|^{-2}_{\mathscr{L}(V,V^*)}$. Then

$$\langle AX_\lambda, Y_\lambda\rangle \geq \frac{1}{2}C\|X_\lambda\|^2_V - \frac{1}{2\varepsilon}\|B\cdot W\|^2_V,$$

so that

$$2\int_0^t \bigl\langle AX_\lambda(s), Y_\lambda(s)\bigr\rangle\,ds \geq C\int_0^t \|X_\lambda(s)\|^2_V\,ds - \frac{1}{\varepsilon}\int_0^t \|B\cdot W(s)\|^2_V\,ds$$

¶Whenever we refer to Itô's formula, we shall always mean the version in [46].

and

$$\begin{aligned}\frac{1}{2}\|X_\lambda(t)\|_H^2 + C\int_0^t \|X_\lambda(s)\|_V^2\,ds + 2\int_0^t \bigl\langle \beta_\lambda(X_\lambda(s)), Y_\lambda(s)\bigr\rangle\,ds \\ \leq \|X_0\|_H^2 + \|B\cdot W(t)\|_H^2 + \frac{1}{\varepsilon}\int_0^t \|B\cdot W(s)\|_V^2\,ds.\end{aligned} \tag{2.4.6}$$

Let j_λ be the Moreau-Yosida regularization of j, that is

$$j_\lambda(x) := \inf_{y\in\mathbb{R}}\Bigl(j(y) + \frac{|x-y|^2}{2\lambda}\Bigr), \qquad \lambda > 0.$$

We recall that j_λ is a convex, proper differentiable function, with $j'_\lambda = \beta_\lambda$, that converges pointwise to j from below. In particular,

$$\beta_\lambda(x)(x-y) \geq j_\lambda(x) - j_\lambda(y) \geq j_\lambda(x) - j(y) \qquad \forall x, y \in \mathbb{R}.$$

This implies

$$\begin{aligned}\int_0^t \bigl\langle \beta_\lambda(X_\lambda(s)), Y_\lambda(s)\bigr\rangle\,ds &= \int_0^t\!\!\int_D \beta_\lambda(X_\lambda(s,x))(X_\lambda(s,x) - B\cdot W(s,x))\,dx\,ds \\ &\geq \int_0^t\!\!\int_D j_\lambda(X_\lambda(s,x))\,dx\,ds - \int_0^t\!\!\int_D j(B\cdot W(s,x))\,dx\,ds,\end{aligned}$$

hence also

$$\begin{aligned}\frac{1}{2}\|X_\lambda(t)\|_H^2 + C\int_0^t \|X_\lambda(s)\|_V^2\,ds + 2\int_0^t\!\!\int_D j_\lambda(X_\lambda(s,x))\,dx\,ds \\ \leq \|X_0\|_H^2 + \|B\cdot W(t)\|_H^2 + \frac{1}{\varepsilon}\int_0^t \|B\cdot W(s)\|_V^2\,ds \\ + 2\int_0^t\!\!\int_D j(B\cdot W(s,x))\,dx\,ds.\end{aligned}$$

Taking the supremum with respect to t yields

$$\begin{aligned}&\bigl\|X_\lambda\bigr\|^2_{C([0,T];H)} + \bigl\|X_\lambda\bigr\|^2_{L^2(0,T;V)} + \bigl\|j_\lambda(X_\lambda)\bigr\|_{L^1(0,T;L^1(D))} \\ &\lesssim \bigl\|X_0\bigr\|^2_H + \bigl\|B\cdot W\bigr\|^2_{C([0,T];H)} + \bigl\|B\cdot W\bigr\|^2_{L^2(0,T;V)} + \bigl\|j(B\cdot W)\bigr\|_{L^1(0,T;L^1(D))},\end{aligned}$$

where the implicit constant depends only on the operator norm of A. It follows by Itô's isometry and Doob's inequality that

$$\bigl\|B\cdot W\bigr\|_{L^2(\Omega;C([0,T];V_0))} \lesssim \bigl\|B\bigr\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,V_0)))},$$

where the right-hand side is finite by assumption, hence, recalling that V_0 is continuously embedded in V,

$$\bigl\|B\cdot W\bigr\|_{C([0,T];H)} + \bigl\|B\cdot W\bigr\|_{L^2(0,T;V)} \lesssim_T \bigl\|B\cdot W\bigr\|_{C([0,T];V_0)}.$$

Analogously, denoting the norm of the continuous embedding $\iota : V_0 \to L^\infty(D)$ by k, one has, recalling that j is symmetric and increasing on $\mathbb{R}_+$,

$$\bigl\|j(B\cdot W(t)\bigr\|_{L^1(D)} \lesssim_{|D|} j\bigl(\|B\cdot W(t)\|_{L^\infty(D)}\bigr) \leq j\bigl(k\|B\cdot W(t)\|_{V_0}\bigr),$$

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

$$\big\|j(B \cdot W)\big\|_{L^1(0,T;L^1(D))} \lesssim_{|D|,T} j\big(k\|B \cdot W\|_{C([0,T];V_0)}\big).$$

We conclude choosing $\Omega' \subset \Omega$ such that $\|X_0(\omega)\|_H$ and $\|B \cdot W(\omega)\|_{C([0,T];V_0)}$ are finite for all $\omega \in \Omega'$, and defining $M : \Omega' \to \mathbb{R}$ as

$$M := \big\|X_0\big\|_H^2 + \big\|B \cdot W\big\|^2_{C([0,T];H)} + \big\|B \cdot W\big\|^2_{L^2(0,T;V)} + \big\|j(B \cdot W)\big\|_{L^1(0,T;L^1(D))}. \qquad \square$$

Remark 2.4.4. The above estimates can be obtained by purely deterministic arguments, without invoking Itô's formula. In fact, note that equation (2.4.5) can equivalently be written as

$$Y_\lambda(t) + \int_0^t \big(AX_\lambda(s) + \beta_\lambda(X_\lambda(s))\big)\, ds = 0.$$

One has $Y_\lambda \in L^2(0,T;V)$, which follows at once by the properties of X_λ and by $B \cdot W \in L^2(\Omega; C([0,T]; V_0))$. Similarly, since AX_λ and $\beta_\lambda(X_\lambda)$ belong to $L^2(\Omega; L^2(0,T;V^*))$, one also has, by the previous identity, $Y'_\lambda \in L^2(0,T;V^*)$. In particular, there exists $\Omega' \subset \Omega$, with $\mathbb{P}(\Omega') = 1$, such that

$$Y_\lambda(\omega) \in L^2(0,T;V), \qquad Y'_\lambda(\omega) \in L^2(0,T;V^*) \qquad \forall \omega \in \Omega'.$$

Lemma 1.4.2 then yields

$$\frac{1}{2}\big\|Y_\lambda(t)\big\|_H^2 + \int_0^t \big\langle AX_\lambda(s), Y_\lambda(s)\big\rangle\, ds + \int_0^t \big\langle \beta_\lambda(X_\lambda), Y_\lambda(s)\big\rangle\, ds = \frac{1}{2}\big\|X_0\big\|_H^2.$$

Lemma 2.4.5. *There exists a constant $N > 0$ such that*

$$\begin{aligned}\big\|X_\lambda\big\|^2_{L^2(\Omega;C([0,T];H))} &+ \big\|X_\lambda\big\|^2_{L^2(\Omega;L^2(0,T;V)} + \big\|\beta_\lambda(X_\lambda)X_\lambda\big\|_{L^1(\Omega;L^1(0,T;L^1(D)))}\\ &< N\Big(\big\|X_0\big\|^2_{L^2(\Omega;H)} + \big\|B\big\|^2_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}\Big).\end{aligned}$$

Proof. Itô's formula yields

$$\begin{aligned}\big\|X_\lambda(t)\big\|_H^2 &+ 2\int_0^t \big\langle AX_\lambda(s), X_\lambda(s)\big\rangle\, ds + 2\int_0^t \big\langle \beta_\lambda(X_\lambda(s)), X_\lambda(s)\big\rangle\, ds\\ &= \big\|X_0\big\|_H^2 + 2\int_0^t X_\lambda(s)B(s)\, dW(s) + \frac{1}{2}\int_0^t \big\|B(s)\big\|^2_{\mathscr{L}^2(U,H)}\, ds,\end{aligned}$$

where X_λ in the stochastic integral on the right-hand side has to be interpreted as taking values in $H^* \simeq H$. The coercivity of A and the monotonicity of β_λ readily imply, after taking supremum in time and expectation,

$$\begin{aligned}\mathbb{E}\big\|X_\lambda\big\|^2_{C([0,T];H)} &+ 2C\,\mathbb{E}\big\|X_\lambda\big\|^2_{L^2(0,T;V)} + \mathbb{E}\int_0^T \big\langle \beta_\lambda(X_\lambda(s)), X_\lambda(s)\big\rangle\, ds\\ &\lesssim \mathbb{E}\big\|X_0\big\|_H^2 + \mathbb{E}\big\|B\big\|^2_{L^2(0,T;\mathscr{L}^2(U,H))} + \mathbb{E}\sup_{t\in[0,T]}\bigg|\int_0^t X_\lambda(s)B(s)\, dW(s)\bigg|,\end{aligned}$$

where, by Lemma 1.5.1,

$$\mathbb{E}\sup_{t\in[0,T]}\bigg|\int_0^t X_\lambda(s)B(s)\, dW(s)\bigg| \leq \varepsilon\,\mathbb{E}\big\|X_\lambda\big\|^2_{C([0,T];H)} + N(\varepsilon)\,\mathbb{E}\int_0^T \big\|B(s)\big\|^2_{\mathscr{L}^2(U,H)}\, ds$$

for any $\varepsilon > 0$, whence the result follows choosing ε small enough. $\square$

We now establish weak compactness properties for the sequence $(\beta_\lambda(X_\lambda))$.

Lemma 2.4.6. *The sequence $(\beta_\lambda(X_\lambda))$ is relatively weakly compact in $L^1(\Omega \times (0,T) \times D)$. Moreover, there exists a set $\Omega'' \subset \Omega$, with $\mathbb{P}(\Omega'') = 1$, such that $(\beta_\lambda(X_\lambda(\omega, \cdot))$ is weakly relatively compact in $L^1((0,T) \times D)$ for all $\omega \in \Omega''$.*

Proof. Recalling that, for any y, $r \in \mathbb{R}$, $j(y) + j^*(r) = ry$ if and only if $r \in \partial j(y) = \beta(y)$, one has

$$j\big((I + \lambda\beta)^{-1}x\big) + j^*\big(\beta_\lambda(x)\big) = \beta_\lambda(x)(I + \lambda\beta)^{-1}x \leq \beta_\lambda(x)x \qquad \forall x \in \mathbb{R}. \tag{2.4.7}$$

In fact, since $\beta_\lambda \in \beta \circ (I + \lambda\beta)^{-1}$, it follows from $\beta = \partial j$ that $\beta_\lambda(x) \in \partial j\big((I + \lambda\beta)^{-1}x\big)$. Moreover, $\beta\big((I+\lambda\beta)^{-1}x\big)(I+\lambda\beta)^{-1}x \geq 0$ by monotonicity of β, hence the inequality in (2.4.7) follows since $(I + \lambda\beta)^{-1}$ is a contraction. The previous lemma thus implies, thanks to the symmetry of j^*, that there exists a constant N, independent of λ, such that, setting

$$\bar{N}(X_0, B) := N\Big(\|X_0\|^2_{L^2(\Omega;H)} + \|B\|^2_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}\Big),$$

one has

$$\mathbb{E}\int_0^T\!\!\int_D j^*\big(|\beta_\lambda(X_\lambda)|\big) \leq \mathbb{E}\int_0^T\!\!\int_D \beta_\lambda(X_\lambda)X_\lambda < \bar{N}(X_0, B).$$

Since j^* is superlinear at infinity, the sequence $(\beta_\lambda(X_\lambda))$ is uniformly integrable on $\Omega \times (0,T) \times D$ by the de la Vallée-Poussin criterion, hence weakly relatively compact in $L^1(\Omega \times (0,T) \times D)$ by a well-known theorem of Dunford and Pettis. The first assertion is thus proved.

By (2.4.6), since $Y_\lambda = X_\lambda - B \cdot W$, it follows that

$$\begin{aligned}\int_0^t \big\langle \beta_\lambda(X_\lambda(s)), X_\lambda(s)\big\rangle\, ds \lesssim \|X_0\|_H^2 + \|B \cdot W(t)\|_H^2 + \int_0^t \|B \cdot W(s)\|_V^2\, ds \\ + \int_0^t \big\langle \beta_\lambda(X_\lambda(s)), B \cdot W(s)\big\rangle\, ds,\end{aligned}$$

where, by Young's inequality and convexity (recalling that $j^*(0) = 0$),

$$\int_0^t \big\langle \beta_\lambda(X_\lambda(s)), B \cdot W(s)\big\rangle\, ds \leq \frac{1}{2}\int_0^t\!\!\int_D j^*\big(\beta_\lambda(X_\lambda)\big) + \int_0^t\!\!\int_D j(2B \cdot W).$$

Rearranging terms and proceeding as in the (end of the) proof of Lemma 2.4.3, we infer that there exists a set $\Omega'' \subset \Omega$, with $\mathbb{P}(\Omega'') = 1$, and a function $M : \Omega'' \to \mathbb{R}$ such that

$$\int_0^T \big\langle \beta_\lambda(X_\lambda(\omega, s)), X_\lambda(\omega, s)\big\rangle\, ds < M(\omega) \qquad \forall \omega \in \Omega''. \tag{2.4.8}$$

The symmetry of j^* and (2.4.7) yield that, for any $\omega \in \Omega''$, $(\beta_\lambda(X_\lambda(\omega, \cdot)))$ is weakly relatively compact in $L^1((0,T) \times D)$. $\square$

In order to pass to the limit as $\lambda \to 0$, we are going to use Simon's compactness criterion, i.e. Lemma 1.4.3, and Brézis' Lemma 1.3.14.

Proposition 2.4.7. *There exists $\Omega' \subseteq \Omega$, with $\mathbb{P}(\Omega') = 1$, such that, for any $\omega \in \Omega'$, there* *exists a subsequence $\lambda' = \lambda'(\omega)$ of λ such that, as $\lambda' \to 0$,*

$$\begin{aligned}
&X_{\lambda'}(\omega,\cdot) \overset{*}{\rightharpoonup} X(\omega,\cdot) && \textit{in } L^\infty(0,T;H),\\
&X_{\lambda'}(\omega,\cdot) \rightharpoonup X(\omega,\cdot) && \textit{in } L^2(0,T;V),\\
&X_{\lambda'}(\omega,\cdot) \longrightarrow X(\omega,\cdot) && \textit{in } L^2(0,T;H),\\
&\beta_{\lambda'}(X_{\lambda'}(\omega,\cdot)) \rightharpoonup \xi(\omega,\cdot) && \textit{in } L^1((0,T)\times D).
\end{aligned}$$

Proof. The first two convergence statements follow by Lemma 2.4.3, and the fourth one follows by Lemma 2.4.6. Let us show that the third convergence statement holds. In the following we omit the indication of ω, as no confusion can arise. Setting $Y_\lambda = X_\lambda - B \cdot W$, (2.4.5) can equivalently be written as the deterministic equation (with random coefficients) on V^*

$$Y'_\lambda + AX_\lambda + \beta_\lambda(X_\lambda) = 0,$$

where

$$\begin{aligned}
\|AX_\lambda\|_{L^1(0,T;V_0^*)} \lesssim \|AX_\lambda\|_{L^1(0,T;V^*)} \lesssim \|X_\lambda\|_{L^1(0,T;V)},\\
\|\beta_\lambda(X_\lambda)\|_{L^1(0,T;V_0^*)} \lesssim \|\beta_\lambda(X_\lambda)\|_{L^1(0,T;V^*)} \lesssim \|\beta_\lambda(X_\lambda)\|_{L^1(0,T;L^1(D))},
\end{aligned}$$

hence, again by Lemmata 2.4.3 and 2.4.6, $\|Y'_\lambda\|_{L^1(0,T;V_0^*)}$ is bounded uniformly over λ. Moreover, since $B \cdot W \in L^2(\Omega; C([0,T];V_0))$ and

$$\|Y_\lambda\|_{L^2(0,T;V)} \le \|X_\lambda\|_{L^2(0,T;V)} + \|B\cdot W\|_{L^2(0,T;V)},$$

we conclude that (Y_λ) is bounded in $L^2(0,T;V)$. Simon's compactness criterion then implies that Y_λ, hence also X_λ, is relatively compact in $L^2(0,T;H)$. Since $X_{\lambda'} \rightharpoonup X$ in $L^2(0,T;V)$, it follows that

$$X_{\lambda'}(\omega,\cdot) \longrightarrow X(\omega,\cdot) \qquad \text{in } L^2(0,T;H),$$

thus completing the proof. □

We are now going to show that the couple (X,ξ) just constructed is indeed the unique solution to the equation with "smoothed" noise (2.4.3).

Proof of Proposition 2.4.1. In spite of the above preparations, the argument is quite long, so we subdivide it into several steps.

STEP 1. In the notation of Proposition 2.4.7, let $\omega \in \Omega'$ be arbitrary but fixed. Note that $X_{\lambda'} \to X$ in $L^2(0,T;H)$ implies that, passing to a further subsequence of λ', denoted with the same symbol for simplicity, $X_{\lambda'}(t) \to X(t)$ in H for almost all $t \in [0,T]$. Moreover, $X_{\lambda'} \rightharpoonup X$ in $L^2(0,T;V)$ implies that

$$\int_0^t AX_\lambda(s)\,ds \rightharpoonup \int_0^t AX(s)\,ds \qquad \text{in } V^*$$

for all $t \in [0,T]$. In fact, taking $\phi_0 \in V$ and $\phi := s \mapsto 1_{[0,t]}(s)\phi_0 \in L^2(0,t;V)$, one obviously has $A^*\phi \in L^2(0,t;V^*)$ and

$$\begin{aligned}\int_0^t \langle AX_\lambda(s), \phi_0\rangle\, ds = \int_0^T \langle AX_\lambda(s), \phi(s)\rangle\, ds = \int_0^T \langle X_\lambda(s), A^*\phi(s)\rangle\, ds \\ \longrightarrow \int_0^T \langle X(s), A^*\phi(s)\rangle\, ds = \int_0^t \langle AX(s), \phi_0\rangle\, ds.\end{aligned}$$

Similarly, $\beta_{\lambda'}(X_{\lambda'}) \rightharpoonup \xi$ in $L^1((0,T)\times D)$ implies

$$\int_0^t \beta_{\lambda'}(X_{\lambda'}(s))\, ds \rightharpoonup \int_0^t \xi(s)\, ds \qquad \text{in } L^1(D)$$

for all $t \in [0,T]$. In particular, passing to the limit as $\lambda' \to 0$ in the regularized equation (2.4.5) yields

$$X(t) + \int_0^t AX(s)\, ds + \int_0^t \xi(s)\, ds = X_0 + B\cdot W(t) \qquad \text{in } V_0^* \text{ for a.a. } t \in [0,T].$$

Since $AX \in L^2(0,T;V^*) \hookrightarrow L^1(0,T;V_0^*)$ and $\xi \in L^1(0,T;L^1(D)) \hookrightarrow L^1(0,T;V_0^*)$, recalling that $B\cdot W \in C([0,T];V_0)$, we infer that $X \in C([0,T];V_0^*)$, hence the previous identity is true for all $t \in [0,T]$. Moreover, it follows from $X \in L^\infty(0,T;H)$ that $X \in C_w([0,T];H)$, thanks Lemma 1.4.1. Note also that all terms expect the second one on the left-hand side take values in $L^1(D)$, and all terms except the third one on the left-hand side take values in V^*, hence the above identity holds true also in $L^1(D)\cap V^*$.

Let us now show that $\xi \in \beta(X)$ a.e. in $(0,T)\times D$: $X_{\lambda'} \to X$ in $L^2(0,T;H)$ implies that, passing to a subsequence of λ', still denoted by the same symbol, $X_{\lambda'} \to X$ a.e. in $(0,T)\times D$, hence also $(I+\lambda'\beta)^{-1}X_{\lambda'} \to X$ a.e. in $(0,T)\times D$. Since $\beta_{\lambda'}(X_{\lambda'}) \in \beta((I+\lambda'\beta)^{-1}X_{\lambda'})$ a.e. in $(0,T)\times D$ and $\beta_{\lambda'}(X_{\lambda'})(I+\lambda'\beta)^{-1}X_{\lambda'}$ is bounded in $L^1((0,T)\times D)$ by (2.4.8), Brézis' Lemma 1.3.14 implies the claim. These relations and the weak convergence $\beta_{\lambda'}(X_{\lambda'}) \rightharpoonup \xi$ in $L^1((0,T)\times D)$ also imply, by the weak lower semicontinuity of convex integrals, that

$$\begin{aligned}\int_0^T\!\!\int_D \big(j(X)+j^*(\xi)\big) &\le \liminf_{\lambda'\to 0} \int_0^T\!\!\int_D \big(j((I+\lambda' A)^{-1}X_{\lambda'}) + j^*(\beta_{\lambda'}(X_{\lambda'}))\big) \\ &= \liminf_{\lambda'\to 0} \int_0^T\!\!\int_D \beta_{\lambda'}(X_{\lambda'})(I+\lambda' A)^{-1}X_{\lambda'} \le N,\end{aligned}$$

where N is a constant that depends on ω.

Step 2. Still keeping ω fixed as in the previous step, we are going to show that the limits X and ξ constructed above are unique. Suppose there exist (X_i,ξ_i), $\xi_i \in \beta(X_i)$ a.e. in $(0,T)\times D$, $i=1,2$, such that

$$X_i(t) + \int_0^t AX_i(s)\, ds + \int_0^t \xi_i(s)\, ds = X_0 + B\cdot W(t)$$

in $L^1(D)\cap V^*$ for all $t \in [0,T]$. Setting $X = X_1 - X_2$ and $\xi = \xi_1 - \xi_2$, it is enough to show that

$$X(t) + \int_0^t AX(s)\, ds + \int_0^t \xi(s)\, ds = 0 \tag{2.4.9}$$

in $L^1(D)\cap V^*$ for all $t \in [0,T]$ implies $X = 0$ and $\xi = 0$. By the hypotheses on A, there exists

$m \in \mathbb{N}$ such that $(I+\delta A)^{-m}$ maps $L^1(D)$ in $L^\infty(D)$. Therefore, setting

$$X^\delta := (I+\delta A)^{-m}X, \qquad \xi^\delta := (I+\delta A)^{-m}\xi,$$

one has

$$X^\delta(t) + \int_0^t AX^\delta(s)\,ds + \int_0^t \xi^\delta(s)\,ds = 0$$

for all $t \in [0,T]$, for which Itô's formula and monotonicity of A yield

$$\frac{1}{2}\big\|X^\delta(t)\big\|_H^2 + \int_0^t\!\!\int_D \xi^\delta(s,x)X^\delta(s,x)\,dx\,ds \leq 0.$$

We can now take the limit as $\delta \to 0$. Since $(I+\delta A)^{-m}$ converges, in the strong operator topology, to the identity in $\mathscr{L}(H)$, one has $\|X^\delta(t)\|_H \to \|X(t)\|_H$ for all $t \in [0,T]$. Passing to a subsequence of δ, still denoted by the same symbol, we also have $X^\delta \to X$ and $\xi^\delta \to \xi$ a.e. in $(0,T)\times D$, hence $X^\delta\xi^\delta \to X\xi$ a.e. in $(0,T)\times D$. Let us show that $(X^\delta\xi^\delta)$ is uniformly integrable: by the symmetry of j and j^*, and the abstract Jensen inequality of Lemma 1.3.15, we have

$$|X_\delta\xi_\delta| \leq j(X_\delta) + j^*(\xi_\delta) \leq (I+\delta A)^{-m}\big(j(X)+j^*(\xi)\big),$$

where the term on the right-hand side converges to $j(X)+j^*(\xi)$ in $L^1((0,T)\times D)$ as $\delta \to 0$, hence $(X^\delta\xi^\delta)$ is indeed uniformly integrable on $(0,T)\times D$. It follows by Vitali's convergence theorem that, for any $t \in [0,T]$,

$$\int_0^t\!\!\int_D X^\delta\xi^\delta \to \int_0^t\!\!\int_D X\xi,$$

hence also

$$\frac{1}{2}\big\|X(t)\big\|_H^2 + \int_0^t\!\!\int_D X(s,x)\xi(s,x)\,dx\,ds \leq 0.$$

The monotonicity of β immediately implies that $X(t) = 0$ for all $t \in [0,T]$. Substituing in (2.4.9), we are left with $\int_0^t \xi(s)\,ds = 0$ in $L^1(D)$ for all $t \in [0,T]$, so that also $\xi = 0$, and uniqueness is proved.

Step 3. The solution (X,ξ) does *not* have, a priori, any measurability in ω, because of the way it has been constructed. We are going to show that in fact X and ξ are predictable processes. The reasoning for X is simple: with ω fixed, we have proved that from any subsequence of λ one can extract a further subsequence λ', depending on ω, such that the convergences of Proposition 2.4.7 take place, and the limit (X,ξ) is unique. This implies, by a well-known criterion of classical analysis, that the same convergences hold along the original sequence λ, which does *not* depend on ω. The convergence of $X_\lambda(\omega,\cdot)$ to $X(\omega,\cdot)$ in $L^2(0,T;H)$ implies that $X : \Omega \to L^2(0,T;H)$ is measurable and $X_\lambda(\omega,t)$ converges to $X(\omega,t)$ in H in $\mathbb{P}\otimes dt$-measure, hence $X_{\bar\lambda}(\omega,t) \to X(\omega,t)$ in H $\mathbb{P}\otimes dt$-a.e. along a subsequence $\bar\lambda$ of λ. Since X_λ is predictable, being adapted with continuous trajectories in H, we infer that X is predictable. Unfortunately a similar reasoning does *not* work for ξ, because $\xi_\lambda(\omega) := \beta_\lambda(X_\lambda(\omega))$ converges only weakly in $L^1((0,T)\times D)$ for $\mathbb{P}$-a.a. $\omega \in \Omega$.[‖] We shall prove instead that a subsequence of $\xi_\lambda := \beta_\lambda(X_\lambda)$ converges weakly to ξ in $L^1(\Omega\times(0,T)\times D)$. In fact, let $g \in L^\infty((0,T)\times D)$ be arbitrary but

[‖] One may indeed deduce, using Mazur's lemma, that there exists, for each ω in a set of probability one, a sequence $(\tilde\xi_{\mu(\omega)}(\omega))_{\mu(\omega)}$ in the convex envelope of $(\xi_\lambda(\omega))_\lambda$ that converges to $\xi(\omega)$. However, the map $\omega \mapsto \tilde\xi_{\mu(\omega)}(\omega)$ needs not be measurable, hence we cannot infer measurability of its limit ξ.

fixed. Then, setting

$$F_\lambda(\omega) := \int_0^T\int_D \xi_\lambda(\omega,s,x)g(s,x)\,dx\,ds, \qquad F(\omega) := \int_0^T\int_D \xi(\omega,s,x)g(s,x)\,dx\,ds,$$

we have $F_\lambda \to F$ in probability, and we claim that $F_\lambda \to F$ weakly in $L^1(\Omega)$. Let $h \in L^\infty(\Omega)$ be arbitrary but fixed, and introduce the even convex function

$$j_0 := j^*(\cdot/M), \qquad M := \frac{1}{\big(\|g\|_{L^\infty((0,T)\times D)} \vee 1\big)\big(\|h\|_{L^\infty(\Omega)} \vee 1\big)}.$$

Then, by Jensen's inequality,

$$\begin{aligned}
\mathbb{E}\, j_0(F_\lambda h) &= \mathbb{E}\, j_0\bigg(\int_0^T\int_D \xi_\lambda(\omega,s,x)g(s,x)h(\omega)\,dx\,ds\bigg) \\
&\lesssim_{T,|D|} \mathbb{E}\int_0^T\int_D j_0\big(\xi_\lambda(\omega,s,x)g(s,x)h(\omega)\big)\,dx\,ds \\
&\le \mathbb{E}\int_0^T\int_D j^*\big(\xi_\lambda(\omega,s,x)\big)\,dx\,ds,
\end{aligned}$$

where the last term is bounded by a constant independent of λ, as proved in Lemma 2.4.6. Since j_0 inherits the superlinearity at infinity of j^*, the criterion of de la Vallée Poussin implies that $F_\lambda h$ is uniformly integrable, hence, since $F_\lambda h \to Fh$ in probability, that $F_\lambda h \to Fh$ strongly in $L^1(\Omega)$ by Vitali's theorem. As h was arbitrary, this implies that $F_\lambda \to F$ weakly in $L^1(\Omega)$, thus also that $\xi_\lambda \to \xi$ weakly in $L^1(\Omega\times(0,T)\times D)$ by arbitrariness of g. By the canonical identification of $L^1(\Omega\times(0,T)\times D)$ with $L^1(\Omega\times(0,T);L^1(D))$ and Mazur's lemma (see, e.g., [20, 7), p. 360]), there exists a sequence $(\zeta_n)_{n\in\mathbb{N}}$ of convex combinations of (ξ_λ) that converges strongly to ξ in $L^1(D)$ in $\mathbb{P}\otimes dt$-measure, hence $\mathbb{P}\otimes dt$-a.e. passing to a subsequence of n. Since ξ_λ, hence ζ_n, are predictable for all λ and n, respectively, it follows that ξ is a predictable $L^1(D)$-valued process and $\xi : \Omega \to L^1((0,T)\times D))$ is measurable. Moreover, since $X_\lambda(\omega,\cdot) \to X(\omega,\cdot)$ in $L^2(0,T;H)$ for $\mathbb{P}$-a.a. ω and $(X_\lambda)_\lambda$ is bounded in $L^2(\Omega;L^2(0,T;V))$, it follows that $X_\lambda \rightharpoonup X$ in $L^2(\Omega;L^2(0,T;V))$. Therefore, an entirely analogous argument based on Mazur's lemma yields that $X : \Omega \to L^2(0,T;V)$ is measurable.

Step 4. As last step, we are going to show that X and ξ satisfy also estimates in expectation. In particular, the weak and weak* lower semicontinuity of the norm ensures that, for $\mathbb{P}$-almost all $\omega \in \Omega$,

$$\begin{aligned}
\big\|X(\omega,\cdot)\big\|_{L^2(0,T;V)} &\le \liminf_{\lambda\to 0}\big\|X_\lambda(\omega,\cdot)\big\|_{L^2(0,T;V)}, \\
\big\|X(\omega,\cdot)\big\|_{L^\infty(0,T;H)} &\le \liminf_{\lambda\to 0}\big\|X_\lambda(\omega,\cdot)\big\|_{L^\infty(0,T;H)}, \\
\big\|\xi(\omega,\cdot)\big\|_{L^1(Q)} &\le \liminf_{\lambda\to 0}\big\|\beta_\lambda(X_\lambda(\omega,\cdot))\big\|_{L^1(Q)}.
\end{aligned}$$

Taking expectations and recalling Lemmata 2.4.5 and 2.4.6, it follows by Fatou's lemma that, for a constant N,

$$\begin{aligned}
\mathbb{E}\big\|X\big\|^2_{L^2(0,T;V)} &\le \mathbb{E}\liminf_{\lambda\to 0}\big\|X_\lambda\big\|^2_{L^2(0,T;V)} \le \liminf_{\lambda\to 0}\mathbb{E}\big\|X_\lambda\big\|^2_{L^2(0,T;V)} < N, \\
\mathbb{E}\big\|X\big\|^2_{L^\infty(0,T;H)} &\le \mathbb{E}\liminf_{\lambda\to 0}\big\|X_\lambda\big\|^2_{L^\infty(0,T;H)} \le \liminf_{\lambda\to 0}\mathbb{E}\big\|X_\lambda\big\|^2_{L^\infty(0,T;H)} < N, \\
\mathbb{E}\big\|\xi\big\|_{L^1(0,T;L^1(D))} &\le \mathbb{E}\liminf_{\lambda\to 0}\big\|\xi_\lambda\big\|_{L^1(0,T;L^1(D))} \le \liminf_{\lambda\to 0}\mathbb{E}\big\|\xi_\lambda\big\|_{L^1(0,T;L^1(D))} < N,
\end{aligned}$$

i.e.

$$\begin{aligned}&X \in L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V)),\\ &\xi \in L^1(\Omega \times (0,T) \times D).\end{aligned}$$

The proof is thus complete. □

We conclude this section with a corollary that will be used in the following.

Corollary 2.4.8. *There exists a constant N such that*

$$\mathbb{E}\int_0^T\int_D \big(j(X)+j^*(\xi)\big) < N\Big(\big\|X_0\big\|^2_{L^2(\Omega;H)} + \big\|B\big\|^2_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}\Big).$$

Proof. Thanks to Step 3 in the previous proof, there exists a sequence λ, independent of ω, such that $X_\lambda \to X$ a.e. in $(0,T)\times D$ and $\beta_\lambda(X_\lambda) \to \xi$ weakly in $L^1((0,T)\times D)$. Proceeding as in the first part of the proof of Lemma 2.4.6, Lemma 2.4.5 implies that there exists a constant N such that

$$\mathbb{E}\int_0^T\int_D \big(j(I+\lambda\beta)^{-1}X_\lambda) + j^*(\beta_\lambda(X_\lambda))\big)\,dx\,ds < \bar{N}(X_0,B),$$

where $\bar{N}(X_0,B) := N\big(\|X_0\|^2_{L^2(\Omega;H)} + \|B\|^2_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}\big)$. Therefore, in analogy to Step 4 of the previous proof, two applications of Fatou's lemma yield

$$\mathbb{E}\int_0^T\int_D j(X) \leq \liminf_{\lambda\to 0}\mathbb{E}\int_0^T\int_D j((I+\lambda\beta)^{-1}X_\lambda) < \bar{N}(X_0,B),$$

as well as, by the weak lower semicontinuity of convex integrals and Fatou's lemma again,

$$\mathbb{E}\int_0^T\int_D j^*(\xi) \leq \liminf_{\lambda\to 0}\mathbb{E}\int_0^T\int_D j^*(\beta_\lambda(X_\lambda)) < \bar{N}(X_0,B).$$ □

2.5 Well-posedness with additive noise

In this section we prove well-posedness for the equation

$$dX(t) + AX(t)\,dt + \beta(X(t))\,dt \ni B(t)\,dW(t), \qquad X(0) = X_0, \tag{2.5.10}$$

where B is an $\mathscr{L}^2(U,H)$-valued process. Note that this is just equation (2.1.1) with additive noise.

Proposition 2.5.1. *Assume that $X_0 \in L^2(\Omega,\mathscr{F}_0,\mathbb{P};H)$ and that*

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

is measurable and adapted. Then equation (2.5.10) *is well posed in $\mathscr{J}$. Moreover, $X(\omega,\cdot) \in C_w([0,T];H)$ for $\mathbb{P}$-almost all $\omega \in \Omega$.*

Proof. We shall proceed in several steps: first we approximate the coefficient B in such a way that the corresponding equation can be uniquely solved by the methods of the previous section. Then we pass to the limit in an appropriate way, obtaining a solution to (2.5.10), which is then shown to be unique.

STEP 1. By Assumption A(iv), there exists $m \in \mathbb{N}$ such that $(I+A)^{-m}$ maps continuously L^1 to L^∞. The space $V_0 := \mathsf{D}(A^m)$, endowed with inner product

$$\langle u, v\rangle_{V_0} := \langle u, v\rangle_H + \langle A^m u, A^m v\rangle_H, \qquad u, v \in \mathsf{D}(A^m),$$

is a Hilbert space densely and continuously embedded in V. Moreover, the diagram

$$\mathsf{D}(A^m) \xrightarrow{(I+A)^m} L^1(D) \xrightarrow{(I+A)^{-m}} L^\infty(D)$$

immediately shows that V_0 is also continuously embedded in $L^\infty(D)$. In particular, all hypotheses on V_0 of the previous section are met. Moreover, by the ideal property of Hilbert-Schmidt operators, setting, for any $\varepsilon > 0$,

$$B^\varepsilon := (I + \varepsilon A)^{-m} B,$$

we have $B^\varepsilon \in L^2(\Omega; L^2(0,T; \mathscr{L}^2(U, V_0)))$. Then it follows by Proposition 2.4.1 that, for any $\varepsilon > 0$, there exist predictable processes

$$\begin{aligned} X^\varepsilon &\in L^2(\Omega; L^\infty(0,T;H)) \cap L^2(\Omega; L^2(0,T;V)), \\ \xi^\varepsilon &\in L^1(\Omega \times (0,T) \times D), \end{aligned}$$

with $X^\varepsilon(\omega, \cdot) \in C_w([0,T]; H)$ for $\mathbb{P}$-almost all $\omega \in \Omega$, such that

$$X^\varepsilon(t) + \int_0^t AX^\varepsilon(s)\,ds + \int_0^t \xi^\varepsilon(s)\,ds = X_0 + \int_0^t B^\varepsilon(s)\,dW(s) \tag{2.5.11}$$

in $V^* \cap L^1(D)$ for all $t \in [0,T]$. Moreover, $\xi^\varepsilon \in \beta(X^\varepsilon)$ a.e. in $(0,T) \times D$ and $j(X^\varepsilon) + j^*(\xi^\varepsilon) \in L^1((0,T) \times D)$ $\mathbb{P}$-almost surely.

STEP 2. For any $\varepsilon > 0$, the equation in V^*

$$X^\varepsilon_\lambda(t) + \int_0^t AX^\varepsilon_\lambda(s)\,ds + \int_0 \beta_\lambda(X^\varepsilon_\lambda(s))\,ds = X_0 + \int_0^t B^\varepsilon(s)\,dW(s)$$

admits a unique (variational) strong solution X^ε_λ. Taking into account the coercivity of A and the monotonicity of β_λ, Itô's formula yields, for any $\delta > 0$,

$$\begin{aligned} &\big\|X^\varepsilon_\lambda(t) - X^\delta_\lambda(t)\big\|_H^2 + \int_0^t \big\|X^\varepsilon_\lambda(s) - X^\delta_\lambda(s)\big\|_V^2\,ds \\ &\lesssim \int_0^t \big(X^\varepsilon_\lambda(s) - X^\delta_\lambda(s)\big)\big(B^\varepsilon(s) - B^\delta(s)\big)\,dW(s) + \int_0^t \big\|B^\varepsilon(s) - B^\delta(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds. \end{aligned}$$

Taking supremum in time and expectation, it easily follows from Lemma 1.5.1 that

$$\begin{aligned} \big\|X^\varepsilon_\lambda - X^\delta_\lambda\big\|_{L^2(\Omega; L^\infty(0,T;H))} &+ \big\|X^\varepsilon_\lambda - X^\delta_\lambda\big\|_{L^2(\Omega; L^2(0,T;V))} \\ &\lesssim \big\|B^\varepsilon - B^\delta\big\|_{L^2(\Omega; L^2(0,T; \mathscr{L}^2(U,H)))}. \end{aligned}$$

On the other hand, the proof of Proposition 2.4.1 shows that there exists a sequence λ, independent of ε, such that, for $\mathbb{P}$-almost all $\omega \in \Omega$,

$$
\begin{aligned}
X^\varepsilon_\lambda(\omega,\cdot) &\overset{*}{\rightharpoonup} X^\varepsilon(\omega,\cdot) && \text{in } L^\infty(0,T;H),\\
X^\varepsilon_\lambda(\omega,\cdot) &\rightharpoonup X^\varepsilon(\omega,\cdot) && \text{in } L^2(0,T;V),\\
\beta_\lambda(X^\varepsilon_\lambda(\omega,\cdot)) &\rightharpoonup \xi^\varepsilon(\omega,\cdot) && \text{in } L^1((0,T)\times D)
\end{aligned}
$$

as $\lambda \to 0$. Since the weak* limit in $L^\infty(0,T;H)$ as $\lambda \to 0$ of $X^\varepsilon_\lambda - X^\delta_\lambda$ is $X^\varepsilon - X^\delta$, the weak* lower semicontinuity of the norm implies

$$
\big\|X^\varepsilon - X^\delta\big\|_{L^\infty(0,T;H)} \leq \liminf_{\lambda\to 0}\big\|X^\varepsilon_\lambda - X^\delta_\lambda\big\|_{L^\infty(0,T;H)},
$$

thus also, by Fatou's lemma,

$$
\begin{aligned}
\mathbb{E}\big\|X^\varepsilon - X^\delta\big\|^2_{L^\infty(0,T;H)} &\leq \mathbb{E}\liminf_{\lambda\to 0}\big\|X^\varepsilon_\lambda - X^\delta_\lambda\big\|^2_{L^\infty(0,T;H)}\\
&\lesssim \mathbb{E}\big\|B^\varepsilon - B^\delta\big\|^2_{L^2(0,T;\mathscr{L}^2(U,H))}.
\end{aligned}
$$

An entirely similar argument yields

$$
\mathbb{E}\big\|X^\varepsilon - X^\delta\big\|^2_{L^2(0,T;V)} \lesssim \mathbb{E}\big\|B^\varepsilon - B^\delta\big\|^2_{L^2(0,T;\mathscr{L}^2(U,H))},
$$

so that

$$
\begin{aligned}
&\big\|X^\varepsilon - X^\delta\big\|_{L^2(\Omega;L^\infty(0,T;H))} + \big\|X^\varepsilon - X^\delta\big\|_{L^2(\Omega;L^2(0,T;V))}\\
&\qquad\lesssim \big\|B^\varepsilon - B^\delta\big\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}.
\end{aligned}
$$

Taking into account that $\big\|B^\varepsilon - B\big\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))} \to 0$ as $\varepsilon \to 0$, it follows that (X^ε) is a Cauchy sequence in $E := L^2(\Omega;L^\infty(0,T;H)) \cap L^2(\Omega;L^2(0,T;V))$, hence there exists $X \in E$ such that X^ε converges (strongly) to X in E as $\varepsilon \to 0$. In particular, the limit process X is predictable. Moreover, by Corollary 2.4.8, there exists a constant N such that

$$
\begin{aligned}
\mathbb{E}\int_0^T\!\!\int_D \big(j(X^\varepsilon) + j^*(\xi^\varepsilon)\big)\,dx\,ds &< N\Big(\big\|X_0\big\|^2_{L^2(\Omega;H)} + \big\|B^\varepsilon\big\|^2_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}\Big)\\
&\leq N\Big(\big\|X_0\big\|^2_{L^2(\Omega;H)} + \big\|B\big\|^2_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}\Big),
\end{aligned}
\tag{2.5.12}
$$

as it follows by the ideal property of Hilbert-Schmidt operators and the contractivity of $(I + \varepsilon A)^{-1}$. The criterion by de la Vallée Poussin then implies that (ξ^ε) is uniformly integrable on $\Omega \times (0,T) \times D$, hence, by the Dunford-Pettis theorem, (ξ^ε) is weakly relatively compact in $L^1(\Omega \times (0,T) \times D)$. Therefore, passing to a subsequence of ε, denoted by the same symbol, there exists ξ belonging to the latter space such that $\xi^\varepsilon \to \xi$ therein in the weak topology. In particular, by an argument based on Mazur's lemma, entirely analogous to that used in Step 3 of the proof of Proposition 2.4.1, one infers that ξ is a predictable process.

Step 3. We can now pass to the limit as $\varepsilon \to 0$ in Equation (2.5.11), by a reasoning analogous to the one use in Step 1 of the proof of Proposition 2.4.1. As proved in the previous step, X^ε converges strongly to X in $L^2(\Omega;L^\infty(0,T;H))$, hence

$$
\operatorname*{ess\,sup}_{t\in[0,T]}\big\|X^\varepsilon(t) - X(t)\big\|_H \to 0
$$

in probability as $\varepsilon \to 0$. Let $\phi_0 \in V_0$ be arbitrary. Since $V_0 \hookrightarrow L^\infty(D)$, one has

$$\langle X^\varepsilon(t), \phi_0 \rangle \to \langle X(t), \phi_0 \rangle$$

in probability for almost all $t \in [0,T]$. Let us set, for an arbitrary but fixed $t \in [0,T]$, $\phi : s \mapsto 1_{[0,t]}(s)\phi_0 \in L^2(0,T;V)$, so that $A\phi \in L^2(0,T;V^*)$. Recalling that $X^\varepsilon \to X$ (strongly, hence also weakly) in $L^2(\Omega;L^2(0,T;V))$, it follows immediately that $X^\varepsilon \rightharpoonup X$ in $L^2(0,T;V)$ in measure, hence

$$\begin{aligned}\int_0^t \langle AX^\varepsilon, \phi_0\rangle \, ds = \int_0^T \langle AX^\varepsilon(s), \phi(s)\rangle \, ds &= \int_0^T \langle X^\varepsilon(s), A\phi(s)\rangle \, ds \\ \to \int_0^T \langle X(s), A\phi(s)\rangle \, ds &= \int_0^t \langle AX(s), \phi_0\rangle \, ds\end{aligned}$$

in probability as $\varepsilon \to 0$. A completely analogous reasoning shows that

$$\int_0^t \langle \xi^\varepsilon(s), \phi_0\rangle \, ds \to \int_0^t \langle \xi(s), \phi_0\rangle \, ds$$

in probability as $\varepsilon \to 0$. Doob's maximal inequality and the convergence

$$\big\| B^\varepsilon - B \big\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))} \xrightarrow{\varepsilon\to 0} 0$$

readily yield also that $B^\varepsilon \cdot W(t) \to B \cdot W(t)$ in H in probability for all $t \in [0,T]$. In particular, since $\phi_0 \in V_0$ and $t \in [0,T]$ are arbitrary, we infer that

$$X(t) + \int_0^t AX(s)\, ds + \int_0^t \xi(s)\, ds = X_0 + \int_0^t B(s)\, dW(s)$$

holds in V_0^* for almost all t. Recalling that $\xi \in L^1(0,T;L^1(D)) \hookrightarrow L^1(0,T;V_0^*)$, so that all terms except the first on the left-hand side have trajectories in the space $C([0,T];V_0^*)$, we conclude that the identity holds for all $t \in [0,T]$. Moreover, thanks to Lemma 1.4.1, $X \in C([0,T];V_0^*)$ and $X \in L^\infty(0,T;H)$ imply that $X \in C_w([0,T];H)$. Note also that all terms bar the second [third] one on the left-hand side are $L^1(D)$-valued [V^*-valued], hence the identity holds in $L^1(D) \cap V^*$ for all $t \in [0,T]$.

Step 4. Convergence of $X^\varepsilon \to X$ in $L^2(\Omega;L^\infty(0,T;H))$ implies convergence in measure in $\Omega \times (0,T) \times D$, hence, by Fatou's lemma, (2.5.12) yields

$$\mathbb{E}\int_0^T\!\!\int_D j(X) < \bar{N}(X_0,B),$$

where $\bar{N}(X_0,B)$ is the constant appearing in the last term of (2.5.12). Similarly, since $\xi^\varepsilon \to \xi$ weakly in $L^1(\Omega \times (0,T) \times D)$, (2.5.12) and the weak lower semicontinuity of convex integrals yield

$$\mathbb{E}\int_0^T\!\!\int_D j^*(\xi) < \bar{N}(X_0,B).$$

To complete the proof of existence, we only need to show that $\xi \in \beta(X)$ a.e. in $\Omega \times (0,T) \times D$. Note that, passing to a subsequence of ε, still denoted by the same symbol, we have $X^\varepsilon \to X$ a.e. in $\Omega \times (0,T) \times D$. Recalling that $\xi^\varepsilon \in \beta(X^\varepsilon)$ a.e. in $\Omega \times (0,T) \times D$, (2.5.12) again implies

$$\mathbb{E}\int_0^T\!\!\int_D X^\varepsilon \xi^\varepsilon = \mathbb{E}\int_0^T\!\!\int_D \big(j(X^\varepsilon) + j^*(\xi^\varepsilon)\big) < \bar{N}(X_0,B).$$

It follows by monotonicity that $X^\varepsilon \xi^\varepsilon \geq 0$, hence $X^\varepsilon \xi^\varepsilon \in L^1(\Omega \times (0,T) \times D)$. Brézis' Lemma 1.3.14 then yields $\xi \in \beta(X)$ a.e. in $\Omega \times (0,T) \times D$.

Uniqueness and continuous dependence of the solution on the initial datum is an immediate consequence of the next result. $\square$

We first need to introduce weighted (in time) versions of some spaces of processes. For any $p \in [1,\infty]$ and $\alpha \geq 0$, we shall denote by $L^p_\alpha(0,T)$ the space $L^p(0,T)$ endowed with the norm $\|f\|_{L^p_\alpha(0,T)} := \|t \mapsto e^{-\alpha t} f(t)\|_{L^p(0,T)}$. It is clear that $L^p(0,T)$ and $L^p_\alpha(0,T)$, for different values of α, are all isomorphic (their norms are equivalent). Completely similar notation will be used for vector-valued L^p and L^p_α spaces. For typographical economy, restricted only to the formulation of the following proposition, let us define the Banach space

$$F_\alpha := L^2(\Omega; L^\infty_\alpha(0,T;H)) \cap L^2(\Omega; L^2_\alpha(0,T;V)),$$

endowed with the norm

$$\|\cdot\|_{F_\alpha} := \|\cdot\|_{L^2(\Omega;L^\infty_\alpha(0,T;H)) \cap L^2(\Omega;L^2_\alpha(0,T;V))} + \sqrt{\alpha}\|\cdot\|_{L^2(\Omega;L^2_\alpha(0,T;H))}$$

Proposition 2.5.2. *Let (X_1,ξ_1), $(X_2,\xi_2) \in \mathscr{J}$ be solutions to* (2.5.10) *with initial values X_{01}, $X_{02} \in L^2(\Omega,\mathscr{F}_0;H)$ and diffusion coefficients B_1, $B_2 \in L^2(\Omega; L^2(0,T;\mathscr{L}^2(U,H)))$, respectively. Then, for any $\alpha \geq 0$,*

$$\|X_1 - X_2\|_{F_\alpha} \lesssim \|X_{01} - X_{02}\|_{L^2(\Omega;H)} + \|B_1 - B_2\|_{L^2(\Omega;L^2_\alpha(0,T;\mathscr{L}^2(U,H)))}.$$

In particular, there is a unique solution $(X,\xi) \in \mathscr{J}$ to (2.5.10).

Proof. Setting

$$Y := X_1 - X_2, \qquad Y_0 := X_{01} - X_{02}, \qquad G := B_1 - B_2,$$

one has

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

in $V^* \cap L^1(D)$, where $\zeta := \xi_1 - \xi_2$, and ξ_1, ξ_2 are defined in the obvious way. By the hypotheses on A, there exists $m \in \mathbb{N}$ such that, using the notation $h^\delta := (I+\delta A)^{-m}h$ for any h for which it makes sense,

$$AY^\delta,\ \zeta^\delta \in L^1(\Omega; L^1(0,T;H)),$$

while Y_0^δ and G^δ have the same integrability properties of Y, Y_0 and G, respectively. In particular, we have

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

in V^*. Let $\alpha > 0$ be arbitrary but fixed, and add a superscript α to any process that is multiplied pointwise by the function $t \mapsto e^{-\alpha t}$. The integration by parts formula yields

$$Y^{\delta,\alpha}(t) + \int_0^t (A+\alpha I)Y^{\delta,\alpha}(s)\,ds + \int_0^t \zeta^{\delta,\alpha}(s)\,ds = Y_0^\delta + \int_0^t G^{\delta,\alpha}(s)\,dW(s),$$

to which we can apply Itô's formula for the square of the norm in H, obtaining, using the coercivity of A,

$$\begin{aligned}\big\|Y^{\delta,\alpha}(t)\big\|_H^2 + 2\alpha\int_0^t \big\|Y^{\delta,\alpha}(s)\big\|_H^2\,ds + 2C\int_0^t \big\|Y^{\delta,\alpha}(s)\big\|_V^2\,ds\\ + 2\int_0^t \big\langle Y^{\delta,\alpha}(s),\zeta^{\delta,\alpha}(s)\big\rangle\,ds\\ \leq \big\|Y_0^\delta\big\|_H^2 + \int_0^t Y^{\delta,\alpha}(s)G^{\delta,\alpha}(s)\,dW(s) + \int_0^t \big\|G^{\delta,\alpha}(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds.\end{aligned}$$

We are now going to pass to the limit as $\delta\to 0$: the first term on the left-hand side and on the right-hand side clearly converge to $\|Y^\alpha(t)\|_H^2$ and $\|Y_0\|_H^2$, respectively. Since $(I+\delta A)^{-1}$ converges to the identity in H as well as in V in the strong operator topology, the dominated convergence theorem yields

$$\begin{aligned}\int_0^t \big\|Y^{\delta,\alpha}(s)\big\|_V^2\,ds &\longrightarrow \int_0^t \big\|Y^\alpha(s)\big\|_V^2\,ds,\\ \int_0^t \big\|G^{\delta,\alpha}(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds &\longrightarrow \int_0^t \big\|G^\alpha(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds\end{aligned}$$

as $\delta\to 0$ for all $t\in[0,T]$. Defining the real local martingales

$$M^{\delta,\alpha} := (Y^{\delta,\alpha}G^{\delta,\alpha})\cdot W, \qquad M^\alpha := (Y^\alpha G^\alpha)\cdot W,$$

in order to establish convergence in probability (uniformly on compact sets) of the sequence $M^{\delta,\alpha}$ to M^α as $\delta\to 0$, it is sufficient to show that $[M^{\delta,\alpha}-M^\alpha, M^{\delta,\alpha}-M^\alpha]_T$ converges to zero in probability. To this purpose, note that

$$\begin{aligned}[M^{\delta,\alpha}-M^\alpha, M^{\delta,\alpha}-M^\alpha]_T^{1/2} &= \big\|Y^{\delta,\alpha}G^{\delta,\alpha}-Y^\alpha G^\alpha\big\|_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))}\\ &\leq \big\|Y^{\delta,\alpha}G^{\delta,\alpha}-Y^{\delta,\alpha}G^\alpha\big\|_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))}\\ &\quad + \big\|Y^{\delta,\alpha}G^\alpha-Y^\alpha G^\alpha\big\|_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))},\end{aligned}$$

where

$$\big\|Y^{\delta,\alpha}(t)G^{\delta,\alpha}(t)-Y^{\delta,\alpha}(t)G^\alpha(t)\big\|_{\mathscr{L}^2(U,\mathbb{R}))} \leq \big\|Y^\alpha(t)\big\|_H\big\|G^{\delta,\alpha}(t)-G^\alpha(t)\big\|_{\mathscr{L}^2(U,H))}$$

for all $t\in[0,T]$. Since the right-hand side converges to 0 as $\delta\to 0$ and it is bounded by $2\|Y^\alpha\|_{L^\infty(0,T;H)}\|G^\alpha(t)\|_{\mathscr{L}^2(U,H)}$, and $G^\alpha\in L^2(0,T;\mathscr{L}^2(U,H))$, the dominated convergence theorem yields

$$\big\|Y^{\delta,\alpha}G^{\delta,\alpha}-Y^{\delta,\alpha}G^\alpha\big\|_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))}\to 0$$

as $\delta\to 0$. Similarly, $\big\|Y^{\delta,\alpha}G^\alpha-Y^\alpha G^\alpha\big\|_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))}$ tends to 0 as $\delta\to 0$ by completely analogous argument.

We are now going to show that $Y^{\delta,\alpha}\zeta^{\delta,\alpha}\to Y^\alpha\zeta^\alpha$ in $L^1(\Omega\times(0,T)\times D)$, which clearly implies that

$$\int_0^t\int_D Y^{\delta,\alpha}\zeta^{\delta,\alpha}\to\int_0^t\int_D Y^\alpha\zeta^\alpha$$

in probability for all $t\in[0,T]$. Since $Y^{\delta,\alpha}\to Y^\alpha$ and $\zeta^{\delta,\alpha}\to\zeta^\alpha$ in measure in $\Omega\times(0,T)\times D$, Vitali's theorem implies strong convergence in L^1 if the sequence $(Y^{\delta,\alpha}\zeta^{\delta,\alpha})$ is uniformly integrable in $\Omega\times(0,T)\times D$. In turn, the latter is certainly true if $\big(|Y^{\delta,\alpha}\zeta^{\delta,\alpha}|\big)$ is dominated by a sequence that converges strongly in L^1. In order to prove this property, note that j and j^*

are increasing on $\mathbb{R}_+$, hence

$$\begin{aligned}\frac{1}{4}\bigl|Y^{\delta,\alpha}(\omega,t,x)\zeta^{\delta,\alpha}(\omega,t,x)\bigr| &\leq j\bigl(e^{-\alpha t}|Y^{\delta}(\omega,t,x)|/2\bigr) + j^*\bigl(e^{-\alpha t}|\zeta^{\delta}(\omega,t,x)|/2\bigr)\\ &\leq j\bigl(|Y^{\delta}(\omega,t,x)|/2\bigr) + j^*\bigl(|\zeta^{\delta}(\omega,t,x)|/2\bigr),\end{aligned}$$

so that, by the symmetry of j and j^*, and by the Jensen inequality of Lemma 1.3.15,

$$\frac{1}{4}\bigl|Y^{\delta,\alpha}\zeta^{\delta,\alpha}\bigr| \leq j(Y^\delta/2) + j^*(\zeta^\delta/2) \leq (I+\delta A)^{-m}\bigl(j(Y/2) + j^*(\zeta/2)\bigr),$$

where, by convexity and symmetry,

$$j(Y/2) = j\Bigl(\frac{1}{2}X_1 + \frac{1}{2}(-X_2)\Bigr) \leq \frac{1}{2}\bigl(j(X_1) + j(X_2)\bigr) \in L^1(\Omega\times(0,T)\times D),$$

and, completely analogously,

$$j^*(\zeta/2) \leq \frac{1}{2}\bigl(j^*(\xi_1) + j^*(\xi_2)\bigr) \in L^1(\Omega\times(0,T)\times D),$$

hence

$$\bigl|Y^{\delta,\alpha}\zeta^{\delta,\alpha}\bigr| \lesssim (I+\delta A)^{-m}\bigl(j(X_1) + j(X_2) + j^*(\xi_1) + j^*(\xi_2)\bigr).$$

Since the right-hand side of this expression converges strongly in $L^1(\Omega\times(0,T)\times D)$ as $\delta\to 0$, it is, a fortiori, uniformly integrable, and so is the left-hand side.

We have thus obtained

$$\begin{aligned}&\bigl\|Y^\alpha(t)\bigr\|_H^2 + 2\alpha\int_0^t \bigl\|Y^\alpha(s)\bigr\|_H^2\,ds + 2\int_0^t \mathscr{E}\bigl(Y^\alpha(s),Y^\alpha(s)\bigr)\,ds\\ &\quad + 2\int_0^t\!\!\int_D Y^\alpha(s,x)\zeta^\alpha(s,x)\,dx\,ds\\ &\qquad \leq \bigl\|Y_0\bigr\|_H^2 + \int_0^t Y^\alpha(s)G^\alpha(s)\,dW(s) + \int_0^t \bigl\|G^\alpha(s)\bigr\|^2_{\mathscr{L}^2(U,H)}\,ds,\end{aligned}$$

where, by monotonicity, $Y^\alpha\zeta^\alpha = e^{-2\alpha\cdot}(X_1-X_2)(\xi_2-\xi_2)\geq 0$, hence, taking the $L^\infty(0,T)$ norm and expectation on both sides,

$$\begin{aligned}&\bigl\|Y^\alpha\bigr\|_{L^2(\Omega;L^\infty(0,T;H))} + \sqrt{\alpha}\bigl\|Y^\alpha\bigr\|_{L^2(\Omega;L^2(0,T;H))} + \bigl\|Y^\alpha\bigr\|_{L^2(\Omega;L^2(0,T;V))}\\ &\qquad \lesssim \bigl\|Y_0\bigr\|_{L^2(\Omega;H)} + \biggl(\mathbb{E}\sup_{t\leq T}\biggl|\int_0^t Y^\alpha(s)G^\alpha(s)\,dW(s)\biggr|\biggr)^{1/2}\\ &\qquad\quad + \bigl\|G^\alpha\bigr\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}.\end{aligned}$$

By Lemma 1.5.1, one has

$$\begin{aligned}\biggl(\mathbb{E}\sup_{t\leq T}\biggl|\int_0^t Y^\alpha(s)G^\alpha(s)\,dW(s)\biggr|\biggr)^{1/2} &\leq \varepsilon\bigl\|Y^\alpha\bigr\|_{L^2(\Omega;L^\infty(0,T;H))}\\ &\quad + N(\varepsilon)\bigl\|G^\alpha\bigr\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))},\end{aligned}$$

with $\varepsilon>0$ arbitrary. Choosing ε sufficiently small and rearranging terms, one obtains

$$\|X_1 - X_2\|_{F_\alpha} \lesssim \|X_{01} - X_{02}\|_{L^2(\Omega;H)} + \|B_1 - B_2\|_{L^2(\Omega;L^2_\alpha(0,T;\mathscr{L}^2(U,H)))}$$

as claimed.

Choosing $\alpha = 0$, $X_{01} = X_{02}$, and $B_1 = B_2$, one gets immediately $X_1 = X_2$, hence also, by substitution,

$$\int_0^t (\xi_1(s) - \xi_2(s))\,ds = 0 \qquad \forall t \in [0,T],$$

which implies uniqueness of ξ. $\square$

2.6 Proof of the main result

For every progressively measurable process $Y \in L^2(\Omega; L^2(0,T;H))$ and initial datum $X_0 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P}; H)$, we consider the equation

$$dX(t) + AX(t)\,dt + \beta(X(t))\,dt \ni B(t, Y(t))\,dW(t), \qquad X(0) = X_0. \tag{2.6.13}$$

Since $B(\cdot, Y)$ is U-measurable, adapted, and belongs to $L^2(\Omega; L^2(0,T; \mathscr{L}^2(U,H)))$, the above equation is well-posed in $\mathscr{J}$ by Proposition 2.5.1, hence one can define a map

$$\begin{aligned} \Gamma : L^2(\Omega; H) \times L^2(\Omega; L^2(0,T;H)) &\longrightarrow L^2(\Omega; L^2(0,T;H)) \times L^1(\Omega \times (0,T) \times D) \\ (X_0, Y) &\longmapsto (X, \xi), \end{aligned}$$

where (X,ξ) is the unique process in $\mathscr{J}$ solving (2.6.13). Denote the $L^2(\Omega; L^2(0,T;H))$-valued component of Γ by Γ_1 and the $L^1(\Omega \times (0,T) \times D)$-valued component by Γ_2: we are going to show that $Y \mapsto \Gamma_1(X_0, Y)$ is a (strict) contraction of $L^2(\Omega; L^2(0,T;H))$, if endowed with a suitably chosen equivalent norm. Let $X_i = \Gamma_1(X_{0i}, Y_i)$, $i = 1,2$, with obvious meaning of the symbols. For any $\alpha \geq 0$, Proposition 2.5.2 yields

$$\begin{aligned} &\big\|X_1 - X_2\big\|_{L^2(\Omega; L^\infty_\alpha(0,T;H)) \cap L^2(\Omega; L^2_\alpha(0,T;V))} + \sqrt{\alpha}\big\|X_1 - X_2\big\|_{L^2(\Omega; L^2_\alpha(0,T;H))} \\ &\lesssim \big\|X_{01} - X_{02}\big\|_{L^2(\Omega;H)} + \big\|B(\cdot, Y_1) - B(\cdot, Y_2)\big\|_{L^2(\Omega; L^2_\alpha(0,T;\mathscr{L}^2(U,H)))}, \end{aligned} \tag{2.6.14}$$

in particular, by the Lipschitz continuity of B,

$$\begin{aligned} \big\|X_1 - X_2\big\|_{L^2(\Omega; L^2_\alpha(0,T;H))} &\lesssim \frac{1}{\sqrt{\alpha}}\big\|X_{01} - X_{02}\big\|_{L^2(\Omega;H)} \\ &\quad + \frac{1}{\sqrt{\alpha}}\big\|B(\cdot, Y_1) - B(\cdot, Y_2)\big\|_{L^2(\Omega; L^2_\alpha(0,T;\mathscr{L}^2(U,H)))} \\ &\lesssim \frac{1}{\sqrt{\alpha}}\Big(\big\|X_{01} - X_{02}\big\|_{L^2(\Omega;H)} + \big\|Y_1 - Y_2\big\|_{L^2(\Omega; L^2_\alpha(0,T;H))}\Big), \end{aligned} \tag{2.6.15}$$

where the implicit constant does not depend on α. In particular, if $X_{01} = X_{02}$, choosing α large enough, one has that, for any $X_0 \in L^2(\Omega, H)$, $Y \mapsto \Gamma_1(X_0, Y)$ is a contraction of $L^2(\Omega; L^2_\alpha(0,T;H))$. It follows by the Banach fixed-point theorem that $\Gamma_1(X_0, \cdot)$ has a unique fixed point X therein, hence also in $L^2(\Omega; L^2(0,T;H))$ by equivalence of norms. Setting $\xi := \Gamma_2(X_0, X)$, by definition of the map Γ, (X,ξ) is a solution to (2.1.1) and it belongs to $\mathscr{J}$.

Let X_{01}, $X_{02} \in L^2(\Omega, \mathscr{F}_0; H)$ and X_1, X_2 be the unique fixed points of the maps $\Gamma_1(X_{0i}, \cdot)$, $i = 1,2$, respectively, and $\xi_i := \Gamma_2(X_{0i}, X_i)$, $i = 1,2$. Replacing Y_i with $X_i = \Gamma_1(X_{0i}, X_i)$, $i = 1,2$, in (2.6.15) yields

$$\big\|X_1 - X_2\big\|_{L^2(\Omega; L^2_\alpha(0,T;H))} \leq C_1\big\|X_{01} - X_{02}\big\|_{L^2(\Omega;H)} + C_2\big\|X_1 - X_2\big\|_{L^2(\Omega; L^2_\alpha(0,T;H))},$$

with $C_1 > 0$, $C_2 \in]0,1[$, hence, by equivalence of norms,

$$\|X_1 - X_2\|_{L^2(\Omega;L^2(0,T;H))} \lesssim \|X_{01} - X_{02}\|_{L^2(\Omega;H)}.$$

This implies, substituting Y_i with $X_i = \Gamma(X_{0i}, X_i)$, $i = 1, 2$, in (2.6.14), with $\alpha = 0$,

$$\begin{aligned}
&\|X_1 - X_2\|_{L^2(\Omega;L^\infty(0,T;H)) \cap L^2(\Omega;L^2(0,T;V))} \\
&\qquad \lesssim \|X_{01} - X_{02}\|_{L^2(\Omega;H)} + \|B(\cdot, X_1) - B(\cdot, X_2)\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))} \\
&\qquad \lesssim \|X_{01} - X_{02}\|_{L^2(\Omega;H)} + \|X_1 - X_2\|_{L^2(\Omega;L^2(0,T;H))} \\
&\qquad \lesssim \|X_{01} - X_{02}\|_{L^2(\Omega;H)}.
\end{aligned}$$

Choosing $\alpha = 0$ and $X_{01} = X_{02}$, one gets immediately $X_1 = X_2$, hence also, by substitution,

$$\int_0^t (\xi_1(s) - \xi_2(s))\, ds = 0 \qquad \forall t \in [0, T],$$

which implies uniqueness of ξ.

Chapter 3

Singular semilinear equations: refined well-posedness

In this chapter, we prove existence/uniqueness of solutions to stochastic semilinear evolution equations with monotone nonlinear drift and multiplicative noise, assuming the initial datum to be only measurable and allowing the diffusion coefficient to be locally Lipschitz-continuous. Moreover, we show how the finiteness of the p-th moment of solutions depends on the integrability of the initial datum, in the whole range $p \in]0, \infty[$. Lipschitz continuity of the solution map in p-th moment is established, under a Lipschitz continuity assumption on the diffusion coefficient, in the even larger range $p \in [0, \infty[$.

The results presented in this chapter are part of a joint work with Carlo Marinelli: see [63].

3.1 The problem: literature and main goals

We study semilinear stochastic partial differential equations on a smooth bounded domain $D \subseteq \mathbb{R}^d$ of the form

$$dX_t + AX_t\,dt + \beta(X_t)\,dt \ni B(t, X_t)\,dW_t, \qquad X(0) = X_0, \tag{3.1.1}$$

where A is a linear coercive maximal monotone operator on (a subspace of) $H := L^2(D)$, β is a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ defined everywhere, W is a cylindrical Wiener process on a separable Hilbert space U, and B is a process taking values in the space of Hilbert-Schmidt operators from U to $L^2(D)$ satisfying a (local) Lipschitz continuity condition. Precise assumptions on the data of the problem are given in §3.2 below.

Assuming that the initial datum X_0 has finite second moment and the diffusion coefficient B is globally Lipschitz continuous, we proved in Chapter 2 that equation (3.1.1) admits a unique solution, in a generalized variational sense, whose trajectories are weakly continuous in H. The contribution of this chapter is to extend these results in several directions. As a first step we show that the solution X is pathwise strongly continuous in H, rather than just weakly continuous. This is possible thanks to an Itô-type formula, interesting in its own right, for the square of the H-norm of processes satisfying minimal integrability conditions, in a variational setting extending the classical one by Pardoux [72]. The strong pathwise continuity allows us to prove that existence and uniqueness of solutions to (3.1.1) continues to hold under much weaker assumptions on the initial datum and on the diffusion coefficient. In particular, X_0 needs only be measurable and B can be locally Lipschitz-continuous with linear growth. Denoting by Ω

the underlying probability space, the solution map $X_0 \mapsto X$ is thus defined on $L^0(\Omega; H)$, with codomain contained in $L^0(\Omega; E)$, where E is a suitable path space. By the results of Chapter 2 we also have that the solution map restricted to $L^2(\Omega; H)$ has codomain contained in $L^2(\Omega; E)$. As a further result, we extrapolate these mapping properties to the whole range of exponents $p \in [0, \infty[$, that is, we show that if $X_0 \in L^p(\Omega; H)$ then $X \in L^p(\Omega; E)$ for every positive finite p, and we provide an explicit upper bound on the $L^p(\Omega; E)$-norm of the solution in terms of the $L^p(\Omega; H)$-norm of the initial datum. If, in addition, B is Lipschitz-continuous, we show that the solution map is Lipschitz-continuous from $L^p(\Omega; H)$ to $L^p(\Omega; E)$ for all $p \in [0, \infty[$. In the particular case $p = 0$, this implies that solutions converge uniformly on $[0, T]$ in probability if the corresponding initial data converge in probability.

In the classical variational theory of SPDEs, existence and uniqueness of solutions under a local Lipschitz condition on B and measurability of X_0 were obtained by Pardoux in [72]. Our results do not follow from his, however, as equation (3.1.1) cannot be cast in the usual variational setting. Stochastic equations where *all* nonlinear terms are locally Lipschitz-continuous have been considered in the semigroup approach (see, e.g., [47] and references therein), but our existence results are not covered, as β can be discontinuous and have arbitrary growth. Moreover, the properties of the solution map between $L^p(\Omega; H)$ and $L^p(\Omega; E)$ do not seem to have been addressed even in the classical variational setting. On the other hand, the continuity of the solution map in the case $p = 0$ for ordinary SDEs in $\mathbb{R}^n$ with Lipschitz coefficients has been studied, also with very general semimartingale noise (see, e.g., [35]).

The chapter is organized as follows. In §3.2 we state the main assumptions and we recall the well-posedness result for (3.1.1) obtained in Chapter 2. In §3.3 we prove a generalized Itô formula for the square of the norm, as well as the strong pathwise continuity of solutions. In §3.4 we prove existence and uniqueness of strong variational solutions to (3.1.1) assuming first that B is locally Lipschitz-continuous with linear growth and that X_0 is square integrable, hence removing the latter assumption in a second step, allowing X_0 to be merely measurable. While in the former case solutions have finite second moment, in the latter case one needs to work with processes that are just measurable (in ω), so that uniqueness has to be proved in a much larger space. This is achieved by a suitable application of the Itô formula of §3.3 and stopping arguments. In §3.5 we show that X_0 having finite p-th moment implies that the solution belongs to a space of processes with finite p-moment as well, with explicit control of its norm. The Lipschitz continuity of the solution map is then established in a particular case. Further regularity of the solution and of invariant measures is obtained in the last section, under additional regularity assumptions on X_0 and B.

3.2 Assumptions and preliminaries

Let D be a bounded domain in $\mathbb{R}^d$ with smooth boundary, and V a real separable Hilbert space densely, continuously, and compactly embedded in $H := L^2(D)$. The scalar product and the norm of H will be denoted by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$, respectively. Identifying H with its dual H', the triple (V, H, V') is a so-called Gelfand triple: the duality form between V and V' extends the scalar product of H, i.e. $\langle v, w \rangle = {}_V\langle v, w \rangle_{V'}$ for any v, $w \in H$. For this reason, we shall simply denote the duality form of V and V' by the same symbol used for the scalar product in H.

The following assumptions on the linear operator $A \in \mathscr{L}(V, V')$ will be tacitly assumed to hold throughout the whole text:

(i) there exists $C > 0$ such that $\langle Av, v \rangle \geq C\|v\|_V^2$ for every $v \in V$;

(ii) the part of A in H can be 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;

(iv) there exists $m \in \mathbb{N}$ such that $(I + \delta A_1)^{-m} \in \mathscr{L}(L^1(D), L^\infty(D))$.

We shall occasionally refer to hypothesis (i) as coercivity of A, and to hypothesis (iv) as ultracontractivity of the resolvent of A_1.

Let us now state the assumptions on the nonlinear part of the drift: $\beta \subset \mathbb{R} \times \mathbb{R}$ is a maximal monotone graph such that $0 \in \beta(0)$ and $\mathsf{D}(\beta) = \mathbb{R}$. Let $j : \mathbb{R} \to [0, +\infty)$ be the unique convex lower-semicontinuous function such that $j(0) = 0$ and $\beta = \partial j$, where ∂ stands for the subdifferential in the sense of convex analysis. We assume that

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

Denoting the Moreau-Fenchel conjugate of j by j^*, the fact that $\mathsf{D}(\beta) = \mathbb{R}$ is equivalent to the superlinearity of j^* at infinity, i.e. to

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

For a comprehensive treatment of maximal monotone operators and their connection with convex analysis we refer to, e.g., [10]. Here we limit ourselves to recalling that, for any maximal monotone graph γ on a Hilbert space E, its resolvent and Yosida approximation of γ are defined as $(I + \lambda\gamma)^{-1}$ and

$$\gamma_\lambda := \frac{1}{\lambda}\big(I - (I + \lambda\gamma)^{-1}\big),$$

respectively, that both are continuous operators on E, and that the former is a contraction, while the latter is Lipschitz-continuous with Lipschitz constant bounded by $1/\lambda$.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, endowed with a right-continuous and completed filtration $(\mathscr{F}_t)_{t\in[0,T]}$, on which a cylindrical Wiener process W on a real separable Hilbert space U is defined. The diffusion coefficient

$$B : \Omega \times [0, T] \times H \to \mathscr{L}^2(U, H)$$

is assumed to be such that $B(\cdot, \cdot, x)$ is progressively measurable for every $x \in H$, and to grow at most linearly in its third argument, uniformly with respect to the others. That is, we assume that there exists a constant N such that

$$\big\|B(t, \omega, x)\big\|_{\mathscr{L}^2(U,H)} \le N\big(1 + \|x\|\big)$$

for all $(\omega, t, x) \in \Omega \times [0, T] \times H$. In addition to this, we shall consider either of two different assumptions, namely

(B1) B is Lipschitz continuous in its third argument, uniformly with respect to the others, i.e.

$$\big\|B(\omega, t, x) - B(\omega, t, y)\big\|_{\mathscr{L}^2(U,H)} \le N\|x - y\|$$

for all $(\omega, t) \in \Omega \times [0, T]$ and x, $y \in H$.

(B2) B is locally Lipschitz continuous in its third argument, uniformly with respect to the others, i.e. there exists a function $R \mapsto N_R : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\big\|B(\omega,t,x) - B(\omega,t,y)\big\|_{\mathscr{L}^2(U,H)} \leq N_R\|x-y\|$$

for all $(\omega,t) \in \Omega \times [0,T]$ and $x,\, y \in H$ with $\|x\|,\, \|y\| \leq R$.

Finally, X_0 is assumed to be an H-valued $\mathscr{F}_0$-measurable random variable.

Let us now define the concept of solution to equation (3.1.1).

Definition 3.2.1. *A strong solution to* (3.1.1) *is a pair* (X,ξ)*, where* X *is a* V*-valued adapted process and* ξ *is an* $L^1(D)$*-valued predictable process, such that,* $\mathbb{P}$*-almost surely,*

$$X \in L^\infty(0,T;H) \cap L^2(0,T;V), \qquad \xi \in L^1(0,T;L^1(D)),$$
$$\xi \in \beta(X) \quad \textit{a.e. in } (0,T) \times D,$$

and

$$X(t) + \int_0^t AX(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s,X(s))\,dW(s)$$

in $V' \cap L^1(D)$ *for all* $t \in [0,T]$.

It is convenient to introduce the family of sets $(\mathscr{J}_p)_{p\geq 0}$ as follows:

$$\mathscr{J}_p \subset \Big(L^p(\Omega;C([0,T];H)) \cap L^p(\Omega;L^2(0,T;V))\Big) \times L^{p/2}(\Omega;L^1((0,T)\times D)$$

formed by processes (ϕ,ψ) such that ϕ is adapted with values in V, ψ is predictable with values in $L^1(D)$, $\psi \in \beta(\phi)$ a.e. in $\Omega \times (0,T) \times D$, and $j(\phi) + j^*(\psi) \in L^{p/2}(\Omega;L^1((0,T)\times D)$.

The following well-posedness result has been proved in Chapter 2. Just for the purposes of this statement, we shall denote the space $\mathscr{J}_2$ with $L^\infty(0,T;H)$ in place of $C([0,T];H)$ by $\tilde{\mathscr{J}}_2$.

Theorem 3.2.2. *If* $X_0 \in L^2(\Omega,\mathscr{F}_0;H)$ *and* B *satisfies the global Lipschitz condition* (B1)*, then there exists a unique strong solution* (X,ξ) *to* (3.1.1) *belonging to* $\tilde{\mathscr{J}}_2$*. Furthermore, the trajectories of* X *are weakly continuous in* H *and the solution map*

$$\begin{aligned} L^2(\Omega;H) &\longrightarrow L^2(\Omega;L^\infty(0,T;H)) \cap L^2(\Omega;L^2(0,T;V)) \\ X_0 &\longmapsto X \end{aligned}$$

is Lipschitz-continuous.

Our main result in this chapter is the following far-reaching extension of Theorem 3.2.2: under the more general local Lipschitz continuity assumption (B2), for any $X_0 \in L^p(\Omega,\mathscr{F}_0,\mathbb{P};H)$, $p \in [0,\infty[$, there exists a strong solution (X,ξ) belonging to $\mathscr{J}_p$, which is unique in $\mathscr{J}_0$. In particular, the trajectories of X are strongly continuous in H. Precise statements and proofs are given in §3.4.

3.3 Pathwise continuity via a generalized Itô formula

In this section we prove that, under the assumptions of Theorem 3.2.2, the unique strong solution (X,ξ) in $\mathscr{J}_2$ to (3.1.1) is such that X admits a modification with strongly continuous trajectories in H, rather than just weakly continuous. To this purpose, we need a generalized Itô's formula for the square of the norm under minimal integrability assumptions, that will play a fundamental role throughout.

We first need some preparations. Let us recall that the part of A in H is the linear (unbounded) operator on H defined by $A_2 := A \cap (V \times H)$. In particular,

$$\mathsf{D}(A_2) = \{u \in V : Au \in H\} \qquad \text{and} \qquad A_2 u = Au \quad \forall u \in \mathsf{D}(A_2).$$

It is well known (see, e.g., [7]) that A_2 is closed and that $\mathsf{D}(A_2)$ is a Banach space with respect to the graph norm

$$\|u\|^2_{\mathsf{D}(A_2)} := \|u\|^2 + \|Au\|^2.$$

Moreover, $\mathsf{D}(A_2)$ is continuously and densely embedded in V.

Lemma 3.3.1. *Let $v \in V$ and $v_\lambda := (I + \lambda A_1)^{-1} v$. Then $v_\lambda \to v$ in V as $\lambda \to 0$.*

Proof. Let $v \in V$ and $\varepsilon > 0$: since $\mathsf{D}(A_2)$ is densely embedded in V, we can choose $u \in \mathsf{D}(A_2)$ such that $\|v - u\|_V < \varepsilon$. Setting $u_\lambda := (I + \lambda A_1)^{-1} u$, we have

$$\|v - v_\lambda\|_V \le \|v - u\|_V + \|u - u_\lambda\|_V + \|u_\lambda - v_\lambda\|_V.$$

Since $u, v \in V$, we have $u_\lambda - v_\lambda = (I + \lambda A_2)^{-1}(u - v)$, and recalling that A_2 is the part of A in H we have

$$(u_\lambda - v_\lambda) + \lambda A(u_\lambda - v_\lambda) = u - v,$$

where the identity holds in V as well. Taking the duality product with $A(u_\lambda - v_\lambda) \in V'$, by coercivity and boundedness of A it follows that

$$\begin{aligned} {}_{V'}\big\langle A(u_\lambda - v_\lambda), u_\lambda - v_\lambda\big\rangle_V + \lambda\, {}_{V'}\big\langle A(u_\lambda - v_\lambda), A(u_\lambda - v_\lambda)\big\rangle_V& \\ \ge C\big\|u_\lambda - v_\lambda\big\|_V^2 + \lambda\big\|A(u_\lambda - v_\lambda)\big\|^2& \end{aligned}$$

and

$${}_{V'}\langle A(u_\lambda - v_\lambda), u\rangle_V \le \|A\|_{\mathscr{L}(V,V')}\|u_\lambda - v_\lambda\|_V \|u\|_V,$$

hence

$$C\|u_\lambda - v_\lambda\|_V^2 + \lambda\|A(u_\lambda - v_\lambda)\|^2 \le \|A\|_{\mathscr{L}(V,V')}\|u_\lambda - v_\lambda\|_V \|u\|_V,$$

which implies that there exists a constant $N > 0$, independent of λ, such that

$$\|u_\lambda - v_\lambda\|_V \le N\|u - v\|_V,$$

or, equivalently, that $(I + \lambda A_1)^{-1}$ is uniformly bounded in V with respect to λ. This implies that

$$\|u_\lambda - v_\lambda\|_V \le N\|u - v\|_V \le N\varepsilon.$$

It remains to estimate the term $\|u - u_\lambda\|_V$. Since $u \in \mathsf{D}(A_2)$ and

$$u_\lambda := (I + \lambda A_1)^{-1} u = (I + \lambda A_2)^{-1} u,$$

one has $u_\lambda \in \mathsf{D}(A_2^2)$, hence, recalling that A_2 is the part of A in H,

$$Au_\lambda + \lambda A(Au_\lambda) = Au$$

in $H \hookrightarrow V'$. Taking the duality pairing with $Au_\lambda \in \mathsf{D}(A_2) \hookrightarrow V$, one has

$${}_{V'}\big\langle Au_\lambda, Au_\lambda\big\rangle_V + \lambda\, {}_{V'}\big\langle A(Au_\lambda), Au_\lambda\big\rangle_V = {}_{V'}\big\langle Au, Au_\lambda\big\rangle_V,$$

where

$$
{}_{V'}\langle Au_\lambda, Au_\lambda\rangle_V = \|Au_\lambda\|^2, \qquad {}_{V'}\langle A(Au_\lambda), Au_\lambda\rangle_V \geq C\|Au_\lambda\|_V^2,
$$
$$
{}_{V'}\langle Au, Au_\lambda\rangle_V = \langle Au, Au_\lambda\rangle \leq \frac{1}{2}\|Au\|^2 + \frac{1}{2}\|Au_\lambda\|^2,
$$

hence

$$
\|Au_\lambda\|^2 + \lambda C\|Au_\lambda\|_V^2 \leq \frac{1}{2}\|Au\|^2 + \frac{1}{2}\|Au_\lambda\|^2,
$$

which implies that $\sqrt{\lambda}\|Au_\lambda\|_V \leq N\|Au\|$, with a constant N independent of λ. Therefore, since $u \in \mathsf{D}(A_2)$,

$$
\|u_\lambda - u\|_V = \lambda\|Au_\lambda\|_V \leq N\sqrt{\lambda}\|Au\|.
$$

Choosing λ such that $N\sqrt{\lambda}\|Au\| < \varepsilon$, one has then

$$
\|v_\lambda - v\|_V < (2+N)\varepsilon,
$$

from which the conclusion follows by arbitrariness of ϵ. □

We recall that (see, e.g., [45]) if two Banach spaces F and G are continuously embedded in a separated topological vector space E, their sum $F+G$ is defined as the subspace of E

$$
F + G := \{u \in E : \exists f \in F,\, g \in G :\ u = f + g\}.
$$

Endowed with the norm

$$
\|u\|_{F+G} := \inf_{u=f+g}\big(\|f\|_F + \|g\|_G\big),
$$

$F+G$ is a Banach space. Similarly, the intersection $F \cap G$ is also a Banach space if endowed with the norm

$$
\|u\|_{F\cap G} := \|u\|_F + \|u\|_G.
$$

Moreover, if $F \cap G$ is dense in both F and G, then F' and G' are continuously embedded in $(F\cap G)'$, and $(F+G)' = F' \cap G'$. In the following we shall deal with $F := L^1(0,T;H)$ and $G := L^2(0,T;V')$, so that as ambient space E one can simply take $L^1(0,T;V')$. In this case $F \cap G$ is dense in both F and G, hence, by reflexivity of V,

$$
\big(L^1(0,T;H) + L^2(0,T;V')\big)' = L^\infty(0,T;H) \cap L^2(0,T;V).
$$

Theorem 3.3.2. *Let Y, v and g be adapted processes such that*

$$
\begin{aligned}
Y &\in L^0(\Omega; L^\infty(0,T;H) \cap L^2(0,T;V)),\\
v &\in L^0(\Omega; L^1(0,T;H) + L^2(0,T;V')),\\
g &\in L^0(\Omega; L^1(0,T;L^1(D))),\\
\exists\, \alpha > 0 :\quad & j(\alpha Y) + j^*(\alpha g) \in L^0(\Omega; L^1((0,T)\times D)).
\end{aligned}
$$

Moreover, let $Y_0 \in L^0(\Omega, \mathscr{F}_0; H)$ and G be a progressive $\mathscr{L}^2(U,H)$-valued process such that

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

If

$$
Y(t) + \int_0^t v(s)\,ds + \int_0^t g(s)\,ds = Y_0 + \int_0^t G(s)\,dW(s) \qquad \forall t \in [0,T] \quad \mathbb{P}\text{-a.s.}
$$

in $V' \cap L^1(D)$, then

$$\frac{1}{2}\|Y(t)\|^2 + \int_0^t \langle v(s), Y(s)\rangle\,ds + \int_0^t\int_D g(s,x)Y(s,x)\,dx\,ds = \frac{1}{2}\|Y_0\|^2 + \frac{1}{2}\int_0^t \|G(s)\|^2_{\mathscr{L}^2(U,H)}\,ds + \int_0^t Y(s)G(s)\,dW(s)$$

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

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

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

where $g^\delta \in L^1(0,T;H)$, hence the classical Itô's formula yields, for every $\delta > 0$,

$$\frac{1}{2}\|Y^\delta(t)\|^2 + \int_0^t \langle v^\delta(s), Y^\delta(s)\rangle\,ds + \int_0^t\int_D g^\delta(s,x)Y^\delta(s,x)\,dx\,ds = \frac{1}{2}\|Y_0^\delta\|^2 + \frac{1}{2}\int_0^t \|G^\delta(s)\|^2_{\mathscr{L}^2(U,H)}\,ds + \int_0^t Y^\delta(s)G^\delta(s)\,dW(s).$$

Let us pass to the limit as $\delta \to 0$. Since the resolvent of A_1 coincides on H with the resolvent of A_2, which converges to the identity in $\mathscr{L}(H)$ in the strong operator topology, we immediately infer that

$$\begin{aligned} Y^\delta(t) &\longrightarrow Y(t) && \text{in } H \quad \forall t \in [0,T],\\ g^\delta &\longrightarrow g && \text{in } L^1(0,T;L^1(D)),\\ Y_0^\delta &\longrightarrow Y_0 && \text{in } H,\\ G^\delta &\longrightarrow G && \text{in } L^2(0,T;\mathscr{L}^2(U,H)) \end{aligned}$$

where the last statement, which follows by well-known continuity properties of Hilbert-Schmidt operators, also implies

$$\int_0^t \|G^\delta(s)\|^2_{\mathscr{L}^2(U,H)}\,ds \longrightarrow \int_0^t \|G(s)\|^2_{\mathscr{L}^2(U,H)}\,ds.$$

Moreover, by the previous lemma we have

$$Y^\delta \longrightarrow Y \quad \text{in } L^2(0,T;V),$$

and $Y \in L^\infty(0,T;H)$ and the contractivity in H of the resolvent of A_1 immediately imply, by the dominated convergence theorem, that $Y^\delta \to Y$ weakly* in $L^\infty(0,T;H)$. Therefore, by reflexivity of V,

$$Y^\delta \longrightarrow Y \quad \text{weakly* in } L^\infty(0,T;H) \cap L^2(0,T;V).$$

Since $v \in L^1(0,T;H) + L^2(0,T;V')$, we have that $v = v_1 + v_2$, with $v_1 \in L^1(0,T;H)$ and $v_2 \in L^2(0,T;V')$. In this case v^δ has to be interpreted as

$$v^\delta := (I+\delta A_1)^{-m}v_1 + (I+\delta A)^{-m}v_2.$$

Note that this is very natural since A_1 and A coincide on $\mathsf{D}(A_1)\cap V$. By the properties of the resolvent it easily follows that

$$v_1^\delta \longrightarrow v_1 \quad \text{in } L^1(0,T;H).$$

Moreover, since $A^{-1}v_2 \in L^2(0,T;V)$ and $A^{-1}v_2^\delta = (I+\delta A_2)^{-m}A^{-1}v_2$, thanks to Lemma 3.3.1 we have that $A^{-1}v_2^\delta \to A^{-1}v_2$ in $L^2(0,T;V)$, hence also, by continuity of A,

$$v_2^\delta \longrightarrow v_2 \quad \text{in } L^2(0,T;V').$$

The convergences of v^δ and Y^δ just proved thus imply

$$\int_0^t \big\langle v^\delta(s), Y^\delta(s)\big\rangle\, ds \longrightarrow \int_0^t \big\langle v(s), Y(s)\big\rangle\, ds$$

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

We are now going to prove that $\big((Y^\delta G^\delta)\cdot W - (YG)\cdot W\big)^*_T \to 0$ in probability. Setting $M_\delta := (Y^\delta G^\delta)\cdot W$ and $M := (YG)\cdot W$, it is well known that it suffices to show that the quadratic variation of $M_\delta - M$ converges to 0 in probability. One has

$$\begin{aligned}
\big[M_\delta - M, M_\delta - M\big] &= \big\|Y^\delta G^\delta - YG\big\|^2_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))}\\
&\le \big\|Y^\delta G^\delta - Y^\delta G\big\|^2_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))} + \big\|Y^\delta G - YG\big\|^2_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))}\\
&\le \|Y\|^2_{L^\infty(0,T;H)}\|G^\delta - G\|^2_{L^2(0,T;\mathscr{L}^2(U,H))} + \|Y^\delta G - YG\|^2_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))},
\end{aligned}$$

where the convergence to zero of the first term in the last expression has already been proved, and

$$\|Y^\delta G - YG\|^2_{L^2(0,T;\mathscr{L}^2(U,\mathbb{R}))} \le \int_0^T \big\|Y^\delta(s) - Y(s)\big\|^2 \big\|G(s)\big\|^2_{\mathscr{L}^2(U,H)}\, ds \longrightarrow 0,$$

by the dominated convergence theorem, because $Y^\delta \to Y$ pointwise in H and $\|Y^\delta - Y\| \le 2\|Y\| \in L^\infty(0,T)$. We have thus shown that

$$\int_0^\cdot Y^\delta(s)G^\delta(s)\, dW(s) \longrightarrow \int_0^\cdot Y(s)G(s)\, dW(s)$$

in probability, hence $\mathbb{P}$-a.s. along a subsequence of δ.

Finally, it is clear that $Y^\delta g^\delta \to Yg$ in measure in $(0,T)\times D$, and that, thanks to the assumptions on j,

$$\pm\alpha^2 Y^\delta g^\delta \le j(\pm\alpha Y^\delta) + j^*(\alpha g^\delta) \lesssim 1 + j(\alpha Y^\delta) + j^*(\alpha g^\delta),$$

where the second inequality follows from the fact that, thanks to the assumption on the growth of j at ∞, there exists a constant $M > 0$ such that

$$j(r) \le M\big(1 + j(-r)\big) \qquad \forall r \in \mathbb{R}.$$

Jensen's inequality for sub-Markovian operators (see, e.g., [41]) thus yields

$$j(\alpha Y^\delta) + j^*(\alpha g^\delta) \le (I+\delta A_1)^{-m}\big(j(\alpha Y) + j^*(\alpha g)\big),$$

so that

$$\alpha^2 \big|Y^\delta g^\delta\big| \lesssim 1 + (I + \delta A_1)^{-m}\big(j(\alpha Y) + j^*(\alpha g)\big).$$

Since $j(\alpha Y)+j^*(\alpha g) \in L^1((0,T)\times D)$ by assumption, the contractivity of the resolvent in $L^1(D)$ and the dominated convergence theorem imply that the right-hand side in the last inequality is convergent in $L^1((0,T) \times D)$. Hence $(Y^\delta g^\delta)_\delta$ is uniformly integrable and, by Vitali's theorem,

$$\int_0^t \int_D g^\delta(s,x) Y^\delta(s,x)\,dx\,ds \longrightarrow \int_0^t \int_D g(s,x)Y(s,x)\,dx\,ds$$

for all $t \in [0,T]$. The proof is thus completed. □

As a first important consequence of the generalized Itô formula we show that (the first component of) strong solutions are pathwise strongly continuous in H.

Theorem 3.3.3. *Let (X,ξ) be the unique strong solution to* (3.1.1) *belonging to $\mathscr{J}_2$. Then X has strongly continuous paths in H, i.e. there exists $\Omega' \in \mathscr{F}$ with $\mathbb{P}(\Omega') = 1$ such that*

$$X(\omega) \in C([0,T];H) \qquad \forall \omega \in \Omega'.$$

Proof. Let $r \in [0,T]$. We have to prove that $X(t) \to X(r)$ in H as $t \to r$, $t \in [0,T]$. It follows from Theorem 3.3.2 that for every $t \in [0,T]$ there exists $\Omega' \in \mathscr{F}_0$ with $\mathbb{P}(\Omega') = 1$ such that

$$\begin{aligned}\frac{1}{2}\|X(t)\|^2 - \frac{1}{2}\|X(r)\|^2 = &- \int_r^t \langle AX(s), X(s)\rangle\,ds - \int_r^t\int_D \xi(s)X(s)\,ds \\ &+ \frac{1}{2}\int_r^t \big\|B(s)\big\|^2_{\mathscr{L}^2(U,H)} + \int_r^t X(s)B(s,X(s))\,dW(s)\end{aligned}$$

everywhere on Ω'. By the definition of strong solution, we can assume that $X \in L^\infty(0,T;H)$, $AX \in L^2(0,T;V')$, $j(X) + j^*(\xi) \in L^1((0,T) \times D)$, and that $B(\cdot,X) \in L^2(0,T;\mathscr{L}^2(U,H))$, everywhere on Ω'. Since $X\xi = j(X) + j^*(\xi)$, it follows that the process

$$[0,T] \ni s \longmapsto \psi(s) := -\langle AX(s), X(s)\rangle - \int_D \xi(s)X(s) + \frac{1}{2}\big\|B(s,X(s))\big\|^2_{\mathscr{L}^2(U,H)}$$

belongs to $L^1(0,T)$ everywhere on Ω'. Therefore, writing

$$\frac{1}{2}\|X(t)\|^2 - \frac{1}{2}\|X(r)\|^2 = \int_r^t \phi(s)\,ds + \int_r^t X(s)B(s,X(s))\,dW(s),$$

since $\psi \in L^1(0,T)$ and the stochastic integral has continuous trajectories, we have, as $t \to r$,

$$\|X(t)\|^2 - \|X(r)\|^2 \to 0,$$

so that $\|X(t)\| \to \|X(r)\|$. Furthermore, $X(t) \to X(r)$ weakly in H as $t \to r$ by Theorem 3.2.2, hence, since H is uniformly convex, we conclude that $X(t) \to X(r)$ in H (cf., e.g., [23, Proposition 3.32]). □

3.4 Existence and uniqueness

We begin with a simple estimate that will be used several times.

Lemma 3.4.1. *Let F and G be progressively measurable processes with values in the spaces H and $\mathscr{L}^2(U,H)$, respectively, such that FG is integrable with respect to W. For any numbers p,*

$\varepsilon > 0$ *and any stopping time* S *one has*

$$\big\|\big((FG)\cdot W\big)^*_S\big\|_{L^p(\Omega)} \lesssim \varepsilon\big\|F^*_S\big\|^2_{L^{2p}(\Omega)} + \frac{1}{\varepsilon}\big\|G1\!\!1_{[\![0,S]\!]}\big\|^2_{L^{2p}(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))}$$

Proof. The Burkholder-Davis-Gundy inequality asserts that

$$\big\|\big((FG)\cdot W\big)^*_S\big\|_{L^p(\Omega)} \eqsim \big\|[(FG)\cdot W,(FG)\cdot W]^{1/2}_S\big\|_{L^p(\Omega)},$$

where, by the ideal property of Hilbert-Schmidt operators and Young's inequality,

$$\begin{aligned}
[(FG)\cdot W,(FG)\cdot W]^{1/2}_S &= \Big(\int_0^S \big\|F(t)G(t)\big\|^2_{\mathscr{L}^2(U,\mathbb{R})}\,dt\Big)^{1/2}\\
&\le \Big(\int_0^S \big\|F(t)\big\|^2\big\|G(t)\big\|^2_{\mathscr{L}^2(U,H)}\,dt\Big)^{1/2}\\
&\le F^*_S\Big(\int_0^S \big\|G(t)\big\|^2_{\mathscr{L}^2(U,H)}\,dt\Big)^{1/2}\\
&\le \varepsilon F^{*2}_S + \frac{1}{\varepsilon}\int_0^S \big\|G(t)\big\|^2_{\mathscr{L}^2(U,H)}\,dt.
\end{aligned}$$

Therefore, taking the $L^p(\Omega)$-(quasi)norm on both sides,

$$\big\|\big((FG)\cdot W\big)^*_S\big\|_{L^p(\Omega)} \lesssim \varepsilon\big\|F^*_S\big\|^2_{L^{2p}(\Omega)} + \frac{1}{\varepsilon}\big\|G1\!\!1_{[\![0,S]\!]}\big\|^2_{L^{2p}(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))} \qquad \square$$

Let (X,ξ) and $(Y,\eta) \in \mathscr{J}_0$ be strong solutions, in the sense of Definition 3.2.1, to the equation

$$dX + AX\,dt + \beta(X)\,dt \ni B(\cdot,X)\,dW$$

with initial conditions X_0 and Y_0, both elements of $L^0(\Omega,\mathscr{F}_0,\mathbb{P};H)$, respectively. Here and throughout this section we assume that B is locally Lipschitz-continuous in the sense of assumption (B2).

Let us also introduce the sequence of stopping times $(T_n)_{n\in\mathbb{N}}$ defined as

$$T_n := \inf\big\{t \ge 0 : \|X_\Gamma(t)\| \ge n \ \text{ or } \ \|Y_\Gamma(t)\| \ge n\big\} \wedge T.$$

Here and in the following, for any $\Gamma \in \mathscr{F}_0$, we shall denote multiplication by $1\!\!1_\Gamma$ by a subscript Γ. Even though the stopping times T_n depend on Γ, we shall not indicate this explicitly to avoid making the notation too cumbersome.

The stopping times T_n are well defined because, by definition of $\mathscr{J}_0$, X and Y have continuous paths with values in H. Moreover, $T_n \neq 0$ for sufficiently large n.

The estimate in the following lemma is an essential tool, from which, for instance, uniqueness and a local property of solutions will follow as easy corollaries.

Lemma 3.4.2. *Let* $\Gamma \in \mathscr{F}_0$ *be such that* $X_{0\Gamma}, Y_{0\Gamma} \in L^2(\Omega,\mathscr{F}_0,\mathbb{P};H)$. *One has, for every* $n \in \mathbb{N}$,

$$\mathbb{E}\big(X_\Gamma - Y_\Gamma\big)^{*2}_{T_n} \lesssim \mathbb{E}\big\|X_{0\Gamma} - Y_{0\Gamma}\big\|^2,$$

with implicit constant depending on T *and on the Lipschitz constant of* B *in the ball in* H *of radius* n.

Proof. One has

$$(X-Y)+\int_0^t A(X-Y)\,ds+\int_0^t(\xi-\eta)\,ds=X_0-Y_0+\int_0^t(B(X)-B(Y))\,dW.$$

We recall that, for any $\mathscr{F}_0$-measurable random variable ζ and any stochastically integrable process K, one has $\zeta(K\cdot W)=(\zeta K)\cdot W$. Therefore

$$\begin{aligned}(X-Y)_\Gamma+\int_0^t A(X-Y)_\Gamma\,ds+\int_0^t(\xi-\eta)_\Gamma\,ds\\ =(X_0-Y_0)_\Gamma+\int_0^t(B(X)-B(Y))_\Gamma\,dW.\end{aligned}$$

The Itô formula of Theorem 3.3.2 yields

$$\begin{aligned}&\big\|X_\Gamma-Y_\Gamma\big\|^2(t\wedge T_n)+2\int_0^{t\wedge T_n}\big\langle A(X_\Gamma-Y_\Gamma),X_\Gamma-Y_\Gamma)\big\rangle\,ds\\&\quad+2\int_0^{t\wedge T_n}\int_D((X-Y)(\xi-\eta))_\Gamma\,ds\\&=\big\|X_{0\Gamma}-Y_{0\Gamma}\big\|^2+\int_0^{t\wedge T_n}\big\|(B(X)-B(Y))_\Gamma\big\|^2_{\mathscr{L}^2(U,H)}\,ds\\&\quad+2\int_0^{t\wedge T_n}(X-Y)_\Gamma(B(X)-B(Y))_\Gamma\,dW,\end{aligned}$$

where (a) the second and term terms on the left-hand side are positive by monotonicity of A and β, and by the assumption that $\xi\in\beta(X)$, $\eta\in\beta(Y)$ a.e. in $\Omega\times(0,T)\times D$; (b) one has

$$(B(X)-B(Y))_\Gamma=1\!\!1_\Gamma\big(B(X_\Gamma)-B(Y_\Gamma)\big),$$

hence

$$1\!\!1_{[\![0,T_n]\!]}\big\|(B(X)-B(Y))_\Gamma\big\|^2_{\mathscr{L}^2(U,H)}\lesssim_n 1\!\!1_{[\![0,T_n]\!]}\,1\!\!1_\Gamma\big\|X_\Gamma-Y_\Gamma\big\|.$$

Taking supremum in time and expectation,

$$\begin{aligned}\mathbb{E}\big(X_\Gamma^{T_n}-Y_\Gamma^{T_n}\big)^{*2}_t\lesssim\mathbb{E}\big\|X_{0\Gamma}-Y_{0\Gamma}\big\|^2+\int_0^t\mathbb{E}\big(X_\Gamma^{T_n}-Y_\Gamma^{T_n}\big)^{*2}_s\,ds\\+\mathbb{E}\sup_{s\le t}\int_0^{s\wedge T_n}(X-Y)_\Gamma(B(X)-B(Y))_\Gamma\,dW,\end{aligned}$$

where, by Lemma 3.4.1, the last term on the right-hand side is bounded by

$$\begin{aligned}&\varepsilon\,\mathbb{E}\big(X_\Gamma^{T_n}-Y_\Gamma^{T_n}\big)^{*2}_t+N(\varepsilon)\,\mathbb{E}\int_0^{t\wedge T_n}\big\|(B(X)-B(Y))_\Gamma\big\|^2_{\mathscr{L}^2(U,H)}\,ds\\&\quad\le\varepsilon\,\mathbb{E}\big(X_\Gamma^{T_n}-Y_\Gamma^{T_n}\big)^{*2}_t+N(\varepsilon,n)\int_0^t\mathbb{E}\big(X_\Gamma^{T_n}-Y_\Gamma^{T_n}\big)^{*2}_s\,ds.\end{aligned}$$

Choosing ε small enough, it follows by Gronwall's inequality that

$$\mathbb{E}\big(X_\Gamma-Y_\Gamma\big)^{*2}_{T_n}=\mathbb{E}\big(X_\Gamma^{T_n}-Y_\Gamma^{T_n}\big)^{*2}_T\lesssim\mathbb{E}\big\|X_{0\Gamma}-Y_{0\Gamma}\big\|^2,$$

with an implicit constant that depends on T and on the Lipschitz constant of B on the ball in H of radius n. □

Corollary 3.4.3. *Uniqueness of strong solutions in* $\mathscr{J}_0$ *holds for* (3.1.1).

Proof. Let (X,ξ), $(Y,\eta) \in \mathscr{J}_0$ be strong solutions to (3.1.1). For any $\Gamma \in \mathscr{F}_0$ such that $X_{0\Gamma} \in L^2(\Omega; H)$ the previous lemma yields $X_\Gamma^{T_n} = Y_\Gamma^{T_n}$ for all $n \in \mathbb{N}$, hence $X_\Gamma = Y_\Gamma$. Writing

$$\Omega = \bigcup_{k\in\mathbb{N}} \Omega_k, \qquad \Omega_k := \big\{\omega \in \Omega : \|X_0(\omega)\| \le k\big\},$$

and choosing Γ as Ω_k, it follows that $X\mathbb{1}_{\Omega_k} = Y\mathbb{1}_{\Omega_k}$ for all k, hence $X = Y$. By comparison, $\xi = \eta$ a.e. in $\Omega \times (0,T) \times D$. □

Remark 3.4.4. To prove the corollary, by inspection of the proof of Lemma 3.4.2 it is evident that one may directly take $\Gamma = \Omega$, as in this case $X_0 - Y_0 = 0$, whose second moment is obviously finite. This immediately implies $X^{T_n} = Y^{T_n}$ for all $n \in \mathbb{N}$, hence $X = Y$.

Corollary 3.4.5. *Let $\Gamma \in \mathscr{F}_0$. If $X_{0\Gamma} = Y_{0\Gamma}$, then $X_\Gamma = Y_\Gamma$, and $\xi_\Gamma = \eta_\Gamma$ a.e. in $\Omega\times(0,T)\times D$.*

Proof. Write $\Omega = \bigcup_{k\in\mathbb{N}} \Omega_k$, where

$$\Omega_k := \big\{\omega \in \Omega : \|X_0(\omega)\| \le k\big\} \cap \big\{\omega \in \Omega : \|Y_0(\omega)\| \le k\big\}.$$

Then $X_0\mathbb{1}_{\Gamma\cap\Omega_k}$, $Y_0\mathbb{1}_{\Gamma\cap\Omega_k} \in L^2(\Omega; H)$, and Lemma 3.4.2 implies that $X_{\Gamma\cap\Omega_k} = Y_{\Gamma\cap\Omega_k}$ for all $k \in \mathbb{N}$, hence $X_\Gamma = Y_\Gamma$, as well as, again by comparison, $\xi_\Gamma = \eta_\Gamma$ a.e. in $\Omega \times (0,T) \times D$. □

Now that uniqueness is cleared, we turn to the question of existence of strong solutions. For this we need some preparations. For $R > 0$, let us consider the truncation operator $\sigma_R : H \to H$ defined as

$$\sigma_R : x \longmapsto \begin{cases} x, & \|x\| \le R, \\ Rx/\|x\|, & \|x\| > R. \end{cases}$$

We shall then define

$$\begin{aligned} B_R : \Omega \times [0,T] \times H &\longrightarrow \mathscr{L}^2(U,H) \\ (\omega,t,x) &\longmapsto B(\omega,t,\sigma_R(x)). \end{aligned}$$

Let us check that B_R is Lipschitz-continuous for every $R > 0$. The progressive measurability of B_R follows from the one of B and the fact that $\sigma_R : H \to H$ is (Lipschitz) continuous. Moreover, since σ_R is 1-Lipschitz continuous, thanks to the local Lipschitz continuity and the linear growth of B, for every $\omega \in \Omega$, $t \in [0,T]$ and $x, y \in H$ one has

$$\|B_R(\omega,t,x) - B_R(\omega,t,y)\|_{\mathscr{L}^2(U,H)} \le N_R\|\sigma_R(x) - \sigma_R(y)\| \le N_R\|x - y\|$$

as well as

$$\|B_R(\omega,t,x)\|_{\mathscr{L}^2(U,H)} \le N(1 + \|\sigma_R(x)\|) \le N(1 + \|x\|).$$

Thanks to Theorems 3.2.2 and 3.3.3, as well as Lemma 3.4.2, the equation

$$dX_n + AX_n\,dt + \beta(X_n)\,dt = B_n(X_n)\,dW, \qquad X_n(0) = X_0, \tag{3.4.2}$$

admits a strong solution (X_n, ξ_n), which belongs to $\mathscr{J}_2$ and is unique in $\mathscr{J}_0$, for every $n \in \mathbb{N}$.* Moreover, by the strong continuity of the paths of X_n, one can define the increasing sequence

*Note that Theorem 3.2.2 only shows that (X_n, ξ_n) is unique in $\mathscr{J}_2$, while Lemma 3.4.2 yields uniqueness in the larger space $\mathscr{J}_0$.

of stopping times $(\tau_n)_{n\in\mathbb{N}}$ by

$$\tau_n := \inf\bigl\{t \in [0,T] : \|X_n(t)\| \geq n\bigr\},$$

as well as the stopping time

$$\tau := \lim_{n\to\infty} \tau_n = \sup_{n\in\mathbb{N}} \tau_n.$$

As first step we show that the sequence of processes (X_n, ξ_n) satisfies a sort of consistency condition.

Lemma 3.4.6. *One has $X_{n+1}^{\tau_n} = X_n^{\tau_n}$ for all $n \in \mathbb{N}$, as well as $\xi_n 1\!\!1_{[\![0,\tau_n]\!]} = \xi_{n+1} 1\!\!1_{[\![0,\tau_n]\!]}$ in $L^0(\Omega \times (0,T) \times D)$.*

Proof. Itô's formula yields, in view of the monotonicity of A and β,

$$\begin{aligned}\bigl\|X_{n+1} - X_n\bigr\|^2(t\wedge\tau_n) \lesssim \int_0^{t\wedge\tau_n} (X_{n+1} - X_n)\bigl(B_{n+1}(X_{n+1}) - B_n(X_n)\bigr)(s)\,dW(s)\\ + \int_0^{t\wedge\tau_n} \bigl\|B_{n+1}(X_{n+1}(s)) - B_n(X_n(s))\bigr\|^2_{\mathscr{L}^2(U,H)}\,ds.\end{aligned}$$

Note that $B_{n+1} = B_n$ on the ball of radius n in H, hence $B_n(X_n) = B_{n+1}(X_n)$ on $[\![0,\tau_n]\!]$. Therefore, since B_{n+1} is Lipschitz continuous,

$$\begin{aligned}\mathbb{E}\bigl(X_{n+1}^{\tau_n} - X_n^{\tau_n}\bigr)^{*2}_t \lesssim \mathbb{E}\bigl(((X_{n+1} - X_n)(B_{n+1}(X_{n+1}) - B_{n+1}(X_n)))\cdot W\bigr)^*_{t\wedge\tau_n}\\ + \int_0^t \mathbb{E}\bigl(X_{n+1}^{\tau_n} - X_n^{\tau_n}\bigr)^{*2}_s\,ds,\end{aligned}$$

where the first term on the right-hand side can be estimated, thanks to the BDG inequality and the ideal property of Hilbert-Schmidt operators, by

$$\begin{aligned}&\mathbb{E}\biggl(\int_0^{t\wedge\tau_n} \bigl\|X_{n+1} - X_n\bigr\|^2(s)\bigl\|B_{n+1}(X_{n+1}(s)) - B_{n+1}(X_n(s))\bigr\|^2_{\mathscr{L}^2(U,H)}\,ds\biggr)^{1/2}\\ &\qquad\lesssim_n \mathbb{E}\bigl(X_{n+1}^{\tau_n} - X_n^{\tau_n}\bigr)^*_t\biggl(\int_0^t \bigl\|X_{n+1}^{\tau_n} - X_n^{\tau_n}\bigr\|^2(s)\biggr)^{1/2}\\ &\qquad\lesssim_n \varepsilon\,\mathbb{E}\bigl(X_{n+1}^{\tau_n} - X_n^{\tau_n}\bigr)^{*2}_t + \frac{1}{\varepsilon}\int_0^t \mathbb{E}\bigl(X_{n+1}^{\tau_n} - X_n^{\tau_n}\bigr)^{*2}_s\,ds.\end{aligned}$$

Choosing ε small enough, Gronwall's inequality implies

$$\mathbb{E}\bigl(X_{n+1}^{\tau_n} - X_n^{\tau_n}\bigr)^{*2}_t = 0$$

for all $t \leq T$, hence $X_{n+1}^{\tau_n} = X_n^{\tau_n}$. The first claim is thus proved.

In order to prove the second claim, note that it holds

$$\begin{aligned}X_{n+1}^{\tau_n}(t) + \int_0^{t\wedge\tau_n} AX_{n+1}\,ds + \int_0^{t\wedge\tau_n} \xi_{n+1}\,ds = X_0 + \int_0^{t\wedge\tau_n} B_{n+1}(X_{n+1})\,dW,\\ X_n^{\tau_n}(t) + \int_0^{t\wedge\tau_n} AX_n\,ds + \int_0^{t\wedge\tau_n} \xi_n\,ds = X_0 + \int_0^{t\wedge\tau_n} B_n(X_n)\,dW,\end{aligned}$$

where $B_n(X_n)$ on the right-hand side of the second identity can be replaced by $B_{n+1}(X_{n+1})$ because the paths of $X_{n+1}^{\tau_n}$ remain within a ball of radius n in H and $X_{n+1}^{\tau_n} = X_n^{\tau_n}$. This identity also yields, by comparison,

$$\int_0^t \xi_{n+1} 1_{[\![0,\tau_n]\!]}\,ds = \int_0^{t\wedge\tau_n} \xi_{n+1}\,ds = \int_0^{t\wedge\tau_n} \xi_n\,ds = \int_0^t \xi_n 1_{[\![0,\tau_n]\!]}\,ds,$$

which implies the second claim.[†] □

The lemma implies that one can define processes X and ξ on $[\![0,\tau]\!]$ by the prescriptions $X := X_n$ and $\xi := \xi_n$ on $[\![0,\tau_n]\!]$ for all $n \in \mathbb{N}$, or equivalently (but perhaps less tellingly), as $X = \lim_{n\to\infty} X_n$ and $\xi = \lim_{n\to\infty} \xi_n$.

We are now going to show that the linear growth assumption on B implies that $\tau = T$. We shall first establish a priori estimates for the solution to equation (3.4.2).

Lemma 3.4.7. *There exists a constant $N > 0$, independent of n, such that*

$$\mathbb{E}\big\|X_n\big\|^2_{C([0,T];H)} + \mathbb{E}\big\|X_n\big\|^2_{L^2(0,T;V)} + \mathbb{E}\big\|\xi_n X_n\big\|_{L^1((0,T)\times D)} < N\big(1 + \mathbb{E}\|X_0\|^2\big).$$

Proof. The Itô formula of Theorem 3.3.2 yields

$$\begin{aligned}\|X_n(t)\|^2 &+ 2\int_0^t \langle AX_n(s), X_n(s)\rangle\,ds + 2\int_0^t\int_D \xi_n(s)X_n(s)\,dx\,ds\\ &= \|X_0\|^2 + \int_0^t \big\|B_n(s,X_n(s))\big\|^2_{\mathscr{L}^2(U,H)}\,ds + 2\int_0^t X_n(s)B_n(s,X_n(s))\,dW(s),\end{aligned}$$

where, recalling that $B_n = B(\cdot,\cdot,\sigma_n(\cdot))$ and σ_n is a contraction in H, and that B grows at most linearly,

$$\int_0^t \big\|B_n(s,X_n(s))\big\|^2_{\mathscr{L}^2(U,H)} \lesssim T + \int_0^T \|X_n(s)\|^2\,ds.$$

Denoting the stochastic integral on the right-hand side by M_n, taking supremum in time and expectation we get, by the coercivity of A,

$$\begin{aligned}\mathbb{E}\big\|X_n\big\|^2_{C([0,T];H)} &+ \mathbb{E}\big\|X_n\big\|^2_{L^2(0,T;V)} + \mathbb{E}\big\|\xi_n X_n\big\|_{L^1((0,T)\times D)}\\ &\lesssim 1 + \mathbb{E}\|X_0\|^2 + \mathbb{E}\int_0^T \|X_n(s)\|^2\,ds + \mathbb{E}\,M_T^{*2},\end{aligned}$$

where the implicit constant depends on T. By Lemma 3.4.1 we have, for any $\varepsilon > 0$,

$$\mathbb{E}\,M_T^{*2} \lesssim \varepsilon\,\mathbb{E}\big\|X_n\big\|^2_{C([0,T];H)} + N(\varepsilon)\,\mathbb{E}\int_0^T \|X_n(s)\|^2\,ds,$$

therefore, choosing ε sufficiently small,

$$\begin{aligned}\mathbb{E}\big\|X_n\big\|^2_{C([0,T];H)} &+ \mathbb{E}\big\|X_n\big\|^2_{L^2(0,T;V)} + \mathbb{E}\big\|\xi_n X_n\big\|_{L^1((0,T)\times D)}\\ &\lesssim 1 + \mathbb{E}\|X_0\|^2 + \mathbb{E}\int_0^T \|X_n(s)\|^2\,ds.\end{aligned}$$

Since this inequality holds also with T replaced by any $t \in]0,T]$, we also have

$$\mathbb{E}\big\|X_n\big\|^2_{C([0,t];H)} \lesssim 1 + \mathbb{E}\|X_0\|^2 + \int_0^T \mathbb{E}\big\|X_n\big\|^2_{C([0,s];H)}\,ds,$$

[†] The argument in fact proves the following slightly stronger statement: setting $\Xi_n := \int_0^\cdot \xi_n\,ds$, the processes $\Xi_{n+1}^{\tau_n}$ and $\Xi_n^{\tau_n}$ are indistinguishable for all n.

hence, by Gronwall's inequality,

$$\mathbb{E}\big\|X_n\big\|^2_{C([0,T];H)} \lesssim 1 + \mathbb{E}\|X_0\|^2,$$

with implicit constant depending on T. Since $C([0,T];H) \hookrightarrow L^2(0,T;H)$, one easily deduces

$$\mathbb{E}\big\|X_n\big\|^2_{C([0,T];H)} + \mathbb{E}\big\|X_n\big\|^2_{L^2(0,T;V)} + \mathbb{E}\big\|\xi_n X_n\big\|_{L^1((0,T)\times D)} \lesssim 1 + \mathbb{E}\|X_0\|^2. \qquad \square$$

Lemma 3.4.8. *One has*

$$\mathbb{P}\Big(\limsup_{n\to\infty}\{\tau_n \le T\}\Big) = 0.$$

In particular, $\tau = T$.

Proof. By Markov's inequality and the previous lemma,

$$\mathbb{P}\big(\|X_n\|_{C([0,T];H)} \ge n\big) \le \frac{1}{n^2}\,\mathbb{E}\|X_n\|^2_{C([0,T];H)} \lesssim \frac{1}{n^2}\big(1 + \mathbb{E}\|X_0\|^2\big).$$

Since the event $\{\|X_n\|_{C([0,T];H)} \ge n\}$ coincides with $\{\tau_n \le T\}$, one has

$$\sum_{n=1}^{\infty} \mathbb{P}\big(\tau_n \le T\big) < \infty,$$

thus also, by the Borel-Cantelli lemma,

$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}}\bigcup_{k\ge n}\{\tau_k \le T\}\right) = 0.$$

In other words, the sequence (τ_n) is ultimately constant: for each ω in a subset of Ω of $\mathbb{P}$-measure one, there exists $m = m(\omega)$ such that $\tau_n(\omega) = T$ for all $n > m$. In particular, $\tau = T$ $\mathbb{P}$-almost surely. $\square$

This lemma implies that the processes X and ξ defined immediately after the proof of Lemma 3.4.6 are indeed defined on the whole interval $[0,T]$.

We can now prove the first existence result.

Theorem 3.4.9. *Assume that $X_0 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P}; H)$. Then equation* (3.1.1) *admits a unique strong solution, which belongs to $\mathscr{J}_2$.*

Proof. Uniqueness of strong solutions is proved, in more generality, by Corollary 3.4.3. Let us prove existence. By stopping at τ_n, one has

$$X_n^{\tau_n}(t) + \int_0^{t\wedge\tau_n} AX_n(s)\,ds + \int_0^{t\wedge\tau_n} \xi_n(s)\,ds = X_0 + \int_0^{t\wedge\tau_n} B_n(X_n(s))\,ds,$$

where, by definition of X, $X_n^{\tau_n} = X^{\tau_n}$, as well as, by definition of B_n,

$$\int_0^{t\wedge\tau_n} B_n(X_n(s))\,ds = \int_0^{t\wedge\tau_n} B(X(s))\,ds.$$

Similarly, by definition of ξ it follows that

$$\int_0^{t\wedge\tau_n} \xi_n(s)\,ds = \int_0^{t\wedge\tau_n} \xi(s)\,ds,$$

hence that

$$X^{\tau_n}(t) + \int_0^{t\wedge\tau_n} AX(s)\,ds + \int_0^{t\wedge\tau_n} \xi(s)\,ds = X_0 + \int_0^{t\wedge\tau_n} B(X(s))\,ds.$$

Since this identity holds for all $n \in \mathbb{N}$ and $\tau_n \to T$ as $n \to \infty$, we infer that

$$X(t) + \int_0^t AX(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(X(s))\,ds$$

for all $t \in [0,T]$ $\mathbb{P}$-a.s.. Moreover, for every $n \in \mathbb{N}$, $\xi_n \in \beta(X_n)$ a.e. in $\Omega \times (0,T) \times D$, hence $\xi_n 1\!\!1_{[\![0,\tau_n]\!]} \in \beta(X_n) 1\!\!1_{[\![0,\tau_n]\!]}$, thus also $\xi 1\!\!1_{[\![0,\tau_n]\!]} \in \beta(X) 1\!\!1_{[\![0,\tau_n]\!]}$ a.e. in $\Omega \times (0,T) \times D$. Recalling that $\tau_n \to T$ as $n \to \infty$, this in turn implies $\xi \in \beta(X)$ a.e. in $\Omega \times (0,T) \times D$.

Moreover, since (X, ξ) is the almost sure limit of (X_n, ξ_n), we immediately infer that X and ξ are predictable H-valued and $L^1(D)$-valued processes, respectively. The a priori estimates of Lemma 3.4.7 and Fatou's lemma then yield

$$X \in L^2(\Omega; C([0,T]; H)) \cap L^2(\Omega; L^2(0,T;V)), \qquad \xi \in L^1(\Omega \times (0,T) \times D).$$

Similarly, $\xi_n \in \beta(X_n)$ implies $X_n \xi_n = j(X_n) + j^*(\xi_n)$, hence

$$\mathbb{E}\int_0^T \int_D \bigl(j(X_n) + j^*(\xi_n)\bigr) \lesssim 1 + \mathbb{E}\|X_0\|^2$$

for all $n \in \mathbb{N}$, and again by Fatou's lemma, as well as by the lower-semicontinuity of convex integrals, one obtains

$$j(X) + j^*(\xi) \in L^1(\Omega; L^1(0,T; L^1(D))).$$

We have thus proved that $(X, \xi) \in \mathscr{J}_2$, so the proof is completed. $\square$

The second existence result, which allows X_0 to be merely $\mathscr{F}_0$-measurable, follows by a further "gluing" procedure.

Theorem 3.4.10. *Assume that $X_0 \in L^0(\Omega, \mathscr{F}_0, \mathbb{P}; H)$. Then equation* (3.1.1) *admits a unique strong solution.*

Proof. Uniqueness of strong solutions has already been proved in Corollary 3.4.3. It is hence enough to prove existence. Let us define the sequence $(\Gamma_n)_{n\in\mathbb{N}}$ of elements of $\mathscr{F}_0$ as

$$\Gamma_n := \bigl\{\omega \in \Omega : \|X_0\| \le n\bigr\}.$$

It is evident that (Γ_n) is a sequence increasing to Ω, and that $X_{0\Gamma_n} = X_0 1\!\!1_{\Gamma_n} \in L^2(\Omega; H)$. Therefore, by the previous theorem, for each $n \in \mathbb{N}$ there exists a unique strong solution (X_n, ξ_n) to (3.1.1) with initial condition $X_{0\Gamma_n}$. By the local property of solutions established in Corollary 3.4.5, we have that $X_{n+1} 1\!\!1_{\Gamma_n}$ and $X_n 1\!\!1_{\Gamma_n}$ are indistinguishable, and $\xi_{n+1} 1\!\!1_{\Gamma_n} = \xi_n 1\!\!1_{\Gamma_n}$ a.e. in $\Omega \times (0,T) \times D$. Since (Γ_n) is increasing, it makes sense to define the processes X and ξ by

$$X 1\!\!1_{\Gamma_n} = X_n 1\!\!1_{\Gamma_n}, \quad \xi 1\!\!1_{\Gamma_n} = \xi_n 1\!\!1_{\Gamma_n}$$

for all $n \in \mathbb{N}$. This amounts to saying that X and ξ are the $\mathbb{P}$-a.s. limits of X_n and ξ_n, respectively, which immediately implies that X and ξ are predictable processes with values in H and $L^1(D)$, respectively. Moreover, by construction, we also have

$$X \in L^0(\Omega; C([0,T]; H) \cap L^2(0,T;V)), \qquad \xi \in L^0(\Omega; L^1(0,T; L^1(D)))$$

In fact, writing $E := C([0,T];H) \cap L^2(0,T;V)$ for compactness of notation, by the previous theorem we have $X_n \in L^2(\Omega;E)$ and $\xi \in L^1(\Omega \times (0,T) \times D)$, and for any arbitrary but fixed ω in a subset of Ω of probability one, there exists $n = n(\omega)$ such that $(X(\omega),\xi(\omega)) = (X_n(\omega),\xi_n(\omega)) \in E \times L^1((0,T) \times D)$. Furthermore, since $\xi_n \in \beta(X_n)$ a.e. for all $n \in \mathbb{N}$, it is easy to see that

$$\xi 1\!\!1_{\Gamma_n} = \xi_n 1\!\!1_{\Gamma_n} \in \beta(X_n)1\!\!1_{\Gamma_n} = \beta(X_n 1\!\!1_{\Gamma_n})1\!\!1_{\Gamma_n} = \beta(X)1\!\!1_{\Gamma_n}$$

for all $n \in \mathbb{N}$, so that $\xi \in \beta(X)$ a.e. because $\Gamma_n \uparrow \Omega$. Similarly,

$$j(X_n)1\!\!1_{\Gamma_n} = j(1\!\!1_{\Gamma_n} X_n)1\!\!1_{\Gamma_n} = j(X)1\!\!1_{\Gamma_n}$$

as well as, by the same reasoning, $j^*(\xi_n)1\!\!1_{\Gamma_n} = j^*(\xi)1\!\!1_{\Gamma_n}$. Since, by the previous theorem, $j(X_n) + j^*(\xi_n) \in L^1(\Omega; L^1((0,T) \times D)$ for all $n \in \mathbb{N}$, it follows that

$$\big(j(X) + j^*(\xi)\big)1\!\!1_{\Gamma_n} \in L^1(\Omega; L^1((0,T) \times D) \qquad \forall n \in \mathbb{N},$$

hence $j(X) + j^*(\xi) \in L^0(\Omega; L^1((0,T) \times D)$. □

3.5 Moment estimates and dependence on the initial datum

We are now going to show that the integrability of the solution is determined by the integrability of the initial condition.

Theorem 3.5.1. *Let $p \geq 0$. If $X_0 \in L^p(\Omega, \mathscr{F}_0, \mathbb{P}; H)$, then the unique strong solution to equation* (3.1.1) *belongs to $\mathscr{J}_p$.*

Proof. Itô's formula yields

$$\begin{aligned}\|X(t)\|^2 &+ 2\int_0^t \langle AX(s), X(s)\rangle\,ds + 2\int_0^t\!\!\int_D \xi(s)X(s)\,dx\,ds\\ &= \|X_0\|^2 + \int_0^t \big\|B(s,X(s))\big\|^2_{\mathscr{L}^2(U,H)}\,ds + 2\int_0^t X(s)B(s,X(s))\,dW(s).\end{aligned}$$

For any $\alpha > 0$, it follows by the integration-by-parts formula that

$$\begin{aligned}e^{-2\alpha t}\|X(t)\|^2 &+ 2\alpha\int_0^t e^{-2\alpha s}\|X(s)\|^2\,ds + 2\int_0^t e^{-2\alpha s}\langle AX(s), X(s)\rangle\,ds\\ &+ 2\int_0^t\!\!\int_D e^{-2\alpha s}\xi(s)X(s)\,dx\,ds\\ &= \|X_0\|^2 + \int_0^t e^{-2\alpha s}\big\|B(s,X(s))\big\|^2_{\mathscr{L}^2(U,H)}\,ds\\ &+ 2\int_0^t e^{-2\alpha s}X(s)B(s,X(s))\,dW(s).\end{aligned}$$

Let M denote the stochastic integral on the right-hand side, and $Y(t) := e^{-\alpha t}X(t)$. Since X has continuous paths in H, one can introduce the sequence of stopping times $(T_n)_{n\in\mathbb{N}}$, increasing to T, as

$$T_n := \inf\big\{t \geq 0 : \|X(t)\| \geq n\big\} \wedge T.$$

It follows by the local Lipschitz-continuity property of B that

$$
\begin{aligned}
&\|Y^{T_n}(t)\|^2 + 2\alpha \int_0^{t\wedge T_n} \|Y(s)\|^2\,ds + 2C\int_0^{t\wedge T_n} \|Y(s)\|_V^2\,ds \\
&\qquad + 2\int_0^{t\wedge T_n}\int_D e^{-2\alpha s}\xi(s)X(s)\,dx\,ds \\
&\qquad \le \|X_0\|^2 + \int_0^{t\wedge T_n} e^{-2\alpha s}\big\|B_n(s,X(s))\big\|^2_{\mathscr{L}^2(U,H)}\,ds + 2M^{T_n}(t).
\end{aligned}
$$

Recalling that $B_n = B(\cdot,\cdot,\sigma_n(\cdot))$ and σ_n is a contraction in H, and that B grows at most linearly, one has

$$
e^{-2\alpha s}\big\|B_n(s,X(s))\big\|^2_{\mathscr{L}^2(U,H)} \lesssim e^{-2\alpha s} + \|Y(s)\|^2,
$$

hence

$$
\int_0^{t\wedge T_n} e^{-2\alpha s}\big\|B_n(s,X(s))\big\|^2_{\mathscr{L}^2(U,H)}\,ds \lesssim \frac{1}{2\alpha} + \int_0^{t\wedge T_n}\|Y(s)\|^2\,ds. \tag{3.5.3}
$$

Taking supremum in time and the $L^{p/2}(\Omega)$-(quasi)norm, recalling the BDG inequality and the fact that $e^{-\alpha t}\xi_n X_n \ge e^{-\alpha T}\xi_n X_n$, we are left with

$$
\begin{aligned}
&\big\|Y^*_{T_n}\big\|^2_{L^p(\Omega)} + \alpha\big\|Y\mathbb{1}_{[\![0,T_n]\!]}\big\|^2_{L^p(\Omega;L^2(0,T;H))} + \big\|Y\mathbb{1}_{[\![0,T_n]\!]}\big\|^2_{L^p(\Omega;L^2(0,T;V))} \\
&\qquad + e^{-\alpha T}\big\|\xi X\mathbb{1}_{[\![0,T_n]\!]}\big\|_{L^{p/2}(\Omega;L^1((0,T)\times D))} \\
&\lesssim \big\|X_0\big\|^2_{L^p(\Omega;H)} + \frac{1}{2\alpha} + \big\|Y\mathbb{1}_{[\![0,T_n]\!]}\big\|^2_{L^p(\Omega;L^2(0,T;H))} + \big\|[M,M]^{1/2}_{T_n}\big\|_{L^{p/2}(\Omega)}.
\end{aligned}
$$

Lemma 3.4.1 and (3.5.3) yield

$$
[M,M]^{1/2}_{T_n} \lesssim \varepsilon Y^{*2}_{T_n} + \frac{1}{\varepsilon}\bigg(\frac{1}{2\alpha} + \big\|Y\mathbb{1}_{[\![0,T_n]\!]}\big\|^2_{L^2(0,T;H)}\bigg),
$$

hence

$$
\big\|[M,M]^{1/2}_{T_n}\big\|_{L^{p/2}(\Omega)} \lesssim \varepsilon\big\|Y^*_{T_n}\big\|^2_{L^p(\Omega)} + \frac{1}{\varepsilon}\big\|Y\mathbb{1}_{[\![0,T_n]\!]}\big\|^2_{L^p(\Omega;L^2(0,T;H))} + \frac{1}{2\alpha\varepsilon},
$$

where the implicit constant is independent of α and of an arbitrary $\varepsilon > 0$ to be chosen later. We thus have

$$
\begin{aligned}
&\big\|Y^*_{T_n}\big\|^2_{L^p(\Omega)} + \alpha\big\|Y\mathbb{1}_{[\![0,T_n]\!]}\big\|^2_{L^p(\Omega;L^2(0,T;H))} + \big\|Y\mathbb{1}_{[\![0,T_n]\!]}\big\|^2_{L^p(\Omega;L^2(0,T;V))} \\
&\qquad + e^{-\alpha T}\big\|\xi X\mathbb{1}_{[\![0,T_n]\!]}\big\|_{L^{p/2}(\Omega;L^1((0,T)\times D))} \\
&\lesssim \big\|X_0\big\|^2_{L^p(\Omega;H)} + \varepsilon\big\|(Y^{T_n})^*_T\big\|^2_{L^p(\Omega)} + (1+1/\varepsilon)\big\|Y\mathbb{1}_{[\![0,T_n]\!]}\big\|^2_{L^p(\Omega;L^2(0,T;H))} \\
&\qquad + \frac{1}{2\alpha}(1+1/\varepsilon).
\end{aligned}
$$

Since the implicit constant is independent of α and ε, one can take ε small enough and α large enough so that

$$
\begin{aligned}
&\big\|Y^*_{T_n}\big\|^2_{L^p(\Omega)} + \big\|Y\mathbb{1}_{[\![0,T_n]\!]}\big\|^2_{L^p(\Omega;L^2(0,T;V))} + \big\|\xi X\mathbb{1}_{[\![0,T_n]\!]}\big\|_{L^{p/2}(\Omega;L^1((0,T)\times D))} \\
&\qquad \lesssim 1 + \big\|X_0\big\|^2_{L^p(\Omega;H)}.
\end{aligned}
$$

As the implicit constant is independent of n and T_n increases to T, we get

$$\begin{aligned}\|Y\|^2_{L^p(\Omega;C([0,T];H))} + \|Y\|^2_{L^p(\Omega;L^2(0,T;V))} + \|\xi X\|_{L^{p/2}(\Omega;L^1((0,T)\times D))} \\ \lesssim 1 + \|X_0\|^2_{L^p(\Omega;H)}.\end{aligned}$$

The proof is completed noting that, for $E := C([0,T];H) \cap L^2(0,T;V)$,

$$\|X\|_E \leq e^{\alpha T}\|Y\|_E. \qquad \square$$

If B is Lipschitz-continuous, related arguments show that the solution map is Lipschitz-continuous between spaces with finite p-th moment in the whole range $p \in [0,\infty[$. We consider the cases $p > 0$ and $p = 0$ separately.

Proposition 3.5.2. *Let $p > 0$. If B is Lipschitz-continuous in the sense of assumption* (B1)*, then the solution map*

$$\begin{aligned}L^p(\Omega;H) &\longrightarrow L^p(\Omega;C([0,T];H)) \cap L^p(\Omega;L^2(0,T;V)) \\ X_0 &\longmapsto X\end{aligned}$$

is Lipschitz-continuous.

Proof. Let X_0, $Y_0 \in L^p(\Omega;H)$. The previous theorem asserts that the (unique) strong solutions (X,ξ) and (Y,η) to (3.1.1) with initial condition X_0 and Y_0, respectively, belong to $L^p(\Omega;E)$, where, as before, E stands for $C([0,T];H) \cap L^2(0,T;V)$. By Itô's formula,

$$\begin{aligned}\|X-Y\|^2 + 2\int_0^t \langle A(X-Y), X-Y\rangle\,ds + 2\int_0^t\int_D (\xi-\eta)(X-Y)\,ds \\ = \|X_0-Y_0\|^2 + \int_0^t \big\|B(X)-B(Y)\big\|^2_{\mathscr{L}^2(U,H)}\,ds \\ + 2\int_0^t (X-Y)(B(X)-B(Y))\,dW,\end{aligned}$$

where the third term on the left-hand side is positive by monotonicity of β. Let $\alpha > 0$ be a constant to be chosen later, and set $X_\alpha := Xe^{-\alpha\cdot}$, $Y_\alpha := Ye^{-\alpha\cdot}$. It follows by the integration-by-parts formula, in complete analogy to the proof of the previous theorem, by the Lipschitz continuity of B, and by the coercivity of A, that

$$\begin{aligned}\|X_\alpha - Y_\alpha\|^2 + \alpha\int_0^t \|X_\alpha-Y_\alpha\|^2\,ds + \int_0^t \|X_\alpha-Y_\alpha\|_V^2\,ds \\ \lesssim \|X_0-Y_0\|^2 + \int_0^t \big\|X_\alpha-Y_\alpha\big\|^2\,ds + M,\end{aligned}$$

where $M := \big(e^{-2\alpha\cdot}(X-Y)(B(X)-B(Y))\big)\cdot W$. Taking supremum in time and the $L^{p/2}(\Omega)$-(quasi)norm yields

$$\begin{aligned}\big\|X_\alpha - Y_\alpha\big\|^2_{L^p(\Omega;C([0,T];H))} + \alpha\big\|X_\alpha-Y_\alpha\big\|^2_{L^p(\Omega;L^2(0,T;H))} \\ + \big\|X_\alpha-Y_\alpha\big\|^2_{L^p(\Omega;L^2(0,T;V))} \\ \lesssim \big\|X_0-Y_0\big\|^2_{L^p(\Omega;H)} + \big\|X_\alpha-Y_\alpha\big\|^2_{L^p(\Omega;L^2(0,T;H))} + \big\|M_T^*\big\|_{L^{p/2}(\Omega)},\end{aligned}$$

where, by Lemma 3.4.1,

$$\big\|M_T^*\big\|_{L^{p/2}(\Omega)} \lesssim \varepsilon\big\|X_\alpha - Y_\alpha\big\|^2_{L^p(\Omega;C([0,T];H))} + N(\varepsilon)\big\|X_\alpha - Y_\alpha\big\|^2_{L^p(\Omega;L^2(0,T;H))}$$

for any $\varepsilon > 0$. Choosing first ε small enough, then α sufficiently large, we obtain

$$\big\|X_\alpha - Y_\alpha\big\|^2_{L^p(\Omega;C([0,T];H))} + \big\|X_\alpha - Y_\alpha\big\|^2_{L^p(\Omega;L^2(0,T;V))} \lesssim \big\|X_0 - Y_0\big\|^2_{L^p(\Omega;H)},$$

which completes the proof noting that $\|X - Y\|_E \leq e^{\alpha T}\|X_\alpha - Y_\alpha\|_E$. □

Lipschitz continuity of the solution map can also be obtained in the case $p = 0$. As already seen, the space $E := C([0,T];H) \cap L^2(0,T;V)$, equipped with the norm

$$\big\|u\big\|_E := \big\|u\big\|_{C([0,T];H)} + \big\|u\big\|_{L^2(0,T;V)},$$

is a Banach space. Then $L^0(\Omega;E)$, endowed with the topology of convergence in probability, is a complete metrizable topological vector space. In particular, the distance

$$d(f,g) := \mathbb{E}\big(\|f-g\|_E \wedge 1\big)$$

generates its topology.

Proposition 3.5.3. *If B is Lipschitz-continuous in the sense of assumption* (B1)*, then the solution map*

$$\begin{aligned} L^0(\Omega;H) &\longrightarrow L^0(\Omega;E) \\ X_0 &\longmapsto X \end{aligned}$$

is Lipschitz-continuous.

Proof. Let X_0, $Y_0 \in L^0(\Omega,\mathscr{F}_0,\mathbb{P};H)$, and (X,ξ), (Y,η) the unique solutions in $\mathscr{J}_0$ to equation (3.1.1) with initial datum X_0 and Y_0, respectively. The stopping time

$$T_1 := \inf\Big\{t \geq 0 :\ (X-Y)_t^* + \Big(\int_0^t \|X(s) - Y(s)\|_V^2\,ds\Big)^{1/2} \geq 1\Big\} \wedge T.$$

is well defined thanks to the pathwise continuity of X and Y. For every $\alpha > 0$, using the same notation as in the previous proof, Theorem 3.3.2 yields, by monotonicity of β and coercivity of A,

$$\begin{aligned} \big(X_\alpha - Y_\alpha\big)_t^{*2} &+ \int_0^t \|X_\alpha(s) - Y_\alpha(s)\|_V^2\,ds + \alpha\int_0^t \|X_\alpha(s) - Y_\alpha(s)\|^2\,ds \\ &\lesssim \|X_0 - Y_0\|^2 + \int_0^t \big\|(B(X(s)) - B(Y(s)))_\alpha\big\|^2_{\mathscr{L}^2(U,H)}\,ds \\ &\quad + \big((X_\alpha - Y_\alpha)(B(X) - B(Y))_\alpha \cdot W\big)_t^* \end{aligned}$$

Raising to the power 1/2, stopping at T_1, and taking expectation, we get, by the Lipschitz continuity of B,

$$\begin{aligned}
&\mathbb{E}\bigl(X_\alpha - Y_\alpha\bigr)^*_{T_1} + \mathbb{E}\biggl(\int_0^{T_1} \|X_\alpha(s) - Y_\alpha(s)\|_V^2\,ds\biggr)^{1/2}\\
&\qquad + \sqrt{\alpha}\,\mathbb{E}\biggl(\int_0^{T_1} \|X_\alpha(s) - Y_\alpha(s)\|^2\,ds\biggr)^{1/2}\\
&\qquad \lesssim \mathbb{E}\,\mathbb{1}_{[\![0,T_1]\!]}\|X_0 - Y_0\| + \mathbb{E}\biggl(\int_0^{T_1} \|X_\alpha(s) - Y_\alpha(s)\|^2\,ds\biggr)^{1/2}\\
&\qquad + \mathbb{E}\bigl((X_\alpha - Y_\alpha)(B(X) - B(Y))_\alpha \cdot W\bigr)^{*1/2}_{T_1},
\end{aligned}$$

where, by Lemma 3.4.1 and Lipschitz continuity of B, the last term on the right-hand side is bounded by

$$\varepsilon\,\mathbb{E}\bigl(X_\alpha - Y_\alpha\bigr)^*_{T_1} + N(\varepsilon)\,\mathbb{E}\biggl(\int_0^{T_1} \|X_\alpha(s) - Y_\alpha(s)\|^2\,ds\biggr)^{1/2}$$

for every $\varepsilon > 0$. Therefore, choosing ε small enough and α large enough, we are left with

$$\mathbb{E}\bigl(X - Y\bigr)^*_{T_1} + \mathbb{E}\biggl(\int_0^{T_1} \|X(s) - Y(s)\|_V^2\,ds\biggr)^{1/2} \lesssim \mathbb{E}\,\mathbb{1}_{[\![0,T_1]\!]}\|X_0 - Y_0\|.$$

The proof is concluded noting that, by definition of T_1,

$$\bigl(X - Y\bigr)^*_{T_1} + \biggl(\int_0^{T_1} \|X(s) - Y(s)\|_V^2\,ds\biggr)^{1/2} = \bigl\|X - Y\bigr\|_{C([0,T];H)\cap L^2(0,T;V)} \wedge 1,$$

and $\|X_0 - Y_0\| \leq 1$ on $[\![0, T_1]\!]$, hence

$$\mathbb{E}\,\mathbb{1}_{[\![0,T_1]\!]}\|X_0 - Y_0\| = \mathbb{E}\,\mathbb{1}_{[\![0,T_1]\!]}\bigl(\|X_0 - Y_0\| \wedge 1\bigr) \leq \mathbb{E}\bigl(\|X_0 - Y_0\| \wedge 1\bigr). \qquad \square$$

Chapter 4

Singular semilinear equations: long-time behaviour

In this chapter, 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.

The results presented in this chapter are part of a joint work with Carlo Marinelli: see [62].

4.1 The problem: literature and main goals

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

$$dX_t + AX_t\,dt + \beta(X_t)\,dt \ni B(X_t)\,dW_t, \qquad X(0) = X_0. \tag{4.1.1}$$

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 §4.2 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 §4.2 below. Well-posedness of equation (4.1.1) in the strong (variational) sense has been obtained in Chapters 2 and 3 by a combination of classical results by Pardoux and Krylov-Rozovskiĭ (see [46, 72]) 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 [46, 72]. For this reason, questions such as ergodicity and existence of invariant measures for (4.1.1) cannot be addressed using the results by Barbu and Da Prato in [11], which appear to be the only ones available for equations in the variational setting (cf. also [68]). 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 [28, 29, 78].

Even in this case, however, we are not aware of results on equations with a drift term as general as in (4.1.1). Our results thus considerably extend, or at least complement, those on reaction-diffusion equations in [26, 28, 29], 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 [37, 38], 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 [37]). These results allow us to show that the Kolmogorov operator associated to the stochastic equation (4.1.1) 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 chapter and of the main results. In Section 4.2 we state the basic assumptions which are in force throughout the paper, and recall the well-posedness result for equation (4.1.1) obtained in Chapter 2. Section 4.3 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 equation (4.1.1): 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 [72, p. 62–ff]). Furthermore, we establish results on the first and second-order differentiability, both in the Gâteaux and Fréchet sense, of (variational) solutions to semilinear equations with regular drift with respect to the initial datum. In Section 4.4 we prove that the transition semigroup P generated by the solution to (4.1.1) 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 4.5 deals with the Kolmogorov equation associated to (4.1.1). 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 4.3 play a key role.

4.2 General assumptions and well-posedness

Before stating the hypotheses on the coefficients and on the initial datum of equation (4.1.1) that will be in force throughout the paper, let us fix some notation.

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. 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 $\langle Av, v\rangle \geq C\|v\|_V^2$ 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;

(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 $\beta = \partial j$, in the sense of convex analysis. We assume that

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

This hypothesis is obviously satisfied if j (or, equivalently, β) is symmetric. 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, 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.* We assume that

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

$$\begin{aligned} \|B(x) - B(y)\|_{\mathscr{L}^2(U,H)} &\leq L_B\|x - y\| &\qquad \forall x, y \in H, \\ \|B(x)\|_{\mathscr{L}^2(U,H)} &\leq L_B(1 + \|x\|) &\qquad \forall x \in H. \end{aligned}$$

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 equation (4.1.1) has been proved in Chapter 2, allowing the coefficient B to be also random and time-dependent.

Theorem 4.2.1. *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*

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

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, X is $\mathbb{P}$-a.s. pathwise weakly continuous from $[0, T]$ to H, and the solution map

$$\begin{aligned} L^2(\Omega; H) &\longrightarrow L^2(\Omega; L^\infty(0, T; H)) \cap L^2(\Omega; L^2(0, T; V)) \\ X_0 &\longmapsto X \end{aligned}$$

is Lipschitz-continuous.

The main result of this chapter will be stated at the end of Section 4.5 and will follow after some intermediate steps.

*Expressions involving random elements are always meant to hold $\mathbb{P}$-a.s. unless otherwise stated.

4.3 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.

4.3.1 Itô formulas

The following version of Itô's formula for the square of the norm is proved in Chapter 3.

Proposition 4.3.1. *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

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

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 $(I + \delta A_1)^{-k}$, we have

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

where $g^\delta \in L^1(0,T;H)$, hence the classical Itô formula yields, for every $\delta > 0$,

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

We are now going to pass to the limit as $\delta \to 0$. By the assumptions on A and the regularity properties of Y, g, Y_0, and G, one has

$$\begin{aligned}
Y^\delta(t) &\to Y(t) && \text{in } H \quad \forall t \in [0,T],\\
Y^\delta &\to Y && \text{in } L^2(0,T;V),\\
AY^\delta &\to AY && \text{in } L^2(0,T;V'),\\
g^\delta &\to g && \text{in } L^1(0,T;L^1(D)),\\
Y_0^\delta &\to Y_0 && \text{in } H,\\
G^\delta &\to G && \text{in } L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H))).
\end{aligned}$$

This implies

$$\int_0^t \langle AY^\delta(s), Y^\delta(s)\rangle\,ds \longrightarrow \int_0^t \langle AY(s), Y(s)\rangle\,ds$$

and

$$\int_0^t \|G^\delta(s)\|^2_{\mathscr{L}^2(U,H)}\,ds \longrightarrow \int_0^t \|G(s)\|^2_{\mathscr{L}^2(U,H)}\,ds.$$

Using the dominated convergence theorem, it is not difficult to check that $\|Y^\delta G^\delta - YG\|^2_{\mathscr{L}^2(U,\mathbb{R})}$ converges to zero in probability, hence also (along a subsequence)

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

Finally, the symmetry assumption on j ensures that $(g^\delta Y^\delta)$ is uniformly integrable on $(0,T)\times D$, so that

$$\int_0^t\int_D g^\delta(s,x)Y^\delta(s,x)\,dx\,ds \to \int_0^t\int_D g(s,x)Y(s,x)\,dx\,ds. \qquad \square$$

We shall also need a simplified version of an Itô formula in the variational setting, due to Pardoux, for functions more general than the square of the H-norm. For its proof (in a more general context) we refer to [72, p. 62-ff.].

Proposition 4.3.2. *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

$$\begin{aligned}
F(Y(t)) = F(Y_0) &+ \int_0^t DF(Y(s))v(s)\,ds + \int_0^t DF(Y(s))G(s)\,dW(s)\\
&+ \frac{1}{2}\int_0^t \operatorname{Tr}\big(G^*(s)D^2F(Y(s))G(s)\big)\,ds
\end{aligned}$$

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 4.3.1.

Proposition 4.3.3. *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 $F \in C^2_b(H) \cap C^1_b(V') \cap C^1_b(L^1(D))$, one has

$$\begin{aligned} F(Y(t)) = F(Y_0) &+ \int_0^t \langle Av(s), DF(Y(s))\rangle\,ds + \int_0^t\int_D g(s)DF(Y(s))\,ds \\ &+ \int_0^t DF(Y(s))G(s)\,dW(s) + \frac{1}{2}\int_0^t \mathrm{Tr}\big(G^*(s)D^2F(Y(s))G(s)\big)\,ds \end{aligned}$$

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 4.3.2, for every $\delta > 0$ we have

$$\begin{aligned} F(Y^\delta(t)) = F(Y_0^\delta) &+ \int_0^t \langle Av^\delta(s), DF(Y^\delta(s))\rangle\,ds + \int_0^t\int_D g^\delta(s)DF(Y^\delta(s))\,ds \\ &+ \int_0^t DF(Y^\delta(s))G^\delta(s)\,dW(s) + \frac{1}{2}\int_0^t \mathrm{Tr}\big((G^\delta)^*(s)D^2F(Y^\delta(s))G^\delta(s)\big)\,ds \end{aligned}$$

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 $v^\delta + \delta Av^\delta = v$ 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 $\int_D g^\delta DF(Y^\delta) \le \|DF\|_{C_b(H,L^\infty(D))}\|g\|_{L^1(D)} \in L^1(0,T)$, 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 $D^2F \in C(H, \mathscr{L}(H))$ and $G^\delta \to G$ in $L^2(\Omega; L^2(0,T;\mathscr{L}^2(U,H)))$, we have (possibly along a subsequence)

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

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

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

Finally, by the Davis inequality and the ideal property of Hilbert-Schmidt operators, we have

$$\begin{aligned}
&\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t DF(Y^\delta(s))G^\delta(s)\, dW(s) - \int_0^t DF(Y(s))G(s)\, dW(s)\right| \\
&\lesssim \mathbb{E}\left(\int_0^T \big\|DF(Y^\delta(s))G^\delta(s) - DF(Y(s))G(s)\big\|^2_{\mathscr{L}(U,\mathbb{R})}\, ds\right)^{1/2} \\
&\lesssim \mathbb{E}\left(\int_0^T \|DF(Y^\delta(s))\|^2 \|G^\delta(s) - G(s)\|^2_{\mathscr{L}^2(U,H)}\, ds\right)^{1/2} \\
&\qquad + \mathbb{E}\left(\int_0^T \|DF(Y^\delta(s)) - DF(Y(s))\|^2 \|G^\delta(s)\|^2_{\mathscr{L}^2(U,H)}\, ds\right)^{1/2} \\
&\le \|DF\|_{C(H,H)}\|G^\delta - G\|_{L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,H)))} \\
&\qquad + \mathbb{E}\left(\int_0^T \|G(s)\|^2_{\mathscr{L}^2(U,H)} \|DF(Y^\delta(s)) - DF(Y(s))\|^2\, ds\right)^{1/2},
\end{aligned}$$

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

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

4.3.2 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 4.2, $G \in \mathscr{L}^2(U, H)$, and $x \in H$.

For compactness of notation we 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., [26, 29, 59]), 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 4.3.4. *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*

$$Y_h' + AY_h = g'(X^x)Y_h, \qquad Y_h(0) = h, \tag{4.3.2}$$

in the variational sense.

Proof. Classical (deterministic) results imply that (4.3.2) 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 (4.3.2), 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

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

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

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

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

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

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

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

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

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

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

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

we conclude that

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

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 (4.3.2). 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

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

from which one infers

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

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

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

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

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

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

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

is continuous. This implies, by a well-known criterion (see, e.g., [3, 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

$$Y^n_h(t) - Y_h(t) + \int_0^t A(Y^n_h(s) - Y_h(s))\, ds = x_n - x + \int_0^t \bigl(g'(X^n)Y^n_h - g'(X)Y_h\bigr)(s)\, ds,$$

for which the integration-by-parts formula yields

$$\begin{aligned} \frac{1}{2}\bigl\|Y^n_h(t) - Y_h(t)\bigr\|^2 &+ C\int_0^t \bigl\|Y^n_h(s) - Y_h(s)\bigr\|^2_V\, ds \\ &\leq \frac{1}{2}\|x^n - x\|^2 + \int_0^t \bigl\langle g'(X^n)Y^n_h - g'(X)Y_h, Y^n_h - Y_h\bigr\rangle(s)\, ds, \end{aligned}$$

where

$$\begin{aligned} \bigl\langle g'(X^n)Y^n_h - g'(X)Y_h, Y^n_h - Y_h\bigr\rangle &= \bigl\langle g'(X_n)(Y^n_h - Y_h), Y^n_h - Y_h\bigr\rangle \\ &\quad + \bigl\langle (g'(X^n) - g'(X))Y_h, Y^n_h - Y_h\bigr\rangle, \end{aligned}$$

so that, by elementary estimates,

$$\begin{aligned} \bigl\|Y^n_h(t) - Y_h(t)\bigr\|^2 &+ 2C\int_0^t \bigl\|Y^n_h(s) - Y_h(s)\bigr\|^2_V\, ds \\ &\leq \|x_n - x\|^2 + \bigl(2\|g\|_{\dot{C}^{0,1}} + 1\bigr)\int_0^t \bigl\|Y^n_h(s) - Y_h(s)\bigr\|^2\, ds \\ &\quad + \int_0^t \bigl\|\bigl(g'(X^n(s)) - g'(X(s))\bigr)Y_h(s)\bigr\|^2\, ds. \end{aligned}$$

Taking the supremum in time, Gronwall's inequality implies

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

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

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

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 $g \in C^2_b$ (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 $\|Y^x_h\|_{L^p(\Omega;E)} \lesssim \|h\|$ for all $p \geq 0$ with a constant independent of x, we conclude that $x \mapsto X^x$ is of class C^1_b 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

$$Z'_{hk} + AZ_{hk} = g'(X)Z_{hk} + g''(X)Y_hY_k, \qquad Z_{hk}(0) = 0, \tag{4.3.3}$$

where $h, k \in H$ and Y_h, Y_k are the solutions to (4.3.2) 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 (4.3.3) 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., [50]).

Lemma 4.3.5. *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, for every $t \in [0,T]$,

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

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

$$
\begin{aligned}
\|y_n(t) - y_m(t)\|^2 &+ 2C\int_0^t \|y_n(s) - y_m(s)\|_V^2\,ds \\
&\le 2\|\ell\|_{L^\infty((0,T)\times D)} \int_0^t \|y_n(s) - y_m(s)\|^2\,ds \\
&\quad + 2\int_0^t \langle f_n(s) - f_m(s), y_n(s) - y_m(s)\rangle\,ds \\
&\le 2\|\ell\|_{L^\infty((0,T)\times D)} \int_0^t \|y_n(s) - y_m(s)\|^2\,ds \\
&\quad + 2\|y_n - y_m\|_{C([0,t];H)}\|f_n - f_m\|_{L^1(0,T;H)}
\end{aligned}
$$

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

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

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

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

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. □

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 $d \le 4$. In fact, by the Sobolev embedding theorem, $H_0^1 \hookrightarrow L^{2^*}$, where

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

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

We proceed as follows: first we establish well-posedness for equation (4.3.3), and then we show that its unique solution identifies D^2X.

Proposition 4.3.6. *Assume that V is continuously embedded in $L^4(D)$. Then equation* (4.3.3) *admits a unique variational solution Z_{hk} for any h, $h \in H$. Moreover, the map*

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

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

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

so that $g''(X)Y_hY_k \in L^1(0,T;H)$ since $Y_h, Y_k \in L^2(0,T;V)$. Hence, Lemma 4.3.5 implies that there is a unique

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

such that, for every $t \in [0,T]$,

$$Z_{hk}(t) + \int_0^t AZ_{hk}(s)\,ds = \int_0^t g'(X(s))Z_{hk}(s)\,ds + \int_0^t g''(X(s))Y_h(s)Y_k(s)\,ds.$$

Let us show that $(h,k) \mapsto Z_{hk}$ is a continuous bilinear map. The bilinearity is clear from equation (4.3.3). Moreover, testing by Z_{hk} and using the coercivity of A we have that

$$\begin{aligned}
&\|Z_{hk}(t)\|^2 + \int_0^t \|Z_{hk}(s)\|_V^2\,ds \\
&\qquad \lesssim \|g\|_{C_b^1}\int_0^t \|Z_{hk}(s)\|^2\,ds + \|g\|_{C_b^2}\int_0^t \|Y_h(s)\|_V\|Y_k(s)\|_V\,ds \\
&\qquad \leq \|g\|_{C_b^1}\int_0^t \|Z_{hk}(s)\|^2\,ds + \|g\|_{C_b^2}\|Y_h\|_{L^2(0,T;V)}\|Y_k\|_{L^2(0,T;V)} \\
&\qquad \lesssim_T \|g\|_{C_b^1}\int_0^t \|Z_{hk}(s)\|^2\,ds + \|g\|_{C_b^2}\|h\|\|k\|
\end{aligned}$$

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 4.3.7. *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

$$z^\varepsilon : (h,k) \longmapsto z_{hk}^\varepsilon := \frac{1}{\varepsilon}\left(Y_h^{x+\varepsilon k} - Y_h^x\right) - Z_{hk}^x.$$

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

$$\begin{aligned}
&z_{hk}^\varepsilon(t) + \int_0^t Az_{hk}^\varepsilon(s)\,ds \\
&\qquad = \int_0^t \Big(\frac{g'(X^\varepsilon)Y_h^\varepsilon - g'(X)Y_h}{\varepsilon} - g'(X)Z_{hk} - g''(X)Y_hY_k\Big)(s)\,ds,
\end{aligned}$$

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

$$\begin{aligned} R_\varepsilon &= \Bigl(\frac{g'(X^\varepsilon) - g'(X)}{\varepsilon} - g''(X)Y_k\Bigr)Y_h, \\ S_\varepsilon &= \Bigl(g'(X^\varepsilon)\frac{Y_h^\varepsilon - Y_h}{\varepsilon} - g'(X)Z_{hk}\Bigr). \end{aligned}$$

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

$$\begin{aligned} R'_\varepsilon &:= \Bigl(\frac{g'(X+\varepsilon Y_k) - g'(X)}{\varepsilon} - g''(X)Y_k\Bigr)Y_h, \\ R''_\varepsilon &:= \frac{g'(X^\varepsilon) - g'(X+\varepsilon Y_k)}{\varepsilon}Y_h, \\ S'_\varepsilon &:= g'(X)z^\varepsilon_{hk}, \\ S''_\varepsilon &:= \bigl(g'(X^\varepsilon) - g'(X)\bigr)\frac{Y_h^\varepsilon - Y_h}{\varepsilon}. \end{aligned}$$

The integration-by-parts formula and obvious estimates yield

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

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

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

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

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

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

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

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 4.3.4,

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

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

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

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

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

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

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

We thus conclude that

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

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

$$\begin{aligned} x &\longmapsto Z^x \\ H &\longrightarrow \mathscr{L}_2(H;L^2(\Omega;E)) \end{aligned}$$

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,

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

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

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

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

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

where

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

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

$$\begin{aligned}\int_0^t \langle R_i, Z^n_{hk} - Z_{hk}\rangle(s)\,ds &\leq \bigl\|Z^n_{hk} - Z_{hk}\bigr\|_{C([0,t];H)}\bigl\|R_i\bigr\|_{L^1(0,t;H)}\\ &\leq \delta\bigl\|Z^n_{hk} - Z_{hk}\bigr\|^2_{C([0,t];H)} + \frac{1}{\delta}\bigl\|R_i\bigr\|^2_{L^1(0,t;H)}.\end{aligned}$$

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

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

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

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

from which it follows

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

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

$$\begin{aligned}\bigl\|R_3\bigr\|_{L^1(0,T;H)} \leq \bigl\|Y^n_h - Y_h\bigr\|_{L^2(0,T;V)}&\bigl\|Y_k\bigr\|_{L^2(0,T;V)}\\ &+ \bigl\|Y^n_k - Y_k\bigr\|_{L^2(0,T;V)}\bigl\|Y^n_h\bigr\|_{L^2(0,T;V)},\end{aligned}$$

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

$$\bigl\|Y^n_h Y^n_k - Y_h Y_k\bigr\|_{L^1(0,T;H)} \lesssim \|h\|\|k\|$$

then implies, by the dominated convergence theorem,

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

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,

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

hence the dominated convergence theorem yields

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

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

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

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. □

4.4 Invariant measures

Throughout this section, we consider equation (4.1.1) 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 (4.1.1) is Markovian. Let $P = (P_t)_{t\geq 0}$ be the transition semigroup defined by

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

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

$$\langle Ax, x\rangle \geq \frac{1}{2}\|B(x)\|^2_{\mathscr{L}^2(U,H)} + C\|x\|^2_V - C_0 \qquad \forall x \in V, \tag{4.4.4}$$

with $C_0 > 0$ a constant.

Theorem 4.4.1. *The set of invariant measures for the transition semigroup $(P_t)_t$ is not empty.*

Proof. Let (X, ξ) be the unique strong solution to (4.1.1). For every $t \geq 0$ one has, by Proposition 4.3.1,

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

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 $\mathbb{E}[M,M]^{1/2}_T$ is finite: one has, by the ideal property of Hilbert-Schmidt operators and the Cauchy-Schwarz inequality,

$$\begin{aligned}\mathbb{E}[M,M]^{1/2}_T &= \mathbb{E}\bigg(\int_0^T \big\|XB(X)\big\|^2_{\mathscr{L}^2(U,\mathbb{R})}\,ds\bigg)^{1/2}\\ &\leq \mathbb{E}\big\|X\big\|_{L^\infty(0,T;H)}\bigg(\int_0^T \big\|B(X)\big\|^2_{\mathscr{L}^2(U,H)}\,ds\bigg)^{1/2}\\ &\leq \Big(\mathbb{E}\big\|X\big\|^2_{L^\infty(0,T;H)}\Big)^{1/2}\bigg(\mathbb{E}\int_0^T \big\|B(X)\big\|^2_{\mathscr{L}^2(U,H)}\,ds\bigg)^{1/2},\end{aligned}$$

where the last term is finite thanks to Theorem 4.2.1 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 (4.4.4) into account,

$$C\,\mathbb{E}\int_0^t \|X(s)\|^2_V\,ds + \mathbb{E}\int_0^t\int_D j(X(s))\,ds + \mathbb{E}\int_0^t\int_D j^*(\xi(s))\,ds \leq \frac{1}{2}\|x\|^2 + C_0 t \tag{4.4.5}$$

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 (4.4.5) yield

$$\begin{aligned}\mu_t(B_n^c) &= \frac{1}{t}\int_0^t \pi_s(B_n^c)\,ds = \frac{1}{t}\int_0^t \mathbb{P}\big(\|X(s)\|_V^2 > n^2\big)\,ds\\ &\leq \frac{1}{tn^2}\int_0^t \mathbb{E}\|X(s)\|_V^2\,ds \leq \frac{1}{Ctn^2}\,C_0 t = \frac{C_0}{Cn^2},\end{aligned}$$

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. □

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 4.4.2. *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 4.3.2 to the process X and the function $G_\delta : x \mapsto g_\delta\big(\|x\|^2\big)$, 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

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

whence

$$\begin{aligned}g_\delta\big(\|X(t)\|^2\big) &+ 2\int_0^t g_\delta'\big(\|X(s)\|^2\big)\big(\langle AX(s), X(s)\rangle + \langle \xi(s), X(s)\rangle\big)\,ds\\ &- 2\int_0^t g_\delta''\big(\|X(s)\|^2\big)\big\|X(s)B(X(s))\big\|^2_{\mathscr{L}^2(U,\mathbb{R})}\,ds\\ &= g_\delta\big(\|x\|^2\big) + 2\int_0^t g_\delta'\big(\|X(s)\|^2\big)X(s)B(X(s))\,dW(s)\\ &\quad + \int_0^t g_\delta'\big(\|X(s)\|^2\big)\big\|B(X(s))\big\|^2_{\mathscr{L}^2(U,\mathbb{R})}\,ds.\end{aligned}$$

Since $g_\delta' > 0$ and $g_\delta'' < 0$, the coercivity condition (4.4.4) and the monotonicity of β imply

$$\begin{aligned} g_\delta\big(\|X(t)\|^2\big) + 2\int_0^t g_\delta'\big(\|X(s)\|^2\big)\big(C\|X(s)\|_V^2 - C_0\big)\,ds \\ \leq g_\delta\big(\|x\|^2\big) + 2\int_0^t g_\delta'\big(\|X(s)\|^2\big) X(s)B(X(s))\,dW(s). \end{aligned}$$

Taking into account that $|g_\delta'| \leq 1$, the stochastic integral is a martingale, exactly as in the proof of Theorem 4.4.1, hence has expectation zero, so that

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

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

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

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

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

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

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

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

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

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

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

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

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

whose Borel measurability will be proved in Lemma 4.4.4 below.

Theorem 4.4.3. *Let μ be an invariant measure for the transition semigroup P. Then one has*

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

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 Φ, Ψ, $\Psi_* : H \to \mathbb{R}_+ \cup \{+\infty\}$ defined as

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

as well as their approximations Φ_n, Ψ_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)

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

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

and

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

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 ($\Phi_n \geq 0$ 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

$$\begin{aligned}
\mathbb{E}\, \Phi_n(X(s)) &= \mathbb{E}\big(\|X(s)\|_V^2 \wedge n^2\big) \leq \mathbb{E}\|X(s)\|_V^2,\\
\mathbb{E}\, \Psi_n(X(s)) &= \mathbb{E}\Big(n \wedge \int_D j(X(s))\Big) \leq \mathbb{E} \int_D j(X(s)),\\
\mathbb{E}\, \Psi_{*n}(X(s)) &= \mathbb{E}\Big(n \wedge \int_D j^*(\beta_{1/n}(X(s)))\Big) \leq \mathbb{E} \int_D j^*(\xi(s)),
\end{aligned}$$

where, in the last inequality, we have used the fact that for every $r \in \mathsf{D}(\beta) = \mathbb{R}$ 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 $|\beta^0(r)| \leq |y|$ for every $y \in \beta(r)$ (note that the uniqueness of $\beta^0(r)$ follows from the maximal monotonicity of β). Thanks to estimate (4.4.5) we have, by Tonelli's theorem,

$$
\begin{aligned}
C\int_0^1 &\big(\mathbb{E}\,\Phi_n(X(s)) + \mathbb{E}\,\Psi_n(X(s)) + \mathbb{E}\,\Psi_{*n}(X(s))\big)\,ds\\
&\leq C\,\mathbb{E}\int_0^1 \|X(s)\|_V^2\,ds + \mathbb{E}\int_0^1\int_D j(X(s))\,ds + \mathbb{E}\int_0^1\int_D j^*(\xi(s))\,ds\\
&\leq \frac{1}{2}\|x\|^2 + C_0,
\end{aligned}
$$

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

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

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

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

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 $\mu(J^*) = 1$. □

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

Lemma 4.4.4. *The sets*

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

belong to the Borel σ-algebra of H.

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

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

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 $J^* \in \mathscr{B}(H)$. □

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

Theorem 4.4.5. *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., [2, Thm. 19.25]). Let $\mathscr{I}$ be the set of all invariant measures for P: by Theorem 4.4.1, 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 4.4.3, 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 4.4.1, 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. □

Under a very mild growth condition on the drift one can also obtain uniqueness.

Theorem 4.4.6. *If β is superlinear, i.e. if there exists $c > 0$ and $\delta > 0$ such that*

$$(y_1 - y_2)(x_1 - x_2) \geq c|x_1 - x_2|^{2+\delta} \qquad \forall (x_i, y_i) \in \beta, \quad i = 1, 2,$$

then there exists a unique invariant measure μ for the transition semigroup P. Moreover, μ is strongly mixing.

Proof. For any $x, y \in H$, by Itô's formula, the monotonicity of A, the superlinearity of β, and Jensen inequality we have

$$\mathbb{E}\|X(t;0,x) - X(t;0,y)\|^2 + \tilde{c}\int_0^t \Bigl(\mathbb{E}\|X(t;0,x) - X(t;0,y)\|^2\Bigr)^{1+\frac{\delta}{2}} \leq \|x-y\|^2$$

for a positive constant $\tilde{c}$. Denoting by $y(\cdot;y_0)$ the solution to the Cauchy problem

$$y' + y^{1+\frac{\delta}{2}} = 0, \qquad y(0) = y_0 \geq 0,$$

one can easily check that

$$c(t) := \sup_{y_0 \geq 0} y(t;y_0) \to 0 \quad \text{as } t \to \infty$$

and that $c(t) \geq 0$ for every $t \geq 0$. We deduce that

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

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

$$\begin{aligned}\left|P_t\varphi(x) - \int_H \varphi(y)\,\mu(dy)\right|^2 &\leq \|D\varphi\|_\infty^2 \int_H \mathbb{E}\|X(t;0,x) - X(t;0,y)\|^2\,\mu(dy)\\ &\leq \|D\varphi\|_\infty^2 c(t)\end{aligned}$$

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

$$\left|P_t\varphi(x) - \int_H \varphi(y)\,\mu(dy)\right|^2 \to 0$$

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. □

4.5 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 (4.1.1), 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 (4.1.1), i.e.

$$[L_0 f](x) = -\frac{1}{2}\operatorname{Tr}\bigl(D^2 f(x) B(x) B^*(x)\bigr) + \langle Ax, Df(x)\rangle + \langle \beta(x), Df(x)\rangle,$$

for $x \in V \cap J^*$, 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

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

Here $f \in C^1_b(V')$ 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, $f \in C^1_b(L^1(D))$ 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 $|Df(x)k| \leq N\|k\|_{L^1(D)}$ and $x \mapsto Df(x) \in C(L^1(D), L^\infty(D))$. For any $f \in C^2_b(H)$, one has, recalling the linear growth condition on B,

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

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 $f \in C^1_b(V')$,

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

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

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

and recalling that $x \mapsto \|j^*(\beta(x))\|_{L^1(D)} \in L^1(H,\mu)$ by Theorem 4.4.3, it is enough to consider the second term on the right-hand side: for any $f \in C^1_b(L^1(D))$, $\sup_{x\in V}\|Df(x)\|_{L^\infty(D)}$ is finite, hence, recalling that $j \in C(\mathbb{R})$, we infer that $j(Df(x))$ is bounded pointwise in D, thus also in $L^1(D)$, uniformly over $x \in V$. In particular, we have that $x \mapsto \|j(Df(x))\|_{L^1(D)} \in L^1(H,\mu)$.

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 4.3.3, for every $g \in \mathsf{D}(L_0)$ we have

$$\begin{aligned}
g(X^x(t)) &+ \int_0^t \langle AX^x(s), Dg(X^x(s))\rangle\, ds + \int_0^t \langle \beta(X^x(s)), Dg(X^x(s))\rangle\, ds \\
&= g(x) + \frac{1}{2}\int_0^t \mathrm{Tr}[B^*(X^x(s))D^2g(X^x(s))B(X^x(s))]\, ds \\
&\quad + \int_0^t Dg(X^x(s))B(X^x(s))\, dW(s),
\end{aligned}$$

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

$$dX_{\lambda n} + AX_{\lambda n}\,dt + \beta_{\lambda n}(X_{\lambda n})\,dt = B(X_{\lambda n})\,dW(t), \qquad X_{\lambda n}(0) = x. \tag{4.5.6}$$

Since $\beta_{\lambda n}$ is Lipschitz-continuous, equation (4.5.6) admits a unique strong (variational) solution $X^x_{\lambda n} \in L^2(\Omega; E)$, where, as before, $E := C([0,T]; H) \cap L^2(0,T; V)$. Furthermore, the generator of the transition semigroup $P^{\lambda n} = (P^{\lambda n}_t)_{t\geq 0}$ on $B_b(H)$ defined by $P^{\lambda n}_t f(x) := \mathbb{E}\, f(X^x_{\lambda n}(t))$, restricted to $C^1_b(V') \cap C^2_b(H)$, is given by $-L^{\lambda n}_0$, where

$$[L^{\lambda n}_0 f](x) = -\frac{1}{2}\operatorname{Tr}\bigl(D^2 f(x) B(x) B^*(x)\bigr) + \langle Ax, Df(x)\rangle + \langle \beta_{\lambda n}(x), Df(x)\rangle, \quad x \in V.$$

This follows arguing as in the case of L_0 (even using the simpler Itô formula of Proposition 4.3.2, rather than the one of Proposition 4.3.3).

Let us now consider the stationary Kolmogorov equation

$$\alpha v + L^{\lambda n}_0 v = g, \qquad g \in \mathsf{D}(L_0), \quad \alpha > 0. \tag{4.5.7}$$

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

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

which is the natural candidate to solve (4.5.7). If we show that $v_{\lambda n} \in C^1_b(V') \cap C^2_b(H)$, then an application of Itô's formula (in the version of Proposition 4.3.2) shows that indeed $v_{\lambda n}$ solves (4.5.7). 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 equation (4.5.6). From now on we shall restrict our considerations to the case of additive noise, i.e. we assume that $B \in \mathscr{L}^2(U, H)$ is non-random. Moreover, we shall assume that V is continuously embedded in $L^4(D)$. The latter assumption is needed to apply the second-order differentiability results of §4.3.2. We recall that, thanks to Theorems 4.3.4 and 4.3.7, 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

$$Y'_h(t) + AY_h(t) + \beta'_{\lambda n}(X_{\lambda n}(t))Y_h(t) = 0, \qquad Y_h(0) = h. \tag{4.5.8}$$

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

$$Z'_{hk}(t) + AZ_{hk}(t) + \beta'_{\lambda n}(X_{\lambda n}(t))Z_{hk}(t) + \beta''_{\lambda n}(X_{\lambda n}(t))Y_h(t)Y_k(t) = 0, \quad Z_{hk}(0) = 0. \tag{4.5.9}$$

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 4.5.1. *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*

$$u' + Au = F(\cdot,u), \quad u(0) = u_0, \qquad \text{and} \qquad v' + A_2v = F(\cdot,v), \quad v(0) = u_0,$$

respectively, then $u = v$.

Proof. Let us first assume that $u' + Au = f$ and $v' + A_2v = f$, 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)

$$v_\varepsilon' + A_2v_\varepsilon = f_\varepsilon, \qquad v_\varepsilon(0) = u_0^\varepsilon$$

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

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

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 4.3.5 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

$$u' + Au = F(\cdot,u), \qquad v' + A_2v = F(\cdot,v), \qquad u(0) = v(0) = u_0,$$

respectively. Setting $f := F(\cdot,v)$, thanks to the assumptions on F we have that $f \in L^2(0,T;H)$, hence v is a mild solution to $v' + A_2v = f$, $v(0) = u_0$. 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 $u = v$. □

The following estimates are crucial.

Proposition 4.5.2. *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 $H = \mathsf{D}((\eta I + A)^\delta)$. 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 (4.5.8) 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 (4.5.8) and integrating in time yields

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

for all $t \in [0,T]$, and the first estimate is thus proved. The second estimate follows directly from Proposition 4.3.7. 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 $t > 0$,

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

Since, as $k \to \infty$, $\sigma_k(Y_{h\varepsilon}^x)$ converges a.e. to a measurable function $w_\varepsilon \in \operatorname{sgn}(Y_{h\varepsilon}^x)$ 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 (4.5.8), by Lemma 4.5.1 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 $t \in [0,1]$,

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

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 4.5.3. *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 (4.5.7). Moreover, there exists a positive constant M, independent of λ and n, such that

$$\|v_{\lambda n}\|_{C_b^1(H)\cap C_b^1(L^1(D))} \leq M \tag{4.5.10}$$

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 4.5.2,

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

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

$$Dv_{\lambda n}(x)h = \mathbb{E}\int_0^{+\infty} e^{-\alpha t} Dg(X_{\lambda n}^x(t))Y_h^x(t)\,dt. \tag{4.5.11}$$

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 4.5.2, we have, for every $k \in H$,

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

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

$$D^2v_{\lambda n}(x)(h,k) = \mathbb{E}\int_0^{+\infty} e^{-\alpha t}\left(D^2g(X_{\lambda n}^x(t))Y_h^x(t)Y_k^x(t) + Dg(X_{\lambda n}^x(t))Z_{hk}^x(t)\right)dt. \tag{4.5.12}$$

Moreover, using the third estimate of Proposition 4.5.2 and the fact that $g \in C_b^1(L^1(D))$, it follows by Hölder's inequality and (4.5.11) that

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

which implies that $v_{\lambda n} \in C_b^1(L^1(D))$ and the estimate (4.5.10). Finally, by the last estimate of Proposition 4.5.2 and the fact that $g \in C_b^1(V')$, we have

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

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 (4.5.7). Indeed, since $g \in C_b^2(H)\cap C_b^1(V')$, by Itô's formula in the version of Proposition 4.3.2 we get

$$
\begin{aligned}
g(X_{\lambda n}^x(t)) &+ \int_0^t \langle AX_{\lambda n}^x(s), Dg(X_{\lambda n}^x(s))\rangle\,ds + \int_0^t \langle \beta_{\lambda n}(X_{\lambda n}^x(s)), Dg(X_{\lambda n}^x(s))\rangle\,ds\\
&= g(x) + \frac{1}{2}\int_0^t \operatorname{Tr}[B^*(X_{\lambda n}^x(s))D^2g(X_{\lambda n}^x(s))B(X_{\lambda n}^x(s))]\,ds\\
&\quad + \int_0^t Dg(X_{\lambda n}^x(s))B(X_{\lambda n}^x(s))\,dW(s)
\end{aligned}
$$

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 4.5.4. *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

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

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

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

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

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

Since $Dv_{\lambda n}(x)$ is bounded in $L^\infty(D)$ uniformly over λ, n and x by estimate (4.5.10), one has

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

hence

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

in $L^0(D)$ as $\lambda\to 0$ for every $x\in V$. Recalling the definition of η in §4.2, 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 $|\beta_\lambda|\le|\beta|$,

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

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

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

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

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

where the right-hand side belongs to $L^1(H,\mu)$ by Theorem 4.4.3. A further application of the dominated convergence theorem thus yields

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

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 equation (4.1.1), 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 §4.2, 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 4.5.5. *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 $\lambda \in (0,1)$,

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

if follows by Lemma 4.5.3 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 4.5.4, there exist $\lambda_0 > 0$ and $n_0 \in \mathbb{N}$ such that

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

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

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

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, [34, Theorem 1.12]). □

Chapter 5

Singular semilinear equations: regularity

In this chapter, we prove a regularity result for the equation

$$dX(t) + AX(t)\,dt + \beta(X(t))\,dt \ni B(t,X(t))\,dW(t)\,, \qquad X(0) = X_0. \tag{5.0.1}$$

In particular, we show how the smoothness of the solution improves (as well as of invariant measures, if they exist) if the initial datum and the diffusion coefficient are smoother, without any further regularity assumption on the (possibly singular) monotone drift term β. For example, if A (better said, the part of A in H) is self-adjoint, the solution has paths belonging to the domain of A in H if X_0 and B, roughly speaking, take values in the domain of $A^{1/2}$. This implies that X is a strong solution in the classical sense, not just in the variational one. The results of this chapter are part of the joint work [63] with Carlo Marinelli.

We are going to show that the regularity of the solution to equation (5.0.1) improves, if the initial datum and the diffusion coefficient are smoother, irrespective of the (possible) singularity of the drift coefficient β. In particular, we provide sufficient conditions implying that the variational solution to (5.0.1) is also an analytically strong solution, in the sense that it takes values in the domain of the part of A in H (see §3.3). If the solution to (5.0.1) generates a Markovian semigroup on $C_b(H)$ admitting an invariant measure, we also show that improved regularity of the solution carries over to further regularity of the invariant measure, in the sense that its support is made of smoother functions.

Theorem 5.0.1. *Assume that the hypotheses of §3.2 are satisfied, that A is symmetric and that*

$$X_0 \in L^2(\Omega,\mathscr{F}_0,\mathbb{P};V), \qquad B(\cdot,X) \in L^2(\Omega;L^2(0,T;\mathscr{L}^2(U,V))). \tag{5.0.2}$$

Then the unique solution (X,ξ) to the equation (5.0.1) satisfies

$$X \in L^2(\Omega;C([0,T];H)) \cap L^2(\Omega;L^\infty(0,T;V)) \cap L^2(\Omega;L^2(0,T;\mathsf{D}(A_2))).$$

For the proof we need the following positivity result.

Lemma 5.0.2. *Let A_λ and β_λ be the Yosida approximations of A_2 and β, respectively. One has*

$$\big\langle A_\lambda u, \beta_\lambda(u)\big\rangle \geq 0 \qquad \forall u \in H.$$

Proof. Let $j_\lambda : \mathbb{R} \to \mathbb{R}$ be the positive convex function defined as $j_\lambda(x) := \int_0^x \beta_\lambda(y)\,dy$. Then, for any u, $v \in L^2(D)$,

$$j_\lambda(v) - j_\lambda(u) \geq j'_\lambda(u)(v-u)$$

a.e. in D, hence, integrating over D,

$$\int_D j_\lambda(v) - \int j_\lambda(u) \geq \big\langle \beta_\lambda(u), v-u \big\rangle.$$

Choosing $v = (I+\lambda A_2)^{-1}u$, one has $u - v = \lambda A_\lambda u$, thus also

$$\lambda \big\langle \beta_\lambda(u), A_\lambda u \big\rangle \geq \int j_\lambda(u) - \int j_\lambda((I+\lambda A_2)^{-1}u).$$

Since A_1 is an extension of A_2 and $u \in L^1(D)$, Jensen's inequality for sub-Markovian operators and accretivity of A_1 in $L^1(D)$ imply

$$\int j_\lambda((I+\lambda A_2)^{-1}u) \leq \int (I+\lambda A_2)^{-1} j_\lambda(u) \leq \int j_\lambda(u). \qquad \square$$

Proof of Theorem 5.0.1. For any $\lambda > 0$, let J_λ and A_λ be the resolvent and the Yosida approximations of A_2, the part of A in H, as defined in §3.3. That is,

$$J_\lambda := (I+\lambda A_2)^{-1}, \qquad A_\lambda := \frac{1}{\lambda}(I - J_\lambda).$$

We recall that J_λ and A_λ are bounded linear operators on H, that J_λ is a contraction, and that $A_\lambda = AJ_\lambda$.

Setting $G := B(\cdot, X)$, let us consider the equation

$$dX_\lambda(t) + A_\lambda X_\lambda(t)\,dt + \beta_\lambda(X_\lambda(t))\,dt = G(t)\,dW(t), \qquad X_\lambda(0) = X_0.$$

Since A_λ is bounded and β_λ is Lipschitz-continuous, it admits a unique strong solution

$$X_\lambda \in L^2(\Omega; C([0,T]; H)),$$

for which Itô's formula for the square of the H-norm yields

$$\begin{aligned}
\frac{1}{2}\|X_\lambda(t)\|^2 + \int_0^t \big\langle A_\lambda X_\lambda(s), X_\lambda(s) \big\rangle\,ds + \int_0^t\!\!\int_D \beta_\lambda(X_\lambda(s))X_\lambda(s)\,ds \\
= \frac{1}{2}\|X_0\|^2 + \frac{1}{2}\int_0^t \|G(s)\|^2_{\mathscr{L}^2(U,H)}\,ds + \int_0^t X_\lambda(s)G(s)\,dW(s)
\end{aligned}$$

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

$$X_\lambda = J_\lambda X_\lambda + X_\lambda - J_\lambda X_\lambda = J_\lambda X_\lambda + \lambda A_\lambda X_\lambda$$

and recalling that $A_\lambda = AJ_\lambda$ and that A is coercive, we have, after taking supremum in time and expectation,

$$\begin{aligned}
&\frac{1}{2}\,\mathbb{E}\big\|X_\lambda\big\|^2_{C([0,T];H)} + C\,\mathbb{E}\int_0^T \big\|J_\lambda X_\lambda(s)\big\|^2_V\,ds \\
&\qquad + \lambda\,\mathbb{E}\int_0^T \|A_\lambda X_\lambda(s)\|^2\,ds + \mathbb{E}\int_0^T\!\!\int_D \beta_\lambda(X_\lambda(s))X_\lambda(s)\,ds \\
&\qquad \le \frac{1}{2}\,\mathbb{E}\|X_0\|^2 + \frac{1}{2}\,\mathbb{E}\int_0^T \big\|G(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds + \mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t X_\lambda(s)G(s)\,dW(s)\right|,
\end{aligned}$$

where, by Lemma 3.4.1, the last term on the right-hand side is bounded by

$$\varepsilon\,\mathbb{E}\|X_\lambda\|^2_{C([0,T];H)} + C_\varepsilon\,\mathbb{E}\int_0^T \big\|G(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds,$$

so that, rearranging terms and choosing ε small enough, we deduce that there exists a constant $N > 0$ independent of λ such that

$$\big\|X_\lambda\big\|^2_{L^2(\Omega;C([0,T];H))} + \big\|J_\lambda X_\lambda\big\|^2_{L^2(\Omega;L^2(0,T;V))} + \big\|\beta_\lambda(X_\lambda)X_\lambda\big\|_{L^1(\Omega\times(0,T)\times D)} < N. \tag{5.0.3}$$

Moreover, let us introduce the function

$$\begin{aligned}
\varphi_\lambda : H &\longrightarrow [0,+\infty[,\\
u &\longmapsto \frac{1}{2}\langle A_\lambda u, u\rangle.
\end{aligned}$$

The linearity and the boundedness of A_λ immediately implies that $\varphi_\lambda \in C^2(H)$ with $D\varphi_\lambda(u) = A_\lambda$, and, by linearity, $D^2\varphi_\lambda(u) = A_\lambda$, for all $u \in H$ (in the latter statement A_λ has to be considered as an element of $\mathscr{L}_2(H)$, the space of bounded bilinear maps on H). Itô's formula applied to $\varphi_\lambda(X_\lambda)$ then yields

$$\begin{aligned}
&\varphi_\lambda(X_\lambda(t)) + \int_0^t \big\|A_\lambda X_\lambda(s)\big\|^2\,ds + \int_0^t \big\langle A_\lambda X_\lambda(s), \beta_\lambda(X_\lambda(s))\big\rangle\,ds \\
&\qquad = \varphi_\lambda(X_0) + \frac{1}{2}\int_0^t \operatorname{Tr}\big(G^*(s)D^2\varphi_\lambda(X_\lambda(s))G(s)\big)\,ds + \int_0^t A_\lambda X_\lambda(s)G(s)\,dW(s)
\end{aligned}$$

for every $t \in [0,T]$ $\mathbb{P}$-almost surely. Writing, as before, $X_\lambda = J_\lambda X_\lambda + \lambda A_\lambda X_\lambda$, the coercivity of A implies that

$$\varphi_\lambda(X_\lambda) = \frac{1}{2}\big\langle A_\lambda X_\lambda, X_\lambda\big\rangle \ge \frac{C}{2}\big\|J_\lambda X_\lambda\big\|^2_V + \frac{1}{2}\lambda\big\|A_\lambda X_\lambda\big\|^2 \gtrsim \big\|J_\lambda X_\lambda\big\|^2_V.$$

The continuity of J_λ on V (see Lemma 3.3.1) instead implies

$$\varphi_\lambda(X_0) = \big\langle AJ_\lambda X_0, X_0\big\rangle \le \big\|A\big\|_{\mathscr{L}(V,V')}\big\|J_\lambda X_0\big\|_V\big\|X_0\big\|_V \lesssim \big\|A\big\|_{\mathscr{L}(V,V')}\big\|X_0\big\|^2_V.$$

Denoting a complete orthonormal basis of U by $(u_k)_k$, we have, recalling again the continuity of J_λ on V and that $D^2\varphi_\lambda(u) = A_\lambda$ for all $u \in H$,

$$\begin{aligned}\operatorname{Tr}\bigl(G^*D^2\varphi_\lambda(X_\lambda)G\bigr) &= \sum_{k=0}^{\infty}\bigl\langle G^*D^2\varphi_\lambda(X_\lambda)Gu_k, u_k\bigr\rangle_U = \sum_{k=0}^{\infty}\bigl\langle A_\lambda Gu_k, Gu_k\bigr\rangle \\ &= \sum_{k=0}^{\infty}\bigl\langle AJ_\lambda Gu_k, Gu_k\bigr\rangle \le \|A\|_{\mathscr{L}(V,V')}\sum_{k=0}^{\infty}\|J_\lambda Gu_k\|_V\|Gu_k\|_V \\ &\lesssim \|A\|_{\mathscr{L}(V,V')}\sum_{k=0}^{\infty}\|Gu_k\|_V^2 = \|A\|_{\mathscr{L}(V,V')}\|G\|^2_{\mathscr{L}^2(U,V)}.\end{aligned}$$

Moreover, by Lemma 3.4.1,

$$\begin{aligned}\mathbb{E}\bigl((A_\lambda X_\lambda G)\cdot W\bigr)^*_T &\lesssim \varepsilon\,\mathbb{E}\sup_{t\in[0,T]}\|A_\lambda X_\lambda(t)\|^2_{V'} + N(\varepsilon)\,\mathbb{E}\int_0^T\|G(s)\|^2_{\mathscr{L}^2(U,V)}\,ds \\ &\le \varepsilon\|A\|^2_{\mathscr{L}(V,V')}\,\mathbb{E}\sup_{t\in[0,T]}\|J_\lambda X_\lambda(t)\|^2_V + N(\varepsilon)\,\mathbb{E}\int_0^T\|G(s)\|^2_{\mathscr{L}^2(U,V)}\,ds\end{aligned}$$

for every $\varepsilon > 0$. Taking supremum in time and expectations in the Itô formula for $\varphi_\lambda(X_\lambda)$, choosing ε small enough we obtain, thanks to the previous lemma and hypothesis (5.0.2), that there exists a constant $N > 0$ independent of λ, such that

$$\|J_\lambda X_\lambda\|^2_{L^2(\Omega;L^\infty(0,T;V))} + \|A_\lambda X_\lambda\|^2_{L^2(\Omega;L^2(0,T;H))} < N. \tag{5.0.4}$$

Reasoning as in Chapter 2, it follows by (5.0.3) that

$$\begin{aligned} X_\lambda &\longrightarrow X \quad \text{weakly in } L^2(\Omega;L^2(0,T;H)),\\ J_\lambda X_\lambda &\longrightarrow X \quad \text{weakly in } L^2(\Omega;L^2(0,T;V)),\\ \beta_\lambda(X_\lambda) &\longrightarrow \xi \quad\ \text{weakly in } L^1(\Omega\times(0,T)\times D),\end{aligned}$$

where (X,ξ) is the unique solution to (3.1.1). Moreover, by (5.0.4) we have

$$J_\lambda X_\lambda - X_\lambda = \lambda A_\lambda X_\lambda \longrightarrow 0 \qquad \text{in } L^2(\Omega;L^2(0,T;H)),$$

hence, $\mathbb{P}$-almost surely and for almost every $t \in (0,T)$, $J_\lambda X_\lambda(t)$ converges to $X(t)$ in H. Since the function $\|\cdot\|_V^2$ is lower semicontinuous on H, we infer that

$$\|X(t)\|_V^2 \le \liminf_{\lambda\to 0}\|J_\lambda X_\lambda(t)\|_V^2$$

for almost every t. Hence, taking supremum in time and expectation, we deduce that

$$X \in L^2(\Omega;L^\infty(0,T;V)).$$

Moreover, by (5.0.4) we also have

$$A_\lambda X_\lambda \longrightarrow \eta \quad \text{weakly in } L^2(\Omega;L^2(0,T;H)),$$

hence, since $J_\lambda X_\lambda \to X$ weakly in $L^2(\Omega;L^2(0,T;V))$, by the continuity and the linearity of A we necessarily have $\eta = AX$, hence $X \in L^2(\Omega;L^2(0,T;\mathsf{D}(A_2)))$. □

As last result we show that if the solution to (5.0.1) generates a Markovian semigroup $P = (P_t)_{t\ge0}$ on $C_b(H)$ admitting an invariant measure, then the improved regularity of solutions given by Theorem 5.0.1 implies better integrability properties also for the invariant measures. Existence and uniqueness of invariant measures, ergodicity, and the Kolmogorov equation associated to (5.0.1) were studied in Chapter 4. In particular, we recall the following result. The set of invariant measures of P will be denoted by $\mathscr{M}$.

Proposition 5.0.3. *Assume that the hypotheses of §3.2 are satisfied, that $X_0 \in H$ is non-random, and that B is non-random and time-independent. Then the solution X to* (5.0.1) *is Markovian and its associated transition semigroup P admits an ergodic invariant measure. Moreover, there exists a positive constant N such that*

$$\int_H \|u\|_V^2 \,\mu(du) + \int_H\int_D j(u)\,\mu(du) + \int_H\int_D j^*(\beta^0(u))\,\mu(du) < N \qquad \forall \mu \in \mathscr{M}.$$

If β is superlinear, there exists a unique invariant measure for P, which is strongly mixing as well.

Theorem 5.0.4. *Assume that the hypotheses of §3.2 are satisfied and that*

$$B : H \to \mathscr{L}^2(U,V), \qquad \|B(v)\|_{\mathscr{L}^2(U,V)} \lesssim 1 + \|v\|_V.$$

If A is symmetric, then there exists a positive constant N such that

$$\int_H \|Au\|^2 \,\mu(du) < N \qquad \forall \mu \in \mathscr{M}.$$

In particular, every $\mu \in \mathscr{M}$ is concentrated on $\mathsf{D}(A_2)$, i.e. $\mu(\mathsf{D}(A_2)) = 1$.

Proof. For every $x \in V$, let (X^x, ξ^x) be the unique strong solution to (3.1.1) with initial datum x. Setting $G := B(X^x)$, Itô's formula for $\varphi_\lambda(X_\lambda)$ as in the proof of the previous theorem yields

$$\begin{aligned}\varphi_\lambda(X_\lambda(t)) &+ \int_0^t \big\|A_\lambda X_\lambda(s)\big\|^2\,ds + \int_0^t \big\langle A_\lambda X_\lambda(s), \beta_\lambda(X_\lambda(s))\big\rangle\,ds \\ &= \varphi_\lambda(x) + \frac{1}{2}\int_0^t \operatorname{Tr}\big(G^*(s)D^2\varphi_\lambda(X_\lambda(s))G(s)\big)\,ds + \int_0^t A_\lambda X_\lambda(s)G(s)\,dW(s).\end{aligned}$$

Since $A_\lambda X_\lambda \in L^2(\Omega; L^\infty(0,T;H))$ and $G \in L^2(\Omega; L^2(0,T;\mathscr{L}^2(U,H)))$, the last term on the right hand side is a martingale; hence, taking expectations and recalling that $\varphi_\lambda(x) \lesssim \|x\|_V^2$, it follows by Lemma 5.0.2 and by the estimates obtained in the proof of the previous theorem that

$$\mathbb{E}\int_0^t \big\|A_\lambda X_\lambda(s)\big\|^2\,ds \lesssim \|x\|_V^2 + \mathbb{E}\|G\|^2_{L^2(0,t;\mathscr{L}^2(U,V))}.$$

Since this holds for every $\lambda > 0$, letting $\lambda \to 0$ and recalling that, as in the proof of the previous theorem, $A_\lambda X_\lambda$ converges to AX weakly in $L^2(\Omega; L^2(0,T;H))$, a weak lower semicontinuity argument and the linear growth assumption on B yield

$$\mathbb{E}\int_0^t \big\|AX^x(s)\big\|^2\,ds \lesssim 1 + \|x\|_V^2 \tag{5.0.5}$$

for every $t \in [0,T]$ and $x \in V$. Let us introduce the function $F : H \to [0,+\infty]$ defined as

$$F(u) := \begin{cases} \|Au\|^2 & \text{if } u \in \mathsf{D}(A_2), \\ +\infty & \text{if } u \in H \setminus \mathsf{D}(A_2), \end{cases}$$

and the sequence of functions $(F_n)_{n\in\mathbb{N}}$, $F_n : H \to [0,+\infty)$, defined as

$$F_n(u) := \left\|A_{1/n}u\right\|^2 \wedge n^2.$$

It is easily seen that $F_n \in C_b(H)$ for all $n \in \mathbb{N}$ and that F_n converges pointwise to F from below. Therefore, for any invariant measure μ, it follows by Fubini's theorem that

$$\begin{aligned}
\int_H F_n(x)\,\mu(dx) = \int_0^1\int_H F_n(x)\,\mu(dx)\,ds &= \int_0^1\int_H P_sF_n(x)\,\mu(dx)\,ds \\
&= \int_H\int_0^1 \mathbb{E}\big(\left\|A_{1/n}X^x(s)\right\|^2 \wedge n^2\big)\,ds\,\mu(dx) \\
&\le \int_H \mathbb{E}\int_0^1 \left\|A_{1/n}X^x(s)\right\|^2\,ds\,\mu(dx)
\end{aligned}$$

Recalling that $\|A_\lambda u\| \le \|Au\|$ for all $u \in H$, it follows by (5.0.5) that

$$\int_H F_n(x)\,\mu(dx) \lesssim 1 + \int_H \|x\|_V^2\,\mu(dx).$$

Since $\|\cdot\|_V^2 \in L^1(H,\mu)$ by Theorem 4.4.3, we get

$$\int_H F_n(x)\,\mu(dx) \lesssim N$$

for a positive constant N, independent of n and μ. Letting $n \to \infty$, by the monotone convergence theorem we deduce that $F \in L^1(H,\mu)$, hence F is finite μ-almost everywhere in H, and in particular $\mu(\mathsf{D}(A_2)) = 1$. $\square$

Chapter 6

Divergence-type equations with singular reaction term

We prove well-posedness for doubly nonlinear parabolic stochastic partial differential equations of the form

$$dX_t - \operatorname{div} \gamma(\nabla X_t)\, dt + \beta(X_t)\, dt \ni B(t, X_t)\, dW_t,$$

where γ and β are the two nonlinearities, assumed to be multivalued maximal monotone operators everywhere defined on $\mathbb{R}^d$ and $\mathbb{R}$ respectively, and W is a cylindrical Wiener process. Using variational techniques, uniform estimates (both pathwise and in expectation) and compactness results, well-posedness is proved under the classical Leray-Lions conditions on γ and with no restrictive smoothness or growth assumptions on β. The operator B is assumed to be Hilbert-Schmidt and to satisfy some classical Lipschitz conditions in the second variable.

The results presented in this chapter are part of the work [75], recently published in *Journal of Differential Equations*.

6.1 The problem: literature and main goals

In this chapter, we consider the boundary value problem with homogeneous Dirichlet conditions associated to a doubly nonlinear parabolic stochastic partial differential equation on an smooth bounded domain $D \subseteq \mathbb{R}^d$ of the type

$$dX_t - \operatorname{div} \gamma(\nabla X_t)\, dt + \beta(X_t)\, dt \ni B(t, X_t)\, dW_t \quad \text{in } D \times (0,T)\,, \tag{6.1.1}$$

$$X(0) = X_0 \quad \text{in } D\,, \tag{6.1.2}$$

$$X = 0 \quad \text{on } \partial D \times (0,T)\,, \tag{6.1.3}$$

where γ and β are two maximal monotone operators everywhere defined on $\mathbb{R}^d$ and $\mathbb{R}$, respectively, W is a cylindrical Wiener process, and B is a random time-dependent Hilbert-Schmidt operator (we will state the complete assumptions on the data in the next section). We prove existence of global solutions as well as a continuous dependence result using variational techniques (see e.g. the classical works [46, 71, 72] about the variational approach to SPDEs).

The problem (6.1.1)–(6.1.3) is very interesting from the mathematical point of view: as a matter of fact, the equation presents two strong nonlinearities. The first one is represented by γ within the divergence operator: in this case, we will need to assume some classical growth assumptions (the so-called Leray-Lions conditions) in order to recover a suitable coercivity on a natural Sobolev space. The other nonlinearity is represented by the operator β: this is treated as generally as possible, with no restriction on the growth and regularity. Because of this generality, the concept of solution and the appropriate estimates are more difficult to achieve, as we will see. We point also out that dealing with maximal monotone graphs makes our analysis absolutely exhaustive. As a matter of fact, in this way we include in our treatment any continuous increasing function β (with any order of growth), as well as every increasing function with a countable number of jumps: indeed, it is a standard matter to see that if β is an increasing function on $\mathbb{R}$ with jumps in $\{x_n\}_{n\in\mathbb{N}}$, one can obtain a maximal monotone graph by setting $\beta(x_n) = [\beta_-(x_n), \beta_+(x_n)]$. Finally, very mild assumptions on the noise are required, so that our results fit to any reasonable random time-dependent Hilbert-Schmidt operator B; in the case of multiplicative noise, only classical Lipschitz continuity hypotheses are in order.

The noteworthy feature of the results contained in this chapter is that problem (6.1.1)–(6.1.3) is very general and embraces a wide variety of specific sub-problems which are interesting on their own: consequently, we provide with our treatment a unifying analysis to several cases of parabolic SPDEs. Let us mention now some of these and the main related literature.

If γ is the identity on $\mathbb{R}^d$, the resulting equation is the classical semilinear SPDE driven by the Laplace operator $dX - \Delta X\,dt + \beta(X)\,dt \ni B\,dW_t$, which has been widely studied. For example, in Chapter 2, global existence results of solutions are provided in the semilinear case, with the laplacian being generalized to any suitable linear operator: here, the idea is to doubly approximate the problem, in order to recover more regularity on β and B, to find then appropriate estimates on the approximated solutions and finally to pass to the limit in the equation. Moreover, in [28], mild solutions are obtained under the strong hypotheses that β is a polynomial of odd degree $m > 1$ and B can be written as $(-\Delta)^{-\frac{s}{2}}$ for a suitable s; in [9], existence of mild solutions is proved with no restrictive hypotheses on the growth of β, but imposing some strong continuity assumptions on the stochastic convolution. In [58], well-posedness is established for the semilinear problem in a L^q setting, with β having polynomial growth.

If γ is the monotone function on $\mathbb{R}^d$ given by $\gamma(x) = |x|^{p-2}x$, $x \in \mathbb{R}^d$, for a certain $p \geq 2$, then the term represented by the divergence in (6.1.1) is the usual p-Laplacian: in this case, our equation becomes $dX - \Delta_p X\,dt + \beta(X)\,dt \ni B\,dW_t$, where $\Delta_p\cdot := \operatorname{div}(|\nabla\cdot|^{p-2}\nabla\cdot)$. This problem is far more interesting and complex than the semilinear case since $-\Delta_p$ is nonlinear for any $p > 2$ and consequently (6.1.1) becomes doubly nonlinear in turn. Among the extensive literature dealing with this problem, we can mention [54] for example, where the stochastic p-Laplace equation is studied in the singular case $p \in [1, 2)$, and [55] as well.

Let us now briefly outline the structure of the chapter.

In section 6.2, we state the precise assumptions of the work and we accurately describe the general setting: here, the main hypotheses are stated and the variational setting is presented. Furthermore, we outline the four main results: the first theorem ensures that problem (6.1.1)–(6.1.3) admits global solutions in a suitable weak variational way in the case of additive noise, the second one is the very natural continuous dependence property with respect to the initial datum and B, the third is the existence result in case of multiplicative noise and the last one states the continuous dependence property with respect to the initial datum in case of multiplicative noise.

Section 6.3 contains the proof of the existence theorem with additive noise: the main idea is to introduce two approximations on the problem. The first approximation depends on a parameter λ and it is made on the maximal monotone operators β and γ, considering the Yosida approximations, as usual; moreover, a correction term is added in order to recover a suitable coercivity when λ is fixed, and that is going to vanish when taking the limit as $\lambda \searrow 0$. The second approximation depends on a parameter ε and is made on the operator B in order to gain more regularity on the noise. The double approximation is very similar to the one performed in Chapter 2. The general idea is that given a fixed approximation in ε, the approximated noise is regular enough to allow us to pass to the limit pathwise in λ: once this first step is carried out, suitable probability estimates allow us to pass to the limit also in ε. More specifically, the proof of existence consists in obtaining uniform estimates on the approximated solutions, independently of the approximations, and then passing to the limit in the approximated problem. To this purpose, we will recover pathwise estimates which are uniform in λ (but not in ε), and global estimates also in expectation which are uniform both in λ and in ε. The passage to the limit is carried out in two steps: the first is on λ and it is made pathwise, while the second is made on ε and is made globally also in probability. The main idea is to use Itô's formula and some sharp testings to obtain L^1 estimates on the nonlinear terms in β and rely on the Dunford-Pettis theorem to recover a weak compactness, being inspired in this sense by some calculations performed in [9].

Section 6.4 is devoted to proving the continuous dependence result for the additive noise case, which easily follows from the definition of solution itself and a generalized Itô formula, which is accurately proved in the Section 6.7.

Section 6.5 contains the proof of the main result, which ensures that the problem with multiplicative noise is well-posed: here, we build the global solutions step-by-step, proving at each iteration accurate contraction estimates and using classical fixed-points arguments. The continuous dependence follows from the generalized version of Itô's formula contained in Section 6.7.

The Sections 6.6 and 6.7 contain a version of a variational integration-by-parts formula and the generalized Itô formula, which are widely used throughout: the first one is made pathwise and it is used when passing to the limit on λ in order to identify the limit of the nonlinearity in γ, while the second is a direct generalization of the classical Itô formula in a variational setting, and it is needed in the passage to the limit on ε and in the proof of the continuous dependence. The idea of the proof is to identify accurate approximations on the processes which have to satisfy appropriate conditions, such as linearity, smoothness properties and suitable asymptotical behaviours: in this sense, appropriate elliptic approximations are performed.

6.2 Setting and main results

In this section we state the precise assumptions on the data of the problem and the concept of solution. Moreover, we present the main results which will be proved in the subsequent sections.

In the sequel, $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space, where the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is assumed to satisfy the so-called "usual conditions" (i.e. it is saturated and right continuous) and $T > 0$ is the fixed final time; moreover, $D \subseteq \mathbb{R}^d$ is a smooth bounded domain and $Q := D\times(0, T)$ is the corresponding space-time cylinder. Furthermore, we set

$$H := L^2(D)$$

and we use the symbol $(\cdot, \cdot)$ for the standard inner product of H. We write "$\cdot$" for the usual scalar product in $\mathbb{R}^d$.

We can now specify the main hypotheses of this chapter. First of all, we introduce

$$
\begin{gathered}
\gamma : \mathbb{R}^d \to 2^{\mathbb{R}^d} \quad \text{maximal monotone}\,, \quad D(\gamma) = \mathbb{R}^d\,, \quad 0 \in \gamma(0) \\
\beta : \mathbb{R} \to 2^{\mathbb{R}} \quad \text{maximal monotone}\,, \quad D(\beta) = \mathbb{R}\,, \quad 0 \in \beta(0) \\
W \quad \text{cylindrical Wiener process on } U\,,
\end{gathered}
$$

where U is a suitable separable Hilbert space. Now, it is clear that the function

$$
j : \mathbb{R} \to [0, +\infty) \quad \text{proper, convex, lower semicontinuous}\,, \quad \partial j = \beta\,, \quad j(0) = 0
$$

is well defined; furthermore, we make the assumption that also γ is a subdifferential, i.e. that there exists

$$
k : \mathbb{R}^d \to [0, +\infty) \quad \text{proper, convex, lower semicontinuous}\,, \quad \partial k = \gamma\,, \quad k(0) = 0\,.
$$

We denote by k^* and j^* the convex conjugate functions of k and j, respectively, we also assume that j is even, i.e.

$$
j(x) = j(-x) \quad \text{for every } x \in \mathbb{R}.
$$

Remark 6.2.1. The symmetry hypothesis on j is needed in order to prove the generalized Itô formula for the solutions of our problem, which will be strongly used throughout the proofs. However, this can be weakened: the main point is that we only need j to grow at the same rate both at $+\infty$ and at $-\infty$ (cf. [12, p. 429]), i.e.

$$
\limsup_{|x| \to +\infty} \frac{j(x)}{j(-x)} < +\infty\,.
$$

Now, for every $\delta \in (0, 1)$, we introduce the resolvents and the Yosida approximations of γ and β as

$$
\begin{aligned}
J_\delta &:= (I_d + \delta\gamma)^{-1}\,, \qquad & R_\delta &:= (I_1 + \delta\beta)^{-1}\,, \\
\gamma_\delta &:= \frac{I_d - J_\delta}{\delta}\,, \qquad & \beta_\delta &:= \frac{I_1 - R_\delta}{\delta}\,,
\end{aligned}
$$

where the symbol I_m stands for the identity in $\mathbb{R}^m$ for any $m \in \mathbb{N}$.

As we have anticipated, we need to make some assumptions on the growth of γ, namely the so-called Leray-Lions conditions, which are widely required in the classical literature on elliptic and parabolic PDEs (the reader can refer here to [16–18] for classical examples). More in detail, we suppose that there are positive constants K, D_1, D_2 and an exponent $p \in [2, +\infty)$ such that

$$
\begin{gathered}
\sup\{|y| : y \in \gamma(r)\} \leq D_1 \left(1 + |r|^{p-1}\right) \quad \text{for every } r \in \mathbb{R}^d\,, \\
y \cdot r \geq K|r|^p - D_2 \quad \text{for every } r \in \mathbb{R}^d\,,\ y \in \gamma(r)\,.
\end{gathered}
$$

In the sequel, we will write $q := \frac{p}{p-1} \in (1, 2]$ for the conjugate exponent of p.

Finally, we set

$$
V := W_0^{1,p}(D)
$$

and define the divergence operator in the variational sense:

$$
-\operatorname{div} : L^q(D)^d \to V^*\,, \quad \langle -\operatorname{div} u, v\rangle := \int_D u \cdot \nabla v\,, \quad u \in L^q(D)^d\,,\ v \in V\,,
$$

where we have used the symbol $\langle\cdot,\cdot\rangle$ for the duality pairing between V and V^*. Here and in the sequel, we make the natural identification $H \cong H^*$, so that H is continuously embedded in V^*: for every $u \in H$ and $v \in V$, we have $\langle u, v\rangle = (u, v)$. Taking these remarks into account, we have

$$V \overset{c}{\hookrightarrow} H \hookrightarrow V^*\,,$$

where the first inclusion is also dense. Moreover, we set

$$V_0 := H^k_0(D)\,, \quad k := \left[\max\left\{\frac{d}{2}, 1 + \frac{d}{2} - \frac{d}{p}\right\}\right] + 1\ :$$

note that with this particular choice of k, the classical results on Sobolev embeddings (see [10, Thm. 1.5] and [43, Thm. 219]) ensure that

$$V_0 \hookrightarrow V \quad \text{densely}\,, \quad V_0 \hookrightarrow L^\infty(D)\,,$$

so that we have

$$V_0 \hookrightarrow V \cap L^\infty(D)\,, \qquad V^*, L^1(D) \hookrightarrow V_0^*\,.$$

We can now state the four main results of the chapter, which ensure that problem (6.1.1)–(6.1.2) is well-posed, both with additive and multiplicative noise.

Theorem 6.2.2. *Assume that*

$$X_0 \in L^2\left(\Omega, \mathcal{F}_0, \mathbb{P}; H\right)\,, \tag{6.2.4}$$

$$B \in L^2\big(\Omega \times (0,T); \mathscr{L}^2(U,H)\big) \quad \textit{progressively measurable}\,, \tag{6.2.5}$$

$$\gamma \quad \textit{is single-valued}\ ; \tag{6.2.6}$$

then there exist

$$X \in L^2\left(\Omega; L^\infty(0,T;H)\right) \cap L^p\left(\Omega \times (0,T); V\right)\,, \qquad X \in C_w\left([0,T]; H\right) \quad \mathbb{P}\text{-}a.s.\,, \tag{6.2.7}$$

$$\eta \in L^q\left(\Omega \times (0,T) \times D\right)^d\,, \tag{6.2.8}$$

$$\xi \in L^1\left(\Omega \times (0,T) \times D\right)\,, \tag{6.2.9}$$

where X and ξ are predictable, η is adapted, and the following relations hold:

$$X(t) - \int_0^t \operatorname{div} \eta(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s)\,dW_s \qquad \textit{in } L^1(D) \cap V^*, \quad \forall\, t \in [0,T]\,, \quad \mathbb{P}\text{-}a.s.\,, \tag{6.2.10}$$

$$\eta \in \gamma(\nabla X) \quad a.e.\ in\ \Omega \times (0,T) \times D\,, \tag{6.2.11}$$

$$\xi \in \beta(X) \quad a.e.\ in\ \Omega \times (0,T) \times D\,, \tag{6.2.12}$$

$$j(X) + j^*(\xi) \in L^1\left(\Omega \times (0,T) \times D\right)\,. \tag{6.2.13}$$

Furthermore, if hypothesis (6.2.6) *is not assumed, then the same conclusion is true replacing conditions* (6.2.7) *and* (6.2.10) *with, respectively,*

$$X \in L^\infty\big(0,T; L^2(\Omega;H)\big) \cap L^p\left(\Omega \times (0,T); V\right) \cap C_w\big([0,T]; L^2(\Omega;H)\big)\,, \tag{6.2.14}$$

$$X(t) - \int_0^t \operatorname{div} \eta(s)\, ds + \int_0^t \xi(s)\, ds = X_0 + \int_0^t B(s)\, dW_s \qquad \text{in } L^1(D) \cap V^*, \quad \forall\, t \in [0,T], \quad \mathbb{P}\text{-}a.s. \tag{6.2.15}$$

Remark 6.2.3. The integral equation (6.2.10) is satisfied in the dual space V_0^*, but X is not V_0-valued, so that the results provided are not a direct generalization of the classical concept of variational solution (cf. [56]): we can define them as a weaker type of variational solution, in which the integral expression holds in a dual space V_0^*, but the solution takes values only in a space larger than V_0 (V in our case). Nevertheless, the integral formulation (6.2.10) can be seen as an identity in $L^1(D)$, so that the choice of V_0 turns out to be only a technical device *a posteriori*. The fact that one cannot expect classical variational solutions for this type of problem is due to fact that no hypotheses on the growth of β are assumed (in contrast to a large part of the literature).

Remark 6.2.4. Let us comment on hypothesis (6.2.6). The fact that γ is single-valued (thus a continuous function) is needed in order to prove uniqueness for our problem, which in turn ensures some reasonable measurability properties for the processes X, η and ξ, as we will show later on. On the other side, if we do not require (6.2.6), the measurability of the solutions cannot be shown using the same argument, but it has to be recovered in a different way: however, in this case, the formulation that one obtains is weaker than the previous one, since the passage to the limit has to be carried out in $\Omega \times D$, with $t \in [0,T]$ fixed, and the solution X is found is a larger space.

Theorem 6.2.5. *Assume that*

$$X_0^1, X_0^2 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)\,, \tag{6.2.16}$$

$$B_1, B_2 \in L^2\left(\Omega \times (0,T); \mathscr{L}^2(U,H)\right) \quad \textit{progressively measurable}. \tag{6.2.17}$$

If hypothesis (6.2.6) *holds and* (X_1, η_1, ξ_1), (X_2, η_2, ξ_2) *are any two corresponding solutions satisfying* (6.2.7)–(6.2.13), *then there is a constant* $C > 0$ *(independent of the above quantities) such that*

$$\begin{aligned} &\|X_1 - X_2\|_{L^2(\Omega; L^\infty(0,T;H))} \\ &\quad \le C \left\|X_0^1 - X_0^2\right\|_{L^2(\Omega;H)} + C \|B_1 - B_2\|_{L^2(\Omega \times (0,T); \mathscr{L}^2(U,H))}\,. \end{aligned} \tag{6.2.18}$$

In this setting, if $X_0^1 = X_0^2$ *and* $B_1 = B_2$, *then* $X_1 = X_2$, $\eta_1 = \eta_2$ *and* $\xi_1 = \xi_2$. *Moreover, if hypothesis* (6.2.6) *is not assumed and* (X_1, η_1, ξ_1), (X_2, η_2, ξ_2) *are any two corresponding solutions satisfying* (6.2.8)–(6.2.9) *and* (6.2.11)–(6.2.15), *then*

$$\begin{aligned} &\|X_1 - X_2\|_{L^\infty(0,T; L^2(\Omega;H))} \\ &\quad \le \left\|X_0^1 - X_0^2\right\|_{L^2(\Omega;H)} + \|B_1 - B_2\|_{L^2(\Omega \times (0,T); \mathscr{L}^2(U,H))}\,. \end{aligned} \tag{6.2.19}$$

In this setting, if $X_0^1 = X_0^2$ *and* $B_1 = B_2$, *then* $X_1 = X_2$ *and* $-\operatorname{div} \eta_1 + \xi_1 = -\operatorname{div} \eta_2 + \xi_2$.

Remark 6.2.6. The uniqueness result strongly depends on the assumption (6.2.6). Indeed, if (6.2.6) is in order, uniqueness holds for the three solution components, separately; on the other side, if we do not assume (6.2.6), we can only recover uniqueness for X and the joint process $-\operatorname{div} \eta + \xi$. Moreover, note that the nonlinearity γ prevents us from finding a continuous dependence estimate also in the space $L^p(\Omega \times (0,T); V)$ for any $p > 2$. Nevertheless, if $p = 2$ and γ is the identity, the operator $-\Delta$ is linear and we can recover continuous dependence also in $L^2(\Omega \times (0,T); V)$, for which we refer to Chapter 2.

Theorem 6.2.7. *Assume that*

$$X_0 \in L^2\left(\Omega, \mathcal{F}_0, \mathbb{P}; H\right), \tag{6.2.20}$$

$$B : \Omega \times [0,T] \times H \to \mathscr{L}^2(U,H) \quad \textit{progressively measurable}, \tag{6.2.21}$$

$$\begin{gathered} \exists\, L_B > 0 \ : \quad \|B(\omega,t,x_1) - B(\omega,t,x_2)\|_{\mathscr{L}^2(U,H)} \leq L_B \|x_1 - x_2\|_H \\ \textit{for every} \quad (\omega,t) \in \Omega \times [0,T], \ \ x_1, x_2 \in H, \end{gathered} \tag{6.2.22}$$

$$\begin{gathered} \exists\, R_B > 0 \ : \quad \|B(\omega,t,x)\|_{\mathscr{L}^2(U,H)} \leq R_B \left(1 + \|x\|_H\right) \\ \textit{for every} \quad (\omega,t,x) \in \Omega \times [0,T] \times H. \end{gathered} \tag{6.2.23}$$

If hypothesis (6.2.6) *holds, then there exists a triplet* (X, η, ξ) *satisfying conditions* (6.2.7)–(6.2.9), (6.2.11)–(6.2.13) *and*

$$\begin{gathered} X(t) - \int_0^t \operatorname{div} \eta(s)\, ds + \int_0^t \xi(s)\, ds = X_0 + \int_0^t B(s, X(s))\, dW_s \\ \textit{in } L^1(D) \cap V^*, \quad \forall\, t \in [0,T], \quad \mathbb{P}\textit{-a.s.} \end{gathered} \tag{6.2.24}$$

If hypothesis (6.2.6) *is not assumed, then the same conclusion is true replacing* (6.2.7) *with* (6.2.14)*, and condition* (6.2.24) *with*

$$\begin{gathered} X(t) - \int_0^t \operatorname{div} \eta(s)\, ds + \int_0^t \xi(s)\, ds = X_0 + \int_0^t B(s, X(s))\, dW_s \\ \textit{in } L^1(D) \cap V^*, \quad \forall\, t \in [0,T], \quad \mathbb{P}\textit{-a.s.} \end{gathered} \tag{6.2.25}$$

Theorem 6.2.8. *Let* X_0^1, X_0^2 *satisfy condition* (6.2.16)*. If* (6.2.6) *holds, and* (X_1, η_1, ξ_1) *and* (X_2, η_2, ξ_2) *are any two corresponding solutions satisfying* (6.2.7)–(6.2.9), (6.2.11)–(6.2.13) *and* (6.2.24)*, then there is a constant* $C > 0$ *(independent of the above quantities) such that*

$$\|X_1 - X_2\|_{L^2(\Omega; L^\infty(0,T;H))} \leq C \left\|X_0^1 - X_0^2\right\|_{L^2(\Omega;H)}. \tag{6.2.26}$$

In this setting, if $X_0^1 = X_0^2$*, then* $X_1 = X_2$*,* $\eta_1 = \eta_2$ *and* $\xi_1 = \xi_2$*. Moreover, if hypothesis* (6.2.6) *is not assumed and* (X_1, η_1, ξ_1)*,* (X_2, η_2, ξ_2) *are any two corresponding solutions satisfying* (6.2.8)–(6.2.9), (6.2.11)–(6.2.14) *and* (6.2.25)*, then there is a constant* $C > 0$ *(independent of the above quantities) such that*

$$\|X_1 - X_2\|_{L^\infty(0,T; L^2(\Omega;H))} \leq C \left\|X_0^1 - X_0^2\right\|_{L^2(\Omega;H)}. \tag{6.2.27}$$

In this setting, if $X_0^1 = X_0^2$*, then* $X_1 = X_2$ *and* $-\operatorname{div} \eta_1 + \xi_1 = -\operatorname{div} \eta_2 + \xi_2$*.*

Remark 6.2.9. It is worth recalling the classical approach to problem (6.1.1)–(6.1.3) in the deterministic case and the main differences with the stochastic case. The corresponding deterministic problem is

$$\frac{\partial u}{\partial t} - \operatorname{div} \gamma(\nabla u) + \beta(u) \ni f, \quad u(0) = u_0,$$

with homogeneous boundary conditions for u: here, the classical approach consists in proving that the sum of the two operators $-\operatorname{div}(\nabla \cdot)$ and $\beta(\cdot)$ is m-accretive in a suitable space. To this end, it is well-known that if (i) E is a Banach space with uniformly convex dual E^*, (ii) A and B are two m-accretive sets in $E \times E$, (iii) $D(A) \cap D(B) \neq \emptyset$, (iv) $\langle Au, J(B_\lambda u)\rangle_E \geq 0$ for every $u \in D(A)$ and $\lambda \in (0,1)$ (where $J : E \to E^*$ is the duality mapping of E and B_λ is the Yosida approximation of B), then $A + B$ is m-accretive in $E \times E$ (see [10, Prop. 3.8]). If we take for example $E = L^s(D)$ for $1 < s < +\infty$, $A = -\operatorname{div}\gamma(\nabla\cdot)$, $B = \beta(\cdot)$ with their natural domains, we only need to check (iv), since (i)–(iii) are clearly satisfied. To this aim, we need to handle the term

$$\int_D -\operatorname{div}\gamma(\nabla u)\phi(\beta_\lambda(u))\,,$$

where $\phi(r) = |r|^{s-2}r$, $r \in \mathbb{R}$, using integration by parts. The first problem occurs if $s < 2$, since in this case the derivative of ϕ explodes at 0; if $s \geq 2$, we can proceed formally and recover

$$\int_D \phi'(\beta_\lambda(u))\beta'_\lambda(u)\gamma(\nabla u)\cdot\nabla u \geq 0\,.$$

The main difficulty is that β_λ is not differentiable, so that one needs to rely on some generalized chain-rules for Lipschitz functions or suitable mollifications of β_λ. The problem can be seen then as a particular case of the general one

$$\frac{\partial u}{\partial t} + Au \ni f\,,$$

with A purely nonlinear (multivalued) operator, for which one can rely on several classical well-posedness results. However, the corresponding general problem in the stochastic case, i.e.

$$du + Au\,dt \ni B\,dW_t\,,$$

does not have a direct counterpart in terms of existence and uniqueness: as a consequence, in our case the proof of m-accretivity is not sufficient to ensure well-posedness, so that one needs to deal with the problem "by hand". To this end, the variational approach is in order.

6.3 Existence with additive noise

In this section we prove the two existence results contained in Theorem 6.2.2: as already mentioned, we are going to approximate the problem using two different parameters. Uniform estimates are then proved and we obtain global solutions to the original problem by passing to the limit in a suitable topology.

6.3.1 The approximated problem

Thanks to (6.2.5), for every $\varepsilon \in (0,1)$ there exists an operator

$$B^\varepsilon \in L^2\left(\Omega\times(0,T);\mathscr{L}^2(U,V_0)\right) \tag{6.3.28}$$

such that:

$$B^\varepsilon \to B \quad \text{in } L^2\left(\Omega\times(0,T);\mathscr{L}^2(U,H)\right) \quad \text{as } \varepsilon\searrow 0\,, \tag{6.3.29}$$

$$\|B^\varepsilon\|_{L^2(\Omega\times(0,T);\mathscr{L}^2(U,H))} \leq \|B\|_{L^2(\Omega\times(0,T);\mathscr{L}^2(U,H))} \quad \text{for every } \varepsilon\in(0,1)\,. \tag{6.3.30}$$

Indeed, if k is chosen as in the definition of V_0 in the previous section, then the operator $(I-\varepsilon\Delta)^{-k}$ maps H into V_0 for every $\varepsilon > 0$, so that it suffices to take $B^\varepsilon := (I-\varepsilon\Delta)^{-k}B$. With this particular choice, using the fact that the operator $(I-\varepsilon\Delta)^{-k} : H \to H$ is a linear contraction converging to the identity in the strong operator topology as $\varepsilon\searrow 0$ and the ideal property of $\mathscr{L}^2(U;H)$ in $\mathscr{L}(U,H)$, we have that (6.3.28)–(6.3.30) are satisfied.

For every $\lambda \in (0,1)$ and $\varepsilon \in (0,1)$, let us consider the approximated problem

$$dX_\lambda^\varepsilon - \operatorname{div}[\gamma_\lambda(\nabla X_\lambda^\varepsilon) + \lambda \nabla X_\lambda^\varepsilon]\,dt + \beta_\lambda(X_\lambda^\varepsilon)\,dt = B^\varepsilon\,dW_t \quad \text{in } D \times (0,T)\,,$$
$$X_\lambda^\varepsilon(0) = X_0 \quad \text{in } D\,,$$

whose integral formulation is given by

$$\begin{aligned} X_\lambda^\varepsilon(t) - \int_0^t \operatorname{div}[\gamma_\lambda(\nabla X_\lambda^\varepsilon(s))]\,ds - \lambda \int_0^t \Delta X_\lambda^\varepsilon(s)\,ds + \int_0^t \beta_\lambda(X_\lambda^\varepsilon(s))\,ds \\ = X_0 + \int_0^t B^\varepsilon(s)\,dW_s \end{aligned} \tag{6.3.31}$$

in $H^{-1}(D)$, for every $t \in [0,T]$, $\mathbb{P}$-almost surely, where here the divergence operator $-\operatorname{div}: L^2(D)^d \to H^{-1}(D)$ and the laplacian is intended in the usual variational way, i.e.

$$-\Delta : H_0^1(D) \to H^{-1}(D)\,, \quad \langle -\Delta u, v\rangle_{H_0^1(D)} := \int_D \nabla u \cdot \nabla v\,, \quad u, v \in H_0^1(D)\,.$$

A unique solution to the approximated problem (6.3.31) can be easily obtained using the classical results contained in [46] (see also [56, Thm. 4.2.4]). In fact, the operator

$$A_\lambda : H_0^1(D) \to H^{-1}(D)\,, \quad A_\lambda : \phi \mapsto -\operatorname{div}[\gamma_\lambda(\nabla\phi) + \lambda\nabla\phi] + \beta_\lambda(\phi)\,, \tag{6.3.32}$$

is well-defined thanks to the Lipschitz continuity of β_λ and γ_λ, and problem (6.3.31) is the variational formulation with respect to the Gelfand triple $H_0^1(D) \hookrightarrow H \hookrightarrow H^{-1}(D)$ of the following:

$$dX_\lambda^\varepsilon + A_\lambda X_\lambda^\varepsilon\,dt = B^\varepsilon\,dW_t \quad \text{in } (0,T) \times D\,, \tag{6.3.33}$$
$$X_\lambda^\varepsilon(0) = X_0 \quad \text{in } D\,. \tag{6.3.34}$$

In this setting, we need to check that the operator A_λ satisfies the classical properties of hemicontinuity, monotonicity, coercivity and boundedness, in order to recover solutions of (6.3.31). The following lemma is straightforward.

Lemma 6.3.1. *The following conditions are satisfied for every $\lambda \in (0,1)$.*

(H1) (Hemicontinuity). For all $u, v, x \in H_0^1(D)$, the following map is continuous:

$$s \mapsto \langle A_\lambda(u + sv), x\rangle_{H_0^1(D)}\,, \quad s \in \mathbb{R}\,.$$

(H2) (Monotonicity). For all $u, v \in H_0^1(D)$,

$$\langle A_\lambda u - A_\lambda v, u - v\rangle_{H_0^1(D)} \geq 0\,.$$

(H3) (Coercivity). There exists $C_1 > 0$ such that, for all $v \in H_0^1(D)$,

$$\langle A_\lambda v, v\rangle_{H_0^1(D)} \geq C_1 \|v\|_{H_0^1(D)}^2\,.$$

(H4) (Boundedness). There exists $C_2 > 0$ such that, for all $v \in H_0^1(D)$,

$$\|A_\lambda v\|_{H^{-1}(D)} \leq C_2 \|v\|_{H_0^1(D)}\,.$$

Proof. For all $u, v, x \in H_0^1(D)$ we have

$$\begin{aligned}&\langle A_\lambda(u+sv), x\rangle_{H_0^1(D)}\\&\qquad= \int_D \gamma_\lambda(\nabla(u+sv))\cdot\nabla x + \lambda\int_D \nabla(u+sv)\cdot\nabla x + \int_D \beta_\lambda(u+sv)x\,,\end{aligned}$$

so that (H1) is satisfied thanks to the Lipschitz continuity of γ_λ and β_λ. Secondly, (H2) trivially holds using the monotonicity of γ_λ and β_λ. Moreover, for all $v \in H_0^1(D)$, thanks to the monotonicity of γ_λ and β_λ, and the fact that $\gamma(0) \ni 0$ and $\beta(0) \ni 0$, we have

$$\langle A_\lambda v, v\rangle_{H_0^1(D)} = \int_D \gamma_\lambda(\nabla v)\cdot\nabla v + \lambda\int_D |\nabla v|^2 + \int_D \beta_\lambda(v)v \geq \lambda\int_D |\nabla v|^2\,,$$

so that (H3) holds true thanks to the Poincaré inequality. Finally, using the Lipschitz continuity of β_λ and γ_λ and the Hölder inequality, we have for all $u, v \in H_0^1(D)$

$$\begin{aligned}&\langle A_\lambda v, u\rangle_{H_0^1(D)} = \int_D \gamma_\lambda(\nabla v)\cdot\nabla u + \lambda\int_D \nabla v\cdot\nabla u + \int_D \beta_\lambda(v)u\\&\leq \left(\frac{1}{\lambda}+\lambda\right)\|\nabla v\|_H\|\nabla u\|_H + \frac{1}{\lambda}\|v\|_H\|u\|_H \leq \left(\frac{2}{\lambda}+\lambda\right)\|v\|_{H_0^1(D)}\|u\|_{H_0^1(D)}\,,\end{aligned}$$

from which (H4) follows. □

Lemma 6.3.1 ensures that, for all $\varepsilon, \lambda \in (0,1)$, there exists a unique adapted process

$$X_\lambda^\varepsilon \in L^2\left(\Omega; C([0,T];H)\right) \cap L^2\left(\Omega\times(0,T); H_0^1(D)\right) \tag{6.3.35}$$

such that

$$\begin{aligned}&X_\lambda^\varepsilon(t) - \int_0^t \operatorname{div}[\gamma_\lambda(\nabla X_\lambda^\varepsilon(s))]\,ds - \lambda\int_0^t \Delta X_\lambda^\varepsilon(s)\,ds + \int_0^t \beta_\lambda(X_\lambda^\varepsilon(s))\,ds\\&\qquad= X_0 + \int_0^t B^\varepsilon(s)\,dW_s\end{aligned} \tag{6.3.36}$$

in $H^{-1}(D)$, for every $t \in [0,T]$, $\mathbb{P}$-almost surely.

6.3.2 A priori estimates I

Here we prove uniform pathwise estimates on X_λ^ε, independent of λ (but not of ε), which will allow us to pass to the limit as $\lambda \searrow 0$ in the approximated problem (6.3.36) with ε fixed.

Let us define, for any $\varepsilon \in (0,1)$,

$$W_B^\varepsilon(t) := \int_0^t B^\varepsilon(s)\,dW_s\,, \quad t \in [0,T]\,. \tag{6.3.37}$$

Thanks to the Burkholder-Davis-Gundy inequality and condition (6.2.5) we deduce

$$W_B^\varepsilon \in L^2\left(\Omega; L^\infty(0,T;V_0)\right)\,. \tag{6.3.38}$$

In particular, recalling that $V_0 \hookrightarrow V \cap L^\infty(D)$, we have that

$$W_B^\varepsilon(\omega) \in L^p(0,T;V)\cap L^\infty(Q) \quad \text{for } \mathbb{P}\text{-almost every } \omega \in \Omega\,. \tag{6.3.39}$$

Equation (6.3.36) can be rewritten as

$$\partial_t \left(X_\lambda^\varepsilon - W_B^\varepsilon\right)(t) - \operatorname{div}\left[\gamma_\lambda(\nabla X_\lambda^\varepsilon(t)) + \lambda \nabla X_\lambda^\varepsilon(t)\right] + \beta_\lambda(X_\lambda^\varepsilon(t)) = 0 \quad \text{in } H^{-1}(D)$$

for every $t \in [0,T]$, for any ω out of a set of probability 0 (the symbol ∂_t for the derivative with respect to time makes sense only if applied to the difference $X_\lambda^\varepsilon - W_B^\varepsilon$). Fix now ω and test by $X_\lambda^\varepsilon(t) - W_B^\varepsilon(t)$ (see [8, §1.3]): we obtain

$$\begin{aligned}
&\frac{1}{2}\left\|X_\lambda^\varepsilon(t) - W_B^\varepsilon(t)\right\|_H^2 + \int_0^t \int_D \gamma_\lambda(\nabla X_\lambda^\varepsilon(s)) \cdot \nabla(X_\lambda^\varepsilon(s) - W_B^\varepsilon(s))\, ds \\
&\quad + \lambda \int_0^t \int_D \nabla X_\lambda^\varepsilon(s) \cdot \nabla\left(X_\lambda^\varepsilon(s) - W_B^\varepsilon(s)\right)\, ds \\
&\quad + \int_0^t \int_D \beta_\lambda(X_\lambda^\varepsilon(s))(X_\lambda^\varepsilon(s) - W_B^\varepsilon(s))\, ds = \frac{1}{2}\left\|X_0\right\|_H^2 .
\end{aligned} \tag{6.3.40}$$

Using the identity $I_d = \lambda\gamma_\lambda + J_\lambda$ and rearranging terms in the previous relation, we have

$$\begin{aligned}
&\frac{1}{2}\left\|X_\lambda^\varepsilon(t) - W_B^\varepsilon(t)\right\|_H^2 + \int_0^t \int_D \gamma_\lambda(\nabla X_\lambda^\varepsilon(s)) \cdot J_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\, ds \\
&\quad + \lambda \int_0^t \int_D \left|\gamma_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\right|^2\, ds + \lambda \int_0^t \int_D \left|\nabla X_\lambda^\varepsilon(s)\right|^2 ds \\
&\quad + \int_0^t \int_D \beta_\lambda(X_\lambda^\varepsilon(s))(X_\lambda^\varepsilon(s) - W_B^\varepsilon(s))\, ds \\
&= \frac{1}{2}\left\|X_0\right\|_H^2 + \int_0^t \int_D \gamma_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right) \cdot \nabla W_B^\varepsilon(s)\, ds + \lambda \int_0^t \int_D \nabla X_\lambda^\varepsilon(s) \cdot \nabla W_B^\varepsilon(s)\, ds .
\end{aligned}$$

Using the generalized Young inequality of the form $ab \le \delta \frac{p-1}{p} a^{\frac{p}{p-1}} + C_{\delta,p} b^p$ (for any $a, b, \delta > 0$ and a certain $C_{\delta,p} > 0$) on the second term on the right-hand side, thanks also to hypotheses on γ and the properties of the resolvent we deduce for every $t \in [0,T]$ that

$$\begin{aligned}
&\frac{1}{2}\left\|X_\lambda^\varepsilon(t) - W_B^\varepsilon(t)\right\|_H^2 + K \int_0^t \left\|J_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\right\|_{L^p(D)}^p ds + \lambda \int_0^t \left\|\gamma_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\right\|_H^2 ds \\
&\quad + \lambda \int_0^t \left\|\nabla X_\lambda^\varepsilon(s)\right\|_H^2 ds + \int_0^t \int_D \beta_\lambda(X_\lambda^\varepsilon(s))(X_\lambda^\varepsilon(s) - W_B^\varepsilon(s))\, ds \\
&\le C' + \frac{1}{2}\left\|X_0\right\|_H^2 + \delta \frac{(p-1)D_1}{p} \int_0^t \left\|J_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\right\|_{L^p(D)}^p ds \\
&\quad + C_{\delta,p} \int_0^t \left\|\nabla W_B^\varepsilon(s)\right\|_{L^p(D)}^p ds + \frac{\lambda}{2} \int_0^t \left\|\nabla X_\lambda^\varepsilon(s)\right\|_H^2 ds + \frac{\lambda}{2} \int_0^t \left\|\nabla W_B^\varepsilon(s)\right\|_H^2 ds
\end{aligned}$$

for a positive constants C' independent of λ and ε. Hence, choosing $\delta = \frac{Kp}{2D_1(p-1)}$, we get that, for every $t \in [0,T]$,

$$\begin{aligned}
&\left\|X_\lambda^\varepsilon(t)\right\|_H^2 + \frac{K}{2} \int_0^t \left\|J_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\right\|_{L^p(D)}^p ds + \lambda \int_0^t \left\|\gamma_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\right\|_H^2 ds \\
&\quad + \frac{\lambda}{2} \int_0^t \left\|\nabla X_\lambda^\varepsilon(s)\right\|_H^2 ds + \int_0^t \int_D \beta_\lambda(X_\lambda^\varepsilon(s))(X_\lambda^\varepsilon(s) - W_B^\varepsilon(s))\, ds \\
&\le C' + \frac{1}{2}\left\|X_0\right\|_H^2 + C_p \left\|W_B^\varepsilon\right\|_{L^p(0,T;V)}^p + \frac{1}{2}\left\|W_B^\varepsilon\right\|_{L^\infty(0,T;H)}^2 \\
&\quad + \frac{1}{2}\left\|W_B^\varepsilon\right\|_{L^2(0,T;H_0^1(D))}^2
\end{aligned} \tag{6.3.41}$$

for a positive constant C_p independent of λ and ε. Denoting by $j_\lambda : \mathbb{R} \to [0,+\infty)$ the proper, convex, lower semicontinuous function such that $\beta_\lambda = \partial j_\lambda$ and $j_\lambda(0) = 0$, one has that $j_\lambda \le j$ and $j_\lambda(x) \nearrow j(x)$ for every $x \in \mathbb{R}$ (recall that $\mathbb{R} = D(\beta) \subseteq D(j)$). Hence, for every $x, y \in \mathbb{R}$ we have that

$$\beta_\lambda(x)(x-y) \ge j_\lambda(x) - j_\lambda(y) \ge j_\lambda(x) - j(y)\,.$$

Applying this inequality to the last term on the left-hand side of (6.3.41), we deduce that, for every $t \in [0,T]$,

$$\begin{aligned}
&\|X^\varepsilon_\lambda(t)\|_H^2 + \frac{K}{2}\int_0^t \|J_\lambda(\nabla X^\varepsilon_\lambda(s))\|^p_{L^p(D)}\,ds + \lambda\int_0^t \|\gamma_\lambda(\nabla X^\varepsilon_\lambda(s))\|_H^2\,ds \\
&\qquad + \frac{\lambda}{2}\int_0^t \|\nabla X^\varepsilon_\lambda(s)\|_H^2\,ds + \int_0^t\int_D j_\lambda(X^\varepsilon_\lambda(s))\,ds \\
&\lesssim 1 + \|X_0\|_H^2 + \|W^\varepsilon_B\|^2_{L^p(0,T;V)} + \|W^\varepsilon_B\|^2_{L^\infty(0,T;H)} \\
&\qquad + \|W^\varepsilon_B\|^2_{L^2(0,T;H^1_0(D))} + \int_Q j(W^\varepsilon_B)\,.
\end{aligned}$$

Note that all the terms on the right-hand side are finite $\mathbb{P}$-almost surely: for the first five, this is immediate thanks to (6.2.4) and (6.3.39), while $j(W^\varepsilon_B) \in L^1(Q)$ since $W^\varepsilon_B \in L^\infty(Q)$. Using the positivity of j_λ we deduce that for $\mathbb{P}$-almost every $\omega \in \Omega$ there exists a positive constant $M = M_{\omega,\varepsilon}$, independent of λ, such that, for every $\lambda \in (0,1)$,

$$\|X^\varepsilon_\lambda(\omega)\|_{L^\infty(0,T;H)} \le M_{\omega,\varepsilon}\,, \tag{6.3.42}$$

$$\|J_\lambda(\nabla X^\varepsilon_\lambda(\omega))\|_{L^p(Q)} \le M_{\omega,\varepsilon}\,, \tag{6.3.43}$$

$$\lambda^{1/2}\|\gamma_\lambda(\nabla X^\varepsilon_\lambda(\omega))\|_{L^2(Q)} \le M_{\omega,\varepsilon}\,, \tag{6.3.44}$$

$$\lambda^{1/2}\|\nabla X^\varepsilon_\lambda(\omega)\|_{L^2(Q)} \le M_{\omega,\varepsilon}\,. \tag{6.3.45}$$

Finally, by the hypotheses on γ and the properties of the resolvent we also have

$$\int_Q |\gamma_\lambda(\nabla X^\varepsilon_\lambda)|^q \le D_1 \int_Q (1 + |J_\lambda(\nabla X^\varepsilon_\lambda)|)^p\,,$$

so that by (6.3.43) it follows (possibly redefining $M_{\omega,\varepsilon}$) that, for every $\lambda \in (0,1)$,

$$\|\gamma_\lambda(\nabla X^\varepsilon_\lambda(\omega))\|_{L^q(Q)} \le M_{\omega,\varepsilon}\,. \tag{6.3.46}$$

6.3.3 A priori estimates II

In this section we prove some estimates in expectation on X^ε_λ independent both of λ and ε. The main tool is a version of Itô's formula in a variational framework.

Thanks to conditions (6.2.4)–(6.2.5) and (6.3.35)–(6.3.36), we can apply Itô's formula (see [56]), obtaining

$$\begin{aligned}
&\frac{1}{2}\|X^\varepsilon_\lambda(t)\|_H^2 + \int_0^t\int_D \gamma_\lambda(\nabla X^\varepsilon_\lambda(s))\cdot\nabla X^\varepsilon_\lambda(s)\,ds \\
&\qquad + \lambda\int_0^t\int_D |\nabla X^\varepsilon_\lambda(s)|^2\,ds + \int_0^t\int_D \beta_\lambda(X^\varepsilon_\lambda(s))X^\varepsilon_\lambda(s)\,ds \qquad (6.3.47)\\
&= \frac{1}{2}\|X_0\|_H^2 + \frac{1}{2}\int_0^t \|B^\varepsilon(s)\|^2_{\mathscr{L}^2(U,H)}\,ds + \int_0^t (X^\varepsilon_\lambda(s), B^\varepsilon(s)\,dW_s)
\end{aligned}$$

for every $t \in [0,T]$, $\mathbb{P}$-almost surely, which yields, by definition of γ_λ,

$$\begin{aligned}
&\frac{1}{2}\left\|X_\lambda^\varepsilon(t)\right\|_H^2 + K\int_0^t \left\|J_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\right\|_{L^p(D)}^p ds + \lambda\int_0^t \left\|\gamma_\lambda\left(\nabla X_\lambda^\varepsilon(s)\right)\right\|_H^2 ds\\
&\quad + \lambda\int_0^t \left\|\nabla X_\lambda^\varepsilon(s)\right\|_H^2 ds + \int_0^t\int_D \beta_\lambda(X_\lambda^\varepsilon(s))X_\lambda^\varepsilon(s)\,ds\\
&\leq C'' + \frac{1}{2}\left\|X_0\right\|_H^2 + \frac{1}{2}\left\|B^\varepsilon(s)\right\|_{L^2(0,T;\mathscr{L}^2(U,H))}^2 + \sup_{t\in[0,T]}\left|\int_0^t \left(X_\lambda^\varepsilon(s), B^\varepsilon(s)\,dW_s\right)\right|
\end{aligned}$$

for a constant $C'' > 0$, independent of ε and λ. Thanks to Davis' inequality, the Hölder and Young inequalities, and condition (6.3.30), we have for some $c, \tilde{c} > 0$

$$\begin{aligned}
\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t \left(X_\lambda^\varepsilon(s), B^\varepsilon(s)\,dW_s\right)\right| &\leq c\mathbb{E}\left[\left(\int_0^T \left\|X_\lambda^\varepsilon(s)\right\|_H^2 \left\|B^\varepsilon(s)\right\|_{\mathscr{L}^2(U,H)}^2 ds\right)^{1/2}\right]\\
&\leq c\mathbb{E}\left[\left\|X_\lambda^\varepsilon\right\|_{L^\infty(0,T;H)}\left\|B^\varepsilon\right\|_{L^2(0,T;\mathscr{L}^2(U,H))}\right]\\
&\leq \frac{1}{4}\left\|X_\lambda^\varepsilon\right\|_{L^2(\Omega;L^\infty(0,T;H))}^2 + \tilde{c}\left\|B\right\|_{L^2(\Omega\times(0,T);\mathscr{L}^2(U,H))}^2\,;
\end{aligned}$$

consequently, taking the supremum in $t \in [0,T]$ and expectations, we obtain

$$\begin{aligned}
&\frac{1}{4}\left\|X_\lambda^\varepsilon\right\|_{L^2(\Omega;L^\infty(0,T;H))}^2 + K\left\|J_\lambda\left(\nabla X_\lambda^\varepsilon\right)\right\|_{L^p(\Omega\times(0,T)\times D)}^p\\
&\quad + \lambda\left\|\gamma_\lambda\left(\nabla X_\lambda^\varepsilon\right)\right\|_{L^2(\Omega\times(0,T)\times D)}^2 + \lambda\left\|\nabla X_\lambda^\varepsilon\right\|_{L^2(\Omega\times(0,T)\times D)}^2 + \int_{\Omega\times Q}\beta_\lambda(X_\lambda^\varepsilon)X_\lambda^\varepsilon \qquad (6.3.48)\\
&\leq C'' + \frac{1}{2}\left\|X_0\right\|_{L^2(\Omega;H)}^2 + \frac{3}{2}\left\|B\right\|_{L^2(\Omega\times(0,T);\mathscr{L}^2(U,H))}^2\,.
\end{aligned}$$

We infer that there exists a constant $N > 0$, independent of λ and ε, such that

$$\left\|X_\lambda^\varepsilon\right\|_{L^2(\Omega;L^\infty(0,T;H))} \leq N\,, \qquad (6.3.49)$$

$$\left\|J_\lambda\left(\nabla X_\lambda^\varepsilon\right)\right\|_{L^p(\Omega\times(0,T)\times D)} \leq N\,, \qquad (6.3.50)$$

$$\lambda^{1/2}\left\|\gamma_\lambda\left(\nabla X_\lambda^\varepsilon\right)\right\|_{L^2(\Omega\times(0,T)\times D)} \leq N\,, \qquad (6.3.51)$$

$$\lambda^{1/2}\left\|\nabla X_\lambda^\varepsilon\right\|_{L^2(\Omega\times(0,T)\times D)} \leq N\,, \qquad (6.3.52)$$

for every $\varepsilon, \lambda \in (0,1)$. Finally, by the assumptions on γ we also have

$$\int_{\Omega\times Q}\left|\gamma_\lambda(\nabla X_\lambda^\varepsilon)\right|^q \leq D_1\int_{\Omega\times Q}\left(1 + \left|J_\lambda(\nabla X_\lambda^\varepsilon)\right|\right)^p,$$

so that by (6.3.50) it follows (possibly redefining N) that, for every $\varepsilon, \lambda \in (0,1)$,

$$\left\|\gamma_\lambda(\nabla X_\lambda^\varepsilon)\right\|_{L^q(\Omega\times(0,T)\times D)} \leq N\,. \qquad (6.3.53)$$

6.3.4 A priori estimates III

In this section we prove uniform estimates on the term $\beta_\lambda(X_\lambda^\varepsilon)$, independent of λ (with ε fixed), which are useful to recover a suitable weak compactness. We rely on some computations performed in [9] to obtain some L^1 estimates, the classical results by de la Vallée-Poussin about uniform integrability and on the Dunford-Pettis theorem.

Firstly, let us fix $\omega \in \Omega$. The properties of the resovent and the monotonicity of β_λ imply that

$$j(R_\lambda X_\lambda^\varepsilon) + j^*(\beta_\lambda(X_\lambda^\varepsilon)) = \beta_\lambda(X_\lambda^\varepsilon) R_\lambda X_\lambda^\varepsilon \le |\beta_\lambda(X_\lambda^\varepsilon)|\,|X_\lambda^\varepsilon| = \beta_\lambda(X_\lambda^\varepsilon) X_\lambda^\varepsilon\,.$$

Consequently, from inequality (6.3.41) evaluated at time T and the previous relation, recalling (6.3.39) and using the generalized Young inequality of the form $ab \le j(2a) + j^*(b/2)$ for any $a, b \in \mathbb{R}$, we deduce that $\mathbb{P}$-almost surely we have

$$\begin{aligned}
\int_Q j^*(\beta_\lambda(X_\lambda^\varepsilon)) &\le \int_Q \beta_\lambda(X_\lambda^\varepsilon) X_\lambda^\varepsilon \\
&\le C' + \frac{1}{2}\|X_0\|_H^2 + C_p \|W_B^\varepsilon\|_{L^p(0,T;V)}^p + \frac{1}{2}\|W_B^\varepsilon\|_{L^\infty(0,T;H)}^2 \\
&\quad + \frac{1}{2}\|W_B^\varepsilon\|_{L^2(0,T;H_0^1(D))}^2 + \int_Q \beta_\lambda(X_\lambda^\varepsilon) W_B^\varepsilon \\
&\le C' + \frac{1}{2}\|X_0\|_H^2 + C_p \|W_B^\varepsilon\|_{L^p(0,T;V)}^p + \frac{1}{2}\|W_B^\varepsilon\|_{L^\infty(0,T;H)}^2 \\
&\quad + \frac{1}{2}\|W_B^\varepsilon\|_{L^2(0,T;H_0^1(D))}^2 + \|j\left(2W_B^\varepsilon\right)\|_{L^1(Q)} + \frac{1}{2}\int_Q j^*\left(\beta_\lambda(X_\lambda^\varepsilon)\right).
\end{aligned}$$

All the terms on the right hand side are finite thanks to (6.2.4) and (6.3.39): hence, since j^* is even by assumption, we have proved that

$$\|j^*\left(|\beta_\lambda(X_\lambda^\varepsilon(\omega))|\right)\|_{L^1(Q)} = \|j^*\left(\beta_\lambda(X_\lambda^\varepsilon(\omega))\right)\|_{L^1(Q)} \le \int_Q \beta_\lambda(X_\lambda^\varepsilon(\omega)) X_\lambda^\varepsilon(\omega) \le M_{\omega,\varepsilon} \qquad (6.3.54)$$

for $\mathbb{P}$-almost every $\omega \in \Omega$; moreover, since $D(\beta) = \mathbb{R}$, we have that

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

Hence, using then the criterion by de la Vallée-Poussin for uniform integrability combined with the Dunford-Pettis theorem, we deduce that, for $\mathbb{P}$-almost every $\omega \in \Omega$ and for every $\varepsilon \in (0,1)$,

$$\{\beta_\lambda(X_\lambda^\varepsilon)(\omega)\}_{\lambda\in(0,1)} \quad \text{is weakly relatively compact in } L^1(Q)\,. \qquad (6.3.55)$$

Finally, let us obtain the corresponding information also in expectation. It easily follows from (6.3.48) that there exists a constant $N > 0$, independent of λ and ε, such that

$$\|\beta_\lambda(X_\lambda^\varepsilon) X_\lambda^\varepsilon\|_{L^1(\Omega\times(0,T)\times D)} \le N \quad \text{for every } \varepsilon, \lambda \in (0,1)\,;$$

hence, in analogy to the derivation of (6.3.54), we get

$$\int_{\Omega\times Q} j^*(\beta_\lambda(X_\lambda^\varepsilon)) \le \|\beta_\lambda(X_\lambda^\varepsilon) X_\lambda^\varepsilon\|_{L^1(\Omega\times(0,T)\times D)} \le N \quad \text{for every } \varepsilon, \lambda \in (0,1)\,. \qquad (6.3.56)$$

Since j^* is even and superlinear at infinity, the criterion by de la Vallée-Poussin and the Dunford-Pettis theorem imply that

$$\{\beta_\lambda(X_\lambda^\varepsilon)\}_{\varepsilon,\lambda\in(0,1)} \quad \text{is weakly relatively compact in } L^1(\Omega\times(0,T)\times D)\,. \qquad (6.3.57)$$

6.3.5 Passage to the limit as $\lambda \searrow 0$

In this section, we pass to the limit as $\lambda \searrow 0$ in the approximated problem (6.3.36) with $\varepsilon \in (0,1)$ being fixed: the idea is to pass to the limit pathwise as $\lambda \searrow 0$. Throughout the section, $\varepsilon \in (0,1)$ and $\omega \in \Omega$ are fixed.

First of all, conditions (6.3.42)–(6.3.46) and (6.3.55) ensure that there exist

$$
\begin{gathered}
X^\varepsilon(\omega) \in L^\infty(0,T;H)\,,\\
Y^\varepsilon(\omega) \in L^p(Q)^d\,,\\
\eta^\varepsilon(\omega) \in L^q(Q)^d\,,\\
\xi^\varepsilon(\omega) \in L^1(Q)
\end{gathered}
$$

and a sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ (which clearly depends on ε and ω as well) such that as $n \to \infty$

$$X^\varepsilon_{\lambda_n}(\omega) \overset{*}{\rightharpoonup} X^\varepsilon(\omega) \quad \text{in } L^\infty(0,T;H)\,, \tag{6.3.58}$$

$$J_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}(\omega)\right) \rightharpoonup Y^\varepsilon(\omega) \quad \text{in } L^p(Q)^d\,, \tag{6.3.59}$$

$$\gamma_{\lambda_n}(\nabla X^\varepsilon_\lambda(\omega)) \rightharpoonup \eta^\varepsilon(\omega) \quad \text{in } L^q(Q)^d\,, \tag{6.3.60}$$

$$\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega)) \rightharpoonup \xi^\varepsilon(\omega) \quad \text{in } L^1(Q) \tag{6.3.61}$$

and also as $\lambda \searrow 0$ that

$$\lambda\gamma_\lambda(\nabla X^\varepsilon_\lambda(\omega)) \to 0 \quad \text{in } L^2(Q)^d\,, \tag{6.3.62}$$

$$\lambda\nabla X^\varepsilon_\lambda(\omega) \to 0 \quad \text{in } L^2(Q)^d\,. \tag{6.3.63}$$

In particular, since $\lambda^2|\gamma_\lambda(\nabla X^\varepsilon_\lambda)|^2 = |\nabla X^\varepsilon_\lambda - J_\lambda(\nabla X^\varepsilon_\lambda)|^2$, from (6.3.62) we have that

$$\int_Q |\nabla X^\varepsilon_\lambda - J_\lambda(\nabla X^\varepsilon_\lambda)|^2(\omega) \to 0 \quad \text{as } \lambda \searrow 0\,,$$

which together with (6.3.59) implies that $\nabla X^\varepsilon_{\lambda_n}(\omega) \rightharpoonup Y^\varepsilon$ in $L^2(Q)^d$; hence, we deduce

$$X^\varepsilon(\omega) \in L^p(0,T;V)\,,$$

$Y^\varepsilon = \nabla X^\varepsilon$ and as a consequence (possibly renominating $\{\lambda_n\}_{n\in\mathbb{N}}$)

$$J_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}(\omega)\right) \rightharpoonup \nabla X^\varepsilon(\omega) \quad \text{in } L^p(Q)^d\,, \tag{6.3.64}$$

$$\nabla X^\varepsilon_{\lambda_n}(\omega) \rightharpoonup \nabla X^\varepsilon(\omega) \quad \text{in } L^2(Q)^d\,. \tag{6.3.65}$$

The second step is to prove a strong convergence for X^ε_λ. To this purpose, equation (6.3.36) can be rewritten on the path starting from ω as

$$\partial_t\left(X^\varepsilon_\lambda - W^\varepsilon_B\right)(t) - \operatorname{div}\gamma_\lambda(\nabla X^\varepsilon_\lambda(t)) - \lambda\Delta X^\varepsilon_\lambda(t) + \beta_\lambda(X^\varepsilon_\lambda(t)) = 0 \quad \text{in } H^{-1}(D)$$

for every $t \in [0,T]$: we estimate the different terms of the previous relation in the larger space $L^1(0,T;V_0^*)$. Recalling that $L^1(D), H^{-1}(D), V^* \hookrightarrow V_0^*$, using the fact that $\|-\operatorname{div} v\|_{V^*} \leq \|v\|_{L^q(D)}$ for every $v \in L^q(D)^d$ (thanks to definition of divergence) and that $\|-\Delta v\|_{H^{-1}(D)} \leq \|\nabla v\|_{L^2(D)}$ for every $v \in H^1_0(D)$, using conditions (6.3.45)–(6.3.46) and (6.3.55), we deduce that for every $\lambda \in (0,1)$

$$
\begin{aligned}
\|-\operatorname{div}\gamma_\lambda(\nabla X^\varepsilon_\lambda(\omega))\|_{L^1(0,T;V_0^*)} \leq c\,\|\gamma_\lambda(\nabla X^\varepsilon_\lambda(\omega))\|_{L^q(Q)} \leq M_{\omega,\varepsilon}\,,\\
\|-\lambda\Delta X^\varepsilon_\lambda\|_{L^1(0,T;V_0^*)} \leq c\lambda\,\|\nabla X^\varepsilon_\lambda\|_{L^2(Q)} \leq M_{\omega,\varepsilon}\,,\\
\|\beta_\lambda(X^\varepsilon_\lambda(\omega))\|_{L^1(0,T;V_0^*)} \leq c\,\|\beta_\lambda(X^\varepsilon_\lambda(\omega))\|_{L^1(Q)} \leq M_{\omega.\varepsilon}\,,
\end{aligned}
$$

for a certain constant $c > 0$ and renominating the constant $M_{\omega,\varepsilon}$ at each passage. Hence, we deduce by difference that

$$\|\partial_t\left(X^\varepsilon_\lambda - W^\varepsilon_B\right)(\omega)\|_{L^1(0,T;V_0^*)} \leq M_{\omega,\varepsilon} \quad \text{for every } \lambda \in (0,1)\,. \tag{6.3.66}$$

At this point, we can recover a strong convergence using some classical compactness results with $\omega \in \Omega$ being fixed. Since by (6.3.65) the family $\{X^\varepsilon_{\lambda_n}(\omega)\}_{n\in\mathbb{N}}$ is bounded in $L^2(0,T;H^1_0(D))$, thanks also to (6.3.66) we can apply Lemma 1.4.3 to recover that the set F is relatively compact in $L^2(0,T;H)$. Hence, there exists $X^\varepsilon_B(\omega) \in L^2(0,T;H)$ such that

$$(X^\varepsilon_{\lambda_n} - W^\varepsilon_B)(\omega) \to X^\varepsilon_B(\omega) \quad \text{in } L^2(0,T;H) \quad \text{as } n \to \infty\,,$$

possibly updating the sequence $\{\lambda_n\}_{n\in\mathbb{N}}$. Using condition (6.3.58) and the fact that W^ε_B is fixed with respect to λ, we infer that

$$(X^\varepsilon_{\lambda_n} - W^\varepsilon_B)(\omega) \overset{*}{\rightharpoonup} (X^\varepsilon - W^\varepsilon_B)(\omega) \quad \text{in } L^\infty(0,T;H) \quad \text{as } n \to \infty\,,$$

and for uniqueness of the weak limit we have $X^\varepsilon_B(\omega) = (X^\varepsilon - W^\varepsilon_B)(\omega)$ a.e. in Q. As a consequence, we have that

$$X^\varepsilon_{\lambda_n}(\omega) \to X^\varepsilon(\omega) \quad \text{in } L^2(0,T;H) \quad \text{as } n \to \infty\,. \tag{6.3.67}$$

We are now ready to pass to the limit as $\lambda \searrow 0$ in (6.3.36): in particular, we are going to show that for every $\varepsilon \in (0,1)$ we have

$$X^\varepsilon(t) - \int_0^t \operatorname{div}\eta^\varepsilon(s)\,ds + \int_0^t \xi^\varepsilon(s)\,ds = X_0 + \int_0^t B^\varepsilon(s)\,dW_s \\ \text{in } V_0^*, \quad \forall\, t \in [0,T]\,, \quad \mathbb{P}\text{-a.s.}\,, \tag{6.3.68}$$

$$\eta^\varepsilon \in \gamma(\nabla X^\varepsilon) \quad \text{a.e. in } Q\,, \quad \mathbb{P}\text{-almost surely}\,, \tag{6.3.69}$$

$$\xi^\varepsilon \in \beta(X) \quad \text{a.e. in } Q\,, \quad \mathbb{P}\text{-almost surely}\,, \tag{6.3.70}$$

$$j(X^\varepsilon) + j^*(\xi^\varepsilon) \in L^1(Q)\,, \quad \mathbb{P}\text{-almost surely}\,. \tag{6.3.71}$$

Firstly, let $\varepsilon \in (0,1)$ and $\omega \in \Omega$ be fixed as usual. Let $w \in V_0$ and recall the fact that $V_0 \hookrightarrow L^\infty(D) \cap V$: then, thanks to (6.3.58), (6.3.60), (6.3.63) and (6.3.61), for almost every $t \in (0,T)$ we have

$$\int_D X^\varepsilon_{\lambda_n}(t)w \to \int_D X^\varepsilon(t)w\,,$$

$$\int_0^t\int_D \gamma_{\lambda_n}(\nabla X^\varepsilon_{\lambda_n}(s)) \cdot \nabla w\,ds \to \int_0^t\int_D \eta^\varepsilon(s) \cdot \nabla w\,ds\,,$$

$$\lambda_n \int_0^t\int_D \nabla X^\varepsilon_{\lambda_n}(s) \cdot \nabla w\,ds \to 0\,,$$

$$\int_0^t\int_D \beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(s))w\,ds \to \int_0^t\int_D \xi^\varepsilon(s)w\,ds\,,$$

as $n \to \infty$. Hence, taking these remarks into account, letting $n \to \infty$ in equation (6.3.36) evaluated with λ_n, we obtain exactly

$$X^\varepsilon(t) - \int_0^t \operatorname{div}\eta^\varepsilon(s)\,ds + \int_0^t \xi^\varepsilon(s)\,ds = X_0 + \int_0^t B^\varepsilon(s)\,dW_s \qquad \text{in } V_0^*$$
$$\text{for almost every } t\in(0,T)\,, \quad \mathbb{P}\text{-almost surely}\,.$$

Since all the terms except the first are continuous with respect to time, we deduce a posteriori that $X^\varepsilon(\omega)\in C([0,T];V_0^*)$ $\mathbb{P}$-almost surely. Since also $X^\varepsilon(\omega)\in L^\infty(0,T;H)$, by Lemma 1.4.1 we deduce that

$$X^\varepsilon \in C_w([0,T];H) \quad \mathbb{P}\text{-almost surely}\,. \tag{6.3.72}$$

Hence, the last integral relation holds for every $t\in[0,T]$ and (6.3.68) is proved.

Secondly, let us show (6.3.70). By (6.3.67) we can assume that $X^\varepsilon_{\lambda_n}(\omega)\to X^\varepsilon(\omega)$ a.e. in Q as $n\to\infty$, from which, since R_{λ_n} is a contraction, we deduce also that $R_{\lambda_n}X^\varepsilon_{\lambda_n}(\omega)\to X^\varepsilon(\omega)$ a.e. in Q. Moreover, by (6.3.61), we also know that $\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega))\in\beta(R_{\lambda_n}X^\varepsilon_{\lambda_n}(\omega))$ and $\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega))\rightharpoonup\xi^\varepsilon(\omega)$ in $L^1(Q)$. Consequently, since $\{\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega))X^\varepsilon_{\lambda_n}(\omega)\}_{n\in\mathbb{N}}$ is bounded in $L^1(Q)$ thanks to (6.3.54), we can apply Lemma 1.3.14, with the choices $Y=Q$, μ the Lebesgue measure on Q, $y_n=X^\varepsilon_{\lambda_n}$ and $g_n=R_{\lambda_n}X^\varepsilon_{\lambda_n}$, to infer (6.3.70).

Furthermore, by definition of β_{λ_n} we have $X^\varepsilon - R_{\lambda_n}X^\varepsilon_{\lambda_n} = (X^\varepsilon - X^\varepsilon_{\lambda_n}) + \lambda_n\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})$, so that thanks to (6.3.61) and (6.3.67) we deduce that $R_{\lambda_n}X^\varepsilon_{\lambda_n}(\omega)\to X^\varepsilon(\omega)$ in $L^1(Q)$: hence, by the weak lower semicontinuity of the convex integrals and conditions (6.3.61) and (6.3.54), we have that

$$\begin{aligned}
&\int_Q \left[j(X^\varepsilon(\omega)) + j^*(\xi^\varepsilon(\omega))\right] \le \liminf_{n\to\infty}\int_Q \left[j(R_{\lambda_n}X^\varepsilon_{\lambda_n}(\omega)) + j^*(\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})(\omega))\right]\\
&= \liminf_{n\to\infty}\int_Q R_{\lambda_n}X^\varepsilon_{\lambda_n}(\omega)\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega)) \le \liminf_{n\to\infty}\int_Q X^\varepsilon_{\lambda_n}(\omega)\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega)) \le M_{\omega,\varepsilon}\,,
\end{aligned}$$

so that also (6.3.71) is proved. Let us also point out that condition (6.3.70) implies $\xi^\varepsilon X^\varepsilon = j(X^\varepsilon)+j^*(\xi^\varepsilon)$ almost everywhere on Q, so that from the very last calculations, using the fact that R_{λ_n} is a contraction and the monotonicity of β_λ, we have

$$\xi^\varepsilon(\omega)X^\varepsilon(\omega)\in L^1(Q)\,, \qquad \int_Q \xi^\varepsilon(\omega)X^\varepsilon(\omega) \le \liminf_{n\to\infty}\int_Q \beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega))X^\varepsilon_{\lambda_n}(\omega)\,. \tag{6.3.73}$$

Finally, let us show that (6.3.69) holds: in the next passages, we will omit to write ω to simplify notations. From equation (6.3.40) evaluated at time T, recalling conditions (6.3.58), (6.3.60), (6.3.61), (6.3.63), (6.3.73) and (6.3.38), we get that

$$\begin{aligned}
&\limsup_{n\to\infty}\int_Q \gamma_{\lambda_n}(\nabla X^\varepsilon_{\lambda_n})\cdot\nabla X^\varepsilon_{\lambda_n} = \frac12\left\|X_0\right\|_H^2 + \lim_{n\to\infty}\int_Q \gamma_{\lambda_n}(\nabla X^\varepsilon_{\lambda_n})\cdot\nabla W^\varepsilon_B\\
&+ \lim_{n\to\infty}\lambda_n\int_Q \nabla X^\varepsilon_{\lambda_n}\cdot\nabla W^\varepsilon_B + \lim_{n\to\infty}\int_Q \beta_{\lambda_n}(X^\varepsilon_{\lambda_n})W^\varepsilon_B\\
&- \frac12\liminf_{n\to\infty}\left\|X^\varepsilon_{\lambda_n}(T) - W^\varepsilon_B(T)\right\|_H^2 - \lim_{n\to\infty}\lambda_n\left\|\nabla X^\varepsilon_{\lambda_n}\right\|_H^2 - \liminf_{n\to\infty}\int_Q \beta_{\lambda_n}(X^\varepsilon_{\lambda_n})X^\varepsilon_{\lambda_n}\\
&\le \frac12\left\|X_0\right\|_H^2 + \int_Q \eta^\varepsilon\cdot\nabla W^\varepsilon_B + \int_Q \xi^\varepsilon W^\varepsilon_B - \frac12\left\|X^\varepsilon(T) - W^\varepsilon_B(T)\right\|_H^2 - \int_Q \xi^\varepsilon X^\varepsilon\,.
\end{aligned}$$

At this point, thanks to conditions (6.3.68)–(6.3.71), we can prove that the following testing formula holds:

$$\frac{1}{2}\left\|X^\varepsilon(T)-W_B^\varepsilon(T)\right\|_H^2+\int_Q \eta^\varepsilon\cdot\nabla(X^\varepsilon-W_B^\varepsilon)+\int_Q \xi^\varepsilon(X^\varepsilon-W_B^\varepsilon)=\frac{1}{2}\left\|X_0\right\|_H^2\,. \tag{6.3.74}$$

Remark 6.3.2. The proof of (6.3.74) relies on sharp approximations of elliptic type and is very technical: hence, we omit it here in order not to make the treatment heavier. The reader can refer to Section 6.6 for a complete and rigorous proof of (6.3.74).

Hence, thanks to (6.3.74), the last set of inequalities can be read as

$$\limsup_{n\to\infty}\int_Q \gamma_{\lambda_n}(\nabla X_{\lambda_n}^\varepsilon)\cdot\nabla X_{\lambda_n}^\varepsilon\le\int_Q \eta^\varepsilon\cdot\nabla X^\varepsilon\,,$$

from which, using the definition of γ_{λ_n} and condition (6.3.62) we deduce that

$$\begin{aligned}
&\limsup_{n\to\infty}\int_Q \gamma_{\lambda_n}(\nabla X_{\lambda_n}^\varepsilon)\cdot J_{\lambda_n}\left(\nabla X_{\lambda_n}^\varepsilon\right)\\
&=\limsup_{n\to\infty}\int_Q\left[\gamma_{\lambda_n}(\nabla X_{\lambda_n}^\varepsilon)\cdot\nabla X_{\lambda_n}^\varepsilon-\lambda_n|\gamma_{\lambda_n}(\nabla X_{\lambda_n}^\varepsilon)|^2\right]\\
&=\limsup_{n\to\infty}\int_Q \gamma_{\lambda_n}(\nabla X_{\lambda_n}^\varepsilon)\cdot\nabla X_{\lambda_n}^\varepsilon-\lim_{n\to\infty}\lambda_n\int_Q\left|\gamma_{\lambda_n}(\nabla X_{\lambda_n}^\varepsilon)\right|^2\le\int_Q \eta^\varepsilon\cdot\nabla X^\varepsilon\,.
\end{aligned}$$

This last inequality together with (6.3.59) and (6.3.60) implies condition (6.3.69) thanks to the usual tools of monotone analysis.

6.3.6 Measurability properties of the solutions

In this section, we show that the solution components X^ε, η^ε and ξ^ε constructed in the previous section have also some regularity with respect to ω. Moreover, we prove uniform estimates with respect to ε: to this purpose, we will use the results of Sections 6.3.3 and 6.3.4, as well as natural lower semicontinuity properties.

First of all, note that, a priori, X^ε, η^ε and ξ^ε are not even measurable processes, because of the way they have been build (the sequence λ_n could depend on ω as well). To show measurability, we need to prove uniqueness for problem (6.3.68)–(6.3.71). Hence, let $(X_1^\varepsilon,\eta_1^\varepsilon,\xi_1^\varepsilon)$ and $(X_2^\varepsilon,\eta_2^\varepsilon,\xi_2^\varepsilon)$ satisfy conditions (6.3.68)–(6.3.71): taking the difference of (6.3.68) and setting $Y^\varepsilon:=X_1^\varepsilon-X_2^\varepsilon$, $\zeta^\varepsilon:=\eta_1^\varepsilon-\eta_2^\varepsilon$ and $\psi^\varepsilon:=\xi_1^\varepsilon-\xi_2^\varepsilon$ we have

$$Y^\varepsilon(t)-\int_0^t \operatorname{div}\zeta^\varepsilon(s)\,ds+\int_0^t\psi^\varepsilon(s)\,ds=0\quad\text{for every } t\in[0,T]\,,\quad\mathbb{P}\text{-a.s.}$$

By convexity we have $j(Y^\varepsilon/2)+j^*(\psi^\varepsilon/2)\le\frac{1}{2}\left(j(X_1^\varepsilon)+j(X_2^\varepsilon)+j^*(\xi_1^\varepsilon)+j^*(\xi_2^\varepsilon)\right)$, where the right-hand side is in $L^1(Q)$: hence, using the same argument as in Section 6.6 with $X_0=0$ and $B=0$, we infer that

$$\frac{1}{2}\left\|Y^\varepsilon(t)\right\|_H^2+\int_0^t\int_D\zeta^\varepsilon(s)\cdot\nabla Y^\varepsilon(s)\,ds+\int_0^t\int_D\psi^\varepsilon(s)Y^\varepsilon(s)\,ds=0\,.$$

The monotonicity of γ and β implies that $Y^\varepsilon=0$. Moreover, in view of (6.2.6), γ is a continuous function. This implies that $\zeta^\varepsilon=0$ and the first integral expression becomes $\int_0^t\psi^\varepsilon(s)\,ds=0$ for every $t\in[0,T]$, so that also $\psi^\varepsilon=0$ and uniqueness is proved.

At this point, we are ready to prove that the sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ constructed in the previous section can be chosen independent of ω: more precisely, we can prove that for any sequence

$\{\lambda_n\}_{n\in\mathbb{N}}$ decreasing to 0, conditions (6.3.58)–(6.3.61) and (6.3.64)–(6.3.65) hold. Indeed, let $\{\lambda_n\}_{n\in\mathbb{N}}$ be any sequence decreasing to 0 and fix $\omega \in \Omega$: then, for every subsequence of $\{\lambda_n\}_{n\in\mathbb{N}}$ (which we still denote with the same symbol for sake of simplicity), the estimates (6.3.42)–(6.3.46) hold. Proceeding as in Section 6.3.5 and invoking the uniqueness, we can then extract a further sub-subsequence (depending on ω) along which the same weak convergences to X^ε, η^ε and ξ^ε hold. This implies that the convergences (6.3.58)–(6.3.61) and (6.3.64)–(6.3.65) are true for the original sequence $\{\lambda_n\}_{n\in\mathbb{N}}$, which does not depend on ω.

Now, let us prove some measurability properties of the processes X^ε, η^ε and ξ^ε. First of all, since $X^\varepsilon_{\lambda_n} \to X^\varepsilon$ in $L^2(0,T;H)$ $\mathbb{P}$-almost surely, it is clear that X^ε is predictable (since so are $X^\varepsilon_{\lambda_n}$ for every $n \in \mathbb{N}$). Secondly, let us focus on ξ^ε: we prove that $\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}) \rightharpoonup \xi^\varepsilon$ in $L^1(\Omega\times(0,T)\times D)$. To this aim, for any $g \in L^\infty(Q)$, setting

$$F^\varepsilon_{\lambda_n} := \int_Q \beta_{\lambda_n}(X^\varepsilon_{\lambda_n})g\,, \qquad F^\varepsilon := \int_Q \xi^\varepsilon g\,,$$

we know that $F^\varepsilon_{\lambda_n} \to F^\varepsilon$ $\mathbb{P}$-almost surely: let us show that $F^\varepsilon_{\lambda_n} \rightharpoonup F^\varepsilon$ in $L^1(\Omega)$. Indeed, for any $h \in L^\infty(\Omega)$, if we define

$$j_0^*(\cdot) := j^*\left(\cdot/M\right)\,, \quad M := \frac{1}{(1\vee\|g\|_{L^\infty(Q)})(1\vee\|h\|_{L^\infty(\Omega)})}\,,$$

by the Jensen inequality we have that

$$\begin{aligned}\mathbb{E}\left[j_0^*(F^\varepsilon_{\lambda_n}h)\right] &= \mathbb{E}\left[j_0^*\left(\int_Q \beta_{\lambda_n}(X^\varepsilon_{\lambda_n})gh\right)\right]\\ &\leq C_{T,|D|}\mathbb{E}\int_Q j_0^*(\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})gh) \leq \int_{\Omega\times Q} j^*(|\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})|)\,,\end{aligned}$$

where the last term is bounded uniformly in n by (6.3.56). Consequently, since j_0^* is still superlinear at infinity, by the de la Vallée-Poussin criterion, we deduce that $\{F^\varepsilon_{\lambda_n}h\}_{n\in\mathbb{N}}$ is uniformly integrable on Ω: taking also into account that $F^\varepsilon_{\lambda_n}h \to F^\varepsilon h$ $\mathbb{P}$-almost surely, Vitali's convergence theorem ensures that $F^\varepsilon_{\lambda_n}h \to F^\varepsilon h$ in $L^1(\Omega)$. Since this is true for any h and g, this implies that $\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}) \rightharpoonup \xi^\varepsilon$ in $L^1(\Omega\times(0,T)\times D)$. By Mazur's Lemma there is a sequence made up of convex combinations of $\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})$ which converge strongly ξ^ε in $L^1(Q)$, $\mathbb{P}$-almost surely. This ensures that ξ^ε is predictable (since so are $\beta_\lambda(X^\varepsilon_\lambda)$ for every n). Finally, using a similar argument, one can show also that η^ε is adapted.

It is now time to prove some uniform estimates with respect to ε. By (6.3.58)–(6.3.61), (6.3.64) and the estimates (6.3.49)–(6.3.53), using the lower semicontinuity of the norm, we have

$$\begin{aligned}\|X^\varepsilon(\omega)\|_{L^\infty(0,T;H)} &\leq \liminf_{n\to\infty}\left\|X^\varepsilon_{\lambda_n}(\omega)\right\|_{L^\infty(0,T;H)}\,,\\ \|\nabla X^\varepsilon(\omega)\|_{L^p(Q)} &\leq \liminf_{n\to\infty}\left\|J_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}(\omega)\right)\right\|_{L^p(Q)}\,,\\ \|\eta^\varepsilon(\omega)\|_{L^q(Q)} &\leq \liminf_{n\to\infty}\left\|\gamma_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}(\omega)\right)\right\|_{L^q(Q)}\,,\\ \|\xi^\varepsilon(\omega)\|_{L^1(Q)} &\leq \liminf_{n\to\infty}\left\|\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}(\omega))\right\|_{L^1(Q)}\,.\end{aligned}$$

Taking expectations and using (6.3.49)–(6.3.53) and (6.3.57), the Fatou's lemma implies

$$\begin{aligned}
\mathbb{E}\left\|X^\varepsilon\right\|^2_{L^\infty(0,T;H)} &\leq \liminf_{n\to\infty}\left\|X^\varepsilon_{\lambda_n}\right\|^2_{L^2(\Omega;L^\infty(0,T;H))} \leq N\,,\\
\mathbb{E}\left\|\nabla X^\varepsilon\right\|^p_{L^p(Q)} &\leq \liminf_{n\to\infty}\left\|J_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}\right)\right\|^p_{L^p(\Omega\times Q)} \leq N\,,\\
\mathbb{E}\left\|\eta^\varepsilon\right\|^q_{L^q(Q)} &\leq \liminf_{n\to\infty}\left\|\gamma_{\lambda_n}\left(\nabla X^\varepsilon_{\lambda_n}\right)\right\|^q_{L^q(\Omega\times Q)} \leq N\,,\\
\mathbb{E}\left\|\xi^\varepsilon\right\|_{L^1(Q)} &\leq \liminf_{n\to\infty}\left\|\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})\right\|_{L^1(\Omega\times Q)} \leq N\,,
\end{aligned}$$

for a certain positive constant N independent of ε. Hence, we have also proved that

$$X^\varepsilon \in L^2\left(\Omega; L^\infty(0,T;H)\right) \cap L^p\left(\Omega\times(0,T);V\right)\,, \tag{6.3.75}$$

$$\eta^\varepsilon \in L^q\left(\Omega\times(0,T)\times D\right)^d\,, \quad \xi^\varepsilon \in L^1\left(\Omega\times(0,T)\times D\right) \tag{6.3.76}$$

and that the following estimates hold:

$$\|X^\varepsilon\|_{L^2(\Omega;L^\infty(0,T;H))\cap L^p(\Omega\times(0,T);V)} \leq N \quad \text{for every } \varepsilon\in(0,1)\,, \tag{6.3.77}$$

$$\|\eta^\varepsilon\|_{L^q(\Omega\times(0,T)\times D)} \leq N \quad \text{for every } \varepsilon\in(0,1)\,, \tag{6.3.78}$$

$$\|\xi^\varepsilon\|_{L^1(\Omega\times(0,T)\times D)} \leq N \quad \text{for every } \varepsilon\in(0,1)\,. \tag{6.3.79}$$

Moreover, since $\beta_{\lambda_n}(X^\varepsilon_{\lambda_n}) \rightharpoonup \xi^\varepsilon$ in $L^1(Q)$ as $n\to\infty$, $\mathbb{P}$-almost surely, by the weak lower semicontinuity of the convex integral we have

$$\int_Q j^*(\xi^\varepsilon) \leq \liminf_{n\to\infty}\int_Q j^*\left(\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})\right) \quad \mathbb{P}\text{-almost surely :}$$

hence, thanks to the Fatou lemma and condition (6.3.56), we deduce that

$$\int_{\Omega\times Q} j^*(\xi^\varepsilon) \leq \liminf_{n\to\infty}\int_{\Omega\times Q} j^*\left(\beta_{\lambda_n}(X^\varepsilon_{\lambda_n})\right) \leq N\,,$$

where N is independent of ε. Consequently, since j^* is even, we have that $\{j^*(\xi^\varepsilon)\}_{\varepsilon\in(0,1)}$ is bounded in $L^1(\Omega\times Q)$: hence, since j^* is superlinear at ∞, the classical results by de la Vallée-Poussin and the Dunford-Pettis theorem ensure that

$$\{\xi^\varepsilon\}_{\varepsilon\in(0,1)} \quad \text{is weakly relatively compact in } L^1(\Omega\times(0,T)\times D)\,. \tag{6.3.80}$$

Similarly, $R_{\lambda_n}X^\varepsilon_{\lambda_n} \to X^\varepsilon$ in $L^1(Q)$ and $j(R_\lambda X^\varepsilon_\lambda) \leq j(R_\lambda X^\varepsilon_\lambda) + j^*(\beta_\lambda(X^\varepsilon_\lambda)) = \beta_\lambda(X^\varepsilon_\lambda)X^\varepsilon_\lambda$: hence, the weak lower semicontinuity of the convex integrals, Fatou's lemma and condition (6.3.56) imply

$$\int_{\Omega\times Q} j(X^\varepsilon) \leq \liminf_{n\to\infty}\int_{\Omega\times Q} j\left(R_{\lambda_n}X^\varepsilon_{\lambda_n}\right) \leq \sup_{\varepsilon,\lambda\in(0,1)}\|\beta_\lambda(X^\varepsilon_\lambda)X^\varepsilon_\lambda\|_{L^1(\Omega\times Q)} \leq N\,.$$

Taking these remarks into account, we have also obtained that

$$\|j(X^\varepsilon)\|_{L^1(\Omega\times(0,T)\times D)} + \|j^*(\xi^\varepsilon)\|_{L^1(\Omega\times(0,T)\times D)} \leq N \quad \text{for every } \varepsilon\in(0,1)\,. \tag{6.3.81}$$

6.3.7 Passage to the limit as $\varepsilon \searrow 0$

In this section, we pass to the limit as $\varepsilon \searrow 0$ in the sub-prolem (6.3.68)–(6.3.71) and we recover global solutions to the original problem: to this end, the passage to the limit takes place also in probability, as we have already anticipated.

First of all, thanks to (6.3.77)–(6.3.79), we deduce that there exist

$$X \in L^\infty\left(0,T;L^2(\Omega;H)\right) \cap L^p\left(\Omega\times(0,T);V\right) ,$$
$$\eta \in L^q\left(\Omega\times(0,T)\times D\right)^d , \qquad \xi \in L^1\left(\Omega\times(0,T)\times D\right) ,$$

and a sequence $\{\varepsilon_n\}_{n\in\mathbb{N}}$ with $\varepsilon_n \searrow 0$ as $n\to\infty$ such that

$$X^{\varepsilon_n} \overset{*}{\rightharpoonup} X \quad \text{in } L^\infty\left(0,T;L^2(\Omega;H)\right) , \tag{6.3.82}$$
$$X^{\varepsilon_n} \rightharpoonup X \quad \text{in } L^p\left(\Omega\times(0,T);V\right)) , \tag{6.3.83}$$
$$\eta^{\varepsilon_n} \rightharpoonup \eta \quad \text{in } L^q\left(\Omega\times(0,T)\times D\right)^d , \tag{6.3.84}$$
$$\xi^{\varepsilon_n} \rightharpoonup \xi \quad \text{in } L^1\left(\Omega\times(0,T)\times D\right) . \tag{6.3.85}$$

Let us prove a strong convergence for X^ε: given $\varepsilon,\delta\in(0,1)$, consider equation (6.3.68) evaluated for ε and δ. Then, taking the difference we have

$$\begin{aligned} X^\varepsilon(t) - X^\delta(t) - \int_0^t \operatorname{div}(\eta^\varepsilon(s)-\eta^\delta(s))\,ds + \int_0^t \left(\xi^\varepsilon(s)-\xi^\delta(s)\right)\,ds& \\ = \int_0^t \left(B^\varepsilon(s)-B^\delta(s)\right)dW_s \quad \text{in } V_0^* \quad \text{for every } t\in[0,T]\,, \quad \mathbb{P}\text{-a.s}\,.& \end{aligned}$$

Now, notice that thanks to the symmetry and the convexity of j and j^*, we have

$$j\left(\frac{X^\varepsilon - X^\delta}{2}\right) + j^*\left(\frac{\xi^\varepsilon-\xi^\delta}{2}\right) \le \frac12\left(j(X^\varepsilon)+j(X^\delta)+j^*(\xi^\varepsilon)+j^*(\xi^\delta)\right) ,$$

where the term on the right hand side is in $L^1(\Omega\times(0,T)\times D)$ thanks to (6.3.81): hence, recalling also condition (6.3.72) we can apply Proposition 6.7.1 with the choices $Y = X^\varepsilon - X^\delta$, $f=\eta^\varepsilon-\eta^\delta$, $g=\xi^\varepsilon-\xi^\delta$, $T=B^\varepsilon-B^\delta$ and $\alpha=1/2$ to infer that

$$\begin{aligned} \frac12\left\|X^\varepsilon(t)-X^\delta(t)\right\|_H^2 &+ \int_0^t\int_D \left(\eta^\varepsilon(s)-\eta^\delta(s)\right)\cdot\left(\nabla X^\varepsilon(s)-\nabla X^\delta(s)\right)\,ds \\ &+ \int_0^t\int_D \left(\xi^\varepsilon(s)-\xi^\delta(s)\right)\left(X^\varepsilon(s)-X^\delta(s)\right)\,ds \\ = \frac12\int_0^t \left\|B^\varepsilon(s)-B^\delta(s)\right\|^2_{\mathscr{L}^2(U,H)}\,ds &+ \int_0^t \left((X^\varepsilon-X^\delta)(s),(B^\varepsilon-B^\delta)(s)\,dW_s\right) \end{aligned}$$

for every $t\in[0,T]$, $\mathbb{P}$-almost surely. Now, proceeding exactly as in Section 6.3.3, we take the supremum in t and expectations, use the monotonicity of γ and β together with (6.3.69)–(6.3.70) and the Davis inequality, so that we have

$$\left\|X^\varepsilon - X^\delta\right\|^2_{L^2(\Omega;L^\infty(0,T;H))} \le c\left\|B^\varepsilon-B^\delta\right\|^2_{L^2(\Omega\times(0,T);\mathscr{L}^2(U,H))}$$

for every $\varepsilon,\delta\in(0,1)$, for a positive constant c independent of ε: taking into account (6.3.29), this implies that the sequence $\{X^\varepsilon\}_{\varepsilon\in(0,1)}$ is Cauchy in the space $L^2(\Omega;L^\infty(0,T;H))$, so that by (6.3.82) we deduce

$$X \in L^2\left(\Omega;L^\infty(0,T;H)\right) \tag{6.3.86}$$

and

$$X^\varepsilon \to X \quad \text{in } L^2\left(\Omega;L^\infty(0,T;H)\right) , \quad \text{as } \varepsilon\searrow 0\,. \tag{6.3.87}$$

We are now ready to pass to the limit in equation (6.3.68): to this purpose, fix $w \in V_0$ (recall that $V_0 \hookrightarrow L^\infty(D) \cap V$). Then, thanks to (6.3.87), (6.3.83)–(6.3.85) and (6.3.29), for every $t \in [0,T]$ we have as $n \to \infty$ that

$$\begin{gathered}
\mathbb{E}\left[\operatorname{ess}\sup_{t\in(0,T)}\left|\int_D X^{\varepsilon_n}(t)w - \int_D X(t)w\right|\right] \to 0\,, \\
\mathbb{E}\left[\int_0^t\int_D \eta^{\varepsilon_n}\cdot\nabla w\,ds\right] \to \mathbb{E}\left[\int_0^t\int_D \eta\cdot\nabla w\,ds\right]\,, \\
\mathbb{E}\left[\int_0^t\int_D \xi^{\varepsilon_n}(s)w\,ds\right] \to \mathbb{E}\left[\int_0^t\int_D \xi(s)w\,ds\right]\,, \\
\mathbb{E}\left[\int_0^t (w, B^{\varepsilon_n}(s)\,dW_s)\right] \to \mathbb{E}\left[\int_0^t (w, B(s)\,dW_s)\right]\,,
\end{gathered}$$

so that evaluating (6.3.68) with ε_n and letting $n \to \infty$, we deduce

$$X(t) - \int_0^t \operatorname{div}\eta(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s)\,dW_s \qquad \text{in } V_0^*\,,$$
$$\text{for almost every } t \in (0,T)\,, \quad \mathbb{P}\text{-almost surely}\,.$$

Since all the terms except the first have $\mathbb{P}$-almost surely continuous paths in V_0^*, we have *a posteriori* that $X \in C([0,T];V_0^*)$ $\mathbb{P}$-almost surely. Moreover, it is not difficult to check that the fact that $X^\varepsilon \in C_w([0,T];H)$ for every ε together with (6.3.87) readily implies

$$X \in C_w([0,T];H) \quad \mathbb{P}\text{-almost surely}\,, \tag{6.3.88}$$

so that the integral relation holds for every $t \in [0,T]$ and (6.2.7)–(6.2.10) are proved. Furthermore, for every $t \in [0,T]$ and $\mathbb{P}$-almost surely, all the terms in (6.2.10) except $\int_0^t \eta(s)\,ds$ are in $L^1(D)$ and all the terms except $\int_0^t \xi(s)\,ds$ are in V^*, so that by difference the integral relation holds in $L^1(D) \cap V^*$.

At this point, let us focus on (6.2.12) and (6.2.13). By (6.3.87), we may assume that $X^{\varepsilon_n} \to X$ almost everywhere in $\Omega \times Q$; moreover, by (6.3.70) and (6.3.81) we have

$$\int_{\Omega\times Q} \xi^\varepsilon X^\varepsilon = \int_{\Omega\times Q} (j(X^\varepsilon) + j^*(\xi^\varepsilon)) \leq N\,,$$

where $N > 0$ is independent of ε. Hence, $\{\xi^\varepsilon X^\varepsilon\}_{\varepsilon\in(0,1)}$ is bounded in $L^1(\Omega \times Q)$, and recalling also (6.3.85) we can apply Lemma 1.3.14, with the choices $Y = \Omega \times Q$, $\mu = \mathbb{P} \otimes \mathrm{Leb}_Q$, $y_n = X^{\varepsilon_n}$ and $g_n = \xi^{\varepsilon_n}$, to infer that (6.2.12) holds. Moreover, thanks to conditions (6.3.87), (6.3.85) and (6.3.81), using the weak lower semicontinuity of the convex integrals we have that

$$\int_{\Omega\times Q} (j(X) + j^*(\xi)) \leq \liminf_{n\to\infty} \int_{\Omega\times Q} (j(X^{\varepsilon_n}) + j^*(\xi^{\varepsilon_n})) \leq N\,,$$

so that (6.2.13) is proved. Let us also point out that from the last inequality, thanks to (6.2.12) and (6.3.70) we obtain

$$\int_{\Omega\times Q} \xi X \leq \liminf_{n\to\infty} \int_{\Omega\times Q} \xi^{\varepsilon_n} X^{\varepsilon_n}\,. \tag{6.3.89}$$

The next thing that we need to prove is condition (6.2.11). To this end, thanks to the regularities that we have found on the solutions, we can apply Proposition 6.7.1 to infer that for every $t \in [0,T]$

$$\frac{1}{2}\left\|X^{\varepsilon_n}(t)\right\|^2_{L^2(\Omega;H)} + \int_0^t\int_{\Omega\times D}\eta^{\varepsilon_n}(s)\cdot\nabla X^{\varepsilon_n}(s)\,ds + \int_0^t\int_{\Omega\times D}\xi^{\varepsilon_n}(s)X^{\varepsilon_n}(s)\,ds$$
$$= \frac{1}{2}\left\|X_0\right\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^t\left\|B^{\varepsilon_n}(s)\right\|^2_{L^2(\Omega;\mathscr{L}^2(U,H))}\,ds\,,$$

from which, thanks to (6.3.87), (6.3.89) and (6.3.30), we have $\mathbb{P}$-almost surely that

$$\begin{aligned}\limsup_{n\to\infty}\int_{\Omega\times Q}\eta^{\varepsilon_n}\cdot\nabla X^{\varepsilon_n} &= \frac{1}{2}\left\|X_0\right\|^2_{L^2(\Omega;H)} + \frac{1}{2}\lim_{n\to\infty}\left\|B^{\varepsilon_n}\right\|^2_{L^2(\Omega\times(0,T);\mathscr{L}^2(U,H))}\\ &\quad - \frac{1}{2}\liminf_{n\to\infty}\left\|X^{\varepsilon_n}(T)\right\|^2_{L^2(\Omega;H)} - \liminf_{n\to\infty}\int_{\Omega\times Q}\xi^{\varepsilon_n}X^{\varepsilon_n}\\ &\leq \frac{1}{2}\left\|X_0\right\|^2_{L^2(\Omega;H)} + \frac{1}{2}\left\|B\right\|^2_{L^2(\Omega\times(0,T);\mathscr{L}^2(U,H))} - \frac{1}{2}\left\|X(T)\right\|^2_{L^2(\Omega;H)} - \int_{\Omega\times Q}\xi X\,.\end{aligned}$$

Now, we apply a second time Proposition 6.7.1 with the choices $Y = X$, $f = \eta$, $g = \xi$ and $T = B$: hence, the right hand side of the last set of inequality is exactly $\int_{\Omega\times Q}\eta\cdot\nabla X$, so that we have

$$\limsup_{n\to\infty}\int_{\Omega\times Q}\eta^{\varepsilon_n}\cdot\nabla X^{\varepsilon_n} \leq \int_{\Omega\times Q}\eta\cdot\nabla X\,.$$

This condition together with (6.3.83)–(6.3.84) and (6.3.69) implies exactly (6.2.11).

Finally, let us show that X and ξ are predictable processes, and η is adapted. At the end of Section 6.3.6 we checked that X^ε and ξ^ε are predictable, and η^ε is adapted, for every $\varepsilon \in (0,1)$. Now, from (6.3.87) it immediately follows that also X is predictable. Moreover, by conditions (6.3.84)–(6.3.85) and Mazur's Lemma we can recover strong convergences for some suitable convex combinations of $\{\eta^{\varepsilon_n}\}$ and $\{\xi^{\varepsilon_n}\}$: since these are still adapted and predicable, respectively, we can easily infer that η is adapted and ξ is predictable. This completes the proof.

6.3.8 The further existence result

In this section we prove the last part of Theorem 6.2.2, in which condition (6.2.6) is not assumed anymore. The idea is to to pass to the limit in a different way, using only the estimates in expectations and avoiding the pathwise arguments.

For any $\lambda \in (0,1)$, consider the approximated problem

$$dX_\lambda - \operatorname{div}\gamma_\lambda(\nabla X_\lambda)\,dt - \lambda\Delta X_\lambda\,dt + \beta_\lambda(X_\lambda)\,dt \ni B\,dW_t\ :$$

the classical variational approach in the Gelfand triple $H^1_0(D) \hookrightarrow H \hookrightarrow H^{-1}(D)$ ensures the existence of the approximated solutions

$$X_\lambda \in L^2\left(\Omega; C([0,T];H)\right) \cap L^2\left(\Omega\times(0,T);H^1_0(D)\right)\,.$$

Using Itô's formula and proceeding as in Sections 6.3.3 and 6.3.4, it is not difficult to prove that there exist a positive constant N, independent of λ, such that

$$\|X_\lambda\|_{L^2(\Omega;L^\infty(0,T;H))} \le N\,, \qquad \|J_\lambda(\nabla X_\lambda)\|_{L^p(\Omega\times(0,T)\times D)} \le N\,,$$
$$\|\gamma_\lambda(\nabla X_\lambda)\|_{L^q(\Omega\times(0,T)\times D)} \le N\,,$$
$$\{\beta_\lambda(X_\lambda)\}_{\lambda\in(0,1)} \quad \text{is weakly relatively compact in } L^1(\Omega\times(0,T)\times D)\,,$$
$$\|j(X_\lambda)\|_{L^1(\Omega\times(0,T)\times D)} + \|j^*(\beta_\lambda(X_\lambda))\|_{L^1(\Omega\times(0,T)\times D)} \le N\,,$$
$$\lambda^{1/2}\,\|\nabla X_\lambda\|_{L^2(\Omega\times(0,T);H)} \le N\,,$$
$$\lambda^{1/2}\,\|\gamma_\lambda(\nabla X_\lambda)\|_{L^2(\Omega\times(0,T)\times D)} \le N\,.$$

We deduce that there exist

$$X \in L^\infty\left(0,T;L^2(\Omega;H)\right) \cap L^p\left(\Omega\times(0,T);V\right)\,,$$
$$\eta \in L^q\left(\Omega\times(0,T)\times D\right)^d\,, \qquad \xi \in L^1\left(\Omega\times(0,T)\times D\right)\,,$$

and a sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ decreasing to 0 such that, as $n\to\infty$,

$$X_{\lambda_n} \overset{*}{\rightharpoonup} X \quad \text{in } L^\infty\left(0,T;L^2(\Omega;H)\right)\,,$$
$$J_{\lambda_n}(\nabla X_{\lambda_n}) \rightharpoonup \nabla X \quad \text{in } L^p\left(\Omega\times(0,T)\times D\right)^d\,,$$
$$\gamma_{\lambda_n}(\nabla X_{\lambda_n}) \rightharpoonup \eta \quad \text{in } L^q\left(\Omega\times(0,T)\times D\right)^d\,,$$
$$\beta_{\lambda_n}(X_{\lambda_n}) \rightharpoonup \xi \quad \text{in } L^1\left(\Omega\times(0,T)\times D\right)\,.$$

Fix $w \in L^\infty(\Omega;V_0)$: then, since the four last convergences imply that $X_{\lambda_n}(t) \rightharpoonup X(t)$ in $L^2(\Omega;H)$ for almost every $t\in(0,T)$, we have, as $n\to\infty$,

$$\int_{\Omega\times D} X_{\lambda_n}(t)w \to \int_{\Omega\times D} X(t)w\,,$$
$$\int_0^t\int_{\Omega\times D} \gamma_{\lambda_n}(\nabla X_{\lambda_n})\cdot\nabla w \to \int_0^t\int_{\Omega\times D} \eta\cdot\nabla w\,,$$
$$\int_0^t\int_{\Omega\times D} \beta_{\lambda_n}(X_{\lambda_n})w \to \int_0^t\int_{\Omega\times D} \xi w$$

for almost every $t\in(0,T)$. Hence, letting $n\to\infty$, we get, for almost every $t\in(0,T)$,

$$X(t) - \int_0^t \operatorname{div}\eta(s)\,ds + \int_0^t \xi(s)\,ds = X_0 + \int_0^t B(s)\,dW_s \quad \text{in } V_0^*\,, \quad \mathbb{P}\text{-almost surely}\;:$$

since all the terms except the first are continuous with values in $L^1(\Omega;V_0^*)$, we infer also that $X\in C([0,T];L^1(\Omega;V_0^*))$ and the integral relation holds for every $t\in[0,T]$. Moreover, since we also have $X\in L^\infty(0,T;L^2(\Omega;H))$, by Lemma 1.4.1 we can infer that $X\in C_w([0,T];L^2(\Omega;H))$.

Secondly, using the weak lower semicontinuity of the convex integrals and the estimates on $j(X_\lambda)$ and $j^*(\beta_\lambda(X_\lambda))$, it is immediate to check that $j(X)+j^*(\xi)\in L^1(\Omega\times Q)$. Furthermore, as we did at the end of Section 6.3.7, using Mazur's lemma, we deduce also that X and ξ are predictable, and η is adapted.

The last thing that we have to check is that $\eta\in\gamma(\nabla X)$ and $\xi\in\beta(X)$ a.e. in $\Omega\times Q$. To this aim, by the second part of Proposition 6.7.1, using the notation $\eta_\lambda := \gamma_\lambda(\nabla X_\lambda)$, we have that, for every $t \in [0,T]$,

$$\frac{1}{2}\|X_\lambda(t)\|^2_{L^2(\Omega;H)} + \int_0^t\int_{\Omega\times D}\eta_\lambda(s)\cdot\nabla X_\lambda(s)\,ds + \int_0^t\int_{\Omega\times D}\beta_\lambda(X_\lambda)(s)X_\lambda(s)\,ds$$
$$= \frac{1}{2}\|X_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^t\|B(s)\|^2_{L^2(\Omega;\mathscr{L}^2(U,H))}\,ds$$

and

$$\frac{1}{2}\|X(t)\|^2_{L^2(\Omega;H)} + \int_0^t\int_{\Omega\times D}\eta(s)\cdot\nabla X(s)\,ds + \int_0^t\int_{\Omega\times D}\xi(s)X(s)\,ds$$
$$= \frac{1}{2}\|X_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^t\|B(s)\|^2_{L^2(\Omega;\mathscr{L}^2(U,H))}\,ds\,.$$

We deduce that

$$\begin{aligned}
&\limsup_{n\to\infty}\left[\int_{\Omega\times Q}\eta_{\lambda_n}\cdot\nabla X_{\lambda_n} + \int_{\Omega\times Q}\beta_{\lambda_n}(X_{\lambda_n})X_{\lambda_n}\right]\\
&= \frac{1}{2}\|X_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^T\|B(s)\|^2_{L^2(\Omega;\mathscr{L}^2(U,H))}\,ds - \frac{1}{2}\liminf_{n\to\infty}\|X_{\lambda_n}(T)\|^2_{L^2(\Omega;H)}\\
&\le \frac{1}{2}\|X_0\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^T\|B(s)\|^2_{L^2(\Omega;\mathscr{L}^2(U,H))}\,ds - \frac{1}{2}\|X(T)\|^2_{L^2(\Omega;H)}\\
&= \int_{\Omega\times Q}\eta\cdot\nabla X + \int_{\Omega\times Q}\xi X\,.
\end{aligned}$$

Let us identify $\mathbb{R}^d\times\mathbb{R}$ with $\mathbb{R}^{d+1}$, indicate the generic element in $\mathbb{R}^{d+1}$ as a couple (x,y), where $x\in\mathbb{R}^d$ and $y\in\mathbb{R}$, and use the symbol $\bullet$ for the usual scalar product in $\mathbb{R}^{d+1}$. Consider the proper, convex and lower semicontinuous function $\Phi:\mathbb{R}^{d+1}\to[0,+\infty)$ given by $\Phi(x,y):=k(x)+j(y)$, $(x,y)\in\mathbb{R}^{d+1}$: then the subdifferential of Φ is the operator $\Xi:\mathbb{R}^{d+1}\to 2^{\mathbb{R}^{d+1}}$ given by $\Xi(x,y)=\{(u,v)\in\mathbb{R}^{d+1}: u\in\gamma(x), v\in\beta(y)\}$. Hence, recalling that $\beta_\lambda(X_\lambda)R_\lambda X_\lambda = \beta_\lambda(X_\lambda)X_\lambda - \lambda|\beta_\lambda(X_\lambda)|^2 \le \beta_\lambda(X_\lambda)X_\lambda$ and similarly $\eta_\lambda\cdot J_\lambda(\nabla X_\lambda) = \eta_\lambda\cdot\nabla X_\lambda - \lambda|\eta_\lambda|^2$, we have proved that

$$\limsup_{n\to\infty}\int_{\Omega\times Q}(\eta_{\lambda_n},\beta_{\lambda_n}(X_{\lambda_n}))\bullet(J_{\lambda_n}(\nabla X_{\lambda_n}),R_{\lambda_n}X_{\lambda_n}) \le \int_{\Omega\times Q}(\eta,\xi)\bullet(\nabla X,X)\,,$$

allowing us to infer that $(\eta,\xi)\in\Xi(\nabla X,X)$, i.e. that $\eta\in\gamma(\nabla X)$ and $\xi\in\beta(X)$ a.e. in $\Omega\times Q$, thanks to the classical results of convex analysis.

6.4 Continuous dependence on the initial datum with additive noise

This section is devoted to the proof of the continuous dependence and uniqueness results contained in Theorem 6.2.5. The main tool that we use is the generalized Itô formula contained in Proposition 6.7.1.

We start assuming (6.2.6): let (X_0^1,B_1), (X_0^2,B_2), (X_1,η_1,ξ_1), (X_2,η_2,ξ_2) be as in Theorem 6.2.5. Then, writing relation (6.2.10) for $(X_1,\eta_1,\xi_1,X_0^1,B_1)$ and $(X_2,\eta_2,\xi_2,X_0^2,B_2)$ and taking the difference, $\mathbb{P}$-almost surely we obtain

$$\begin{aligned} X_1(t) - X_2(t) - \int_0^t \operatorname{div}\left[\eta_1(s) - \eta_2(s)\right]\, ds + \int_0^t (\xi_1(s) - \xi_2(s))\, ds \\ = X_0^1 - X_0^2 + \int_0^t (B_1(s) - B_2(s))\, dW_s \qquad \text{for every } t \in [0,T]\,. \end{aligned}$$

Now, we note that thanks to (6.2.13) and the symmetry of j, for $i = 1, 2$ we have

$$j\left(\frac{X_1 - X_2}{2}\right) + j^*\left(\frac{\xi_1 - \xi_2}{2}\right) \le \frac{1}{2}\left[j(X_1) + j(X_2) + j^*(\xi_1) + j(\xi_2)\right]\,,$$

where the right hand side is in $L^1(\Omega \times (0,T) \times D)$: hence, we can apply Proposition 6.7.1 with the choices $Y = X_1 - X_2$, $f = \eta_1 - \eta_2$, $g = \xi_1 - \xi_2$, $T = B_1 - B_2$ and $\alpha = 1/2$ in order to infer that for every $t \in [0,T]$

$$\begin{aligned} &\frac{1}{2}\left\|X_1(t) - X_2(t)\right\|_H^2 + \int_0^t \int_D (\eta_1(s) - \eta_2(s)) \cdot (\nabla X_1(s) - \nabla X_2(s))\, ds \\ &\quad + \int_0^t \int_D (\xi_1(s) - \xi_2(s))\,(X_1(s) - X_2(s))\, ds \\ &= \frac{1}{2}\left\|X_0^1 - X_0^2\right\|_H^2 + \frac{1}{2}\int_0^t \left\|(B_1 - B_2)(s)\right\|^2_{\mathscr{L}^2(U,H)}\, ds \\ &\quad + \int_0^t ((X_1 - X_2)(s), (B_1 - B_2)(s)\, dW_s)\,. \end{aligned}$$

Hence, taking into account (6.2.11)–(6.2.12) and the monotonicity of γ and β, we obtain

$$\begin{aligned} \left\|X_1(t) - X_2(t)\right\|_H^2 \le \left\|X_0^1 - X_0^2\right\|_H^2 + \int_0^t \left\|B_1(s) - B_2(s)\right\|^2_{\mathscr{L}^2(U,H)}\, ds \\ + 2\sup_{t\in[0,T]}\left|\int_0^t ((X_1 - X_2)(s), (B_1 - B_2)(s)\, dW_s)\right|\,; \end{aligned}$$

moreover, proceeding exactly as in Section 6.3.3, taking the supremum in $t \in [0,T]$ in the last expression and then expectations, thanks to the Davis inequality and the Young inequality, we easily obtain

$$\begin{aligned} \left\|X_1 - X_2\right\|^2_{L^2(\Omega;L^\infty(0,T;H))} \le \left\|X_0^1 - X_0^2\right\|^2_{L^2(\Omega;H)} + c\left\|B_1 - B_2\right\|^2_{L^2(\Omega\times(0,T);\mathscr{L}^2(U,H))} \\ + \frac{1}{2}\left\|X_1 - X_2\right\|^2_{L^2(\Omega;L^\infty(0,T;H))} \end{aligned}$$

for a positive constant c, from which (6.2.18) follows. Finally, if $X_0^1 = X_0^2$ and $B_1 = B_2$, we immediately get $X_1 = X_2$: substituting in the difference of the respective equations (6.2.10) we have $\int_0^t (-\operatorname{div}(\eta_1(s) - \eta_2(s)) + (\xi_1(s) - \xi_2(s)))\, ds = 0$ for every t. Relying now on hypothesis (6.2.6) and proceeding as in Section 6.3.6, we easily get also $\eta_1 = \eta_2$ and $\xi_1 = \xi_2$.

Let us prove now the second part of Theorem 6.2.5, in which condition (6.2.6) is not assumed. By the second part of Theorem 6.2.2, we have that, for every $t \in [0,T]$,

$$\begin{aligned} X_1(t) - X_2(t) - \int_0^t \operatorname{div}\left[\eta_1(s) - \eta_2(s)\right]\, ds + \int_0^t (\xi_1(s) - \xi_2(s))\, ds \\ = X_0^1 - X_0^2 + \int_0^t (B_1(s) - B_2(s))\, dW_s \qquad \mathbb{P}\text{-almost surely}\,: \end{aligned}$$

hence, using the second part of Proposition 6.7.1, we infer that, for every $t \in [0,T]$,

$$
\begin{aligned}
&\frac{1}{2}\left\|X_1(t)-X_2(t)\right\|^2_{L^2(\Omega;H)} + \int_0^t\int_{\Omega\times D}(\eta_1(s)-\eta_2(s))\cdot(\nabla X_1(s)-\nabla X_2(s))\,ds \\
&\qquad + \int_0^t\int_{\Omega\times D}(\xi_1(s)-\xi_2(s))\,(X_1(s)-X_2(s))\,ds \\
&= \frac{1}{2}\left\|X_0^1-X_0^2\right\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^t \left\|(B_1-B_2)(s)\right\|^2_{L^2(\Omega;\mathscr{L}^2(U,H))}\,ds\,,
\end{aligned}
$$

which together with the monotonicity of γ and β implies (6.2.19). Finally, if we have $X_0^1 = X_0^2$ and $B_1 = B_2$, it is clear that $X_1 = X_2$ and, by comparison in the equation itself, also $\int_0^t (-\operatorname{div}(\eta_1(s)-\eta_2(s)) + (\xi_1(s)-\xi_2(s)))\,ds = 0$ for every t, as before, so that $-\operatorname{div}\eta_1+\xi_1 = -\operatorname{div}\eta_2+\xi_2$.

6.5 Well-posedness with multiplicative noise

In this section, we prove the main theorem of the work, which ensures that the original problem is well-posed also with multiplicative noise. Let us describe the approach that we will follow.

The main idea is to prove existence of solutions proceeding step-by-step: we introduce a parameter $\tau > 0$, we prove using contraction estimates that we are able to recover some solutions on each subinterval $[0,\tau], [\tau, 2\tau], \ldots [n\tau, (n+1)\tau], \ldots$ provided that τ is chosen sufficiently small, and finally we paste together each solution on the whole interval $[0,T]$. In this sense, the main point of the argument is to prove that such a value of τ can be chosen uniformly with respect to n, so that the procedure stops when we reach the final time T (in a finite number of steps).

6.5.1 Existence

In this section we prove the two existence results contained in Theorem 6.2.7. We start from the first one, i.e. assuming (6.2.6). First of all, for every $a, b \in [0,T]$ with $b > a$ and for any progressively measurable process $Y \in L^2(\Omega\times(0,T)\times D)$, condition (6.2.23) implies that $B(\cdot,\cdot,Y) \in L^2(\Omega\times(a,b);\mathscr{L}^2(U,H))$: hence, for every $X_a \in L^2(\Omega,\mathcal{F}_a,\mathbb{P};H)$, thanks to Theorem 6.2.2 we know that there exist

$$X_{a,b} \in L^2\left(\Omega; L^\infty(a,b;H)\right) \cap L^p\left(\Omega\times(a,b);V\right)\,, \tag{6.5.90}$$

$$\eta_{a,b} \in L^q\left(\Omega\times(a,b)\times D\right)^d\,, \qquad \xi_{a,b} \in L^1\left(\Omega\times(a,b)\times D\right)\,, \tag{6.5.91}$$

such that $X_{a,b}$ is adapted with $\mathbb{P}$-almost surely weakly continuous paths in H and the following relations hold:

$$
\begin{gathered}
X_{a,b}(t) - \int_a^t \operatorname{div}\eta_{a,b}(s)\,ds + \int_a^t \xi_{a,b}(s)\,ds = X_a + \int_a^t B(s,Y(s))\,dW_s \\
\text{in } V_0^*\,, \quad \text{for every } t\in[a,b]\,, \quad \mathbb{P}\text{-a.s.}\,,
\end{gathered} \tag{6.5.92}
$$

$$\eta_{a,b} \in \gamma(\nabla X_{a,b}) \quad \text{a.e. in } \Omega\times(a,b)\times D\,, \tag{6.5.93}$$

$$\xi_{a,b} \in \beta(X_{a,b}) \quad \text{a.e. in } \Omega\times(a,b)\times D\,, \tag{6.5.94}$$

$$j(X_{a,b}) + j^*(\xi_{a,b}) \in L^1\left(\Omega\times(a,b)\times D\right)\,, \tag{6.5.95}$$

where $X_{a,b}$ is unique in the sense of Theorem 6.2.5. Now, we need the following lemma.

Lemma 6.5.1. *For every $\tau > 0$ and $n \in \mathbb{N}$ fixed, let $X_{n\tau} \in L^2(\Omega, \mathcal{F}_{n\tau}, \mathbb{P}; H)$ and $Y_1, Y_2 \in L^2(\Omega \times (n\tau, (n+1)\tau) \times D)$ be progressively measurable: then, if (X_1, η_1, ξ_1) and (X_2, η_2, ξ_2) are any respective solutions to (6.5.90)–(6.5.95) with the choices $a = n\tau$, $b = (n+1)\tau$ and same initial value $X_a = X_{n\tau}$, we have the following estimate:*

$$\|X_1 - X_2\|_{L^2(\Omega\times(n\tau,(n+1)\tau)\times D)} \le \sqrt{\tau L_B}\, \|Y_1 - Y_2\|_{L^2(\Omega\times(n\tau,(n+1)\tau)\times D)} \, . \tag{6.5.96}$$

Proof. Taking the difference of equations (6.5.92) evaluated with $i = 1, 2$ and recalling the generalized Itô formula (6.7.116), setting $X := X_1 - X_2$, $\eta := \eta_1 - \eta_2$ and $\xi := \xi_1 - \xi_2$, we easily get that for every $t \in [m\tau, (m+1)\tau]$

$$\frac{1}{2}\|X(t)\|^2_{L^2(\Omega\times D)} + \int_0^t \int_{\Omega\times D} \eta(s)\cdot\nabla X(s)\, ds + \int_0^t \int_{\Omega\times D} \xi(s) X(s)\, ds \\ = \frac{1}{2}\|B(Y_1) - B(Y_2)\|^2_{L^2(\Omega\times(m\tau,(m+1)\tau);\mathscr{L}^2(U,H))} \, .$$

Hence, using the Lipschitz continuity of B and the monotonicity of γ and β we have

$$\frac{1}{2}\|X_1 - X_2\|^2_{L^\infty(m\tau,(m+1)\tau;L^2(\Omega\times D))} \le \frac{L_B}{2}\|Y_1 - Y_2\|^2_{L^2(\Omega\times(m\tau,(m+1)\tau)\times D)} \, ,$$

from which (6.5.96) follows. □

Now, let us build some solutions X, η and ξ in each sub-interval. To this purpose, we choose $\tau > 0$ such that the constant appearing in (6.5.96) is less than 1, for example

$$\tau := \frac{1}{2L_B} \, .$$

Firstly, we focus on $[0, \tau]$: taking into account the remarks that we have just made, it is well defined the function

$$\Phi_0 : L^2(\Omega\times(0,\tau)\times D) \to L^2(\Omega\times(0,\tau)\times D) \, , \quad \Phi_0(Y) := X \, , \tag{6.5.97}$$

where X is the unique solution to (6.5.90)–(6.5.95) with the choices $a = 0$ and $b = \tau$, with X_0 given by (6.2.20). It is clear that X is a solution of problem (6.2.24) in $[0, \tau]$ if and only if it is a fixed point for Φ_0. Thanks to the estimate (6.5.96) and the choice of τ, Φ_0 is a contraction: hence, it has a fixed point

$$X^{(0)} \in L^2(\Omega; L^\infty(0,\tau;H)) \cap L^p(\Omega\times(0,\tau); V) \, ,$$

with $\mathbb{P}$-almost surely weakly continuous paths in H, which solves (6.2.24) with certain

$$\eta^{(0)} \in L^q(\Omega\times(0,\tau)\times D)^d \, , \qquad \xi^{(0)} \in L^1(\Omega\times(0,\tau)\times D) \, .$$

Secondly, let us focus on $[\tau, 2\tau]$, set $X_\tau := X^{(0)}(\tau)$ (which is in $L^2(\Omega, \mathcal{F}_\tau, \mathbb{P}; H)$ since $X^{(0)}$ is adapted) and define the function

$$\Phi_1 : L^2(\Omega\times(\tau,2\tau)\times D) \to L^2(\Omega\times(\tau,2\tau)\times D) \, , \quad \Phi_1(Y) := X \, , \tag{6.5.98}$$

where X is the solution to (6.5.90)–(6.5.95) with the choices $a = \tau$ and $b = 2\tau$. As we have already done, thanks to the estimate (6.5.96) and the same choice of τ, Φ_1 is a contraction:

hence, it has a fixed point

$$X^{(1)} \in L^2\left(\Omega; L^\infty(\tau, 2\tau; H)\right) \cap L^p\left(\Omega \times (\tau, 2\tau); V\right) ,$$

with $\mathbb{P}$-almost surely weakly continuous paths in H, which is a solution of (6.2.24) with certain

$$\eta^{(1)} \in L^q\left(\Omega \times (\tau, 2\tau) \times D\right)^d , \qquad \xi^{(1)} \in L^1\left(\Omega \times (\tau, 2\tau) \times D\right) .$$

Suppose by induction that we have built the triplets until step $m-1$, i.e. $(X^{(0)}, \eta^{(0)}, \xi^{(0)}), \ldots, (X^{(m-1)}, \eta^{(m-1)}, \xi^{(m-1)})$. To proceed, we focus on the interval $[m\tau, (m+1)\tau]$, set $X_{m\tau} := X^{(m-1)}(m\tau)$ (which is in $L^2(\Omega, \mathcal{F}_{m\tau}, \mathbb{P}; H)$ since $X^{(m-1)}$ is adapted) and define the function

$$\Phi_m : L^2\left(\Omega \times (m\tau, (m+1)\tau) \times D\right) \to L^2\left(\Omega \times (m\tau, (m+1)\tau) \times D\right) , \tag{6.5.99}$$

which maps Y into X, where X is the solution to (6.5.90)–(6.5.95) with the choices $a = m\tau$ and $b = (m+1)\tau$. Now, Φ_m is a contraction thanks to (6.5.96) and the choice of τ, so it has a fixed point

$$X^{(m)} \in L^2\left(\Omega; L^\infty(m\tau, (m+1)\tau; H)\right) \cap L^p\left(\Omega \times (m\tau, (m+1)\tau); V\right) :$$

with $\mathbb{P}$-almost surely weakly continuous paths in H, which is a solution of (6.2.24) with certain

$$\eta^{(m)} \in L^q\left(\Omega \times (m\tau, (m+1)\tau) \times D\right)^d , \qquad \xi^{(m)} \in L^1\left(\Omega \times (m\tau, (m+1)\tau) \times D\right) .$$

In this way, we can define the triplet (X, η, ξ) by setting, as it is natural, $(X, \eta, \xi) := (X^{(m)}, \eta^{(m)}, \xi^{(m)})$ in $\Omega \times [m\tau, (m+1)\tau) \times D$ for every $m \in \mathbb{N}$ until we reach T: bearing in mind how we have built $(X^{(m)}, \eta^{(m)}, \xi^{(m)})$, it is clear that X, η and ξ are well-defined and solve the problem with multiplicative noise.

Finally, if we do not assume (6.2.6), it is clear that, using the same argument, the respective solutions constructed in this way are well-defined and satisfy conditions (6.2.14) and (6.2.25) instead of (6.2.7) and (6.2.24), respectively.

6.5.2 Continuous dependence on the initial datum

We present here the proof of the continuous dependence results contained in the last part of Theorem 6.2.7. Here, we repeat exactly the same argument of Section 6.4 with the choices $B_1 := B(\cdot, X_1)$ and $B_2 := (\cdot, X_2)$.

If (6.2.6) is assumed, for any given $\tau > 0$, the same computations on the interval $(0, \tau)$ get us to

$$\begin{aligned}
&\left\|X_1 - X_2\right\|^2_{L^2(\Omega; L^\infty(0,\tau;H))} \\
&\qquad \le c \left\|X_0^1 - X_0^2\right\|^2_{L^2(\Omega;H)} + c\left\|B(X_1) - B(X_2)\right\|^2_{L^2(\Omega \times (0,\tau); \mathscr{L}^2(U,H))}
\end{aligned}$$

for a constant $c > 0$ independent of τ; using the Lipschitz continuity of B we obtain

$$\left\|X_1 - X_2\right\|^2_{L^2(\Omega; L^\infty(0,\tau;H))} \le c \left\|X_0^1 - X_0^2\right\|^2_{L^2(\Omega;H)} + c\tau \left\|X_1 - X_2\right\|^2_{L^2(\Omega; L^\infty(0,\tau;H))} .$$

Hence, choosing for example $\tau = \frac{c}{2}$, we get the desired relation on the interval $[0, \tau]$. The idea is clearly to iterate the procedure on the following intervals $[\tau, 2\tau]$, $[2\tau, 3\tau]$, $\ldots$ until we reach the final time T, so that (6.2.26) is proved. The important point that we have to check is that the choice of τ can be made uniformly with respect to each sub-interval, but this is not difficult: as a matter of fact, for any $n \geq 1$, performing the same computations on $[n\tau, (n+1)\tau]$ we obtain

$$\begin{aligned}\|X_1 - X_2\|^2_{L^2(\Omega; L^\infty(n\tau,(n+1)\tau;H))} &\leq c\,\|X_1(n\tau) - X_2(n\tau)\|^2_{L^2(\Omega;H)} \\ &\quad + c\tau\,\|X_1 - X_2\|^2_{L^2(\Omega; L^\infty(n\tau,(n+1)\tau;H))}\,,\end{aligned}$$

for the same constant c, from which we deduce that the choice of τ is independent of n, and one can easily conclude by induction on n. As we did in Section 6.4, if $X_0^1 = X_0^1$, then by (6.2.26) we have $X_1 = X_2$, and hypothesis (6.2.6) also ensures $\eta_1 = \eta_2$ and $\xi_1 = \xi_2$.

Secondly, if (6.2.6) is not assumed, proceeding as in the final part of Section 6.4 we get for every $t \in [0,T]$ that

$$\begin{aligned}&\|X_1(t) - X_2(t)\|^2_{L^2(\Omega;H)} \\ &\quad\leq \left\|X_0^1 - X_0^2\right\|^2_{L^2(\Omega;H)} + \int_0^t \|(B(X_1) - B(X_2))(s)\|^2_{L^2(\Omega;\mathscr{L}^2(U,H))}\,ds\,,\end{aligned}$$

from which (6.2.27) follows using the Lipschitz continuity of B and the Gronwall lemma. Finally, if $X_0^1 = X_0^1$, then by (6.2.27) $X_1 = X_2$ and consequently $-\operatorname{div}\eta_1 + \xi_1 = -\operatorname{div}\eta_2 + \xi_2$.

6.6 An integration-by-parts formula

The aim of this section is to give a complete proof of the generalized testing formula contained in equation (6.3.74): throughout the section, we assume to work with the notations and setting of Section 6.3.5. Here, $\varepsilon \in (0,1)$ and $\omega \in \Omega$ are fixed as usual.

The main point is that we cannot directly test equation (6.3.68) by $X^\varepsilon - W_B^\varepsilon$, as we did in Section 6.3.2, since the regularity of X^ε is not sufficient: more specifically, $\partial_t(X^\varepsilon - W_B^\varepsilon)$ is only intended in V_0^* and we would need that $X^\varepsilon - W_B^\varepsilon$ takes values in V_0, but this is not the case. However, by condition (6.3.73) and the regularities of X^ε, W_B^ε and η^ε, all the terms in (6.3.74) make sense: hence, the intuitive idea is that (6.3.74) holds at least in a formal way. To give a rigorous proof of it, a natural way could be to try to pass to the limit as $\lambda \searrow 0$ in (6.3.40): however, it is not necessarily true in our framework that equation (6.3.40) converges to (6.3.74) as $\lambda \searrow 0$, so this approach does not work. Hence, the idea is to see (6.3.74) as a limit problem as $\delta \searrow 0$, for another parameter δ, such that the approximations in δ have good smoothing properties and behave better that the approximations in λ. In this sense, a similar approach was presented in [12], where the approximations were built using suitable powers of the resolvent of the Laplacian. However, in our case we have to approximate also elements in $W^{-1,q}(D)$ (namely, $-\operatorname{div}\eta^\varepsilon$) and the resolvent of the laplacian does not work since $-\Delta$ is not coercive on V: the idea is thus to identify another suitable space, in which (6.3.68) can be intended, and to define appropriate approximations on it. To this purpose, we need some preparatory work.

First of all, note that the operator $-\operatorname{div} : L^q(D)^d \to V^*$ is linear, continuous and satisfies $\|-\operatorname{div} u\|_{V^*} \leq \|u\|_{L^q(D)}$ for every $u \in L^q(D)^d$. Let us define the space

$$V^*_{div} := \left\{-\operatorname{div} u : u \in L^q(D)^d\right\} \subseteq V^*\,.$$

Secondly, we introduce the space $V^*_{div} + L^1(D)$ as the subspace of V_0^* given by all the formal linear combinations of elements in V^*_{div} and $L^1(D)$. With this notations, we can note that equation (6.3.68) actually holds in $V_{div}^* + L^1(D)$: in other words, for every $t \in [0,T]$, we have

$$(X^\varepsilon - W_B^\varepsilon)(t) + \int_0^t (-\operatorname{div}\eta^\varepsilon(s) + \xi^\varepsilon(s))\, ds = X_0 \quad \text{in } V_{div}^* + L^1(D)\,. \tag{6.6.100}$$

Hence, the idea is that it is sufficient to identify a way to approximate only elements in $V_{div}^* + L^1(D)$, and not any element of V_0^*, which would be much more demanding.

To this end, for every $\delta \in (0,1)$, let $\mathcal{R}_\delta := (I - \delta\Delta)^{-1}$ be the resolvent of the Laplace operator. It is well-known that for every $r \in [1,+\infty)$, $\mathcal{R}_\delta : L^r(D) \to L^r(D)$ is a linear contraction converging to the identity as $\delta \searrow 0$ in the strong operator topology (the reader can refer to [10,15,24]). In this setting, we define the operator $\mathbf{R}_\delta : L^r(D)^d \to L^r(D)^d$ extending $\mathcal{R}_\delta$ component-by-component: consequently, we easily deduce that also $\mathbf{R}_\delta$ is a linear contraction on $L^r(D)^d$ converging to the identity as $\delta \searrow 0$. With this notations, we have the following result.

Lemma 6.6.1. *For every $u \in L^q(D)^d$ such that $-\operatorname{div} u \in L^1(D)$ (in the distributional sense), we have*

$$-\operatorname{div}\mathbf{R}_\delta u = \mathcal{R}_\delta\left(-\operatorname{div} u\right)\,.$$

Moreover, for every $f \in H^1(D)$, we have

$$\nabla\mathcal{R}_\delta f = \mathbf{R}_\delta \nabla f\,.$$

Proof. Let us first assume that $u \in (C_c^\infty(D))^d$: then, using the definition of $\mathbf{R}_\delta$ and $\mathcal{R}_\delta$, integration by parts and the fact that $\mathcal{R}_\delta$ commutes with Δ, for every $\varphi \in C_c^\infty(D)$ we have

$$\begin{aligned}
\int_D (-\operatorname{div} u)\,\varphi = \int_D u\cdot\nabla\varphi = \sum_{i=1}^d \int_D u_i \frac{\partial\varphi}{\partial x_i} &= \sum_{i=1}^d \int_D \left(\mathcal{R}_\delta u_i - \delta\Delta\mathcal{R}_\delta u_i\right)\frac{\partial\varphi}{\partial x_i} \\
&= \int_D \mathbf{R}_\delta u\cdot\nabla\varphi + \delta\int_D \Delta(\operatorname{div}\mathbf{R}_\delta u)\varphi \\
&= \int_D \left[-\operatorname{div}\mathbf{R}_\delta u - \delta\Delta(-\operatorname{div}\mathbf{R}_\delta u)\right]\varphi\,.
\end{aligned}$$

Hence, by definition of the resolvent, we deduce that $-\operatorname{div}\mathbf{R}_\delta u = \mathcal{R}_\delta(-\operatorname{div} u)$ for every $u \in (C_c^\infty(D))^d$. At this point, if $u \in L^q(D)^d$ and $-\operatorname{div} u \in L^1(D)$, the first thesis follows by approximating u with a sequence $\{u_n\}_{n\in\mathbb{N}} \subseteq (C^\infty(D))^d$ such that $u_n \to u$ in $L^q(D)^d$ and $-\operatorname{div} u_n \to -\operatorname{div} u$ in $L^1(D)$. Finally, in a similar way, the second assertion is clearly true for every $f \in C^\infty(\overline{D})$: hence, given $f \in H^1(D)$, we can conclude by density approximating f with a sequence $\{f_n\}_{n\in\mathbb{N}} \subseteq C^\infty(\overline{D})$. □

Now, for every $\delta \in (0,1)$, we introduce the operator

$$\Lambda_\delta^1 : V_{div}^* \to V_{div}^*$$

in the following way: for any given $f \in V_{div}^*$, with $f = -\operatorname{div} u$ for a certain $u \in L^q(D)^d$, we set $\Lambda_\delta^1 f := -\operatorname{div}\mathbf{R}_\delta u$. Note that Λ_δ^1 is well-defined: indeed, if $f = -\operatorname{div} u_1 = -\operatorname{div} u_2$, we have $-\operatorname{div}(u_1 - u_2) = 0$ and by Lemma 6.6.1 we deduce that $0 = \mathcal{R}_\delta(-\operatorname{div}(u_1 - u_2)) = -\operatorname{div}(\mathbf{R}_\delta(u_1 - u_2))$, so that $-\operatorname{div}\mathbf{R}_\delta u_1 = -\operatorname{div}\mathbf{R}_\delta u_2$. Secondly, we set

$$\Lambda_\delta^2 : L^1(D) \to L^1(D)\,, \quad \Lambda_\delta^2 := \mathcal{R}_\delta\,.$$

The first part of Lemma 6.6.1 ensures that $\Lambda^1_\delta = \Lambda^2_\delta$ on the intersection $V^*_{div} \cap L^1(D)$: hence, it is well-defined the operator

$$\Lambda_\delta := \Lambda^1_\delta + \Lambda^2_\delta : V^*_{div} + L^1(D) \to V^*_{div} + L^1(D) \tag{6.6.101}$$

such that

$$\Lambda_\delta(-\operatorname{div} u) = -\operatorname{div} \mathbf{R}_\delta u\,, \quad \Lambda_\delta(f) = \mathcal{R}_\delta f \quad \forall\, u \in L^q(D)^d\,,\ f \in L^1(D)\,, \tag{6.6.102}$$

which is automatically linear.

We are now ready to build the approximations. First of all, we choose $k \in \mathbb{N}$ as in the defintion of V_0, so that the k-th power $\mathcal{R}^k_\delta$ maps H into $V_0 \subseteq V \cap L^\infty(D)$. At this point, we define

$$X^\varepsilon_\delta := \mathcal{R}^k_\delta X^\varepsilon\,, \quad W^\varepsilon_\delta := \mathcal{R}^k_\delta W^\varepsilon_B\,, \quad \eta^\varepsilon_\delta := \mathbf{R}^k_\delta \eta^\varepsilon\,, \quad \xi^\varepsilon_\delta := \mathcal{R}^k_\delta \xi^\varepsilon\,, \quad X^\delta_0 := \mathcal{R}^k_\delta X_0\ :$$

then, taking into account the properties of $\mathcal{R}_\delta$ and $\mathbf{R}_\delta$ and the second part of Lemma 6.6.1, by the regularities of the processes in play we have as $\delta \searrow 0$ that

$$X^\varepsilon_\delta(t) \to X^\varepsilon(t) \quad \text{in } H \quad \forall\, t \in [0,T]\,, \quad X^\varepsilon_\delta \to X^\varepsilon \quad \text{in } L^p\left(0,T;V\right) \tag{6.6.103}$$

$$W^\varepsilon_\delta(t) \to W^\varepsilon(t) \quad \text{in } H \quad \forall\, t \in [0,T]\,, \quad W^\varepsilon_\delta \to W^\varepsilon_B \quad \text{in } L^p\left(0,T;V\right)\,, \tag{6.6.104}$$

$$\eta^\varepsilon_\delta \to \eta^\varepsilon \quad \text{in } L^q\left(Q\right)^d\,, \qquad \xi^\varepsilon_\delta \to \xi^\varepsilon \quad \text{in } L^1(Q)\,, \tag{6.6.105}$$

$$X^\delta_0 \to X_0 \quad \text{in } H\,. \tag{6.6.106}$$

Now, applying the operator Λ^k_δ to equation (6.6.100), we get for every $t \in [0,T]$ that

$$\left(X^\varepsilon_\delta - W^\varepsilon_\delta\right)(t) - \int_0^t \operatorname{div} \eta^\varepsilon_\delta(s)\,ds + \int_0^t \xi^\varepsilon_\delta(s)\,ds = X^\delta_0\,. \tag{6.6.107}$$

With our choice of k, it now makes sense to test by $X^\varepsilon_\delta - W^\varepsilon_\delta$: it easily follows that

$$\frac{1}{2}\left\|X^\varepsilon_\delta(T) - W^\varepsilon_\delta(T)\right\|^2_H + \int_Q \nabla\eta^\varepsilon_\delta \cdot \nabla\left(X^\varepsilon_\delta - W^\varepsilon_\delta\right) + \int_Q \xi^\varepsilon_\delta\left(X^\varepsilon_\delta - W^\varepsilon_\delta\right) = \frac{1}{2}\left\|X^\delta_0\right\|^2_H\,, \tag{6.6.108}$$

from which, taking into account (6.6.103)–(6.6.106), we deduce that

$$\lim_{\delta\searrow 0} \int_Q \xi^\varepsilon_\delta\left(X^\varepsilon_\delta - W^\varepsilon_\delta\right) = \frac{1}{2}\left\|X_0\right\|^2_H - \frac{1}{2}\left\|(X^\varepsilon - W^\varepsilon_B)(T)\right\|^2_H - \int_Q \nabla\eta^\varepsilon \cdot \nabla\left(X^\varepsilon - W^\varepsilon_B\right). \tag{6.6.109}$$

In order to evaluate the limit in the previous expression, we take advantage of Vitali convergence theorem: to this purpose, thanks to (6.6.103)–(6.6.105), we can assume with no restriction that $\xi^\varepsilon_\delta \to \xi^\varepsilon$ and $X^\varepsilon_\delta - W^\varepsilon_\delta \to X^\varepsilon - W^\varepsilon_B$ almost everywhere in Q. Let us show that $\{\xi^\varepsilon_\delta(X^\varepsilon_\delta - W^\varepsilon_\delta)\}_{\delta\in(0,1)}$ is uniformly integrable in Q: thanks to the generalized Jensen inequality for the positive operator R_δ (see [40, 41] for references), we have

$$\begin{aligned}\pm\xi^\varepsilon_\delta(X^\varepsilon_\delta - W^\varepsilon_\delta) &\le j\left(\pm(X^\varepsilon_\delta - W^\varepsilon_\delta)\right) + j^*(\xi^\varepsilon_\delta) = j\left(X^\varepsilon_\delta - W^\varepsilon_\delta\right) + j^*(\xi^\varepsilon_\delta)\\ &\le \mathcal{R}^k_\delta\left[j\left(X^\varepsilon - W^\varepsilon_B\right) + j^*(\xi^\varepsilon)\right] \quad \text{a.e. in } Q\,.\end{aligned}$$

Now, since $j(X^\varepsilon - W^\varepsilon_B), j^*(\xi^\varepsilon) \in L^1(Q)$ thanks to (6.3.71) and (6.3.39), the right hand side of the previous expression converges in $L^1(Q)$ and consequently it is uniformly integrable in Q: we deduce that also $\{\xi^\varepsilon_\delta(X^\varepsilon_\delta - W^\varepsilon_\delta)\}_{\delta\in(0,1)}$ is uniformly integrable in Q. Hence, by Vitali convergence theorem, we infer that

$$\xi_\delta^\varepsilon \left(X_\delta^\varepsilon - W_\delta^\varepsilon\right) \to \xi^\varepsilon \left(X^\varepsilon - W_B^\varepsilon\right) \quad \text{in } L^1(Q) \quad \text{as } \delta \searrow 0\,,$$

so that passing to the limit in (6.6.109) we recover exactly (6.3.74).

6.7 The generalized Itô formula

In this section, we prove a generalized Itô formula, which is widely used in Sections 6.3.7 and 6.4: we collect the general result in the following proposition.

Proposition 6.7.1. *Assume the following conditions:*

$$Y_0 \in L^2\left(\Omega, \mathcal{F}_0, \mathbb{P}; H\right), \tag{6.7.110}$$

$$T \in L^2\big(\Omega \times (0,T); \mathscr{L}^2(U,H)\big) \quad \textit{progressively measurable}, \tag{6.7.111}$$

$$Y \in L^2\left(\Omega; L^\infty(0,T;H)\right) \cap L^p\left(\Omega \times (0,T); V\right), \qquad Y \in C_w\left([0,T];H\right) \quad \mathbb{P}\text{-}a.s., \tag{6.7.112}$$

$$f \in L^q\left(\Omega \times (0,T) \times D\right)^d, \qquad g \in L^1\left(\Omega \times (0,T) \times D\right), \tag{6.7.113}$$

$$\exists\, \alpha > 0 : \quad j(\alpha Y) + j^*(\alpha g) \in L^1\left(\Omega \times (0,T) \times D\right), \tag{6.7.114}$$

$$Y(t) - \int_0^t \operatorname{div} f(s)\,ds + \int_0^t g(s)\,ds = Y_0 + \int_0^t T(s)\,dW_s \quad \textit{in } V_0^* \tag{6.7.115}$$

for every $t \in [0,T]$, $\mathbb{P}$-almost surely. Then, the following Itô formula holds

$$\begin{aligned} &\frac{1}{2}\left\|Y(t)\right\|_H^2 + \int_0^t \int_D f(s) \cdot \nabla Y(s)\,ds + \int_0^t \int_D g(s) Y(s)\,ds \\ &\qquad = \frac{1}{2}\left\|Y_0\right\|_H^2 + \frac{1}{2}\int_0^t \left\|T(s)\right\|_{\mathscr{L}^2(U,H)}^2\,ds + \int_0^t \left(Y(s), T(s)\,dW_s\right) \end{aligned} \tag{6.7.116}$$

for every $t \in [0,T]$, $\mathbb{P}$-almost surely. Furthermore, if hypothesis (6.7.112) *is replaced by the weaker condition*

$$Y \in L^\infty\big(0,T; L^2(\Omega;H)\big) \cap L^p\left(\Omega \times (0,T); V\right) \cap C_w\big([0,T]; L^2(\Omega;H)\big), \tag{6.7.117}$$

then instead of (6.7.116) *we have the following for every $t \in [0,T]$:*

$$\begin{aligned} &\frac{1}{2}\left\|Y(t)\right\|_{L^2(\Omega;H)}^2 + \int_0^t \int_{\Omega \times D} f(s) \cdot \nabla Y(s)\,ds + \int_0^t \int_{\Omega \times D} g(s) Y(s)\,ds \\ &\qquad = \frac{1}{2}\left\|Y_0\right\|_{L^2(\Omega;H)}^2 + \frac{1}{2}\int_0^t \left\|T(s)\right\|_{L^2(\Omega;\mathscr{L}^2(U,H))}^2\,ds\,. \end{aligned} \tag{6.7.118}$$

Proof. We proceed exactly in the same way as in Section 6.6. If k is given by the definition of V_0 and for every $\delta \in (0,1)$, $\mathcal{R}_\delta$ and $\mathbf{R}_\delta$ are as in Section 6.6, we define

$$Y_\delta := \mathcal{R}_\delta^k Y\,, \quad T_\delta := \mathcal{R}_\delta^k T\,, \quad f_\delta := \mathbf{R}_\delta^k f\,, \quad g_\delta := \mathcal{R}_\delta^k g\,, \quad Y_0^\delta := \mathcal{R}_\delta^k Y_0 \;:$$

hence, thanks to (6.7.110)–(6.7.113) and Lemma 6.6.1 we have as $\delta \searrow 0$

$$Y_\delta(t) \to Y(t) \quad \text{in } H \quad \text{for every } t \in [0,T] , \quad \mathbb{P}\text{-almost surely} , \tag{6.7.119}$$

$$Y_\delta \to Y \quad \text{in } L^p\left(\Omega \times (0,T); V\right) \tag{6.7.120}$$

$$T_\delta \to T \quad \text{in } L^2\left(\Omega \times (0,T); \mathscr{L}^2(U,H)\right) , \tag{6.7.121}$$

$$f_\delta \to f \quad \text{in } L^q\left(\Omega \times Q\right)^d , \qquad g_\delta \to g \quad \text{in } L^1(\Omega \times Q) , \tag{6.7.122}$$

$$Y_0^\delta \to Y_0 \quad \text{in } L^2(\Omega; H) . \tag{6.7.123}$$

Consequently, if we apply the operator Λ_δ^k to (6.7.115), taking definition (6.6.101)–(6.6.102) into account, we have $\mathbb{P}$-almost surely that

$$Y_\delta(t) - \int_0^t \operatorname{div} f_\delta(s)\, ds + \int_0^t g_\delta(s)\, ds = Y_0^\delta + \int_0^t T_\delta(s)\, dW_s \quad \text{in } H , \quad \forall\, t \in [0,T] .$$

Now, with our choice of k, we can apply the classical Itô formula (see [56] for example) to recover that $\mathbb{P}$-almost surely, for every $t \in [0,T]$,

$$\begin{aligned} &\frac{1}{2} \|Y_\delta(t)\|_H^2 + \int_0^t \int_D f_\delta(s) \cdot \nabla Y_\delta(s)\, ds + \int_0^t \int_D g_\delta(s) Y_\delta(s)\, ds \\ &\qquad = \frac{1}{2} \left\|Y_0^\delta\right\|_H^2 + \frac{1}{2} \int_0^t \|T_\delta(s)\|_{L_2(U,H)}^2\, ds + \int_0^t \left(Y_\delta(s), T_\delta(s)\, dW_s\right) . \end{aligned} \tag{6.7.124}$$

Now, let us focus on the stochastic integral: we have

$$\begin{aligned} &\int_0^t \left(Y_\delta(s), T_\delta(s)\, dW_s\right) - \int_0^t \left(Y(s), T(s)\, dW_s\right) \\ &\qquad = \int_0^t \left(Y_\delta(s), (T_\delta - T)(s)\, dW_s\right) + \int_0^t \left((Y_\delta - Y)(s), T(s)\, dW_s\right) , \end{aligned}$$

where thanks to the Davis inequality and (6.7.120)–(6.7.121) we have (renominating the positive constant c)

$$\begin{aligned} &\mathbb{E} \sup_{t\in[0,T]} \left| \int_0^t \left(Y_\delta(s), (T_\delta - T)(s)\, dW_s\right) \right| \\ &\qquad \le c\mathbb{E} \left[\left(\int_0^T \|Y_\delta(s)\|_H^2 \|(T_\delta - T)(s)\|_{\mathscr{L}^2(U,H)}^2\, ds \right)^{1/2} \right] \\ &\qquad \le c \|T_\delta - T\|_{L^2(\Omega\times(0,T);\mathscr{L}^2(U,H))} \to 0 \end{aligned}$$

and, by the dominated convergence theorem, also

$$\begin{aligned} &\mathbb{E} \sup_{t\in[0,T]} \left| \int_0^t \left((Y_\delta - Y)(s), T_s\, dW_s\right) \right|^2 \\ &\qquad \le c\mathbb{E} \left[\left(\int_0^T \|(Y_\delta - Y)(s)\|_H^2 \|T_s\|_{\mathscr{L}^2(U,H)}^2\, ds \right)^{1/2} \right] \to 0 . \end{aligned}$$

Hence, we have $\int_0^\cdot (Y_\delta(s), T_\delta(s)\, dW_s) \to \int_0^\cdot (Y(s), T(s)\, dW_s)$ in $L^2(\Omega; L^\infty(0,T))$, so that consequently (at least for a subsequence)

$$\int_0^t \left(Y_\delta(s), T_\delta(s)\, dW_s\right) \to \int_0^t \left(Y(s), T(s)\, dW_s\right) \quad \text{for every } t \in [0,T] , \quad \mathbb{P}\text{-a.s.}$$

Hence, letting $\delta \searrow 0$ and taking into account (6.7.119)–(6.7.123), $\mathbb{P}$-almost surely we have

$$\begin{aligned}\lim_{\delta\searrow 0}\int_{(0,t)\times D} g_\delta Y_\delta = \frac{1}{2}\left\|Y_0\right\|_H^2 + \frac{1}{2}\int_0^t \left\|T(s)\right\|^2_{\mathscr{L}^2(U,H)}\,ds + \int_0^t \left(Y(s), T(s)\,dW_s\right)\\ - \frac{1}{2}\left\|Y(t)\right\|_H^2 - \int_{(0,t)\times D} f_\delta\cdot\nabla Y \quad \text{for every } t\in[0,T]\ :\end{aligned} \tag{6.7.125}$$

we evaluate the limit on the left hand side using Vitali's convergence theorem. To this purpose, by (6.7.119) and (6.7.122) we can assume with no restriction that $Y_\delta \to Y$ and $g_\delta \to g$ almost everywhere in $\Omega\times(0,t)\times D$; moreover, thanks to the generalized Jensen inequality for positive operators (see [40, 41]), we have

$$\pm\alpha^2 g_\delta Y_\delta \leq j\left(\pm\alpha Y_\delta\right) + j^*\left(\alpha g_\delta\right) = j\left(\alpha Y_\delta\right) + j^*\left(\alpha g_\delta\right) \leq \mathcal{R}_\delta^k\left[j(\alpha Y) + j^*(\alpha g)\right]\,.$$

Thanks to (6.7.114) and the properties of $\mathcal{R}_\delta$, the term on the right hand side converges in $L^1(\Omega\times(0,t)\times D)$, hence it is uniformly integrable: consequently, we deduce that also $\{g_\delta Y_\delta\}_{\delta\in(0,1)}$ is uniformly integrable, and Vitali's convergence theorem implies that

$$g_\delta Y_\delta \to gY \quad \text{in } L^1\left(\Omega\times(0,t)\times D\right)\,, \quad \text{as } \delta\searrow 0\,,$$

so that passing to the limit in (6.7.125) we obtain (6.7.116).
To show (6.7.118), we proceed in a very similar way: note that since (6.7.112) is replaced by (6.7.117), then instead of (6.7.119) we have

$$Y_\delta(t) \to Y(t) \quad \text{in } L^2(\Omega; H)\,, \quad \text{for every } t\in[0,T]\,.$$

Once we have obtained (6.7.124) as before, we observe that the stochastic integral in (6.7.124) is a local martingale, so that there exists a sequence of increasing stopping times $\{\tau_n\}_{n\in\mathbb{N}}$ such that $\tau_n \nearrow \infty$ and the corresponding stopped processes are martingales: hence, stopping (6.7.124) at time τ_n, taking expectations and then letting $n\to\infty$, thanks to dominated convergence theorem we directly obtain for every $t\in[0,T]$

$$\begin{aligned}\frac{1}{2}\left\|Y_\delta(t)\right\|^2_{L^2(\Omega;H)} + \int_0^t\int_{\Omega\times D} f_\delta(s)\cdot\nabla Y_\delta(s)\,ds + \int_0^t\int_{\Omega\times D} g_\delta(s)Y_\delta(s)\,ds\\ = \frac{1}{2}\left\|Y_0^\delta\right\|^2_{L^2(\Omega;H)} + \frac{1}{2}\int_0^t \left\|T_\delta(s)\right\|^2_{L^2(\Omega;\mathscr{L}^2(U,H))}\,ds\,.\end{aligned}$$

At this point, (6.7.118) follows as before letting $\delta\searrow 0$ in the previous equation. □

Chapter 7

Singular equations in divergence form

In this chapter, we prove existence and uniqueness of strong solutions, as well as continuous dependence on the initial datum, for a class of fully nonlinear second-order stochastic PDEs with drift in divergence form. Due to rather general assumptions on the growth of the nonlinearity in the drift, which, in particular, is allowed to grow faster than polynomially, existing techniques are not applicable. A well-posedness result is obtained through a combination of a priori estimates on regularized equations, interpreted both as stochastic equations as well as deterministic equations with random coefficients, and weak compactness arguments. The result is essentially sharp, in the sense that no extra hypotheses are needed, bar continuity of the nonlinear function in the drift, with respect to the deterministic theory.

The results presented in this chapter are part of the joint work [67] with Carlo Marinelli, to appear on *Stochastics & Partial Differential Equations: Analysis and Computations.*

7.1 The problem: literature and main goals

Let us consider the nonlinear stochastic partial differential equation

$$du(t) - \operatorname{div}\gamma(\nabla u(t))\,dt = B(t,u(t))\,dW(t), \qquad u(0) = u_0, \tag{7.1.1}$$

on $L^2(D)$, where $D \subset \mathbb{R}^d$ is a bounded domain with smooth boundary. Here γ is the gradient of a continuously differentiable convex function on $\mathbb{R}^d$ growing faster than linearly at infinity, the divergence is interpreted in the usual variational sense, W is a cylindrical Wiener process, and B is a map with values in the space of Hilbert-Schmidt operators satisfying suitable Lipschitz continuity hypotheses. Precise assumptions on the data of the problem are given in §7.2 below.

Our main result is the well-posedness of (7.1.1), in the strong probabilistic sense, without any polynomial growth condition on γ nor any boundedness assumption on the noise (see Theorem 7.2.2 below). The lack of growth and coercivity assumptions on γ makes it impossible to apply the variational approach by Pardoux and Krylov-Rozovskiĭ (see [46, 72]), which is the only known general technique to solve nonlinear stochastic PDEs without linear terms in the drift such as (7.1.1), with the possible exception of viscosity solutions, a theory of which, however, does not seem to be available for such equations. On the other hand, we recall that, if γ is coercive and has polynomial growth, the results in *op. cit.* provide a fully satisfactory well-posedness result for (7.1.1).

The available literature dealing with stochastic equations in divergence form such as (7.1.1) is very limited and, to the best of our knowledge, entirely focused on the case where γ satisfies the above-mentioned coercivity and growth assumptions: see, e.g., [54] and the bibliography of [56] for results on the p-Laplace equation, which corresponds to the case $\gamma(x) = |x|^{p-1}x$, and Chapter 6 on stochastic equations in divergence form with doubly nonlinear drift. The main novelty of this chapter is thus to provide a satisfactory well-posedness result in the strong sense for such divergence-form equations under neither coercivity nor growth assumptions on γ. On the other hand, it is worth recalling that well-posedness results are available for other classes of monotone SPDEs with nonlinearities satisfying no coercivity and growth conditions, most notably the stochastic porous media equation: see, e.g., [12]. However, the structure of divergence-form equations such as (7.1.1) is radically different. Indeed, as is well-known, the porous media operator is quasilinear, while the divergence-type operator in (7.1.1) is fully nonlinear. Moreover, the monotonicity properties (hence the dynamics associated to the the solutions) are different: the porous media operator is monotone in H^{-1}, whereas the divergence-form operator is monotone in L^2.

As is often the case in the treatment of evolution equations of monotone type, the first step consists in the regularization of (7.1.1), replacing γ with its Yosida approximation (a monotone Lipschitz-continuous function), thus obtaining a family of equations for which well-posedness is known to hold (in our case, we also need to add a "small" elliptic term in the drift as well as to smooth the diffusion coefficient B). In a second step, one proves that the solutions to the regularized equations are compact in suitable topologies, so that, by passage to the limit in the regularization parameters (roughly speaking), a process can be constructed that, in a final step, is shown to actually be the unique solution to (7.1.1) and to depend continuously on the initial datum. It is well known that the last two steps are the more challenging ones, and our problem is no exception.

The approach we follow combines elements of the variational method and *ad hoc* arguments, most notably a priori estimates on the solutions to regularized equations, weak compactness techniques, and a generalized version of Itô's formula for the square of the norm under minimal integrability assumptions. A crucial role is played by a mix of pathwise and "averaged"* a priori estimates. Even though the approach is reminiscent of that in Chapter 2, the problem we consider here is of a completely different nature, and, correspondingly, new ideas are needed. In particular, the absence of a linear term in the drift precludes the possibility of applying a wealth of techniques available for semi-linear problems. For instance, the strong pathwise compactness criteria used in Chapter 2 are no longer available, so that we have to rely on weak compactness arguments only. This way one can construct a limit process, but its identification as a solution expectedly presents major new issues with respect to the case where stronger compactness is available. Moreover, a rather subtle measurability problem arises from the fact that the divergence is not injective, which is the reason for assuming γ to be a continuous monotone map, and not just a maximal monotone graph on $\mathbb{R}^d \times \mathbb{R}^d$. A (less regular) solution to the more general problem when γ satisfies only the latter condition will appear elsewhere. We remark that the results obtained here hold under hypotheses that are as general as those of the deterministic theory, except for the continuity assumption on γ (see, e.g., [10, pp. 207–ff.]).

*That is, in expectation.

7.2 Main result

Given a positive real number T, let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space, fixed throughout, satisfying the so-called "usual conditions". We shall denote a cylindrical Wiener process on a separable Hilbert space H by W.

Let D be a smooth bounded domain of $\mathbb{R}^d$, and assume that a map

$$B : \Omega \times [0,T] \times L^2(D) \longrightarrow \mathscr{L}^2(H, L^2(D))$$

is given such that, for a constant $C > 0$,

$$\big\|B(\omega,t,x) - B(\omega,t,y)\big\|_{\mathscr{L}^2(H,L^2(D))} \leq C\big\|x-y\big\|_{L^2(D)}$$

for all $\omega \in \Omega$, $t \in [0,T]$, $x, y \in L^2(D)$. To avoid trivial situations, we also assume that, for an $x_0 \in L^2(D)$, $B(\omega,t,x_0) < C$ for all ω and t. This implies that B grows at most linearly in x, uniformly over ω and t. Furthermore, the map $(\omega,t) \mapsto B(\omega,t,x)h$ is assumed to be measurable and adapted for all $x \in L^2(D)$ and $h \in H$.

We assume that γ is the subdifferential of a continuously differentiable convex function $k : \mathbb{R}^d \to \mathbb{R}_+$ such that $k(0) = 0$,

$$\lim_{|x|\to\infty} \frac{k(x)}{|x|} = +\infty$$

(i.e. k is superlinear at infinity), and

$$\limsup_{|x|\to\infty} \frac{k(-x)}{k(x)} < \infty.$$

Then $\gamma : \mathbb{R}^d \to \mathbb{R}^d$ is a continuous maximal monotone map, i.e.

$$\big(\gamma(x) - \gamma(y)\big) \cdot (x-y) \geq 0 \qquad \forall x, y \in \mathbb{R}^n$$

(the centered dot stands for the Euclidean scalar product in $\mathbb{R}^d$), and (the graph of) γ is maximal with respect to the order by inclusion. Moreover, the convex conjugate function $k^* : \mathbb{R}^d \to \mathbb{R}_+$ of k, defined as

$$k^*(y) = \sup_{r\in\mathbb{R}^d} \big(y \cdot r - k(r)\big),$$

is itself convex and superlinear at infinity. For these facts of convex analysis, as well as those used in the sequel, we refer to, e.g., [42].

All assumptions on B and γ (hence also on k) are assumed to be in force from now on.

Definition 7.2.1. *Let u_0 be an L^2-valued $\mathscr{F}_0$-measurable random variable. A* strong solution *to equation* (7.1.1) *is a process $u : \Omega \times [0,T] \to L^2(D)$ satisfying the following properties:*

(i) u is measurable, adapted and

$$u \in L^1(0,T; W_0^{1,1}(D))$$

(ii) $B(\cdot,u)h$ is measurable and adapted for all $h \in H$ and

$$B(\cdot,u) \in L^2(0,T; \mathscr{L}^2(H, L^2(D))) \qquad \mathbb{P}\text{-}a.s.;$$

(iii) $\gamma(\nabla u)$ *is an* $L^1(D)^d$*-valued measurable adapted process with*

$$\gamma(\nabla u) \in L^1(0,T;L^1(D)^d) \qquad \mathbb{P}\text{-a.s.};$$

(iv) one has, as an equality in $L^2(D)$,

$$u(t) - \int_0^t \operatorname{div}\gamma(\nabla u(s))\,ds = u_0 + \int_0^t B(s,u(s))\,dW(s) \qquad \mathbb{P}\text{-a.s.} \tag{7.2.2}$$

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

Since $\gamma(\nabla u)$ is only assumed to take values in $L^1(D)^d$, the second term on the left-hand side of (7.2.2) does not belong, a priori, to $L^2(D)$. The identity (7.2.2) has to be interpreted to hold in the sense of distributions, so that the term containing $\gamma(\nabla u)$ takes values in $L^2(D)$ by difference. In fact, the conditions on B in (i) imply that the stochastic integral in (7.2.2) is an $L^2(D)$-valued local martingale.

Let $\mathscr{K}$ be the set of measurable adapted processes $\phi : \Omega \times [0,T] \to L^2(D)$ such that

$$\begin{aligned}
&\mathbb{E}\sup_{t\le T}\bigl\|\phi(t)\bigr\|^2_{L^2(D)} + \mathbb{E}\int_0^T \bigl\|\phi(t)\bigr\|_{W_0^{1,1}(D)}\,dt < \infty,\\
&\mathbb{E}\int_0^T\!\!\int_D \bigl|\gamma(\nabla\phi(t,x))\bigr|\,dx\,dt < \infty,\\
&\mathbb{E}\int_0^T\!\!\int_D \bigl(k(\nabla\phi(t,x)) + k^*(\gamma(\nabla\phi(t,x)))\bigr)\,dx\,dt < \infty.
\end{aligned}$$

Our main result is the following.

Theorem 7.2.2. *Let* $u_0 \in L^2(\Omega;L^2(D))$ *be* $\mathscr{F}_0$*-measurable. Then* (7.1.1) *admits a strong solution* u*, which is unique within* $\mathscr{K}$*. Moreover,* u *has weakly continuous paths in* $L^2(D)$ *and the solution map* $u_0 \mapsto u$ *is Lipschitz-continuous from* $L^2(\Omega;L^2(D))$ *to* $L^2(\Omega;L^\infty(0,T;L^2(D)))$.

We do not know whether well-posedness continues to hold also without the condition that the solution belongs to $\mathscr{K}$. This assumption, in fact, plays a crucial role in the proof of uniqueness.

Abbreviated notation for function spaces will be used from now on: Lebesgue and Sobolev spaces on D will be denoted without explicit mention of D itself; for any $p \in [1,\infty]$, $L^p(\Omega)$ will be denoted by $\mathbb{L}^p$, $L^p(0,T)$ by L^p_t, and $L^p(D)$ sometimes by L^p_x. Mixed-norm spaces will be denoted just by juxtaposition, e.g. $\mathbb{L}^p L^q_t L^r_x$ to mean $L^p(\Omega;L^q(0,T;L^r(D)))$ and $L^1_{t,x}$ to mean $L^1([0,T]\times D)$.

7.3 An Itô formula for the square of the norm

We prove an Itô formula for the square of the L^2-norm of a class of processes with minimal integrability conditions. This is an essential tool to prove uniqueness of strong solutions and their continuous dependence on the initial datum in Sections 7.5 and 7.6 below, and it is interesting in its own right.

Proposition 7.3.1. *Assume that*

$$y(t) + \alpha\int_0^t y(s)\,ds - \int_0^t \operatorname{div}\zeta(s)\,ds = y_0 + \int_0^t C(s)\,dW(s)$$

holds in L^2 for all $t \in [0,T]$ $\mathbb{P}$-a.s., where $\alpha \geq 0$ is a constant,

$$y : \Omega \times [0,T] \to L^2, \qquad \zeta : \Omega \times [0,T] \to L^1, \qquad C : \Omega \times [0,T] \to \mathscr{L}^2(H, L^2)$$

are measurable adapted processes such that

$$y \in \mathbb{L}^2 L_t^\infty L_x^2 \cap \mathbb{L}^1 L_t^1 W_0^{1,1}, \qquad \zeta \in \mathbb{L}^1 L_{t,x}^1, \qquad C \in \mathbb{L}^2 L_t^2 \mathscr{L}^2(H, L^2),$$

and y_0 is an $\mathscr{F}_0$-measurable L^2-valued random variable with $\mathbb{E}\|y_0\|^2 < \infty$. If there exists a constant $c > 0$ such that

$$\mathbb{E} \int_0^T \int_D \big(k(c\nabla y) + k^*(c\zeta)\big) < \infty,$$

then

$$\begin{aligned}\frac{1}{2}\|y(t)\|^2 + \alpha \int_0^t \|y(s)\|^2\,ds + \int_0^t \int_D \zeta(s,x) \cdot \nabla y(s,x)\,dx\,ds \\ = \frac{1}{2}\|y_0\|^2 + \frac{1}{2}\int_0^t \|C(s)\|^2_{\mathscr{L}^2(H,L^2)}\,ds + \int_0^t y(s)C(s)\,dW(s)\end{aligned}$$

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

Proof. Note that $\operatorname{div}\zeta \in (W_0^{1,\infty})'$, hence, by Sobolev embedding theorems and duality, there exists a positive integer r such that $\operatorname{div}\zeta \in H^{-r}$. Therefore, denoting the Dirichlet Laplacian on $L^2(D)$ by Δ, there also exists a positive integer m such that $(I - \delta\Delta)^{-m}$, $\delta > 0$, maps H^{-r} and (a fortiori) L^2 to $H_0^1 \cap W^{1,\infty}$. Using the notation $h^\delta := (I - \delta\Delta)^{-m} h$, it is readily seen that

$$y^\delta(t) + \alpha \int_0^t y^\delta(s)\,ds - \int_0^t \operatorname{div}\zeta^\delta(s)\,ds = y_0^\delta + \int_0^t T^\delta(s)\,dW(s)$$

for all $t \in [0,T]$ $\mathbb{P}$-a.s. as an identity in L^2, for which Itô's formula yields

$$\begin{aligned}\frac{1}{2}\big\|y^\delta(t)\big\|^2 + \alpha \int_0^t \big\|y^\delta(s)\big\|^2\,ds + \int_0^t \int_D \zeta^\delta \cdot \nabla y^\delta \\ = \frac{1}{2}\big\|y_0^\delta\big\|^2 + \frac{1}{2}\int_0^t \big\|C^\delta(s)\big\|^2_{\mathscr{L}^2(H,L^2)}\,ds + \int_0^t y^\delta(s)C^\delta(s)\,dW(s)\end{aligned}$$

for all $t \in [0,T]$ $\mathbb{P}$-almost surely. We are going to pass to the limit as $\delta \to 0$ in this identity. The dominated convergence theorem immediately implies that, $\mathbb{P}$-a.s.,

$$\begin{aligned}\big\|y^\delta(t)\big\|^2 &\longrightarrow \big\|y(t)\big\|^2, \\ \int_0^t \big\|y^\delta(s)\big\|^2\,ds &\longrightarrow \int_0^t \big\|y(s)\big\|^2\,ds, \\ \int_0^t \big\|C^\delta(s)\big\|^2_{\mathscr{L}^2(H,L^2)}\,ds &\longrightarrow \int_0^t \big\|C(s)\big\|^2_{\mathscr{L}^2(H,L^2)}\,ds\end{aligned}$$

for all $t \in [0,T]$, and $\|y_0^\delta\|^2 \to \|y_0\|^2$, as $\delta \to 0$. Defining the real local martingales

$$M^\delta := (y^\delta C^\delta) \cdot W, \qquad M := (yC) \cdot W,$$

we are going to show that

$$\mathbb{E}\sup_{t \leq T}\big|M^\delta(t) - M(t)\big| \longrightarrow 0$$

as $\delta \to 0$. In fact, Davis' inequality for local martingales (see, e.g., [61]) yields

$$\begin{aligned}\mathbb{E}\sup_{t\le T}\bigl|M^\delta(t)-M(t)\bigr| &\lesssim \mathbb{E}\bigl[M^\delta - M, M^\delta - M\bigr]_T^{1/2}\\ &= \mathbb{E}\biggl(\int_0^T \bigl\|y^\delta(t)C^\delta(t)-y(t)C(t)\bigr\|^2_{\mathscr{L}^2(H,\mathbb{R})}\,dt\biggr)^{1/2},\end{aligned}$$

and one has, identifying $\mathscr{L}^2(H,\mathbb{R})$ with H and recalling that $(I-\delta\Delta)^{-m}$ is contractive in L^2,

$$\begin{aligned}\bigl\|y^\delta C^\delta - yC\bigr\|_H &\le \bigl\|y^\delta C^\delta - y^\delta C\bigr\|_H + \bigl\|y^\delta C - yC\bigr\|_H\\ &\le \Bigl(\sup_{t\le T}\|y(t)\|\Bigr)\bigl\|C^\delta - C\bigr\|_{\mathscr{L}^2(H,L^2)} + \bigl\|y^\delta C - yC\bigr\|_H,\end{aligned}$$

so that

$$\begin{aligned}&\mathbb{E}\biggl(\int_0^T \bigl\|y^\delta(t)C^\delta(t)-y(t)C(t)\bigr\|^2_H\,dt\biggr)^{1/2}\\ &\qquad\lesssim \mathbb{E}\sup_{t\le T}\|y(t)\|\biggl(\int_0^T \bigl\|C^\delta(t)-C(t)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,dt\biggr)^{1/2}\\ &\qquad\quad + \mathbb{E}\biggl(\int_0^T \bigl\|(y^\delta(t)-y(t))C(t)\bigr\|^2_H\,dt\biggr)^{1/2}.\end{aligned}$$

It follows by the Cauchy-Schwarz inequality that the first term on the right-hand side is dominated by

$$\Bigl(\mathbb{E}\sup_{t\le T}\|y(t)\|^2\Bigr)^{1/2}\biggl(\mathbb{E}\int_0^T \bigl\|C^\delta(t)-C(t)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,dt\biggr)^{1/2},$$

which converges to zero by properties of Hilbert-Schmidt operators and the dominated convergence theorem. Moreover,

$$\bigl\|(y^\delta(t)-y(t))C(t)\bigr\|^2_H \lesssim \bigl\|y(t)\bigr\|^2\bigl\|C(t)\bigr\|^2_{\mathscr{L}^2(H,L^2)}$$

and $y \in L^\infty_t L^2_x$, $C \in L^2_t\mathscr{L}(H,L^2_x)$ $\mathbb{P}$-a.s. imply, by dominated convergence, that

$$\int_0^T \bigl\|(y^\delta(t)-y(t))C(t)\bigr\|^2_H\,dt \longrightarrow 0$$

$\mathbb{P}$-a.s. as $\delta \to 0$. Since

$$\biggl(\int_0^T \bigl\|(y^\delta(t)-y(t))C(t)\bigr\|^2_H\,dt\biggr)^{1/2} \lesssim \sup_{t\le T}\|y(t)\|\biggl(\int_0^T \bigl\|C(t)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,dt\biggr)^{1/2}$$

and, by the Cauchy-Schwarz inequality,

$$\begin{aligned}&\mathbb{E}\sup_{t\le T}\|y(t)\|\biggl(\int_0^T \bigl\|C(t)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,dt\biggr)^{1/2}\\ &\qquad\le \Bigl(\mathbb{E}\sup_{t\le T}\|y(t)\|^2\Bigr)^{1/2}\biggl(\mathbb{E}\int_0^T \bigl\|C(t)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,dt\biggr)^{1/2} < \infty,\end{aligned}$$

again by dominated convergence it follows that

$$\mathbb{E}\biggl(\int_0^T \bigl\|(y^\delta(t)-y(t))C(t)\bigr\|^2_H\,dt\biggr)^{1/2} \longrightarrow 0$$

as $\delta \to 0$. We have thus shown that $\mathbb{E}\sup_{t\leq T}\bigl|M^\delta(t) - M(t)\bigr| \to$ as $\delta \to 0$, hence, in particular, that

$$\int_0^t y^\delta(s) C^\delta(s)\, dW(s) \longrightarrow \int_0^t y(s) C(s)\, dW(s)$$

in probability as $\delta \to 0$ for all $t \in [0,T]$.

To complete the proof, we are going to show that $\nabla Y^\delta \cdot \zeta^\delta \to \nabla Y \cdot \zeta$ in $\mathbb{L}^1 L^1_{t,x}$, which readily implies that

$$\int_0^t\!\!\int_D \nabla y^\delta(s,x) \cdot \zeta^\delta(s,x)\, dx\, ds \longrightarrow \int_0^t\!\!\int_D \nabla y(s,x) \cdot \zeta(s,x)\, dx\, ds$$

in probability for all $t \in [0,T]$. Since $\nabla y^\delta \to \nabla y$ and $\zeta^\delta \to \zeta$ in measure in $\Omega \times (0,T) \times D$, in view of Vitali's theorem, it suffices to prove that the sequence $(\nabla y^\delta \cdot \zeta^\delta)$ is uniformly integrable in $\Omega \times (0,T) \times D$. One has

$$\begin{aligned} c^2\bigl(\nabla y^\delta \cdot \zeta^\delta\bigr) &\leq k\bigl(c\nabla y^\delta\bigr) + k^*\bigl(c\zeta^\delta\bigr), \\ -c^2\bigl(\nabla y^\delta \cdot \zeta^\delta\bigr) &\leq k\bigl(c(-\nabla y^\delta)\bigr) + k^*\bigl(c\zeta^\delta\bigr) \end{aligned}$$

hence

$$\begin{aligned} c^2\bigl|\nabla y^\delta \cdot \zeta^\delta\bigr| &\lesssim k\bigl(c\nabla y^\delta\bigr) + k\bigl(c(-\nabla y^\delta)\bigr) + k^*\bigl(c\zeta^\delta\bigr) \\ &\lesssim 1 + k\bigl(c\nabla y^\delta\bigr) + k^*\bigl(c\zeta^\delta\bigr), \end{aligned}$$

where the second inequality follows by the hypothesis $\limsup_{|x|\to\infty} k(-x)/k(x) < \infty$. By Jensen's inequality for sub-Markovian operators (see [41, Theorem 3.4]) we also have

$$\begin{aligned} k\bigl(c\nabla y^\delta\bigr) &= k\bigl((I-\delta\Delta)^{-m} c\nabla y\bigr) \leq (I-\delta\Delta)^{-m} k\bigl(c\nabla y\bigr), \\ k^*\bigl(c\zeta^\delta\bigr) &= k^*\bigl((I-\delta\Delta)^{-m} c\zeta\bigr) \leq (I-\delta\Delta)^{-m} k^*\bigl(c\zeta\bigr), \end{aligned}$$

hence

$$c^2\bigl|\nabla y^\delta \cdot \zeta^\delta\bigr| \lesssim 1 + (I-\delta\Delta)^{-m}\bigl(k(c\nabla y) + k^*(c\zeta)\bigr),$$

where the right-hand side is uniformly integrable because it converges in $\mathbb{L}^1 L^1_{t,x}$ as $\delta \to 0$. This yields that $(\nabla y^\delta \cdot \zeta^\delta)$ is uniformly integrable as well, thus concluding the proof. $\square$

7.4 Well-posedness for an auxiliary SPDE

Let V_0 be a separable Hilbert space, densely and continuously embedded in H^1_0, and continuously embedded in $W^{1,\infty}$. The Sobolev embedding theorem easily implies that such a space indeed exists.

We are going to prove that the auxiliary equation

$$du(t) - \operatorname{div} \gamma(\nabla u(t))\, dt = G(t)\, dW(t), \qquad u(0) = u_0, \tag{7.4.3}$$

where G is an $\mathscr{L}^2(U, V_0)$-valued process, is well posed.

Proposition 7.4.1. *Assume that $u_0 \in \mathbb{L}^2(L^2)$ is $\mathscr{F}_0$-measurable and that $G : \Omega \times [0,T] \to$*

$\mathscr{L}^2(U,V_0)$ is measurable and adapted, with

$$\mathbb{E}\int_0^T \big\|G(t)\big\|^2_{\mathscr{L}^2(U,V_0)}\,dt < \infty.$$

Then equation (7.4.3) admits a unique strong solution u such that

$$\begin{aligned}
&\mathbb{E}\sup_{t\le T}\|u(t)\|^2 + \mathbb{E}\int_0^T \big\|u(t)\big\|_{W_0^{1,1}}\,dt < \infty,\\
&\mathbb{E}\int_0^T \big\|\gamma(\nabla u(t))\big\|_{L^1}\,dt < \infty,\\
&\int_0^T \big(\big\|k(\nabla u(t))\big\|_{L^1} + \big\|k^*(\gamma(\nabla u(t)))\big\|_{L^1}\,dt\big) < \infty \quad \mathbb{P}\text{-almost surely.}
\end{aligned}$$

Moreover, the paths of u are $\mathbb{P}$-a.s. weakly continuous with values in L^2.

The assumptions of Proposition 7.4.1 are (tacitly) assumed to hold throughout the section. The proof will be given after some preliminary results.

Let $\gamma_\lambda : \mathbb{R}^d \to \mathbb{R}^d$, $\lambda > 0$, be the Yosida regularization of γ, i.e.

$$\gamma_\lambda := \frac{1}{\lambda}\big(I - (I+\lambda\gamma)^{-1}\big), \qquad \lambda > 0,$$

and consider the regularized equation

$$du_\lambda(t) - \operatorname{div}\gamma_\lambda(\nabla u_\lambda(t))\,dt - \lambda\Delta u_\lambda(t)\,dt = G(t)\,dW(t), \qquad u_\lambda(0) = u_0.$$

Since γ_λ is monotone and Lipschitz-continuous, it is not difficult to check that the operator

$$v \longmapsto -\big(\operatorname{div}\gamma_\lambda(\nabla v) + \lambda\Delta v\big)$$

satisfies the conditions of the classical variational approach by Pardoux, Krylov and Rozovskiĭ [46, 72] on the Gelfand triple $H_0^1 \hookrightarrow L^2 \hookrightarrow H^{-1}$, hence there exists a unique adapted process u_λ with values in H_0^1 such that

$$\mathbb{E}\big\|u_\lambda\big\|^2_{C_tL^2_x} + \mathbb{E}\int_0^T \big\|u_\lambda(t)\big\|^2_{H_0^1}\,dt < \infty$$

and

$$u_\lambda(t) - \int_0^t \operatorname{div}\gamma_\lambda(\nabla u_\lambda(s))\,ds - \lambda\int_0^t \Delta u_\lambda(s)\,ds = u_0 + \int_0^t G(s)\,dW(s) \tag{7.4.4}$$

in H^{-1} for all $t \in [0,T]$.

7.4.1 A priori estimates

We are now going to establish several a priori estimates for u_λ and related processes, both pathwise and in expectation.

We begin with a simple maximal estimate for stochastic integrals that will be used several times in the sequel.

Lemma 7.4.2. *Let U, H, K be separable Hilbert spaces. If*

$$F : \Omega\times[0,T] \to \mathscr{L}(H,K), \qquad G : \Omega\times[0,T] \to \mathscr{L}^2(U,H)$$

are measurable and adapted processes such that

$$\mathbb{E}\sup_{t\le T}\big\|F(t)\big\|^2_{\mathscr{L}(H,K)} + \mathbb{E}\int_0^T \big\|G(t)\big\|^2_{\mathscr{L}^2(U,H)}\,dt < \infty,$$

then, for any $\varepsilon > 0$,

$$\begin{aligned}
&\mathbb{E}\sup_{t\le T}\Big\|\int_0^t F(s)G(s)\,dW(s)\Big\|_K\\
&\qquad\le \varepsilon\,\mathbb{E}\sup_{t\le T}\big\|F(t)\big\|^2_{\mathscr{L}(H,K)} + N(\varepsilon)\,\mathbb{E}\int_0^T \big\|G(t)\big\|^2_{\mathscr{L}^2(U,H)}\,dt.
\end{aligned}$$

Proof. By the ideal property of Hilbert-Schmidt operators (see, e.g., [20, p. V.52]), one has

$$\begin{aligned}
\big\|F(s)G(s)\big\|_{\mathscr{L}^2(U,K)} &\le \big\|F(s)\big\|_{\mathscr{L}(H,K)}\big\|G(s)\big\|_{\mathscr{L}^2(U,H)}\\
&\le \sup_{s\le T}\big\|F(s)\big\|_{\mathscr{L}(H,K)}\big\|G(s)\big\|_{\mathscr{L}^2(U,H)}
\end{aligned}$$

for all $s \in [0,T]$, hence

$$\int_0^T \big\|F(s)G(s)\big\|^2_{\mathscr{L}^2(U,K)}\,ds \le \sup_{s\le T}\big\|F(s)\big\|^2_{\mathscr{L}(H,K)}\int_0^T \big\|G(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds,$$

where the right-hand side is finite $\mathbb{P}$-a.s. thanks to the assumptions on F and G. Then $(FG)\cdot W$ is a K-valued local martingale, for which Davis' inequality yields

$$\begin{aligned}
\mathbb{E}\sup_{t\le T}\Big\|\int_0^t F(s)G(s)\,dW(s)\Big\|_K &\lesssim \mathbb{E}\big[(FG)\cdot W,(FG)\cdot W\big]_T^{1/2}\\
&= \mathbb{E}\Big(\int_0^T \big\|F(s)G(s)\big\|^2_{\mathscr{L}^2(U,K)}\,ds\Big)^{1/2}\\
&\le \mathbb{E}\sup_{s\le T}\big\|F\big\|_{\mathscr{L}(H,K)}\Big(\int_0^T \big\|G(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds\Big)^{1/2}.
\end{aligned}$$

The proof is finished invoking the elementary inequality

$$ab \le \frac{1}{2}\big(\varepsilon a^2 + \frac{1}{\varepsilon}b^2\big) \qquad \forall a,b\in\mathbb{R},\ \varepsilon>0,$$

and choosing ε properly. □

The estimate in the previous lemma will be used only in the case $K = \mathbb{R}$. The more general proof we have given is not more complicated than in the simpler case actually needed.

Lemma 7.4.3. *There exists a constant* N *such that*

$$\begin{aligned}
&\big\|u_\lambda\big\|_{\mathbb{L}^2C_tL^2_x} + \lambda^{1/2}\big\|\nabla u_\lambda\big\|_{\mathbb{L}^2L^2_{t,x}} + \big\|\gamma_\lambda(\nabla u_\lambda)\cdot\nabla u_\lambda\big\|_{\mathbb{L}^1L^1_{t,x}}\\
&\qquad< N\Big(\big\|u_0\big\|_{\mathbb{L}^2L^2_x} + \big\|G\big\|_{\mathbb{L}^2L^2_t\mathscr{L}^2(H,L^2_x)}\Big).
\end{aligned}$$

Proof. Itô's formula yields

$$\begin{aligned}\big\|u_\lambda(t)\big\|^2 + 2\int_0^t\!\!\int_D \gamma(\nabla u_\lambda(s))\cdot\nabla u_\lambda(s)\,dx\,ds + 2\lambda\int_0^t \big\|\nabla u_\lambda(s)\big\|^2\,ds\\ = \big\|u_0\big\|^2 + 2\int_0^t u_\lambda(s)G(s)\,dW(s) + \frac{1}{2}\int_0^t \big\|G(s)\big\|^2_{\mathscr{L}^2(H,L^2)}\,ds,\end{aligned}$$

where u_λ in the stochastic integral on the right-hand side has to be interpreted as taking values in $\mathscr{L}(L^2,R)\simeq L^2$. Taking supremum in time and expectation we get

$$\begin{aligned}\mathbb{E}\big\|u_\lambda\big\|^2_{C_tL^2_x} + \mathbb{E}\int_0^T\!\!\int_D \gamma_\lambda(\nabla u_\lambda(s))\cdot\nabla u_\lambda(s)\,dx\,ds + \lambda\,\mathbb{E}\big\|\nabla u_\lambda\big\|^2_{L^2_{t,x}}\\ \lesssim \mathbb{E}\big\|u_0\big\|^2 + \mathbb{E}\big\|G\big\|^2_{L^2_t\mathscr{L}^2(H,L^2))} + \mathbb{E}\sup_{t\in[0,T]}\bigg|\int_0^t u_\lambda(s)G(s)\,dW(s)\bigg|,\end{aligned}$$

where, by Lemma 1.5.1,

$$\mathbb{E}\sup_{t\in[0,T]}\bigg|\int_0^t u_\lambda(s)G(s)\,dW(s)\bigg| \le \varepsilon\,\mathbb{E}\big\|u_\lambda\big\|^2_{C_tL^2_x} + N(\varepsilon)\,\mathbb{E}\int_0^T \big\|G(s)\big\|^2_{\mathscr{L}^2(H,L^2)}\,ds$$

for any $\varepsilon > 0$. The proof is completed choosing ε small enough and recalling that γ_λ is monotone. □

Lemma 7.4.4. *The families* (∇u_λ) *and* $(\gamma_\lambda(\nabla u_\lambda))$ *are relatively weakly compact in* $\mathbb{L}^1L^1_{t,x}$.

Proof. Recall that, for any y, $r\in\mathbb{R}^d$, ones has $k(y)+k^*(r) = r\cdot y$ if and only if $r\in\partial k(y) = \gamma(y)$. Therefore, since

$$\gamma_\lambda(x)\in\partial k\big((I+\lambda\gamma)^{-1}x\big) = \gamma\big((I+\lambda\gamma)^{-1}x\big) \qquad \forall x\in\mathbb{R}^n,$$

we deduce, by the definition of γ_λ, that

$$\begin{aligned}k\big((I+\lambda\gamma)^{-1}x\big) + k^*\big(\gamma_\lambda(x)\big) &= \gamma_\lambda(x)\cdot(I+\lambda\gamma)^{-1}x\\ &= \gamma_\lambda(x)\cdot x - \lambda\big|\gamma_\lambda(x)\big|^2 \le \gamma_\lambda(x)\cdot x \qquad \forall x\in\mathbb{R}^d.\end{aligned} \tag{7.4.5}$$

By Lemma 7.4.3 we infer that there exists a constant N, independent of λ, such that

$$\mathbb{E}\int_0^T\!\!\int_D k^*\big(\gamma_\lambda(\nabla u_\lambda)\big) \le \mathbb{E}\int_0^T\!\!\int_D \gamma_\lambda(\nabla u_\lambda)\cdot\nabla u_\lambda < N.$$

Since k^* is superlinear at infinity, the family $(\gamma_\lambda(\nabla u_\lambda))$ is uniformly integrable on $\Omega\times(0,T)\times D$ by the de la Vallée Poussin criterion (see the remark on uniform integrability at the end of Chapter), hence relatively weakly compact in $\mathbb{L}^1L^1_{t,x}$ by a well-known theorem of Dunford and Pettis.

Similarly, Lemma 7.4.3 and (7.4.5) imply that there exists a constant N, independent of λ, such that

$$\mathbb{E}\int_0^T\!\!\int_D k\big((I+\lambda\gamma)^{-1}\nabla u_\lambda\big) \le \mathbb{E}\int_0^T\!\!\int_D \gamma_\lambda(\nabla u_\lambda)\cdot\nabla u_\lambda < N.$$

Since k is superlinear at infinity, the criteria by de la Vallée Poussin and Dunford-Pettis imply that the sequence $(I+\lambda\gamma)^{-1}\nabla u_\lambda$ is uniformly integrable on $\Omega\times(0,T)\times D$, hence relatively weakly compact in $\mathbb{L}^1L^1_{t,x}$. Moreover, since

$$\nabla u_\lambda = (I+\lambda\gamma)^{-1}\nabla u_\lambda + \lambda\gamma_\lambda(\nabla u_\lambda),$$

the relative weak compactness of (∇u_λ) immediately follows by the same property of $(\gamma_\lambda(\nabla u_\lambda))$ proved above. $\square$

From now on we shall assume, without loss of generality, that $\lambda \in \,]0,1]$.

Lemma 7.4.5. *There exists $\Omega' \subseteq \Omega$ with $\mathbb{P}(\Omega') = 1$ and $M : \Omega' \to \mathbb{R}$ such that*

$$\big\|u_\lambda(\omega)\big\|_{L_t^\infty L_x^2} + \sqrt{\lambda}\,\big\|\nabla u_\lambda(\omega)\big\|_{L^2_{t,x}} + \big\|k_\lambda(\nabla u_\lambda(\omega))\big\|_{L^1_{t,x}} < M(\omega)$$

for all $\omega \in \Omega'$.

Proof. Setting $v_\lambda := u_\lambda - G \cdot W$, equation (7.4.4) can be written as

$$v_\lambda(t) - \int_0^t \operatorname{div}\big(\gamma_\lambda(\nabla u_\lambda(s)) + \lambda \nabla u_\lambda(s)\big)\,ds = u_0,$$

or, equivalently, as

$$v_\lambda' - \operatorname{div}\big(\gamma_\lambda(\nabla u_\lambda) + \lambda\nabla u_\lambda\big) = 0, \qquad v_\lambda(0) = u_0. \tag{7.4.6}$$

By Itô's isometry and Doob's inequality, one has

$$\mathbb{E}\sup_{t\le T}\bigg\|\int_0^t G(s)\,dW(s)\bigg\|^2_{V_0} \lesssim \mathbb{E}\int_0^T \big\|G(s)\big\|^2_{\mathscr{L}(H,V_0)}\,ds < \infty,$$

hence $G\cdot W \in \mathbb{L}^2 L_t^\infty H_0^1$, because $V_0 \hookrightarrow H_0^1$. In particular, since $u_\lambda \in \mathbb{L}^2 L_t^\infty H_0^1$, it follows that $v_\lambda \in \mathbb{L}^2 L_t^\infty H_0^1$. Moreover, since $\operatorname{div}\gamma_\lambda(\nabla u_\lambda)$ and Δu_λ belong to $\mathbb{L}^2 L_t^2 H^{-1}$, by the previous identity we also deduce that $v_\lambda'(\omega) \in L_t^2 H^{-1}$ for $\mathbb{P}$-a.a. $\omega \in \Omega$. In particular, taking into account the hypotheses on u_0 and G, there exists $\Omega' \subset \Omega$, with $\mathbb{P}(\Omega') = 1$, such that

$$\begin{gathered} u_0(\omega) \in L_x^2, \quad G\cdot W(\omega,\cdot) \in L_t^\infty V_0, \\ v_\lambda(\omega) \in L_t^2 H_0^1, \quad v_\lambda'(\omega) \in L_t^2 H^{-1} \end{gathered}$$

for all $\omega \in \Omega'$. Let us consider from now on a fixed but arbitrary $\omega \in \Omega'$. Taking the duality pairing of (7.4.6) by v_λ and integrating (more precisely, applying Lemma 1.4.2) implies that, for all $t \in [0,T]$,

$$\begin{aligned} \frac{1}{2}\|v_\lambda(t)\|^2 + \int_0^t\!\!\int_D \gamma_\lambda(\nabla u_\lambda(s)) \cdot \nabla v_\lambda(s)\,dx\,ds& \\ + \lambda \int_0^t\!\!\int_D \nabla u_\lambda(s)\cdot\nabla v_\lambda(s)\,dx\,ds &= \frac{1}{2}\|u_0\|^2, \end{aligned}$$

where $\|u_\lambda\| \le \|v_\lambda\| + \|G\cdot W\|$, hence $\|u_\lambda\|^2 \le 2\big(\|v_\lambda\|^2 + \|G\cdot W\|^2\big)$, as well as

$$\|v_\lambda\|^2 \ge \frac{1}{2}\|u_\lambda\|^2 - \|G\cdot W\|^2.$$

Moreover, Young's inequality yields

$$\begin{aligned} \int_D \nabla u_\lambda \cdot \nabla v_\lambda &= \big\|\nabla u_\lambda\big\|^2 - \int_D \nabla u_\lambda \cdot \nabla(G\cdot W) \\ &\ge \frac{1}{2}\big\|\nabla u_\lambda\big\|^2 - \frac{1}{2}\big\|\nabla(G\cdot W)\big\|^2, \end{aligned}$$

hence also, taking into account the previous estimate,

$$\begin{aligned}\frac{1}{2}\big\|u_\lambda(t)\big\|^2 &+ 2\int_0^t\!\!\int_D \gamma_\lambda(\nabla u_\lambda(s))\cdot\nabla v_\lambda(s)\,dx\,ds + \lambda\int_0^t \big\|\nabla u_\lambda(s)\big\|^2\,ds\\ &\leq \big\|u_0\big\|^2 + \big\|G\cdot W(t)\big\|^2 + \lambda\int_0^t \big\|\nabla(G\cdot W(s))\big\|^2\,ds.\end{aligned} \tag{7.4.7}$$

Let k_λ be the Moreau-Yosida regularization of k, i.e.

$$k_\lambda(x) := \inf_{y\in\mathbb{R}^d}\Big(k(y) + \frac{|x-y|^2}{2\lambda}\Big), \qquad \lambda > 0.$$

As is well known, k_λ is a proper convex function that converges pointwise to k from below, and $\partial k_\lambda = \gamma_\lambda$. Therefore, it follows from

$$\gamma_\lambda(x)\cdot(x-y) \geq k_\lambda(x) - k_\lambda(y) \geq k_\lambda(x) - k(y) \qquad \forall x,y\in\mathbb{R}^d$$

that

$$\begin{aligned}\int_0^t\!\!\int_D &\gamma_\lambda(\nabla u_\lambda(s))\cdot\nabla v_\lambda(s)\,dx\,ds\\ &= \int_0^t\!\!\int_D \gamma_\lambda(\nabla u_\lambda(s,x))(\nabla u_\lambda(s,x) - \nabla(G\cdot W(s,x)))\,dx\,ds\\ &\geq \int_0^t\!\!\int_D k_\lambda(\nabla u_\lambda(s,x))\,dx\,ds - \int_0^t\!\!\int_D k(\nabla(G\cdot W(s,x)))\,dx\,ds,\end{aligned}$$

hence also

$$\begin{aligned}\frac{1}{2}\big\|u_\lambda(t)\big\|^2 &+ 2\int_0^t\!\!\int_D k_\lambda(\nabla u_\lambda(s,x))\,dx\,ds + \lambda\int_0^t \big\|\nabla u_\lambda(s)\big\|^2\,ds\\ &\leq \big\|u_0\big\|^2 + \big\|G\cdot W(t)\big\|^2 + \lambda\int_0^t \big\|\nabla(G\cdot W(s))\big\|^2\,ds\\ &\quad + 2\int_0^t\!\!\int_D k(\nabla(G\cdot W(s,x)))\,dx\,ds.\end{aligned}$$

Taking the supremum with respect to t yields

$$\begin{aligned}\big\|u_\lambda\big\|^2_{C_tL^2_x} &+ \big\|k_\lambda(\nabla u_\lambda)\big\|_{L^1_{t,x}} + \lambda\big\|\nabla u_\lambda\big\|^2_{L^2_{t,x}}\\ &\lesssim \big\|u_0\big\|^2_{L^2_x} + \big\|G\cdot W\big\|^2_{L^\infty_t L^2_x} + \big\|G\cdot W\big\|^2_{L^2_t H^1_0} + \big\|k(\nabla(G\cdot W))\big\|_{L^1_{t,x}}.\end{aligned}$$

As already observed above, the first three terms on the right-hand side are clearly finite. Moreover, since $V_0 \hookrightarrow W^{1,\infty}$, one has

$$\big\|k(\nabla(G\cdot W))\big\|_{L^1_{t,x}} \lesssim_{T,D} \big\|k(\nabla(G\cdot W))\big\|_{L^\infty_{t,x}} < \infty$$

by the continuity of k. Since ω was chosen arbitrarily in Ω', the proof is completed. □

Lemma 7.4.6. *There exists a set Ω', with $\mathbb{P}(\Omega') = 1$, such that, for all $\omega\in\Omega'$, the families $(\gamma_\lambda(\nabla u_\lambda))$ and (∇u_λ) are relatively weakly compact in $L^1_{t,x}$.*

Proof. Let Ω' be defined as in the proof of Lemma 7.4.5, and fix an arbitrary $\omega\in\Omega'$. By

(7.4.7), since $v_\lambda = u_\lambda - G\cdot W$, it follows that

$$\begin{aligned}\int_0^t\int_D &\gamma_\lambda(\nabla u_\lambda(s))\cdot\nabla u_\lambda(s)\,dx\,ds\\ &\leq \frac{1}{2}\|u_0\|^2+\frac{1}{2}\|G\cdot W(t)\|^2+\frac{1}{2}\int_0^t\|G\cdot W(s)\|^2_{H^1_0}\,ds\\ &\quad+\int_0^t\int_D\gamma_\lambda(\nabla u_\lambda(s))\cdot\nabla(G\cdot W(s))\,dx\,ds\end{aligned}$$

for all $t\leq T$. Thanks to Young's inequality, convexity of k^*, and $k^*(0)=0$, one has

$$\begin{aligned}\gamma_\lambda(\nabla u_\lambda)\cdot\nabla(G\cdot W) &= \frac{1}{2}\gamma_\lambda(\nabla u_\lambda)\cdot 2\nabla(G\cdot W)\\ &\leq \frac{1}{2}k^*\big(\gamma_\lambda(\nabla u_\lambda)\big)+k(2\nabla(G\cdot W)).\end{aligned}$$

Recalling that $k^*(\gamma_\lambda(x))\leq\gamma_\lambda(x)\cdot x$ for all $x\in\mathbb{R}^n$, rearranging terms one gets

$$\begin{aligned}\int_0^T\int_D k^*(\nabla u_\lambda(s))\,dx\,ds &\lesssim \|u_0\|^2+\|G\cdot W(T)\|^2+\int_0^T\|G\cdot W(t)\|^2_{H^1_0}\,ds\\ &\quad+\int_0^T\int_D k\big(2\nabla(G\cdot W(s))\big)\,dx\,ds,\end{aligned}$$

where all terms on the right-hand side are finite, as already established in the proof of Lemma 7.4.5. Appealing again to the criteria by de la Vallée Poussin and Dunford-Pettis, we immediately infer that $(\gamma_\lambda(\nabla u_\lambda(\omega,\cdot)))$ is relatively weakly compact in $L^1_{t,x}$.

Denoting by M (a constant depending on ω) the right-hand side of the previous inequality, the above estimates also yield

$$\big\|\gamma_\lambda(\nabla u_\lambda)\cdot\nabla u_\lambda\big\|_{L^1_{t,x}}\lesssim M,$$

hence also, recalling that $k((I+\lambda\gamma)^{-1}x)\leq\gamma_\lambda(x)\cdot x$,

$$\big\|k\big((I+\lambda\gamma)^{-1}\nabla u_\lambda\big)\big\|_{L^1_{t,x}}\lesssim M.$$

This implies, in complete analogy to the previous case, that $\big((I+\lambda\gamma)^{-1}\nabla u_\lambda\big)$ is relatively weakly compact in $L^1_{t,x}$. Since

$$\nabla u_\lambda=\lambda\gamma_\lambda(\nabla u_\lambda)+(I+\lambda\gamma)^{-1}\nabla u_\lambda,$$

the relative weak compactness of $(\nabla u_\lambda(\omega,\cdot))$ in $L^1_{t,x}$ follows immediately. □

7.4.2 Proof of Proposition 7.4.1

Let $\omega\in\Omega'$ be arbitrary but fixed, where Ω' is a subset of Ω with probability one, chosen as in the proof of Lemma 7.4.5. The relative weak compactness of $(\gamma_\lambda(\nabla u_\lambda))$ in $L^1_{t,x}$, proved in Lemma 7.4.6, implies that there exists $\eta\in L^1_{t,x}$ such that $\gamma_\mu(\nabla u_\mu)\to\eta$ weakly in $L^1_{t,x}$, where μ is a subsequence of λ. This in turn implies that

$$\int_0^t\operatorname{div}\gamma_\mu(\nabla u_\mu(s))\,ds\longrightarrow\int_0^t\operatorname{div}\eta(s)\,ds\qquad\text{weakly in } V_0'$$

for all $t \in [0,T]$. In fact, for any $\phi_0 \in V_0$, setting $\phi := s \mapsto 1_{[0,t]}(s)\phi_0 \in L^\infty_t V_0$, recalling that $V_0 \hookrightarrow W^{1,\infty}$, we have

$$\begin{aligned}\int_0^t \big\langle -\operatorname{div}\gamma_\mu(\nabla u_\mu(s)), \phi_0\big\rangle_{V_0}\, ds &= \int_0^T \big\langle -\operatorname{div}\gamma_\mu(\nabla u_\mu(s)), \phi(s)\big\rangle_{V_0}\, ds\\ &= \int_0^T\!\!\int_D \gamma_\mu(\nabla u_\mu(s))\cdot\nabla\phi(s)\, ds\\ &\longrightarrow \int_0^T\!\!\int_D \eta(s)\cdot\nabla\phi(s)\, ds = \int_0^t \big\langle -\operatorname{div}\eta(s), \phi_0\big\rangle\, ds\end{aligned}$$

as $\mu \to 0$. Moreover, $\sqrt{\lambda}u_\lambda$ is bounded in $L^2_t H^1_0$ thanks to Lemma 7.4.5, hence, recalling that Δ is an isomorphism of H^1_0 and H^{-1}, $\lambda\Delta u_\lambda \to 0$ in $L^2_t H^{-1}$ as $\lambda \to 0$, in particular

$$\lambda\int_0^t \Delta u_\lambda(s)\, ds \longrightarrow 0 \qquad \text{in } H^{-1}$$

for all $t \in [0,T]$ as $\lambda \to 0$. Therefore, considering the regularized equation

$$u_\mu(t) - \int_0^t \operatorname{div}\gamma_\mu(\nabla u_\mu(s))\, ds - \mu\int_0^t \Delta u_\mu(s)\, ds = u_0 + G\cdot W(t)$$

and passing to the limit as $\mu \to 0$, we infer that $u_\mu(t) \to u(t)$ weakly in V_0' for all $t \in [0,T]$, hence one can write

$$u(t) - \int_0^t \operatorname{div}\eta(s)\, ds = u_0 + G\cdot W(t) \qquad \text{in } V_0' \tag{7.4.8}$$

for all $t \in [0,T]$. Since $\operatorname{div}\eta \in L^1_t V_0'$ and $G\cdot W \in L^\infty_t V_0$, it immediately follows that $u \in C_t V_0'$. Moreover, since, thanks to Lemma 7.4.5, $(u_\mu(t))$ is bounded in L^2, we also have $u_\mu(t) \to u(t)$ weakly in L^2. In fact, let $\varepsilon > 0$ and $\psi \in L^2$ be arbitrary. Since V_0 is dense in L^2, there exists $\phi \in V_0$ with $\|\psi - \phi\| < \varepsilon$, and one can write

$$\big|\langle u_\mu(t) - u_\nu(t), \psi\rangle\big| \leq \big|\langle u_\mu(t) - u_\nu(t), \psi - \phi\rangle\big| + \big|\langle u_\mu(t) - u_\nu(t), \phi\rangle\big|,$$

where the second term on the right-hand side converges to zero as μ, $\nu \to 0$, and

$$\big|\langle u_\mu(t) - u_\nu(t), \psi - \phi\rangle\big| \leq \big\|u_\mu(t) - u_\nu(t)\big\|\,\big\|\psi - \phi\big\| < N\varepsilon,$$

so that, recalling that Hilbert spaces are weakly sequentially complete, $u_\mu(t)$ converges weakly in L^2, necessarily to $u(t)$, for all $t \in [0,T]$. This also immediately implies that $u \in L^\infty_t L^2_x$. From this, together with $u \in C_t V_0'$, it follows in turn that $u \in C_w([0,T];L^2)$ by a criterion due to Strauss (see [79, Theorem 2.1] – here and below $C_w([0,T];E)$ stands for the space of space of weakly continuous functions from $[0,T]$ to a Banach space E). Furthermore, since all terms in (7.4.8) except the second one on the left-hand side take values in L^2, it follows that (7.4.8) is satisfied also as an identity in L^2.

Let us show that $u \in L^1_t W^{1,1}_0$: the relative weak compactness of (∇u_λ) in $L^1_{t,x}$, proved in Lemma 7.4.6, implies that there exists $v \in L^1_{t,x}$ such that, along a subsequence of λ which can be assumed to coincide with μ, $\nabla u_\mu \to v$ weakly in $L^1_{t,x}$. Taking into account that $u_\mu \in H^1_0$ for all μ and that $u_\mu \to u$ weakly* in $L^\infty_t L^2_x$, it easily follows that $v = \nabla u$ a.e. in $[0,T]\times D$ and that $u \in L^1_t W^{1,1}_0$.

As a next step, we are going to show that $\eta = \gamma(\nabla u)$ a.e. in $(0,T)\times D$. For this we shall need the "energy" identity proved in the following lemma.

Lemma 7.4.7. *Assume that*

$$y(t) - \int_0^t \operatorname{div}\zeta(s)\,ds = y_0 + f(t) \qquad \textit{in } L^2 \quad \forall t \in [0,T],$$

where $y_0 \in L^2_x$, $y \in L^\infty_t L^2_x \cap L^1_t W^{1,1}_0$, $\zeta \in L^1_{t,x}$, *and* $f \in L^2_t V_0$ *with* $f(0) = 0$. *Furthermore, assume that there exists* $c > 0$ *such that*

$$k(c\nabla y) + k^*(c\zeta) \in L^1_{t,x}.$$

Then

$$\big\|y(t) - f(t)\big\|^2 + 2\int_0^t\!\!\int_D \zeta(s,x)\cdot\nabla\big(y(s,x) - f(s,x)\big)\,dx\,ds = \big\|y_0\big\|^2 \qquad \forall t \in [0,T].$$

Proof. The proof if analogous to that of Proposition 7.3.1, of which we borrow the notation and the setup. In particular, let $m \in \mathbb{N}$ be such that

$$y^\delta(t) - \int_0^t \operatorname{div}\zeta^\delta(s)\,ds = y_0^\delta + f^\delta(t) \qquad \text{in } L^2 \quad \forall t \in [0,T],$$

hence, by Lemma 1.4.2,

$$\big\|y^\delta(t) - f^\delta(t)\big\|^2 + 2\int_0^t\!\!\int_D \zeta^\delta\cdot\nabla\big(y^\delta - f^\delta\big) = \big\|y_0^\delta\big\|^2 \qquad \forall t \in [0,T],$$

where, as $\delta \to 0$, $\big\|y^\delta(t) - f^\delta(t)\big\|^2 \to \big\|y(t) - f(t)\big\|^2$ for all $t \in]0,T]$ and $\big\|y_0^\delta\big\|^2 \to \big\|y_0\big\|^2$. Moreover, since $y^\delta - f^\delta \to y - f$ in $L^1_t W^{1,1}_0$ and $\zeta^\delta \to \zeta$ in $L^1_{t,x}$, we have that, up to selecting a subsequence,

$$\zeta^\delta\cdot\nabla\big(y^\delta - f^\delta\big) \longrightarrow \zeta\cdot\nabla\big(y - f\big)$$

almost everywhere in $[0,T]\times D$. Therefore, taking Vitali's theorem into account, the lemma is proved if we show that $\zeta^\delta\cdot\nabla(y^\delta - f^\delta)$ is uniformly integrable: one has, by Young's inequality and convexity,

$$\begin{aligned}\frac{c^2}{2}\zeta^\delta\cdot\nabla(y^\delta - f^\delta) &\le k\big(c/2(\nabla y^\delta - \nabla f^\delta)\big) + k^*\big(c\zeta^\delta\big)\\ &\le \frac12 k\big(c\nabla y^\delta\big) + \frac12 k\big(c(-\nabla f^\delta)\big) + k^*\big(c\zeta^\delta\big),\end{aligned}$$

as well as

$$\begin{aligned}-\frac{c^2}{2}\zeta^\delta\cdot\nabla(y^\delta - f^\delta) &\le k\big(c/2(-\nabla y^\delta + \nabla f^\delta)\big) + k^*\big(c\zeta^\delta\big)\\ &\le \frac12 k\big(c(-\nabla y^\delta)\big) + \frac12 k\big(c\nabla f^\delta\big) + k^*\big(c\zeta^\delta\big),\end{aligned}$$

hence

$$\begin{aligned}c^2\big|\zeta^\delta\cdot\nabla(y^\delta - f^\delta)\big| &\le k\big(c\nabla y^\delta\big) + k\big(c(-\nabla y^\delta)\big)\\ &\quad + k\big(c\nabla f^\delta\big) + k\big(c(-\nabla f^\delta)\big) + 4k^*\big(c\zeta^\delta\big).\end{aligned}$$

It follows by Jensen's inequality for sub-Markovian operators, recalling that $(I - \delta\Delta)^{-m}$ and

∇ commute, that

$$c^2\big|\zeta^\delta \cdot \nabla(y^\delta - f^\delta)\big| \le (I-\delta\Delta)^{-m}\Big(k(c\nabla y) + k\big(c(-\nabla y)\big) \\ + k(c\nabla f) + k\big(c(-\nabla f)\big) + 4k^*(c\zeta)\Big),$$

where $k(c\nabla y)$ and $k^*(c\zeta)$ belong to $L^1_{t,x}$ by assumption, and the same holds for $k(c\nabla f) + k(c(-\nabla f))$ because $f \in W^{1,\infty}$. Moreover, note that the hypothesis $\limsup_{|x|\to\infty} k(-x)/k(x) < \infty$ implies that

$$\int_0^T\!\!\int_D k(c(-\nabla y)) \lesssim 1 + \int_0^T\!\!\int_D k(\nabla y) < \infty,$$

therefore, taking into account that $(I-\delta\Delta)^{-m}$ is a contraction in L^1, we obtain that $c^2|\zeta^\delta \cdot \nabla(y^\delta - f^\delta)|$ is dominated by a sequence that converges in $L^1_{t,x}$, which immediately implies that $\zeta^\delta \cdot \nabla(y^\delta - f^\delta)$ is uniformly integrable in $[0,T]\times D$. □

As in the proof of Lemma 7.4.5, it follows from (7.4.6) and Lemma 1.4.2 that

$$\frac12\|v_\lambda(t)\|^2 + \int_0^t\!\!\int_D \gamma_\lambda(\nabla u_\lambda(s))\cdot\nabla v_\lambda(s)\,dx\,ds \\ + \lambda\int_0^t\!\!\int_D \nabla u_\lambda(s)\cdot\nabla v_\lambda(s)\,dx\,ds = \frac12\|u_0\|^2$$

for all $t\in[0,T]$, where $v_\lambda = u_\lambda - G\cdot W$. This immediately implies

$$\begin{aligned}\frac12\|v_\lambda(t)\|^2 &+ \int_0^t\!\!\int_D \gamma_\lambda(\nabla u_\lambda(s))\cdot\nabla u_\lambda(s)\,dx\,ds \\ &\le \frac12\|u_0\|^2 + \int_0^t\!\!\int_D \gamma_\lambda(\nabla u_\lambda(s))\cdot\nabla(G\cdot W(s))\,dx\,ds \\ &\quad + \lambda\int_0^t\!\!\int_D \nabla u_\lambda(s)\cdot\nabla(G\cdot W(s))\,dx\,ds,\end{aligned} \tag{7.4.9}$$

where

$$\liminf_{\mu\to0}\big\|v_\mu(t)\big\| \ge \big\|u(t) - G\cdot W(t)\big\| \qquad \forall t\in[0,T]$$

by the weak lower semicontinuity of the norm and the weak convergence of $u_\mu(t)$ to $u(t)$ in L^2. Moreover, recalling that $\gamma_\mu(\nabla u_\mu)\to\eta$ weakly in $L^1_{t,x}$ and $\nabla(G\cdot W)\in L^\infty_{t,x}$, as $V_0 \hookrightarrow W^{1,\infty}$, we have

$$\int_0^t\!\!\int_D \gamma_\mu(\nabla u_\mu(s))\cdot\nabla(G\cdot W(s))\,dx\,ds \longrightarrow \int_0^t\!\!\int_D \eta(s)\cdot\nabla(G\cdot W(s))\,dx\,ds.$$

The last term on the right-hand side of (7.4.9) converges to zero as $\mu\to0$ because (∇u_μ) is bounded in $L^1_{t,x}$ and $\nabla(G\cdot W)\in L^\infty_{t,x}$. We have thus obtained

$$\limsup_{\mu\to0}\int_0^T\!\!\int_D \gamma_\mu(\nabla u_\mu(s))\cdot\nabla u_\mu(s)\,dx\,ds \\ \le \frac12\big\|u_0\big\|^2 - \frac12\big\|u(T) - G\cdot W(T)\big\|^2 + \int_0^t\!\!\int_D \eta(s)\cdot\nabla(G\cdot W(s))\,dx\,ds.$$

By Lemma 7.4.7 we have

$$\begin{aligned}\frac{1}{2}\big\|u_0\big\|^2 - \frac{1}{2}\big\|u(T) - G\cdot W(T)\big\|^2 + \int_0^T\!\!\int_D \eta(s)\cdot\nabla(G\cdot W(s))\,dx\,ds\\ = \int_0^T\!\!\int_D \eta(s)\cdot\nabla u(s)\,dx\,ds,\end{aligned}$$

which implies that

$$\limsup_{\mu\to 0}\int_0^T\!\!\int_D \gamma_\mu(\nabla u_\mu)\cdot\nabla u_\mu\,dx\,ds \le \int_0^T\!\!\int_D \eta\cdot\nabla u\,dx\,ds.$$

Moreover, since

$$\gamma_\mu(x)\cdot(I+\mu\gamma)^{-1}x = \gamma_\mu(x)\cdot x - \mu|\gamma_\mu(x)|^2 \le \gamma_\mu(x)\cdot x$$

for all $x\in\mathbb{R}^n$, we obtain

$$\limsup_{\mu\to 0}\int_0^T\!\!\int_D \gamma_\mu(\nabla u_\mu)\cdot(I+\mu\gamma)^{-1}\nabla u_\mu\,dx\,ds \le \int_0^T\!\!\int_D \eta\cdot\nabla u\,dx\,ds,$$

where $(I+\mu\gamma)^{-1}\nabla u_\mu \to \nabla u$ and $\gamma_\mu(\nabla u_\mu)\to\eta$ weakly in $L^1_{t,x}$. In particular, the weak lower semicontinuity of convex integrals yields

$$\begin{aligned}&\int_0^T\!\!\int_D \big(k(\nabla u) + k^*(\eta)\big)\\ &\qquad\le \liminf_{\mu\to 0}\int_0^T\!\!\int_D \big(k((I+\mu\gamma)^{-1}\nabla u_\mu) + k^*(\gamma_\mu(\nabla u_\mu))\big)\,dx\,dt\\ &\qquad= \liminf_{\mu\to 0}\int_0^T\!\!\int_D \gamma_\mu(\nabla u_\mu)\cdot(I+\mu\gamma)^{-1}\nabla u_\mu\,dx\,dt < N,\end{aligned}$$

where $N=N(\omega)$ is a constant. Recalling that $\gamma_\mu\in\gamma((I+\mu\gamma)^{-1})$ and $\gamma=\partial k$, we have

$$k((I+\mu\gamma)^{-1}\nabla u_\mu) + \gamma_\mu(\nabla u_\mu)\cdot(z-(I+\mu\gamma)^{-1}\nabla u_\mu) \le k(z) \qquad \forall z\in\mathbb{R}^n.$$

From this it follows, again by the weak lower semicontinuity of convex integrals, that

$$\int_0^T\!\!\int_D k(\nabla u) + \int_0^T\!\!\int_D \eta\cdot(\zeta-\nabla u) \le \int_0^T\!\!\int_D k(\zeta) \qquad \forall\zeta\in L^\infty_{t,x}.$$

Let A be an arbitrary Borel subset of $(0,T)\times D$, $z_0\in\mathbb{R}^n$, $R>0$ a constant, and

$$\zeta_R := z_0 1_A + T_R(\nabla u)1_{A^c},$$

where $T_R:\mathbb{R}^n\to\mathbb{R}^n$, is the truncation operator

$$T_R : x \longmapsto \begin{cases} x, & |x|\le R,\\ Rx/|x|, & |x|>R.\end{cases}$$

Then $\zeta_R \in L^\infty_{t,x}$, and

$$\begin{aligned}\int_A k(\nabla u) + \int_A \eta\cdot(z_0-\nabla u) &\leq \int_A k(z_0)\\ &+ \int_{A^c}\bigl(k(T_R(\nabla u)) - k(\nabla u)\bigr) + \int_{A^c}\eta\cdot\bigl(T_R(\nabla u)-\nabla u\bigr),\end{aligned}$$

where $T_R(\nabla u)\to\nabla u$ and $k(T_R(\nabla u))\to k(\nabla u)$ a.e. in $(0,T)\times D$ as $R\to\infty$, as well as

$$\bigl|T_R(\nabla u)-\nabla u\bigr| \leq 2\bigl|\nabla u\bigr|, \qquad \bigl|k(T_R(\nabla u))-k(\nabla u)\bigr| \lesssim 1+k(\nabla u)$$

(the latter inequality follows by the assumptions on the behavior of k at infinity). Since $k(\nabla u)$, $k^*(\eta)\in L^1_{t,x}$, the dominated convergence theorem implies that

$$\int_A k(\nabla u) + \int_A \eta\cdot(z_0-\nabla u) \leq \int_A k(z_0)$$

for arbitrary z_0 and A, hence also that

$$k(\nabla u) + \eta\cdot(z_0-\nabla u) \leq k(z_0)$$

a.e. in $(0,T)\times D$ for all $z_0\in\mathbb{R}^n$. By definition of subdifferential it follows immediately that $\eta=\gamma(\nabla u)$ a.e. in $(0,T)\times D$.

Let us now show, still keeping ω fixed, that the limit u constructed above is unique. In particular, since $\eta=\gamma(\nabla u)$, it is also unique. Assume that there exist u_1, u_2 such that

$$u_i(t) - \int_0^t \operatorname{div}\gamma(\nabla u_i(s))\,ds = u_0 + G\cdot W(t), \qquad i=1,2,$$

in L^2 for all $t\in[0,T]$. Setting $v=u_1-u_2$ and $\zeta=\gamma(\nabla u_1)-\gamma(\nabla u_2)$, it is enough to show that

$$v(t) - \int_0^t \operatorname{div}\zeta(s)\,ds = 0$$

in L^2 for all $t\in[0,T]$ implies $v=0$. To this aim, it suffices to note that, by Lemma 7.4.7,

$$\frac{1}{2}\bigl\|v(t)\bigr\|^2 + \int_0^t\!\!\int_D \zeta\cdot\nabla v = 0$$

for all $t\in[0,T]$. The monotonicity of γ immediately implies $v=0$, i.e. $u_1=u_2$, so that uniqueness of u is proved.

The process u has been constructed for each ω in a set of probability one via limiting procedures along sequences that depend on ω itself. Of course such a construction, in general, does not produce a measurable process. In our situation, however, uniqueness of u allows us to even prove that u is predictable. The following simple observation is crucial: we have proved that from any subsequence of λ one can extract a further subsequence μ, depending on ω, such that u_μ converges to u as $\mu\to 0$, in several topologies, and that the limit u is unique. This implies, by a classical criterion, that the same convergences hold along the original sequence λ, which does *not* depend on ω. In particular, $u_\lambda(\omega,t)\to u(\omega,t)$ weakly in L^2 for all $t\in[0,T]$ and for $\mathbb{P}$-a.s. ω. Let us show that $u_\lambda\to u$ weakly in $\mathbb{L}^1L^1_tL^2_x$: for an arbitrary $\phi\in\mathbb{L}^\infty L^\infty_tL^2_x$, we have

$$\bigl\langle u_\lambda(\omega,t),\phi(\omega,t)\bigr\rangle \longrightarrow \bigl\langle u(\omega,t),\phi(\omega,t)\bigr\rangle$$

a.e. in $\Omega \times [0,T]$, as well as

$$\begin{aligned}\mathbb{E}\int_0^T \langle u_\lambda(\omega,t),\phi(\omega,t)\rangle^2\,dt &\leq \mathbb{E}\int_0^T \|u_\lambda(\omega,t)\|^2\|\phi(\omega,t)\|^2\,dt\\ &\leq \|\phi\|^2_{\mathbb{L}^\infty L^\infty_t L^2_x}\,\mathbb{E}\int_0^T \|u_\lambda(\omega,t)\|^2\,dt < N\end{aligned}$$

for a constant N independent of λ, since (u_λ) is bounded in $\mathbb{L}^2L^2_{t,x}$ by Lemma 7.4.3. Then $\langle u_\lambda,\phi\rangle$ is uniformly integrable in $\Omega\times[0,T]$ by the criterion of de la Vallée Poussin, hence $\langle u_\lambda,\phi\rangle \to \langle u,\phi\rangle$ in $\mathbb{L}^1L^1_t$ by Vitali's theorem. Since $\phi \in \mathbb{L}^\infty L^\infty_t L^2_x$ is arbitrary, it follows that $u_\lambda \to u$ weakly in $\mathbb{L}^1L^1_tL^2_x$. Mazur's lemma (see, e.g., [20, p. 360]) implies that there exists a sequence (ζ_n) of convex combinations of u_λ such that $\zeta_n(\omega,t)\to u(\omega,t)$ in L^2 in $\mathbb{P}\otimes dt$-measure, hence a.e. in $\Omega\times[0,T]$ along a subsequence. Since (u_λ) is a collection of L^2-valued predictable processes, the same holds for (ζ_n), so that the $\mathbb{P}\otimes dt$-a.e. pointwise limit u of (a subsequence of) ζ_n is an L^2-valued predictable process as well. We also have that $u\in\mathbb{L}^2L^\infty_tL^2_x$, as it follows by $u_\lambda\to u$ in $\mathbb{L}^1L^1_tL^2_x$ and the boundedness of (u_λ) in $\mathbb{L}^2L^\infty_tL^2_x$.

Moreover, recalling that $\nabla u_\lambda \to \nabla u$ and $\gamma_\lambda(\nabla u_\lambda)\to\eta$ weakly in $L^1_{t,x}$ $\mathbb{P}$-a.s., and that, by Lemma 7.4.4, (∇u_λ) and $(\gamma_\lambda(\nabla u_\lambda))$ are bounded in $\mathbb{L}^1L^1_{t,x}$, an entirely analogous argument shows that $\nabla u_\lambda\to\nabla u$ and $\gamma_\lambda(\nabla u_\lambda)\to\eta=\gamma(\nabla u)$ weakly in $\mathbb{L}^1L^1_{t,x}$. This implies that η is a measurable adapted process, as well as, by weak lower semicontinuity of the norm,

$$\begin{aligned}\mathbb{E}\|\nabla u\|_{L^1_{t,x}} &\leq \liminf_{\lambda\to0}\mathbb{E}\|\nabla u_\lambda\|_{L^1_{t,x}}<\infty,\\ \mathbb{E}\|\eta\|_{L^1_{t,x}} &\leq \liminf_{\lambda\to0}\mathbb{E}\|\gamma_\lambda(\nabla u_\lambda)\|_{L^1_{t,x}}<\infty.\end{aligned}$$

We can hence conclude that

$$\begin{aligned}&u\in\mathbb{L}^2L^\infty_tL^2_x\,\cap\,\mathbb{L}^1L^1_tW^{1,1}_0,\\ &\eta=\gamma(\nabla u)\in\mathbb{L}^1L^1_{t,x}.\end{aligned}$$

Finally, Lemma 7.4.3 and (7.4.5) yield

$$\begin{aligned}&\mathbb{E}\int_0^T\int_D\big(k((I+\lambda\gamma)^{-1}\nabla u_\lambda)+k^*(\gamma_\lambda(\nabla u_\lambda))\big)\\ &\qquad< N\Big(\mathbb{E}\|u_0\|^2+\mathbb{E}\int_0^T\|G(s)\|^2_{\mathscr{L}(H,L^2)}\,ds\Big),\end{aligned}$$

where, by the weak lower semicontinuity of convex integrals and $(I+\lambda\gamma)^{-1}\nabla u_\lambda \to \nabla u$, $\gamma_\lambda(\nabla u_\lambda)\to\eta$ weakly in $L^1_{t,x}$ $\mathbb{P}$-a.s., one has

$$\int_0^T\int_D\big(k(\nabla u)+k^*(\eta)\big)\leq\liminf_{\lambda\to0}\int_0^T\int_D\big(k((I+\lambda\gamma)^{-1}\nabla u_\lambda)+k^*(\gamma_\lambda(\nabla u_\lambda))\big)$$

$\mathbb{P}$-a.s., hence, by Fatou's lemma,

$$\begin{aligned}\mathbb{E}\int_0^T\int_D\big(k(\nabla u)+k^*(\eta)\big)&\leq\liminf_{\lambda\to0}\mathbb{E}\int_0^T\int_D\big(k((I+\lambda\gamma)^{-1}\nabla u_\lambda)+k^*(\gamma_\lambda(\nabla u_\lambda))\big)\\ &< N\Big(\mathbb{E}\|u_0\|^2+\mathbb{E}\int_0^T\|G(s)\|^2_{\mathscr{L}(H,L^2)}\,ds\Big)<\infty.\end{aligned}\tag{7.4.10}$$

Remark 7.4.8. The proof of uniqueness of u does *not* depend on γ being single-valued. In particular, all results on u obtained thus far, including the predictability of u, can be obtained under the more general assumption that γ is an everywhere defined maximal monotone graph on $\mathbb{R}^n \times \mathbb{R}^n$, with $\gamma = \partial k$. However, in this more general framework, the uniqueness of η does *not* follow, because the divergence is not injective. This implies that we would not be able even to prove that η is a measurable process (with respect to the product σ-algebra of $\mathscr{F}$ and the Borel σ-algebra of $[0,T]$).

7.5 Well-posedness with additive noise

We are now going to prove well-posedness for the equation

$$du(t) - \operatorname{div} \gamma(\nabla u(t))\, dt = G(t)\, dW(t), \qquad u(0) = u_0, \tag{7.5.11}$$

where G is no longer supposed to take values in $\mathscr{L}^2(H, V_0)$, as in the previous section, but just in $\mathscr{L}^2(H, L^2)$. In other words, we are considering equation (7.1.1) with additive noise.

Proposition 7.5.1. *Assume that $u_0 \in \mathbb{L}^2 L^2_x$ is $\mathscr{F}_0$-measurable and that $G : \Omega \times [0,T] \to \mathscr{L}^2(H, L^2)$ is measurable and adapted. Then equation* (7.4.3) *is well posed in $\mathscr{K}$. Moreover, its solution is pathwise weakly continuous with values in L^2.*

Proof. Since one has $(I - \varepsilon\Delta)^{-m} : L^2 \to H^{2m} \cap H^1_0$ for any $m \in \mathbb{N}$, choosing $m > 1/2 + n/4$, the Sobolev embedding theorem yields $H^{2m} \hookrightarrow W^{1,\infty}$, hence $V_0 := H^{2m} \cap H^1_0$ satisfies all hypotheses stated at the beginning of the previous section. In particular, setting

$$G^\varepsilon := (I - \varepsilon\Delta)^{-m} G,$$

the ideal property of Hilbert-Schmidt operators implies that G^ε is an $\mathscr{L}^2(H, V_0)$-valued measurable and adapted process such that

$$\mathbb{E} \int_0^T \big\|G^\varepsilon(s)\big\|^2_{\mathscr{L}^2(H,V_0)}\, ds \lesssim \mathbb{E} \int_0^T \big\|G(s)\big\|^2_{\mathscr{L}^2(H,L^2)}\, ds < \infty.$$

It follows by Proposition 7.4.1 that, for any $\varepsilon > 0$, there exists a unique predictable process

$$u^\varepsilon \in \mathbb{L}^2 L^\infty_t L^2_x \cap \mathbb{L}^1 L^1_t W^{1,1}_0$$

such that

$$\begin{gathered} \eta^\varepsilon = \gamma(u^\varepsilon) \in \mathbb{L}^1 L^1_{t,x}, \\ k(\nabla u^\varepsilon) + k^*(\eta^\varepsilon) \in L^1_{t,x} \quad \mathbb{P}\text{-a.s.}, \\ u^\varepsilon \in C_w([0,T]; L^2) \quad \mathbb{P}\text{-a.s.}, \end{gathered}$$

satisfying

$$u^\varepsilon(t) - \int_0^t \operatorname{div} \eta^\varepsilon(s)\, ds = u_0 + \int_0^t G^\varepsilon(s)\, dW(s) \tag{7.5.12}$$

in L^2 for all $t \in [0,T]$.

In complete analogy to the previous section, the equation in H^{-1}

$$u^\varepsilon_\lambda(t) - \int_0^t \operatorname{div} \gamma_\lambda(\nabla u^\varepsilon_\lambda(s))\, ds - \lambda \int_0 \Delta u^\varepsilon_\lambda(s)\, ds = u_0 + \int_0^t G^\varepsilon(s)\, dW(s)$$

admits a unique (variational) strong solution u^ε_λ for any $\varepsilon > 0$ and $\lambda > 0$. Taking into account the monotonicity of γ_λ, Itô's formula yields, for any $\delta > 0$,

$$
\begin{aligned}
&\bigl\|u^\varepsilon_\lambda(t) - u^\delta_\lambda(t)\bigr\|^2 + \lambda \int_0^t \bigl\|\nabla(u^\varepsilon_\lambda(s) - u^\delta_\lambda(s))\bigr\|^2\,ds \\
&\lesssim \int_0^t \bigl(u^\varepsilon_\lambda(s) - u^\delta_\lambda(s)\bigr)\bigl(G^\varepsilon(s) - G^\delta(s)\bigr)\,dW(s) + \int_0^t \bigl\|G^\varepsilon(s) - G^\delta(s)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,ds.
\end{aligned}
$$

Taking supremum in time and expectation, it easily follows from Lemma 7.4.2 that

$$
\mathbb{E}\sup_{t\le T}\bigl\|u^\varepsilon_\lambda(t) - u^\delta_\lambda(t)\bigr\|^2 \lesssim \mathbb{E}\int_0^T \bigl\|G^\varepsilon(t) - G^\delta(t)\bigr\|^2_{\mathscr{L}^2(H,L^2)}.
$$

For arbitrary fixed ε, $\delta > 0$, the proof of Proposition 7.4.1 shows that

$$
\begin{aligned}
u^\varepsilon_\lambda &\longrightarrow u^\varepsilon && \text{weakly* in } L^\infty_t L^2_x,\\
\nabla u^\varepsilon_\lambda &\longrightarrow \nabla u^\varepsilon && \text{weakly in } L^1_{t,x},\\
\gamma_\lambda(\nabla u^\varepsilon_\lambda) &\longrightarrow \eta^\varepsilon && \text{weakly in } L^1_{t,x}
\end{aligned}
$$

$\mathbb{P}$-a.s. as $\lambda \to 0$, and the same holds replacing ε with δ. In particular, on a set of probability one, $u^\varepsilon_\lambda - u^\delta_\lambda \to u^\varepsilon - u^\delta$ weakly* in $L^\infty_t L^2_x$ as $\lambda \to 0$, hence the weak* lower semicontinuity of the norm and Fatou's lemma imply

$$
\begin{aligned}
\mathbb{E}\bigl\|u^\varepsilon - u^\delta\bigr\|^2_{L^\infty_t L^2_x} &\le \liminf_{\lambda\to 0}\mathbb{E}\bigl\|u^\varepsilon_\lambda - u^\delta_\lambda\bigr\|^2_{L^\infty_t L^2_x}\\
&\lesssim \mathbb{E}\int_0^T \bigl\|G^\varepsilon(s) - G^\delta(s)\bigr\|^2_{\mathscr{L}^2(H,L^2))}\,ds.
\end{aligned}
$$

It follows by the ideal property of Hilbert-Schmidt operators, the contractivity of $(I - \varepsilon\Delta)^{-m}$, and the dominated convergence theorem, that

$$
\mathbb{E}\int_0^T \bigl\|G^\varepsilon(s) - G(s)\bigr\|^2_{\mathscr{L}^2(H,L^2))}\,ds \longrightarrow 0
$$

as $\varepsilon \to 0$. This implies that (u^ε) is a Cauchy sequence in $\mathbb{L}^2 L^\infty_t L^2_x$, hence there exists a predictable L^2-valued process u such that u^ε converges (strongly) to u in $\mathbb{L}^2 L^\infty_t L^2_x$ as $\varepsilon \to 0$. Moreover, by (7.4.10) there exists a constant N, independent of ε, such that

$$
\begin{aligned}
&\mathbb{E}\int_0^T\int_D \bigl(k(\nabla u^\varepsilon) + k^*(\eta^\varepsilon)\bigr)\,dx\,ds\\
&\qquad < N\Bigl(\mathbb{E}\bigl\|u_0\bigr\|^2 + \mathbb{E}\int_0^T \bigl\|G^\varepsilon(s)\bigr\|^2_{\mathscr{L}(H,L^2)}\,ds\Bigr)\\
&\qquad \le N\Bigl(\mathbb{E}\bigl\|u_0\bigr\|^2 + \mathbb{E}\int_0^T \bigl\|G(s)\bigr\|^2_{\mathscr{L}(H,L^2)}\,ds\Bigr),
\end{aligned}
\tag{7.5.13}
$$

where we have used again the ideal property of Hilbert-Schmidt operators and the contractivity of $(I-\varepsilon\Delta)^{-m}$ in the last step. The sequences (∇u^ε) and $(\gamma(\nabla u^\varepsilon))$ are hence uniformly integrable on $\Omega \times [0,T] \times D$ by the criterion of de la Vallée Poussin, hence relatively weakly compact in $\mathbb{L}^1(L^1_{t,x})$ by the Dunford-Pettis theorem. Therefore, passing to a subsequence of ε, denoted by the same symbol, there exist v and η such that $\nabla u^\varepsilon \to v$ and $\gamma(\nabla u^\varepsilon) \to \eta$ weakly in $\mathbb{L}^1 L^1_{t,x}$ as

$\varepsilon \to 0$. It is then straightforward to check that $v = \nabla u$ and

$$u \in \mathbb{L}^1 L^1_t W^{1,1}_0.$$

An argument based on Mazur's lemma, entirely analogous to the one used in the proof of Proposition 7.4.1, shows that η is an L^1-valued adapted process.

We can now pass to the limit as $\varepsilon \to 0$ in (7.5.12). The strong convergence of u^ε to u in $\mathbb{L}^2 L^\infty_t L^2_x$ implies that

$$\operatorname*{ess\,sup}_{t\in[0,T]} \big\| u^\varepsilon(t) - u(t) \big\| \to 0$$

in probability as $\varepsilon \to 0$. Let $\phi_0 \in V_0$ be arbitrary. Since $V_0 \hookrightarrow L^\infty$, one has

$$\big\langle u^\varepsilon(t), \phi_0 \big\rangle \to \big\langle u(t), \phi_0 \big\rangle$$

in probability for almost all $t \in [0,T]$. Let us set, for an arbitrary but fixed $t \in [0,T]$, $\phi : s \mapsto 1_{[0,t]}(s)\phi_0 \in L^\infty_t V_0$. Recalling that $\eta^\varepsilon = \gamma(\nabla u^\varepsilon) \to \eta$ weakly in $\mathbb{L}^1 L^1_{t,x}$, it follows immediately that

$$\begin{aligned} -\int_0^t \langle \operatorname{div} \eta^\varepsilon, \phi_0 \rangle \, ds &= \int_0^T \int_D \eta^\varepsilon(s) \cdot \phi(s) \, ds \\ &\to \int_0^T \int_D \eta(s) \cdot \nabla\phi(s) \, ds = -\int_0^t \langle \operatorname{div} \eta(s), \phi_0 \rangle \, ds \end{aligned}$$

weakly in $\mathbb{L}^1$ as $\varepsilon \to 0$. Doob's maximal inequality and the convergence

$$\mathbb{E} \int_0^T \big\| G^\varepsilon(t) - G(t) \big\|_{\mathscr{L}^2(H,L^2)} \longrightarrow 0$$

as $\varepsilon \to 0$ readily yield also that $G^\varepsilon \cdot W(t) \to G \cdot W(t)$ in L^2 in probability for all $t \in [0,T]$. In particular, since $\phi_0 \in V_0$ and $t \in [0,T]$ are arbitrary, we infer that

$$u(t) - \int_0^t \operatorname{div} \eta(s) \, ds = u_0 + \int_0^t B(s) \, dW(s)$$

holds in V_0' for almost all t. Recalling that $\eta \in L^1_{t,x}$, which implies in turn that $\operatorname{div} \eta \in L^1_t V_0'$, it follows that all terms except the first on the left-hand side have trajectories in $C_t V_0'$, hence that the identity holds for all $t \in [0,T]$. Moreover, thanks to Strauss' weak continuity criterion, $u \in C_t V_0'$ and $u \in L^\infty_t L^2_x$ imply $u \in C_w([0,T]; L^2)$. Note also that all terms bar the second one on the left-hand side are L^2-valued, hence the identity holds also in L^2 for all $t \in [0,T]$.

The weak convergences $\nabla u^\varepsilon \to \nabla u$ and $\eta^\varepsilon \to \eta$ in $\mathbb{L}^1 L^1_{t,x}$ and the weak lower semicontinuity of convex integrals yield, taking (7.5.13) into account,

$$\mathbb{E} \int_0^T \int_D \big(k(\nabla u) + k^*(\eta) \big) < N \Big(\mathbb{E} \big\| u_0 \big\|^2 + \mathbb{E} \int_0^T \big\| G(s) \big\|^2_{\mathscr{L}^2(H,L^2)} \, ds \Big).$$

To complete the proof of existence, we only need to show that $\eta = \gamma(\nabla u)$ a.e. in $\Omega \times (0,T) \times D$. Note that, by Proposition 7.3.1, we have

$$\begin{aligned} &\frac{1}{2} \big\| u^\varepsilon(T) \big\|^2 + \int_0^T \int_D \eta^\varepsilon \cdot \nabla u^\varepsilon \\ &\qquad = \frac{1}{2} \big\| u_0 \big\|^2 + \frac{1}{2} \int_0^T \big\| G^\varepsilon(s) \big\|^2_{\mathscr{L}^2(H,L^2)} \, ds + \int_0^T u^\varepsilon(s) G^\varepsilon(s) \, dW(s) \end{aligned}$$

and

$$\begin{aligned}
&\frac{1}{2}\bigl\|u(T)\bigr\|^2 + \int_0^T\!\!\int_D \eta\cdot\nabla u\\
&\qquad = \frac{1}{2}\bigl\|u_0\bigr\|^2 + \frac{1}{2}\int_0^T \bigl\|G(s)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,ds + \int_0^T u(s)G(s)\,dW(s),
\end{aligned}$$

where, as $\varepsilon \to 0$, $\|u^\varepsilon(T)\| \to \|u(T)\|$ in $\mathbb{L}^2$, thanks to the strong convergence of u^ε to u in $\mathbb{L}^2 L^\infty_t L^2_x$;

$$\int_0^T \bigl\|G^\varepsilon(s)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,ds \longrightarrow \int_0^T \bigl\|G(s)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,ds$$

in $\mathbb{L}^2$ by an (already seen) argument involving the ideal property of Hilbert-Schmidt operators;

$$\int_0^T u^\varepsilon(s)G^\varepsilon(s)\,dW(s) \longrightarrow \int_0^T u(s)G(s)\,dW(s)$$

in $\mathbb{L}^1$ as it follows by Lemma 7.4.2. In particular, we infer

$$\begin{aligned}
&\limsup_{\varepsilon\to 0}\int_0^T\!\!\int_D \gamma(\nabla u^\varepsilon)\cdot\nabla u^\varepsilon\\
&\qquad \le \frac{1}{2}\bigl\|u_0\bigr\|^2 - \frac{1}{2}\bigl\|u(T)\bigr\|^2 + \frac{1}{2}\int_0^t \bigl\|G(s)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,ds + \int_0^t u(s)G(s)\,dW(s)\\
&\qquad = \int_0^t\!\!\int_D \eta\cdot\nabla u,
\end{aligned}$$

hence also, by Fatou's lemma,

$$\limsup_{\varepsilon\to 0}\mathbb{E}\int_0^T\!\!\int_D \gamma(\nabla u^\varepsilon)\cdot\nabla u^\varepsilon \le \mathbb{E}\int_0^t\!\!\int_D \eta\cdot\nabla u.$$

Since $\nabla u^\varepsilon \to \nabla u$ and $\gamma(\nabla u^\varepsilon) \to \eta$ weakly in $\mathbb{L}^1 L^1_{t,x}$, recalling that γ is maximal monotone, it follows that $\eta \in \gamma(\nabla u)$ a.e. in $\Omega\times(0,T)\times D$ (see, e.g., [10, Lemma 2.3, p. 38]).

Let u_{01}, $u_{02} \in \mathbb{L}^2 L^2_x$ be $\mathscr{F}_0$-measurable, and G_1, $G_2 : \Omega\times[0,T] \to \mathscr{L}^2(H,L^2)$ be measurable adapted processes such that

$$\mathbb{E}\int_0^T \bigl\|G_i(s)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,ds < \infty, \qquad i = 1,2.$$

If $u_i \in \mathscr{K}$, $i = 1,2$, are solutions to

$$du_i - \operatorname{div}\gamma(\nabla u_i)\,dt = G_i\,dW, \qquad u_i(0) = u_{0i},$$

we are going to show that

$$\mathbb{E}\sup_{t\le T}\bigl\|u_1(t) - u_2(t)\bigr\|^2 \lesssim \mathbb{E}\bigl\|u_{01} - u_{02}\bigr\|^2 + \mathbb{E}\int_0^T \bigl\|G_1(s) - G_2(s)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,ds, \tag{7.5.14}$$

from which uniqueness and Lipschitz-continuous dependence on the initial datum follow immediately. We shall actually obtain this estimate as a special case of a more general one that will be useful in the next section: setting

$$y(t) := u_1(t) - u_2(t), \qquad y_0 := u_{01} - u_{02}, \qquad F(t) := G_1(t) - G_2(t),$$

one has

$$y(t) - \int_0^t \operatorname{div} \zeta(s)\, ds = y_0 + \int_0^t F(s)\, dW(s),$$

where $\zeta = \gamma(\nabla u_1) - \gamma(\nabla u_2)$. Setting, for any $\alpha \geq 0$,

$$y^\alpha(t) := e^{-\alpha t} y(t), \qquad \zeta(t) := e^{-\alpha t}\zeta(t), \qquad F^\alpha(t) := e^{-\alpha t}F(t),$$

the integration by parts formula yields

$$y^\alpha(t) + \int_0^t \bigl(\alpha y^\alpha(s) - \operatorname{div} \zeta^\alpha(s)\bigr)\, ds = y_0 + \int_0^t F^\alpha(s)\, dW(s),$$

from which, by Proposition 7.3.1, we deduce

$$\begin{aligned}
&\bigl\|y^\alpha(t)\bigr\|^2 + 2\alpha \int_0^t \bigl\|y^\alpha(s)\bigr\|^2\, ds + 2\int_0^t\!\!\int_D \zeta^\alpha(s)\cdot\nabla y^\alpha(s)\, ds\\
&\qquad \leq \bigl\|y_0\bigr\|^2 + 2\int_0^t y^\alpha(s)F^\alpha(s)\, dW(s) + \int_0^t \bigl\|F^\alpha(s)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\, ds,
\end{aligned}$$

where, by monotonicity of γ, $\zeta^\alpha\cdot\nabla y^\alpha = e^{-2\alpha\cdot}\bigl(\gamma(\nabla u_1)-\gamma(\nabla u_2)\bigr)\cdot(\nabla u_1 - \nabla u_2) \geq 0$. Therefore, taking the supremum in t and expectation on both sides, one has

$$\begin{aligned}
&\mathbb{E}\sup_{t\leq T}\bigl\|y^\alpha(t)\bigr\|^2 + \alpha\,\mathbb{E}\int_0^T \bigl\|y^\alpha(s)\bigr\|^2\, ds\\
&\qquad \lesssim \mathbb{E}\bigl\|y_0\bigr\|^2 + \mathbb{E}\sup_{t\leq T}\biggl|\int_0^t y^\alpha(s)F^\alpha(s)\, dW(s)\biggr| + \mathbb{E}\int_0^T \bigl\|F^\alpha(s)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\, ds\\
&\qquad \lesssim \mathbb{E}\bigl\|y_0\bigr\|^2 + \mathbb{E}\int_0^T \bigl\|F^\alpha(s)\bigr\|^2_{\mathscr{L}^2(H,L^2)}\, ds, \qquad (7.5.15)
\end{aligned}$$

where the second inequality follows by an application of Lemma 1.5.1. Estimate (7.5.14) is just the special case $\alpha = 0$. □

7.6 Proof of the main result

Thanks to the results established thus far, we are now in the position to prove Theorem 7.2.2. Let $v : \Omega\times[0,T]\to L^2$ be a measurable adapted process such that

$$\mathbb{E}\int_0^T \bigl\|v(s)\bigr\|^2\, ds < \infty,$$

and consider the equation

$$du(t) - \operatorname{div}\gamma(\nabla u(t))\, dt = B(t, v(t))\, dW(t), \qquad u(0) = u_0,$$

where u_0 is an $\mathscr{F}_0$-measurable L^2-valued random variable with finite second moment. The assumptions on B imply that $B(\cdot, v)$ is measurable, adapted, and such that

$$\mathbb{E}\int_0^T \bigl\|B(s,v(s))\bigr\|^2_{\mathscr{L}^2(H,L^2)}\, ds < \infty,$$

hence the above equation is well-posed in $\mathscr{K}$ by Proposition 7.5.1, which allows one to define a map $\Gamma : (u_0, v) \mapsto u$. Let $u_i = \Gamma(u_{0i}, v_i)$, $i = 1, 2$, where u_{0i} and v_i satisfy the same measurability and integrability assumptions on u_0 and v, respectively. For any $\alpha \geq 0$, (7.5.15) and the Lipschitz continuity of B yield

$$
\begin{aligned}
&\mathbb{E}\sup_{t\leq T}\bigl(e^{-2\alpha t}\bigl\|u_1(t)-u_2(t)\bigr\|^2\bigr) + \mathbb{E}\int_0^T e^{-2\alpha s}\bigl\|u_1(s)-u_2(s)\bigr\|^2\,ds\\
&\quad\lesssim \frac{1}{\alpha}\,\mathbb{E}\bigl\|u_{01}-u_{02}\bigr\|^2 + \frac{1}{\alpha}\,\mathbb{E}\int_0^T e^{-2\alpha s}\bigl\|B(s,v_1(s))-B(s,v_2(s))\bigr\|^2_{\mathscr{L}^2(H,L^2)}\,ds\\
&\quad\lesssim \frac{1}{\alpha}\,\mathbb{E}\bigl\|u_{01}-u_{02}\bigr\|^2 + \frac{1}{\alpha}\,\mathbb{E}\int_0^T e^{-2\alpha s}\bigl\|v_1(s)-v_2(s)\bigr\|^2\,ds.
\end{aligned}
$$

Choosing α large enough, it follows that, for any u_0 as above, the map $v \mapsto \Gamma(u_0, v)$ is strictly contractive in the Banach space E_α of measurable adapted processes v such that

$$
\|v\|_{E_\alpha} := \left(\mathbb{E}\int_0^T e^{-2\alpha s}\|v(s)\|^2\,ds\right)^{1/2}.
$$

By the Banach fixed point theorem, the map $v \mapsto \Gamma(u_0, v)$ admits a unique fixed point u in E_α. Since all E_α-norms are equivalent for different values of α, u belongs to E_0 and, by definition of Γ, u also belongs to $\mathscr{K}$ and solves (7.1.1). Taking into account that any solution to (7.1.1) is necessarily a fixed point of $v \mapsto \Gamma(u_0, v)$, it immediately follows that u is the unique solution to (7.1.1) in $\mathscr{K}$. Lipschitz continuity of the solution map follows from the above estimate, which manifestly implies

$$
\mathbb{E}\int_0^T \bigl\|u_1(s)-u_2(s)\bigr\|^2\,ds \eqsim \mathbb{E}\int_0^T e^{-2\alpha s}\bigl\|u_1(s)-u_2(s)\bigr\|^2\,ds \lesssim \mathbb{E}\bigl\|u_{01}-u_{02}\bigr\|^2.
$$

and

$$
\begin{aligned}
\mathbb{E}\sup_{t\leq T}\bigl\|u_1(t)-u_2(t)\bigr\|^2 &\eqsim \mathbb{E}\sup_{t\leq T}\bigl(e^{-2\alpha t}\bigl\|u_1(t)-u_2(t)\bigr\|^2\bigr)\\
&\lesssim \mathbb{E}\int_0^T e^{-2\alpha s}\bigl\|u_1(s)-u_2(s)\bigr\|^2\,ds,
\end{aligned}
$$

thus completing the proof.

7.7 A remark on uniform integrability

The classical characterization of uniform integrability by de la Vallée Poussin states that, in the setting of a measure space $(X, \mathcal{A})$ endowed with a finite measure μ, a bounded subset $\mathscr{G}$ of $L^1(X, \mu; \mathbb{R}^n)$ is uniformly integrable if and only if there exists a continuous increasing convex function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$, with $\varphi(0) = 0$ and $\lim_{x\to\infty} \varphi(x)/x = \infty$, such that

$$
\int_A \varphi(|g|)\,d\mu < 1 \qquad \forall g \in \mathscr{G}
$$

(see, e.g., [4, p. 12]).

The following criterion for uniform integrability can be proved in the same way (the proof is included for completeness).

Lemma 7.7.1. *Let $F : \mathbb{R}^d \to \mathbb{R}_+$ be a continuous convex function such that $F(0) = 0$ and*

$$
\lim_{|x|\to\infty} \frac{F(x)}{|x|} = \infty.
$$

Let $\mathscr{G}$ be a subset of $L^0(X, \mu; \mathbb{R}^d)$ such that $F(\mathscr{G})$ is bounded in $L^1(X, \mu)$. Then $\mathscr{G}$ is uniformly integrable.

Proof. We have to prove that $\mathscr{G}$ is bounded in $L^1(X, \mu)$ and that for any $\varepsilon > 0$ there exists $\delta > 0$ such that, for any $A \in \mathcal{A}$ with $\mu(A) < \delta$,

$$\int_A |g| \, d\mu < \varepsilon \qquad \forall g \in \mathscr{G}.$$

By definition of limit, for any $M > 0$ there exists R (depending on M) such that $|x| < F(x)/M$ for all $x \in \mathbb{R}^d$ such that $|x| > R$. Then

$$\begin{aligned} \int_A |g| \, d\mu &= \int_{A \cap \{|g| \le R\}} |g| \, d\mu + \int_{A \cap \{|g| > R\}} |g| \, d\mu \\ &\le R\mu(A) + \frac{1}{M} \int_X F(g) \, d\mu \end{aligned}$$

for all $g \in \mathscr{G}$. Choosing $A = X$, this proves that $\mathscr{G}$ is bounded in $L^1(X, \mu)$. Let $\varepsilon > 0$ be arbitrary, and choose M such that the second-term on the right-hand side of the last inequality is smaller than $\varepsilon/2$. Then $\delta := \varepsilon/(2R)$ satisfies the required condition. $\square$

Chapter 8

Singular equations in divergence form: an alternative approach

In this chapter, we prove existence and uniqueness of strong solutions for a class of second-order stochastic PDEs with multiplicative Wiener noise and drift of the form $\operatorname{div}\gamma(\nabla\cdot)$, where γ is a maximal monotone graph in $\mathbb{R}^d \times \mathbb{R}^d$ obtained as the subdifferential of a convex function satisfying very mild assumptions on its behavior at infinity. The well-posedness result complements the corresponding one in Chapter 7 where, under the additional assumption that γ is single-valued, a solution with better integrability and regularity properties is constructed. The proof given here, however, is self-contained.

The results presented in this chapter are part of the joint work [64] with Carlo Marinelli, to appear in *Springer Proceedings in Mathematics & Statistics*.

8.1 The problem: literature and main results

Let us consider the stochastic partial differential equation

$$du(t) - \operatorname{div}\gamma(\nabla u(t))\,dt \ni B(t,u(t))\,dW(t), \qquad u(0) = u_0, \tag{8.1.1}$$

posed on $L^2(D)$, with D a bounded domain of $\mathbb{R}^d$ with smooth boundary. The following assumptions will be in force: (a) γ is the subdifferential of a lower semicontinuous convex function $k : \mathbb{R}^d \to \mathbb{R}_+$ with $k(0) = 0$ and such that

$$\lim_{|x|\to\infty} \frac{k(x)}{|x|} = +\infty, \qquad \limsup_{|x|\to\infty} \frac{k(-x)}{k(x)} < +\infty$$

(in particular, γ is a maximal monotone graph in $\mathbb{R}^d \times \mathbb{R}^d$ whose domain coincides with $\mathbb{R}^d$); (b) W is a cylindrical Wiener process on a separable Hilbert space H, supported by a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,T]}, \mathbb{P})$ satisfying the "usual conditions"; (c) B is a map from $\Omega \times [0,T] \times L^2(D)$ to $\mathscr{L}^2(H, L^2(D))$, the space of Hilbert-Schmidt operators from H to $L^2(D)$, that is Lipschitz-continuous and has linear growth with respect to its third argument, uniformly with respect to the other two, and is such that $B(\cdot,\cdot,a)$ is measurable and adapted for all $a \in L^2(D)$.

Under the additional assumption that γ is a (single-valued) continuous function, we proved in Chapter 7 that (8.1.1) admits a strong solution u, which is unique within a set of processes satisfying mild integrability conditions. The solution of Chapter 7 is constructed pathwise, i.e. for each $\omega \in \Omega$, so that, as is natural to expect, measurability problems arise with respect to the usual σ-algebras on $\Omega \times [0,T]$ used in the theory of stochastic processes. Precisely because of such an issue we needed to assume γ to be single-valued.

The purpose of this chapter is to provide an alternative approach to establish the well-posedness of (8.1.1) that, avoiding pathwise constructions, is simpler than that of Chapter 7 and does not need any extra assumption on γ. The price to pay is that the solution we obtain here is less regular than that of Chapter 7. We also refer to Chapter 6 for a related result obtained by analogous methods.

Let us define the concept of solution to (8.1.1) we shall be working with.

Definition 8.1.1. *Let u_0 be an $L^2(D)$-valued $\mathscr{F}_0$-measurable random variable. A* strong solution *to equation* (8.1.1) *is a couple (u,η) satisfying the following properties:*

(i) u is a measurable and adapted $L^2(D)$-valued process such that

$$u \in L^1(0,T;W_0^{1,1}(D)), B(\cdot,u) \in L^2(0,T;\mathscr{L}^2(U,L^2(D))) \quad \mathbb{P}\text{-}a.s.;$$

(ii) η is a measurable and adapted $L^1(D)^d$-valued process such that

$$\eta \in L^1(0,T;L^1(D)^d) \quad \mathbb{P}\text{-}a.s., \qquad \eta \in \gamma(\nabla u) \quad a.e.\ in\ \Omega \times (0,T) \times D;$$

(iii) one has, as an equality in $L^2(D)$,

$$u(t) - \int_0^t \operatorname{div}\eta(s)\,ds = u_0 + \int_0^t B(s,u(s))\,dW(s) \quad \mathbb{P}\text{-}a.s. \quad \forall\, t \in [0,T]. \tag{8.1.2}$$

Note that (8.1.2) has to be intended in the sense of distributions. In particular, since $\eta \in L^1(D)^d$, the integrand in the second term of (8.1.2) does not, in general, take values in $L^2(D)$. However, the conditions on B imply that the stochastic integral in (8.1.2) is an $L^2(D)$-valued local martingale, hence the term involving the divergence of η turns out to be $L^2(D)$-valued by comparison.

We can now state our main result. Here and in the following $k^* : \mathbb{R}^d \to \mathbb{R}_+$ is the convex conjugate of k.

Theorem 8.1.2. *Let $u_0 \in L^2(\Omega,\mathscr{F}_0;L^2(D))$. Then equation* (8.1.1) *admits a unique strong solution (u,η) such that*

$$\begin{aligned}
&\sup_{t\le T} \mathbb{E}\|u(t)\|^2_{L^2(D)} + \mathbb{E}\int_0^T \|u(t)\|_{W_0^{1,1}(D)}\,dt < \infty,\\
&\mathbb{E}\int_0^T \|\eta(t)\|_{L^1(D)^d}\,dt < \infty,\\
&\mathbb{E}\int_0^T \big(\|k(\nabla u(t))\|_{L^1(D)} + \|k^*(\eta(t))\|_{L^1(D)}\big)\,dt < \infty.
\end{aligned}$$

Moreover, the solution map $u_0 \mapsto u$ is Lipschitz-continuous from the space $L^2(\Omega;L^2(D))$ to $L^\infty(0,T;L^2(\Omega;L^2(D)))$, and u is weakly continuous as a function on $[0,T]$ with values in $L^2(\Omega;L^2(D))$.

Under the extra assumption of γ being single-valued, the solution obtained in Chapter 7 is more regular in the sense that $\mathbb{E}\sup_{t\le T}\|u(t)\|^2_{L^2(D)}$ is finite, the solution map is Lipschitz-continuous from $L^2(\Omega;L^2(D))$ to $L^2(\Omega;L^\infty(0,T;L^2(D)))$, and $u(\omega,\cdot)$ is weakly continuous as a function on $[0,T]$ with values in $L^2(D)$ for $\mathbb{P}$-a.a. $\omega\in\Omega$.

8.2 Well-posedness of an auxiliary equation

The goal of this section is to prove well-posedness of a version of (8.1.1) with additive noise. Namely, we consider the initial value problem

$$du(t)-\operatorname{div}\gamma(\nabla u(t))\,dt\ni G(t)\,dW(t),\qquad u(0)=u_0, \tag{8.2.3}$$

where $G\in L^2(\Omega\times[0,T];\mathscr{L}^2(H,L^2(D)))$ is a measurable and adapted process.

Proposition 8.2.1. *Equation* (8.2.3) *admits a unique strong solution* (u,η) *satisfying the same integrability and weak continuity conditions of Theorem 8.1.2.*

We introduce the regularized equation

$$du_\lambda(t)-\operatorname{div}\gamma_\lambda(\nabla u_\lambda(t))\,dt-\lambda\Delta u_\lambda(t)\,dt=G(t)\,dW(t),\qquad u_\lambda(0)=u_0,$$

where $\gamma_\lambda:\mathbb{R}^d\to\mathbb{R}^d$, $\gamma_\lambda:=\frac{1}{\lambda}(I-(I+\lambda\gamma)^{-1})$, for any $\lambda>0$, is the Yosida approximation of γ, and $\Delta:H^1_0(D)\to H^{-1}(D)$ is the (variational) Dirichlet Laplacian. Since γ_λ is monotone and Lipschitz-continuous, the classical variational approach (see [46, 72] as well as [56]) yields the existence of a unique predictable process u_λ with values in $H^1_0(D)$ such that

$$\mathbb{E}\|u_\lambda\|^2_{C([0,T];L^2(D))}+\mathbb{E}\int_0^T\|u_\lambda(t)\|^2_{H^1_0(D)}\,dt<\infty$$

and

$$u_\lambda(t)-\int_0^t\operatorname{div}\gamma_\lambda(\nabla u_\lambda(s))\,ds-\lambda\int_0^t\Delta u_\lambda(s)\,ds=u_0+\int_0^t G(s)\,dW(s) \tag{8.2.4}$$

$\mathbb{P}$-a.s. in $H^{-1}(D)$ for all $t\in[0,T]$.

We are now going to prove a priori estimates and weak compactness in suitable topologies for u_λ and related processes. These will allow us to pass to the limit as $\lambda\to 0$ in (8.2.4).

For notational parsimony, we shall often write, for any $p\ge 0$, L^p_ω, L^p_t, and L^p_x in place of $L^p(\Omega)$, $L^p(0,T)$, and $L^p(D)$, respectively, and C_t to denote $C([0,T])$. Other similar abbreviations are self-explanatory. The $L^2(D)$-norm will be denoted simply by $\|\cdot\|$. If a function $f:D\to\mathbb{R}^d$ is such that each component f^j, $j=1,\dots,d$, belongs to $L^p(D)$, we shall just write $f\in L^p(D)$ rather than $f\in L^p(D)^d$.

Lemma 8.2.2. *There exists a constant N such that*

$$\begin{aligned}\|u_\lambda\|_{L^2_\omega C_t L^2_x}+\lambda^{1/2}\|\nabla u_\lambda\|_{L^2_{t,\omega,x}}+\|\gamma_\lambda(\nabla u_\lambda)\cdot\nabla u_\lambda\|_{L^1_{t,\omega,x},}\\ <N\big(\|u_0\|_{L^2_{\omega,x}}+\|G\|_{L^2_{t,\omega}\mathscr{L}^2(H,L^2_x)}\big).\end{aligned}$$

Proof. Itô's formula for the square of the norm in L^2_x yields

$$\begin{aligned}\|u_\lambda(t)\|^2 + 2\int_0^t\int_D \gamma(\nabla u_\lambda(s))\cdot\nabla u_\lambda(s)\,dx\,ds + 2\lambda\int_0^t \|\nabla u_\lambda(s)\|^2\,ds \\ = \|u_0\|^2 + 2\int_0^t u_\lambda(s)G(s)\,dW(s) + \int_0^t \|G(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds,\end{aligned}$$

hence, taking the supremum in time and expectation,

$$\begin{aligned}\mathbb{E}\|u_\lambda\|^2_{C_tL^2_x} + \mathbb{E}\int_0^T\int_D \gamma_\lambda(\nabla u_\lambda(s))\cdot\nabla u_\lambda(s)\,dx\,ds + \lambda\,\mathbb{E}\|\nabla u_\lambda\|^2_{L^2_{t,x}} \\ \lesssim \mathbb{E}\|u_0\|^2 + \mathbb{E}\|G\|^2_{L^2_t\mathscr{L}^2(H,L^2_x)} + \mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t u_\lambda(s)G(s)\,dW(s)\right|,\end{aligned}$$

where, by Davis' inequality (see, e.g., [61]), the ideal property of Hilbert-Schmidt operators (see, e.g., [20, p. V.52]), and the elementary inequality $ab \leq \varepsilon a^2 + b^2/\varepsilon$ $\forall a, b \geq 0$, $\varepsilon > 0$,

$$\begin{aligned}\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t u_\lambda(s)G(s)\,dW(s)\right| &\lesssim \mathbb{E}\biggl(\int_0^T \|u_\lambda(s)G(s)\|^2_{\mathscr{L}^2(H,\mathbb{R})}\,ds\biggr)^{1/2} \\ &\leq \varepsilon\,\mathbb{E}\|u_\lambda\|^2_{C_tL^2_x} + N(\varepsilon)\,\mathbb{E}\int_0^T \|G(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds\end{aligned}$$

for any $\varepsilon > 0$. To conclude it suffices to choose ε small enough. □

Lemma 8.2.3. *The families* (∇u_λ) *and* $(\gamma_\lambda(\nabla u_\lambda))$ *are relatively weakly compact in* $L^1(\Omega \times (0,T) \times D)$.

Proof. Recall that, for any $y, r \in \mathbb{R}^d$, ones has $k(y)+k^*(r) = r\cdot y$ if and only if $r \in \partial k(y) = \gamma(y)$. Therefore, since

$$\gamma_\lambda(x) \in \partial k\bigl((I+\lambda\gamma)^{-1}x\bigr) = \gamma\bigl((I+\lambda\gamma)^{-1}x\bigr) \qquad \forall x \in \mathbb{R}^d,$$

we deduce by the definition of γ_λ that

$$\begin{aligned}k\bigl((I+\lambda\gamma)^{-1}x\bigr) + k^*\bigl(\gamma_\lambda(x)\bigr) &= \gamma_\lambda(x)\cdot(I+\lambda\gamma)^{-1}x \\ &= \gamma_\lambda(x)\cdot x - \lambda\,|\gamma_\lambda(x)|^2 \leq \gamma_\lambda(x)\cdot x \qquad \forall x \in \mathbb{R}^d.\end{aligned} \tag{8.2.5}$$

(See, e.g., [42] for all necessary facts from convex analysis used in this note.) Hence, taking Lemma 8.2.2 into account, there exists a constant $N > 0$, independent of λ, such that

$$\mathbb{E}\int_0^T\int_D k^*\bigl(\gamma_\lambda(\nabla u_\lambda)\bigr) \leq \mathbb{E}\int_0^T\int_D \gamma_\lambda(\nabla u_\lambda)\cdot\nabla u_\lambda < N.$$

The assumptions on k imply that its convex conjugate k^* is also convex, lower semicontinuous and such that $\lim_{|y|\to\infty} k^*(y)/|y| = +\infty$. Therefore a simple modification of the criterion by de la Vallée Poussin implies that $(\gamma_\lambda(\nabla u_\lambda))$ is uniformly integrable on $\Omega\times(0,T)\times D$, hence that it is relatively weakly compact in $L^1_{t,\omega,x}$ by the Dunford-Pettis theorem. A completely analogous argument shows that

$$\mathbb{E}\int_0^T\int_D k\bigl((I+\lambda\gamma)^{-1}\nabla u_\lambda\bigr) \leq \mathbb{E}\int_0^T\int_D \gamma_\lambda(\nabla u_\lambda)\cdot\nabla u_\lambda < N,$$

hence that $(I+\lambda\gamma)^{-1}\nabla u_\lambda$ is relatively weakly compact in $L^1_{t,\omega,x}$. Moreover, since $\nabla u_\lambda =$

$(I+\lambda\gamma)^{-1}\nabla u_\lambda+\lambda\gamma_\lambda(\nabla u_\lambda)$, it also follows that (∇u_λ) is relatively weakly compact in $L^1_{t,\omega,x}$. $\square$

Thanks to Lemmata 8.2.2 and 8.2.3, there exists a subsequence of λ, denoted by the same symbol, and processes $u \in L^\infty_t L^2_{\omega,x} \cap L^1_{t,\omega} W^{1,1}_0$ and $\eta \in L^1_{t,\omega,x}$ such that

$$\begin{aligned}
u_\lambda &\longrightarrow u && \text{weakly* in } L^\infty_t L^2_{\omega,x},\\
u_\lambda &\longrightarrow u && \text{weakly in } L^1_{t,\omega} W^{1,1}_0,\\
\gamma_\lambda(\nabla u_\lambda) &\longrightarrow \eta && \text{weakly in } L^1_{t,\omega,x},\\
\lambda u_\lambda &\longrightarrow 0 && \text{weakly in } L^2_{t,\omega} H^1_0.
\end{aligned}$$

as $\lambda \to 0$. Let $t \in [0,T]$ be arbitrary but fixed. The fourth convergence above implies

$$\lambda \int_0^t \Delta u_\lambda(s)\,ds \longrightarrow 0 \qquad \text{in } L^2_\omega H^{-1},$$

while the third yields, for any $\varphi \in L^\infty_\omega W^{1,\infty}$,

$$\mathbb{E}\int_0^t\!\!\int_D \gamma_\lambda(\nabla u_\lambda(s))\cdot\nabla\varphi\,dx\,ds \longrightarrow \mathbb{E}\int_0^t\!\!\int_D \eta(s)\cdot\nabla\varphi\,dx\,ds,$$

hence $\mathbb{E}\displaystyle\int_0^t \langle \operatorname{div}\gamma_\lambda(\nabla u_\lambda(s)),\varphi\rangle\,ds \longrightarrow \mathbb{E}\int_0^t \langle \operatorname{div}\eta(s),\varphi\rangle\,ds$. Therefore, recalling (8.2.4), by difference we deduce that

$$\mathbb{E}\langle u_\lambda(t),\varphi\rangle \longrightarrow \mathbb{E}\langle u(t),\varphi\rangle.$$

Consequently, since $u_\lambda(t)$ is bounded in $L^2_\omega L^2_x$, we also have that $u_\lambda(t) \to u(t)$ weakly in $L^2_\omega L^2_x$. Taking the limit as $\lambda \to 0$ in (8.2.4) thus yields

$$u(t) - \int_0^t \operatorname{div}\eta(s)\,ds = u_0 + \int_0^t G(s)\,dW(s) \qquad \text{in } L^1_\omega V_0',$$

where V_0' is the (topological) dual of a separable Hilbert space V_0 embedded continuously and densely in H^1_0, and continuously in $W^{1,\infty}$. The identity immediately implies that $u \in C_t L^1_\omega V_0'$. Since $u \in L^\infty_t L^2_\omega L^2_x$, it follows by Lemma 1.4.1 that u is a weakly continuous function on $[0,T]$ with values in $L^2_\omega L^2_x$.

By Mazur's lemma there exist sequences of convex combinations of $\gamma_\lambda(\nabla u_\lambda)$ that converge η in (the norm topology of) $L^1_{t,\omega,x}$, thus also, passing to a subsequence, $\mathbb{P}\otimes dt$-almost everywhere in L^1_x. Similarly, since $u_\lambda \to u$ weakly* in $L^\infty_t L^2_{\omega,x}$ implies that $u_\lambda \to u$ weakly in $L^2_{t,\omega,x}$, there exist sequences of convex combinations of u_λ that converge to u $\mathbb{P}\otimes dt$-almost everywhere in L^2_x. Since convex combinations of (u_λ) and of $(\gamma_\lambda(\nabla u_\lambda))$ are (at least) predictable and adapted, respectively, it follows that u is predictable and η is measurable and adapted. Moreover, thanks to the weak lower semicontinuity of convex integrals, one has

$$\mathbb{E}\int_0^T\!\!\int_D \big(k(\nabla u) + k^*(\eta)\big) < \infty.$$

In order to show that $\eta \in \gamma(\nabla u)$ for a.a. (ω,t,x), we need the following "energy identity".

Lemma 8.2.4. *Assume that*

$$y(t) + \alpha\int_0^t y(s)\,ds - \int_0^t \operatorname{div}\zeta(s)\,ds = y_0 + \int_0^t C(s)\,dW(s)$$

in L^2_x $\mathbb{P}$-a.s. for all $t \in [0,T]$, where $\alpha \in \mathbb{R}$, $y_0 \in L^2_{\omega,x}$ is $\mathscr{F}_0$-measurable, and

$$y \in L^\infty_t L^2_{\omega,x} \cap L^1_{t,\omega} W^{1,1}_0, \qquad \zeta \in L^1_{t,\omega,x}, \qquad C \in L^2_{t,\omega}\mathscr{L}^2(H, L^2_x)$$

are measurable and adapted processes such that $k(c\nabla y) + k^*(c\zeta) \in L^1_{t,\omega,x}$ for a constant $c > 0$. Then

$$\begin{aligned}\mathbb{E}\|y(t)\|^2 + 2\alpha\,\mathbb{E}\int_0^t \|y(s)\|^2\,ds + 2\,\mathbb{E}\int_0^t\int_D \zeta\cdot\nabla y\,dx\,ds \\ = \mathbb{E}\|y_0\|^2 + \mathbb{E}\int_0^t \|C(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds \qquad \forall t \in [0,T].\end{aligned}$$

Proof. Let $m \in \mathbb{N}$ be such that such that $(I - \delta\Delta)^{-m}$ maps L^1_x into $H^1_0 \cap W^{1,\infty}$, and use the notation $h^\delta := (I - \delta\Delta)^{-m} h$ for any h taking values in L^1_x. One has

$$y^\delta(t) + \alpha\int_0^t y^\delta(s)\,ds - \int_0^t \operatorname{div}\zeta^\delta(s)\,ds = y_0^\delta + \int_0^t C^\delta(s)\,dW(s) \tag{8.2.6}$$

$\mathbb{P}$-a.s. for all $t \in [0,T]$, as an equality in L^2_x, for which Itô's formula yields

$$\begin{aligned}\|y^\delta(t)\|^2 + 2\alpha\int_0^t \|y^\delta(s)\|^2\,ds + 2\int_0^t\int_D \zeta^\delta\cdot\nabla y^\delta\,dx\,ds \\ = \|y_0^\delta\|^2 + \int_0^t \|C^\delta(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds + \int_0^t y^\delta(s)C^\delta(s)dW(s).\end{aligned}$$

It is evident from (8.2.6) that y^δ is a continuous L^2_x-valued process, hence the stochastic integral $(y^\delta C^\delta)\cdot W$ on the right-hand side of the above identity is a continuous local martingale. Let (T_n) be a localizing sequence, and multiply the previous identity by $1_{[0,T_n]}$, to obtain, thanks to $\mathbb{E}(y^\delta C^\delta)\cdot W(\cdot\wedge T_n) = 0$,

$$\begin{aligned}\mathbb{E}\|y^\delta(t\wedge T_n)\|^2 + 2\alpha\,\mathbb{E}\int_0^{t\wedge T_n} \|y^\delta(s)\|^2\,ds + 2\,\mathbb{E}\int_0^{t\wedge T_n}\int_D \zeta^\delta\cdot\nabla y^\delta\,dx\,ds \\ = \mathbb{E}\|y_0^\delta\|^2 + \mathbb{E}\int_0^{t\wedge T_n} \|C^\delta(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds.\end{aligned}$$

Letting n tend to ∞, the dominated convergence theorem yields

$$\begin{aligned}\mathbb{E}\|y^\delta(t)\|^2 + 2\alpha\,\mathbb{E}\int_0^t \|y^\delta(s)\|^2\,ds + 2\,\mathbb{E}\int_0^t\int_D \zeta^\delta\cdot\nabla y^\delta\,dx\,ds \\ = \mathbb{E}\|y_0^\delta\|^2 + \mathbb{E}\int_0^t \|C^\delta(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds\end{aligned}$$

for all $t \in [0,T]$. We are now going to pass to the limit as $\delta \to 0$: the first and second terms on the left-hand side and the first on the right-hand side clearly converge to $\mathbb{E}\|y(t)\|^2$, $2\alpha\,\mathbb{E}\int_0^t \|y(s)\|^2\,ds$ and $\mathbb{E}\|y_0\|^2$, respectively. Properties of Hilbert-Schmidt operators and the dominated convergence theorem also yield

$$\lim_{\delta\to 0}\mathbb{E}\int_0^t \|C^\delta(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds = \mathbb{E}\int_0^t \|C(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds$$

for all $t \in [0,T]$. To conclude it then suffices to show that $\nabla y^\delta\cdot\zeta^\delta \to \nabla y\cdot\zeta$ in $L^1_{t,\omega,x}$. Since $\nabla y^\delta \to \nabla y$ and $\zeta^\delta \to \zeta$ in measure in $\Omega\times(0,t)\times D$, Vitali's theorem implies strong convergence in $L^1_{t,\omega,x}$ if the sequence $(\nabla y^\delta\cdot\zeta^\delta)$ is uniformly integrable in $\Omega\times(0,t)\times D$. In turn, the latter is certainly true if $(|\nabla y^\delta \cdot \zeta^\delta|)$ is dominated by a sequence that converges strongly in $L^1_{t,\omega,x}$. Indeed, using the assumptions on the behavior of k at infinity as well as the generalized Jensen inequality for sub-Markovian operators (see [41]), one has

$$\pm c^2 \zeta^\delta \cdot \nabla y^\delta \lesssim 1 + k(c\nabla y^\delta) + k^*(c\zeta^\delta) \leq 1 + (I - \delta\Delta)^{-m} \left(k(c\nabla y) + k^*(c\zeta)\right),$$

where the sequence on the right-hand side converges in $L^1_{t,\omega,x}$ as $\delta \to 0$ because, by assumption, $k(c\nabla y) + k^*(c\zeta) \in L^1_{t,\omega,x}$. □

Itô's formula yields

$$\begin{aligned}
\mathbb{E}\|u_\lambda(t)\|^2 &+ 2\,\mathbb{E}\int_0^t\!\!\int_D \gamma_\lambda(\nabla u_\lambda)\cdot\nabla u_\lambda + 2\lambda\,\mathbb{E}\int_0^t \|\nabla u_\lambda\|^2 \\
&= \mathbb{E}\|u_0\|^2 + \mathbb{E}\int_0^t \|G(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds
\end{aligned}$$

and, by Lemma 8.2.4,

$$\mathbb{E}\|u(t)\|^2 + 2\,\mathbb{E}\int_0^t\!\!\int_D \eta\cdot\nabla u = \mathbb{E}\|u_0\|^2 + \mathbb{E}\int_0^t \|G(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds.$$

One then has

$$\begin{aligned}
2\limsup_{\lambda\to 0}\mathbb{E}&\int_0^T\!\!\int_D \gamma_\lambda(\nabla u_\lambda(s))\cdot\nabla u_\lambda(s)\,dx\,ds \\
&\leq \mathbb{E}\|u_0\|^2 + \mathbb{E}\int_0^T \|G(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds - \liminf_{\lambda\to 0}\mathbb{E}\|u_\lambda(T)\|^2 \\
&\leq \mathbb{E}\|u_0\|^2 + \mathbb{E}\int_0^T \|G(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds - \mathbb{E}\|u(T)\|^2 \\
&= \mathbb{E}\int_0^T\!\!\int_D \eta(s)\cdot\nabla u(s)\,dx\,ds.
\end{aligned}$$

Since $\nabla u_\lambda \to \nabla u$ and $\gamma_\lambda(\nabla u_\lambda) \to \eta$ weakly in $L^1_{t,\omega,x}$, this implies that $\eta \in \gamma(\nabla u)$ a.e. in $\Omega \times (0,T) \times D$. We have thus proved the existence and weak continuity statements of Proposition 8.2.1.

In order to show that the solution is unique, we are going to prove that *any* solution depends continuously on (u_0, G). Let (u_i, η_i), $i = 1, 2$, satisfy

$$u_i(t) - \int_0^t \operatorname{div}\eta_i(s)\,ds = u_0 + \int_0^t G_i(s)\,ds$$

in the sense of Definition 8.1.1, as well as the integrability conditions (on u and η) of Theorem 8.1.2. Setting $y := u_1 - u_2$, $y_0 := u_{01} - u_{02}$, $\zeta := \eta_1 - \eta_2$, and $F := G_1 - G_2$, one has

$$y(t) - \int_0^t \operatorname{div}\zeta(s)\,ds = y_0 + \int_0^t F(s)\,dW(s)$$

$\mathbb{P}$-a.s. in $L^2(D)$ for all $t \in [0,T]$. For any process h, let us use the notation $h^\alpha(t) := e^{-\alpha t}h(t)$. For any $\alpha > 0$, the integration-by-parts formula yields

$$y^\alpha(t) + \int_0^t \left(-\operatorname{div}\zeta^\alpha(s) + \alpha y^\alpha(s)\right)ds = y_0 + \int_0^t F^\alpha(s)\,dW(s),$$

hence also, thanks to Lemma 8.2.4,

$$\begin{aligned}\mathbb{E}\|y^\alpha(t)\|^2 + 2\alpha\,\mathbb{E}\int_0^t \|y^\alpha(s)\|^2\,ds + 2\,\mathbb{E}\int_0^t\!\!\int_D \zeta^\alpha(s)\cdot\nabla y^\alpha(s)\,dx\,ds\\ \leq \mathbb{E}\|y_0\|^2 + \mathbb{E}\int_0^t \|F^\alpha(s)\|^2_{\mathscr{L}^2(H,L^2_x)}\,ds,\end{aligned}$$

where $\zeta^\alpha\cdot\nabla y^\alpha \geq 0$ by monotonicity. Therefore, taking the L^∞_t norm,

$$\|y^\alpha\|_{L^\infty_t L^2_{\omega,x}} + \sqrt{\alpha}\|y^\alpha\|_{L^2_{t,\omega,x}} \lesssim \|y_0\|_{L^2_{\omega,x}} + \|F^\alpha\|_{L^2_{t,\omega}\mathscr{L}^2(H,L^2_x)},$$

that is, using the notation $L^p_t(\alpha) := L^p([0,T], e^{-\alpha t}dt)$ for any $p \geq 0$,

$$\begin{aligned}\|u_1 - u_2\|_{L^\infty_t(\alpha)L^2_{\omega,x}} + \sqrt{\alpha}\|u_1 - u_2\|_{L^2_t(\alpha)L^2_{\omega,x}}\\ \lesssim \|u_{01} - u_{02}\|_{L^2_{\omega,x}} + \|G_1 - G_2\|_{L^2_t(\alpha)L^2_\omega\mathscr{L}^2(H,L^2_x)}.\end{aligned}\tag{8.2.7}$$

Taking $\alpha = 0$ and $G_1 = G_2$ immediately yields the uniqueness of solutions (as well as Lipschitz-continuous dependence on the initial datum). The proof of Proposition 8.2.1 is thus complete.

8.3 Proof of Theorem 8.1.2

For any $v \in L^2_{t,\omega,x}$ measurable and adapted, and any $\mathscr{F}_0$-measurable random variable $u_0 \in L^2_{\omega,x}$, the process $B(\cdot, v)$ is measurable, adapted, and belongs to $L^2_{t,\omega}\mathscr{L}^2(H, L^2_x)$, hence the equation

$$du(t) - \operatorname{div}\gamma(\nabla u(t))\,dt \ni B(t, v(t))\,dW(t), \qquad u(0) = u_0,$$

is well-posed in the sense of Proposition 8.2.1. Moreover, for any v_1, v_2 and u_{01}, u_{02} satisfying the same hypotheses on v and u_0, respectively, (8.2.7) yields

$$\begin{aligned}\|u_1 - u_2\|_{L^\infty_t(\alpha)L^2_{\omega,x}} + \sqrt{\alpha}\|u_1 - u_2\|_{L^2_t(\alpha)L^2_{\omega,x}}\\ \lesssim \|u_{01} - u_{02}\|_{L^2_{\omega,x}} + \|B(\cdot, v_1) - B(\cdot, v_2)\|_{L^2_t(\alpha)L^2_\omega\mathscr{L}^2(H,L^2_x)}.\end{aligned}$$

It hence follows by the Lipschitz-continuity of B that

$$\|u_1 - u_2\|_{L^2_t(\alpha)L^2_{\omega,x}} \lesssim \frac{1}{\sqrt{\alpha}}\Big(\|u_{01} - u_{02}\|_{L^2_{\omega,x}} + \|v_1 - v_2\|_{L^2_t(\alpha)L^2_{\omega,x}}\Big),\tag{8.3.8}$$

where the implicit constant does not depend on α. In particular, denoting by Γ the map $(u_0, v) \mapsto u$, one has that $v \mapsto \Gamma(u_0, v)$ is a strict contraction of $L^2_t(\alpha)L^2_{\omega,x}$ for α large enough. Therefore, by equivalence of norms, $v \mapsto \Gamma(u_0, v)$ admits a unique fixed point in $L^2_{t,\omega,x}$, which solves (8.1.1) and satisfies all integrability conditions. Such solution is unique as any solution is a fixed point of $v \mapsto \Gamma(u_0, v)$.

Let us show that the solution map $u_0 \mapsto u$ is Lipschitz-continuous: (8.3.8) yields, choosing α large enough,

$$\|u_1 - u_2\|_{L^2_t(\alpha)L^2_{\omega,x}} \leq N_1\|u_{01} - u_{02}\|_{L^2_{\omega,x}} + N_2\|u_1 - u_2\|_{L^2_t(\alpha)L^2_{\omega,x}}$$

with constants $N_1 > 0$ and $0 < N_2 < 1$, hence, by equivalence of norms,

$$\|u_1 - u_2\|_{L^2_t L^2_{\omega,x}} \lesssim \|u_{01} - u_{02}\|_{L^2_{\omega,x}}.$$

This in turn implies, in view of (8.2.7) (with $\alpha = 0$) and the Lipschitz-continuity of B,

$$\begin{aligned}\|u_1 - u_2\|_{L_t^\infty L^2_{\omega,x}} &\lesssim \|u_{01} - u_{02}\|_{L^2_{\omega,x}} + \|B(\cdot,u_1) - B(\cdot,u_2)\|_{L^2_{t,\omega}\mathscr{L}^2(H,L^2_x)} \\ &\lesssim \|u_{01} - u_{02}\|_{L^2_{\omega,x}} + \|u_1 - u_2\|_{L^2_t L^2_{\omega,x}} \lesssim \|u_{01} - u_{02}\|_{L^2_{\omega,x}},\end{aligned}$$

which completes the proof.

Remark. A priori estimates entirely analogous to those of Lemma 8.2.2, as well as weak compactness results exactly as in Lemma 8.2.3, can be proved for the regularized equation obtained by replacing γ with $\gamma_\lambda + \lambda\nabla$ directly in (8.1.1). It is however not immediately clear how to pass to the limit as $\lambda \to 0$ in the stochastic integrals appearing in such regularized equations with multiplicative noise, i.e. to show that $B(u_\lambda) \cdot W$ converges to $B(u) \cdot W$ in a suitable sense.

Chapter 9

Doubly singular equations in divergence form

In this chapter, we prove well-posedness for a class of second-order SPDEs with multiplicative Wiener noise and doubly nonlinear drift of the form $-\operatorname{div}\gamma(\nabla\cdot)+\beta(\cdot)$, where γ is the subdifferential of a convex function on $\mathbb{R}^d$ and β is a maximal monotone graph everywhere defined on $\mathbb{R}$, on which neither growth nor continuity assumptions are imposed. These results provide an effective generalization and give a unifying treatment to the theory presented in Chapters 6–7–8.

The results presented in this chapter are part of the joint work [66] with Carlo Marinelli, to appear in *Atti Accademia Nazionale Lincei. Rendiconti Lincei. Matematica e Applicazioni.*

9.1 The problem: literature and main goals

Let D be a bounded domain of $\mathbb{R}^d$ with smooth boundary and $T>0$ a fixed number. We shall establish well-posedness in the strong sense for stochastic partial differential equations of the type

$$\begin{cases} du(t)-\operatorname{div}\gamma(\nabla u(t))\,dt+\beta(u(t))\,dt\ni B(t,u(t))\,dW(t) & \text{in } (0,T)\times D,\\ u=0 & \text{in } (0,T)\times\partial D,\\ u(0)=u_0 & \text{in } D,\end{cases}\tag{9.1.1}$$

where $\gamma\subset\mathbb{R}^d\times\mathbb{R}^d$ and $\beta\subset\mathbb{R}\times\mathbb{R}$ are everywhere-defined maximal monotone graphs, the first one of which is assumed to be the subdifferential of a convex function $k:\mathbb{R}^d\to\mathbb{R}$. Furthermore, 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 §9.2 below.

Equations with drift in divergence type, both in deterministic and stochastic settings, have a long history and are thoroughly studied, especially because of their physical significance. From a mathematical point of view, they are particularly interesting because they are fully nonlinear, in the sense that they do not contain any “leading” linear term. For stochastic equations, the first well-posedness result is most likely due to Pardoux, as an application of his general results in [72] on monotone stochastic evolution equations in the variational setting (see also [46] for improved results under more general assumptions on B). In this case one needs to assume

$\beta = 0$ and

$$\gamma(x) \cdot x \gtrsim |x|^p - 1, \qquad |\gamma(x)| \lesssim |x|^{p-1} - 1 \qquad \forall x \in \mathbb{R}^d,$$

with $p > 1$ (the centered dot stands for the usual Euclidean scalar product in $\mathbb{R}^d$). These are precisely the classical Leray-Lions conditions, well known in the deterministic theory (cf. [49]). In some special cases a simple polynomial-type β can be added: for instance, if γ corresponds to the p-Laplacian, i.e. $\gamma(x) = |x|^{p-2}x$, $p \geq 2$, one may consider $\beta(x) = |x|^{p-2}x$ (cf. [56, p. 83]). However, it is well known that if two nonlinear operators satisfy the conditions needed in the variational setting, their sum in general does not. This phenomenon already gives rise to severe restrictions on the class of semilinear equations with polynomial nonlinearities that can be solved by such methods.

In some recent works we have obtained well-posedness results for (9.1.1) under much more general hypotheses than those mentioned above. In particular, in Chapter 6 we assume that γ still satisfies the classical Leray-Lions assumptions, but we impose no growth restriction on β: a very mild symmetry-like condition on its behavior at infinity is shown to suffice. On the other hand, in Chapter 7 we consider the case $\beta = 0$, with no hypotheses on the growth of γ, but with the additional requirement that γ is single-valued (a symmetry-like assumption on γ is needed in this case as well). Equations with more general, possibly multivalued γ, are treated in Chapter 8, where, however, less regular solutions are obtained.

Our goal is to unify and extend the above-mentioned well-posedness results for equation (9.1.1), thus treating the case where both γ and β can be multivalued, without any restriction on their rate of growth. We shall also show that we can do so without loosing any regularity of solutions with respect to the results of Chapter 7. The approach we take consists in a combination of (deterministic and stochastic) variational techniques and weak compactness in L^1 spaces. A key feature is the construction of a candidate solution as pathwise limit, in suitable topologies, of solutions to regularized equations. In particular, due to this type of construction, in order to obtain measurability properties of solutions, uniqueness of limits is crucial. Roughly speaking, we can prove that $-\operatorname{div}\gamma(\nabla u) + \beta(u)$ is unique, hence that it is measurable, but showing that each one of them is unique (hence measurable) seems difficult, if not impossible. This is the reason why γ was assumed to be single-valued in Chapters 6 and 7. In the general setting of this work we thus need different ideas: let u_λ, γ_λ, and β_λ be suitable regularizations of u, γ, and β, respectively, and set $\eta_\lambda := \gamma_\lambda(\nabla u_\lambda)$ and $\xi_\lambda := \beta_\lambda(u_\lambda)$. Comparing weak limits, obtained in different ways, of the image of the pair $(\eta_\lambda, \xi_\lambda)$ under a continuous linear map, we are going to prove that there exist two limiting processes η and ξ, "sections" of $\gamma(\nabla u)$ and $\beta(u)$, respectively, that are indeed predictable and satisfy suitable uniqueness properties. One may say that we restore uniqueness working in a suitable quotient space, although quotient spaces do not appear explicitly.

The well-posedness result obtained here may be interesting also in the deterministic setting, as our results extend to the doubly nonlinear case the sharpest results available for equations with $\beta = 0$ and $B = 0$, whose hypotheses on γ are identical to ours (cf. [10, p. 207-ff])

The chapter is organized as follows: in Section 9.2 we state the assumptions and the main result, which is then proved in Section 9.3.

9.2 Main result

Before stating the main result, we fix notation and introduce the necessary assumptions.

As already mentioned, D stands for a bounded domain in $\mathbb{R}^d$ with smooth boundary. We shall denote the Hilbert space $L^2(D)$ by H, its norm and scalar product by $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$, respectively. We shall denote the Dirichlet Laplacian on $L^1(D)$ (as well as on $L^2(D)$, without notationally distinguish them) by Δ.

Let $(\Omega,\mathscr{F},\mathbb{P})$ be a probability space, endowed with a filtration $(\mathscr{F}_t)_{t\in[0,T]}$ satisfying the so-called usual conditions, on which all random elements will be defined. Equality of stochastic processes is meant to be in the sense of indistinguishability, unless otherwise stated. We assume that the diffusion coefficient

$$B : \Omega \times [0,T] \times H \to \mathscr{L}^2(U,H)$$

is such that $B(\cdot,\cdot,h)$ is progressively measurable for all $h \in H$, and there exists a positive constant N_B such that

$$\begin{aligned}\big\|B(\omega,t,x)\big\|_{\mathscr{L}^2(U,H)} &\le N_B\big(1+\|x\|\big),\\ \big\|B(\omega,t,x)-B(\omega,t,y)\big\|_{\mathscr{L}^2(U,H)} &\le N_B\|x-y\|\end{aligned}$$

for all $(\omega,t)\in\Omega\times[0,T]$ and $x,y\in H$. Moreover, let the initial datum u_0 be $\mathscr{F}_0$-measurable with finite second moment, i.e. $u_0\in L^2(\Omega,\mathscr{F}_0;H)$.

Let $k:\mathbb{R}^d\to\mathbb{R}_+$ be a convex function with $k(0)=0$ such that

$$\limsup_{|x|\to+\infty}\frac{k(x)}{k(-x)}<+\infty, \qquad \lim_{|x|\to+\infty}\frac{k(x)}{|x|}=+\infty$$

(we shall call the second condition superlinearity at infinity). Then its subdifferential $\gamma:=\partial k$ is a maximal monotone graph in $\mathbb{R}^d\times\mathbb{R}^d$. We assume that the domain of γ coincides with $\mathbb{R}^d$, which implies that k^*, the convex conjugate of k, is superlinear at infinity as well. Moreover, let $j:\mathbb{R}\to\mathbb{R}_+$ be a further convex function with $j(0)=0$ such that

$$\limsup_{|x|\to+\infty}\frac{j(x)}{j(-x)}<+\infty,$$

whose subdifferential $\beta:=\partial j$ is an everywhere defined maximal monotone graph in $\mathbb{R}\times\mathbb{R}$, so that j^* is superlinear at infinity. All notions of convex analysis and from the theory of maximal monotone operators used thus far and in the sequel are standard and are treated in detail, for instance, in [10].

We can now give the notion of solution to (9.1.1) that we are going to work with. Throughout the work, V_0 is a separable Hilbert space continuously embedded in both $W^{1,\infty}(D)$ and $H^1_0(D)$: for instance one can take, thanks to Sobolev embedding theorems, $V_0:=H^k_0(D)$ for $k\in\mathbb{N}$ sufficiently large. Moreover, the divergence operator is defined as

$$\begin{aligned}\operatorname{div}: L^1(D)^d &\longrightarrow V_0'\\ f &\longmapsto \big[g\mapsto -\langle f,\nabla g\rangle\big],\end{aligned}$$

which is thus linear and bounded. In fact, for any $f\in L^1(D)^d$ and $g\in V_0$,

$$\big|\langle f,\nabla g\rangle\big| \le \big\|f\big\|_{L^1(D)}\big\|g\big\|_{W^{1,\infty}} \lesssim \big\|f\big\|_{L^1(D)}\big\|g\big\|_{V_0}$$

because V_0 is continuously embedded in $W^{1,\infty}$.

Definition 9.2.1. *A strong solution to* (9.1.1) *is a triplet* (u,η,ξ)*, where* u*,* η*, and* ξ *are* *adapted processes taking values in* $W_0^{1,1}(D) \cap H$, $L^1(D)^d$, *and* $L^1(D)$, *respectively, such that* $\eta \in \gamma(\nabla u)$ *and* $\xi \in \beta(u)$ *a.e. in* $\Omega \times (0,T) \times D$,

$$\begin{aligned} u &\in L^0(\Omega; C([0,T]; H)) \cap L^0(\Omega; L^1(0,T; W_0^{1,1}(D))), \\ \eta &\in L^0(\Omega; L^1((0,T) \times D)^d), \\ \xi &\in L^0(\Omega; L^1((0,T) \times D)), \\ \nabla u \cdot \eta + u\xi &\in L^0(\Omega; L^1((0,T) \times D)), \end{aligned}$$

and

$$\langle u, \phi \rangle + \int_0^{\cdot} \langle \eta(s), \nabla\phi \rangle \, ds + \int_0^{\cdot} \langle \xi(s), \phi \rangle \, ds = \langle u_0, \phi \rangle + \left\langle \int_0^{\cdot} B(s, u(s)) \, dW(s), \phi \right\rangle$$

for all $\phi \in V_0$.

The last identity in the above definition is equivalent to the validity in the dual of V_0 of the equality

$$u - \int_0^{\cdot} \operatorname{div} \eta(s) \, ds + \int_0^{\cdot} \xi(s) \, ds = u_0 + \int_0^{\cdot} B(s, u(s)) \, dW(s).$$

Note that u, u_0 and the stochastic integrals take values in H and the third term on the left-hand side takes values in $L^1(D)$, hence also the second term on the right-hand side belongs to $L^1(D)$, so that the equality holds also in $L^1(D)$. The same reasoning implies that the sum of the second and third terms on the left-hand side take values in H, so that the above equality can also be seen as valid in H.

The main result of the chapter is the following. The proof is given in §9.3 below.

Theorem 9.2.2. *There exists a strong solution* (u, η, ξ) *to equation* (9.1.1). *It is predictable and satisfies the following properties:*

$$\begin{aligned} u &\in L^2(\Omega; C([0,T]; H)) \cap L^1(\Omega; L^1(0,T; W_0^{1,1}(D))), \\ \eta &\in L^1(\Omega \times (0,T) \times D)^d, \\ \xi &\in L^1(\Omega \times (0,T) \times D), \\ \nabla u \cdot \eta &\in L^1(\Omega \times (0,T) \times D), \\ u\xi &\in L^1(\Omega \times (0,T) \times D). \end{aligned}$$

Moreover, the solution map

$$\begin{aligned} L^2(\Omega, \mathscr{F}_0; H) &\longrightarrow L^2(\Omega; C([0,T]; H)) \\ u_0 &\longmapsto u \end{aligned}$$

is Lipschitz-continuous. In particular, if (u_1, η_1, ξ_1) *and* (u_2, η_2, ξ_2) *are any two strong solutions satisfying the properties above, then* $u_1 = u_2$ *and* $-\operatorname{div} \eta_1 + \xi_1 = -\operatorname{div} \eta_2 + \xi_2$ *in* $L^2(\Omega; C([0,T]; H))$ *and* $L^1(\Omega; L^1(0,T; V_0'))$, *respectively.*

9.3 Proof of Theorem 9.2.2

9.3.1 Itô's formula for the square of the H-norm

We establish a version of Itô's formula for the square of the H-norm in a generalized variational setting, which will play an important role in the sequel. The result is interesting in its own right, as it does not follow from the classical ones in [46, 72], and is apparently new for Itô processes containing a drift term in divergence form with minimal integrability properties.

Proposition 9.3.1. *Let Y, f, and g be measurable adapted processes with values in $H \cap W_0^{1,1}(D)$, $L^1(D)^d$, and $L^1(D)$, respectively, such that*

$$\begin{aligned}
&Y \in L^0(\Omega; L^\infty(0,T;H)) \cap L^0(\Omega; L^1(0,T;W_0^{1,1}(D))),\\
&f \in L^0(\Omega; L^1((0,T) \times D)^d),\\
&g \in L^0(\Omega; L^1((0,T) \times D)),
\end{aligned}$$

and there exists constants a, $b > 0$ such that

$$k(a\nabla u) + k^*(af) + j(bu) + j^*(bg) \in L^0(\Omega; L^1((0,T) \times D)).$$

Moreover, let $Y_0 \in L^0(\Omega, \mathscr{F}_0; H)$ and G be an $\mathscr{L}^2(U,H)$-valued progressively measurable process such that $G \in L^0(\Omega; L^2(0,T;\mathscr{L}^2(U,H)))$. If

$$Y - \int_0^\cdot \operatorname{div} f(s)\,ds + \int_0^\cdot g(s)\,ds = Y_0 + \int_0^\cdot G(s)\,dW(s)$$

as an identity in V_0', then

$$\begin{aligned}
&\frac{1}{2}\|Y\|^2 + \int_0^\cdot\!\!\int_D f(s)\cdot\nabla Y(s)\,ds + \int_0^\cdot\!\!\int_D g(s)Y(s)\,ds\\
&\qquad = \frac{1}{2}\|Y_0\|^2 + \frac{1}{2}\int_0^\cdot \|G(s)\|^2_{\mathscr{L}^2(U,H)}\,ds + \int_0^\cdot Y(s)G(s)\,dW(s).
\end{aligned}$$

Proof. The proof is essentially a combination of arguments described in great detail in Chapters 3 and 7, hence we shall limit ourselves to a sketch only. Using a superscript δ to denote the action of $(I-\delta\Delta)^{-m}$, for a sufficiently large $m \in \mathbb{N}$, we have, thanks to Sobolev embedding theorems and classical elliptic regularity results,

$$Y^\delta - \int_0^\cdot \operatorname{div} f^\delta(s)\,ds + \int_0^\cdot g^\delta(s)\,ds = Y_0^\delta + \int_0^\cdot G^\delta(s)\,dW(s)$$

as an identity of H-valued processes. Itô's formula for Hilbert-space valued continuous semimartingales thus yields

$$\begin{aligned}
&\frac{1}{2}\|Y^\delta\|^2 + \int_0^\cdot\!\!\int_D f^\delta(s)\cdot\nabla Y^\delta(s)\,ds + \int_0^\cdot\!\!\int_D g^\delta(s)Y^\delta(s)\,ds\\
&\qquad = \frac{1}{2}\|Y_0^\delta\|^2 + \frac{1}{2}\int_0^\cdot \|G^\delta(s)\|^2_{\mathscr{L}^2(U,H)}\,ds + \int_0^\cdot Y^\delta(s)G^\delta(s)\,dW(s).
\end{aligned} \tag{9.3.2}$$

Thanks to the assumptions on Y, f, g ad G, it easily follows that, $\mathbb{P}$-a.s.,

$$\begin{aligned}
Y_0^\delta &\longrightarrow Y_0 && \text{in } H,\\
Y^\delta(t) &\longrightarrow Y(t) && \text{in } H \quad \forall\, t \in [0,T],\\
f^\delta &\longrightarrow f && \text{in } L^1((0,T)\times D)^d,\\
g^\delta &\longrightarrow g && \text{in } L^1((0,T)\times D),\\
G^\delta &\longrightarrow G && \text{in } L^2(0,T;\mathscr{L}^2(U,H)).
\end{aligned}$$

Similarly, using simple properties of Hilbert-Schmidt operators and the dominated convergence theorem, it is not difficult to verify that the quadratic variation of $(Y^\delta G^\delta - YG)\cdot W$ converges to zero in probability, so that

$$\int_0^\cdot Y^\delta G^\delta \, dW \longrightarrow \int_0^\cdot YG\, dW$$

uniformly (with respect to time) in probability. Furthermore, thanks to the hypotheses on k and j, the families $(\nabla u^\delta \cdot Y^\delta)$ and $(g^\delta Y^\delta)$ are uniformly integrable in $(0,T)\times D$ $\mathbb{P}$-a.s., hence by Vitali's theorem we also have that, $\mathbb{P}$-a.s.,

$$\begin{aligned}
f^\delta \cdot \nabla Y^\delta &\longrightarrow f\cdot \nabla Y && \text{in } L^1((0,T)\times D),\\
g^\delta Y^\delta &\longrightarrow gY && \text{in } L^1((0,T)\times D).
\end{aligned}$$

The proof is completed passing to the limit as $\delta \to 0$ in (9.3.2), in complete analogy to Sections 4.3 and 7.3. □

Corollary 9.3.2. *Under the assumptions of the previous proposition, one has*

$$Y \in L^0(\Omega; C([0,T];H)).$$

Proof. Since $Y \in L^\infty(0,T;H) \cap C([0,T];V_0')$, the trajectories of Y are weakly continuous in H (see, e.g., [79]). Moreover, by Itô's formula one has

$$\begin{aligned}
\frac12 \|Y(t)\|^2 - \frac12 \|Y(r)\|^2 &+ \int_r^t\!\int_D f(s)\cdot\nabla Y(s)\,ds + \int_r^t\!\int_D g(s)Y(s)\,ds\\
&= \frac12 \int_r^t \|G(s)\|^2_{\mathscr{L}^2(U,H)}\,ds + \int_r^t Y(s)G(s)\,dW(s)
\end{aligned}$$

for every r, $t \in [0,T]$. This implies, by an argument analogous to the one used in Chapter 3.3, that the function $t \mapsto \|Y(t)\|$ is continuous on $[0,T]$. By a well-known criterion we thus conclude that Y has strongly continuous trajectories in H. □

9.3.2 Well-posedness in a special case

As a first step we prove existence of solutions to (9.1.1) assuming that the noise is of additive type and that

$$B \in L^2(\Omega; L^2(0,T;\mathscr{L}^2(U,V_0))).$$

For any $\lambda > 0$, let γ_λ and β_λ denote the Yosida approximations of γ and β, respectively, and consider the regularized equation

$$du_\lambda(t) - \lambda\Delta u_\lambda(t)\,dt - \operatorname{div}\gamma_\lambda(\nabla u_\lambda(t))\,dt + \beta_\lambda(u_\lambda(t))\,dt = B(t)\,dW(t), \quad u_\lambda(0) = u_0.$$

Since γ_λ and β_λ are monotone and Lipschitz-continuous, it is not difficult to check that the operator

$$\phi \longmapsto -\lambda\Delta\phi - \operatorname{div}\gamma_\lambda(\nabla\phi) + \beta_\lambda(\phi)$$

is hemicontinuous, monotone, coercive and bounded on $(H^1_0(D), H, H^{-1}(D))$, so that the classical results by Pardoux [72] provide existence and uniqueness of a variational solution

$$u_\lambda \in L^2(\Omega; C([0,T]; H)) \cap L^2(\Omega; L^2(0,T; H^1_0(D))).$$

The a priori estimates on the solution u_λ contained in the next lemma can be obtained essentially as in Chapters 2, 6 and 7.

Lemma 9.3.3. *There exists a constant N independent of λ such that*

$$\begin{aligned} &\|u_\lambda\|^2_{L^2(\Omega;C([0,T];H))} + \lambda\|\nabla u_\lambda\|^2_{L^2(\Omega;L^2(0,T;H))} \\ &\qquad + \|\gamma_\lambda(\nabla u_\lambda)\cdot\nabla u_\lambda\|_{L^1(\Omega\times(0,T)\times D)} + \|\beta_\lambda(u_\lambda)u_\lambda\|_{L^1(\Omega\times(0,T)\times D)} < N \end{aligned}$$

for all $\lambda \in (0,1)$. Furthermore, there exists $\Omega' \in \mathscr{F}$ with $\mathbb{P}(\Omega') = 1$ such that, for every $\omega \in \Omega'$, there exists a constant $M(\omega)$ independent of λ such that

$$\begin{aligned} &\|u_\lambda(\omega)\|^2_{C([0,T];H)} + \lambda\|\nabla u_\lambda(\omega)\|^2_{L^2(0,T;H)} \\ &+ \|\gamma_\lambda(\nabla u_\lambda(\omega))\cdot\nabla u_\lambda(\omega)\|_{L^1((0,T)\times D)} + \|\beta_\lambda(u_\lambda(\omega))u_\lambda(\omega)\|_{L^1((0,T)\times D)} < M(\omega) \end{aligned}$$

for all $\lambda \in (0,1)$.

Proof. It is an immediate consequence of the (proofs of the) Lemmata 7.4.3–7.4.6, for the part involving γ, and Lemmata 2.4.3–2.4.6, for the part involving β. □

Since

$$\begin{aligned} k^*(\gamma_\lambda(\nabla u_\lambda)) \le k^*(\gamma_\lambda(\nabla u_\lambda)) + k((I+\lambda\gamma)^{-1}\nabla u_\lambda) &= \gamma_\lambda(\nabla u_\lambda)\cdot(I+\lambda\gamma)^{-1}\nabla u_\lambda \\ &\le \gamma_\lambda(\nabla u_\lambda)\cdot\nabla u_\lambda \end{aligned}$$

and

$$j^*(\beta_\lambda(u_\lambda)) \le j^*(\beta_\lambda(u_\lambda)) + j((I+\lambda\beta)^{-1}u_\lambda) = \beta_\lambda(u_\lambda)(I+\lambda\beta)^{-1}u_\lambda \le \beta_\lambda(u_\lambda)u_\lambda,$$

we infer that the families $(k^*(\gamma_\lambda(\nabla u_\lambda)))$ and $(j^*(\beta_\lambda(u_\lambda)))$ are uniformly bounded in $L^1(\Omega \times (0,T) \times D)$. Therefore, recalling that k^* and j^* are superlinear, thanks to the de la Vallée-Poussin criterion and the Dunford-Pettis theorem we deduce that the families $(\gamma_\lambda(u_\lambda))$ and $(\beta_\lambda(u_\lambda))$ are relatively weakly compact in $L^1(\Omega\times(0,T)\times D)^d$ and $L^1(\Omega\times(0,T)\times D)$, respectively. Analogously, the families $(\gamma_\lambda(u_\lambda(\omega)))$ and $(\beta_\lambda(u_\lambda(\omega)))$ are relatively weakly compact in $L^1((0,T)\times D)^d$ and $L^1((0,T)\times D)$, respectively, for $\mathbb{P}$-a.e. $\omega\in\Omega$.

Let Ω' be as in the previous lemma and take $\omega \in \Omega'$. Then we have, along a subsequence λ'

of λ depending on ω,

$$\begin{aligned}
u_{\lambda'}(\omega) &\longrightarrow u(\omega) && \text{weakly* in } L^\infty(0,T;H),\\
\nabla u_{\lambda'}(\omega) &\longrightarrow \nabla u(\omega) && \text{weakly in } L^1((0,T)\times D)^d,\\
\lambda' u_{\lambda'}(\omega) &\longrightarrow 0 && \text{in } L^2(0,T;H^1_0(D)),\\
\gamma_{\lambda'}(u_{\lambda'}(\omega)) &\longrightarrow \eta(\omega) && \text{weakly in } L^1((0,T)\times D)^d,\\
\beta_{\lambda'}(u_{\lambda'}(\omega)) &\longrightarrow \xi(\omega) && \text{weakly in } L^1((0,T)\times D),
\end{aligned}$$

hence, by passage to the weak limit in the regularized equation taking test functions in V_0, we have

$$u - \int_0^\cdot \operatorname{div}\eta(s)\,ds + \int_0^\cdot \xi(s)\,ds = u_0 + \int_0^\cdot B(s)\,dW(s). \tag{9.3.3}$$

Moreover, by the lower semicontinuity of convex integrals, it also follows that

$$k(\nabla u(\omega)) + k^*(\eta(\omega)) + j(u(\omega)) + j^*(\xi(\omega)) \in L^1((0,T)\times D).$$

Arguing as in Chapters 2 and 7, one can show that the process u constructed in this way is unique in the space $L^2(\Omega;C([0,T];H))$. This ensures in turn that the convergences of (u_λ) to u hold along the entire sequence λ, which is independent of ω. In particular, we have that

$$u_\lambda(\omega) \longrightarrow u(\omega) \quad \text{weakly in } L^2(0,T;H) \quad \forall\,\omega\in\Omega'.$$

Since (u_λ) is bounded in $L^2(\Omega\times(0,T)\times D)$, we deduce that u_λ converges weakly to u also in $L^2(\Omega\times(0,T);H)$. Hence, by a direct application of Mazur's lemma, we infer that u is a predictable process with values in H. Unfortunately a similar argument does not apply to η and ξ. In fact, by uniqueness of u, we can only infer from (9.3.3) that $-\operatorname{div}\eta+\xi$ is unique: namely, assume that $(\eta_i(\omega),\xi_i(\omega))$, $i=1,2$, are weak limits in $L^1(0,T;L^1(D))^{d+1}$ of $\big(\gamma_\lambda(\nabla u_\lambda(\omega)),\beta_\lambda(u_\lambda)\big)$ along two subsequences of λ (depending on ω). Then

$$\int_0^t \big(-\operatorname{div}(\eta_1-\eta_2)+(\xi_1-\xi_2)\big)\,ds = 0 \qquad \forall t\in[0,T],$$

hence $-\operatorname{div}(\eta_1-\eta_2)+(\xi_1-\xi_2)=0$, or, equivalently, $-\operatorname{div}\eta_1+\xi_1 = -\operatorname{div}\eta_2+\xi_2$ in V_0' for a.a. $t\in[0,T]$. However, this allows us to claim, setting $\eta_\lambda := \gamma_\lambda(\nabla u_\lambda)$ and $\xi_\lambda := \beta_\lambda(u_\lambda)$, that

$$-\operatorname{div}\eta_\lambda + \xi_\lambda \longrightarrow -\operatorname{div}\eta+\xi \quad \text{weakly in } L^1(0,T;V_0') \quad \forall\omega\in\Omega'$$

along the whole sequence λ, thanks to the same uniqueness argument already used for u. In fact, let us set, for notational convenience,

$$\begin{aligned}
\Phi : L^1(D)^{d+1} &\longrightarrow V_0'\\
(v,f) &\longmapsto -\operatorname{div} v + f
\end{aligned}$$

and $\zeta_\lambda := (\eta_\lambda,\xi_\lambda)$, $\zeta := (\eta,\xi)$. Note that Φ, being a linear bounded operator, can be extended to a linear bounded operator from $L^1((0,T)\times D)^{d+1} \simeq L^1(0,T;L^1(D)^{d+1})$ to $L^1(0,T;V_0')$, also when both spaces are endowed with the weak topology. Then $\zeta_\lambda\to\zeta$ weakly in $L^1((0,T)\times D)^{d+1}$ implies that $\Phi\zeta_\lambda\to\Phi\zeta$ weakly in $L^1(0,T;V_0')$ for all $\omega\in\Omega'$. Such a convergence, however, does not allow to infer that $-\operatorname{div}\eta+\xi$ is predictable as a V_0'-valued process. The reason is that we may certainly find, by Mazur's lemma, a convex combination of $-\operatorname{div}\eta_\lambda+\xi_\lambda$ converging strongly to $-\operatorname{div}\eta+\xi$ in $L^1(0,T;V_0')$ for all $\omega\in\Omega'$, but such a convex combination would depend on ω, bringing us back to the same problem we are trying to solve.* In order to show that $-\operatorname{div}\eta+\xi$ is indeed predictable, we are first going to prove that

$$-\operatorname{div}\eta_\lambda+\xi_\lambda \longrightarrow -\operatorname{div}\eta+\xi \qquad \text{weakly in } L^1(\Omega\times(0,T);V_0').$$

We have just shown that

$$\int_0^T \bigl\langle \Phi\zeta_\lambda(\omega,t),\phi(t)\bigr\rangle\,dt \longrightarrow \int_0^T \bigl\langle \Phi\zeta_\lambda(\omega,t),\phi(t)\bigr\rangle\,dt$$

for all $\phi\in L^\infty(0,T;V_0)$, for all $\omega\in\Omega'$, where $\langle\cdot,\cdot\rangle$ stands for the duality between V_0' and $V_0''=V_0$. Let $\psi\in L^\infty(\Omega\times(0,T);V_0)$. Then $\psi(\omega,\cdot)\in L^\infty(0,T;V_0)$ for $\mathbb{P}$-a.e. $\omega\in\Omega$. Indeed, the set

$$A:=\bigl\{(\omega,t)\in\Omega\times[0,T]:\ \|\psi(\omega,t)\|_{V_0}>\|\psi\|_{L^\infty(\Omega\times(0,T);V_0)}\bigr\}$$

belongs to $\mathscr{F}\otimes\mathscr{B}([0,T])$, and, by Tonelli's theorem,

$$|A|=\int_\Omega\int_0^T 1_A\,dt\,d\mathbb{P}=\int_\Omega \operatorname{Leb}(A_\omega)\,\mathbb{P}(d\omega),$$

where $|A|$ denotes the measure of A and A_ω stands for the section of A at ω, i.e.

$$A_\omega:=\bigl\{t\in[0,T]:\ (\omega,t)\in A\bigr\},$$

which belongs to $\mathscr{B}([0,T])$ for $\mathbb{P}$-a.e. $\omega\in\Omega$. Since $|A|=0$, it follows that $|A_\omega|=0$ for $\mathbb{P}$-a.e. $\omega\in\Omega$. This implies, by definition of A, that $\psi(\omega,\cdot)\in L^\infty(0,T)$ for $\mathbb{P}$-a.e. $\omega\in\Omega$. Consequently, we have

$$\int_0^T \bigl\langle \Phi\zeta_\lambda(\omega,t),\psi(\omega,t)\bigr\rangle\,dt \longrightarrow \int_0^T \bigl\langle \Phi\zeta(\omega,t),\psi(\omega,t)\bigr\rangle\,dt$$

for $\mathbb{P}$-a.e. $\omega\in\Omega$. To complete the argument it is then enough to show that the left-hand side, as a subset of $L^0(\Omega)$ indexed by λ, is uniformly integrable. To this end, we collect some simple facts about uniform integrability in the following lemma.

Lemma 9.3.4. *Let $(X,\mathscr{A},m)$ be a finite measure space and I an arbitrary index set.*

(a) *Let $(f_i)_{i\in I}$, $(g_i)_{i\in I}\subset L^0(X;\mathbb{R}^n)$ be such that $|f_i|\le|g_i|$ for all $i\in I$ and assume that (g_i) is uniformly integrable. Then (f_i) is uniformly integrable.*

(b) *Let $(f_i)\subset L^0(X;\mathbb{R}^n)$ be uniformly integrable and $\phi\in L^\infty(X;\mathbb{R}^n)$. Then $(\phi\cdot f_i)\subset L^0(X)$ is uniformly integrable.*

(c) *Let $F:\mathbb{R}^n\to\mathbb{R}$ with $F(0)=0$ be convex and superlinear at infinity, and $(f_i)\subset L^0(X;\mathbb{R}^n)$ be such that $(F\circ f_i)$ is bounded in $L^1(X)$. Then (f_i) is uniformly integrable.*

(d) *Let $(Y,\mathscr{B},n)$ be a further finite measure space. If $(f_i)\subset L^0(X\times Y,\mathscr{A}\otimes\mathscr{B},m\otimes n;\mathbb{R}^n)$ is uniformly integrable, then $(g_i)\subset L^0(X;\mathbb{R}^n)$ defined by*

$$g_i:=\int_Y f_i(\cdot,y)\,n(dy)$$

is uniformly integrable.

*We could just say that $-\operatorname{div}\eta+\xi$ is weakly measurable with respect to $\mathscr{F}$ and the Borel σ-algebra of $L^1(0,T;V_0')$. Since this space is separable, by Pettis' theorem we also have strong measurability. This observation, however, does not seem to imply the desired result.

Proof. (a) is an immediate consequence of the definition of uniform integrability.

(b) Let $\varepsilon > 0$. By assumption, there exists $\delta = \delta(\varepsilon) > 0$ such that

$$\int_A \bigl|f_i\bigr|_{\mathbb{R}^n} \, dm < \frac{\varepsilon}{\|\phi\|_{L^\infty}} \qquad \forall A \in \mathscr{A},\ m(A) < \delta.$$

Then

$$\int_A \bigl|\phi \cdot f_i\bigr| \, dm \leq \|\phi\|_{L^\infty} \int_A \bigl|f_i\bigr|_{\mathbb{R}^n} \, dm < \varepsilon.$$

(c) is a variation of the classical criterion by de la Vallée-Poussin. A detailed proof (which is nonetheless very close to the one in the standard one-dimensional case) can be found in Chapter 7.

(d) Let $\varepsilon > 0$. By assumption, there exists $\delta' = \delta'(\varepsilon) > 0$ such that

$$\int_C \bigl|f_i\bigr|_{\mathbb{R}^n} \, dm \otimes n < \varepsilon \qquad \forall C \in \mathscr{A} \otimes \mathscr{B},\ m \otimes n(C) < \delta'.$$

Let $\delta := \delta'/n(Y)$ and $A \in \mathscr{A}$ with $m(A) < \delta$. Then

$$\int_A \biggl|\int_Y f_i(x,y)\, n(dy)\biggr|_{\mathbb{R}^n} m(dx) \leq \int_{A\times Y} \bigl|f_i(x,y)\bigr|_{\mathbb{R}^n} m(dx)\, n(dy) < \varepsilon$$

because $m \otimes n(A \times Y) = m(A)n(Y) < \delta n(Y) = \delta'$. $\square$

Let us now resume with the main reasoning. Since

$$\int_0^T \bigl\langle \Phi\zeta_\lambda, \psi \bigr\rangle \lesssim \|\psi\|_{L^\infty(\Omega\times(0,T);V_0)} \biggl(\int_0^T\!\!\int_D |\eta_\lambda| + \int_0^T\!\!\int_D |\xi_\lambda|\biggr),$$

by parts (a), (b) and (d) of the previous lemma it is sufficient to show that (η_λ) and (ξ_λ) are uniformly integrable in $\Omega\times(0,T)\times D$. But this is true, in view of part (c) of the previous lemma, because $k^*(\eta_\lambda)$ and $j^*(\xi_\lambda)$ are uniformly bounded in $L^1(\Omega\times(0,T)\times D)$. Vitali's theorem then yields

$$\int_0^T \bigl\langle \Phi\zeta_\lambda(\omega,t), \psi(\omega,t)\bigr\rangle \, dt \longrightarrow \int_0^T \bigl\langle \Phi\zeta(\omega,t), \psi(\omega,t)\bigr\rangle \, dt \qquad \text{in } L^1(\Omega),$$

hence, in particular,

$$\Phi(\eta_\lambda, \xi_\lambda) \longrightarrow \Phi(\eta, \xi) \quad \text{weakly in } L^1(\Omega\times(0,T);V_0').$$

Furthermore, from the uniform integrability of (η_λ) and (ξ_λ) in $\Omega\times(0,T)\times D$ it also follows that, along a subsequence μ of λ,

$$(\eta_\mu, \xi_\mu) \longrightarrow (\bar\eta, \bar\xi) \quad \text{weakly in } L^1(\Omega\times(0,T)\times D)^{d+1},$$

hence also

$$\Phi(\eta_\mu, \xi_\mu) \longrightarrow \Phi(\bar\eta, \bar\xi) \quad \text{weakly in } L^1(\Omega\times(0,T);V_0').$$

An application of Mazur's lemma yields, in complete analogy to the case of u, that $\bar\eta$ and $\bar\xi$ are predictable processes with values in $L^1(D)^d$ and $L^1(D)$, respectively. Since μ is a subsequence of λ, by uniqueness of the weak limit we have that $\Phi(\eta,\xi) = \Phi(\bar\eta,\bar\xi)$, i.e.

$$-\operatorname{div}\eta + \xi = -\operatorname{div}\bar\eta + \bar\xi.$$

This implies that the identity (9.3.3) remains true with η and ξ replaced by $\bar{\eta}$ and $\bar{\xi}$, respectively. In other words, modulo relabelling, we can just assume, without loss of generality, that η and ξ in (9.3.3) are predictable and that

$$(\eta_\lambda, \xi_\lambda) \longrightarrow (\eta, \xi) \qquad \text{weakly in } L^1(\Omega \times (0,T) \times D)^{d+1}.$$

By weak lower semicontinuity and Lemma 9.3.3, this also implies, arguing as in Chapters 2, 6 and 7, that

$$\begin{aligned} u &\in L^2(\Omega; L^\infty(0,T;H)) \cap L^1(\Omega; L^1(0,T;W_0^{1,1}(D))),\\ \eta &\in L^1(\Omega \times (0,T) \times D)^d,\\ \xi &\in L^1(\Omega \times (0,T) \times D),\\ k(\nabla u) + k^*(\eta) = \nabla u \cdot \eta &\in L^1(\Omega \times (0,T) \times D),\\ j(u) + j^*(\xi) = u\xi &\in L^1(\Omega \times (0,T) \times D). \end{aligned}$$

In order to show that $\eta \in \gamma(\nabla u)$ and $\xi \in \beta(u)$ a.e. in $\Omega \times (0,T) \times D$, it suffices to prove, by the maximal monotonicity of γ and β, that

$$\limsup_{\lambda \to 0} \mathbb{E} \int_0^T \!\! \int_D \big(\eta_\lambda \cdot \nabla u_\lambda + \xi_\lambda u_\lambda\big) \leq \mathbb{E} \int_0^T \!\! \int_D \big(\eta \cdot \nabla u + \xi u\big). \tag{9.3.4}$$

To this purpose, note that the ordinary Itô formula and Proposition 9.3.1 yield

$$\frac{1}{2}\,\mathbb{E}\|u_\lambda(T)\|^2 + \mathbb{E}\int_0^T\!\!\int_D \big(\eta_\lambda \cdot \nabla u_\lambda + \xi_\lambda u_\lambda\big) = \frac{1}{2}\,\mathbb{E}\|u_0\|^2 + \frac{1}{2}\,\mathbb{E}\int_0^T \big\|B(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds$$

and

$$\frac{1}{2}\,\mathbb{E}\|u(T)\|^2 + \mathbb{E}\int_0^T\!\!\int_D \big(\eta \cdot \nabla u + \xi u\big) = \frac{1}{2}\,\mathbb{E}\|u_0\|^2 + \frac{1}{2}\,\mathbb{E}\int_0^T \big\|B(s)\big\|^2_{\mathscr{L}^2(U,H)}\,ds,$$

respectively (the stochastic integrals appearing in both versions of Itô's formula are in fact martingales, not just local martingales, hence their expectation is zero). Since $u_\lambda(T) \rightharpoonup u(T)$ in $L^2(\Omega;H)$, one has $\mathbb{E}\|u(T)\|^2 \leq \liminf_{\lambda\to 0} \mathbb{E}\|u_\lambda(T)\|^2$, hence, by comparison, (9.3.4) follows.

Finally, the strong pathwise continuity (in H) of u is an immediate consequence of the corollary to Proposition 9.3.1.

Remark 9.3.5. Another way to "restore" uniqueness of limits for the pair $\zeta_\lambda = (\eta_\lambda, \xi_\lambda)$ is to view it as element of the quotient space $L^1(D)^{d+1}/M$, where $M := \ker \Phi$. Note that M is a closed subset of L^1 (we suppress the superscript as well as the indication of the domain just within this remark), as the inverse image of the closed set $\{0\}$ through a continuous linear map, hence L^1/M is a Banach space. However, working with the spaces $L^1(0,T;L^1/M)$ and $L^1(\Omega \times (0,T); L^1/M)$ present technical difficulties due to the fact that their dual spaces are hard to characterize. A bit more precisely, this has to do with the fact that the dual of $L^1(0,T;E)$ is $L^\infty(0,T;E')$ if and only if E has the Radon-Nikodym property. This property is enjoyed by reflexive spaces, but not by L^1 spaces (see, e.g., [33]).

9.3.3 Well-posedness in the general case

Let us consider now equation (9.1.1) with general additive noise, i.e. with

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

Thanks to classical elliptic regularity results, there exists $m \in \mathbb{N}$ such that the $(I - \delta\Delta)^{-m}$ is a continuous linear map from $L^1(D)$ to $W^{1,\infty}(D) \cap H_0^1(D)$ for every $\delta > 0$. Setting then $V_0 := (I-\Delta)^{-m}(H)$ and $B^\delta := (I-\delta\Delta)^{-m}B$, we have $B^\delta \in L^2(\Omega; L^2(0,T; \mathscr{L}^2(U, V_0)))$, hence, by the well-posedness results already obtained, the equation

$$du^\delta - \operatorname{div}\gamma(\nabla u^\delta)\,dt + \beta(u^\delta)\,dt \ni B^\delta\,dW, \qquad u^\delta(0) = u_0,$$

admits a strong solution $(u^\delta, \eta^\delta, \xi^\delta)$. Arguing as in Chapters 2, 6 and 7, one can show using Itô's formula that (u^δ) is a Cauchy sequence in $L^2(\Omega; C([0,T]; H))$ and that (∇u^δ), (η^δ), and (ξ^δ) are relatively weakly compact in $L^1(\Omega \times (0,T) \times D)$, so that

$$\begin{aligned}
u^\delta &\longrightarrow u \quad && \text{in } L^2(\Omega; C([0,T]; H)),\\
u^\delta &\longrightarrow u && \text{weakly in } L^1(\Omega \times (0,T); W_0^{1,1}(D)),\\
\eta^\delta &\longrightarrow \eta && \text{weakly in } L^1(\Omega \times (0,T) \times D)^d,\\
\xi^\delta &\longrightarrow \xi && \text{weakly in } L^1(\Omega \times (0,T) \times D),
\end{aligned}$$

from which it follows that (u, η, ξ) solves the original equation. Moreover, the strong-weak closure of β readily implies that $\xi \in \beta(u)$ a.e. in $\Omega \times (0,T) \times D$. Finally, arguing as in the previous subsection, by weak lower semicontinuity of convex integrals and Itô's formula one can show that

$$\limsup_{\lambda \to 0} \mathbb{E} \int_0^T \int_D \eta_\lambda \cdot \nabla u_\lambda \leq \mathbb{E} \int_0^T \int_D \eta \cdot \nabla u,$$

so that $\eta \in \gamma(\nabla u)$ a.e. in $\Omega \times (0,T) \times D$ as well.

Continuous dependence on the initial datum is a consequence of Itô's formula and the monotonicity of γ and β. Finally, the generalization to the case of multiplicative noise follows using the Lipschitz continuity of B and a classical fixed point argument. A detailed exposition of the arguments needed to prove these claims can be found in Chapters 2, 6 and 7.

Bibliography

[1] S. Albeverio, H. Kawabi, and M. Röckner, *Strong uniqueness for both Dirichlet operators and stochastic dynamics to Gibbs measures on a path space with exponential interactions*, J. Funct. Anal. **262** (2012), no. 2, 602–638. MR 2854715

[2] C. D. Aliprantis and K. C. Border, *Infinite dimensional analysis*, third ed., Springer, Berlin, 2006. MR MR2378491

[3] A. Ambrosetti and G. Prodi, *A primer of nonlinear analysis*, Cambridge University Press, Cambridge, 1995. MR MR1336591 (96a:58019)

[4] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford University Press, New York, 2000. MR 1857292 (2003a:49002)

[5] W. Arendt, *Heat kernels*, Lecture Notes of the 9th Internet Seminar on Evolution Equations, 2006, available at `https://www.uni-ulm.de/mawi/iaa/members/arendt/`.

[6] W. Arendt and A. V. Bukhvalov, *Integral representations of resolvents and semigroups*, Forum Math. **6** (1994), no. 1, 111–135. MR 1253180

[7] W. Arendt, R. Chill, C. Seifert, D. Vogt, and J. Voigt, *Form methods for evolution equations and applications*, Lecture Notes of the 18th Internet Seminar on Evolution Equations, 2015, available at `https://www.mat.tuhh.de/isem18/Phase_1:_The_lectures`.

[8] V. Barbu, *Analysis and control of nonlinear infinite-dimensional systems*, Academic Press Inc., Boston, MA, 1993. MR MR1195128 (93j:49002)

[9] ———, *Existence for semilinear parabolic stochastic equations*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **21** (2010), no. 4, 397–403. MR 2746091 (2012d:35424)

[10] ———, *Nonlinear differential equations of monotone types in Banach spaces*, Springer Monographs in Mathematics, Springer, New York, 2010. MR 2582280 (2011d:34001)

[11] V. Barbu and G. Da Prato, *Ergodicity for nonlinear stochastic equations in variational formulation*, Appl. Math. Optim. **53** (2006), no. 2, 121–139. MR MR2172782 (2007d:60030)

[12] V. Barbu, G. Da Prato, and M. Röckner, *Existence of strong solutions for stochastic porous media equation under general monotonicity conditions*, Ann. Probab. **37** (2009), no. 2, 428–452. MR MR2510012

[13] V. Barbu and C. Marinelli, *Strong solutions for stochastic porous media equations with jumps*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **12** (2009), no. 3, 413–426. MR 2572464

[14] A. Bendikov and P. Maheux, *Nash type inequalities for fractional powers of non-negative self-adjoint operators*, Trans. Amer. Math. Soc. **359** (2007), no. 7, 3085–3097 (electronic). MR 2299447

[15] P. Benilan, H. Brezis, and M. G. Crandall, *A semilinear equation in $L^1(R^N)$*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **2** (1975), no. 4, 523–555. MR 0390473

[16] L. Boccardo, *Some nonlinear Dirichlet problems in L^1 involving lower order terms in divergence form*, Progress in elliptic and parabolic partial differential equations (Capri, 1994), Pitman Res. Notes Math. Ser., vol. 350, Longman, Harlow, 1996, pp. 43–57. MR 1430139

[17] L. Boccardo and G. Croce, *Elliptic partial differential equations*, De Gruyter, Berlin, 2014. MR 3154599

[18] L. Boccardo and T. Gallouët, *Strongly nonlinear elliptic equations having natural growth terms and L^1 data*, Nonlinear Anal. **19** (1992), no. 6, 573–579. MR 1183664

[19] N. Bourbaki, *Intégration. Chapitre IX: Intégration sur les espaces topologiques séparés*, Hermann, Paris, 1969. MR 0276436

[20] ———, *Espaces vectoriels topologiques. Chapitres 1 à 5*, new ed., Masson, Paris, 1981. MR 633754

[21] H. Brézis, *Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations*, Contributions to nonlinear functional analysis (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1971), Academic Press, New York, 1971, pp. 101–156. MR MR0394323 (52 #15126)

[22] ———, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Publishing Co., Amsterdam, 1973. MR MR0348562 (50 #1060)

[23] ———, *Functional analysis, Sobolev spaces and partial differential equations*, Springer, New York, 2011. MR 2759829 (2012a:35002)

[24] H. Brézis and W. A. Strauss, *Semi-linear second-order elliptic equations in L^1*, J. Math. Soc. Japan **25** (1973), 565–590. MR 0336050 (49 #826)

[25] Z. Brzeźniak, B. Maslowski, and J. Seidler, *Stochastic nonlinear beam equations*, Probab. Theory Related Fields **132** (2005), no. 1, 119–149. MR 2136869

[26] S. Cerrai, *Second order PDE's in finite and infinite dimension*, Lecture Notes in Mathematics, vol. 1762, Springer-Verlag, Berlin, 2001. MR 2002j:35327

[27] ———, *Stochastic reaction-diffusion systems with multiplicative noise and non-Lipschitz reaction term*, Probab. Theory Related Fields **125** (2003), no. 2, 271–304. MR 1961346 (2004a:60117)

[28] G. Da Prato, *Kolmogorov equations for stochastic PDEs*, Birkhäuser Verlag, Basel, 2004. MR MR2111320 (2005m:60002)

[29] G. Da Prato and J. Zabczyk, *Ergodicity for infinite-dimensional systems*, Cambridge University Press, Cambridge, 1996. MR MR1417491 (97k:60165)

[30] ______, *Stochastic equations in infinite dimensions*, second ed., Cambridge University Press, Cambridge, 2014. MR 3236753

[31] R. Dautray and J.-L. Lions, *Mathematical analysis and numerical methods for science and technology. Vol. 2*, Springer-Verlag, Berlin, 1988. MR 969367

[32] E. B. Davies, *Heat kernels and spectral theory*, Cambridge University Press, Cambridge, 1990. MR MR1103113 (92a:35035)

[33] J. Diestel and J. J. Uhl, Jr., *Vector measures*, American Mathematical Society, Providence, R.I., 1977. MR 0453964

[34] A. Eberle, *Uniqueness and non-uniqueness of semigroups generated by singular diffusion operators*, Lecture Notes in Mathematics, vol. 1718, Springer-Verlag, Berlin, 1999. MR MR1734956 (2001c:60122)

[35] M. Emery, *Équations différentielles stochastiques lipschitziennes: étude de la stabilité*, Séminaire de Probabilités, XIII (Univ. Strasbourg, Strasbourg, 1977/78), Lecture Notes in Math., vol. 721, Springer, Berlin, 1979, pp. 281–293. MR 544801 (81h:60075)

[36] K.-J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*, Springer-Verlag, New York, 2000. MR MR1721989 (2000i:47075)

[37] A. Es-Sarhir and W. Stannat, *Invariant measures for semilinear SPDE's with local Lipschitz drift coefficients and applications*, J. Evol. Equ. **8** (2008), no. 1, 129–154. MR 2383485

[38] ______, *Improved moment estimates for invariant measures of semilinear diffusions in Hilbert spaces and applications*, J. Funct. Anal. **259** (2010), no. 5, 1248–1272. MR 2652188

[39] I. Gentil and C. Imbert, *The Lévy-Fokker-Planck equation: Φ-entropies and convergence to equilibrium*, Asymptot. Anal. **59** (2008), no. 3-4, 125–138. MR 2450356

[40] A. Guessab and G. Schmeisser, *Necessary and sufficient conditions for the validity of Jensen's inequality*, Arch. Math. (Basel) **100** (2013), no. 6, 561–570. MR 3069109

[41] M. Haase, *Convexity inequalities for positive operators*, Positivity **11** (2007), no. 1, 57–68. MR 2297322 (2008d:39034)

[42] J.-B. Hiriart-Urruty and C. Lemaréchal, *Fundamentals of convex analysis*, Springer-Verlag, Berlin, 2001. MR 1865628 (2002i:90002)

[43] H. Kardestuncer and D. H. Norrie (eds.), Chapters 1–3 in *finite element handbook*, McGraw-Hill Book Co., New York, 1987. MR 900813 (89g:65001)

[44] T. Kato, *Perturbation theory for linear operators*, Springer-Verlag, Berlin, 1995, Reprint of the 1980 edition. MR 1335452

[45] S. G. Kreĭn, Yu. Ī. Petunīn, and E. M. Semënov, *Interpolation of linear operators*, Translations of Mathematical Monographs, vol. 54, American Mathematical Society, Providence, R.I., 1982. MR 649411 (84j:46103)

[46] N. V. Krylov and B. L. Rozovskiĭ, *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 MR570795 (81m:60116)

[47] M. Kunze and J. van Neerven, *Continuous dependence on the coefficients and global existence for stochastic reaction diffusion equations*, J. Differential Equations **253** (2012), no. 3, 1036–1068. MR 2922662

[48] Se. Kusuoka and C. Marinelli, *On smoothing properties of transition semigroups associated to a class of SDEs with jumps*, Ann. Inst. Henri Poincaré Probab. Stat. **50** (2014), no. 4, 1347–1370. MR 3269997

[49] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod; Gauthier-Villars, Paris, 1969. MR 0259693

[50] J.-L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications. Vol. 1*, Dunod, Paris, 1968. MR 0247243

[51] ______, *Non-homogeneous boundary value problems and applications. Vol. I*, Springer-Verlag, New York-Heidelberg, 1972, Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181. MR 0350177

[52] ______, *Non-homogeneous boundary value problems and applications. Vol. II*, Springer-Verlag, New York-Heidelberg, 1972, Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 182. MR 0350178

[53] ______, *Non-homogeneous boundary value problems and applications. Vol. III*, Springer-Verlag, New York-Heidelberg, 1973, Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 183. MR 0350179

[54] W. Liu, *On the stochastic p-Laplace equation*, J. Math. Anal. Appl. **360** (2009), no. 2, 737–751. MR 2561270

[55] ______, *Well-posedness of stochastic partial differential equations with Lyapunov condition*, J. Differential Equations **255** (2013), no. 3, 572–592. MR 3053478

[56] W. Liu and M. Röckner, *Stochastic partial differential equations: an introduction*, Springer, Cham, 2015. MR 3410409

[57] Z. M. Ma and M. Röckner, *Introduction to the theory of (nonsymmetric) Dirichlet forms*, Springer-Verlag, Berlin, 1992. MR 1214375

[58] C. Marinelli, *On well-posedness of semilinear stochastic evolution equations on L_p spaces*, arXiv:1512.04323.

[59] C. Marinelli, C. Prévôt, and M. Röckner, *Regular dependence on initial data for stochastic evolution equations with multiplicative Poisson noise*, J. Funct. Anal. **258** (2010), no. 2, 616–649. MR MR2557949

[60] C. Marinelli and Ll. Quer-Sardanyons, *Existence of weak solutions for a class of semilinear stochastic wave equations*, SIAM J. Math. Anal. **44** (2012), no. 2, 906–925. MR 2914254

[61] C. Marinelli and M. Röckner, *On the maximal inequalities of Burkholder, Davis and Gundy*, Expo. Math. **34** (2016), no. 1, 1–26. MR 3463679

[62] C. Marinelli and L. Scarpa, *Ergodicity and Kolmogorov equations for dissipative SPDEs with singular drift: a variational approach*, ArXiv e-prints (2017).

[63] ———, *Refined existence and regularity results for a class of semilinear dissipative SPDEs*, ArXiv e-prints (2017).

[64] ———, *On the well-posedness of SPDEs with singular drift in divergence form*, Stochastic Partial Differential Equations and Related Fields (Cham) (A. Eberle, M. Grothaus, W. Hoh, M. Kassmann, W. Stannat, and G. Trutnau, eds.), Springer International Publishing, 2018, pp. 225–235.

[65] ———, *A variational approach to dissipative SPDEs with singular drift*, Ann. Probab. **46** (2018), no. 3, 1455–1497. MR 3785593

[66] ———, *A note on doubly nonlinear spdes with singular drift in divergence form*, Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl. (in press), arXiv:1712.05595.

[67] ———, *Strong solutions to SPDEs with monotone drift in divergence form*, Stoch. Partial Differ. Equ. Anal. Comput. (in press), arXiv:1612.08260.

[68] C. Marinelli and G. Ziglio, *Ergodicity for nonlinear stochastic evolution equations with multiplicative Poisson noise*, Dyn. Partial Differ. Equ. **7** (2010), no. 1, 1–23. MR 2656416

[69] C. Orrieri and L. Scarpa, *Singular Stochastic Allen-Cahn equations with dynamic boundary conditions*, ArXiv e-prints (2017).

[70] E. M. Ouhabaz, *Analysis of heat equations on domains*, Princeton University Press, Princeton, NJ, 2005. MR 2124040

[71] É. Pardoux, *Sur des équations aux dérivées partielles stochastiques monotones*, C. R. Acad. Sci. Paris Sér. A-B **275** (1972), A101–A103. MR 0312572

[72] ———, *Equations aux derivées partielles stochastiques nonlinéaires monotones*, Ph.D. thesis, Université Paris XI, 1975.

[73] É. Pardoux and A. Răşcanu, *Stochastic differential equations, backward SDEs, partial differential equations*, Springer, Cham, 2014. MR 3308895

[74] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer-Verlag, New York, 1983. MR 85g:47061

[75] L. Scarpa, *Well-posedness for a class of doubly nonlinear stochastic PDEs of divergence type*, J. Differential Equations **263** (2017), no. 4, 2113–2156. MR 3650335

[76] ———, *On the stochastic Cahn-Hilliard equation with a singular double-well potential*, Nonlinear Anal. **171** (2018), 102–133. MR 3778274

[77] J. Simon, *Compact sets in the space $L^p(0,T;B)$*, Ann. Mat. Pura Appl. (4) **146** (1987), 65–96. MR 916688 (89c:46055)

[78] W. Stannat, *Stochastic partial differential equations: Kolmogorov operators and invariant measures*, Jahresber. Dtsch. Math.-Ver. **113** (2011), no. 2, 81–109. MR 2768734

[79] W. A. Strauss, *On continuity of functions with values in various Banach spaces*, Pacific J. Math. **19** (1966), 543–551. MR 0205121 (34 #4956)

[80] J. van Neerven, M. C. Veraar, and L. Weis, *Stochastic evolution equations in UMD Banach spaces*, J. Funct. Anal. **255** (2008), no. 4, 940–993. MR 2433958 (2009h:35465)

[81] N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon, *Analysis and geometry on groups*, Cambridge University Press, Cambridge, 1992. MR 1218884