Stochastic Optimal Control in Infinite Dimensions: Dynamic Programming and HJB Equations

G. Fabbri[1] F. Gozzi[2] and A. Święch[3]

with Chapter 6 by M. Fuhrman[4] and G. Tessitore[5]

[1]Aix-Marseille University (Aix-Marseille School of Economics), CNRS and EHESS. e-mail:giorgio.fabbri@univ-amu.fr

[2]Dipartimento di Economia e Finanza, Università *LUISS - Guido Carli* Roma, e-mail: fgozzi@luiss.it

[3]School of Mathematics, Georgia Institute of Technology, Atlanta, e-mail: swiech@math.gatech.edu

[4]Dipartimento di Matematica, Politecnico di Milano, e-mail: marco.fuhrman@polimi.it

[5]Dipartimento di Matematica, Università di Milano Bicocca, e-mail: gianmario.tessitore@unimib.it

Contents

Preface

The main objective of this book is to give an overview of the theory of Hamilton-Jacobi-Bellman (HJB) partial differential equations (PDE) in infinite dimensional Hilbert spaces and its applications to stochastic optimal control of infinite dimensional processes and related fields. Both areas have been developing very rapidly in the last few decades. While there exist several excellent monographs on this subject in finite dimensional spaces (see e.g. [194, 195, 294, 349, 364, 382, 449]), much less has been written in infinite dimensional spaces. A good account of the infinite dimensional case in the deterministic context can be found in [312]. Other books that touch the subject are [23, 129, 364]. We attempt to fill this gap in the literature. Infinite dimensional diffusion processes appear naturally and are used to model phenomena in physics, biology, chemistry, economics, mathematical finance, engineering and many other areas (see e.g. [93, 127, 130, 282, 444]. This book investigates the PDE approach to their stochastic optimal control, however infinite dimensional PDE can also be used to study other properties of such processes as large deviations, invariant measures, stochastic viability, stochastic differential games for infinite dimensional diffusions, etc. (see [60, 127, 129, 182, 184, 192, 361, 363, 425, 428]).

To illustrate the main theme of the book let us begin with a model distributed parameter stochastic optimal control problem. We want to control a process (called the state) given by an abstract stochastic differential equation in a real, separable Hilbert space H

$$\begin{cases} dX(s) = (AX(s) + b(s, X(s), a(s)))ds + \sigma(s, X(s), a(s))dW(s), \quad s > t \geq 0 \\ X(t) = x \in H, \end{cases}$$

where A is the generator of a C_0 semigroup in H, b, σ are certain bounded functions and W is a so called Q-Wiener process[1] in H. The functions $a(\cdot)$, called controls, are stochastic processes with values in some metric space Λ, which satisfy certain measurability properties. The above abstract stochastic differential equation is very general and includes various semi-linear stochastic PDE, as well as other equations which can be rewritten as stochastic functional evolution equations, for instance stochastic differential delay equations. In a most typical optimal control problem we want to find a control $a(\cdot)$, called optimal, which minimizes a cost functional

$$J(t, x; a(\cdot)) = \mathbb{E}\left[\int_t^T l(s, X(s), a(s))ds + g(X(T))\right].$$

(for some $T > t$) among all admissible controls for some functions $l\colon [0, T] \times H \times \Lambda \to \mathbb{R}$, $g\colon H \to \mathbb{R}$.

The dynamic programming approach to the above problem is based on studying the properties of the so called value function

$$V(t, x) = \inf_{a(\cdot)} J(t, x; a(\cdot))$$

[1] Q is a suitable self-adjoint positive operator in H, the covariance operator for W.

and characterizing it as a solution of a fully nonlinear PDE, the associated HJB equation. Since the state $X(s)$ evolves in the infinite dimensional space H, this PDE is defined in $[0,T] \times H$. The link between the value function V and the HJB equation is established by the Bellman principle of optimality known as the dynamic programming principle (DPP),

$$V(t,x) = \inf_{a(\cdot)} \mathbb{E}\left[\int_t^\eta l\left(s, X(s), a(s)\right) ds + V\left(\eta, X(\eta)\right)\right], \quad \text{for all } \eta \in [t,T].$$

Heuristically the DPP can be used to define a two parameter nonlinear evolution system and the associated HJB equation

$$\begin{cases} V_t + \langle Ax, DV\rangle + \inf_{a\in\Lambda}\Big\{\frac{1}{2}\mathrm{Tr}\left[(\sigma(t,x,a)Q^{\frac{1}{2}})(\sigma(t,x,a)Q^{\frac{1}{2}})^* D^2V\right] \\ \qquad\qquad\qquad\qquad + \langle b(t,x,a), DV\rangle + l(t,x,a)\Big\} = 0, \\ V(T,x) = g(x). \end{cases} \tag{0.1}$$

is its generating equation. Such PDE is called infinite dimensional or a PDE in infinitely many variables. We also call it unbounded since it has a term with an unbounded operator A which is well defined only on the domain of A. Other terms may also be undefined for some values of DV and D^2V, the Fréchet derivatives of V, which we may identify with elements of H and with bounded, self-adjoint operators in H respectively. In particular, the term $\mathrm{Tr}[(\sigma Q^{\frac{1}{2}})(\sigma Q^{\frac{1}{2}})^* D^2V]$ is well defined only if $(\sigma Q^{\frac{1}{2}})(\sigma Q^{\frac{1}{2}})^* D^2V$ is of trace class.

The main idea is to use the HJB equation to study properties of the value function, find conditions for optimality, obtain formulas for synthesis of optimal feedback controls, etc. This approach turned out to be very successful for finite dimensional problems because of its clarity and simplicity and thanks to the developments of the theory of fully nonlinear elliptic and parabolic PDE, in particular the introduction of the notion of viscosity solution and advances in the regularity theory. However even there many open questions remain, especially if the HJB equations are degenerate. We hope the dynamic programming approach will be equally valuable for infinite dimensional problems even though a complete theory is not available yet.

Equation (0.1) is an example of a fully nonlinear second-order PDE of (degenerate) parabolic type. In the book we will deal with more general and different versions of such equations and their degenerate elliptic counterparts. If Λ is a singleton, (0.1) is just a terminal value problem for a linear Kolmogorov equation. If Λ is not a singleton but the diffusion coefficient σ is independent of the control parameter a, (0.1) is semi-linear. The theory of linear equations (and some special semi-linear ones) has been studied by many authors and can be found in books [23, 80, 129, 453]. The emphasis of this book is on semi-linear and fully nonlinear equations.

There are several notions of solution applicable to PDE in Hilbert spaces which are discussed in the book: classical solutions, strong solutions, mild solutions in the space of continuous functions, solutions in $L^2(\mu)$, viscosity solutions. Classical solutions are the most regular ones. This notion of solution requires $C^{1,2}$ regularity in the Fréchet sense and imposes additional conditions so that all terms in the equation make sense pointwise for $(t,x) \in [0,T] \times H$. When classical solutions exist we can apply the classical dynamic programming approach to obtain verification theorems and the synthesis of optimal feedback controls. Unfortunately in almost all interesting cases it is not possible to find such solutions, however they are very useful as a theoretical tool in the theory. The notions of strong solutions, mild solutions in the space of continuous functions, and solutions in $L^2(\mu)$ are introduced and studied only for semi-linear equations and define solutions which have at least first derivative (in some suitable sense). Verification theorems and synthesis of optimal feedback controls can still be developed within their framework. The notion of viscosity solutions is the most general and applies to fully nonlinear equations, however at the current stage there are no results on verification theorems and synthesis of optimal feedback controls.

Infinite dimensional problems present unique challenges, among them are the lack of local compactness and no equivalent of Lebesgue measure. This means that standard finite dimensional elliptic and parabolic techniques which are based on measure theory cannot be carried over to the infinite dimensional case. Moreover, the equations are mostly degenerate and contain unbounded terms which are singular. So the methods to find regular solutions to PDE in infinite dimensions like ours tend to be global and are based on semigroup theory, smoothing properties of transition semigroups (like the Ornstein-Uhlenbeck ones), fixed point techniques, and stochastic analysis. These methods are mostly restricted to equations of semi-linear type. On the other hand, the notion of viscosity solution is perfectly suited for fully nonlinear equations. It is local and it does not require any regularity of solutions except continuity. As in finite dimensions it is based on maximum principle through the idea of "differentiation by parts", i.e. replacing the non existing derivatives of viscosity subsolutions (respectively, supersolutions) by derivatives of smooth test functions at points where their graphs touch the graphs of subsolutions (respectively, supersolutions) from above (respectively, below). However as the readers will see, this idea has to be carried out very carefully in infinite dimensions.

The book contains chapters on the most important topics in HJB equations and the DPP approach to infinite dimensional stochastic optimal control.

Chapter 1 contains basic material on infinite dimensional stochastic calculus which is needed in the subsequent chapters. It is however not intended to be an introduction to stochastic calculus and the readers are expected to have some familiarity with it. Chapter 1 is included to make the book more self-contained. Most of the results presented there are well known, hence we only provide references where the reader can find the proofs and find more information about the concepts, examples, etc. We provide proofs only in cases where we could not find good references in the literature.

In Chapter 2 we introduce a general stochastic optimal control problem and prove a key result in the theory, namely the dynamic programming principle. We formulate it in an abstract and general form so that it can be used in many cases without the need of reproving it. Solutions of stochastic PDE must be interpreted in various ways (strong, mild, variational, etc.) and our formulation of the DPP tries to capture this phenomenon. Our proof of the DPP is based on standard ideas however we tried to avoid heavy probabilistic methods regarding weak uniqueness of solutions of stochastic differential equations. Our proof is thus more analytical.

We also introduce many examples of stochastic optimal control problems which can be studied in the framework of the approach presented in the book. They should give the readers an idea about the range and applicability of the material.

Chapter 3 is devoted to the theory of viscosity solutions. The reader should keep in mind the following principle when it comes to unbounded PDE in infinite dimensions: there is no single definition of viscosity solutions that applies to all equations. This is due to the fact that there are many different PDE which contain different unbounded operators and terms which are continuous in various norms. Also the solutions have to be continuous with respect to weaker topologies. However the main idea of the notion of viscosity solutions is always the same as we described before. What changes is the choice of test functions, spaces, topologies, and the interpretation of various terms in the equation. In this book we focus on the notion of so called B-continuous viscosity solution which was introduced by Crandall and P. L. Lions in [103, 104] for first order equations and later adapted to second order equations in [422]. The key result in the theory is the comparison principle which is very technical. Its main component is the so called maximum principle for semicontinuous functions. The proof of such result in finite dimensions was first obtained in [280] and was later simplified and generalized [99, 100, 101, 270]. It is heavily based on measure theory and is not applicable to infinite dimensions. Thus the theory uses a finite dimensional reduction technique introduced by P. L. Lions in [321]. It restricts the class of equations which can be considered, in particular they have to be highly degenerate in the second order terms. We present three techniques to obtain existence of viscosity solutions. The first and most important for this book is the DPP and the stochastic optimal control interpretation, showing directly that the value function is a viscosity solution. This technique applies to HJB equations. The other techniques are finite dimensional approximations and Perron's method. Both can be applied to more general equations, for instance Isaacs equations associated to two-player, zero-sum stochastic differential games, however they have limitations of their own. Moreover we discuss other topics of the theory of viscosity solutions like consistency, singular perturbations, etc.. Several special equations are also studied in the book because of their importance and because they are good examples to show how the definition of viscosity solutions and some techniques can be adjusted to particular cases. They are the HJB equations for the optimal control of Duncan-Mortensen-Zakai equation, stochastic Navier-Stokes equations, and stochastic boundary control. In particular the last one also contains ideas on how to handle HJB equations which may be non-degenerate, for instance if Q is not of trace class. Finally we present applications to infinite dimensional Black-Scholes-Barenblatt equations of mathematical finance.

Chapter 4 is devoted to the theory of mild and strong solutions in spaces of continuous functions through fixed point techniques based on the smoothing properties of transition semigroups such as Ornstein-Uhlenbeck ones. This theory applies only to semi-linear equations, i.e. when the coefficient σ does not depend on the control parameter a, and historically it was the first approach introduced in the literature. It was first introduced by Barbu and Da Prato [23] and later improved and developed in various papers, see e.g [63, 64, 79, 81, 226, 231, 232, 235].

Chapter 4 is divided into four main parts. In the first one (Sections 4.2-4.3), we present the basic tools needed for the analysis: the theory of generalized gradients and the smoothing of transition semigroups. In the second one (Sections 4.4 to 4.7), we develop the theory for a general type of semi-linear HJB equation (parabolic and elliptic) without connection with optimal control problems. The main idea behind this approach is the following. Consider the HJB equation (0.1) in the semi-linear autonomous case:

$$\begin{cases} V_t + \mathcal{A}V + \inf_{a\in\Lambda}\Big\{ \langle b(x,a), DV\rangle + l(t,x,a)\Big\} = 0, \\ V(T,x) = g(x), \end{cases} \tag{0.2}$$

where $\mathcal{A}$ is the linear operator

$$\mathcal{A}\varphi = \langle Ax, D\varphi\rangle + \frac{1}{2}\mathrm{Tr}\left[(\sigma(x)Q^{\frac{1}{2}})(\sigma(x)Q^{\frac{1}{2}})^* D^2\varphi\right].$$

If such operator generates a semigroup $e^{t\mathcal{A}}$ then, by the variation of constants formula, one can rewrite (0.2) in the integral form as

$$V(t,x) = e^{(T-t)\mathcal{A}}g(x) + \int_t^T \left(e^{(T-s)\mathcal{A}}F(s,\cdot)\right)(x)\, ds$$

where $F(s,x) := \inf_{a\in\Lambda}\{\langle b(x,a), DV\rangle + l(s,x,a)\}$. The solution of this integral equation is called a mild solution and is obtained by fixed point techniques. To define it, the solution must at least have first order spatial derivative. Thus one needs suitable smoothing properties of the semigroup $e^{t\mathcal{A}}$ (which is the Ornstein-Uhlenbeck semigroup in the simplest case). Since this semigroup is not strongly continuous, apart from very special cases, one needs to use the theory of π-semigroups introduced in [386] or the one of weakly continuous (or $\mathcal{K}$-continuous) semigroups [75, 82, 225].

In the third part (Section 4.8), we develop a connection with stochastic optimal control problems. The fact that mild solutions have first order spatial derivative allows to give a meaning to formulae for optimal feedbacks. However the proofs of verification theorems and optimal feedback formulae cannot be done straightforwardly as one needs to apply Ito's formula in infinite dimensions which requires smooth functions. For this reason (following [231]), one introduces the notion of strong solution of the HJB equation (0.2) as a suitable limit of classical solutions and proves that any mild solution is also a strong solution.

The fourth and the last part of the chapter (Sections 4.9-4.10) deals with some special equations. In Section 4.9 we show how the techniques developed in the previous sections can be adapted to HJB equations and analysis of optimal control problems for stochastic Burgers equation, stochastic Navier-Stokes equations and stochastic reaction diffusion equations. In Section 4.10 we discuss some equations for which explicit representations of the solutions can be found. Such cases are always of interest in applications.

Chapter 5 is devoted to a relatively new and promising theory of mild and strong solutions in spaces of L^2 functions with respect to a suitable measure μ (see [223, 3, 4, 94]). The contents of this chapter are similar to the previous one as the main ideas behind the definition of mild and strong solutions of HJB equations are the same. The difference is in the fact that the reference space is not the space of continuous functions but the space of square integrable ones with respect to the measure μ. The results are similar: existence and uniqueness of solutions of HJB equations through fixed point arguments, verification theorem through approximations, existence of optimal feedbacks. The advantage of this approach is that the results require weaker assumptions on the data, thus enlarging the range of possible applications, including for instance the control of delay equations, however at a cost of weaker statements, for example the first order spatial derivative is now defined in a Sobolev weak sense and is not in general a Fréchet derivative. The main tools used here are the theory of invariant measures for infinite dimensional stochastic differential equations and the properties of transition semigroups in the space of integrable functions with respect to such measures.

Chapter 6 is devoted to a different and in many respects complementary technique of Backward Stochastic Differential Equations (BSDE). It is written by two leading expert in the field. BSDE are Itô type equations in which the initial condition is replaced by a final condition and a new unknown process appears corresponding to a suitable martingale term. In the nonlinear, finite dimensional case BSDE have been introduced in [372] while their direct connection with optimal stochastic control was firstly investigated in [154] and [377]. Since then, the general theory of BSDE has developed considerably, see [56, 58, 152, 289, 325, 371]. Besides stochastic control, applications were given to many fields, for instance to optimal stopping, stochastic differential games, nonlinear partial differential equations and many topics related to mathematical finance. Infinite dimensional BSDE have also been considered, see for instance [97, 211, 252, 264, 373]. The interest for us is that BSDE provide an alternative way to represent the value function of an optimal control problem and consequently to study the corresponding HJB equation and to solve the control problem. It turns out that the most suitable notion of solution for the HJB equation is, in this context, the one of mild solution on spaces of continuous functions but, unlike in Chapter 4, the BSDE method seems particularly adapted to treat degenerate cases in which the transition semigroup has no smoothing properties. The price to pay is that normally we need more regular coefficients and a structural condition (imposing, roughly speaking, that the control acts within the image of the noise). If these requirements are satisfied the BSDE techniques reveal to be very flexible. In particular, in Chapter 6 we will show how they allow to treat both parabolic and elliptic HJB equations (see [55, 212, 265, 336]). Moreover we will present extensions to the case of locally Lipschitz Hamiltonian, quadratically growing with respect to the gradient (see [54, 55, 205]), to the case of HJB equations corresponding to ergodic control problems ([206]) and the case of state equations with noise and control on the boundary ([131, 338]). We will finally describe how the regularity requirement of the coefficients can be partially removed introducing a suitable concept of "generalized gradient", see [213].

It is impossible to cover all aspects of the theory of HJB equations in infinite dimensions and its connections to stochastic optimal control. In particular the theory of integro-PDE is an emerging area which is not presented in the book. We do not discuss first order equations and extensions to Banach spaces. Equations in the space of probability measures is another emerging topic. We have chosen a selection of topics which give a broad overview of the field and enough information so that the readers can start exploring the subject on their own. There are already enough important applications to justify the interest in the subject. The readers should not be restricted to the boundaries drawn by the book. We hope the book will spur the interest and research in the field among theoretical and applied mathematicians, and it will be useful to all kinds of scientists and researchers working in areas related to stochastic control.

CHAPTER 1

Preliminaries on stochastic calculus in infinite dimensions

1.1. Basic probability

We recall basic notions of measure theory and give a short introduction to random variables and the theory of Bochner integral.

1.1.1. Probability spaces, σ-fields.

DEFINITION 1.1 (π-system, σ-field) *Consider a set Ω and denote by $\mathcal{P}(\Omega)$ the power set of Ω.*

(i) *A nonempty class of subsets of Ω, $\mathscr{F} \subset \mathcal{P}(\Omega)$, is called a* π-system *if it is closed under finite intersections.*

(ii) *A class of subsets of Ω, $\mathscr{F} \subset \mathcal{P}(\Omega)$ is called a* σ-field in Ω *if $\Omega \in \mathscr{F}$ and $\mathscr{F}$ is closed under complements and countable unions.*

(iii) *A class of subsets of Ω, $\mathscr{F} \subset \mathcal{P}(\Omega)$ is called a* λ-system *if:*
  - $\Omega \in \mathscr{F}$;
  - *if $A, B \in \mathscr{F}, A \subset B$, then $B \setminus A \in \mathscr{F}$;*
  - *if $A_i \in \mathscr{F}, i = 1, 2, ..., A_i \uparrow A$, then $A \in \mathscr{F}$.*

If $\mathscr{G}$ and $\mathscr{F}$ are two σ-fields in Ω and $\mathscr{G} \subset \mathscr{F}$, we say that $\mathscr{G}$ is a sub-σ-field of $\mathscr{F}$. Given a class $\mathscr{C} \subset \mathcal{P}(\Omega)$, the smallest σ-field containing $\mathscr{C}$ is called the *σ-field generated by $\mathscr{C}$*. It is denoted by $\sigma(\mathscr{C})$. A σ-field $\mathscr{F}$ in Ω is said to be *countably generated* if there exists a countable class of subsets $\mathscr{C} \subset \mathcal{P}(\Omega)$ such that $\sigma(\mathscr{C}) = \mathscr{F}$.

If $\mathscr{C} \subset \mathcal{P}(\Omega)$ and $A \subset \Omega$ we denote $\mathscr{C} \cap A := \{B \cap A : B \in \mathscr{C}\}$. We denote by $\sigma_A(\mathscr{C} \cap A)$ the σ-field of subsets of A generated by $\mathscr{C} \cap A$. It is easy to see that $\sigma_A(\mathscr{C} \cap A) = \sigma(\mathscr{C}) \cap A$ (see for instance [12], page 5).

For $A \subset \Omega$ we denote its *complement* by $A^c := \Omega \setminus A$, and for $A, B \subset \Omega$ we denote their *symmetric difference* by $A \Delta B := (A \setminus B) \cup (B \setminus A)$. We will write $\mathbb{R}^+ = [0, +\infty), \overline{\mathbb{R}}^+ = [0, +\infty) \cup \{+\infty\}, \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$.

THEOREM 1.2 *Let $\mathscr{G}$ be a π-system and $\mathscr{F}$ be a λ-system in some set Ω, such that $\mathscr{G} \subset \mathscr{F}$. Then $\sigma(\mathscr{G}) \subset \mathscr{F}$.*

PROOF. See [281], Theorem 1.1, page 2. □

COROLLARY 1.3 *Let $\mathscr{G}$ be a π-system and $\mathscr{F}$ be the smallest family of subsets of Ω such that:*
  - $\mathscr{G} \subset \mathscr{F}$;
  - *if $A \in \mathscr{F}$ then $A^c \in \mathscr{F}$;*
  - *if $A_i \in \mathscr{F}, A_i \cap A_j = \emptyset$ for $i, j = 1, 2, ..., i \neq j$, then $\cup_{i=1}^{\infty} A_i \in \mathscr{F}$.*

Then $\sigma(\mathscr{G}) = \mathscr{F}$.

PROOF. Since $\sigma(\mathscr{G})$ satisfies the three conditions for $\mathscr{F}$, we obviously have $\mathscr{F} \subset \sigma(\mathscr{G})$. For the opposite inclusion it remains to notice that $\mathscr{F}$ is a λ-system. (For a self-contained proof see also [130], Proposition 1.4, pages 17.) □

DEFINITION 1.4 (Measurable space) *If Ω is a set and $\mathscr{F}$ is a σ-field in Ω, the pair $(\Omega, \mathscr{F})$ is called a* measurable space .

DEFINITION 1.5 (Probability measure, probability space) *Consider a measurable space $(\Omega, \mathscr{F})$. A function $\mu : \mathscr{F} \to [0, +\infty) \cup \{+\infty\}$ is called a* measure *on $(\Omega, \mathscr{F})$ if $\mu(\emptyset) = 0$, and whenever $A_i \in \mathscr{F}, A_i \cap A_j = \emptyset$ for $i, j = 1, 2, ..., i \neq j$, then*

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

The triplet $(\Omega, \mathscr{F}, \mu)$ is called a measure space. *If $\mu(\Omega) < +\infty$ we say that μ is a* bounded measure. *If $\Omega = \bigcup_{n=1}^{\infty} A_n$, where $A_n \in \mathscr{F}, \mu(A_n) < +\infty, n = 1, 2, ...,$ we say that μ is a σ-*finite measure. *If $\mu(\Omega) = 1$ we say that μ is a* probability measure. *We will use the symbol $\mathbb{P}$ to denote probability measures. The triplet $(\Omega, \mathscr{F}, \mathbb{P})$ is called a* probability space.

Thus a probability measure is a σ-additive function $\mathbb{P} : \mathscr{F} \to [0, 1]$ such that $\mathbb{P}(\Omega) = 1$.

Given a measure space $(\Omega, \mathscr{F}, \mu)$, we denote $\mathcal{N} := \{F \subset \Omega \ : \ \exists G \in \mathscr{F}, F \subset G, \mu(G) = 0\}$. The elements of $\mathcal{N}$ are called *μ-null sets*. If $\mathcal{N} \subset \mathscr{F}$, the measure space $(\Omega, \mathscr{F}, \mu)$ is said to be *complete*. The σ field $\overline{\mathscr{F}} := \sigma(\mathscr{F}, \mathcal{N})$ is called the *completion* of $\mathscr{F}$ (with respect to μ). It is easy to see that $\sigma(\mathscr{F}, \mathcal{N}) = \{A \cup B : A \in \mathscr{F}, B \in \mathcal{N}\}$. If $\mathscr{G} \subset \mathscr{F}$ is another σ-field then $\sigma(\mathscr{G}, \mathcal{N})$ is called the *augmentation* of $\mathscr{G}$ by the null sets of $\mathscr{F}$. The augmentation of $\mathscr{G}$ may be different from its completion, as the latter is just the augmentation of $\mathscr{G}$ by the subsets of the sets of measure zero in $\mathscr{G}$. We also have $\sigma(\mathscr{G}, \mathcal{N}) = \{A \subset \Omega : A \Delta B \in \mathcal{N} \text{ for some } B \in \mathscr{G}\}$.

LEMMA 1.6 *Let μ_1, μ_2 be two bounded measures on a measurable space $(\Omega, \mathscr{F})$, and let $\mathscr{G}$ be a π-system in Ω such that $\Omega \in \mathscr{G}$ and $\sigma(\mathscr{G}) = \mathscr{F}$. Then $\mu_1 = \mu_2$ if and only if $\mu_1(A) = \mu_2(A)$ for every $A \in \mathscr{G}$.*

PROOF. See [281], Lemma 1.17, page 9. □

Let $\Omega_t, t \in \mathcal{T}$ be a family of sets. We will denote the Cartesian product of the family Ω_t by $\times_{t \in \mathcal{T}} \Omega_t$. If $\mathcal{T}$ is finite ($\mathcal{T} = \{1, ..., n\}$) or countable ($\mathcal{T} = \mathbb{N}$), we will also write $\Omega_1 \times ... \times \Omega_n$, respectively $\Omega_1 \times \Omega_2 \times ....$ If each Ω_t is a topological space, we endow $\times_{t \in \mathcal{T}} \Omega_t$ with the product topology. If each Ω_t has a σ-field $\mathscr{F}_t$, we define the *product σ-field* $\otimes_{t \in \mathcal{T}} \mathscr{F}_t$ in $\times_{t \in \mathcal{T}} \Omega_t$ as the σ-field generated by the one-dimensional cylinder sets $A_t \times (\times_{s \neq t} \Omega_s)$. If $\mathcal{T} = \{1, ..., n\}$ (respectively, $\mathcal{T} = \mathbb{N}$) we will just write $\otimes_{t \in \mathcal{T}} \mathscr{F}_t = \mathscr{F}_1 \otimes ... \otimes \mathscr{F}_n$ (respectively, $\otimes_{t \in \mathcal{T}} \mathscr{F}_t = \mathscr{F}_1 \otimes \mathscr{F}_2 \otimes ...$).

If S is a topological space, the σ-field generated by the open sets of S is called *Borel σ-field*. It will be denoted by $\mathcal{B}(S)$. If S is a metric space, unless stated otherwise, its default σ-field will always be $\mathcal{B}(S)$. It is not difficult to see that if $S_1, S_2, ...$ are separable metric spaces, then

$$\mathcal{B}(S_1 \times S_2 \times ...) = \mathcal{B}(S_1) \otimes \mathcal{B}(S_2) \otimes ....$$

If (S, ρ) is a metric space, $A \subset S$, and we consider (A, ρ) as a metric space, then $\mathcal{B}(A) = A \cap \mathcal{B}(S)$. A complete separable metric space is called a *Polish space*. Also $\mathcal{B}(\overline{\mathbb{R}}^+) = \sigma(\mathcal{B}(\mathbb{R}^+), \{+\infty\}), \mathcal{B}(\overline{\mathbb{R}}) = \sigma(\mathcal{B}(\mathbb{R}), \{-\infty\}, \{+\infty\})$.

A measurable space $(\Omega, \mathscr{F})$ is called *countably determined* (or $\mathscr{F}$ is called countably determined) if there is a countable set $\mathscr{F}_0 \subset \mathscr{F}$ such that any two probability measures on $(\Omega, \mathscr{F})$ that agree on $\mathscr{F}_0$ must be the same. It follows from Lemma 1.6 that if $\mathscr{F}$ is countably generated then $\mathscr{F}$ is countably determined. If S is a Polish space then $\mathcal{B}(S)$ is countably generated.

If $(\Omega_i, \mathscr{F}_i, \mu_i), i = 1, ..., n$ are measure spaces, their product measure on $(\Omega_1 \times ... \times \Omega_n, \mathscr{F}_1 \otimes ... \otimes \mathscr{F}_n)$ is denoted by $\mu_1 \otimes ... \otimes \mu_2$. It is the measure such that

$$\mu_1 \otimes ... \otimes \mu_2(A_1 \times ... \times A_n) = \mu_1(A_1) \cdot ... \cdot \mu_n(A_n)$$

for all $A_i \in \Omega_i, i = 1, ..., n$.

If S is a metric space, a bounded measure μ on $(S, \mathcal{B}(S))$ is called *regular* if

$$\mu(A) = \sup\{\mu(C) : C \subset A, C \text{ closed}\} = \inf\{\mu(U) : A \subset U, U \text{ open}\} \quad \forall A \in \mathcal{B}(S).$$

Every bounded measure on $(S, \mathcal{B}(S))$ is regular (see [374], Chapter II, Theorem 1.2). A bounded measure μ on $(S, \mathcal{B}(S))$ is called *tight* if for every $\epsilon > 0$ there exists a compact set $K_\epsilon \subset S$ such that $\mu(S \setminus K_\epsilon) < \epsilon$. If S is a Polish space then every bounded measure on $(S, \mathcal{B}(S))$ is tight (see [374], Chapter II, Theorem 3.2).

We refer to [42, 44, 199, 281, 374] for more on the general theory of measure and probability.

1.1.2. Random variables.

DEFINITION 1.7 (Random variable) *A measurable map X between two measurable spaces $(\Omega, \mathscr{F})$ and $(\tilde{\Omega}, \mathscr{G})$ is a called a* random variable*. This means that X is a random variable if $X^{-1}(A) \in \mathscr{F}$ for every $A \in \mathscr{G}$. We write it shortly as $X^{-1}(\mathscr{G}) \subset \mathscr{F}$. Sometimes we will just say that X is $\mathscr{F}/\mathscr{G}$ measurable.*
If $\tilde{\Omega} = \mathbb{R}$ (resp. $\mathbb{R}^+$) and $\mathscr{G}$ is the Borel σ-field $\mathcal{B}(\mathbb{R})$ (resp. $\mathcal{B}(\mathbb{R}^+)$) then X is said to be a real random variable *(resp.* positive random variable*).*
If $\Omega, \tilde{\Omega}$ are topological spaces and $\mathscr{F}, \mathscr{G}$ are the Borel σ-fields then X is said to be Borel measurable.

Given a random variable $X : (\Omega, \mathscr{F}) \to (\tilde{\Omega}, \mathscr{G})$ we denote by $\sigma(X)$ the smallest sub-σ-field of $\mathscr{F}$ that makes X measurable, i.e. $\sigma(X) := X^{-1}(\mathscr{G})$. It is called the *$\sigma$-field generated by X*. Given a set of indices I and a family of random variables $X_i : (\Omega, \mathscr{F}) \to (\tilde{\Omega}, \mathscr{G})$, $i \in I$, the σ-field $\sigma\left(\{X_i\}_{i\in I}\right)$ generated by $\{X_i\}_{i\in I}$ is the smallest sub-σ-field of $\mathscr{F}$ that makes all the functions $X_i : \left(\Omega, \sigma\left(\{X_i\}_{i\in I}\right)\right) \to (\tilde{\Omega}, \mathscr{G})$ measurable, i.e. $\sigma\left(\{X_i\}_{i\in I}\right) = \sigma\left(X_i^{-1}(\mathscr{G}) : i \in I\right)$.

LEMMA 1.8 *Let $(\Omega, \mathscr{F})$ be a measurable space. Then:*

(i) If $(\tilde{\Omega}, \mathscr{G})$ is a measurable space, $X : \Omega \to \tilde{\Omega}$, and $\mathscr{C} \subset \mathscr{G}$ is such that $\sigma(\mathscr{C}) = \mathscr{G}$, then X is $\mathscr{F}/\mathscr{G}$ measurable if and only if $X^{-1}(\mathscr{C}) \subset \mathscr{F}$. Moreover, $\sigma(X) = \sigma(X^{-1}(\mathscr{C}))$.

(ii) If $X_n : \Omega \to \overline{\mathbb{R}}, n = 1, 2, ...$ are random variables, then $\sup_n X_n, \inf_n X_n, \limsup_n X_n, \liminf_n X_n$ are random variables.

(iii) Let $X_n : \Omega \to S, n = 1, 2, ...$ be random variables, where S is a metric space. Then:

- *if S is complete then $\{\omega : X_n(\omega) \text{ converges}\} \in \mathscr{F}$;*
- *if $X_n \to X$ on Ω, then X is a random variable.*

(iv) Let $(\Omega_i, \mathscr{F}_i), i = 1, 2$ be measurable spaces, and $X : \Omega_1 \times \Omega_2 \to \Omega$ be $(\mathscr{F}_1 \otimes \mathscr{F}_2)/\mathscr{F}$ measurable. Then, for every $\omega_1 \in \Omega_1$, $X_{\omega_1}(\cdot) = X(\omega_1, \cdot)$ is $\mathscr{F}_2/\mathscr{F}$ measurable, and, for every $\omega_2 \in \Omega_2$, $X_{\omega_2}(\cdot) = X(\cdot, \omega_2)$ is $\mathscr{F}_1/\mathscr{F}$ measurable.

PROOF. See for instance [281], Lemmas 1.4, 1.9, 1.10, and [407], Theorem 7.5, page 138. □

THEOREM 1.9 *Let $(\Omega, \mathscr{F})$ and $(\tilde{\Omega}, \mathscr{G})$ be two measurable spaces and (S, d) a Polish space (i.e. a complete and separable metric space). Let $X : (\Omega, \mathscr{F}) \to (\tilde{\Omega}, \mathscr{G})$ and $\phi : (\Omega, \mathscr{F}) \to (S, \mathcal{B}(S))$ be two random variables. Then ϕ is measurable as a map from $(\Omega, \sigma(X))$ to $(S, \mathcal{B}(S))$ if and only if there exists a measurable map $\eta : (\tilde{\Omega}, \mathscr{G}) \to (S, \mathcal{B}(S))$ such that $\phi = \eta \circ X$.*

PROOF. See [281], Lemma 1.13, page 7, or [449] Theorem 1.7 page 5. □

We refer to [42, 199, 281, 407] for more on measurability and for the general theory of integration.

DEFINITION 1.10 (Borel isomorphism) *Let* $(\Omega, \mathscr{F})$ *and* $(\tilde{\Omega}, \mathscr{G})$ *be two measurable spaces. A bijection* f *from* Ω *onto* $\tilde{\Omega}$ *is called a* Borel isomorphism *if* f *is* $\mathscr{F}/\mathscr{G}$ *measurable and* f^{-1} *is* $\mathscr{G}/\mathscr{F}$ *measurable. We then say that* $(\Omega, \mathscr{F})$ *and* $(\tilde{\Omega}, \mathscr{G})$ *are Borel isomorphic.*

DEFINITION 1.11 (Standard measurable space) *A measurable space* $(\Omega, \mathscr{F})$ *is called* standard *if it is Borel isomorphic to one of the following spaces:*

(i) $(\{1, .., n\}, \mathcal{B}(\{1, .., n\}))$,
(ii) $(\mathbb{N}, \mathcal{B}(\mathbb{N}))$,
(iii) $(\{0,1\}^{\mathbb{N}}, \mathcal{B}(\{0,1\}^{\mathbb{N}}))$,

where we have the discrete topologies in $\{1, .., n\}$ *and* $\mathbb{N}$*, and the product topology in* $\{0,1\}^{\mathbb{N}}$.

The following theorem collects results that can be found in [374] (Chapter I, Theorems 2.8 and 2.12).

THEOREM 1.12 *If* S *is a Polish space, then* $(S, \mathcal{B}(S))$ *is standard. If a Borel subset of* S *is uncountable, then it is Borel isomorphic to* $\{0,1\}^{\mathbb{N}}$*. Two Borel subsets of* S *are Borel isomorphic if and only if they have the same cardinality. If* $(\Omega, \mathscr{F})$ *is standard and* $A \in \mathscr{F}$*, then* $(A, \mathscr{F} \cap A)$ *is standard.*

In particular we have the following result.

THEOREM 1.13 *If* $(\Omega, \mathscr{F})$ *is standard, then it is Borel isomorphic to a closed subset of* $[0,1]$ *(with its induced Borel sigma field).*

DEFINITION 1.14 (Simple random variable) *Let* $(\Omega, \mathscr{F})$ *be a measurable space, and* (S, d) *be a metric space (endowed with the Borel* σ*-field induced by the distance). A random variable* $X\colon (\Omega, \mathscr{F}) \to (S, \mathcal{B}(S))$ *is called simple (or simple function) if it has a finite number of values.*

LEMMA 1.15 *Let* $f\colon (\Omega, \mathscr{F}) \to S$ *be a measurable function between a measurable space* $(\Omega, \mathscr{F})$ *and a separable metric space* (S, d) *(endowed with the Borel* σ*-field induced by the distance). Then there exists a sequence* $f_n : \Omega \to S$ *of simple,* $\mathscr{F}/\mathcal{B}(S)$ *measurable functions, such that* $d\left(f(\omega), f_n(\omega)\right)$ *is monotonically decreasing to* 0 *for every* $\omega \in \Omega$.

PROOF. See [130], Lemma 1.3, page 16. □

LEMMA 1.16 *Let* S *be a Polish space with metric* d*. Let* $(\Omega, \mathscr{F}, \mathbb{P})$ *be a complete probability space and let* $\mathscr{G}_1, \mathscr{G}_2 \subset \mathscr{F}$ *be two* σ*-fields with the following property: for every* $A \in \mathscr{G}_2$ *there exists* $B \in \mathscr{G}_1$ *such that* $\mathbb{P}(A \Delta B) = 0$*. Let* $f : (\Omega, \mathscr{G}_2) \to (S, \mathcal{B}(S))$ *be a measurable function. Then there exists a function* $g : (\Omega, \mathscr{G}_1) \to (S, \mathcal{B}(S))$ *such that* $f = g$*,* $\mathbb{P}$ *a.e., and simple functions* $g_n : (\Omega, \mathscr{G}_1) \to (S, \mathcal{B}(S))$ *such that* $d(f(\omega), g_n(\omega))$ *monotonically decreases to* 0*,* $\mathbb{P}$*-a.e..*

PROOF. The proof follows the lines of the proof of Lemma 1.25, page 13, in [281].
Step 1: Let us assume first that $f = x\mathbf{1}_A$ (the characteristic function) for some $A \in \mathscr{G}_2$ and $x \in S$. By hypothesis, we can find $B \in \mathscr{G}_1$ s.t. $\mathbb{P}(A\Delta B) = 0$ and then the claim is proved if we choose $g_n \equiv g = x\mathbf{1}_B$. The same argument holds for a simple function f.
Step 2: For the case of a general f, thanks to Lemma 1.15 we can find a sequence

of simple, $\mathscr{G}_2$-measurable functions f_n such that $d(f(\omega), f_n(\omega))$ monotonically decreases to 0. By Step 1, we can find simple, $\mathscr{G}_1$-measurable functions g_n such that $f_n = g_n$, $\mathbb{P}$-a.e. Thus the claim follows taking $g(\omega) := \lim g_n(\omega)$ if the limit exists and $g(\omega) = s$ (for some $s \in S$) otherwise. $\square$

LEMMA 1.17 *Let $(\Omega, \mathscr{F})$ be a measurable space, and $V \subset E$ be two real separable Banach spaces such that the embedding of V into E is continuous. Then:*

(i) *$\mathcal{B}(E) \cap V \subset \mathcal{B}(V)$ and $\mathcal{B}(V) \subset \mathcal{B}(E)$.*
(ii) *If $X : \Omega \to V$ is $\mathscr{F}/\mathcal{B}(V)$ measurable, then it is $\mathscr{F}/\mathcal{B}(E)$ measurable.*
(iii) *If $X : \Omega \to E$ is $\mathscr{F}/\mathcal{B}(E)$ measurable, then $X \cdot \mathbf{1}_{\{X \in V\}}$ is $\mathscr{F}/\mathcal{B}(V)$ measurable.*
(iv) *$X : \Omega \to E$ is $\mathscr{F}/\mathcal{B}(E)$ measurable if and only if for every $f \in E^*$, $f \circ X$ is $\mathscr{F}/\mathcal{B}(\mathbb{R})$ measurable.*

PROOF. The embedding of V into E is continuous, so $\mathcal{B}(E) \cap V \subset \mathcal{B}(V)$. Since the embedding is also one-to-one, it follows from [374], Theorem 3.9, page 21, that $\mathcal{B}(V) \subset \mathcal{B}(E)$, which completes the proof of (i). Parts (ii) and (iii) are direct consequences of (i). $f(\Omega)$ is separable because E is separable, so Part (iv) is a particular case of the Pettis theorem, see [381] Theorem 1.1. $\square$

NOTATION 1.18 If E is a Banach space we denote by $|\cdot|_E$ its norm. Given two Banach spaces E and F, we denote by $\mathcal{L}(E,F)$ the Banach space of all continuous linear operators from E to F. If $E = F$ we will usually write $\mathcal{L}(E)$ instead of $\mathcal{L}(E,F)$. If H is a Hilbert space we denote by $\langle \cdot, \cdot \rangle$ its inner product. We will always identify H with its dual. If V, H are two real separable Hilbert spaces, we denote by $\mathcal{L}_2(V,H)$ the space of Hilbert-Schmidt operators from V to H (see Appendix B.3). The space $\mathcal{L}_2(V,H)$ is a real separable Hilbert space with the inner product $\langle \cdot, \cdot \rangle_2$, see Proposition B.25. ■

LEMMA 1.19 *Let $(\Omega, \mathscr{F})$ be a measurable space and V, H be real separable Hilbert spaces. Suppose that $F : \Omega \to \mathcal{L}_2(V,H)$ is a map such that for every $v \in V$, $F(\cdot)v$ is $\mathscr{F}/\mathcal{B}(H)$ measurable. Then F is $\mathscr{F}/\mathcal{B}(\mathcal{L}_2(V,H))$ measurable.*

PROOF. Since $\mathcal{L}_2(V,H)$ is separable, by Lemma 1.17-(iv) it is enough to show that for every $T \in \mathcal{L}_2(V,H)$

$$\omega \mapsto \langle F(\omega), T \rangle_2 = \sum_{k=1}^{+\infty} \langle F(\omega) e_k, T e_k \rangle$$

is $\mathscr{F}/\mathcal{B}(\mathbb{R})$ measurable, where $\{e_k\}$ is any orthonormal basis of V. But this is clear since for every ω

$$\langle F(\omega), T \rangle_2 = \lim_{n \to +\infty} F_n^T(\omega),$$

where

$$F_n^T(\omega) = \sum_{k=1}^{n} \langle F(\omega) e_k, T e_k \rangle$$

and $F_n^T(\omega)$ is $\mathscr{F}/\mathcal{B}(\mathbb{R})$ measurable because it is a finite sum of functions that are $\mathscr{F}/\mathcal{B}(\mathbb{R})$ measurable. $\square$

Let I be an interval in $\mathbb{R}$, E, F be two real Banach spaces, and let E be separable. If $f : I \times E \to F$ is Borel measurable then for every $t \in I$ the function $f(t, \cdot) : E \to F$ is Borel measurable (by Lemma 1.8-(iv)).

Assume now that, for all $t \in I$ and for some $m \geq 0$, $f(t, \cdot) \in B_m(E, F)$ (the space of Borel measurable functions with polynomial growth m, see Appendix A.2

for the precise definition). It is not true in general that the function

$$I \to B_m(E,F), \qquad t \mapsto f(t,\cdot)$$

is Borel measurable. As a counterexample[1] one can take the function

$$[0,1] \times L^2(\mathbb{R}) \to L^2(\mathbb{R}), \qquad (t,x) \mapsto S_t x,$$

where $(S_t)_{t\geq 0}$ is the semigroup of left translations. Indeed the map

$$[0,1] \to \mathcal{L}(L^2(\mathbb{R})), \qquad t \mapsto S_t$$

is not measurable (see e.g. [130], Section 1.2). Since $\mathcal{L}(L^2(\mathbb{R})) \subseteq B_1(L^2(\mathbb{R}), L^2(\mathbb{R}))$ and the norm in $\mathcal{L}(L^2(\mathbb{R}))$ is equivalent to the one induced by $B_1(L^2(\mathbb{R}), L^2(\mathbb{R}))$, the claim follows in a straightforward way.

On the other hand, we have the following useful result.

LEMMA 1.20 *Let I and Λ be two Polish spaces. Let μ be a measure defined on the the Borel σ-field $\mathcal{B}(I)$ and denote by $\overline{\mathcal{B}(I)}$ the completion of $\mathcal{B}(I)$ with respect to μ. Let $f : I \times \Lambda \to \mathbb{R}$ be Borel measurable and such that for every $t \in I$, $f(t,\cdot)$ is bounded from below (respectively, above). Then the function*

$$\underline{f} : I \to \mathbb{R}, \qquad t \mapsto \inf_{a\in\Lambda} f(t,a) \tag{1.1}$$

(respectively, $\overline{f} : I \to \mathbb{R}$, $t \mapsto \sup_{a\in\Lambda} f(t,a)$) is $\overline{\mathcal{F}}/\mathcal{B}(\mathbb{R})$ measurable[2].

In particular if I is an interval in $\mathbb{R}$, E, F are two real Banach spaces with E separable, if $\rho : I \times E \to F$ is Borel measurable and, for all $t \in I$ and for some $m \geq 0$, $\rho(t,\cdot) \in B_m(E,F)$, then the function

$$\rho_1 : I \to \mathbb{R}, \qquad t \mapsto \|f(t,\cdot)\|_{B_m(E,F)} \tag{1.2}$$

is Lebesgue measurable.

PROOF. The first part is Example 7.4.2 in Volume 2 of [44] (recall that Polish spaces are Souslin spaces, see [44], Definition 6.6.1, and so $I \times \Lambda$ is a Souslin space).

For the second claim, observe that since f is Borel measurable, also the function

$$f : I \times E \to \mathbb{R}, \qquad f(t,x) := \frac{|\rho(t,x)|_F}{(1+|x|_E^2)^{m/2}}$$

is Borel measurable (since it is the product of a continuous function with the composition of a continuous function and a Borel measurable one). The result thus follows from part one. □

DEFINITION 1.21 (Independence) *Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let I be a set of indices, and $\mathscr{C}_i \subset \mathscr{F}$ for all $i \in I$. We say that the families $\mathscr{C}_i, i \in I$, are independent if, for every finite subset J of I and every choice of $A_i \in \mathscr{C}_i$, $(i \in J)$, we have*

$$\mathbb{P}\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} \mathbb{P}(A_i).$$

If $\mathscr{C}_i \subset \mathscr{F}$ is, for all $i \in I$, a π-system (resp. σ-field), the definition above gives in particular the notion of *independent π-systems* (resp. *σ-fields*). Random variables are said to be independent if they generate independent σ-fields. A random variable X is independent of some σ-field $\mathscr{G}$ if $\sigma(X)$ and $\mathscr{G}$ are independent σ-fields.

LEMMA 1.22 *Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\mathscr{C}_i \subset \mathscr{F}$, be a π-system for every $i \in I$. If $\mathscr{C}_i$, $i \in I$ are independent, then $\sigma(\mathscr{C}_i), i \in I$, are independent.*

[1]This example has been suggested to us by Mauro Rosestolato.

[2]Observe that $\underline{f}$ is not always Borel measurable, see [44] Volume 2, Exercice 6.10.42(ii), page 59.

Proof. See [281] Lemma 2.6 page 27. □

1.1.3. Bochner integral. Throughout this section $(\Omega, \mathscr{F}, \mu)$ is a measure space where μ is σ-finite, and E is a separable Banach space with norm $|\cdot|_E$. We endow E with the Borel σ-field $\mathcal{B}(E)$.

Lemma 1.23 *Let $X\colon (\Omega, \mathscr{F}) \to E$ be a random variable. Then the real valued function $|X|_E$ is measurable.*

Proof. See [130] Lemma 1.2 page 16. □

Let $p \geq 1$. We denote by $L^p(\Omega, \mathscr{F}, \mu; E)$ the quotient space of the set

$$\tilde{L}^p(\Omega, \mathscr{F}, \mu; E) := \left\{ X\colon (\Omega, \mathscr{F}) \to (E, \mathcal{B}(E)) \text{ measurable} : \int_\Omega |X(\omega)|_E^p \, d\mu(\omega) < +\infty \right\}$$

with respect to the equivalence relation of equality μ-a.e. $L^p(\Omega, \mathscr{F}, \mu; E)$ is a Banach space when endowed with the norm

$$|X|_{L^p(\Omega, \mathscr{F}, \mu; E)} = \left(\int_\Omega |X(\omega)|_E^p \, d\mu(\omega) \right)^{1/p}$$

(see e.g. [138] Theorem 7.17 page 104). We will often denote the norm by $|X|_{L^p}$ when the context is clear. If H is a separable Hilbert space, then $L^2(\Omega, \mathscr{F}, \mu; H)$ is a Hilbert space as well equipped with the scalar product $\langle X, Y \rangle_{L^2(\Omega, \mathscr{F}, \mu; H)} = \int_\Omega \langle X(\omega), Y(\omega) \rangle_H \, d\mu(\omega)$.

The space $L^\infty(\Omega, \mathscr{F}, \mu; E)$ is the quotient space of the space of bounded $\mathscr{F}/\mathcal{B}(E)$ measurable functions with respect to the relation of being equal a.e.. It is a Banach space equipped with the norm

$$|X|_{L^\infty(\Omega, \mathscr{F}, \mu; E)} = \operatorname*{ess\,sup}_\Omega |X(\omega)|_E \, .$$

We will denote $L^p(\Omega, \mathscr{F}, \mu; \mathbb{R})$ simply by $L^p(\Omega, \mathscr{F}, \mu)$ or $L^p(\Omega, \mu)$ when the σ-algebra $\mathscr{F}$ is clear from the context. In the special case when $\Omega = I$ is an interval with endpoints a and b with $a < b$ (which may be $\pm\infty$), $\mathscr{F}$ is the Borel σ-field of I, and μ is the Lebesgue measure on I, we will simply write $L^p(I; E)$ or $L^p(a, b; E)$ for $L^p(I, \mathscr{F}, \mu; E)$. Finally we denote by $L^p_{\text{loc}}(I; E)$ the set of measurable functions $f\colon I \to E$ such that $\int_K |f(s)|_E^p ds$ is finite for every compact subset K of I.

Lemma 1.24 *If $\mathscr{F}$ is countably generated apart from null sets then $L^p(\Omega, \mathscr{F}, \mu; E)$ is a separable Banach space.*

Proof. See [140] p. 92. □

Definition 1.25 (Bochner integral) *Let $X\colon (\Omega, \mathscr{F}, \mu) \to E$ be a simple random variable $X = \sum_{i=1}^N x_i \mathbf{1}_{A_i}$, where $x_i \in E$, $A_i \in \mathscr{F}, \mu(A_i) < +\infty$. The* Bochner integral *of X is defined as*

$$\int_\Omega X(\omega) d\mu(\omega) := \sum_{i=1}^N x_i \mu(A_i).$$

Let X be in $L^1(\Omega, \mathscr{F}, \mu; E)$. The Bochner integral *of X is defined as*

$$\int_\Omega X(\omega) d\mu(\omega) := \lim_{n \to +\infty} \int_\Omega X_n(\omega) d\mu(\omega),$$

where $X_n\colon (\Omega, \mathscr{F}, \mu) \to E$ are simple random variables such that

$$\lim_{n \to +\infty} \int_\Omega |X(\omega) - X_n(\omega)|_E d\mu(\omega) = 0. \tag{1.3}$$

REMARK 1.26 It follows easily from Lemma 1.15, that for $X \in L^1(\Omega, \mathscr{F}, \mu; E)$, there always exists a sequence of simple random variables $X_n \colon (\Omega, \mathscr{F}, \mu) \to E$ as in Definition 1.25, satisfying (1.3). ■

PROPOSITION 1.27 *Let $X \in L^1(\Omega, \mathscr{F}, \mu; E)$. Then the Bochner integral of X is well defined and does not depend on the choice of the sequence. Moreover*

$$\left| \int_\Omega X(\omega) d\mu(\omega) \right|_E \le \int_\Omega |X(\omega)|_E d\mu(\omega). \tag{1.4}$$

PROOF. See [130] Section 1.1 (in particular inequality (1.6) page 19 and the part below Lemma 1.5). The proof there is done for a probability measure μ, but the general case is identical. □

PROPOSITION 1.28 *Assume that $(\Omega, \mathscr{F}, \mu)$ is a complete measure space, E and F are separable Banach spaces and $A : D(A) \subset E \to F$ is a closed operator (see Definition B.3). If $X \in L^1(\Omega, \mathscr{F}, \mu; E)$ and $X \in D(A)$ a.s., then AX is an F-valued random variable, and X is a $D(A)$-valued random variable, where $D(A)$ is endowed with the graph norm of A (see Definition B.3). If moreover $\int_\Omega |AX(\omega)|_F \, d\mu(\omega) < +\infty$, then*

$$A \int_\Omega X(\omega) d\mu(\omega) = \int_\Omega AX(\omega) d\mu(\omega).$$

PROOF. The facts that X is a $D(A)$-valued random variable and AX is an F-valued random variable follow from Lemma 1.17-(ii). For the last part, see the proof of Proposition 1.6, Chapter 1 of [130]. □

COROLLARY 1.29 *Assume that E and F are separable Banach spaces and $T \colon E \to F$ is a continuous linear operator. If $X \in L^1(\Omega, \mathscr{F}, \mu; E)$, then*

$$T \int_\Omega X(\omega) d\mu(\omega) = \int_\Omega TX(\omega) d\mu(\omega).$$

PROOF. It is a particular case of Proposition 1.28. □

DEFINITION 1.30 (Lebesgue point) *Assume that Ω is a metric space with distance d. Let X be in $L^1(\Omega, \mathscr{F}, \mu; E)$. A point $\bar{\omega} \in \Omega$ is said to be a* Lebesgue point *for X if*

$$\lim_{\epsilon \to 0} \frac{1}{\mu(B_\epsilon^\Omega(\bar{\omega}))} \int_{B_\epsilon^\Omega(\bar{\omega})} |X(\omega) - X(\bar{\omega})|_E \, d\mu(\omega) = 0$$

where $B_\epsilon^\Omega(\bar{\omega}) := \{\omega \in \Omega \ : \ d(\omega, \bar{\omega}) \le \epsilon\}$.

THEOREM 1.31 *Let $(\Omega_1, \mathscr{F}_1)$ and $(\Omega_2, \mathscr{F}_2)$ be two measurable spaces and μ_1 (respectively μ_2) be a σ-finite measure on $(\Omega_1, \mathscr{F}_1)$ (respectively on $(\Omega_2, \mathscr{F}_2)$). Then there exists a unique measure $\mu_1 \otimes \mu_2$ on $\mathscr{F}_1 \otimes \mathscr{F}_2$ such that, for every $A \in \mathscr{F}_1$ and $B \in \mathscr{F}_2$ with finite measure,*

$$(\mu_1 \otimes \mu_2)(A \times B) = \mu_1(A)\mu_2(B).$$

The measure $\mu_1 \otimes \mu_2$ is σ-finite.

PROOF. See Theorem 8.2, page 160 in Chapter VI, Section 8 of [306]. □

THEOREM 1.32 (Fubini Theorem) *Let $(\Omega_1, \mathscr{F}_1)$ and $(\Omega_2, \mathscr{F}_2)$ be two measurable spaces and μ_1 (respectively μ_2) be a σ-finite measure on $(\Omega_1, \mathscr{F}_1)$ (respectively on $(\Omega_2, \mathscr{F}_2)$). Let E be a separable Banach space with norm $|\cdot|_E$.*

(i) *Let X be in $L^1(\Omega_1 \times \Omega_2, \mathscr{F}_1 \otimes \mathscr{F}_2, \mu_1 \otimes \mu_2; E)$. Then, for μ_1-almost every $\omega_1 \in \Omega_1$, the function $X(\omega_1, \cdot)$ is in $L^1(\Omega_2, \mathscr{F}_2, \mu_2; E)$, and the function given by*

$$\omega_1 \mapsto \int_{\Omega_2} X(\omega_1, \omega_2) d\mu_2(\omega_2)$$

for μ_1-almost all ω_1 (and defined arbitrarily for other ω_1) is in $L^1(\Omega_1, \mathscr{F}_1, \mu_1; E)$. Moreover, we have

$$\int_{\Omega_1 \times \Omega_2} X(\omega_1, \omega_2) d(\mu_1 \otimes \mu_2)(\omega_1, \omega_2) = \int_{\Omega_1} \int_{\Omega_2} X(\omega_1, \omega_2) d\mu_1(\omega_1) d\mu_2(\omega_2).$$

(ii) *Let $X : \Omega_1 \times \Omega_2 \to E$ be an $\mathscr{F}_1 \otimes \mathscr{F}_2$-measurable map. Assume that, for μ_1-almost every $\omega_1 \in \Omega_1$, the function $X(\omega_1, \cdot)$ is in $L^1(\Omega_2, \mathscr{F}_2, \mu_2; E)$ and that the map given by*

$$\omega_1 \mapsto \int_{\Omega_2} |X(\omega_1, \omega_2)| d\mu_2(\omega_2)$$

for μ_1-almost all ω_1 (and defined arbitrarily for other ω_1) is in $L^1(\Omega_1, \mathbb{R})$. Then X is in $L^1(\Omega_1 \times \Omega_2, \mathscr{F}_1 \otimes \mathscr{F}_2, \mu_1 \otimes \mu_2; E)$ and part (i) of the theorem applies.

PROOF. See Theorem 8.4, page 162, and Theorem 8.7, page 165 in Chapter VI, Section 8 of [306]. □

THEOREM 1.33 *Let E be a Banach space and μ be a bounded measure on $(E, \mathcal{B}(E))$. Then the set of uniformly continuous and bounded functions $UC_b(E)$ is dense in $L^p(E, \mathcal{B}(E), \mu)$ for $1 \le p < +\infty$.*

PROOF. By Lemma 1.15 and the monotone convergence theorem it is enough to prove that every characteristic function $\mathbf{1}_A$ for some $A \in \mathcal{B}(E)$ can be approximated by functions in $UC_b(E)$. Since μ is regular, for every $\epsilon > 0$ we can find a closed set $C, C \subset A$ and an open set $U, A \subset U$ such that $\mu(U \setminus C) < \epsilon^p$. Moreover, considering sets $U_n = \{x \in U : \mathrm{dist}(x : A) > 1/n\}$ if necessary, we can assume that $\mathrm{dist}(C, U) > 0$. Then the function

$$f_\epsilon(x) := \frac{\mathrm{dist}(x, U)}{\mathrm{dist}(x, A) + \mathrm{dist}(x, U)}$$

belongs to $UC_b(E)$ and $|\mathbf{1}_A - f_\epsilon|_{L^p} < \epsilon$. □

1.1.4. Expectation, covariance and correlation. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and E be a separable Banach space with norm $|\cdot|_E$.

DEFINITION 1.34 (Expectation) *Given X in $L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$, we denote by $\mathbb{E}[X]$ the (Bochner) integral $\int_\Omega X(\omega) d\mathbb{P}(\omega)$. $\mathbb{E}[X]$ is said to be the* expectation *(or the* mean*) of X.*

To define the covariance operator, we recall first that if $x \in E$, $y \in F$, where E, F are Hilbert spaces, the operator $x \otimes y : F \to E$ is defined by

$$(x \otimes y)h = x\langle y, h\rangle_F.$$

DEFINITION 1.35 (Covariance operator, correlation) *Given a real, separable Hilbert space H and $X \in L^2(\Omega, \mathscr{F}, \mathbb{P}; H)$, the* covariance operator *of X is defined by*

$$Cov(X) := \mathbb{E}\Big[(X - \mathbb{E}[X]) \otimes (X - \mathbb{E}[X])\Big].$$

For $X, Y \in L^2(\Omega, \mathscr{F}, \mathbb{P}; H)$, the correlation *of X and Y is the operator defined by*

$$Cor(X,Y) := \mathbb{E}\Big[(X - \mathbb{E}[X]) \otimes (Y - \mathbb{E}[Y])\Big].$$

REMARK 1.36 For $X \in L^2(\Omega, \mathscr{F}, \mathbb{P}; H)$, the operator $Cov(X)$ is positive, symmetric and nuclear (see [130] pages 26). ■

1.1.5. Conditional expectation and conditional probability.

THEOREM 1.37 *Consider a separable Banach space E, a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a sub σ-field $\mathscr{G} \subset \mathscr{F}$. There exists a unique contractive linear operator $\mathbb{E}[\cdot|\mathscr{G}]\colon L^1(\Omega, \mathscr{F}, \mathbb{P}; E) \to L^1(\Omega, \mathscr{G}, \mathbb{P}; E)$ such that*

$$\int_A \mathbb{E}[\xi|\mathscr{G}](\omega) d\mathbb{P}(\omega) = \int_A \xi(\omega) d\mathbb{P}(\omega) \qquad \text{for all } A \in \mathscr{G} \text{ and } \xi \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E).$$

If $E = H$ is a Hilbert space the restriction of $\mathbb{E}[\cdot|\mathscr{G}]$ to $L^2(\Omega, \mathscr{F}, \mathbb{P}; H)$ is the orthogonal projection $L^2(\Omega, \mathscr{F}, \mathbb{P}; H) \to L^2(\Omega, \mathscr{G}, \mathbb{P}; H)$.

PROOF. See [130] Proposition 1.10 page 26 and [355] Proposition V-2-5 pages 102-103. □

DEFINITION 1.38 (Conditional expectation) *Given $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$, the random variable $\mathbb{E}[X|\mathscr{G}] \in L^1(\Omega, \mathscr{G}, \mathbb{P}; E)$, defined by Theorem 1.37, is called the* conditional expectation *of X given $\mathscr{G}$.*

The following proposition collects various properties of conditional expectation (see e.g. [380] Proposition 3.15 page 25, and [446] Section 9.7 page 88 for similar properties for real-valued random variables).

DEFINITION 1.39 *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let E be a separable Banach space. A family $\mathcal{H}$ of integrable random variables $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$ is called* uniformly integrable *if*

$$\lim_{R\to\infty} \sup_{X\in\mathcal{H}} \int_{|X|_E \geq R} |X(\omega)|_E d\mathbb{P}(\omega) = 0.$$

PROPOSITION 1.40 *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let E be a separable Banach space. The conditional expectation has the following properties:*

- (i) *If $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$ is $\mathscr{G}$-measurable, then $\mathbb{E}[X|\mathscr{G}] = X$ $\mathbb{P}$-a.s.*
- (ii) *Given $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$ and two σ-fields $\mathscr{G}_1$ and $\mathscr{G}_2$ such that $\mathscr{G}_1 \subset \mathscr{G}_2 \subset \mathscr{F}$,*

$$\mathbb{E}\Big[\mathbb{E}\big[X|\mathscr{G}_1\big]\big|\mathscr{G}_2\Big] = \mathbb{E}\Big[\mathbb{E}\big[X|\mathscr{G}_2\big]\big|\mathscr{G}_1\Big] = \mathbb{E}\big[X|\mathscr{G}_1\big] \qquad \mathbb{P}\text{-a.s.}$$

- (iii) *Let $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$. If X is independent of $\mathscr{G}$, then $\mathbb{E}[X|\mathscr{G}] = \mathbb{E}[X]$ $\mathbb{P}$-a.s.. Moreover, X is independent of $\mathscr{G}$, if and only if, for any bounded, Borel measurable $f : E \to \mathbb{R}$, $\mathbb{E}[f(X)|\mathscr{G}] = \mathbb{E}f(X)$ $\mathbb{P}$-a.s.*
- (iv) *If X is $\mathscr{G}$-measurable and ζ is a real-valued integrable random variable such that $\zeta X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$, then*

$$\mathbb{E}\Big[\zeta X|\mathscr{G}\Big] = X\mathbb{E}\Big[\zeta|\mathscr{G}\Big] \qquad \mathbb{P}\text{-a.s.}$$

- (v) *If $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$, ζ is an integrable, real-valued, $\mathscr{G}$-measurable random variable such that $\zeta X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$, then*

$$\mathbb{E}\Big[\zeta X|\mathscr{G}\Big] = \zeta\mathbb{E}\Big[X|\mathscr{G}\Big] \qquad \mathbb{P}\text{-a.s.}$$

(vi) *If $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$ and $f\colon \mathbb{R} \to \mathbb{R}$ is a convex function such that $\mathbb{E}\left[|f(|X|_E)|\right] < +\infty$, then*

$$f\left(\left|\mathbb{E}\Big[X|\mathscr{G}\Big]\right|_E\right) \le \mathbb{E}\Big[f\left(|X|_E\right)|\mathscr{G}\Big] \qquad \mathbb{P}\text{-}a.s.$$

(vii) *If $X, X_n \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$ for every $n \in \mathbb{N}$, the family $\{X_n\}_{n\in\mathbb{N}}$ is uniformly integrable and $X_n \xrightarrow{n\to\infty} X$, $\mathbb{P}$-a.s., then*

$$\mathbb{E}\Big[X_n|\mathscr{G}\Big] \xrightarrow{n\to\infty} \mathbb{E}\Big[X|\mathscr{G}\Big] \qquad \mathbb{P}\text{-}a.s.$$

(viii) *Let $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$. Assume that $\{\mathscr{G}_n\}_{n\in\mathbb{N}}$ is an increasing family of σ-fields such that $\mathscr{G} = \sigma\left(\mathscr{G}_n : n \in \mathbb{N}\right)$ is a sub-σ-field of $\mathscr{F}$. Then*

$$\mathbb{E}\Big[X|\mathscr{G}_n\Big] \xrightarrow{n\to\infty} \mathbb{E}\Big[X|\mathscr{G}\Big] \qquad \mathbb{P}\text{-}a.s.$$

(ix) *Let Z be a separable Banach space and let $T \in \mathcal{L}(E; Z)$. Then*

$$\mathbb{E}[TX|\mathscr{G}] = T\mathbb{E}[X|\mathscr{G}] \qquad \mathbb{P}\text{-}a.s.$$

PROPOSITION 1.41 *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. Then:*

(i) *If $X, Y \in L^1(\Omega, \mathscr{F}, \mathbb{P}; \mathbb{R})$ and $X \ge Y$, then*

$$\mathbb{E}[X|\mathscr{G}] \ge \mathbb{E}[Y|\mathscr{G}].$$

(ii) *(Conditional Fatou Lemma) If $X_n \in L^1(\Omega, \mathscr{F}, \mathbb{P}; \mathbb{R})$ and $X_n \ge 0$, then*

$$\mathbb{E}[\lim\inf_{n\to\infty} X_n|\mathscr{G}] \le \lim\inf_{n\to\infty} \mathbb{E}[X_n|\mathscr{G}] \qquad \mathbb{P} - a.s.$$

PROOF. See [446], Section 9.7, page 88. □

PROPOSITION 1.42 *Let $(E_1, \mathscr{E}_1)$ and $(E_2, \mathscr{E}_2)$ be two measurable spaces and $\psi : E_1 \times E_2 \to \mathbb{R}$ be a bounded measurable function. Let X_1, X_2 be two random variables in a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with values in $(E_1, \mathscr{E}_1)$ and $(E_2, \mathscr{E}_2)$ respectively, and let $\mathscr{G} \subset \mathscr{F}$ be a σ-field. If X_1 is $\mathscr{G}-$measurable and X_2 is independent of $\mathscr{G}$, then*

$$\mathbb{E}(\psi(X_1, X_2)|\mathscr{G}) = \widehat{\psi}(X_1), \quad \mathbb{P} \text{ a.s.}, \tag{1.5}$$

where

$$\widehat{\psi}(x_1) = \mathbb{E}(\psi(x_1, X_2)), \qquad x_1 \in E_1. \tag{1.6}$$

PROOF. See Proposition 1.12, p. 28 of [130]. □

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and $\mathscr{G}$ be a sub σ-field of $\mathscr{F}$. The conditional probability of $A \in \mathscr{F}$ given $\mathscr{G}$ is defined by

$$\mathbb{P}(A|\mathscr{G})(\omega) := \mathbb{E}[\mathbf{1}_A|\mathscr{G}](\omega).$$

DEFINITION 1.43 *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and $\mathscr{G}$ be a sub σ-field of $\mathscr{F}$. A function $p\colon \Omega \times \mathscr{F} \to [0, 1]$ is called a* regular conditional probability *given $\mathscr{G}$ if it satisfies the following conditions:*

(i) *for each $\omega \in \Omega$, $p(\omega, \cdot)$ is a probability measure on $(\Omega, \mathscr{F})$;*
(ii) *for each $B \in \mathscr{F}$, the function $p(\cdot, B)$ is $\mathscr{G}$-measurable;*
(iii) *for every $A \in \mathscr{F}$, $\mathbb{P}(A|\mathscr{G})(\omega) = p(\omega, A)$, $\mathbb{P}$-a.s..*

It thus follows that, if $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$, where E is a separable Banach space, then

$$\mathbb{E}[X|\mathscr{G}](\omega) = \int_\Omega X(\omega')p(\omega, d\omega') \quad \mathbb{P} \ a.s..$$

THEOREM 1.44 *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, where $(\Omega, \mathscr{F})$ is a standard measurable space. Then, for every sub σ-field $\mathscr{G} \subset \mathscr{F}$, there exists a regular conditional probability $p(\cdot, \cdot)$ given $\mathscr{G}$. Moreover, if $p'(\cdot, \cdot)$ is another regular conditional probability given $\mathscr{G}$, then there exists a set $N \in \mathscr{G}, \mathbb{P}(N) = 0$ such that, if $\omega \notin N$ then $p(\omega, A) = p'(\omega, A)$ for all $A \in \mathscr{F}$.*

Moreover, if $\mathscr{H}$ is a countably determined sub σ-field of $\mathscr{G}$, then there exists a $\mathbb{P}$-null set $N \in \mathscr{G}$ such that, if $\omega \notin N$ then $p(\omega, A) = \mathbf{1}_A(\omega)$ for every $A \in \mathscr{H}$. In particular, if $(\Omega_1, \mathscr{F}_1)$ is a measurable space, $\mathscr{F}_1$ is countably determined, $\{x\} \in \mathscr{F}_1$ for all $x \in \Omega_1$ and $\xi \colon (\Omega, \mathscr{F}) \to (\Omega_1, \mathscr{F}_1)$ is a $\mathscr{G}/\mathscr{F}_1$-random variable, then $p(\omega, \{\omega' : \xi(\omega) = \xi(\omega')\}) = 1$ for $\mathbb{P}$-a.e. ω.

PROOF. See Theorem 8.1, page 147 in [374], or Theorems 3.1, 3.2, and the corollary following them in [267] (see also [449] Proposition 1.9, page 11). □

NOTATION 1.45 If the regular conditional probability exists we will often write $\mathbb{P}(\cdot|\mathscr{G})(\omega)$ or $\mathbb{P}_\omega$ for $p(\omega, \cdot)$. ■

DEFINITION 1.46 (Law of a random variable) *Given a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, a measurable space $(\Omega_1, \mathscr{F}_1)$, and a random variable $X \colon (\Omega, \mathscr{F}) \to (\Omega_1, \mathscr{F}_1)$, the probability measure on $(\Omega_1, \mathscr{F}_1)$ defined by*

$$\mathcal{L}_{\mathbb{P}}(X)(A) := \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$$

is called the law *(or* distribution*)*[3] *of X. We denote the law of X by $\mathcal{L}_{\mathbb{P}}(X)$.*

PROPOSITION 1.47 (Change of variables) *Given a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, a measurable space $(\Omega_1, \mathscr{F}_1)$, a random variable $X \colon (\Omega, \mathscr{F}) \to (\Omega_1, \mathscr{F}_1)$, and a bounded Borel function $\varphi : \Omega_1 \to \mathbb{R}$ we have*

$$\int_\Omega \varphi(X(\omega)) d\mathbb{P}(\omega) = \int_{\Omega_1} \varphi(\omega') d\mathcal{L}_{\mathbb{P}}(X)(\omega').$$

DEFINITION 1.48 (Convergence of random variables) *Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a Polish space (S, d) endowed with the Borel σ-field. Let $X_n \colon \Omega \to S$ and $X \colon \Omega \to S$ be random variables. We say that:*

(i) *X_n converges to X* $\mathbb{P}$-a.s. *(and we write $X_n \to X$ $\mathbb{P}$-a.s.) if $\lim_{n\to\infty} d(X_n(\omega), X(\omega)) = 0$ $\mathbb{P}$-a.s..*
(ii) *X_n converges to X* in probability *if, for every $\epsilon > 0$, $\lim_{n\to+\infty} \mathbb{P}\{\omega \in \Omega : d(X_n(\omega), X(\omega)) > \epsilon\} = 0$.*
(iii) *X_n converges to X* in law *if, for every bounded and continuous $f \colon S \to \mathbb{R}$, $\int_S f(u) d\mathcal{L}_{\mathbb{P}}(X)(u) = \lim_{n\to\infty} \int_S f(u) d\mathcal{L}_{\mathbb{P}}(X_n)(u)$ (i.e. if $\mathbb{E}[f(X)] = \lim_{n\to\infty} \mathbb{E}[f(X_n)]$).*

LEMMA 1.49 *Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a Polish space (S, d) endowed with the Borel σ-field. Let $X_n \colon \Omega \to S$ and $X \colon \Omega \to S$ be random variables.*

(i) *If X_n converges to X* $\mathbb{P}$-a.s. *then X_n converges to X* in probability.
(ii) *If X_n converges to X* in probability *then X_n converges to X* in law.
(iii) *If X_n converges to X* in probability *then it contains a subsequence X_{n_k} such that X_{n_k} converges to X* $\mathbb{P}$-a.s..
(iv) *(*Egoroff's theorem*) If X_n converges to X* $\mathbb{P}$-a.s. *then for every $\epsilon > 0$, there exists $\tilde{\Omega} \in \mathscr{F}$ such that $\mathbb{P}(\Omega \setminus \tilde{\Omega}) < \epsilon$, and X_n converges uniformly to X on $\tilde{\Omega}$.*
(v) *Let $X, X_n \in L^p(\Omega, \mathscr{F}, \mathbb{P}; E), n \in \mathbb{N}, p \geq 1$, and E be a separable Banach space. If X_n converges to X in $L^p(\Omega, \mathscr{F}, \mathbb{P}; E)$, then X_n converges to X* in probability.

[3] In measure theory it is more often called the *push-forward of* $\mathbb{P}$ and denoted by $X_{\#}\mathbb{P}$.

PROOF. For (i), (ii), and (iii) see for instance [281] Lemma 4.2, page 63 and Lemma 4.7, page 66. Part (iv) can be found for instance in [52] Theorem 2, page 170, Section IV.5.4. Property (v) is straightforward. $\square$

LEMMA 1.50 *Let $p > 1$ and $X, X_n \in L^p(\Omega, \mathscr{F}, \mathbb{P}; E), n \in \mathbb{N}$, for some separable Banach space E. Suppose that, for some $M > 0$, $\mathbb{E}\left[|X_n|_E^p\right] \leq M$ for all $n \in \mathbb{N}$. If $X_n \to X$ in probability, then $\mathbb{E}\left[|X - X_n|_E\right] \to 0$.*

PROOF. Since the sequence $\{X_n\}$ is bounded in $L^p(\Omega, \mathscr{F}, \mathbb{P}; E)$, it is uniformly integrable (see e.g. [446] p.127, Section 13.3). The claim follows e.g. from Theorem 13.7, p.131 of [446]. $\square$

1.1.6. Gaussian measures on Hilbert spaces and Fourier transform. In this section we recall the notions of Gaussian measure and Fourier transform for Hilbert-space valued random variables. For an extensive treatment of the subject we refer to [130], Chapter 2, [110], Chapter 1 or [111], Chapter 1.

For a real separable Hilbert space H we denote by $\mathcal{L}_1(H)$ the Banach space of the trace class operators on H, by $\mathcal{L}^+(H)$ the subspace (of $\mathcal{L}(H)$) of all bounded, linear, self-adjoint, positive operators, and we set $\mathcal{L}_1^+(H) := \mathcal{L}_1(H) \cap \mathcal{L}^+(H)$ (see Appendix B.3). We will denote by $M_1(H)$ the set of probability measures on $(H, \mathcal{B}(H))$.

PROPOSITION 1.51 *Consider a real, separable Hilbert space H with the Borel σ-field $\mathcal{B}(H)$ and a probability measure $\mathbb{P}$ on $(H, \mathcal{B}(H))$. If $\int_H |y|\, d\mathbb{P}(y) < +\infty$, then we can define*

$$m := \int_H y\, d\mathbb{P}(y) \in H.$$

If $\int_H |y|^2 d\mathbb{P}(y) < +\infty$, then there exists a unique $Q \in \mathcal{L}_1^+(H)$ such that

$$\langle Qx, y\rangle := \int_H \langle x, h - m\rangle \langle y, h - m\rangle\, d\mathbb{P}(h).$$

PROOF. See [110] Page 7. $\square$

DEFINITION 1.52 (Mean and covariance of a measure on H) *We call m and Q, defined by Proposition 1.51, respectively the* mean *and the* covariance *of $\mathbb{P}$. In other words the mean (respectively covariance) of $\mathbb{P}$ is the mean (respectively covariance) of the identity random variable $I\colon (H, \mathcal{B}(H), \mathbb{P}) \to (H, \mathcal{B}(H))$.*

DEFINITION 1.53 (Fourier transform of a measure) *Let H be a Hilbert space and $\mathcal{B}(H)$ be its Borel σ-field. Given a probability measure $\mathbb{P}$ on $(H, \mathcal{B}(H))$ we define, for $x \in H$,*

$$\hat{\mathbb{P}}(x) := \int_H e^{i\langle y, x\rangle} d\mathbb{P}(y).$$

We call $\hat{\mathbb{P}}\colon H \to \mathbb{C}$ the Fourier transform *of $\mathbb{P}$.*

PROPOSITION 1.54 *Let H be a real, separable Hilbert space, $\mathcal{B}(H)$ be its Borel σ-field, and $\mathbb{P}_1$ and $\mathbb{P}_2$ be two probability measures on $(H, \mathcal{B}(H))$. If $\hat{\mathbb{P}}_1(x) = \hat{\mathbb{P}}_2(x)$ for all $x \in H$, then $\mathbb{P}_1 = \mathbb{P}_2$.*

PROOF. See [110] Proposition 1.7, page 6 or [130], Proposition 2.5, page 35. $\square$

THEOREM 1.55 *Let $X_1, ..., X_n$ be random variables in a real, separable Hilbert space H. The random variables are independent if and only if for every $y_1, ..., y_n \in$*

H

$$\mathbb{E}\left[e^{\left[i\sum_{i=1}^{n}\langle X_i,y_i\rangle\right]}\right]=\prod_{i=1}^{n}\mathbb{E}\left[e^{\left[i\langle X_i,y_i\rangle\right]}\right]. \tag{1.7}$$

PROOF. Obviously if $X_1,...,X_n$ are independent then (1.7) holds. Also Theorem 1.55 is well known if $H=\mathbb{R}^k$. Let now $k\in\mathbb{N}$ and $y_i^j\in H, i=1,...,n, j=1,...,k$, and consider random variables $X_i^k=(\langle X_i,y_i^1\rangle,...,\langle X_i,y_i^k\rangle), i=1,...,n$ in $\mathbb{R}^k$. Therefore, if (1.7) holds then $X_i^k, i=1,...,n$, are independent for every $k\in\mathbb{N}$ and $y_i^j\in H, j=1,...,k$. Since cylindrical sets of the form $\{x:(\langle x,y_i^1\rangle,...,\langle x,y_i^k\rangle)\in A\in\mathcal{B}(\mathbb{R}^k)\}$ generate $\mathcal{B}(H)$ and are a π-system, the collection of sets $\{\omega:(\langle X_i,y_i^1\rangle,...,\langle X_i,y_i^k\rangle)\in A\in\mathcal{B}(\mathbb{R}^k)\}$ over all $k\in\mathbb{N}$ and $y_i^j\in H, i=1,...,n, j=1,...,k, A\in\mathcal{B}(\mathbb{R}^k)$ is a π-system generating $\sigma(X_i)$. Thus, by Lemma 1.22, the sigma algebras $\sigma(X_1),...,\sigma(X_n)$ are independent. □

THEOREM 1.56 *Let H be a real, separable Hilbert space, $\mathcal{B}(H)$ be its Borel σ-field, $a\in H$, and $Q\in\mathcal{L}_1^+(H)$. Then there exists a unique probability measure $\mathbb{P}$ on $(H,\mathcal{B}(H))$ such that*

$$\hat{\mathbb{P}}(x)=e^{i\langle a,x\rangle-\frac{1}{2}\langle Qx,x\rangle}.$$

The measure $\mathbb{P}$ has mean a and covariance Q.

PROOF. See [110] Theorem 1.12 page 12. □

DEFINITION 1.57 (Gaussian measure on H) *Let H be a real, separable Hilbert space, $\mathcal{B}(H)$ be its Borel σ-field, $a\in H$, and $Q\in\mathcal{L}_1^+(H)$. The unique probability measure $\mathbb{P}$ identified by Theorem 1.56 is called the* Gaussian measure *with mean a and covariance Q, and is denoted by $\mathcal{N}(a,Q)$. When $a=0$ we will denote it by $\mathcal{N}_Q$ and call it a centered Gaussian measure.*

PROPOSITION 1.58 *Let $Q\in\mathcal{L}_1^+(H)$. Then for all $y,z\in H$*

$$\int_H \langle x,y\rangle\langle x,z\rangle\mathcal{N}_Q(dx)=\langle Qy,z\rangle \tag{1.8}$$

Define, for $y\in Q^{1/2}(H)$, $\mathcal{Q}_y\in L^2(H,\mathcal{N}_Q)$ as

$$\mathcal{Q}_y(x):=\left\langle Q^{-1/2}y,x\right\rangle, \tag{1.9}$$

where $Q^{-1/2}$ is the pseudoinverse of $Q^{1/2}$ (see Definition B.1). The map (called the "white noise function", see e.g. [111]*[Section 2.5])*

$$y\in Q^{1/2}(H)\to\mathcal{Q}_y\in L^2(H,\mathcal{N}_Q)$$

can be extended to $H_0=\overline{Q^{1/2}(H)}=(\ker Q)^\perp$ and it satisfies

$$\int_H \mathcal{Q}_y(x)\mathcal{Q}_z(x)\mathcal{N}_Q(dx)=\langle y,z\rangle, \qquad y,z\in H_0.$$

Moreover for all $m>0$ we have

$$\int_H |x|^{2m}\mathcal{N}_Q(dx)\le K(m)[\mathrm{Tr}(Q)]^m$$

for some $K(m)>0$, independent of Q.

PROOF. Formula (1.8) follows from [129][Proposition 1.2.4].

The second statement is proved, when $\ker Q=\{0\}$, in [111][Section 2.5.2] (see also Section 1.2.4 of [129]). Since here we do not assume $\ker Q=\{0\}$, we provide a proof. First we observe that $\ker Q=\ker Q^{1/2}$ and that $Q^{1/2}(H)$ is dense in $(\ker Q)^\perp$ since $Q^{1/2}$ is selfadjoint. Moreover by Definition B.1, the pseudoinverse of $Q^{1/2}$ is the operator $Q^{-1/2}:Q^{1/2}(H)\to(\ker Q)^\perp$, hence the map $y\to\mathcal{Q}_y=\left\langle Q^{-1/2}y,x\right\rangle$

is well defined for all $y \in Q^{1/2}(H)$. Furthermore, thanks to formula (1.8), we have, for $y_1, y_2 \in Q^{1/2}(H)$

$$\int_H \left\langle Q^{-1/2}y_1, x\right\rangle \left\langle Q^{-1/2}y_2, x\right\rangle \mathcal{N}_Q(dx) = \left\langle Q(Q^{-1/2}y_1), Q^{-1/2}y_2\right\rangle = \langle y_1, y_2\rangle ,$$

where we used that $Q^{1/2}Q^{-1/2}y = y$ for all $y \in Q^{1/2}(H)$. Hence, for $y_1, y_2 \in Q^{1/2}(H)$,

$$\int_H \mathcal{Q}_{y_1}(x)\mathcal{Q}_{y_2}(x)\mathcal{N}_Q(dx) = \langle y_1, y_2\rangle . \tag{1.10}$$

In view of the above the map $y \to \mathcal{Q}_y = \left\langle Q^{-1/2}y, x\right\rangle$ is an isometry and can be extended to $\overline{Q^{1/2}(H)} = (\ker Q)^\perp$ (endowed with the inner product inherited from H) and (1.10) extends to all $y_1, y_2 \in (\ker Q)^\perp$.

We remark that as pointed out in [111][Section 2.5.2], for a generic $y \in (\ker Q)^\perp$ the image $\mathcal{Q}_y$ is an element of $L^2(H, \mathcal{N}_Q)$, hence an equivalence class of random variables defined $\mathcal{N}_Q$-a.e.; in particular, writing $\mathcal{Q}_y(x) = \left\langle y, Q^{-1/2}x\right\rangle$, $\mathcal{N}_Q$-a.e., would be misleading since, as proved in [111][Proposition 2.22], $\mathcal{N}_Q(Q^{1/2}(H)) = 0$.

Concerning the third claim by Proposition 2.19 page 50 of [130], it holds for $m \in \mathbb{N}$. If $k-1 < m < k$ for $k = 1, 2, ...,$ we use

$$\int_H |x|^{2m}\mathcal{N}_Q(dx) \le \left[\int_H |x|^{2k}\mathcal{N}_Q(dx)\right]^{m/k} .$$

□

We recall a useful result concerning compactness of a family of measures in $M_1(H)$ (see e.g. [130][Section 2.1], [161, 374] for more on this).

Definition 1.59

(i) *A sequence $\{\mathbb{P}_n\}$ in $M_1(H)$ is said to be weakly convergent to some $\mathbb{P} \in M_1(H)$ if, for every $\phi \in C_b(H)$,*

$$\lim_{n\to+\infty} \int_H \phi(x)\mathbb{P}_n(dx) = \int_H \phi(x)\mathbb{P}(dx).$$

(ii) *A family $\Lambda \subseteq M_1(H)$ is said to be compact (respectively, relatively compact) if an arbitrary sequence $\{\mathbb{P}_n\}$ of elements from Λ contains a subsequence $\{\mathbb{P}_{n_k}\}$ weakly convergent to a measure $\mathbb{P} \in \Lambda$ (respectively, to a measure $\mathbb{P} \in M_1(H)$).*

(iii) *A family $\Lambda \subseteq M_1(H)$ is said to be tight if for any $\varepsilon > 0$ there exists a compact set K_ε such that, for every $\mathbb{P} \in \Lambda$,*

$$\mathbb{P}(K_\varepsilon) > 1 - \varepsilon.$$

The following theorem (which also holds when H is a Polish space) is due to Prokhorov.

Theorem 1.60 *Let H be a real separable Hilbert space. A family $\Lambda \subseteq M_1(H)$ is relatively compact if and only if it is tight.*

Proof. See [130], the proof of Theorem 2.3. □

The next theorem gives a useful sufficient condition for compactness.

Theorem 1.61 *Let H be a real separable Hilbert space and let $\{e_i\}_i$ be an orthonormal basis in H. A family $\Lambda \subseteq M_1(H)$ is relatively compact if*

$$\lim_{N\to+\infty} \sup_{\mathbb{P}\in\Lambda} \int_H \sum_{i=N}^{+\infty} \langle x, e_i\rangle^2 \,\mathbb{P}(dx) = 0.$$

PROOF. See [374], the proof of Theorem VI.2.2. □

Concerning Gaussian measures, we have the following result (see Proposition 1.1.5 of [387].

PROPOSITION 1.62 *Let $\mathcal{N}_{Q_n}$ ($n \geq 1$), and $\mathcal{N}_Q$ be centered Gaussian measures on H. If $\lim_{n\to+\infty} \|Q_n - Q\|_{\mathcal{L}_1(H)} = 0$, then the measures $\mathcal{N}_{Q_n}$ converge weakly to $\mathcal{N}_Q$.*

PROOF. Observe that if $\{e_i\}_i$ is an orthonormal basis in H, it follows from (1.8) that for any $N \in \mathbb{N}$,

$$\int_H \sum_{i=N}^{+\infty} \langle x, e_i\rangle^2 \mathcal{N}_{Q_n}(dx) = \sum_{i=N}^{+\infty} \langle Q_n e_i, e_i\rangle .$$

Since $\lim_{n\to+\infty} \|Q_n - Q\|_{\mathcal{L}_1(H)} = 0$, the above formula implies in particular that Theorem 1.61 applies and thus the sequence $\{\mathcal{N}_{Q_n}\}$ is relatively compact.

Moreover, from Theorem 1.56 and Definition 1.57 it is immediate that, as $n \to +\infty$,

$$\widehat{\mathcal{N}_{Q_n}}(x) = e^{-\frac{1}{2}\langle Q_n x, x\rangle} \quad \longrightarrow \quad e^{-\frac{1}{2}\langle Qx,x\rangle} = \widehat{\mathcal{N}_Q}(x), \qquad \forall x \in H.$$

Take now a subsequence $\mathcal{N}_{Q_{n_k}}$ weakly convergent to a probability measure $\mathbb{P}_0$. By Definition 1.53 we must have

$$\widehat{\mathcal{N}_{Q_n}}(x) \to \widehat{\mathbb{P}_0}(x), \qquad \forall x \in H.$$

This implies that $\widehat{\mathbb{P}_0} = \widehat{\mathcal{N}_Q}$ and hence, by Proposition 1.54, that $\mathbb{P}_0 = \mathcal{N}_Q$. Since this is true for any convergent subsequence, the claim now follows by a standard contradiction argument. □

1.2. Stochastic processes and Brownian motion

1.2.1. Stochastic processes.

DEFINITION 1.63 (Filtration, usual conditions) *Let $t \geq 0$. A filtration $\{\mathscr{F}_s^t\}_{s\geq t}$ in a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is a family of σ-fields such that $\mathscr{F}_s^t \subset \mathscr{F}_r^t \subset \mathscr{F}$ whenever $t \leq s \leq r$.*

(i) *We say that $\{\mathscr{F}_s^t\}_{s\geq t}$ is* right continuous *if, for all $s \geq t$, $\mathscr{F}_{s+}^t := \bigcap_{r>s} \mathscr{F}_r^t = \mathscr{F}_s^t$.*
(ii) *We say that $\{\mathscr{F}_s^t\}_{s\geq t}$ is* left continuous *if, for all $s > t$, $\mathscr{F}_{s-}^t := \sigma\left(\bigcup_{r<s} \mathscr{F}_r^t\right) = \mathscr{F}_s^t$. We say that $\{\mathscr{F}_s^t\}_{s\geq t}$ is* continuous *if it is both left and right continuous.*
(iii) *We say that $\{\mathscr{F}_s^t\}_{s\geq t}$ satisfies the* usual conditions *if it is right continuous and complete, i.e. if $\mathscr{F}_s^t$ contains all $\mathbb{P}$-null sets of $\mathscr{F}$ for every $s \geq t$.*

We will often write $\mathscr{F}_s^t$ instead of $\{\mathscr{F}_s^t\}_{s\geq t}$. We also set $\mathscr{F}_{+\infty}^t := \sigma\left(\bigcup_{r<+\infty} \mathscr{F}_r^t\right)$.

Since we will mostly deal with filtrations satisfying the usual conditions we will assume from now on that this property holds unless explicitly stated otherwise. For this reason we include the usual conditions in the definition of a filtered probability space.

DEFINITION 1.64 (Filtered probability space) *Let $\mathscr{F}_s^t$ be a filtration satisfying the usual conditions on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$. The 4-tuple $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P})$ is called a* filtered probability space.

NOTATION 1.65 We use the following convention in this section. When we write $s \in [t,T]$ we mean that $s \in [t,T]$ if $T \in \mathbb{R}$, and $s \in [t,+\infty)$ if $T = +\infty$. So $[t,T]$ is understood to be $[t,+\infty)$ if $T = +\infty$. ■

DEFINITION 1.66 (Stochastic process) *Let $T \in (0,+\infty]$, $t \in [0,T)$ and $(\Omega, \mathscr{F})$ and $(\Omega_1, \mathscr{F}_1)$ be two measurable spaces. A family of random variables $X(\cdot) = \{X(s)\}_{s\in[t,T]}$, $X(s)\colon \Omega \to \Omega_1$, is called a* stochastic process *in $[t,T]$. If $(\Omega_1, \mathscr{F}_1) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ then $X(\cdot)$ is called a real stochastic process.*

DEFINITION 1.67 *Let $\left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s\geq t}, \mathbb{P}\right)$ be a filtered probability space and $(\Omega_1, \mathscr{F}_1)$ be a measurable space. A stochastic process $\{X(s)\}_{s\in[t,T]}\colon [t,T]\times\Omega \to \Omega_1$ is said to be:*

(i) Measurable, *if the map $(s,\omega) \mapsto X(s)(\omega)$ is $\mathcal{B}([t,T]) \otimes \mathscr{F}/\mathscr{F}_1$ measurable.*

(ii) Adapted, *if, for each $s \in [t,T]$, $X(s)\colon \Omega \to \Omega_1$ is an $\mathscr{F}_s^t/\mathscr{F}'$-measurable random variable.*

(iii) Progressively measurable, *if for all $s \in (t,T]$, the restriction of $X(\cdot)$ to $[t,s]\times\Omega$ is $\mathcal{B}([t,s]) \otimes \mathscr{F}_s^t/\mathscr{F}_1$ measurable.*

(iv) Predictable, *if the map $(s,\omega) \mapsto X(s)(\omega)$ is $\mathcal{P}_{[t,T]}/\mathscr{F}_1$ measurable, where $\mathcal{P}_{[t,T]}$ is the σ-field in $[t,T]\times\Omega$ generated by all sets of the form $(s,r]\times A, t \leq s < r \leq T, A \in \mathscr{F}_s^t$ and $\{t\}\times A, A \in \mathscr{F}_t^t$.*

(v) *If E is a separable Banach space (endowed with its Borel σ-field), the process $\{X(s)\}_{s\in[t,T]}\colon [t,T]\times\Omega \to E$ is called* stochastically continuous *at $s \in [t,T]$, if for every $\epsilon, \delta > 0$ there exists $\rho > 0$ such that*

$$\mathbb{P}\left(|X(r) - X(s)| \geq \epsilon\right) \leq \delta, \qquad \textit{for all } r \in (s-\rho, s+\rho)\cap[t,T].$$

(vi) *If (S,d) is a metric space (endowed with its Borel σ-field), the process $\{X(s)\}_{s\in[t,T]}\colon [t,T]\times\Omega \to S$ is called* continuous *(respectively, right-continuous, left-continuous), if for $\mathbb{P}$-a.e. $\omega \in \Omega$, the function $s \mapsto X(s)(\omega)$ is continuous (respectively, right-continuous, left-continuous).*

(vii) *If E is a separable Banach space (endowed with its Borel σ-field), the process $\{X(s)\}_{s\in[t,T]}\colon [t,T]\times\Omega \to E$ is called* integrable, *if $\mathbb{E}[|X(s)|] < +\infty$ for all $s \in [t,T]$. The process is called* uniformly integrable *if it is integrable and the family $\{X(s)\}_{s\in[t,T]}$ is uniformly integrable (see Definition 1.39).*

(viii) *If E is a separable Banach space (endowed with the Borel σ-field induced by the norm), the process $\{X(s)\}_{s\in[t,T]}\colon [t,T]\times\Omega \to E$ is said to be* mean square continuous *if $\mathbb{E}[|X(s)|^2] < +\infty$ for all $s \in [t,T]$ and $\lim_{r\to s}\mathbb{E}[|X(r)-X(s)|^2] = 0$ for all $s \in [t,T]$.*

The concepts of adapted, progressively measurable, and predictable processes can be defined for any filtration $\mathscr{G}_s^t$. In such cases, to emphasize the filtration used, we will refer to the processes as $\mathscr{G}_s^t$-adapted, $\mathscr{G}_s^t$-progressively measurable, and $\mathscr{G}_s^t$-predictable.

Progressive measurability can also be defined using the concept of progressively measurable sets, see e.g. [344], p.4, or [161], page 71. We say that a set $A \subset [t,T]\times\Omega$ is $\mathscr{F}_s^t$-progressively measurable if the function $\mathbf{1}_A$ is a progressively measurable process. Equivalently this means that $A \cap ([t,s]\times\Omega) \in \mathcal{B}([t,s]) \otimes \mathscr{F}_s^t$ for every $s \in [t,T]$. It can be proved that the $\mathscr{F}_s^t$-progressively measurable sets form a σ-field and that a process $X(s)$ is progressively measurable if and only if it is measurable with respect to the σ-field of $\mathscr{F}_s^t$-progressively measurable sets.

DEFINITION 1.68 (Stochastic equivalence, modification) *Let* $(\Omega, \mathscr{F}, \mathbb{P})$ *be a probability space, and* $(\Omega_1, \mathscr{F}_1)$ *be a measurable space. Processes* $X(\cdot), Y(\cdot)\colon [t,T] \times \Omega \to \Omega_1$ *are called* stochastically equivalent, *if for all* $s \in [t,T]$, $\mathbb{P}(X(s) = Y(s)) = 1$. *In this case* $Y(\cdot)$ *is said to be a* modification *of* $X(\cdot)$. *The processes* $X(\cdot)$ *and* $Y(\cdot)$ *are called* indistinguishable, *if* $\mathbb{P}(X(s) = Y(s) : \forall s \in [t,T]) = 1$.

LEMMA 1.69 *Let* $\left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s\geq t}, \mathbb{P}\right)$ *be a filtered probability space and let* $\{X(s)\}_{s\geq t}$ *be a process with values in a Polish space* (S,d), *endowed with the Borel* σ*-field induced by the distance.*

- (i) *If* $X(\cdot)$ *is* $\mathcal{B}([t,T]) \otimes \mathscr{F}/\mathcal{B}(S)$ *measurable and* $\mathscr{F}_s^t$*-adapted, then* $X(\cdot)$ *has an* $\mathscr{F}_s^t$*-progressively measurable modification.*
- (ii) *If* $X(\cdot)$ *is* $\mathscr{F}_s^t$*-adapted and* $X(\cdot)$ *is left- (or right-) continuous for every* ω, *then* $X(\cdot)$ *itself is* $\mathscr{F}_s^t$*-progressively measurable.*

PROOF. *Part (i)*: Since S is Borel isomorphic to a Borel subset A of $\mathbb{R}$, without loss of generality we can consider $X(\cdot)$ to be an $\mathbb{R}$-valued process with values in A. By [346], Theorem T46, page 68, $X(\cdot)$ has an $\mathbb{R}$-valued, $\mathscr{F}_s^t$-progressively measurable modification $\tilde{X}(\cdot)$. Let $a \in A$. We define a process $Y(\cdot)$ by $Y(s) := \tilde{X}(s)\mathbf{1}_{\tilde{X}(s)\in A} + a\mathbf{1}_{\tilde{X}(s)\in(\mathbb{R}\setminus A)}$. The process $Y(\cdot)$ is $\mathscr{F}_s^t$-progressively measurable. Moreover, if $\tilde{X}(s) = X(s)$, then $Y(s) = X(s)$, so $Y(\cdot)$ is a modification of $X(\cdot)$. *Part (ii)*: See [346], Theorem T47, page 70, or [283], Proposition 1.13, page 5. □

1.2.2. Martingales.

NOTATION 1.70 Unless specified otherwise, any Banach space E and any metric space (S,d) will be understood to be endowed with the Borel σ-field induced respectively by the norm and by the distance. ■

DEFINITION 1.71 (Martingale) *Let* $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P})$ *be a filtered probability space, and let* $M(\cdot)$ *be an* $\mathscr{F}_s^t$*-adapted and integrable process with values in a separable Banach space* E. *Then* $M(\cdot)$ *is said to be a* martingale *if, for all* $r, s \in [t,T], s \leq r$,

$$\mathbb{E}\left[M(r)|\mathscr{F}_s^t\right] = M(s) \qquad \mathbb{P}-a.s..$$

If $E = \mathbb{R}$, *we say that* $M(s)$ *is a* submartingale *(respectively,* supermartingale*), if*

$$\mathbb{E}\left[M(r)|\mathscr{F}_s^t\right] \geq M(s), \quad (\textit{respectively, } \mathbb{E}\left[M(r)|\mathscr{F}_s^t\right] \leq M(s))\ \mathbb{P}-a.s..$$

If $\mathbb{E}|M(s)|^2 < +\infty$ for all $s \geq t$ we say that $M(\cdot)$ is *square integrable.*

THEOREM 1.72 (Doob's maximal inequalities) *Let* $T > 0$, $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P})$ *be a filtered probability space, and* H *be a separable Hilbert space. Let* $M(\cdot)$ *be a right-continuous* H*-valued martingale such that* $M(s) \in L^p(\Omega, \mathscr{F}, \mathbb{P}; H)$ *for all* $s \in [t,T]$. *Then:*

- (i) *If* $p \geq 1$, $\mathbb{P}\left(\sup_{s\in[t,T]} |M(s)| > \lambda\right) \leq \frac{1}{\lambda^p}\mathbb{E}\left[|M(T)|^p\right]$, *for all* $\lambda > 0$.
- (ii) *If* $p > 1$, $\mathbb{E}\left[\sup_{s\in[t,T]} |M(s)|^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|M(T)|^p\right]$.

PROOF. We observe that, if $M(\cdot)$ is a right-continuous H-valued martingale such that $M(s) \in L^p(\Omega, \mathscr{F}, \mathbb{P}; H), p \geq 1$, for all $s \in [t,T]$, then by Proposition 1.40-(vi), $|M(\cdot)|^p$ is a right-continuous $\mathbb{R}$-valued submartingale with $|M(s)| \in L^p(\Omega, \mathscr{F}, \mathbb{P}; \mathbb{R})$ for all $s \in [t,T]$. The claims now easily follow from [283] Theorem 3.8 (i) and (iii), pages 13-14. □

In particular we see that a right-continuous E-valued martingale $M(\cdot)$ is square integrable if and only if $\mathbb{E}|M(T)|^2 < +\infty$.

NOTATION 1.73 (Square integrable martingales) Let $T \in (0, +\infty)$, $t \in [0, T)$, let $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P})$ be a filtered probability space, and E be a separable Banach space. The class of all continuous square integrable martingales $M\colon [t, T] \times \Omega \to E$ is denoted by $\mathcal{M}^2_{t,T}(E)$. ■

If H is a separable Hilbert space then $\mathcal{M}^2_{t,T}(H)$ endowed with the scalar product

$$\langle M, N \rangle_{\mathcal{M}^2_{t,T}} := \mathbb{E}\left[\langle M(T), N(T) \rangle\right].$$

is a Hilbert space (see [220], page 22).

THEOREM 1.74 (Angle bracket process, Quadratic variation process) *Let $T > 0$, $t \in [0, T)$, H be a separable Hilbert space, and $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P})$ be a filtered probability space. For every $M \in \mathcal{M}^2_{t,T}(H)$ there exists a unique (real) increasing, adapted, continuous process starting from 0 at t, called the* angle bracket process, *and denoted by $\langle M \rangle_t$, such that $|M_s|^2 - \langle M \rangle_s$ is a continuous martingale. Moreover there exists a unique $\mathcal{L}_1^+(H)$-valued continuous adapted process starting from 0 at t, called the* quadratic variation of M, *and denoted by $\langle\langle M \rangle\rangle_s$, such that, for all $x, y \in H$, the process*

$$\langle M_s, x \rangle \langle M_s, y \rangle - \Big\langle \langle\langle M \rangle\rangle_s (x), y \Big\rangle, \qquad s \in [t, T]$$

is a continuous martingale. Moreover $\langle M \rangle_s = \mathrm{Tr}(\langle\langle M \rangle\rangle_s)$.

PROOF. See [220], Definition 2.9 and Lemma 2.1, page 22. □

THEOREM 1.75 (Burkholder-Davis-Gundy inequality) *Let $T > 0, t \in [0, T)$, H be a separable Hilbert space, and $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P})$ be a filtered probability space. For every $p > 0$ there exists $c_p > 0$ such that, for every $M \in \mathcal{M}^2_{t,T}(H)$ with $M(0) = 0$,*

$$c_p^{-1} \mathbb{E}\left[\langle M \rangle_T^{p/2}\right] \le \mathbb{E}\left[\sup_{s \in [t,T]} |M(s)|^p\right] \le c_p \mathbb{E}\left[\langle M \rangle_T^{p/2}\right].$$

PROOF. See [380], Theorem 3.49, page 37. □

1.2.3. Stopping times.

DEFINITION 1.76 (Stopping time) *Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a filtration $\{\mathscr{F}_s^t\}_{s \ge t}$ on Ω. A random variable $\tau\colon (\Omega, \mathscr{F}) \to [t, +\infty]$ is said to be an* $\mathscr{F}_s^t$-stopping time *if, for all $s \ge t$,*

$$\{\tau \le s\} := \{\omega \in \Omega \,:\, \tau(\omega) \le s\} \in \mathscr{F}_s^t.$$

Given a stopping time τ we denote with $\mathscr{F}_\tau$ the sub-σ-field of $\mathscr{F}$ defined by

$$\mathscr{F}_\tau := \Big\{ A \in \mathscr{F} \,:\, A \cap \{\tau \le s\} \in \mathscr{F}_s^t \ \textit{for all } s \ge t \Big\}.$$

PROPOSITION 1.77 *Let $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P})$ be a filtered probability space.*

(i) *If τ and σ are $\mathscr{F}_s^t$-stopping times, so are $\tau \wedge \sigma$, $\tau \vee \sigma$ and $\tau + \sigma$.*

(ii) *If σ_n (for $n = 1, 2...$) are $\mathscr{F}_s^t$-stopping times, then*

$$\sup_n \sigma_n, \ \inf_n \sigma_n, \ \limsup_n \sigma_n, \ \liminf_n \sigma_n$$

are $\mathscr{F}_s^t$-stopping times.

(iii) *For any $\mathscr{F}_s^t$-stopping time τ there exists a decreasing sequence of discrete-valued $\mathscr{F}_s^t$-stopping times τ_n, such that $\lim_{n\to\infty} \tau_n = \tau$.*

(iv) *Let (S, d) be a metric space (endowed with the Borel σ-field induced by the distance), and $X\colon [t, +\infty) \times \Omega \to S$ be a continuous and $\mathscr{F}_s^t$-adapted process. Let $A \subset S$ be an open or a closed set. Then the* hitting time

$$\tau_A := \inf\{s \geq t \ : \ X(s) \in A\}$$

is a stopping time. (It is understood that $\inf\{\emptyset\} = +\infty$.*)*

Proof. (i) and (ii): see [283], Lemmas 2.9 and 2.11, page 7. (iii): see [281], Lemma 7.4, page 122. (iv): see [449], Example 3.3, page 24, or [349], Proposition 1.3.2., page 12 (there $S = \mathbb{R}^n$, but the proofs are the same). □

Proposition 1.78 *Let $\left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s \geq t}, \mathbb{P}\right)$ be a filtered probability space, $(\Omega_1, \mathscr{F}_1)$ be a measurable space, $X\colon [t, +\infty) \times \Omega \to \Omega_1$ be an $\mathscr{F}_s^t$-progressively measurable process, and τ be an $\mathscr{F}_s^t$-stopping time. Then the random variable $X(\tau)$, (where $X(\tau)(\omega) := X(\tau(\omega), \omega)$), is $\mathscr{F}_\tau$-measurable and the process $X(s \wedge \tau)$ is $\mathscr{F}_s^t$-progressively measurable.*

Proof. See [349], Proposition 1.3.5., page 13, or [449], Proposition 3.5, page 25. □

Theorem 1.79 (Doob's optional sampling theorem) *Let $\left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s \geq t}, \mathbb{P}\right)$ be a filtered probability space, $X\colon [t, +\infty) \times \Omega \to \mathbb{R}$ be a right-continuous $\mathscr{F}_s^t$-submartingale, and τ, σ be two $\mathscr{F}_s^t$-stopping times with τ bounded. Then X_τ is integrable and*

$$\mathbb{E}[X_\tau | \mathscr{F}_\sigma^t] \geq X_{\tau \wedge \sigma}, \quad \mathbb{P} \ a.s..$$

If X^+ (the positive part of the process) is uniformly integrable then the statement extends to unbounded τ.

Proof. See [281], Theorem 7.29, page 135. □

Definition 1.80 (Local martingale) *Let $\left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s \geq t}, \mathbb{P}\right)$ be a filtered probability space. An $\{\mathscr{F}_s^t\}_{s \geq t}$-adapted process $\{X(s)\}_{s \geq t}$ with values in a separable Banach space E is said to be a* local martingale *if there exists an increasing sequence of stopping times $\{\tau_n\}_{n \in \mathbb{N}}$ with $\mathbb{P}(\tau_n \uparrow +\infty) = 1$, such that the process $\{X(s \wedge \tau_n)\}_{s \geq t}$ is a martingale for every $n \in \mathbb{N}$.*

1.2.4. Q-Wiener process.

Definition 1.81 (Real Brownian motion) *Given $t \in \mathbb{R}$ a real stochastic process $\beta\colon [t, +\infty) \times \Omega \to \mathbb{R}$ on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is a* standard (one-dimensional) real Brownian motion *on $[t, +\infty)$ starting at 0, if*

(1) *β is continuous and $\beta(t) = 0$;*
(2) *for all $t \leq t_1 < t_2 < \ldots < t_n$ the random variables $\beta(t_1)$, $\beta(t_2) - \beta(t_1)$, ..., $\beta(t_n) - \beta(t_{n-1})$ are independent;*
(3) *for all $t \leq t_1 \leq t_2$, $\beta(t_2) - \beta(t_1)$ has a Gaussian distribution with mean 0 and covariance $t_2 - t_1$.*

Consider a real, separable Hilbert space Ξ and $Q \in \mathcal{L}^+(\Xi)$. Define $\Xi_0 := Q^{1/2}(\Xi)$ and let $Q^{-1/2}$ be the pseudo-inverse of $Q^{1/2}$ (see Definition B.1). Ξ_0 is a separable Hilbert space when endowed with the inner product $\langle x, y\rangle_{\Xi_0} := \left\langle Q^{-1/2}x, Q^{-1/2}y\right\rangle_{\Xi}$. Let Ξ_1 be an arbitrary real, separable Hilbert space such that $\Xi \subset \Xi_1$ with continuous embedding and $\Xi_0 \subset \Xi_1$ with Hilbert-Schmidt embedding $J\colon \Xi_0 \hookrightarrow \Xi_1$ (see Appendix B.3 on Hilbert-Schmidt operators). The operator

$Q_1 := JJ^*$ belongs to $\mathcal{L}_1^+(\Xi_1)$ and Ξ_0 is identical with the space $Q_1^{\frac{1}{2}}(\Xi_1)$ (see [130] Proposition 4.7, page 85).

THEOREM 1.82 *Consider the setting described above. Let $\{g_k\}_{k\in\mathbb{N}}$ be an orthonormal basis of Ξ_0 and $\{\beta_k\}_{k\in\mathbb{N}}$ be a sequence of mutually independent, standard one-dimensional Brownian motions $\beta_k\colon [t,+\infty)\times\Omega\to\mathbb{R}$ on $[t,+\infty)$ starting at 0. Then for every $s\in[t,+\infty)$ the series*

$$W_Q(s) := \sum_{k=1}^{\infty} g_k\beta_k(s) \tag{1.11}$$

is convergent in $L^2(\Omega,\mathscr{F},\mathbb{P};\Xi_1)$.

PROOF. See [130] Proposition 4.3, page 82, and Proposition 4.7, page 85. □

DEFINITION 1.83 (Q-Wiener process) *The process W_Q defined by (1.11) is called a Q-*Wiener process *on $[t,+\infty)$ starting at 0.*

REMARK 1.84 We will use the notation W_Q to denote a Q-Wiener process. If Q is trace-class, $\Xi_1 = \Xi$ is a canonical choice and it will be understood that W_Q is a Ξ valued process. If Q is not trace-class, writing W_Q and calling it a Q-Wiener process is a slight abuse of notation as it would be more precise to write W_{Q_1} and call it a Q_1-Wiener process with values in Ξ_1. However, even though the construction we have described is not canonical if $\mathrm{Tr}(Q)=+\infty$, and the choice of Ξ_1 is not unique, the class of the integrable processes is independent of the choice of Ξ_1 (see [130] Section 4.1 an in particular Proposition 4.7). Moreover (see [130] Section 4.1.2) for arbitrary $a\in\Xi$ the stochastic process

$$< a, W(s) >:= \sum_{k=1}^{\infty}\langle a, g_k\rangle\beta_k(s), \quad s\geq t,$$

is a real valued Wiener process and

$$\mathbb{E} < a, W(s_1) >< b, W(s_2) >= ((s_1-t)\wedge(s_2-t))\langle Qa, b\rangle,\ a,b\in\Xi.$$

For these reasons, even when $\mathrm{Tr}(Q)=+\infty$, we will still use the notation W_Q and in this case we will also call it a cylindrical Wiener process in Ξ. ■

PROPOSITION 1.85 *Let Ξ be a real, separable Hilbert space, $Q\in\mathcal{L}^+(\Xi)$ and let Ξ_0, Ξ_1 and J be as described above. Let $(\Omega,\mathscr{F},\mathbb{P})$ be a complete probability space and $B\colon[t,+\infty)\times\Omega\to\Xi_1$ be a stochastic process. Denote by $\mathscr{F}_s^{t,0}$ the filtration generated by B i.e.*

$$\mathscr{F}_s^{t,0} = \sigma(B(r)\ :\ t\leq r\leq s),$$

and $\mathscr{F}_s^t := \sigma(\mathscr{F}_s^{t,0},\mathcal{N})$, where $\mathcal{N}$ is the class of the $\mathbb{P}$-null sets. Then B is a Q-Wiener process on $[t,+\infty)$ starting at 0 if and only if:

(1) *$B(t)=0$.*
(2) *B has continuous trajectories.*
(3) *For all $t\leq t_1\leq t_2$ the random variable $B(t_2)-B(t_1)$ is independent of $\mathscr{F}_{t_1}^t$.*
(4) *$\mathcal{L}_{\mathbb{P}}(B(t_2)-B(t_1)) = \mathcal{N}(0,\,(t_2-t_1)Q_1)$, where $Q_1=JJ^*$.*

PROOF. The "only if" part follows from [130], Proposition 4.7, page 85 (observe that in [130] a Wiener process is in fact defined using the four properties (1)-(4)). The "if" part is proved in [130] Proposition 4.3-(ii), page 81 (if $\mathrm{Tr}(Q)=+\infty$ we apply the proposition in the space Ξ_1). □

The existence of a process satisfying conditions $(1)-(4)$ above can also be proved using the Kolmogorov extension theorem (see [130], Proposition 4.4).

REMARK 1.86 If $W_Q(s) = \sum_{k=1}^{\infty} g_k \beta_k(s)$ for some orthonormal basis $\{g_k\}_{k\in\mathbb{N}}$ of Ξ_0, it is easy to see that regardless of the choice of Ξ_1, $\mathscr{F}_s^{t,0} = \sigma\{\beta_k(r) : t \le r \le s, k \in \mathbb{N}\}$. Thus the filtration generated by W_Q does not depend on the choice of Ξ_1. ■

DEFINITION 1.87 (Translated $\mathscr{G}_s^t$-Q-Wiener process) *Let $0 \le t < T \le +\infty$. Let Ξ be a real, separable Hilbert space, $Q \in \mathcal{L}^+(\Xi)$ and let Ξ_0, Ξ_1 and J be as described above. Let $(\Omega, \mathscr{F}, \mathscr{G}_s^t, \mathbb{P})$ be a filtered probability space. We say that a stochastic process $B \colon [t,T] \times \Omega \to \Xi_1$ is a* translated $\mathscr{G}_s^t$-Q-Wiener process *on $[t,T]$ if:*

(1) *B has continuous trajectories.*
(2) *B is adapted to $\mathscr{G}_s^t$.*
(3) *For all $t \le t_1 < t_2 \le T$, $B(t_2) - B(t_1)$ is independent of $\mathscr{G}_{t_1}^t$.*
(4) *$\mathcal{L}_{\mathbb{P}}(B(t_2) - B(t_1)) = \mathcal{N}(0, (t_2 - t_1)Q_1)$, where $Q_1 = JJ^*$.*

If we also have $B(t) = 0$ then we call B a $\mathscr{G}_s^t$-Q-Wiener process *on $[t,T]$.*

We remark that if B is a translated $\mathscr{G}_s^t$-Q-Wiener process, then it is also a translated $\mathscr{F}_s^t$-Q-Wiener process, where $\mathscr{F}_s^t$ is the augmented filtration generated by B. Moreover if W_Q is a Q-Wiener process as in Definition 1.83 then it is also a $\mathscr{F}_s^t$-Q-Wiener process, where $\mathscr{F}_s^t$ is the augmented filtration generated by B.

LEMMA 1.88 *Let $0 \le t < T \le +\infty$. Let Ξ be a real, separable Hilbert space, $Q \in \mathcal{L}^+(\Xi)$ and let Ξ_0 and Ξ_1 be as described above. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space. Let $B \colon [t,T] \times \Omega \to \Xi_1$ be a continuous stochastic process such that $B(t) = 0$. Then B is a Q-Wiener process on $[t,T]$ if and only if, for all $a \in \Xi_1$, $t \le t_1 \le t_2 \le T$, we have*

$$\mathbb{E}\left[e^{i\langle a, B(t_2) - B(t_1)\rangle_{\Xi_1}} | \mathscr{F}_{t_1}^t\right] = e^{-\frac{\langle Q_1 a, a\rangle_{\Xi_1}}{2}(t_2 - t_1)}. \tag{1.12}$$

PROOF. (The proof uses the same arguments as in the finite-dimensional case, see Proposition 1.2.7 of [349]).

The "only if" part: if B is a Q-Wiener process then, by Proposition 1.85-(4), Theorem 1.56 and Definition 1.57,

$$\mathbb{E}\left[e^{i\langle a, B(t_2) - B(t_1)\rangle_{\Xi_1}}\right] = e^{-\frac{\langle Q_1 a, a\rangle_{\Xi_1}}{2}(t_2 - t_1)}.$$

Moreover, since $B(t_2) - B(t_1)$ is independent of $\mathscr{F}_{t_1}^t$,

$$\mathbb{E}\left[e^{i\langle a, B(t_2) - B(t_1)\rangle_{\Xi_1}}\right] = \mathbb{E}\left[e^{i\langle a, B(t_2) - B(t_1)\rangle_{\Xi_1}} | \mathscr{F}_{t_1}^t\right].$$

The "if" part: We have to prove the four conditions in Proposition 1.85: (1) and (2) are already in the assumptions of the lemma. Condition (4) follows easily from (1.12), Theorem 1.56 and Definition 1.57. To prove condition (3), i.e. that $Y := B(t_2) - B(t_1)$ is independent of $\mathscr{F}_{t_1}^t$, observe that, for all $Z \colon \Omega \to \Xi_1$ which are $\mathscr{F}_{t_1}^t$-measurable, one has, for all $a, b \in \Xi_1$,

$$\begin{aligned} \mathbb{E}\left[e^{i\langle a, Y\rangle_{\Xi_1}} e^{i\langle b, Z\rangle_{\Xi_1}}\right] &= \mathbb{E}\left[\mathbb{E}\left[e^{i\langle a, Y\rangle_{\Xi_1}} | \mathscr{F}_{t_1}^t\right] e^{i\langle b, Z\rangle_{\Xi_1}}\right] \\ &= e^{-\frac{\langle Q_1 a, a\rangle_{\Xi_1}}{2}(t_2 - t_1)} \mathbb{E}\left[e^{i\langle b, Z\rangle_{\Xi_1}}\right] = \mathbb{E}\left[e^{i\langle a, Y\rangle_{\Xi_1}}\right] \mathbb{E}\left[e^{i\langle b, Z\rangle_{\Xi_1}}\right]. \end{aligned} \tag{1.13}$$

Since the above holds for all $Z \colon \Omega \to \Xi_1$ which are $\mathscr{F}_{t_1}^t$-measurable, and for all $a, b \in \Xi_1$, we conclude that Y is independent of $\mathscr{F}_{t_1}^t$ by Theorem 1.55. □

LEMMA 1.89 *Let $\mathscr{F}_s^{t,0}$ and $\mathscr{F}_s^t$ be the filtrations defined in Proposition 1.85 for a Q-Wiener process W_Q. Then $\mathscr{F}_s^t$ is right continuous. Moreover, for all $T > t$,*

$\mathscr{F}_T^{t,0}$, and consequently $\mathscr{F}_T^t$, are countably generated up to sets of measure zero. If the trajectories of W_Q are everywhere continuous then

$$\mathscr{F}_T^{t,0} = \mathscr{F}_{T-}^{t,0} = \sigma\left(W_Q(s_i) : i = 1, 2, ...\right), \tag{1.14}$$

where $(s_i), i = 1, 2, ...$ is any dense sequence in $[t, T)$, and hence the filtration $\mathscr{F}_s^{t,0}$ is countably generated and left continuous.

Proof. The proof follows arguments from [402] and [283] (Section 2.7-A). Consider $\tau > s$ and $\epsilon > 0$. Since $W_Q(\tau + \epsilon) - W_Q(s + \epsilon)$ is independent of $\mathscr{F}_{s+}^{t,0}$, for every $A \in \mathscr{F}_{s+}^{t,0}$ and $f \in C_b(\Xi_1)$

$$\mathbb{E}\left(\mathbf{1}_A f(W_Q(\tau+\epsilon) - W_Q(s+\epsilon))\right) = \mathbb{P}(A)\mathbb{E}f(W_Q(\tau+\epsilon) - W_Q(s+\epsilon)).$$

Letting $\epsilon \to 0$ we thus have by the dominated convergence theorem that

$$\mathbb{E}\left(\mathbf{1}_A f(W_Q(\tau) - W_Q(s))\right) = \mathbb{P}(A)\mathbb{E}f(W_Q(\tau) - W_Q(s)). \tag{1.15}$$

Now if $B = \overline{B} \subset \Xi_1$ then there exist functions $f_n \in C_b(\Xi_1), 0 \le f_n \le 1$, such that $f_n(x) \to \mathbf{1}_B(x)$ as $n \to +\infty$ for every $x \in \Xi_1$. Therefore (1.15) implies that

$$\mathbb{P}(A \cap \{W_Q(\tau) - W_Q(s) \in B\}) = \mathbb{P}(A)\mathbb{P}(\{W_Q(\tau) - W_Q(s) \in B\})$$

and since the sets $\{\{W_Q(\tau) - W_Q(s) \in B\} : B = \overline{B} \subset \Xi_1\}$ are a π-system generating $\sigma(W_Q(\tau) - W_Q(s))$, it follows from Lemma 1.22 that $\mathscr{F}_{s+}^{t,0}$ and $\sigma(W_Q(\tau) - W_Q(s))$ are independent.

Now let $s = \tau_0 < \tau_1 < ... < \tau_k \le T$. We have $\sigma(W_Q(\tau_i) - W_Q(s) : i = 1, ..., k) = \sigma(W_Q(\tau_i) - W_Q(\tau_{i-1}) : i = 1, ..., k)$. Let now $A \in \mathscr{F}_{s+}^{t,0}$ and $B_i \in \sigma(W_Q(\tau_i) - W_Q(\tau_{i-1})), i = 1, ..., k$. Since B_i is independent of $A \cap B_1 \cap ... \cap B_{i-1} \in \mathscr{F}_{\tau_{i-1}}^{t,0}, i = 1, ..., k$ and $B_1, ..., B_k$ are independent

$$\begin{aligned}
&\mathbb{P}(A \cap B_1 \cap ... \cap B_k) = \mathbb{P}(A \cap B_1 \cap ... \cap B_{k-1})\mathbb{P}(B_k) = ... \\
&\quad = \mathbb{P}(A \cap B_1)\prod_{i=2}^k \mathbb{P}(B_i) = \mathbb{P}(A)\prod_{i=1}^k \mathbb{P}(B_i) = \mathbb{P}(A)\mathbb{P}(B_1 \cap ... \cap B_k).
\end{aligned}$$

Therefore $\bigcup \sigma(W_Q(\tau_i) - W_Q(s) : i = 1, ..., k)$ (where the union is taken over all partitions $s = \tau_0 < \tau_1 < ... < \tau_k \le T$) is a π-system independent of $\mathscr{F}_{s+}^{t,0}$ and thus $\mathscr{G}_s = \sigma(W_Q(\tau) - W_Q(s) : s \le \tau \le T)$ is independent of $\mathscr{F}_{s+}^{t,0}$.

Since $\mathscr{F}_T^{t,0} = \sigma(\mathscr{F}_s^{t,0}, \mathscr{G}_s)$, the family $\{A_s \cap B_s : A_s \in \mathscr{F}_s^{t,0}, B_s \in \mathscr{G}_s\}$ is a π-system generating $\mathscr{F}_T^{t,0}$. Let now $A \in \mathscr{F}_{s+}^{t,0}$ and let ξ be a version of $\mathbf{1}_A - \mathbb{E}(\mathbf{1}_A|\mathscr{F}_s^{t,0})$. Since ξ is $\mathscr{F}_{s+}^{t,0}$-measurable, it is independent of $\mathscr{G}_s$, so if $A_s \in \mathscr{F}_s^{t,0}, B_s \in \mathscr{G}_s$ then

$$\begin{aligned}
&\mathbb{E}\left(\xi \mathbf{1}_{A_s \cap B_s}\right) = \mathbb{E}\left(\xi \mathbf{1}_{A_s}\mathbf{1}_{B_s}\right) = \mathbb{P}(B_s)\mathbb{E}\left(\xi \mathbf{1}_{A_s}\right) \\
&\quad = \mathbb{P}(B_s)\int_{A_s} \xi d\mathbb{P} = \mathbb{P}(B_s)\left[\int_{A_s} \mathbf{1}_A d\mathbb{P} - \int_{A_s} \mathbb{E}(\mathbf{1}_A|\mathscr{F}_s^{t,0})d\mathbb{P}\right] = 0
\end{aligned}$$

by the definition of conditional expectation. This implies that $\int_D \xi d\mathbb{P} = 0$ for every $D \in \mathscr{F}_T^t$ and thus $\xi = 0, \mathbb{P}$ a.e.. Therefore $\mathbf{1}_A = \mathbb{E}(\mathbf{1}_A|\mathscr{F}_s^{t,0}), \mathbb{P}$ a.e., i.e if $\tilde{A} = \mathbb{E}(\mathbf{1}_A|\mathscr{F}_s^{t,0})^{-1}(1)$ then $\tilde{A} \in \mathscr{F}_s^{t,0}$ and $\mathbb{P}(A \Delta \tilde{A}) = 0$. This shows that $\mathscr{F}_{s+}^{t,0} \subset \mathscr{F}_s^t$.

Now let $A \in \mathscr{F}_{s+}^t$, which means that for every $n \ge 1$, $A \in \mathscr{F}_{s+1/n}^t$ and there exists $B_n \in \mathscr{F}_{s+1/n}^{t,0}$ such that $A \Delta B_n \in \mathcal{N}$. Set

$$B = \bigcap_{m=1}^{+\infty} \bigcup_{n=m}^{+\infty} B_n.$$

Then $B \in \mathscr{F}^{t,0}_{s+} \subset \mathscr{F}^t_s$ and

$$B \setminus A \subset \left(\bigcup_{n=1}^{+\infty} B_n \right) \setminus A = \bigcup_{n=1}^{+\infty} (B_n \setminus A) \in \mathcal{N}.$$

Moreover

$$\begin{aligned} A \setminus B &= A \cap \left(\bigcap_{m=1}^{+\infty} \bigcup_{n=m}^{+\infty} B_n \right)^c = A \cap \left(\bigcup_{m=1}^{+\infty} \bigcap_{n=m}^{+\infty} B_n^c \right) \\ &= \bigcup_{m=1}^{+\infty} \bigcap_{n=m}^{+\infty} (A \cap B_n^c) \subset \bigcup_{m=1}^{+\infty} (A \cap B_m^c) = \bigcup_{m=1}^{+\infty} (A \setminus B_m) \in \mathcal{N}. \end{aligned}$$

Thus $A \Delta B \in \mathcal{N}$ which implies that $A \in \mathscr{F}^t_s$, which completes the proof of the right continuity.

To show that $\mathscr{F}^{t,0}_T$ is countably generated up to sets of measure zero we take a dense sequence $(s_i), i = 1, 2, ...,$ in $[t, T)$. Since $\mathcal{B}(\Xi_1)$ is countably generated (for instance by open balls with rational radii centered at points of a countable dense set), each $\sigma(W_Q(s_i))$ is countably generated and so $\sigma(W_Q(s_i) : i \geq 1)$ is countably generated. It remains to show that for every $s \in (t, T], \sigma(W_Q(s)) \subset \sigma(\mathcal{N}, W_Q(s_i) : s_i < s)$. Let $\Omega_0 \subset \Omega, \mathbb{P}(\Omega_0) = 1$ be such that W_Q has continuous trajectories on $[t, T]$ for $\omega \in \Omega_0$. Let A be an open subset of Ξ_1 and set $A_n = \{x \in A : \mathrm{dist}(x, A^c) > 1/n\}, n = 1, 2, ...$ Then A_n is open, $\overline{A}_n \subset A_{n+1}$, and $\bigcup_{n=1}^{+\infty} A_n = A$. Let s_{i_k} be a sequence of s_i such that $s_{i_k} < s$ and $s_{i_k} \to s$ as $k \to +\infty$. Then, using the continuity of the trajectories of W_Q, it is easy to see that

$$\Omega_0 \cap W_Q(s)^{-1}(A) = \Omega_0 \cap \bigcup_{n=1}^{+\infty} \bigcap_{k=n}^{+\infty} W_Q(s_{i_k})^{-1}(A_n) \in \sigma(\mathcal{N}, W_Q(s_i) : s_i < s).$$

Therefore $W_Q(s)^{-1}(A) \in \sigma(\mathcal{N}, W_Q(s_i) : s_i < s)$ and since the sets $\{W_Q(s)^{-1}(A) :$ A is an open subset of $\Xi_1\}$ generate $\sigma(W_Q(s))$, the result follows. If $\Omega_0 = \Omega$ then we have above

$$W_Q(s)^{-1}(A) = \bigcup_{n=1}^{+\infty} \bigcap_{k=n}^{+\infty} W_Q(s_{i_k})^{-1}(A_n) \in \sigma(W_Q(s_i) : s_i < s).$$

The argument that $\sigma(W_Q(t)) \subset \sigma(W_Q(s_i) : i = 1, 2, ...)$ is similar (or we can just assume that $s_1 = t$). This yields (1.14). □

In fact the above argument shows that if S is a Polish space, $T > t$, and $X : [t, T] \times \Omega \to S$ is a stochastic process with everywhere continuous trajectories, then the filtration generated by X, $\mathscr{F}^X_s := \sigma(X(\tau) : t \leq \tau \leq s)$ is countably generated and left-continuous.

1.2.5. Simple and elementary processes.

DEFINITION 1.90 ($\mathscr{F}^t_s$-simple process) *Let E be a Banach space (endowed with the Borel σ-field) and let $(\Omega, \mathscr{F}, \{\mathscr{F}^t_s\}_{s\in[t,T]}, \mathbb{P})$ be a filtered probability space. A process $X\colon [t, T] \times (\Omega, \mathscr{F}, \mathbb{P}) \to E$ is called* $\mathscr{F}^t_s$-simple *if:*

(i) *Case $T = +\infty$: there exists a sequence of real numbers $\{t_n\}_{n\in\mathbb{N}}$ with $t = t_0 < t_1 < ... < t_n < ...$ and $\lim_{n\to\infty} t_n = +\infty$, a constant $C < +\infty$, and a sequence of random variables $\xi_n\colon \Omega \to E$ with $\sup_{n\geq 0} |\xi_n(\omega)|_E \leq C$ for every $\omega \in \Omega$, such that ξ_n is $\mathscr{F}^t_{t_n}$-measurable for every $n \geq 0$, and*

$$X(s)(\omega) = \begin{cases} \xi_0(\omega) & \text{if } s = t \\ \xi_i(\omega) & \text{if } s \in (t_i, t_{i+1}]. \end{cases}$$

(ii) *Case $T < +\infty$: there exist $t = t_0 < t_1 < ... < t_N = T$, a constant $C < +\infty$, and random variables $\xi_n : \Omega \to E$ for $n = 0, ..., N-1$ with $\sup_{0\le n\le N-1} |\xi_n(\omega)|_E \le C$ for every $\omega \in \Omega$, such that ξ_n is $\mathscr{F}^t_{t_n}$-measurable, and*

$$X(s)(\omega) = \begin{cases} \xi_0(\omega) & \text{if } s = t \\ \xi_i(\omega) & \text{if } s \in (t_i, t_{i+1}]. \end{cases}$$

Definition 1.91 ($\mathscr{F}^t_s$-elementary process) *Let $T \in (0, +\infty)$, $t \in [0, T)$. Let (S, d) be a complete metric space (endowed with the Borel σ-field), and $(\Omega, \mathscr{F}, \{\mathscr{F}^t_s\}_{s\in[t,T]}, \mathbb{P})$ be a filtered probability space. We say that a process $X : [t, T] \times (\Omega, \mathscr{F}, \mathbb{P}) \to S$ is* $\mathscr{F}^t_s$-elementary, *if there exist S-valued random variables $\xi_0, \xi_1, .., \xi_{N-1}$, and a sequence $t = t_0 < t_1 < .. < t_N = T$, such that*

(1) *ξ_i has a finite numbers of values for every $i \in \{0, ..N-1\}$.*
(2) *ξ_i is $\mathscr{F}^t_{t_i}$-measurable for every $i \in \{0, ..N-1\}$.*
(3) *$X(s)(\omega) = \xi_i(\omega)$ for $s \in (t_i, t_{i+1}]$ for $i \in \{0, ..N-1\}$, and $X(t) = \xi_0$.*

Finally we say that a process $X : [t, +\infty) \times (\Omega, \mathscr{F}, \mathbb{P}) \to S$ is $\mathscr{F}^t_s$-elementary *if there exists $T_1 > t$ such that the restriction of X to $[t, T_1]$ is $\mathscr{F}^t_s$-elementary and $X(s) = 0$ for $s > T_1$.*

It is immediate from the definitions that simple and elementary processes are progressively measurable and predictable.

Remark 1.92 In Definitions 1.14, 1.90 and 1.91 we introduced the concepts of *simple* random variable, $\mathscr{F}^t_s$-*simple* processes, and $\mathscr{F}^t_s$-*elementary* process. The reader should be aware that in the literature the use of these terms varies and the same word is often used by different authors to mean different things. ■

Lemma 1.93 *Let E be a separable Banach space endowed with the Borel σ-field, $(\Omega, \mathscr{F}, \mathscr{F}^t_s, \mathbb{P})$ be a filtered probability space and $X : [t, T] \times \Omega \to E$ be a bounded, measurable, $\mathscr{F}^t_s$-adapted process, where $T \in [t, +\infty) \cup \{+\infty\}$. There exists a sequence $X^m(\cdot)$ of $\mathscr{F}^t_s$-elementary E-valued processes on $[t, T]$ such that, for every $1 \le p < +\infty$ and $R > t$,*

$$\lim_{m\to+\infty} \mathbb{E} \int_t^{R\wedge T} |X^m(s) - X(s)|^p_E \, ds = 0. \tag{1.16}$$

The same claim holds if, instead of the Banach space, we consider E to be an interval $[a, b] \subset \mathbb{R}$ or a countable closed subset of $[a, b]$.

Proof. It is enough to prove the result for a single $p \ge 1$. To obtain a sequence of $\mathscr{F}^t_s$-simple processes X^m with the required properties, the proof follows exactly the proof of Lemma 3.2.4, page 132, in [283] with obvious technical modifications as we now have to deal with Bochner integrals in E. We then use Lemma 1.16 to approximate the random variables ξ_i defining X^m by simple random variables to obtain $\mathscr{F}^t_s$-elementary approximating processes.

If E is a countable closed subset of $[a, b]$, we first produce $[a, b]$-valued $\mathscr{F}^t_s$-elementary approximating processes X^m. We then construct an E-valued $\mathscr{F}^t_s$-elementary process Y^m from X^m as follows. Let $X^m(s) = \xi_i$ for $s \in (t_i, t_{i+1}]$ for $i \in \{0, ..N-1\}$, and $X(t) = \xi_0$. Let $\tilde{\xi}_i$ be defined in the following way. If $\xi_i(\omega) \in E$, we set $\tilde{\xi}_i(\omega) = \xi_i(\omega)$. If $\xi_i(\omega) \notin E$, we set $\tilde{\xi}_i(\omega) = \arg\min_{x\in E} |\xi(\omega) - x|$ if $\arg\min_{x\in E} |\xi(\omega) - x|$ is a singleton. If $\arg\min_{x\in E} |\xi(\omega) - x|$ has two points $x_1 < x_2$, we set $\tilde{\xi}_i(\omega) = x_1$. Obviously $\tilde{\xi}_i$ is a simple, $\mathscr{F}^t_{t_i}$-measurable process. We now define $Y^m(s) = \tilde{\xi}_i$ for $s \in (t_i, t_{i+1}]$ for $i \in \{0, ..N-1\}$, and $X(t) = \tilde{\xi}_0$. Then, since X has values in E, it is easy to see that $|Y^m - X| \le 2|X^m - X|$. Therefore the result follows. □

LEMMA 1.94 *Let $\mathscr{F}_s^{t,0}$ and $\mathscr{F}_s^t$ be as in Proposition 1.85, $T \in [t, +\infty) \cup \{+\infty\}$, and let $a(\cdot) : [t, T] \times \Omega \to S$ be an $\mathscr{F}_s^t$-progressively measurable process, where (S, d) is a Polish space endowed with the Borel σ-field. Then there exists an $\mathscr{F}_s^{t,0}$-progressively measurable and $\mathscr{F}_s^{t,0}$-predictable process $a_1(\cdot) : [t, T] \times \Omega \to S$, such that $a(\cdot) = a_1(\cdot)$, $dt \otimes \mathbb{P}$-a.e. on $[t, T] \times \Omega$.*

PROOF. In light of Theorem 1.13 we can assume that $S = [0, 1]$ or S is a countable closed subset of $[0, 1]$. Using Lemma 1.93, we can find an approximating $\mathscr{F}_s^t$-elementary processes $a^n(\cdot)$ on $[t, T]$ of the form

$$a^n(t)(\omega) = \begin{cases} \xi_0^n(\omega) & \text{if } s = t \\ \xi_i^n(\omega) & \text{if } s \in (t_i, t_{i+1}]. \end{cases}$$

such that

$$\sup_{R \geq t} \lim_{n \to \infty} \mathbb{E} \int_t^{R \wedge T} |a(s) - a^n(s)|_{\mathbb{R}}^2 \, ds = 0.$$

Using Lemma 1.16, we can change every ξ_i^n on a null-set to obtain a sequence of $\mathscr{F}_s^{t,0}$-elementary processes $a_1^n(\cdot)$, that still satisfy

$$\sup_{R \geq t} \lim_{n \to \infty} \mathbb{E} \int_t^{R \wedge T} |a(s) - a_1^n(s)|_{\mathbb{R}}^2 \, ds = 0.$$

Obviously the processes $a_1^n(\cdot)$ are $\mathscr{F}_s^{t,0}$-progressively measurable. We can now extract a subsequence (still denoted by $a_1^n(\cdot)$) such that $a_1^n(\cdot) \to a(\cdot)$ $dt \otimes \mathbb{P}$ a.e. on $[t, T] \times \Omega$, and define $a_1(\cdot) := \liminf_{n \to +\infty} a_1^n(\cdot)$. The process $a_1(\cdot)$ is $\mathscr{F}_s^{t,0}$-progressively measurable, $\mathscr{F}_s^{t,0}$-predictable, and $a(\cdot) = a_1(\cdot)$, $dt \otimes \mathbb{P}$ a.e. on $[t, T] \times \Omega$. □

1.3. Stochastic integral

Let $T \in (0, +\infty)$, and $t \in [0, T)$. Throughout the whole section Ξ and H will be two real, separable Hilbert spaces, Q will be an operator in $\mathcal{L}^+(\Xi)$, $\left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s \in [t,T]}, \mathbb{P}\right)$ will be a filtered probability space, and W_Q will be a translated $\mathscr{F}_s^t$-Q-Wiener process on Ω on $[0, T]$. The following concept will be used in Chapter 2.

DEFINITION 1.95 *A 5-tuple $\mu := \left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s \in [t,T]}, \mathbb{P}, W_Q\right)$ described above is called a* generalized reference probability space.

A process $X(\cdot)$ will always be assumed to be defined on Ω, and the expressions "adapted" and "progressively measurable" will always refer to the filtration $\mathscr{F}_s^t$.

1.3.1. Definition of stochastic integral.

In this section we will assume that $\mathrm{Tr}(Q) < +\infty$. If $\mathrm{Tr}(Q) = +\infty$, the construction of the stochastic integral is the same, we just have to consider W_Q as a Ξ_1 valued Wiener process with nuclear covariance Q_1 (see Section 1.2.4). This way W_Q is not uniquely determined but $Q_1^{1/2}(\Xi_1) = \Xi_0 = Q^{1/2}(\Xi), |x|_{\Xi_0} = |Q_1^{-1/2} x|_{\Xi_1}$ for all possible extensions Ξ_1 and the class of integrands and the value of the integrals are independent of the choice of the space Ξ_1 (see [130], Proposition 4.7 and Section 4.1.2).

We recall that we denote by $\mathcal{L}_2(\Xi_0; H)$ the space of Hilbert-Schmidt operators from Ξ_0 to H (see Appendix B.3). It is equipped with its Borel σ field $\mathcal{B}(\mathcal{L}_2(\Xi_0; H))$. $\mathcal{L}_2(\Xi_0; H)$ is a real, separable Hilbert space (see Proposition B.25), and $\mathcal{L}(\Xi; H)$ is dense in $\mathcal{L}_2(\Xi_0; H)$.

DEFINITION 1.96 (The space $\mathcal{N}_Q^p(t,T;H)$) *Given $p \geq 1$, we denote by $\mathcal{N}_Q^p(t,T;H)$ the space of all $\mathcal{L}_2(\Xi_0;H)$-valued, progressively measurable processes $X(\cdot)$, such that*

$$|X(\cdot)|_{\mathcal{N}_Q^p(t,T;H)} := \left(\mathbb{E} \int_t^T \|X(s)\|^p_{\mathcal{L}_2(\Xi_0;H)} ds \right)^{1/p} < \infty.$$

$\mathcal{N}_Q^p(t,T;H)$ is a Banach space if it is endowed with the norm $|\cdot|_{\mathcal{N}_Q^p(t,T;H)}$.

We remark that, as always, two processes in $\mathcal{N}_Q^p(t,T;H)$ are identified if they are equal $\mathbb{P} \otimes dt$ a.e..

REMARK 1.97 In several classical references (see e.g. [130] or [385]), the theory of stochastic integration is developed for predictable processes instead of progressively measurable ones like in our case. However, it follows for instance from Lemma 1.94, that for every $\mathcal{L}_2(\Xi_0;H)$-valued progressively measurable process X there exists a predictable process X_1 which is $\mathbb{P} \otimes dt$ a.e. equal to X. Thus, since we are working with stochastic integrals with respect to Wiener processes (which are continuous), the two concepts coincide. ■

For an $\mathcal{L}(\Xi;H)$-valued, $\mathscr{F}_s^t$-simple process Φ on $[t,T]$, $\Phi(s) = \Phi_0 \mathbf{1}_{\{t\}}(s) + \sum_{i=0}^{i=N-1} \mathbf{1}_{(t_i,t_{i+1}]}(s)\Phi_i$, the stochastic integral with respect to W_Q is defined by

$$\int_t^T \Phi(s) dW_Q(s) := \sum_{i=0}^{N-1} \Phi_i (W_Q(t_{i+1}) - W_Q(t_i)) \in L^2(\Omega;H).$$

Note that if we take Φ to be $\mathcal{L}_2(\Xi_0;H)$-valued, we cannot guarantee that the expression above is well defined, since $\mathcal{L}_2(\Xi_0;H)$ contains unbounded operators in Ξ.

We now extend the stochastic integral to all processes in $\mathcal{N}_Q^2(t,T;H)$ by the following theorem.

THEOREM 1.98 (Itô isometry) *For every $\mathcal{L}(\Xi;H)$-valued, $\mathscr{F}_s^t$-simple process Φ we have*

$$\mathbb{E} \left| \int_t^T \Phi(s) dW_Q(s) \right|_H^2 = \mathbb{E} \int_t^T \|\Phi(s)\|^2_{\mathcal{L}_2(\Xi_0;H)} ds.$$

Thus the stochastic integral is an isometry between the set of $\mathcal{L}(\Xi;H)$-valued, $\mathscr{F}_s^t$-simple processes in $\mathcal{N}_Q^2(t,T;H)$ and its image in $L^2(\Omega;H)$. Moreover, since $\mathcal{L}(\Xi;H)$-valued, $\mathscr{F}_s^t$-simple (and in fact elementary) processes are dense in $\mathcal{N}_Q^2(t,T;H)$, it can be uniquely extended to all processes in $\mathcal{N}_Q^2(t,T;H)$. We denote this unique extension by

$$\int_t^T \Phi(s) dW_Q(s)$$

and call it the stochastic integral *of Φ with respect to W_Q.*

PROOF. See [220], Propositions 2.1, 2.2, and Definition 2.10. See also [130], Proposition 4.22 in the context of predictable processes. □

PROPOSITION 1.99 *For $\Phi \in \mathcal{N}_Q^2(t,T;H)$, consider the process*

$$\begin{cases} I(\Phi) \colon [t,T] \times \Omega \to H \\ I(\Phi)(r) := \int_t^r \Phi(s) dW_Q(s) := \int_t^T \Phi(s) \mathbf{1}_{[t,r]} dW_Q(s). \end{cases}$$

$I(\Phi)$ is a continuous square integrable martingale and $I : \mathcal{N}_Q^2(t,T;H) \to \mathcal{M}_{t,T}^2(H)$ is an isometry. Moreover,

$$\langle\langle I(\Phi)\rangle\rangle_s = \int_t^s \left(\Phi(s)Q^{\frac{1}{2}}\right)\left(\Phi(s)Q^{\frac{1}{2}}\right)^* ds,$$

$$\langle I(\Phi)\rangle_s = \int_t^s \|\Phi(s)\|^2_{\mathcal{L}_2(\Xi_0;H)} ds.$$

PROOF. See [220] Theorem 2.3, page 34. □

The definition of stochastic integral can be further extended to all $\mathcal{L}_2(\Xi_0;H)$-valued progressively measurable processes $\Phi(\cdot)$ such that

$$\mathbb{P}\left(\int_t^T \|\Phi(s)\|^2_{\mathcal{L}_2(\Xi_0;H)} ds < +\infty\right) = 1. \tag{1.17}$$

LEMMA 1.100 *Let $\{\Phi(s)\}_{s\in[t,T]}$ be an $\mathcal{L}_2(\Xi_0;H)$-valued progressively measurable process satisfying (1.17). Then there exists a sequence Φ_n of $\mathcal{L}(\Xi;H)$-valued $\mathscr{F}_s^t$-simple processes such that*

$$\lim_{n\to\infty} \int_t^T \|\Phi(s) - \Phi_n(s)\|^2_{\mathcal{L}_2(\Xi_0;H)} ds = 0 \qquad \mathbb{P} - a.s.. \tag{1.18}$$

Moreover there exists an H-valued random variable, denoted by $\mathcal{I}$, such that

$$\lim_{n\to\infty} \int_t^T \Phi_n(s) dW_Q(s) = \mathcal{I} \qquad \textit{in probability.}$$

$\mathcal{I}$ does not depend on the choice of approximating sequence, more precisely, given Φ_n^1 and Φ_n^2 satisfying (1.18), if $\mathcal{I}_1 := \lim_{n\to\infty} \int_t^T \Phi_n^1(s)dW_Q(s)$ and $\mathcal{I}_2 := \lim_{n\to\infty} \int_t^T \Phi_n^2(s)dW_Q(s)$, then $\mathcal{I}_1 = \mathcal{I}_2$ $\mathbb{P} - a.s.$.

PROOF. See [220], Lemma 2.3, page 39, and Lemma 2.6, page 41. □

The process $\mathcal{I}$ defined by Lemma 1.100 is also called the stochastic integral of Φ with respect to W_Q, and is denoted by $\int_t^T \Phi(s)dW_Q(s)$. We also set $\int_t^r \Phi(s)dW_Q(s) := \int_t^T \Phi(s)\mathbf{1}_{[t,r]}dW_Q(s)$.

PROPOSITION 1.101 *Let $\{\Phi(s)\}_{s\in[t,T]}$ be an $\mathcal{L}_2(\Xi_0;H)$-valued progressively measurable process satisfying (1.17). Then the process*

$$\begin{cases} I(\Phi)\colon [t,T]\times\Omega \to H \\ I(\Phi)(r) := \int_t^r \Phi(s)dW_Q(s). \end{cases}$$

is a continuous local martingale.

PROOF. See [220] pages 42-44. □

Finally we may extend the definition of stochastic integral to all processes that are equivalent to progressively measurable processes satisfying (1.17).

DEFINITION 1.102 *We say that two processes Φ_1 and Φ_2 are equivalent if $\Phi_1 = \Phi_2$, $dt\otimes\mathbb{P}$-a.e.. If Φ belongs to the equivalence class of a progressively measurable process Φ_1 satisfying (1.17), we set*

$$\int_t^T \Phi(s)dW_Q(s) := \int_t^T \Phi_1(s)dW_Q(s).$$

This definition is obviously independent of the choice of a representative process Φ_1. Thus a representative process defines the stochastic integral for the whole equivalence class.

EXAMPLE 1.103 Every $\mathcal{L}_2(\Xi_0; H)$-valued, $\mathscr{F}_s^t$-adapted, and $\overline{\mathcal{B}([t,T]) \otimes \mathscr{F}}$ measurable process Φ satisfying (1.17) is stochastically integrable, where $\overline{\mathcal{B}([t,T]) \otimes \mathscr{F}}$ is the completion of $\mathcal{B}([t,T]) \otimes \mathscr{F}$ with respect to $dt \otimes \mathbb{P}$. To see this we need to find a progressively measurable process Φ_1 which is equivalent to Φ. First, let Φ_2 be a $\mathcal{B}([t,T]) \otimes \mathscr{F}$ measurable process equivalent to Φ (which exists by Lemma 1.16). Since $\mathscr{F}$ is complete, $\Phi_2(s, \cdot)$ is $\mathscr{F}_s^t$ measurable for a.e. s. Thus there exists $A \in \mathcal{B}([t,T])$ of full measure such that $\Phi_2(s, \cdot)$ is $\mathscr{F}_s^t$ measurable for $s \in A$. We then define $\Phi_3 = \Phi_2 \mathbf{1}_A$. Φ_3 is $\mathcal{B}([t,T]) \otimes \mathscr{F}$ measurable and $\mathscr{F}_s^t$-adapted, and thus it has a progressively measurable modification Φ_1 which is clearly equivalent to Φ. ■

THEOREM 1.104 *Let $(E, \mathscr{G}, \mu)$ be a measure space with bounded measure. Let $\Phi : [t,T] \times \Omega \times E \to \mathcal{L}_2(\Xi_0; H)$, be $(\mathcal{B}([t,T]) \otimes \mathscr{F}_T^t \otimes \mathscr{G})/\mathcal{B}(\mathcal{L}_2(\Xi_0; H))$ measurable. Suppose that, for any $x \in E$, $\{\Phi(s, \cdot, x)\}_{s \in [t,T]}$ is progressively measurable and*

$$\int_E |\Phi(\cdot, \cdot, x)|_{\mathcal{N}_Q^2(t,T;H)} d\mu(x) < +\infty.$$

Then:

(i) $\displaystyle\int_t^T \Phi(s, \cdot, x) dW(s)$ *has a $\mathscr{F}_T^t \otimes \mathscr{G}/\mathcal{B}(H)$ measurable version.*

(ii) $\displaystyle\int_G \Phi(\cdot, \cdot, x) d\mu(x)$ *is progressively measurable.*

(iii) *The following equality holds $\mathbb{P}$-a.s.:*

$$\int_G \int_t^T \Phi(s, \cdot, x) dW(s) d\mu(x) = \int_t^T \int_G \Phi(s, \cdot, x) d\mu(x) dW(s).$$

PROOF. See Theorem 2.8, Section 2.2.6, page 57 of [220] and Theorem 4.33, Section 4.5, page 110 of [130]. □

1.3.2. Basic properties and estimates.

LEMMA 1.105 *Let $T > 0$ and $t \in [0,T)$. Assume that Φ is in $\mathcal{N}_Q^2(t,T;H)$ and that τ is an $\mathscr{F}_s^t$-stopping time such that $\mathbb{P}(\tau \le T) = 1$. Then $\mathbb{P}$-a.s.*

$$\int_t^T \mathbf{1}_{[t,\tau]}(r) \Phi(r) dW_Q(r) = \int_t^\tau \Phi(r) dW_Q(r).$$

PROOF. See [220], Lemma 2.7, page 43 (also [130], Lemma 4.24, page 99). □

As a consequence of Theorem 1.75 and Proposition 1.99 we obtain the following theorem (see also e.g. [127], Theorem 5.2.4, page 58).

THEOREM 1.106 (Burkholder-Davis-Gundy inequality for stochastic integrals) *Let $T > 0$ and $t \in [0,T)$. For every $p \ge 2$, there exists a constant c_p such that, for every Φ in $\mathcal{N}_Q^p(t,T;H)$,*

$$\mathbb{E}\left[\sup_{s \in [t,T]} \left| \int_t^s \Phi(r) dW_Q(r) \right|^p \right] \le c_p \mathbb{E}\left[\int_t^T \|\Phi(r)\|^2_{\mathcal{L}_2(\Xi_0,H)} dr \right]^{p/2}$$
$$\le c_p (T-t)^{\frac{p}{2}-1} \mathbb{E}\left[\int_t^T \|\Phi(r)\|^p_{\mathcal{L}_2(\Xi_0,H)} dr \right].$$

PROPOSITION 1.107 *Let $T > 0$ and $t \in [0,T)$. Let A be the generator of a C_0-semigroup $\{e^{rA},\ r \ge 0\}$ on H such that $\|e^{rA}\| \le M e^{\alpha r}$ for every $r \ge 0$ for* *some* $\alpha \in \mathbb{R}$, $M > 0$. *Let* $p > 2$ *and* $\Phi \in \mathcal{N}_Q^p(t,T;H)$. *Let* A_n *be the Yosida approximation of* A. *Then the stochastic convolution process*

$$\Psi(s) := \int_t^s e^{(s-r)A}\Phi(r)dW_Q(r) \tag{1.19}$$

has a continuous modification,

$$\mathbb{E}\left[\sup_{s\in[t,T]}\left|\int_t^s e^{(s-r)A}\Phi(r)dW_Q(r)\right|^p\right] \le C\mathbb{E}\left[\int_t^T \|\Phi(r)\|^p_{\mathcal{L}_2(\Xi_0,H)}dr\right], \tag{1.20}$$

where the constants c *and* C *depend only on* $T-t$, p, M, α, *and*

$$\lim_{n\to\infty}\mathbb{E}\left[\sup_{s\in[t,T]}\left|\int_t^s \left(e^{(s-r)A_n} - e^{(s-r)A}\right)\Phi(r)dW_Q(r)\right|^p\right] = 0. \tag{1.21}$$

If moreover A *generates a* C_0*-pseudo-contraction semigroup (i.e.* $M = 1$ *above, see Appendix B.4) then the claims are also true for* $p = 2$.

Proof. See [220], Lemma 3.3, page 87. The claims for p=2 can be proved repeating the arguments of the proof of Proposition 3.3 of [428] which uses the Unitary Dilation Theorem. □

Proposition 1.108 *Let* A *be the generator of a* C_0*-semigroup on* H, $T > 0$, *and* $t \in [0,T)$. *Assume that* $\Phi : [t,T]\times\Omega \to \mathcal{L}_2(\Xi_0;H)$ *is a progressively measurable process such that* $\Phi(s) \in \mathcal{L}_2(\Xi_0;D(A))$ $\mathbb{P}$*-a.s., for a.e.* $s \in [t,T]$. *Assume that*

$$\mathbb{P}\left(\int_t^T \|\Phi(s)\|^2_{\mathcal{L}_2(\Xi_0;D(A))}ds < +\infty\right) = 1.$$

Then

$$\mathbb{P}\left(\int_t^T \Phi(s)dW_Q(s) \in D(A)\right) = 1 \tag{1.22}$$

and

$$A\int_t^T \Phi(s)dW_Q(s) = \int_t^T A\Phi(s)dW_Q(s), \qquad \mathbb{P}-a.s. \tag{1.23}$$

Proof. We can assume without loss of generality that $Q \in \mathcal{L}_1^+(\Xi)$. The proof follows the proof of Proposition 3.1 (p.76) of [220], however we present it here to clarify a measurability issue. Indeed we first need to show that Φ is an $\mathcal{L}_2(\Xi_0;D(A))$-valued, progressively measurable process. To do this we take $\Psi_n = J_n\Phi$, where $J_n = n(nI-A)^{-1}$ (see Definition B.40). Since $J_n \in \mathcal{L}(H;D(A))$, Ψ_n is an $\mathcal{L}_2(\Xi_0;D(A))$-valued, progressively measurable process. Moreover it is easy to see that if $\Phi(s)(\omega) \in \mathcal{L}_2(\Xi_0;D(A))$, then $\Psi_n(s)(\omega) \to \Phi(s)(\omega)$ in $\mathcal{L}_2(\Xi_0;D(A))$. Therefore, defining $V := \{(s,\omega) : \Psi_n(s)(\omega)$ converges in $\mathcal{L}_2(\Xi_0;D(A))\}$, it follows from Lemma 1.8-(iii) that Φ is equivalent to a progressively measurable process $\lim_{n\to+\infty}\mathbf{1}_V\Psi_n$. The proof is now done in two steps.
Step 1: The claim is true for $\mathscr{F}_s^t$-simple $\mathcal{L}(\Xi;D(A))$-valued processes.
Step 2: If Φ is a $\mathcal{L}_2(\Xi_0;D(A))$-valued progressively measurable process satisfying the hypotheses of this proposition, we take a sequence of $\mathscr{F}_s^t$-simple $\mathcal{L}(\Xi;D(A))$-valued processes Φ_n approximating Φ in the sense of (1.18) so that

$$\lim_{n\to+\infty}\int_t^T \|\Phi(s)-\Phi_n(s)\|^2_{\mathcal{L}_2(\Xi_0;D(A))}ds = 0 \qquad \mathbb{P}-a.s..$$

In particular we have

$$\int_t^T \Phi_n(s)dW_Q(s) \xrightarrow{n\to\infty} \int_t^T \Phi(s)dW_Q(s),$$

$$A \int_t^T \Phi_n(s) dW_Q(s) = \int_t^T A\Phi_n(s) dW_Q(s) \xrightarrow{n\to\infty} \int_t^T A\Phi(s) dW_Q(s)$$

in probability, so the claim follows since A is a closed operator. $\square$

LEMMA 1.109 *Let $T > 0, t \in [0,T)$, and $0 < \alpha < 1$. Let A be the generator of a C_0-semigroup $\{e^{rA},\ r \geq 0\}$ on H, and A_1 be a linear, closed operator such that $A_1 e^{rA}$ is bounded and $A_1 e^{rA} = e^{rA} A_1$ for every $r > 0$. Let $\Phi : [t,T] \times \Omega \to \mathcal{L}_2(\Xi_0; H)$ be progressively measurable. Assume that, for all $s \in [t,T]$,*

$$\int_t^s (s-r)^{\alpha-1} \left(\int_t^r (r-h)^{-2\alpha} \mathbb{E} \left[\left\| A_1 e^{(r-h)A} \Phi(h) \right\|^2_{\mathcal{L}_2(\Xi_0;H)} \right] dh \right)^{1/2} dr < +\infty. \tag{1.24}$$

Then

$$\int_t^s A_1 e^{(s-r)A} \Phi(r) dW_Q(r) = \frac{\sin(\alpha\pi)}{\pi} \int_t^s (s-r)^{\alpha-1} e^{(s-r)A} Y_\alpha^\Phi(r) dr \qquad \mathbb{P}-a.s.$$

for all $s \in [t,T]$, where

$$Y_\alpha^\Phi(r) := \int_t^r (r-h)^{-\alpha} A_1 e^{(r-h)A} \Phi(h) dW_Q(h). \tag{1.25}$$

PROOF. The statement is similar to [127], Theorem 5.2.5, page 58, Section 5.2.1. The proof is based on the factorization method and we give it for completeness.

We use the identity

$$\int_\sigma^t (t-s)^{\alpha-1}(s-\sigma)^{-\alpha} ds = \frac{\pi}{\sin\, \pi\alpha}, \qquad \text{for all } \sigma \leq s \leq t,\ 0<\alpha<1 \tag{1.26}$$

(which can be proved by a simple direct computation). We have

$$\begin{aligned}
&\int_t^s (s-r)^{\alpha-1} e^{(s-r)A} Y_\alpha^\Phi(r) dr \\
&\qquad = \int_t^s (s-r)^{\alpha-1} e^{(s-r)A} \int_t^r (r-h)^{-\alpha} A_1 e^{(r-h)A} \Phi(h) dW_Q(h) dr \\
&\qquad\qquad = \frac{\pi}{\sin\, \pi\alpha} \int_t^s A_1 e^{(s-h)A} \Phi(h) dW_Q(h),
\end{aligned}$$

where in the last line we used the Stochastic Fubini Theorem 1.104 to change the order of integration (see Theorem 4.33, p. 110 in [130] or Theorem 2.8, p. 57 [220], observing that the required hypotheses are satisfied thanks to (1.24)), and (1.26). $\square$

LEMMA 1.110 *Let A be the generator of a C_0-semigroup $\{e^{rA},\ r \geq 0\}$ on H, $T > 0$, $t \in [0,T)$ and $f \in L^p(t,T;H)$, $p \geq 1$. Then:*

(i) *If either $1/p < \alpha \leq 1$, or $p = \alpha = 1$, then the function*

$$G_\alpha f(s) := \int_t^s (s-r)^{\alpha-1} e^{(s-r)A} f(r) dr$$

is in $C([t,T]; H)$.

(ii) *If the semigroup e^{tA} is analytic, $\lambda \in \mathbb{R}$ is such that that $(\lambda I - A)^{-1} \in \mathcal{L}(H)$, $\beta > 0$ and $\alpha > \beta + 1/p$, then the function*

$$G_{\alpha,\beta} f(s) := \int_t^s (s-r)^{\alpha-1} (\lambda I - A)^\beta e^{(s-r)A} f(r) dr$$

is in $C([t,T]; H)$.

PROOF. Part (i): Let $1/p < \alpha \le 1$. Let $t \le s_1 \le s_2 \le T$ and denote $h = s_2 - s_1$. We have

$$\left| \int_t^{s_2} (s_2 - r)^{\alpha-1} e^{(s_2-r)A} f(r) dr - \int_t^{s_1} (s_1 - r)^{\alpha-1} e^{(s_1-r)A} f(r) dr \right|$$
$$\le I_1 + I_2 := \int_t^{t+h} \left| (s_2 - r)^{\alpha-1} e^{(s_2-r)A} f(r) \right| dr$$
$$+ \left| \int_{t+h}^{s_2} (s_2 - r)^{\alpha-1} e^{(s_2-r)A} f(r) dr - \int_t^{s_1} (s_1 - r)^{\alpha-1} e^{(s_1-r)A} f(r) dr \right|.$$

Set $q := \frac{p}{p-1}$ and let $R > 0$ be such that $\left\| e^{sA} \right\| \le R$ for all $s \in [0, T]$. Then

$$I_1 \le R \left(\int_0^h (h - r)^{q(\alpha-1)} dr \right)^{1/q} \left(\int_t^T |f(r)|^p dr \right)^{1/p} \to 0 \text{ as } h \to 0$$

since $0 \ge q(\alpha - 1) > -1$. As regards I_2, after a change of variables we have

$$I_2 \le \int_t^{s_1} (s_1 - r)^{\alpha-1} e^{(s_1-r)A} |f(r + h) - f(r)| dr$$
$$\le R \left(\int_t^T (T - r)^{q(\alpha-1)} dr \right)^{1/q} \left(\int_t^{T-h} |f(r + h) - f(r)|^p dr \right)^{1/p} \to 0 \text{ as } h \to 0.$$

The proof in the case $p = \alpha = 1$ is straightforward.

Part (ii) follows from Proposition A.1.1 in Appendix A, p. 307 of [127]. □

PROPOSITION 1.111 *Let $T > 0$ and $t \in [0, T)$. Let A, A_1, Φ be as in Lemma 1.109. Assume that there exist $0 < \alpha < 1$, $C > 0$ and $p > \frac{1}{\alpha}, p \ge 2$ such that*

$$\int_t^T \mathbb{E} \left(\int_t^s \|(s - r)^{-\alpha} A_1 e^{(s-r)A} \Phi(r)\|^2_{\mathcal{L}_2(\Xi_0;H)} dr \right)^{p/2} ds < C. \tag{1.27}$$

Then

$$\Psi(s) := \int_t^s A_1 e^{(s-r)A} \Phi(r) dW_Q(r)$$

has a continuous modification.

PROOF. We follow the scheme of the proof of Theorem 5.2.6 in [127] (page 59, Section 5.2.1). We give some details because our claim is slightly more general. Observe that using Hölder's and Jensen's inequalities we obtain

$$\int_t^s (s - r)^{\alpha-1} \left(\int_t^r (r - h)^{-2\alpha} \mathbb{E} \left[\left\| A_1 e^{(r-h)A} \Phi(h) \right\|^2_{\mathcal{L}_2(\Xi_0;H)} \right] dh \right)^{1/2} dr$$
$$\le \left(\int_t^s (s - r)^{\frac{(\alpha-1)p}{p-1}} \right)^{\frac{p-1}{p}} \left(\int_t^s \mathbb{E} \left(\int_t^r (r - h)^{-2\alpha} \left\| A_1 e^{(r-h)A} \Phi(h) \right\|^2_{\mathcal{L}_2(\Xi_0;H)} dh \right)^{p/2} \right)^{\frac{1}{p}}$$
$$< +\infty,$$

where we used (1.27) and that $\frac{(1-\alpha)p}{p-1} < 1$ which follows from $p > 1/\alpha$. Therefore the hypotheses of Lemma 1.109 are satisfied and thus we have

$$\int_t^s A_1 e^{(s-r)A} \Phi(r) dW_Q(r) = \frac{\sin(\alpha\pi)}{\pi} \int_t^s (s - r)^{\alpha-1} e^{(s-r)A} Y_\alpha^\Phi(r) dr \qquad \mathbb{P} - a.s.$$

for all $s \in [t, T]$, where $Y_\alpha^\Phi(r)$ is defined by (1.25). The claim will follow from Lemma 1.110-(i) applied to a.e. trajectory. Thus we need to know that the process

$Y_\alpha^\Phi(r)$ has p-integrable trajectories a.s.. This is guaranteed if

$$\mathbb{E}\int_t^T \left(\left[\left|Y_\alpha^\Phi(s)\right|^p\right]\right) ds < +\infty.$$

However, from Theorem 1.106, we have

$$\int_t^T \mathbb{E}\left(\left[\left|Y_\alpha^\Phi(s)\right|^p\right]\right) ds \le c_p \int_t^T \mathbb{E}\left(\int_t^s \|(s-r)^{-\alpha} A_1 e^{(s-r)A}\Phi(r)\|^2_{\mathcal{L}_2(\Xi_0;H)} dr\right)^{p/2} ds \tag{1.28}$$

that is bounded thanks to (1.27). □

1.4. Stochastic differential equations

In this section we take $T > 0$ and take H, Ξ, Q, and a generalized reference probability space $\mu = (\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s\in[0,T]}, \mathbb{P}, W_Q)$ as in Section 1.3 (with $t = 0$). A is the infinitesimal generator of a C_0-semigroup on H, and Λ is a Polish space. We will look at stochastic differential equations (SDE) on the interval $[0, T]$, however all results are the same if, instead of $[0, T]$, we took an interval $[t, T]$, for $0 \le t < T$.

1.4.1. Mild and strong solutions. Let $b\colon [0,T]\times H\times\Omega \to H$ and $\sigma\colon [0,T]\times H\times\Omega \to \mathcal{L}_2(\Xi_0, H)$. We consider the following general stochastic differential equation (SDE)

$$\begin{cases} dX(s) = (AX(s) + b(s, X(s)))ds + \sigma(s, X(s))dW_Q(s) & s \in (0, T] \\ X(0) = \xi, \end{cases} \tag{1.29}$$

where ξ is an H-valued $\mathscr{F}_0$-measurable random variable. To simplify the notation we dropped the ω variable in (1.29) and we use this convention throughout the section.

Definition 1.112 (Strong solution of (1.29)) *An H-valued progressively measurable process $X(\cdot)$ is called a* strong solution *of (1.29) if:*

(i) *For $dt \otimes \mathbb{P}$-a.e. $(s, \omega) \in [0, T] \times \Omega$, $X(s)(\omega) \in D(A)$.*

(ii)

$$\mathbb{P}\left(\int_0^T |X(s)| + |AX(s)| + |b(s, X(s))| ds < +\infty\right) = 1,$$

$$\mathbb{P}\left(\int_0^T \|\sigma(s, X(s))\|^2_{\mathcal{L}_2(\Xi_0,H)} ds < +\infty\right) = 1.$$

(iii) *For every $s \in [0, T]$*

$$X(s) = \xi + \int_0^s AX(r) + b(r, X(r))dr + \int_0^s \sigma(r, X(r))dW_Q(s) \ \ \mathbb{P}\text{-a.e.}.$$

Definition 1.113 (Mild solution of (1.29)) *An H-valued progressively measurable process $X(\cdot)$ is called a* mild solution *of (1.29) if:*

(i)

$$\mathbb{P}\left(\int_0^t |X(s)| + |e^{(t-s)A} b(s, X(s))| ds < +\infty\right) = 1, \qquad \forall t \in [0, T]$$

$$\mathbb{P}\left(\int_0^t \|e^{(t-s)A}\sigma(s, X(s))\|^2_{\mathcal{L}_2(\Xi_0,H)} ds < +\infty\right) = 1, \qquad \forall t \in [0, T].$$

(ii) *For every $s \in [0, T]$*

$$X(s) = e^{sA}\xi + \int_0^s e^{(s-r)A} b(r, X(r))dr + \int_0^s e^{(s-r)A}\sigma(r, X(r))dW_Q(r) \ \ \mathbb{P}\text{-a.e.}.$$

In order for the above definitions to be meaningful, all the processes involved must be well defined and have proper measurability properties so that the integrals that appear in the definitions make sense. We do not want to analyze here the required measurability properties in the most generality. We discuss one case which will frequently appear in applications to optimal control in Remark 1.117 below. Moreover, note that if A_n is the Yosida approximation of A, since by Lemma 1.17-(*i*) $D(A) \in \mathcal{B}(H)$, it follows that the processes $\mathbf{1}_{X(s)\in D(A)} A_n X(s)$ are progressively measurable and they converge as $n \to +\infty$ to $\mathbf{1}_{X(s)\in D(A)} AX(s)$ for every (s, ω). Thus the process $AX(s)$ (understood as $\mathbf{1}_{X(s)\in D(A)} AX(s)$) is progressively measurable.

REMARK 1.114 In the definition of mild solution we assumed that $b\colon [0,T] \times H \times \Omega \to H$ and $\sigma\colon [0,T] \times H \times \Omega \to \mathcal{L}_2(\Xi_0, H)$. However, Definition 1.113 may still make sense even if b and σ do not have values in H and $\mathcal{L}_2(\Xi_0, H)$, provided that the terms $e^{(t-s)A} b(s, X(s))$ and $e^{(t-s)A}\sigma(s, X(s))$ have values in these spaces when they are interpreted properly (see for instance Section 1.5.1 and also Remark 1.117). Therefore in the future when we are dealing with such cases, we will not repeat the definition of mild solution, instead we will just say how to interpret the above terms. ■

DEFINITION 1.115 (Weak mild solution of (1.29)) *A weak mild solution of (1.29) is defined to be any* 6*-tuple* $(\Omega, \mathscr{F}, \mathscr{F}_s, W_Q, \mathbb{P}, X(\cdot))$*, where* $(\Omega, \mathscr{F}, \mathscr{F}_s, \mathbb{P})$ *is a filtered probability space,* W_Q *is a translated* $\mathscr{F}_s$*-*Q*-Wiener process on* Ω*, and* $X(\cdot)$ *is a mild solution for (1.29) in the generalized reference probability space* $(\Omega, \mathscr{F}, \mathscr{F}_s, W_Q, \mathbb{P})$*.*

NOTATION 1.116 In the existing literature, often different authors call the same notion of solution differently, and the same name does not always correspond to the same definition. For instance, the *weak mild solution* introduced above is often called a weak solution and in [130, Chapter 8] it is called a *martingale solution*. ■

REMARK 1.117 Let Λ be a Polish space. Suppose that $\sigma : [0,T] \times H \times \Lambda \to \mathcal{L}(\Xi_0, H)$ is such that for every $u \in \Xi_0$, the map $(t, x, a) \to \sigma(t, x, a)u$ is $\mathcal{B}([0,T]) \otimes \mathcal{B}(H) \otimes \mathcal{B}(\Lambda)/\mathcal{B}(H)$ measurable, and $e^{sA}\sigma(t, x, a) \in \mathcal{L}_2(\Xi_0, H)$ for every (t, x, a) and $s > 0$. It then follows from Lemma 1.19 that, after possibly redefining it at $s = 0$, the map $(s, t, x, a) \to e^{sA}\sigma(t, x, a)$ is $\mathcal{B}([0,T]) \otimes \mathcal{B}([0,T]) \otimes \mathcal{B}(H) \otimes \mathcal{B}(\Lambda)/\mathcal{B}(\mathcal{L}_2(\Xi_0, H))$ measurable. Now, if $X(\cdot) : [0,T] \times \Omega \to H, a(\cdot) : [0,T] \times \Omega \to \Lambda$ are $\mathscr{F}_s$-progressively measurable, then for every $t \in [0,T]$,

$$(s, \omega) \to e^{(t-s)A}\sigma(s, X(s), a(s))$$

is an $\mathcal{L}_2(\Xi_0, H)$-valued $\mathscr{F}_s$-progressively measurable process on $[0,t] \times \Omega$. If this process is in $\mathcal{N}^2_Q(0,t;H)$ for every t then the process

$$(t, \omega) \to \int_0^t e^{(t-s)A}\sigma(s, X(s), a(s))dW_Q(s)$$

is an H-valued $\mathscr{F}_t$ adapted process. Denote

$$G(t, s, \omega) := \mathbf{1}_{[0,t]}(s)e^{(t-s)A}\sigma(s, X(s), a(s)).$$

G is $\mathcal{B}([0,T]) \otimes \mathcal{B}([0,T]) \otimes \mathscr{F}_T/\mathcal{B}(\mathcal{L}_2(\Xi_0, H))$ measurable. Suppose that

$$\int_0^T \left(\mathbb{E} \int_0^T \|G(t, s, \omega)\|^2_{\mathcal{L}_2(\Xi_0,H)} ds \right)^{\frac{1}{2}} dt < +\infty.$$

Then, by (i) of Theorem 1.104, the process

$$(t,\omega) \to \int_0^T G(t,s,\omega)dW_Q(s)$$

has a $\mathcal{B}([0,T]) \otimes \mathscr{F}_T/\mathcal{B}(H)$ measurable modification on $[0,T]\times\Omega$. Therefore, by Lemma 1.69, the process

$$(t,\omega) \to \int_0^t e^{(t-s)A}\sigma(s,X(s),a(s))dW_Q(s) = \int_0^T G(t,s,\omega)dW_Q(s)$$

has an $\mathscr{F}_t$-progressively measurable modification. ∎

1.4.2. Existence and uniqueness of solutions.

DEFINITION 1.118 (The space $M^p_\mu(t,T;E)$) *In this definition $T \in (0,+\infty) \cup \{+\infty\}$. Let $p \geq 1$ and $0 \leq t < T$. Given a Banach space E, we denote by $M^p_\mu(t,T;E)$ the space of all E-valued progressively measurable processes $X(\cdot)$ such that*

$$|X(\cdot)|_{M^p_\mu(t,T;E)} := \left(\mathbb{E}\left(\int_t^T |X(s)|^p ds\right)\right)^{1/p} < +\infty. \tag{1.30}$$

$M^p_\mu(t,T;E)$ is a Banach space endowed with the norm $|\cdot|_{M^p_\mu(t,T;E)}$.

Note that in the notation $M^p_\mu(t,T;E)$ we emphasize the dependence on the generalized reference probability space μ. Processes in $M^p_\mu(t,T;E)$ are identified if they are equal $\mathbb{P}\otimes dt$ a.e..

Let $a\colon [0,T] \to \Lambda$ be an $\mathscr{F}_s$-progressively measurable process (a control process). (Recall that Λ is a Polish space.) We consider the controlled SDE

$$\begin{cases} dX(s) = (AX(s) + b(s,X(s),a(s)))\,ds + \sigma(s,X(s),a(s))dW_Q(s) \\ X(0) = \xi. \end{cases} \tag{1.31}$$

This equation falls into the category of equations (1.29) with $b(s,x,\omega) := b(s,x,a(s,\omega))$ and $\sigma(s,x,\omega) := \sigma(s,x,a(s,\omega))$. Thus strong, mild and weak mild solutions of (1.31) are defined using the definitions for equation (1.29).

HYPOTHESIS 1.119 *The operator A is the generator of a strongly continuous semigroup e^{sA} on H. The function $b\colon [0,T]\times H\times\Lambda \to H$ is $\mathcal{B}([0,T])\otimes\mathcal{B}(H)\otimes\mathcal{B}(\Lambda)/\mathcal{B}(H)$ measurable, $\sigma\colon [0,T]\times H\times\Lambda \to \mathcal{L}_2(\Xi_0,H)$ is $\mathcal{B}([0,T])\otimes\mathcal{B}(H)\otimes\mathcal{B}(\Lambda)/\mathcal{B}(\mathcal{L}_2(\Xi_0,H))$ measurable, and there exists a constant $C>0$ such that*

$$|b(s,x,a) - b(s,y,a)| \leq C|x-y| \qquad \forall x,y\in H,\ s\in[0,T],\ a\in\Lambda, \tag{1.32}$$

$$\|\sigma(s,x,a) - \sigma(s,y,a)\|_{\mathcal{L}_2(\Xi_0,H)} \leq C|x-y| \quad \forall x,y\in H,\ s\in[0,T],\ a\in\Lambda, \tag{1.33}$$

$$|b(s,x,a)| \leq C(1+|x|) \qquad \forall x\in H,\ s\in[0,T],\ a\in\Lambda, \tag{1.34}$$

$$\|\sigma(s,x,a)\|_{\mathcal{L}_2(\Xi_0,H)} \leq C(1+|x|) \qquad \forall x\in H,\ s\in[0,T],\ a\in\Lambda. \tag{1.35}$$

DEFINITION 1.120 (The space $\mathcal{H}^\mu_p(t,T;E)$) *Let $p\geq 1$ and $0\leq t<T$. Given a Banach space E, we denote by $\mathcal{H}^\mu_p(t,T;E)$ the set of all progressively measurable process $X\colon [t,T]\times\Omega \to E$ such that*

$$|X(\cdot)|_{\mathcal{H}^\mu_p(t,T;E)} := \left(\sup_{s\in[t,T]} \mathbb{E}|X(s)|^p\right)^{1/p} < +\infty.$$

It is a Banach space with the norm $|\cdot|_{\mathcal{H}^\mu_p(t,T;E)}$.

Processes in $\mathcal{H}^\mu_p(t,T;E)$ are identified if they are equal $\mathbb{P}\otimes dt$ a.e.. Therefore the *sup* in the definition of $\mathcal{H}^\mu_p(t,T;E)$ must be understood as *esssup*. However we will keep the notation *sup* here and in the all subsequent uses of this space. If the generalized reference probability space μ is clear we will just write $M^p(t,T;E)$ and $\mathcal{H}_p(t,T;E)$ for simplicity.

THEOREM 1.121 *Let $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ for some $p \geq 2$, and let A, b and σ satisfy Hypothesis 1.119. Let $a(\cdot)\colon [0,T] \to \Lambda$ be an $\mathscr{F}_s$-progressively measurable process. Then the SDE (1.31) has a unique, up to a modification, mild solution $X(\cdot) \in \mathcal{H}_p(0,T;H)$. The solution is in fact unique, among all processes such that $\mathbb{P}\left(\int_0^T |X(s)|^2 ds < +\infty\right) = 1$, in particular among the processes in $M^2_\mu(0,T;H)$. $X(\cdot)$ has a continuous modification. Given two continuous versions $X_1(\cdot)$, $X_2(\cdot)$ of the solution, there exists $\tilde{\Omega} \subset \Omega$ with $\mathbb{P}(\tilde{\Omega}) = 1$ s.t. $X_1(s) = X_2(s)$ for all $s \in [0,T]$ and $\omega \in \tilde{\Omega}$.*

PROOF. The proof can be found for instance in [130], Theorem 7.2, page 188 or [220], Theorems 3.3, page 97, and 3.5, page 105. For the last claim, we can take

$$\tilde{\Omega} := \bigcap_{s \in \mathbb{Q} \cap [0,T]} \{\omega \in \Omega \ : \ X_1(s)(\omega) = X_2(s)(\omega)\}.$$

Since $X_1(\cdot)$ is a modification of $X_2(\cdot)$, we have $\mathbb{P}(\tilde{\Omega}) = 1$, and since $X_1(\cdot)$ and $X_2(\cdot)$ are continuous, it follows that $X_1(s)(\omega) = X_2(s)(\omega)$ for all $s \in [0,T]$, $\omega \in \tilde{\Omega}$. □

We will denote the solution of (1.31) by $X(\cdot;\xi,a(\cdot))$ if we want to emphasize the dependence on the initial datum and the control.

COROLLARY 1.122 *Let $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ for some $p \geq 2$, let A, b and σ satisfy Hypothesis 1.119. If $a_1(\cdot), a_2(\cdot)\colon [0,T] \times \Omega \to \Lambda$ are two progressively measurable processes such that $a_1(\cdot) = a_2(\cdot)$, $dt \otimes \mathbb{P}$-a.e. on $[0,T] \times \Omega$, then, $\mathbb{P}-a.e.$,*

$$X\left(s;\xi,a_1(\cdot)\right) = X\left(s;\xi,a_2(\cdot)\right) \ \text{for all } s \in [0,T].$$

PROOF. Denote $X_i(\cdot) := X\left(\cdot;\xi,a_i(\cdot)\right)$. Using Theorem 1.98, Jensen's inequality, and $\sup_{s\in[0,T]} \|e^{sA}\| \leq C$ for some $C \geq 0$, it follows that, for suitable positive C_1 and C_2:

$$\begin{aligned}\mathbb{E}\left[|X_1(s) - X_2(s)|^2\right] &\leq C_1\bigg(\int_0^s \mathbb{E}|b(r,X_1(r),a_1(r)) - b(r,X_2(r),a_2(r))|^2 dr \\ &\quad + \int_0^s \mathbb{E}\|\sigma(r,X_1(r),a_1(r)) - \sigma(r,X_2(r),a_2(r))\|^2_{\mathcal{L}_2(\Xi_0;H)} dr\bigg) \\ &\leq C_2 \int_0^s \mathbb{E}|X_1(r) - X_2(r)|^2 dr, \end{aligned} \quad (1.36)$$

and the claim follows using Gronwall's lemma and the continuity of the trajectories. □

REMARK 1.123 Above we assumed that the σ always takes values in $\mathcal{L}_2(\Xi_0, H)$. Existence and uniqueness results for SDEs with more general σ can be found for instance in [220] Theorem 3.15, page 143, or in [130] Theorem 7.5, page 197. To treat some specific examples we will also prove more general results in Section 1.5. ■

1.4.3. Properties of solutions of SDE.

THEOREM 1.124 *Let $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ for some $p \geq 2$, $a\colon [0,T] \to \Lambda$ be $\mathscr{F}_s$-progressively measurable, and let A, b and σ satisfy Hypothesis 1.119.*

(i) *Let $X(\cdot) = X(\cdot;\xi, a(\cdot))$ be the unique mild solution of (1.31) (provided by Theorem 1.121). Then*

$$\sup_{s\in[0,T]} \mathbb{E}\left[|X(s)|^p\right] \le C_p(T)(1+\mathbb{E}|\xi|^p) \quad \text{if } p \ge 2, \tag{1.37}$$

$$\mathbb{E}\left[\sup_{s\in[0,T]} |X(s)|^p\right] \le C_p(T)(1+\mathbb{E}|\xi|^p) \quad \text{if } p > 2, \tag{1.38}$$

and

$$\mathbb{E}\left[\sup_{r\in[0,s]} |X(r)-\xi|^p\right] \le \omega_\xi(s) \quad \text{if } p > 2, \tag{1.39}$$

where $C_p(T)$ is a constant depending on p, T, C (from Hypothesis 1.119) and M, α (where $\|e^{rA}\| \le Me^{r\alpha}$ for $r \ge 0$), and ω_ξ is a modulus depending on the same constants and on ξ (in particular they are independent of the process $a(\cdot)$ and of the generalized reference probability space).

(ii) *If $\xi, \eta \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ for $p > 2$, and $X(\cdot) = X(\cdot;\xi, a(\cdot)), Y(\cdot) = Y(\cdot;\eta, a(\cdot))$ are the solutions of (1.31), then*

$$\mathbb{E}\left[\sup_{s\in[0,T]} |X(s)-Y(s)|^2\right] \le C_T\left(\mathbb{E}\left[|\xi-\eta|^p\right]\right)^{\frac{2}{p}}, \tag{1.40}$$

where C_T depends only on p, T, C, M, α.

Proof. *Part (i)*: For (1.37) and (1.38) we refer for instance to [130] Theorem 9.1, page 235, or [220], Lemma 3.6, page 102, and Corollary 3.3, page 104. Regarding (1.39), we have that there is a constant c_1 depending only on p and $\sup_{t\in[0,T]}\|e^{tA}\|$, such that

$$\begin{aligned}\mathbb{E}\left[\sup_{r\in[0,s]} |X(r)-\xi|^p\right] \le c\Bigg(&\mathbb{E}\left[\sup_{r\in[0,s]} \left|e^{rA}\xi-\xi\right|^p\right]\\ &+\mathbb{E}\left[\sup_{r\in[0,s]}\left(\int_0^r |b(u,X(u),a(u))|du\right)^p\right]\\ &+\mathbb{E}\left[\sup_{r\in[0,s]}\left|\int_0^r e^{(r-u)A}\sigma(u,X(u),a(u))dW_Q(u)\right|^p\right]\Bigg).\end{aligned}$$

Using Hypothesis 1.119, (1.38), Hölder inequality, and Theorem 1.107, we see that

$$\mathbb{E}\left[\sup_{r\in[0,s]} |X(r)-\xi|^p\right] \le c_2\Bigg(\mathbb{E}\left[\sup_{r\in[0,s]} \left|e^{rA}\xi-\xi\right|^p\right] + \int_0^s (1+\mathbb{E}|\xi|^p)\,dr\Bigg).$$

Since $\sup_{r\in[0,s]}\left|e^{rA}\xi-\xi\right|^p \xrightarrow{s\to 0^+} 0$ a.e., and $\sup_{r\in[0,s]}\left|e^{rA}\xi-\xi\right|^p \le C_1|\xi|^p$, we conclude by the Lebesgue dominated convergence theorem.

Part (ii): See [130] Theorem 9.1, page 235. □

Theorem 1.125 *Let $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ for some $p > 2$, and let A, b and σ satisfy Hypothesis 1.119. Let $a\colon [0,T] \to \Lambda$ be a progressively measurable process. Let $X(\cdot)$ be the unique mild solution of (1.31). Consider the approximating equations*

$$\begin{cases} dX^n(s) = (A_nX^n(s) + b(s,X^n(s),a(s)))\,ds + \sigma(s,X^n(s),a(s))dW_Q(s)\\ X^n(0) = \xi, \end{cases} \tag{1.41}$$

where A_n is the Yosida approximation of A. Let $X_n(\cdot)$ be the solution of (1.41). Then

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{s\in[0,T]} |X^n(s) - X(s)|^p\right] = 0. \tag{1.42}$$

PROOF. See [130] Proposition 7.4, page 196, or [220], Proposition 3.2, page 101. □

The next proposition is a simpler version of Theorem 1.125 which will be useful to simplify the proofs of the results of Section 1.7.

PROPOSITION 1.126 *Let $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$, $f \in M^p_\mu(0,T;H)$, and $\Phi \in \mathcal{N}^p_Q(0,T;H)$ for some $p \geq 2$. Let $X(\cdot)$ be the mild solution of*

$$\begin{cases} dX(s) = (AX(s) + f(s))\, ds + \Phi(s) dW_Q(s) \\ X(0) = \xi \end{cases} \tag{1.43}$$

and $X^n(\cdot)$ be the solution of

$$\begin{cases} dX^n(s) = (A_n X^n(s) + f(s))\, ds + \Phi(s) dW_Q(s) \\ X^n(0) = \xi, \end{cases} \tag{1.44}$$

where A generates a C_0-semigroup and A_n is the Yosida approximation of A. Then, if $p > 2$,

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{s\in[0,T]} |X^n(s) - X(s)|^p\right] = 0. \tag{1.45}$$

Moreover, for $p \geq 2$, there exists $M > 0$, independent of n, such that

$$\sup_{s\in[0,T]} \mathbb{E}\left[|X^n(s)|^p\right] \leq M, \qquad \sup_{s\in[0,T]} \mathbb{E}\left[|X(s)|^p\right] \leq M. \tag{1.46}$$

PROOF. Observe first that the mild solution of (1.43) is well defined thanks to the assumptions on ξ, f and Φ, and

$$X(s) = e^{sA}\xi + \int_0^s e^{(s-r)A} f(r) dr + \int_0^s e^{(s-r)A}\Phi(r) dW_Q(r).$$

The same is true for the mild solution of (1.44) (which is also a strong solution).

To prove (1.45), we write

$$\begin{aligned} X^n(s) - X(s) &= \left(e^{sA_n} - e^{sA}\right)\xi + \int_0^s \left(e^{(s-r)A_n} - e^{(s-r)A}\right) f(r) dr \\ &+ \int_0^s \left(e^{(s-r)A_n} - e^{(s-r)A}\right) \Phi(r) dW_Q(r) =: I_1^n(s) + I_2^n(s) + I_3^n(s). \end{aligned} \tag{1.47}$$

It is enough to show that $\lim_{n\to\infty} \mathbb{E}\left[\sup_{s\in[0,T]} |I_i^n(s)|^p\right] = 0$ for $i \in \{1,2,3\}$. For $i = 3$ this follows from (1.21). To prove it for $i = 2$, we notice that (B.12) implies that if

$$\psi_n(r) := \sup_{s\in[r,T]} \left|\left(e^{(s-r)A_n} - e^{(s-r)A}\right) f(r)\right|,$$

then $\psi_n(r) \xrightarrow{n\to\infty} 0$ a.e. on Ω. Moreover, thanks to (B.11) there exists C_1 such that, for all $t \in [0,T]$ and all n, $\left\|e^{tA_n}\right\| \leq C_1$, so $\psi_n(r) \leq 2C_1 |f(r)|$ for all n. Since $\int_t^T |f(r)| dr < +\infty$ for almost every $\omega \in \Omega$, by the Lebesgue dominated convergence theorem we have

$$\sup_{s\in[0,T]} \left| \int_0^s \left| \left(e^{(s-r)A_n} - e^{(s-r)A} \right) f(r) \right| dr \right|^p$$
$$\le \sup_{s\in[0,T]} \left| \int_0^s \psi_n(r) dr \right|^p \le \left| \int_0^T \psi_n(r) dr \right|^p \xrightarrow{n\to\infty} 0$$

for a.e. $\omega \in \Omega$. Now observe that

$$\sup_{s\in[0,T]} \left| \int_0^s \left| \left(e^{(s-r)A_n} - e^{(s-r)A} \right) f(r) \right| dr \right|^p$$
$$\le \sup_{s\in[0,T]} \int_0^s (2C_1)^p |f(r)|^p \, dr \le \int_0^T (2C_1)^p |f(r)|^p \, dr,$$

and the last expression is integrable (on Ω), since $f \in M^p_\mu(0,T;H)$. Therefore we can apply the Lebesgue dominated convergence theorem, obtaining $\lim_{n\to\infty} \mathbb{E}\left[\sup_{s\in[0,T]} |I_2^n(s)|^p\right] = 0$. The claim for $i = 1$ follows again from (B.12) and the Lebesgue dominated convergence theorem.

Estimates (1.46) are easy consequences of (B.11) and the assumptions on ξ, f, Φ. □

1.4.4. Uniqueness in law.

DEFINITION 1.127 (Finite dimensional distributions) *Let $T > 0$ and $t \in [0,T)$. Consider a measurable space $(\Omega, \mathscr{F})$, two probability spaces $(\Omega_i, \mathscr{F}_i, \mathbb{P}_i)$ for $i = 1, 2$, and two processes $\{X_i(s)\}_{s\in[t,T]} : (\Omega_i, \mathscr{F}_i, \mathbb{P}_i) \to (\Omega, \mathscr{F})$. We say that $X_1(\cdot)$ and $X_2(\cdot)$ have the same* finite dimensional distributions *on $D \subset [t,T]$ if for any $t \le t_1 < t_2 < ... < t_n \le T, t_i \in D$ and $A \in \underbrace{\mathscr{F} \otimes \mathscr{F} \otimes ... \otimes \mathscr{F}}_{n\ times}$, we have*

$$\mathbb{P}_1 \{\omega_1 \ : \ (X_1(t_1), ..X_1(t_n))(\omega_1) \in A\} = \mathbb{P}_2 \{\omega_2 \ : \ (X_2(t_1), ..X_2(t_n))(\omega_2) \in A\}.$$

In this case we write $\mathcal{L}_{\mathbb{P}_1}(X_1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(X_2(\cdot))$ on D. Often we will just write $\mathcal{L}_{\mathbb{P}_1}(X_1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(X_2(\cdot))$ which should be understood that the finite dimensional distributions are the same on some set of full measure.

THEOREM 1.128 *Let H be a separable Hilbert space. Let $(\Omega_i, \mathscr{F}_i, \mathbb{P}_i)$ for $i = 1, 2$ two complete probability spaces, and $(\tilde{\Omega}, \tilde{\mathscr{F}})$ a measurable space. Let $\xi_i : \Omega_i \to \tilde{\Omega}, i = 1, 2$ be two random variables, and $f_i : [t,T] \times \Omega_i \to H, i = 1, 2$ two processes satisfying*

$$\mathbb{P}_1 \left(\int_t^T |f_1(s)| ds < +\infty \right) = \mathbb{P}_2 \left(\int_t^T |f_2(s)| ds < +\infty \right) = 1$$

and, for some subset $D \subset [t,T]$ of full measure,

$$\mathcal{L}_{\mathbb{P}_1}(f_1(\cdot), \xi_1) = \mathcal{L}_{\mathbb{P}_2}(f_2(\cdot), \xi_2) \quad \text{on } D.$$

Then

$$\mathcal{L}_{\mathbb{P}_1} \left(\int_t^{\cdot} f_1(s) ds, \xi_1 \right) = \mathcal{L}_{\mathbb{P}_2} \left(\int_t^{\cdot} f_2(s) ds, \xi_2 \right) \quad \text{on } [t,T]. \tag{1.48}$$

PROOF. See [368] Theorem 8.3, where the theorem was proved for a more general case of Banach space valued processes. □

THEOREM 1.129 *Let* $\left(\Omega_1, \mathscr{F}_1, \mathscr{F}_s^{1,t}, \mathbb{P}_1, W_{Q,1}\right)$ *and* $\left(\Omega_2, \mathscr{F}_2, \mathscr{F}_s^{2,t}, \mathbb{P}_2, W_{Q,2}\right)$ *be two generalized reference probability spaces. Let* $\Phi_i\colon [t,T]\times\Omega_i \to \mathcal{L}_2(\Xi_0; H)$, $i = 1, 2$ *be two* $\mathscr{F}_s^{i,t}$*-progressively measurable processes satisfying*

$$\mathbb{P}_1\left(\int_t^T \|\Phi_1(s)\|^2_{\mathcal{L}_2(\Xi_0;H)} ds < +\infty\right) = \mathbb{P}_2\left(\int_t^T \|\Phi_2(s)\|^2_{\mathcal{L}_2(\Xi_0;H)} ds < +\infty\right) = 1.$$

Let $(\tilde{\Omega}, \tilde{\mathscr{F}})$ *a measurable space and* $\xi_i\colon \Omega_i \to \tilde{\Omega}, i = 1, 2$ *be two random variables. Assume that, for some subset* $D \subset [t,T]$ *of full measure,*

$$\mathcal{L}_{\mathbb{P}_1}\left(\Phi_1(\cdot), W_{Q,1}(\cdot), \xi_1\right) = \mathcal{L}_{\mathbb{P}_2}\left(\Phi_2(\cdot), W_{Q,2}(\cdot), \xi_2\right) \text{ on } D.$$

Then

$$\mathcal{L}_{\mathbb{P}_1}\left(\int_t^{\cdot} \Phi_1(s) dW_{Q,1}(s), \xi_1\right) = \mathcal{L}_{\mathbb{P}_2}\left(\int_t^{\cdot} \Phi_2(s) dW_{Q,2}(s), \xi_2\right) \quad \text{on } [t,T]. \tag{1.49}$$

PROOF. See [368] Theorem 8.6. □

Consider now an operator A and mappings b, σ satisfying Hypothesis 1.119, and $x \in H$. Let $\left(\Omega_1, \mathscr{F}_1, \mathscr{F}_s^{1,t}, \mathbb{P}_1, W_{Q,1}\right)$ and $\left(\Omega_2, \mathscr{F}_2, \mathscr{F}_s^{2,t}, \mathbb{P}_2, W_{Q,2}\right)$ be as in Theorem 1.129. For $i = 1, 2$ consider an $\mathscr{F}_s^{i,t}$-progressively measurable process $a_i\colon [t,T]\times\Omega_i \to \Lambda$.

Let $p > 2$ and let $\zeta_i \in L^p(\Omega_i, \mathscr{F}_t^{i,t}, \mathbb{P}_i)$, $i = 1, 2$. Denote by $\mathcal{H}_{p,i}$ the Banach space of all $\mathscr{F}_s^{i,t}$-progressively measurable processes $Z_i\colon [t,T]\times\Omega_i \to H$ such that

$$\left(\sup_{s\in[t,T]} \mathbb{E}_i |Z_i(s)|^p\right)^{1/p} < +\infty.$$

Let $\mathcal{K}_i\colon \mathcal{H}_{p,i} \to \mathcal{H}_{p,i}$ be the continuous map (see [130] pages 189) defined as

$$\mathcal{K}_i(Z_i(\cdot))(s) := e^{(s-t)A}\zeta_i + \int_t^s e^{(s-r)A} b(r, Z_i(r), a_i(r)) dr + \int_t^s e^{(s-r)A}\sigma(r, Z_i(r), a_i(r)) dW_{Q,i}(r). \tag{1.50}$$

LEMMA 1.130 *Consider the setting described above, and let* $\theta_i : [t,T]\times\Omega_i \to H, i = 1, 2$ *be stochastic processes. If*

$$\mathcal{L}_{\mathbb{P}_1}(Z_1(\cdot), a_1(\cdot), W_{Q,1}(\cdot), \theta_1(\cdot), \zeta_1) = \mathcal{L}_{\mathbb{P}_2}(Z_2(\cdot), a_2(\cdot), W_{Q,2}(\cdot), \theta_2(\cdot), \zeta_2)$$

on some subset $D \subset [t,T]$ *of full measure, then*

$$\mathcal{L}_{\mathbb{P}_1}(\mathcal{K}_1(Z_1(\cdot))(\cdot), a_1(\cdot), W_{Q,1}(\cdot), \theta_1(\cdot), \zeta_1) = \mathcal{L}_{\mathbb{P}_2}(\mathcal{K}_2(Z_2(\cdot))(\cdot), a_2(\cdot), W_{Q,2}(\cdot), \theta_2(\cdot), \zeta_2) \quad \text{on } D. \tag{1.51}$$

PROOF. Observe that, since we only have to check the finite dimensional distributions, the claims of Theorems 1.128 and 1.129 hold even if ξ_1 and ξ_2 are stochastic processes, with (1.48) and (1.49) then being true on some set of full measure. Let us choose a partition $(t_1, .., t_n)$, with $t \le t_1 < t_2 < ... < t_n \le T, t_k \in D, k = 1, ..., n$. We need to show that

$$\mathcal{L}_{\mathbb{P}_1}(\mathcal{K}_1(Z_1(\cdot))(t_k), a_1(t_k), W_{Q,1}(t_k), \theta_1(t_k), \zeta_1 : k = 1, ..., n) = \mathcal{L}_{\mathbb{P}_2}(\mathcal{K}_2(Z_2(\cdot))(t_k), a_2(t_k), W_{Q,1}(t_k), \theta_2(t_k), \zeta_2 : k = 1, ..., n). \tag{1.52}$$

We have, denoting $f^i(r) := \mathbf{1}_{[t,t_1]}(r)e^{(t_1-r)A}b(r,Z_i(r),a_i(r))$ and $\Phi^i(r) := \mathbf{1}_{[t,t_1]}(r)e^{(t_1-r)A}\sigma(r,Z_i(r),a_i(r)), i=1,2,$

$$\mathcal{L}_{\mathbb{P}_1}(f^1(\cdot),\Phi_1(\cdot),Z_1(\cdot),a_1(\cdot),W_{Q,1}(\cdot),\theta_1(\cdot),\zeta_1) \\ = \mathcal{L}_{\mathbb{P}_2}(f^2(\cdot),\Phi_2(\cdot),Z_2(\cdot),a_2(\cdot),W_{Q,2}(\cdot),\theta_2(\cdot),\zeta_2) \text{ on } D,$$

and thus, by Theorem 1.128 applied with

$$\xi_1(\cdot) = (f^1(\cdot),\Phi^1(\cdot),Z_1(\cdot),a_1(\cdot),W_{Q,1}(\cdot),\theta_1(\cdot),\zeta_1),$$
$$\xi_2(\cdot) = (f^2(\cdot),\Phi^2(\cdot),Z_2(\cdot),a_2(\cdot),W_{Q,2}(\cdot),\theta_2(\cdot),\zeta_2),$$

$$\mathcal{L}_{\mathbb{P}_1}\left(\int_t^{t_1} f^1(s)ds, f^1(\cdot),\Phi^1(\cdot),Z_1(\cdot),a_1(\cdot),W_{Q,1}(\cdot),\theta_1(\cdot),\zeta_1\right) \\ = \mathcal{L}_{\mathbb{P}_2}\left(\int_t^{t_1} f^2(s)ds, f^2(\cdot),\Phi^2(\cdot),Z_2(\cdot),a_2(\cdot),W_{Q,2}(\cdot),\theta_2(\cdot),\zeta_2\right) \text{ on } D.$$

Now, applying Theorem 1.129 with

$$\xi_1(\cdot) = \left(\int_t^{t_1} f^1(s)ds, f^1(\cdot),\Phi^1(\cdot),Z_1(\cdot),a_1(\cdot),W_{Q,1}(\cdot),\theta_1(\cdot),\zeta_1\right),$$
$$\xi_2(\cdot) = \left(\int_t^{t_1} f^2(s)ds, f^2(\cdot),\Phi^2(\cdot),Z_2(\cdot),a_2(\cdot),W_{Q,2}(\cdot),\theta_2(\cdot),\zeta_2\right),$$

we obtain

$$\mathcal{L}_{\mathbb{P}_1}\left(\int_t^{t_1} f^1(s)ds, \int_t^{t_1}\Phi^1(s)dW_{Q,1}(s), f^1(\cdot),\Phi^1(\cdot),Z_1(\cdot),a_1(\cdot),W_{Q,1}(\cdot),\theta_1(\cdot),\zeta_1\right) \\ = \mathcal{L}_{\mathbb{P}_2}\left(\int_t^{t_1} f^2(s)ds, \int_t^{t_1}\Phi^2(s)dW_{Q,2}(s), f^2(\cdot),\Phi^2(\cdot),Z_2(\cdot),a_2(\cdot),W_{Q,2}(\cdot),\theta_2(\cdot),\zeta_2\right)$$

on D. (We recall that the stochastic convolution terms in (1.50) and the stochastic integrals above have continuous trajectories a.e.). In particular this implies that

$$\mathcal{L}_{\mathbb{P}_1}(\mathcal{K}_1(Z_1(\cdot))(t_1), f^1(\cdot),\Phi^1(\cdot),Z_1(\cdot),a_1(\cdot),W_{Q,1}(\cdot),\theta_1(\cdot),\zeta_1) \\ = \mathcal{L}_{\mathbb{P}_2}(\mathcal{K}_2(Z_2(\cdot))(t_1), f^2(\cdot),\Phi^2(\cdot),Z_2(\cdot),a_2(\cdot),W_{Q,2}(\cdot),\theta_2(\cdot),\zeta_2) \text{ on } D.$$

We now repeat the above procedure for $t_2, ..., t_n$ which will yield (1.52) as its consequence. □

PROPOSITION 1.131 *Let the operator A and the mappings b, σ satisfy Hypothesis 1.119. Let $\left(\Omega_1,\mathscr{F}_1,\mathscr{F}_s^{1,t},\mathbb{P}_1,W_{Q,1}\right)$ and $\left(\Omega_2,\mathscr{F}_2,\mathscr{F}_s^{2,t},\mathbb{P}_2,W_{Q,2}\right)$ be two generalized reference probability spaces. Let $a_i\colon [t,T]\times\Omega_i\to\Lambda, i=1,2$ be an $\mathscr{F}_s^{i,t}$-progressively measurable process, and let $\zeta_i\in L^p(\Omega_i,\mathscr{F}_t^{i,t},\mathbb{P}_i), i=1,2, p>2$. Let $\mathcal{L}_{\mathbb{P}_1}(a_1(\cdot),W_{Q,1}(\cdot),\zeta_1) = \mathcal{L}_{\mathbb{P}_2}(a_2(\cdot),W_{Q,1}(\cdot),\zeta_2)$ on some subset $D\subset[0,T]$ of full measure. Denote by $X_i(\cdot), i=1,2,$ the unique mild solution of*

$$\begin{cases} dX_i(s) = (AX_i(s)+b(s,X(s),a_i(s)))\,ds + \sigma(s,X_i(s),a_i(s))dW_{Q,i}(s) \\ X_i(t)=\zeta_i \end{cases} \tag{1.53}$$

on $[t,T]$. Then $\mathcal{L}_{\mathbb{P}_1}(X_1(\cdot),a_1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(X_2(\cdot),a_2(\cdot))$ on D.

PROOF. It is known (see [130], proof of Theorem 7.2, pages 188-193) that the map $\mathcal{K}_i$ is a contraction in $\mathcal{H}_{p,i}$ if $[t,T]$ is small enough. Thus if we divide $[t,T]$ into such small intervals $[t,T_1], ...[T_k,T]$, $X_i(\cdot)$ on $[t,T_1]$ is obtained as the limit in $\mathcal{H}_{p,i}$ (restricted to $[t,T_1]$) of the iterates $(\mathcal{K}_i^n(x))(\cdot)$. Therefore, using Lemma 1.130 and passing to the limit as $n\to+\infty$ we obtain

$$\mathcal{L}_{\mathbb{P}_1}(\mathbf{1}_{[t,T_1]}(\cdot)X_1(\cdot),a_1(\cdot),W_{Q,1}(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(\mathbf{1}_{[t,T_1]}(\cdot)X_2(\cdot),a_2(\cdot),W_{Q,1}(\cdot)) \text{ on } D.$$

Without loss of generality we may assume that $T_1 \in D$. The solutions on $[T_1, T_2]$ are obtained as the limits in $\mathcal{H}_{p,i}$ (restricted to $[T_1, T_2]$) of the iterates $(\mathcal{K}_i^n(X_i(T_1)))(\cdot)$, where now

$$\mathcal{K}_i(Z_i(\cdot))(s) := e^{(s-T_1)A}X_i(T_1) + \int_{T_1}^s e^{(s-r)A}b(r, Z_i(r), a_i(r))dr \\ + \int_{T_1}^s e^{(s-r)A}\sigma(r, Z_i(r), a_i(r))dW_{Q,i}(r).$$

Thus, again using Lemma 1.130 and passing to the limit as $n \to +\infty$, it follows that

$$\mathcal{L}_{\mathbb{P}_1}(\mathbf{1}_{[t,T_2]}(\cdot)X_1(\cdot), a_1(\cdot), W_{Q,1}(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(\mathbf{1}_{[t,T_2]}(\cdot)X_2(\cdot), a_2(\cdot), W_{Q,1}(\cdot)) \text{ on } D.$$

We repeat the procedure to obtain the required claim. □

1.5. Further existence and uniqueness results in special cases

Throughout this section $T > 0$ is a fixed constant, H, Ξ, Q, and the generalized reference probability space $\mu = (\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s\in[0,T]}, \mathbb{P}, W_Q)$ are as in Section 1.3 (with $t = 0$), A is the infinitesimal generator of a C_0-semigroup on H, and Λ is a Polish space. As in previous sections we will only consider equations on the interval $[0, T]$, however all results are the same if instead of $[0, T]$ we took an interval $[t, T]$, for $0 \le t < T$.

1.5.1. SDE coming from boundary control problems. In this section we study SDE that include equations coming from optimal control problems with boundary control and noise. To see how they arise the reader can look at the examples in Subsections 2.6.2 and 2.6.3, and Appendix C. We consider the following SDE in H:

$$\begin{cases} dX(s) = \big(AX(s) + b(s, X(s), a(s)) + (\lambda I - A)^{\beta}Ga_b(s)\big)\,ds \\ \qquad\qquad + \sigma(s, X(s), a(s))dW_Q(s), \qquad s \in (0, T] \\ X(0) = \xi. \end{cases} \tag{1.54}$$

HYPOTHESIS 1.132

(i) *A generates an analytic semigroup $\{e^{tA}\}_{t\ge 0}$ and λ is a real constant such that $(\lambda I - A)^{-1} \in \mathcal{L}(H)$.*

(ii) *$a\colon [0, T] \times \Omega \to \Lambda$ is progressively measurable, $b(\cdot, \cdot, \cdot)$ satisfies (1.32) and (1.34).*

(iii) *Λ_b is a Hilbert space and $a_b(\cdot)\colon [0, T] \times \Omega \to \Lambda_b$ is progressively measurable.*

(iv) *$G \in \mathcal{L}(\Lambda_b; H)$.*

(v) *$\beta \in [0, 1)$.*

(vi) *γ is a constant belonging to the interval $\left[0, \frac{1}{2}\right)$, σ is a mapping such that $(\lambda I - A)^{-\gamma}\sigma : [0, T] \times H \times \Lambda_b \to \mathcal{L}_2(\Xi_0; H)$ is continuous. There exists a constant $C > 0$ such that*

$$\|(\lambda I - A)^{-\gamma}\sigma(s, x, a)\|_{\mathcal{L}_2(\Xi_0,H)} \le C(1 + |x|)$$

for all $s \in [0, T],\ x \in H,\ a \in \Lambda$ and

$$\|(\lambda I - A)^{-\gamma}[\sigma(s, x_1, a) - \sigma(s, x_2, a)]\|_{\mathcal{L}_2(\Xi_0;H)} \le C|x_1 - x_2|$$

for all $s \in [0, T],\ x_1, x_2 \in H,\ a \in \Lambda$.

REMARK 1.133 Part (i) of Hypothesis 1.132 implies, thanks to (B.15), that for every $\theta \geq 0$ there exists $M_\theta > 0$ such that

$$|(\lambda I - A)^\theta e^{tA} x| \leq \frac{M_\theta}{t^\theta}|x|, \quad \text{for every } t \in (0,T], \ x \in H. \tag{1.55}$$

■

Following Remark 1.114, the definition of a mild solution of (1.54) is given by Definition 1.113 in which the term

$$\int_0^s e^{(s-r)A}(\lambda I - A)^\beta G a_b(r) dr$$

is interpreted as

$$\int_0^s (\lambda I - A)^\beta e^{(s-r)A} G a_b(r) dr,$$

and the term

$$\int_0^s e^{(s-r)A}\sigma(r, X(r), a(r)) dW_Q(r)$$

as

$$\int_0^s (\lambda I - A)^\gamma e^{(s-r)A}(\lambda I - A)^{-\gamma}\sigma(r, X(r), a(r)) dW_Q(r).$$

This is natural since $(\lambda I - A)^\beta e^{(s-r)A}$ is an extension of $e^{(s-r)A}(\lambda I - A)^\beta$ and $(\lambda I - A)^\gamma e^{(s-r)A}(\lambda I - A)^{-\gamma} = e^{(s-r)A}$.

REMARK 1.134 SDE of type (1.54) appear most frequently in optimal control problems of parabolic equations on a domain $\mathcal{O} \subset \mathbb{R}^n$ with boundary control/noise, see Section 2.6.2. More precisely the cases $\beta \in \left(\frac{3}{4}, 1\right)$ and $\beta \in \left(\frac{1}{4}, \frac{1}{2}\right)$ are related respectively to the Dirichlet and Neumann boundary control problems when one takes $\Lambda_b = L^2(\partial\mathcal{O})$ and $H = L^2(\mathcal{O})$. $\gamma \in \left(\frac{1}{4}, \frac{1}{2}\right)$ arises when one treats problems with boundary noise of Neumann type where again $\Lambda_b = L^2(\partial\mathcal{O})$ and $H = L^2(\mathcal{O})$. $\gamma, \beta \in \left(\frac{1}{2} - \epsilon, \frac{1}{2}\right)$ arise in some specific Dirichlet boundary control/noise problems when one considers $\Lambda_b = L^2(\partial\mathcal{O})$ and a suitable weighted L^2 space as H. ■

THEOREM 1.135 *Assume that Hypothesis 1.132 holds, $p \geq 2$, and let $\alpha := \frac{1}{2} - \gamma$. Suppose that*

$$p > \frac{1}{\alpha} \tag{1.56}$$

and $a_b(\cdot) \in M^q_\mu(0,T;\Lambda_b)$ for some $q \geq p$, $q > \frac{1}{1-\beta}$. Then, for every initial condition $\xi \in L^2(\Omega, \mathscr{F}_0, \mathbb{P})$ and processes $a(\cdot), a_b(\cdot)$, there exists a unique mild solution $X(\cdot) = X(\cdot; 0, \xi, a(\cdot), a_b(\cdot))$ of (1.54) in $\mathcal{H}_2(0,T;H)$ with continuous trajectories $\mathbb{P}$-a.s.. If there exists a constant $C > 0$ such that

$$\|(\lambda I - A)^{-\gamma}\sigma(s,x,a)\|_{\mathcal{L}_2(\Xi_0,H)} \leq C \tag{1.57}$$

for all $s \in [0,T]$, $x \in H$, $a \in \Lambda$, then the solution has continuous trajectories $\mathbb{P}$-a.s. without the restriction $p > \frac{1}{\alpha}$. If $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ then $X(\cdot) \in \mathcal{H}_p(0,T;H)$ and there exists a constant $C_{T,p}$ independent of ξ such that

$$\sup_{s\in[0,T]} \mathbb{E}|X(s)|^p \leq C_{T,p}(1 + \mathbb{E}|\xi|^p). \tag{1.58}$$

PROOF. Assume first that $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ where $p \geq 2$ without the restriction (1.56). Similarly to the proof of Theorem 1.121, we will show that for some

$T_0 \in (0, T]$ the map

$$\begin{cases} \mathcal{K} \colon \mathcal{H}_p(0, T_0) \to \mathcal{H}_p(0, T_0), \\ \mathcal{K} \colon Y \mapsto e^{sA}\xi + \int_0^s e^{(s-r)A} b(r, Y(r), a(r)) dr + \int_0^s (\lambda I - A)^\beta e^{(s-r)A} G a_b(r) dr \\ \qquad\qquad + \int_0^s (\lambda I - A)^\gamma e^{(s-r)A} (\lambda I - A)^{-\gamma} \sigma(r, Y(r), a(r)) dW_Q(r) \end{cases} \tag{1.59}$$

is well defined and is a contraction. The only difference between our case here and that considered in Theorem 1.121 are the last two terms in (1.59).

First we prove that $\mathcal{K}$ maps $\mathcal{H}_p(0, T_0)$ into $\mathcal{H}_p(0, T_0)$. We only show how to deal with the non standard terms. For the third term in (1.59) we can argue as follows. If M_β is the constant from (1.55) for $\theta = \beta$, using (1.55), Hölder and Jensen's inequalities, and $q \geq p$, $q > \frac{1}{1-\beta}$, we obtain

$$\begin{aligned} \sup_{s\in[0,T_0]} \mathbb{E} \left| \int_0^s (\lambda I - A)^\beta e^{(s-r)A} G a_b(r) dr \right|^p & \\ \leq \sup_{s\in[0,T_0]} M_\beta^p \|G\|^p \mathbb{E} \left(\int_0^s \frac{1}{(s-r)^\beta} |a_b(r)| dr \right)^p & \\ \leq M_\beta^p \|G\|^p \left(\int_0^{T_0} \frac{1}{(T_0 - r)^{\frac{\beta q}{q-1}}} dr \right)^{\frac{p(q-1)}{q}} \mathbb{E} \left[\int_0^{T_0} |a_b(r)|^q dr \right]^{\frac{p}{q}} & \\ \leq C_1 \left(\mathbb{E} \left[\int_0^{T_0} |a_b(r)|^q dr \right] \right)^{\frac{p}{q}} < +\infty. & \end{aligned} \tag{1.60}$$

As regards the stochastic integral term, using Theorem 1.106, (1.55), and Hypothesis 1.132-(vi), we estimate

$$\begin{aligned} \sup_{s\in[0,T_0]} \mathbb{E} \left| \int_0^s (\lambda I - A)^\gamma e^{(s-r)A} (\lambda I - A)^{-\gamma} \sigma(r, Y(r), a(r)) dW_Q(r) \right|^p & \\ \leq \sup_{s\in[0,T_0]} C_1 \mathbb{E} \left| \int_0^s \frac{1}{(s-r)^{2\gamma}} \|(\lambda I - A)^{-\gamma} \sigma(r, Y(r), a(r))\|^2_{\mathcal{L}_2(\Xi_0, H)} dr \right|^{\frac{p}{2}} & \\ \leq \sup_{s\in[0,T_0]} C_2 \left(\int_0^{T_0} \frac{1}{(T_0 - r)^{2\gamma}} dr \right)^{\frac{p}{2}-1} \int_0^s \frac{1}{(s-r)^{2\gamma}} \mathbb{E}[(1 + |Y(r)|)^p] dr & \\ \leq C_3 \left(1 + |Y|^p_{\mathcal{H}_p(0,T_0)} \right) & \end{aligned} \tag{1.61}$$

for some constant C_3.

Regarding the proof that, for T_0 small enough, $\mathcal{K}$ is a contraction, the only non-standard term to check is the stochastic convolution term, since the third term in (1.59) does not depend on X. Arguing as before we have that for $X, Y \in \mathcal{H}_p(0, T_0)$, thanks to Theorem 1.106, (1.55), Hypothesis 1.132-(vi), and Jensen's inequality,

$$\sup_{s\in[0,T_0]} \mathbb{E}\left|\int_0^s (\lambda I - A)^\gamma e^{(s-r)A}(\lambda I - A)^{-\gamma}\left[\sigma(r,X(r),a(r)) - \sigma(r,Y(r),a(r))\right] dW_Q(r)\right|^p$$

$$\leq \sup_{s\in[0,T_0]} C_1\mathbb{E}\left(\int_0^s \frac{1}{(s-r)^{2\gamma}}\left\|(\lambda I - A)^{-\gamma}\left[\sigma(r,X(r),a(r)) - \sigma(r,Y(r),a(r))\right]\right\|^2_{\mathcal{L}_2(\Xi_0,H)} dr\right)^{\frac{p}{2}}$$

$$\leq \sup_{s\in[0,T_0]} C_2\mathbb{E}\left(\int_0^s \frac{1}{(s-r)^{2\gamma}}|X(r)-Y(r)|^2 dr\right)^{\frac{p}{2}}$$

$$\leq \sup_{s\in[0,T_0]} C_2\left(\int_0^{T_0} \frac{1}{(T_0-r)^{2\gamma}} dr\right)^{\frac{p}{2}-1}\int_0^s \frac{1}{(s-r)^{2\gamma}}\mathbb{E}[|X(r)-Y(r)|^p]dr$$

$$\leq \omega(T_0)|X-Y|^p_{\mathcal{H}_p(0,T_0)}, \quad (1.62)$$

where $\omega(r) \xrightarrow{r\to 0^+} 0$. So for T_0 small enough (which is independent of the initial condition) we can apply the Banach fixed point theorem in $\mathcal{H}_p(0,T_0)$ as in the proof of Theorem 1.121 (see the proof of [130], Theorem 7.2, page 188). The process can now be reapplied on intervals $[T_0, 2T_0], ..., [kT_0, T]$, where $k = [T/T_0]$, to obtain the existence of a a unique mild solution in $\mathcal{H}_p(0,T)$ in the sense of the integral equality being satisfied for a.e. $s\in[0,T]$.

Estimate (1.58) follows from similar arguments using the growth assumptions on b,σ in Hypothesis 1.132 and a version of Gronwall's lemma, Proposition D.25.

We will now prove the continuity of the trajectories if condition (1.56) is satisfied. We will only prove the continuity of the stochastic convolution term in (1.59) since the continuity of the other terms is easier to show. In particular, the continuity of the trajectories of the third term in (1.59) follows from Lemma 1.110-(ii).

Let now $p > \frac{1}{\alpha}$. Hence there is $0 < \alpha' < \alpha$ such that $p > \frac{1}{\alpha'}$. Then, for $r\in[t,T]$, using (1.55), (1.58), Hypothesis 1.132-(vi), and Jensen's inequality

$$\mathbb{E}\left(\int_0^r (r-h)^{-2\alpha'}\left\|(\lambda I - A)^\gamma e^{(r-h)A}(\lambda I - A)^{-\gamma}\sigma(h,X(h),a(h))\right\|^2_{\mathcal{L}_2(\Xi_0;H)} dh\right)^{\frac{p}{2}}$$

$$\leq \mathbb{E}\left(\int_0^r (r-h)^{-2\alpha'}\|(\lambda I - A)^\gamma e^{(r-h)A}\|^2_{\mathcal{L}(H)}\left\|(\lambda I - A)^{-\gamma}\sigma(h,X(h),a(h))\right\|^2_{\mathcal{L}_2(\Xi_0;H)} ds\right)^{\frac{p}{2}}$$

$$\leq C_1\mathbb{E}\left(\int_0^r (r-h)^{-2\alpha'}(r-h)^{-2\gamma}(1+|X(h)|)^2 dh\right)^{\frac{p}{2}}$$

$$\leq C_1\left(\int_0^T (T-h)^{-2\alpha'}(T-h)^{-2\gamma} dh\right)^{\frac{p}{2}} \sup_{h\in[0,T]}\mathbb{E}[(1+|X(h)|)^p] =: C_2 < +\infty. \quad (1.63)$$

Observe that C_2 does not depend on $r\in[0,T]$. This proves (1.27) and thus the claim follows from Proposition 1.111. When (1.57) holds, estimate (1.63) is easier and can be done for any exponent $p' > 1/\alpha$ in place of p, and thus (1.27) is always satisfied.

Finally we need to discuss the continuity of the trajectories if $\xi\in L^2(\Omega,\mathscr{F}_0,\mathbb{P})$. We argue as in the proof of Theorem 7.2 of [130]. For $n\geq 1$ we define the random variables

$$\xi_n = \begin{cases} \xi & \text{if } |\xi|\leq n \\ 0 & \text{if } |\xi| > n. \end{cases}$$

The solutions $X(\cdot;0,\xi,a(\cdot),a_b(\cdot))$ and $X(\cdot;0,\xi_n,a(\cdot),a_b(\cdot))$ on $[0,T_0]$ are obtained as fixed points in $\mathcal{H}_2(0,T_0)$ and $\mathcal{H}_p(0,T_0)$, with p large enough, of the same contraction map (1.59) with the second map having the term $e^{sA}\xi_n$ in place of $e^{sA}\xi$. Therefore both solutions can be obtained as limits of successive iterations starting from processes $e^{sA}\xi_n$ and $e^{sA}\xi_n$ respectively. It is then easy to see that we have $X(\cdot;0,\xi,a(\cdot),a_b(\cdot)) = X(\cdot;0,\xi_n,a(\cdot),a_b(\cdot))$, $\mathbb{P}$-a.s. on $\{\omega : |\xi(\omega)| \le n\}$. However the solutions $X(\cdot;0,\xi_n,a(\cdot),a_b(\cdot))$ have continuous trajectories. Thus $X(\cdot;0,\xi,a(\cdot),a_b(\cdot))$ has continuous trajectories $\mathbb{P}$-a.s. on $[0,T_0]$ and we can then continue the argument on intervals $[T_0,2T_0],\ldots$. □

Proposition 1.136 *Let the assumptions of Theorem 1.135 be satisfied. Denote the unique mild solution of (1.54) in $\mathcal{H}_p(0,T;H)$ by $X(\cdot) = X(\cdot;0,\xi,a(\cdot),a_b(\cdot))$.*

(i) *If $\xi^1 = \xi^2$ $\mathbb{P}$-a.s., $a^1(\cdot) = a^2(\cdot)$ $dt\otimes d\mathbb{P}$-a.s. $a_b^1(\cdot) = a_b^2(\cdot)$ $dt\otimes d\mathbb{P}$-a.s., then $\mathbb{P}$-a.s., $X(\cdot;0,\xi^1,a^1(\cdot),a_b^1(\cdot)) = X(\cdot;0,\xi^2,a^2(\cdot),a_b^2(\cdot))$ on $[0,T]$.*

(ii) *Let $(\Omega_1,\mathscr{F}_1,\mathscr{F}_s^1,\mathbb{P}_1,W_{Q,1})$ and $(\Omega_2,\mathscr{F}_2,\mathscr{F}_s^2,\mathbb{P}_2,W_{Q,2})$ be two generalized reference probability spaces. Let $\zeta_i \in L^p(\Omega_i,\mathscr{F}_0^i,\mathbb{P}_i), i = 1,2$. Let $(a^i,a_b^i)\colon [0,T]\times\Omega_i \to \Lambda\times\Lambda_b, i = 1,2$ be $\mathscr{F}_s^i$-progressively measurable processes satisfying the assumptions of Theorem 1.135. Suppose that $\mathcal{L}_{\mathbb{P}_1}(a^1(\cdot),a_b^1(\cdot),W_{Q,1}(\cdot),\zeta_1) = \mathcal{L}_{\mathbb{P}_2}(a^2(\cdot),a_b^2(\cdot),W_{Q,1}(\cdot),\zeta_2)$ on some subset $D\subset[t,T]$ of full measure. Then $\mathcal{L}_{\mathbb{P}_1}(X(\cdot;0,\zeta_1,a^1(\cdot),a_b^1(\cdot)),a^1(\cdot),a_b^1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(X(\cdot;0,\zeta_2,a^2(\cdot),a_b^2(\cdot)),a^2(\cdot),a_b^2(\cdot))$ on D.*

(iii) *The solution of (1.54) is unique in $M_\mu^2(0,T;H)$ as well.*

Proof. (i): We argue as in the proof of Corollary 1.122. If $X_i(\cdot) := X(\cdot;0,\xi^i,a^i(\cdot),a_b^i(\cdot))$, using Theorem 1.98, Jensen's inequality and Hypothesis 1.132, we obtain

$$\mathbb{E}\left[|X_1(s)-X_2(s)|^2\right] \le C_T\int_0^s \frac{1}{(s-r)^{2\gamma}}\mathbb{E}|X_1(r)-X_2(r)|^2 dr,$$

and the claim follows using a version of Gronwall's lemma (Proposition D.25), and the continuity of the trajectories.

(ii): The argument is the same as the one used to prove Lemma 1.130 and Proposition 1.131, since in the current case the solution is also found iterating the map $\mathcal{K}$.

(iii): Using the same steps as these in the prof of Theorem 1.135 it is not difficult to see that $\mathcal{K}$ is also a contraction in the space $M_\mu^2(0,T_0;H)$ for some T_0, and this gives the required uniqueness. □

1.5.2. Semilinear SDE with additive noise. In this section we give more precise results for some semilinear SDE with additive noise, i.e. for equation (1.29) when the coefficient σ is constant and we have possible unboundedness in the drift.

Consider first the stochastic convolution process

$$W_A(s) = \int_0^s e^{(s-r)A}\sigma dW_Q(r). \tag{1.64}$$

Hypothesis 1.137

(i) *The linear operator A is the generator of a strongly continuous semigroup $\{e^{tA}, t\ge 0\}$ in H and, for suitable $M\ge 1$ and $\omega\in\mathbb{R}$,*

$$|e^{tA}x| \le Me^{\omega t}|x|, \quad \forall t\ge 0,\ x\in H. \tag{1.65}$$

(ii) *$\sigma\in\mathcal{L}(\Xi,H)$ and the symmetric positive operator*

$$Q_t : H\to H, \qquad Q_t := \int_0^t e^{sA}\sigma Q\sigma^* e^{sA^*} ds,$$

is of trace class for every $t \geq 0$, i.e.

$$\mathrm{Tr}\,[Q_s] < +\infty. \tag{1.66}$$

PROPOSITION 1.138 *Suppose that Hypothesis 1.137 is satisfied. Then the process W_A defined in (1.64) is a Gaussian process with mean 0 and covariance operator Q_s, and is mean square continuous. The trajectories of $W_A(\cdot)$ are $\mathbb{P}$-a.s. square integrable, and $W_A(\cdot) \in M^2_\mu(0,T;H)$. Moreover, if there exists $\gamma > 0$ such that*

$$\int_0^T s^{-\gamma}\mathrm{Tr}\left[e^{sA}\sigma Q\sigma^* e^{sA^*}\right] ds < \infty, \tag{1.67}$$

then $W_A(\cdot)$ has continuous trajectories[4] *and, for $p \geq 2$,*

$$\mathbb{E}\left[\sup_{0\leq s\leq T} |W_A(s)|^p\right] < +\infty.$$

PROOF. See [130] Chapter 5, Theorems 5.2 and 5.11. The last estimate can be found e.g. as a particular case of [211, Proposition 3.2]. □

A completely analogous result also holds for the stochastic convolution starting at a point $t \geq 0$ for $W_A(t,s) := \int_t^s e^{(s-r)A}\sigma dW_Q(r)$.

Let $T > 0$. We consider the SDE

$$\begin{cases} dX(s) = [AX(s) + b(s, X(s))]\, ds + \sigma dW_Q(s), & s > 0 \\ X(0) = \xi. \end{cases} \tag{1.68}$$

HYPOTHESIS 1.139 *$p \geq 2$ and $b(s,x) = b_0(s,x,a_1(s)) + a_2(s)$, where:*

(i) *The process $a_1(\cdot) : [0,T]\times\Omega \to \Lambda$ (where Λ is a given Polish space) is $\mathscr{F}_s$-progressively measurable. The map $b_0 : [0,T]\times H\times\Lambda \to H$ is Borel measurable and there exists a non-negative function $f \in L^1(0,T;\mathbb{R})$ such that*

$$|b_0(s,x,a_1)| \leq f(s)(1+|x|) \qquad \forall s\in[0,T],\ x\in H \ \text{ and } a_1\in\Lambda.$$

$$|b_0(s,x_1,a_1) - b_0(s,x_2,a_1)| \leq f(s)|x_1-x_2| \qquad \forall s\in[0,T],\ x_1,x_2\in H \ \text{ and } a_1\in\Lambda.$$

(ii) *The process $a_2(\cdot)$ is such that for all $t > 0$, the process $(s,\omega)\mapsto e^{(t-s)A}a_2(s,\omega)$, when interpreted properly, is $\mathscr{F}_s$-progressively measurable on $[0,t]\times\Omega$ with values in H, and*

$$|e^{tA}a_2(s,\omega)| \leq t^{-\beta}g(s,\omega) \qquad \forall(s,\omega)\in[0,T]\times\Omega,\ x\in H,$$

for some $\beta\in(0,1)$ and $g\in M^q_\mu(0,T;\mathbb{R})$, where $q\geq p, q > \frac{1}{1-\beta}$.

Hypothesis 1.139 covers some cases which are not standard and for which a separate proof of existence and uniqueness of mild solutions of (1.68) is required.

REMARK 1.140 Hypothesis 1.139-(ii) is satisfied for example when A is the generator of an analytic C_0-semigroup and the process $a_2(\cdot)$ is of the form $a_2(s) = (\lambda I - A)^\beta a_3(s)$, where $\lambda\in\mathbb{R}$ is such that $(\lambda I - A)$ is invertible, $\beta\in(0,1)$, $a_3(\cdot)\in M^q_\mu(0,T;H)$, $q\geq p, q>\frac{1}{1-\beta}$. In such cases the definition of a mild solution of (1.68) is given by Definition 1.113 in which the formal term

$$\int_0^s e^{(s-r)A}a_2(r)dr = \int_0^s e^{(s-r)A}(\lambda I - A)^\beta a_3(r)dr$$

appearing in the definition of mild solution is interpreted as

$$\int_0^s (\lambda I - A)^\beta e^{(s-r)A}a_3(r)dr.$$

[4]Without assuming (1.67) such continuity of trajectories may fail to hold, see e.g. [268].

This is natural since $(\lambda I - A)^\beta e^{(s-r)A}$ is an extension of $e^{(s-r)A}(\lambda I - A)^\beta$. ■

PROPOSITION 1.141 *Let $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ and Hypotheses 1.137 and 1.139 be satisfied. Then equation (1.68) has a unique mild solution $X(\cdot; 0, \xi) \in \mathcal{H}_p^\mu(0, T; H)$. The solution satisfies, for some $C_p(T) > 0$ independent of ξ,*

$$\sup_{s\in[0,T]} \mathbb{E}\left[|X(s; 0, \xi)|^p\right] \le C_p(T)(1 + \mathbb{E}|\xi|^p). \tag{1.69}$$

Moreover, if $\xi_1, \xi_2 \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$, we have, $\mathbb{P}$-a.s.,

$$|X(s; 0, \xi_1) - X(s; 0, \xi_2)| \le Me^{\alpha T}|\xi_1 - \xi_2| e^{Me^{\alpha T}\int_0^s f(r)dr}. \tag{1.70}$$

Finally, if also (1.67) holds for some $\gamma > 0$, then the solution $X(\cdot; 0, \xi)$ has $\mathbb{P}$-a.s. continuous trajectories, and if $\xi = x \in H$ is deterministic we then have

$$\mathbb{E}(\sup_{s\in[0,T]} |X(s)|^p) \le C_p(T)(1 + |x|^p) \tag{1.71}$$

for a suitable constant $C_p(T) > 0$ independent of x. In particular if g in Hypothesis 1.139-(ii) is in $M_\mu^q(0, T; \mathbb{R})$ for any $q \ge 1$, the previous estimate holds for any $p > 0$. The same is true for (1.69) if $\xi = x \in H$.

PROOF. The proof of existence and uniqueness uses the same techniques employed in the Lipschitz case (Theorem 1.121) but contains a small additional difficulty due the presence of the term $a_2(\cdot)$ and possible singularities in s of the Lipschitz norm of $b_0(s, \cdot)$. We will write $\mathcal{H}_p(0, T)$ for $\mathcal{H}_p^\mu(0, T; H)$. For $Y \in \mathcal{H}_p(0, T)$ we set

$$\mathcal{K}(Y)(s) = e^{(s-t)A}\xi + \int_0^s e^{(s-r)A} b_0(r, Y(r), a_1(r))dr + \int_0^s e^{(s-r)A} a_2(r)dr + W_A(s). \tag{1.72}$$

W_A belongs to $\mathcal{H}_p(0, T)$ thanks to Proposition 1.138. Hypotheses 1.139-(i) and 1.139-(ii) ensure respectively that the second and third term in the definition of the map $\mathcal{K}$ belong to $\mathcal{H}_p(0, T)$ as well (see (1.60)). So $\mathcal{K}$ maps $\mathcal{H}_p(0, T)$ into itself. For Y_1, $Y_2 \in \mathcal{H}_p(0, T)$, $s \in [0, T]$,

$$|\mathcal{K}(Y_1)(s) - \mathcal{K}(Y_2)(s)| \le Me^{\omega T} \int_0^s f(r)|Y_1(r) - Y_2(r)|dr,$$

which yields, for $T_0 \in (0, T]$,

$$\begin{aligned} |\mathcal{K}(Y_1) - \mathcal{K}(Y_2)|^p_{\mathcal{H}_p(0,T_0)} &\le Me^{\alpha T} \sup_{s\in[0,T_0]} \mathbb{E}\left[\int_0^s f(r)|Y_1(r) - Y_2(r)|dr\right]^p \\ &\le Me^{\alpha T}\left[\int_0^{T_0} f(r)dr\right]^p \sup_{s\in[0,T_0]} \mathbb{E}|Y_1(s) - Y_2(s)|^p \\ &= Me^{\alpha T}\left[\int_0^{T_0} f(r)dr\right]^p |Y_1 - Y_2|^p_{\mathcal{H}_p(0,T_0)}. \end{aligned} \tag{1.73}$$

Therefore, if T_0 sufficiently small, we can apply the contraction mapping principle to find the unique mild solution of (1.68) in $\mathcal{H}_p(0, T_0)$. The existence and uniqueness of a solution on the whole interval $[0, T]$ follows, as usual, by repeating the procedure a finite number of steps, since the estimate (1.73) does not depend on the initial data, and the number of steps does not blow up since f is integrable. Estimate (1.69) follows from (1.72) applied to the solution X if we perform estimates similar to these above and use Gronwall's Lemma.

To show (1.70) we notice that if $Z(s) = X(s;0,\xi_1) - X(s;0,\xi_2)$, then for $s \in [0,T]$

$$Z(s) = e^{sA}(\xi_1-\xi_2)+\int_0^s e^{(s-r)A}[b_0(r,X(r;0,\xi_1),a_1(r))-b_0(r,X(r;0,\xi_2),a_1(r))]dr.$$

By Hypothesis 1.139 we thus have

$$|Z(s)| \le Me^{\alpha T}|\xi_1-\xi_2| + Me^{\alpha T}\int_0^s f(r)|Z(r)|dr, \quad s\in[0,T]$$

so that, by Gronwall's inequality (see Proposition D.24),

$$|Z(s)| \le Me^{\alpha T}|\xi_1-\xi_2|e^{Me^{\alpha T}\int_0^s f(r)dr}$$

which gives the claim. The continuity of trajectories follows from Proposition 1.138, Hypothesis 1.139 and Lemma 1.110 for the second and fourth terms in (1.72), and from direct computations, together with the dominated convergence theorem, for the $\int_0^s e^{(s-r)A}a_2(r)dr$ term.

The last estimate (1.71) follows by standard arguments (see the proof of (1.38) in Theorem 1.124) if we use Proposition 1.138. This implies that if $g \in M^q_\mu(0,T;\mathbb{R})$ for any $q>0$, (1.71) holds for any $p \ge 2$. For $p\in(0,2)$, denoting $Z_r(s) := \sup_{s\in[0,T]}|X(s)|^r$, we have

$$\mathbb{E}(Z_p(s)) \le [\mathbb{E}(Z_p(s)^{2/p})]^{p/2} \le (C(1+|x|^2))^{p/2} \le C_1(1+|x|^p).$$

$\square$

PROPOSITION 1.142 *Assume that Hypotheses 1.137, 1.139 hold, and let $a_2(\cdot)$ be as in Remark 1.140. Then:*

(i) *Let $\xi_1,\xi_2 \in L^2(\Omega,\mathscr{F}_0,\mathbb{P})$, $\xi_1=\xi_2$ $\mathbb{P}$-a.s.. Let $(a_1^1(\cdot),a_3^1(\cdot)),(a_1^2(\cdot),a_3^2(\cdot))$ be two processes satisfying Hypothesis 1.139, together with Remark 1.140, such that $(a_1^1(\cdot),a_3^1(\cdot)) = (a_1^2(\cdot),a_3^2(\cdot))$, $dt\otimes d\mathbb{P}$-a.s. Then, denoting by $X^i(\cdot;0,\xi_i)$ the solution of (1.68) for $b(s,x) = (\lambda-A)^\beta a_3^i(s) + b_0(s,x,a_1^i(s))$, we have $X^1(\cdot;0,\xi_1) = X^2(\cdot;0,\xi_2)$, $\mathbb{P}$-a.s. on $[0,T]$.*

(ii) *Let $\left(\Omega_1,\mathscr{F}_1,\mathscr{F}_s^1,\mathbb{P}_1,W_{Q,1}\right)$ and $\left(\Omega_2,\mathscr{F}_2,\mathscr{F}_s^2,\mathbb{P}_2,W_{Q,2}\right)$ be two generalized reference probability spaces. Let $\xi_i \in L^2(\Omega_i,\mathscr{F}_0^i,\mathbb{P}_i)$, $i=1,2$. Let $a_1^i(\cdot),a_3^i(\cdot)$, $i=1,2$, be processes on $[0,T]\times\Omega_i$ satisfying Hypothesis 1.139, together with Remark 1.140. Suppose that $\mathcal{L}_{\mathbb{P}_1}(a_1^1(\cdot),a_3^1(\cdot),W_{Q,1}(\cdot),\xi_1) = \mathcal{L}_{\mathbb{P}_2}(a_1^2(\cdot),a_3^2(\cdot),W_{Q,2}(\cdot),\xi_2)$. Then $\mathcal{L}_{\mathbb{P}_1}(X^1(\cdot;0,\xi_1),a_1^1(\cdot),a_3^1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(X(\cdot;0,\xi_2),a_1^2(\cdot),a_3^2(\cdot))$.*

(iii) *If $f\in L^2(0,T;\mathbb{R})$ then the solution of (1.68) ensured by Proposition 1.141 is unique in $M^2_\mu(0,T;H)$ as well.*

PROOF. Parts (i) and (ii) are proved similarly as Proposition 1.136 (i)-(ii). Part (iii) follows from (1.70) which is also true in this case. We also point out that if $f\in L^2(0,T;\mathbb{R})$ then $\mathcal{K}$ maps $M^2_\mu(0,T;H)$ into itself and is a contraction in $M^2_\mu(0,T_0;H)$ for small T_0. $\square$

1.5.3. Semilinear SDE with multiplicative noise. This section contains a result for a class of semilinear SDE with multiplicative noise. Let $T>0$, and let H, Ξ, Λ and a generalized reference probability space $\left(\Omega,\mathscr{F},\{\mathscr{F}_s\}_{s\in[0,T]},\mathbb{P},W\right)$ be as in Section 1.3, where $W(t)$, $t\in[0,T]$, is a cylindrical Wiener process with covariance $Q=I$ (so here $\Xi_0=\Xi$). We consider the following SDE in H:

$$\begin{cases} dX(s) = AX(s)\,ds + b(s,X(s),a(s))\,ds + \sigma(s,X(s),a(s))\,dW(s), & s\in[0,T],\\ X(0)=x\in H. \end{cases} \tag{1.74}$$

HYPOTHESIS 1.143

(i) *The operator A generates a strongly continuous semigroup $\left\{e^{tA}\right\}_{t\geq 0}$ in H.*

(ii) *$a(\cdot)$ is a Λ-valued progressively measurable process.*

(iii) *b is a function such that, for all $s \in (0,T]$, $e^{sA}b : [0,T] \times H \times \Lambda \to H$ is measurable and there exist a non-negative function $f_1 \in L^1(0,T;\mathbb{R})$ such that*

$$|e^{sA}b(t,x,a)| \leq f_1(s)(1+|x|), \tag{1.75}$$

$$|e^{sA}(b(t,x,a) - b(t,y,a))| \leq f_1(s)|x-y|, \tag{1.76}$$

for any $s \in (0,T],\ t \in [0,T],\ x,y \in H,\ a \in \Lambda$.

(iv) *The function $\sigma : [0,T] \times H \times \Lambda \to \mathcal{L}(\Xi, H)$ is such that, for every $v \in \Xi$, the map $\sigma(\cdot,\cdot,\cdot)v : [0,T] \times H \times \Lambda \to H$ is measurable and, for every $s > 0,\ t \in [0,T],\ a \in \Lambda$ and $x \in H$, $e^{sA}\sigma(t,x,a)$ belongs to $\mathcal{L}_2(\Xi,H)$. Moreover there exists a non-negative function $f_2 \in L^2(0,T;\mathbb{R})$ such that*

$$\begin{aligned} |e^{sA}\sigma(t,x,a)|_{\mathcal{L}_2(\Xi,H)} &\leq f_2(s)(1+|x|), & (1.77)\\ |e^{sA}\sigma(t,x,a) - e^{sA}\sigma(t,y,a)|_{\mathcal{L}_2(\Xi,H)} &\leq f_2(s)|x-y|, & (1.78) \end{aligned}$$

for every $s \in (0,T],\ t \in [0,T],\ x,y \in H,\ a \in \Lambda$.

The solution of equation (1.74) is defined in the mild sense of Definition 1.113, where the convolution term

$$\int_0^s e^{(s-r)A}\sigma(r,X(r),a(r))\,dW(r), \qquad s \in [0,T],$$

makes sense thanks to (1.77) and Remark 1.117. Moreover, since $s \mapsto e^{sA}b(t,x,a)$ is continuous on $(0,T]$ for every $t \in [0,T], x \in H, a \in \Lambda$, we can consider $e^{\cdot A}b$ to be $\mathcal{B}([0,T]) \otimes \mathcal{B}([0,T]) \otimes \mathcal{B}(H) \otimes \mathcal{B}(\Lambda)/\mathcal{B}(H)$ measurable.

THEOREM 1.144 *Let Hypothesis 1.143 hold and let $a(\cdot)$ be a Λ-valued, progressively measurable process. Then, the SDE (1.74) has a unique mild solution $X(\cdot)$ in $\mathcal{H}_p(0,T;H)$ for all $p \geq 2$. The solution satisfies*

$$\sup_{s\in[0,T]} \mathbb{E}\left[|X(s)|^p\right] \leq C(1+|x|^p), \quad \textit{for all } p > 0, \tag{1.79}$$

for some constant C depending only on p, f_1, f_2, T, L and $M := \sup_{s\in[0,T]}|e^{sA}|$. Moreover if there exists $\alpha \in (0,1/2)$ such that

$$\int_0^1 s^{-2\alpha} f_2^2(s)ds < +\infty \tag{1.80}$$

then the mild solution $X(\cdot)$ has continuous trajectories, and we have

$$\mathbb{E}\left[\sup_{s\in[0,T]}|X(s)|^p\right] \leq C(1+|x|^p), \quad \textit{for all } p > 0, \tag{1.81}$$

for some C depending on the same quantities as the constant in (1.79).

PROOF. Let $p \geq 2$. The existence of a unique solution is proved using the Banach contraction mapping theorem in $\mathcal{H}_p(0,T_0)$ for some $T_0 \in (0,T)$ small enough. We define $\mathcal{K}\colon \mathcal{H}_p(0,T) \to \mathcal{H}_p(0,T)$ by

$$\mathcal{K}(Y)(s) := e^{sA}x + \int_0^s e^{(s-r)A}b(r,Y(r),a(r))dr + \int_0^s e^{(s-r)A}\sigma(r,Y(r),a(r))dW(r). \tag{1.82}$$

We observe first that this expression belongs to $\mathcal{H}_p(0,T)$. Thanks to (1.75), (1.77) and Theorem 1.106, we have

$$\begin{aligned}
\mathbb{E}\left|\int_0^s e^{(s-r)A}b(r,Y(r),a(r))dr + \int_0^s e^{(s-r)A}\sigma(r,Y(r),a(r))dW(r)\right|^p \\
\le C_p\left(\mathbb{E}\left|\int_0^s [f_1(s-r)(1+|Y(r)|)]\,dr\right|^p\right. \\
\left. + \mathbb{E}\left|\int_0^s e^{(s-r)A}\sigma(r,Y(r),a(r))dW(r)\right|^p\right) \\
\le C_p\left[\int_0^T f_1(r)dr\right]^p \sup_{r\in[0,T]} \mathbb{E}(1+|Y(r)|)^p \\
+ C_p\left[\int_0^T f_2^2(r)dr\right]^{\frac{p}{2}} \sup_{r\in[0,T]} \mathbb{E}(1+|Y(r)|)^p, \quad (1.83)
\end{aligned}$$

where the constant C_p depends only on p. Therefore, for any $Y \in \mathcal{H}_p(0,T)$, $\mathcal{K}(Y) \in \mathcal{H}_p(0,T)$. The estimates showing that $\mathcal{K}$ is a contraction on $\mathcal{H}_p(0,T_0)$ for $T_0 \in (0,T]$ small enough are essentially the same. Using (1.76) and (1.78) instead of (1.75) and (1.77) we obtain, for all $Y_1, Y_2 \in \mathcal{H}_p(0,T_0)$,

$$\begin{aligned}
|\mathcal{K}(Y_1)-\mathcal{K}(Y_2)|^p_{\mathcal{H}_p(0,T_0)} \le C_p\left(\left[\int_0^{T_0} f_1(r)dr\right]^p\right. \\
\left. + \left[\int_0^{T_0} f_2^2(r)dr\right]^{\frac{p}{2}}\right) \sup_{r\in[0,T_0]} \mathbb{E}(|Y_1(r)-Y_2(r)|^p), \quad (1.84)
\end{aligned}$$

and thus $\mathcal{K}$ is a contraction in $\mathcal{H}_p(0,T_0)$ if $T_0 \in (0,T]$ is small enough. The existence and uniqueness of solution in $\mathcal{H}_p(0,T)$ follows, as usual, by repeating the procedure a finite number of steps, since the estimate does not depend on the initial data, and the number of steps does not blow up since the f_1 and f_2^2 are integrable. Estimate (1.79) follows in a standard way by applying estimates like these in (1.83) to the fixed point of the map $\mathcal{K}$ and using Gronwall's Lemma (see also the proof of Theorem 7.5 in [130]). We can then extend it to $0 < p < 2$ as in the proof of Proposition 1.141.

Finally the continuity of the trajectories and (1.81) is proved using the factorization method as in the proof of Proposition 6.9 for $p > 2$. We extend (1.81) to $0 < p \le 2$ in the same way as for (1.79). □

PROPOSITION 1.145 *Assume that Hypothesis 1.143 holds with $f_1(s) = Ls^{-\gamma_1}$ and $f_2(s) = Ls^{-\gamma_2}$ for some $L > 0$, $\gamma_1 \in (0,1)$ and $\gamma_2 \in (0,1/2)$. Let (t_1,x_1), $(t_2,x_2) \in [0,T]\times H$ with $t_1 \le t_2$. Denote by $X(\cdot;t_1,x_1), X(\cdot;t_2,x_2)$ the corresponding mild solutions of (1.74) with the same progressively measurable process $a(\cdot)$ and initial conditions $X(t_i) = x_i$, $i = 1,2$. Then, for all $s \in [t_2,T]$ we have, setting $\gamma_3 := [2(1-\gamma_1)] \wedge [1-2\gamma_2]$,*

$$\begin{aligned}
&\mathbb{E}[|X(s;t_1,x_1)-X(s;t_2,x_2)|^2] \le \\
&\qquad \le C_2\left[|x_1-x_2|^2 + (1+|x_1|^2)|t_2-t_1|^{\gamma_3} + |e^{(t_2-t_1)A}x_1 - x_1|^2\right]
\end{aligned} \quad (1.85)$$

for some constant C_2 depending only on γ_1, γ_2, T, L and $M := \sup_{s\in[0,T]}|e^{sA}|$. Moreover the term $|e^{(t_2-t_1)A}x_1 - x_1|^2$ can be replaced by $|e^{(t_2-t_1)A}x_2 - x_2|^2$.

PROOF. To simplify notation we denote $X_i(s) := X(s;t_i,x_i), b(r,X_i(r)) := b(r,X_i(r),a(r)), \sigma(r,X_i(r)) := \sigma(r,X_i(r),a(r))$, $i = 1,2$. By the definition of a mild solution we have

$$X_i(s) = e^{(s-t_i)A}x_i + \int_{t_i}^{s} e^{(s-r)A} b(r, X_i(r))dr + \int_{t_i}^{s} e^{(s-r)A}\sigma(r, X_i(r))dW(r),$$

hence

$$|X_1(s) - X_2(s)| \le |e^{(s-t_1)A}x_1 - e^{(s-t_2)A}x_2| +$$
$$+ \left|\int_{t_1}^{t_2} e^{(s-r)A} b(r, X_1(r))dr\right| + \left|\int_{t_2}^{s} e^{(s-r)A}\left(b(r, X_1(r)) - b(r, X_2(r))\right)dr\right| +$$
$$+ \left|\int_{t_1}^{t_2} e^{(s-r)A}\sigma(r, X_1(r))dW(r)\right| + \left|\int_{t_2}^{s} e^{(s-r)A}\left(\sigma(r, X_1(r)) - \sigma(r, X_2(r))\right)dW(r)\right|.$$

Therefore

$$\mathbb{E}|X_1(s) - X_2(s)|^2 \le 5|e^{(s-t_1)A}x_1 - e^{(s-t_2)A}x_2|^2$$
$$+ 5\mathbb{E}\left|\int_{t_1}^{t_2} e^{(s-r)A} b(r, X_1(r))dr\right|^2 + 5\mathbb{E}\left|\int_{t_2}^{s} e^{(s-r)A}\left(b(r, X_1(r)) - b(r, X_2(r))\right)dr\right|^2$$
$$+ 5\mathbb{E}\left|\int_{t_1}^{t_2} e^{(s-r)A}\sigma(r, X_1(r))dW(r)\right|^2$$
$$+ 5\mathbb{E}\left|\int_{t_2}^{s} e^{(s-r)A}\left(\sigma(r, X_1(r)) - \sigma(r, X_2(r))\right)dW(r)\right|^2. \quad (1.86)$$

To estimate the second and the third terms we use Jensen's inequality applied to the inner integral. Using Hypothesis 1.143-(ii) and (1.81) we then obtain

$$\mathbb{E}\left|\int_{t_1}^{t_2} e^{(s-r)A} b(r, X_1(r))dr\right|^2 \le L^2\mathbb{E}\left|\int_{t_1}^{t_2} (s-r)^{-\gamma_1}(1 + |X_1(r)|)dr\right|^2$$
$$\le L^2\left(\int_{t_1}^{t_2} (s-r)^{-\gamma_1}dr\right)\int_{t_1}^{t_2} (s-r)^{-\gamma_1}\mathbb{E}(1 + |X_1(r)|)^2dr$$
$$\le 2L^2[1 + C(1 + |x_1|^2)]\left(\int_{t_1}^{t_2} (s-r)^{-\gamma_1}dr\right)^2$$
$$\le 2L^2[1 + C(1 + |x_1|)^2)]\frac{1}{1-\gamma_1}(t_1 - t_2)^{2(1-\gamma_1)}.$$

In the same way we estimate the third term getting, by Hypothesis 1.143-(ii),

$$\mathbb{E}\left|\int_{t_2}^{s} e^{(s-r)A}\left(b(r, X_1(r)) - b(r, X_2(r))\right)dr\right|^2$$
$$\le L^2\left(\int_{t_2}^{s} (s-r)^{-\gamma_1}dr\right)\int_{t_2}^{s} (s-r)^{-\gamma_1}\mathbb{E}|X_1(r) - X_2(r)|^2dr$$
$$\le \frac{L^2(s-t_2)^{1-\gamma_1}}{1-\gamma_1}\int_{t_2}^{s} (s-r)^{-\gamma_1}\mathbb{E}|X_1(r) - X_2(r)|^2dr. \quad (1.87)$$

The fourth and the fifth term of (1.86) are estimated using the isometry formula. We have

$$\mathbb{E}\left|\int_{t_1}^{t_2} e^{(s-r)A}\sigma(r, X_1(r))dW(r)\right|^2 = \int_{t_1}^{t_2} \mathbb{E}|e^{(s-r)A}\sigma(r, X_1(r))|^2dr$$
$$\le L^2\int_{t_1}^{t_2} (s-r)^{-2\gamma_2}\mathbb{E}(1 + |X_1(r)|)^2dr \le 2L^2[1 + C(1 + |x_1|^2)]\int_{t_1}^{t_2} (s-r)^{-2\gamma_2}dr$$
$$\le 2L^2[1 + C(1 + |x_1|^2)]\frac{1}{1-2\gamma_2}(t_1 - t_2)^{1-2\gamma_2}$$

and

$$\mathbb{E}\left|\int_{t_2}^{s} e^{(s-r)A}\left(\sigma(r,X_1(r))-\sigma(r,X_2(r))\right)dW(r)\right|^2$$

$$= \int_{t_2}^{s} \mathbb{E}|e^{(s-r)A}\left(\sigma(r,X_1(r))-\sigma(r,X_2(r))\right|^2 dr$$

$$\leq L^2 \int_{t_2}^{s} (s-r)^{-2\gamma_2}\mathbb{E}|X_1(r)-X_2(r)|^2 dr.$$

Using all these estimates in (1.86) we obtain, for a suitable constant $C_1 > 0$, for $\gamma_3 := [2(1-\gamma_1)] \wedge [1-2\gamma_2]$ and $\gamma_4 := \gamma_1 \vee [2\gamma_2]$,

$$\mathbb{E}|X_1(s)-X_2(s)|^2 \leq 5|e^{(s-t_1)A}x_1 - e^{(s-t_2)A}x_2|^2 + C_1(1+|x_1|^2)|t_2-t_1|^{\gamma_3} +$$

$$+C_1\int_{t_2}^{s}(s-r)^{-\gamma_4}\mathbb{E}|X_1(r)-X_2(r)|^2 dr.$$

Observing that

$$|e^{(s-t_1)A}x_1 - e^{(s-t_2)A}x_2| \leq Me^{\omega T}|x_1-x_2| + |e^{(s-t_2)A}(e^{(t_2-t_1)A}x_1 - x_1)|,$$

we can thus apply a version of Gronwall's Lemma (see Proposition D.25). It gives us

$$\mathbb{E}|X_1(s)-X_2(s)|^2 \leq C_2\left[|x_1-x_2|^2 + (1+|x_1|^2)|t_2-t_1|^{\gamma_3} + |e^{(t_2-t_1)A}x_1 - x_1|^2\right]$$

for some $C_2 > 0$ with the required properties. □

LEMMA 1.146 *Assume that Hypothesis 1.143 holds with $f_1(s) = Ls^{-\gamma_1}$ and $f_2(s) = Ls^{-\gamma_2}$ for some $L > 0$, $\gamma_1 \in [0,1)$ and $\gamma_2 \in [0,1/2)$. Fix a Λ-valued progressively measurable process $a(\cdot)$. Let X be the unique mild solution of (1.74) described in Theorem 1.144. Let $\{e_i\}_{i\in\mathbb{N}}$ an orthonormal basis of Ξ and, for any $k \in \mathbb{N}$, let $P^k\colon \Xi \to \Xi$ be the orthogonal projection onto $\operatorname{span}\{e_1,...,e_k\}$. Let X^k be the unique mild solution of*

$$\begin{cases} dX^k(s) = (AX^k(s) + e^{\frac{1}{k}A}b(s,X^k(s),a(s)))ds + e^{\frac{1}{k}A}\sigma(s,X^k(s),a(s))P^k dW(s) \\ X^k(0) = x. \end{cases} \tag{1.88}$$

Then, for any $p > 0$, there exists $M_p > 0$ such that

$$\sup_{k\in\mathbb{N}} \mathbb{E}\left[\sup_{s\in[0,T]}|X^k(s)|^p\right] \leq M_p. \tag{1.89}$$

Moreover, for every $s \in [0,T]$,

$$\lim_{k\to\infty}\mathbb{E}\left[|X^k(s)-X(s)|^2\right] = 0 \tag{1.90}$$

and, for every $\varphi \in C_m(H)$ $(m \geq 0)$,

$$\lim_{k\to\infty}\mathbb{E}\left[\varphi(X^k(s))\right] = \mathbb{E}\left[\varphi(X(s))\right]. \tag{1.91}$$

PROOF. To simplify the notation we drop the dependence on the variable a in the expressions for b and σ.

It is easy to see that, thanks to Hypothesis 1.143, for any $k \in \mathbb{N}$, the functions $e^{\frac{1}{k}A}b(s,x)$ and $e^{\frac{1}{k}A}\sigma(s,x)P_k$ satisfy 1.143 with $\tilde{f}_i(s) = L_1 f_i(s), i = 1,2$, where $L_1 = \left(\sup_{r\in[0,1]}\left\|e^{rA}\right\|\right)$. Thus we obtain (1.81) for X_k uniformly in k. This gives (1.89).

We now prove (1.90). We have

$$\mathbb{E}\left|X(s)-X^k(s)\right|^2 \le 2\mathbb{E}\left|\int_0^s e^{(s-r)A}\left(b(r,X(r)) - e^{\frac{1}{k}A}b(r,X^k(r))\right)dr\right|^2 + 2\mathbb{E}\left|\int_0^s e^{(s-r)A}\left(\sigma(r,X(r)) - e^{\frac{1}{k}A}\sigma(r,X^k(r))P^k\right)dW(r)\right|^2.$$

Regarding the first term,

$$\mathbb{E}\left|\int_0^s e^{(s-r)A}\left(b(s,X(s)) - e^{\frac{1}{k}A}b(r,X^k(r))\right)dr\right|^2 \le I_1 + I_2 := 2\mathbb{E}\left|\int_0^s e^{\frac{1}{k}A}e^{(s-r)A}\left(b(s,X(s)) - b(r,X^k(r))\right)dr\right|^2 + 2\mathbb{E}\left|\int_0^s e^{(s-r)A}\left(b(r,X(r)) - e^{\frac{1}{k}A}b(r,X(r))\right)dr\right|^2.$$

We denote $C_s := \int_0^s \tilde{f}_1(s-r)dr$ and we set $\mu_s(dr) = \frac{\tilde{f}_1(s-r)dr}{C_s} = \frac{LL_1(s-r)^{-\gamma_1}}{C_s}$. Using Jensen inequality we get

$$I_1 \le 2\mathbb{E}\left[\int_0^s LL_1(s-r)^{-\gamma_1}\left|X(r)-X^k(r)\right|dr\right]^2 \le 2\mathbb{E}C_s^2\left[\int_0^s \left|X(r)-X^k(r)\right|\mu_s(dr)\right]^2$$
$$\le 2\mathbb{E}C_s^2\int_0^s \left|X(r)-X^k(r)\right|^2\mu_s(dr) = 2LL_1C_s\int_0^s (s-r)^{-\gamma_1}\mathbb{E}\left|X(r)-X^k(r)\right|^2 dr.$$

We also have

$$I_2 \le 2\mathbb{E}\left[\int_0^s \left|\left(I - e^{\frac{1}{k}A}\right)e^{(s-r)A}b(r,X(r))\right|dr\right]^2 =: \omega_1(k,s).$$

By (1.81) and the dominated convergence theorem we have that for every k, $\omega_1(k,\cdot)$ is a bounded measurable function, and for every s, $\lim_{k\to\infty}\omega_1(k,s) = 0$.

We now estimate

$$2\mathbb{E}\left|\int_0^s e^{(s-r)A}\left(\sigma(r,X(r)) - e^{\frac{1}{k}A}\sigma(r,X^k(r))P^k\right)dW(r)\right|^2 \le J_1 + J_2 + J_3$$
$$:= 4\mathbb{E}\left|\int_0^s e^{(s-r)A}\sigma(r,X(r))\left(I-P^k\right)dW(r)\right|^2 + 4\mathbb{E}\left|\int_0^s e^{(s-r)A}\left(\sigma(r,X(r)) - e^{\frac{1}{k}A}\sigma(r,X(r))\right)P^k dW(r)\right|^2$$
$$+ 4\mathbb{E}\left|\int_0^s e^{\frac{1}{k}A}e^{(s-r)A}\left(\sigma(r,X(r)) - \sigma(r,X^k(r))\right)P^k dW(r)\right|^2.$$

Recall that $\Xi_0 = \Xi$. To estimate J_1, we set $Q^k := I - P^k$. We have

$$J_1 = 4\mathbb{E}\left|\int_0^s e^{(s-r)A}\sigma(r, X(r))\left(I - P^k\right) dW(r)\right|^2$$
$$= 4\int_0^s \mathbb{E}\left\|e^{(s-r)A}\sigma(r, X(r))Q^k\right\|^2_{\mathcal{L}_2(\Xi,H)} dr$$
$$= 4\int_0^s \mathbb{E}\sum_{i\in\mathbb{N}}\left\langle e^{(s-r)A}\sigma(r, X(r))Q^k e_i, e^{(s-r)A}\sigma(r, X(r))Q^k e_i\right\rangle ds =: \omega_2(k, s).$$

Observe that

$$\sum_{i\in\mathbb{N}}\left\langle e^{(s-r)A}\sigma(r, X(r))Q^k e_i, e^{(s-r)A}\sigma(r, X(r))Q^k e_i\right\rangle$$
$$= \sum_{i=k+1}^{+\infty}\left\langle e^{(s-r)A}\sigma(r, X(r))e_i, e^{(s-r)A}\sigma(r, X(r))e_i\right\rangle$$
$$\le \sum_{i\in\mathbb{N}}\left\langle e^{(s-r)A}\sigma(r, X(r))e_i, e^{(s-r)A}\sigma(r, X(r))e_i\right\rangle = \left\|e^{(s-r)A}\sigma(r, X(r))\right\|^2_{\mathcal{L}_2(\Xi,H)}.$$

Since the series above has nonnegative terms, we obtain

$$\lim_{k\to\infty}\left\|e^{(s-r)A}\sigma(r, X(r))Q^k\right\|^2_{\mathcal{L}_2(\Xi,H)} = 0 \qquad ds\otimes d\mathbb{P} \text{ a.s.}$$

Therefore, thanks to (1.81), Hypothesis 1.143 and the dominated convergence theorem, we obtain that for every $s \in [0, T]$

$$\lim_{k\to\infty} J_1 \le \lim_{k\to\infty}\omega_2(k, s) = 0.$$

The term J_2 is estimated similarly.

For J_3 we estimate

$$J_3 = 4\mathbb{E}\left|\int_0^s e^{\frac{1}{k}A}e^{(s-r)A}\left(\sigma(r, X(r)) - \sigma(r, X^k(r))\right)P^k dW(r)\right|^2$$
$$= 4\mathbb{E}\left[\int_0^s\left\|e^{\frac{1}{k}A}e^{(s-r)A}\left(\sigma(r, X(r)) - \sigma(r, X^k(r))\right)P^k\right\|^2_{\mathcal{L}_2(\Xi,H)} dr\right]$$
$$\le 4L^2L_1^2\int_0^s (s-r)^{-2\gamma_2}\mathbb{E}\left[\left|X(r) - X^k(r)\right|\right]^2 dr.$$

Combining the above estimates we have

$$\mathbb{E}\left|X(s) - X^k(s)\right|^2 \le \omega_3(k, s) + \int_0^s L_2(s-r)^{-\tilde\gamma}\mathbb{E}\left|X(r) - X^k(r)\right|^2 dr,$$

for some $L_2 > 0$, $\tilde\gamma \in [0, 1)$ and some bounded measurable functions $\omega_3(k, \cdot)$ such that for every $s \in [0, T]$ $\lim_{k\to\infty}\omega_3(k, s) = 0$. The claim (1.90) now follows from a version of Gronwall's Lemma (see Proposition D.25).

Thanks to (1.90), for any subsequence of $X^k(s)$ we can extract a sub-subsequence converging to $X(s)$ almost everywhere and then, thanks to (1.89), (1.81) and the dominated convergence theorem, we obtain (1.91) along the sub-subsequence. This implies (1.91) for the whole sequence $X^k(s)$. □

REMARK 1.147 Observe that, thanks to the specific form of the functions f_1 and f_2 considered in Lemma 1.146 and the fact that b and σ satisfy Hypothesis 1.143, the functions $e^{\frac{1}{k}A}b(s, x, a)$ and $e^{\frac{1}{k}A}\sigma(s, x, a)P_k$ satisfy Hypothesis 1.119. ■

Remark 1.148 If in Hypothesis 1.119 we set $W_Q = Q^{1/2}\tilde{W}$ for a suitable cylindrical Wiener process $\tilde{W}$ in Ξ, and we substitute σ with $\tilde{\sigma} := \sigma Q^{1/2}$, it is easy to see that Hypothesis 1.143 is more general. Also Hypothesis 1.143 is more general than Hypotheses 1.137 and 1.139 if we take f bounded and $a_2(\cdot) \equiv 0$ there. ■

1.6. Transition semigroups

Let $T > 0$ and $t \in [0, T]$. Let H, Ξ, Q, and the generalized reference probability space $\mu = (\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s\in[t,T]}, \mathbb{P}, W_Q)$ be the same as in Section 1.3. Consider now the following SDE with non-random coefficients

$$\begin{cases} dX(s) = (AX(s) + b(s, X(s)))\, ds + \sigma(s, X(s)) dW_Q(s) \\ X(t) = x \in H, \end{cases} \tag{1.92}$$

where $b\colon [0, T] \times H \to H$ and $\sigma\colon [0, T] \times H \to \mathcal{L}_2(\Xi_0, H)$. If Hypothesis 1.119, where we drop the dependence on a in all conditions, (respectively, Hypotheses 1.137 and 1.139 with $a_2(\cdot) \equiv 0$ and with no dependence on a_1, respectively, Hypothesis 1.143 with no dependence on a) is satisfied, then Theorem 1.121 (respectively, Proposition 1.141, respectively, Theorem 1.144) ensures that (1.92) has a unique mild solution $X(\cdot; t, x)$.

Notation 1.149 (Transition semigroup) The (two parameter) transition semigroup $P_{t,s}$ associated to equation (1.92) is defined as follows. For a bounded, Borel measurable function $\phi\colon H \to \mathbb{R}$ and $0 \le t \le s \le T$,

$$\begin{cases} P_{t,s}\phi\colon H \to \mathbb{R} \\ P_{t,s}\phi\colon x \mapsto \mathbb{E}[\phi(X(s; t, x))]. \end{cases} \tag{1.93}$$

■

Under our assumptions, by the measurability properties of the solution $X(s; t, x)$ with respect to the initial datum x, $P_{t,s}\phi$ is a bounded measurable function on H for all $\phi \in B_b(H)$. Moreover, thanks to estimates (1.37), (1.69) and (1.79), $P_{t,s}\phi$ is also well defined for any $\phi \in B_m(H)$, $m > 0$.

Definition 1.150 *[Feller and strong Feller property] A transition semigroup $P_{t,s}$ defined in (1.93) is said to possess the* Feller property *if*

$$P_{t,s}(C_b(H)) \subseteq C_b(H)$$

and it is said to possess the strong Feller property *if*

$$P_{t,s}(B_b(H)) \subseteq C_b(H).$$

Theorem 1.151 *[Feller property of the transition semigroup] Assume that Hypothesis 1.143 is satisfied with b and σ independent of a and with $f_1 = L_1 s^{-\gamma_1}$, for $L_1 > 0$ and $\gamma_1 \in (0, 1)$, and $f_2 = L_2 s^{-\gamma_2}$, for $L_2 > 0$ and $\gamma_2 \in (0, 1/2)$. Then for every $\phi \in C_m(H)$ ($m \ge 0$), the function $P_{t,s}\phi\colon H \to \mathbb{R}$ belongs to $C_m(H)$. The same holds if we assume that Hypotheses 1.137 and 1.139 hold without dependence on a_1 and with $a_2(\cdot) = 0$.*

Proof. The result is a consequence of the continuous dependence and growth estimates of Theorem 1.144 and Propositions 1.145, 1.141. □

Theorem 1.152 (Markov property of the transition semigroup) *Let the assumptions of Proposition 1.145 be satisfied with b and σ independent of a. Then for every $\phi \in B_m(H)$ ($m \ge 0$) and $0 \le t \le s \le r \le T$,*

$$\mathbb{E}\phi(X(r; t, x)|\mathscr{F}_s) = P_{s,r}\phi(X(s; t, x)) \quad \mathbb{P} - \textit{almost surely},$$

and

$$P_{t,r}\phi(x) = P_{t,s}(P_{s,r}\phi)(x) \qquad \text{for all } x \in H.$$

The same result is true if Hypotheses 1.137 and 1.139 hold without dependence on a_1 *and with* $a_2(\cdot) = 0$.

PROOF. See [130] Theorem 9.14, page 248, and Corollary 9.15, page 249. The hypothesis in the cylindrical case is a little different from the cylindrical case contained in [130] however the same arguments can be easily adapted using Proposition 1.145 in the last part of the proof. The proof in [130] is given for $\phi \in B_b(H)$ but the argument is exactly the same when $\phi \in B_m(H)$ $(m > 0)$ simply recalling that the operator $P_{t,s}$ is well defined on such functions thanks to estimate (1.81). □

As a particular case we have the following corollary.

COROLLARY 1.153 *Let the assumptions of Proposition 1.145 be satisfied with* b *and* σ *independent of* a *and the time variable* t. *Equation (1.92) then reduces to*

$$\begin{cases} dX(s) = (AX(s) + b(X(s)))\,ds + \sigma(X(s))dW_Q(s) \\ X(0) = x \in H. \end{cases} \tag{1.94}$$

Denote by $X(\cdot;x)$ *the unique mild solution of this equation. In this case, for any* $\phi \in B_m(H)$ $(m \geq 0)$ *and* $s \geq 0$, *we define* $P_s\phi$ *as follows*

$$\begin{cases} P_s\phi\colon H \to \mathbb{R} \\ P_s\phi\colon x \mapsto \mathbb{E}\phi(X(s;x)). \end{cases} \tag{1.95}$$

Then

$$P_{s+r}\phi(x) = P_s(P_r\phi)(x) \qquad \text{for all } x \in H,\ s,r \geq 0.$$

1.7. Itô's and Dynkin's formulae

In this section we assume that $T > 0$, H, Ξ, Q, and the generalized reference probability space $\mu = (\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s\in[0,T]}, \mathbb{P}, W_Q)$ are the same as in Section 1.3. The operator A is the generator of a C_0-semigroup on H, and Λ is a Polish space. The various Itô's and Dynkin's formulae presented in this section are used in proving existence of viscosity solutions (Chapter 3) and verification theorems (Chapters 4 and 5).

THEOREM 1.154 (Itô's Formula) *Assume that* Φ *is a process in* $\mathcal{N}_Q^2(0,T;H)$, f *is an* H*-valued progressively measurable (*$\mathbb{P}$*-a.s.) Bochner integrable process on* $[0,T]$, *and define, for* $s \in [0,T]$,

$$X(s) := X(0) + \int_0^s f(r)dr + \int_0^s \Phi(r)dW_Q(r),$$

where $X(0)$ *is an* $\mathscr{F}_0$*-measurable* H*-valued random variable. Consider* $F\colon [0,T] \times H \to \mathbb{R}$ *and assume that* F *and its derivatives* F_t, DF, D^2F, *are continuous and bounded on bounded subsets of* $[0,T] \times H$. *Then, for* $\mathbb{P}$ *a.e.* ω,

$$\begin{aligned} F(s,X(s)) = F(0,X(0)) + \int_0^s F_t(r,X(r))dr + \int_0^s \langle DF(r,X(r)), f(r)\rangle\, dr + \\ + \int_0^s \langle DF(r,X(r)), \Phi(r)dW_Q(r)\rangle + \\ + \frac{1}{2}\int_0^s \mathrm{Tr}\left[\left(\Phi(r)Q^{1/2}\right)\left(\Phi(r)Q^{1/2}\right)^* D^2F(r,X(r))\right] dr \qquad \text{on } [0,T]. \end{aligned} \tag{1.96}$$

PROOF. See [220], Theorems 2.9 and 2.10. See also, under the assumption of uniform continuity on bounded sets of F and its derivatives, [130] Theorem 4.32 page 106. $\square$

PROPOSITION 1.155 *Let $F\colon [0,T]\times H \to \mathbb{R}$ be such that F and its derivatives F_t, DF, D^2F are continuous in $[0,T]\times H$. Suppose that $DF\colon [0,T]\times H \to D(A^*)$, that A^*DF is continuous in $[0,T]\times H$, and that there exist $C>0$ and $N\geq 1$ such that*

$$|F(s,x)|+|DF(s,x)|+|F_t(s,x)|+\|D^2F(s,x)\|+|A^*DF(s,x)| \leq C(1+|x|)^N \quad (1.97)$$

for all $x\in H$, $s\in[0,T]$. Let $f\in M_\mu^{4N}(0,T;H)$, $\Phi\in\mathcal{N}_Q^{4N}(0,T;H)$ and $x\in H$. Let $X(\cdot)$ be the unique mild solution of (1.43) such that $X(0)=x$. Then, for $\mathbb{P}$ a.e. ω,

$$\begin{aligned} F(s,X(s)) = F(0,x) &+ \int_0^s F_t(r,X(r))dr \\ &+ \int_0^s \langle A^*DF(r,X(r)),X(r)\rangle\, dr + \int_0^s \langle DF(r,X(r)),f(r)\rangle\, dr \\ &+ \frac{1}{2}\int_0^s \mathrm{Tr}\left[\left(\Phi(r)Q^{1/2}\right)\left(\Phi(r)Q^{1/2}\right)^* D^2F(r,X(r))\right] dr \\ &+ \int_0^s \langle DF(r,X(r)),\Phi(r)dW_Q(r)\rangle \qquad \text{on } [0,T]. \quad (1.98)\end{aligned}$$

PROOF. Since both sides of (1.98) are continuous processes, it is enough to prove the formula for a single s. We approximate $X(\cdot)$ by the sequence $X^n(\cdot)$ introduced in Proposition 1.126 (observe that the hypotheses of the proposition are satisfied with $p=4N$). By definition $X^n(\cdot)$ solves the integral equation

$$X^n(s) = \int_0^s (A_nX^n(r)+f(r))\,dr + \int_0^s \Phi(r)dW_Q(r).$$

We can apply Itô's formula (1.96) (A_n^* is the adjoint of A_n) obtaining

$$\begin{aligned} F(s,X^n(s)) = F(0,x) &+ \int_0^s F_t(r,X^n(r))dr \\ &+ \int_0^s \langle A_n^*DF(r,X^n(r)),X^n(r)\rangle\, dr + \int_0^s \langle DF(r,X^n(r)),f(r)\rangle\, dr \\ &+ \frac{1}{2}\int_0^s \mathrm{Tr}\left[\left(\Phi(r)Q^{1/2}\right)\left(\Phi(r)Q^{1/2}\right)^* D^2F(r,X^n(r))\right] dr \\ &+ \int_0^s \langle DF(r,X^n(r)),\Phi(r)dW_Q(r)\rangle\,. \quad (1.99)\end{aligned}$$

Estimates (1.45) and (1.46) yield,

$$\sup_{n\geq 1}\mathbb{E}\left[\sup_{s\in[0,T]}\left(|X^n(s)|^{4N}+|X(s)|^{4N}\right)\right]\leq C, \qquad (1.100)$$

and, up to a subsequence (still denoted by X^n),

$$\lim_{n\to+\infty}\sup_{s\in[0,T]}|X^n(s)-X(s)|=0 \quad \text{for } a.e.\ \omega. \qquad (1.101)$$

Therefore, (1.100), (1.101), (1.97), and the hypotheses on f and Φ guarantee that for a.e. ω, the integrands of the first four integrals of (1.99) converge to the integrands of the corresponding integrals of (1.98), and that the convergence is dominated. (For the second integral we use a well known fact that if $x_n\to x\in D(A^*)$ and $A^*x_n\to A^*x$, then $A_n^*x_n\to A^*x$.) It then follows from the Lebesgue dominated convergence theorem that, up to a subsequence, the first four integrals of (1.99) converge to the corresponding integrals of (1.98). Obviously $F(s, X^n(s))$ also converges to $F(s, X(s))$.

Regarding the stochastic integral, we have

$$\mathbb{E}\left|\int_0^s \langle DF(r, X^n(r)), \Phi(r)dW_Q(r)\rangle - \int_0^s \langle DF(r, X(r)), \Phi(r)dW_Q(r)\rangle\right|^2$$
$$\leq \mathbb{E}\int_0^s |DF(r, X^n(r)) - DF(r, X(r))|^2 \|\Phi(r)\|^2_{\mathcal{L}_2(\Xi_0, H)} dr$$
$$\leq \left(\mathbb{E}\int_0^T |DF(r, X^n(r)) - DF(r, X(r))|^4 dr\right)^{1/2} \|\Phi\|^2_{\mathcal{N}_Q^4(t,T;H)} \to 0 \text{ as } n \to +\infty$$

by (1.100), (1.101), (1.97), and the Lebesgue dominated convergence theorem. Therefore, up to a subsequence, we have that for a.e. ω

$$\int_0^s \langle DF(r, X^n(r)), \Phi(r)dW_Q(r)\rangle \to \int_0^s \langle DF(r, X(r)), \Phi(r)dW_Q(r)\rangle \quad \text{as } n \to +\infty.$$

Thus, passing to the limit along a subsequence in (1.99) yields (1.98). □

PROPOSITION 1.156 *Let b and σ satisfy Hypothesis 1.119 and let $a\colon [t, T] \to \Lambda$ be a progressively measurable process. Let $X(\cdot)$ be the unique mild solution of (1.31) such that $X(0) = x \in H$. Let F satisfy the hypotheses of Proposition 1.155. Then:*

(i) *For $\mathbb{P}$ a.e. ω,*

$$\begin{aligned} F(s, X(s)) &= F(0, x) + \int_0^s F_t(r, X(r))dr \\ &+ \int_0^s \langle A^* DF(r, X(r)), X(r)\rangle\, dr + \int_0^s \langle DF(r, X(r)), b(r, X(r), a(r))\rangle\, dr \\ &+ \frac{1}{2}\int_0^s \mathrm{Tr}\left[\left(\sigma(r, X(r), a(r))Q^{1/2}\right)\left(\sigma(r, X(r), a(r))Q^{1/2}\right)^* D^2F(r, X(r))\right] dr \\ &+ \int_0^s \langle DF(r, X(r)), \sigma(r, X(r), a(r))dW_Q(r)\rangle \qquad \text{on } [0, T]. \end{aligned} \tag{1.102}$$

(ii) *Let η be a real process solving*

$$\begin{cases} d\eta(s) = \tilde{b}(s)ds \\ \eta(0) = \eta_0 \in \mathbb{R}, \end{cases}$$

where $\tilde{b}\colon [0, T] \to \mathbb{R}$ is bounded and progressively measurable. Then, for $\mathbb{P}$ a.e. ω,

$$\begin{aligned} F(s, X(s))\eta(s) &= F(0, x)\eta_0 + \int_0^s (F_t(r, X(r))\eta(r) + F(r, X(r))\tilde{b}(r))dr \\ &+ \int_0^s \langle A^* DF(r, X(r)), X(r)\rangle\, \eta(r)dr + \int_0^s \langle DF(r, X(r)), b(r, X(r), a(r))\rangle\, \eta(r)dr \\ &+ \frac{1}{2}\int_0^s \mathrm{Tr}\left[\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)^* D^2F(r, X(r))\right] \eta(r)dr \\ &+ \int_0^s \langle DF(r, X(r))\eta(r), \sigma(r, X(r), a(r))dW_Q(r)\rangle \qquad \text{on } [0, T]. \end{aligned} \tag{1.103}$$

In particular

$$\mathbb{E}\left[F(s,X(s))\eta(s)\right] = F(0,x)\eta_0 + \mathbb{E}\int_0^s (F_t(r,X(r))\eta(r) + F(r,X(r))\tilde{b}(r))dr$$
$$+\mathbb{E}\int_0^s \langle A^*DF(r,X(r)),X(r)\rangle\,\eta(r)dr + \mathbb{E}\int_0^s \langle DF(r,X(r)),b(r,X(r),a(r))\rangle\,\eta(r)dr$$
$$+\frac{1}{2}\mathbb{E}\int_0^s \mathrm{Tr}\left[\left(\sigma(r,X(r),a(r))Q^{\frac{1}{2}}\right)\left(\sigma(r,X(r),a(r))Q^{\frac{1}{2}}\right)^* D^2F(r,X(r))\right]\eta(r)dr. \tag{1.104}$$

PROOF. Part (i) follows directly from Proposition 1.155 applied with $f(s) := b(s,a(s),X(s))$ and $\Phi(s) := \sigma(s,a(s),X(s))$, noticing that, thanks to (1.34), (1.35) and (1.38), we have $f \in M^p_\mu(0,T;H)$ and $\Phi \in \mathcal{N}^p_Q(0,T;H)$ for every $p \geq 1$.

Part (ii) is a corollary of (i). We introduce the Hilbert space $\hat{H} := H \times \mathbb{R}$ (with the usual inner product), and set

$$\hat{A} = \begin{pmatrix} A \\ 0 \end{pmatrix}, \quad \hat{b} = \begin{pmatrix} b \\ \tilde{b} \end{pmatrix}, \quad \hat{\sigma} = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}.$$

Then the process

$$\hat{X}(s) = \begin{pmatrix} X(s) \\ \eta(s) \end{pmatrix}$$

is the mild solution of the SDE

$$\begin{cases} d\hat{X}(s) = \left(\hat{A}\hat{X}(s) + \hat{b}(s,\hat{X}(s),a(s))\right)ds + \hat{\sigma}(s,\hat{X}(s),a(s))dW_Q(s) \\ \hat{X}(0) = \begin{pmatrix} x \\ \eta_0 \end{pmatrix}. \end{cases}$$

Therefore, (1.103) follows from (1.102) applied to the function $\hat{F}(s,\hat{x}) = F(s,x)\eta_0$, where $\hat{x} = (x,\eta_0)$. Taking expectation in (1.103) we obtain (1.104). □

PROPOSITION 1.157 *Let b and σ satisfy Hypothesis 1.119 and let $a\colon [t,T] \to \Lambda$ be a progressively measurable process. Assume in addition that A is maximal dissipative. Let $X(\cdot)$ be the unique mild solution of (1.31) such that $X(0) = x \in H$. Let $F \in C^{1,2}([0,T] \times H)$ be of the form $F(t,x) = \varphi(t,|x|)$ for some $\varphi(t,r) \in C^{1,2}([0,T] \times \mathbb{R})$, where $\varphi(t,\cdot)$ is even and non-decreasing on $[0,+\infty)$. Moreover suppose that there exist $C > 0$ $N \geq 1$ such that*

$$|F(s,x)| + |DF(s,x)| + |F_t(s,x)| + \|D^2F(s,x)\| \leq C(1+|x|)^N \tag{1.105}$$

for all $x \in H$, $s \in [0,T]$. Then:

(i) *For $\mathbb{P}$ a.e. ω,*

$$F(s,X(s)) \leq F(0,x) + \int_0^s \Bigg[F_t(r,X(r)) + \langle b(r,X(r),a(r)),DF(r,X(r))\rangle$$
$$+\frac{1}{2}\mathrm{Tr}\left[\left(\sigma(r,X(r),a(r))Q^{\frac{1}{2}}\right)\left(\sigma(r,X(r),a(r))Q^{\frac{1}{2}}\right)^* D^2F(r,X(r))\right]\Bigg]dr$$
$$+\int_0^s \langle DF(r,X(r)),b(r,X(r),a(r))dW_Q(r)\rangle \qquad \text{on } [0,T]. \tag{1.106}$$

(ii) *If η is as in part (ii) of Proposition 1.156 and η is positive then, for $\mathbb{P}$ a.e. ω,*

$$
\begin{aligned}
F(s, X(s))\eta(s) \le F(0,x)\eta_0 & \\
&+ \int_0^s (F_t(r, X(r))\eta(r) + F(r, X(r))\tilde{b}(r))dr \\
&+ \int_0^s \langle DF(r, X(r)), b(r, X(r), a(r))\rangle \, \eta(r)dr \\
+ \frac{1}{2}\int_0^s \mathrm{Tr}\Big[\Big(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\Big)&\Big(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\Big)^* D^2F(r, X(r))\Big]\,\eta(r)dr \\
&+ \int_0^s \langle DF(r, X(r))\eta(r), \sigma(r, X(r), a(r))dW_Q(r)\rangle \qquad \text{on } [0,T]. \qquad (1.107)
\end{aligned}
$$

In particular

$$
\begin{aligned}
\mathbb{E}\left[F(s, X(s))\eta(s)\right] \le F(0,x)\eta_0 & \\
&+ \mathbb{E}\int_0^s (F_t(r, X(r))\eta(r) + F(r, X(r))\tilde{b}(r))dr \\
&+ \mathbb{E}\int_0^s \langle DF(r, X(r)), b(r, X(r), a(r))\rangle \, \eta(r)dr \\
+ \frac{1}{2}\mathbb{E}\int_0^s \mathrm{Tr}\Big[\Big(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\Big)&\Big(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\Big)^* D^2F(r, X(r))\Big]\,\eta(r)dr. \qquad (1.108)
\end{aligned}
$$

Proof. (*i*): We set $f(s) := b(s, a(s), X(s))$ and $\Phi(s) := \sigma(s, a(s), X(s))$ and consider the approximation $X^n(\cdot)$ of $X(\cdot)$ as in Proposition 1.126. Observe that, thanks to (1.34), (1.35) and (1.38) we have $f \in M^p_\mu(0,T;H)$ and $\Phi \in \mathcal{N}^p_Q(0,T;H)$ for every $p \ge 1$ so the assumptions of Proposition 1.126 are satisfied.

We observe that $DF(s,x) = \frac{\partial\varphi}{\partial r}(s,|x|)\frac{x}{|x|}$ and, since $\varphi(s,\cdot)$ is non-decreasing on $[0,+\infty)$, $\frac{\partial\varphi}{\partial r}(s,r) \ge 0$. Therefore, since A, and thus A_n, is dissipative,

$$
\langle A_n X^n(s), DF(r, X^n(s))\rangle = \frac{\partial\varphi}{\partial r}(s, |X^n(s)|)\frac{1}{|X^n(s)|}\langle A_n X^n(s), X^n(s)\rangle \le 0 \qquad (1.109)
$$

for every $s \ge 0$.

Hence, applying Itô formula for $X^n(\cdot)$ and using (1.109), we obtain

$$
\begin{aligned}
F(s, X^n(s)) = F(0,x) &+ \int_0^s \Big[F_t(r, X^n(r)) + \langle A_n X^n(r), DF(r, X^n(r))\rangle \\
&+ \langle f(r), DF(r, X^n(r))\rangle + \frac{1}{2}\mathrm{Tr}\Big[\Big(\Phi(r)Q^{\frac{1}{2}}\Big)\Big(\Phi(r)Q^{\frac{1}{2}}\Big)^* D^2F(r, X^n(r))\Big]\Big]dr \\
&+ \int_0^s \langle DF(r, X^n(r)), b(r, X^n(r), a(r))dW_Q(r)\rangle . \\
&\le F(0,x) + \int_0^s \Big[F_t(r, X^n(r)) + \langle f(r), DF(r, X^n(r))\rangle \\
&+ \frac{1}{2}\mathrm{Tr}\Big[\Big(\Phi(r)Q^{\frac{1}{2}}\Big)\Big(\Phi(r)Q^{\frac{1}{2}}\Big)^* D^2F(r, X^n(r))\Big]\Big]dr \\
&+ \int_0^s \langle DF(r, X^n(r)), b(r, X^n(r), a(r))dW_Q(r)\rangle . \qquad (1.110)
\end{aligned}
$$

It remains to pass to the limit as $n \to +\infty$ in (1.110). This can be done following the same arguments as these used in the proof of Proposition 1.155.

(ii): The proof combines the proof of (i) with the arguments used in the proof of Proposition 1.156-(ii). □

REMARK 1.158 Propositions 1.156 and 1.157 are used to work with test functions in Chapter 3. In particular parts (ii) of them are useful when discount factors are present (see e.g. Lemma 3.65). Finally we remark that Theorem 1.154 and Propositions 1.155-1.157 are still true if s there is replaced by a stopping time τ on $[0, T]$. ■

The next two non-standard versions of Dynkin's formula will be used to prove verification theorems in Chapters 4 and 5.

PROPOSITION 1.159 *Let $Q = I$. Assume that Hypothesis 1.143 holds with $f_1(s) = Ls^{-\gamma_1}$ and $f_2(s) = Ls^{-\gamma_2}$ for some $L > 0$, $\gamma_1 \in [0, 1)$ and $\gamma_2 \in [0, 1/2)$. Assume that there exists $\lambda \in \mathbb{R}$ such that $(\lambda I - A)^{-1}b\colon [0, T] \times H \times \Lambda \to H$ is measurable and is continuous in the x variable. Assume that $\sigma : [0, T] \times H \times \Lambda \to \mathcal{L}(\Xi, H)$ and is continuous in the x variable. Suppose moreover that there exists $C > 0$ such that, for all $(t, x, a) \in [0, T] \times H \times \Lambda$,*

$$\begin{cases} |(\lambda I - A)^{-1}b(t, x, a)| \le C(1 + |x|) \\ \|\sigma(t, x, a)\|_{\mathcal{L}(\Xi,H)} \le C(1 + |x|). \end{cases} \tag{1.111}$$

Fix a Λ-valued progressively measurable process $a(\cdot)$. Let X be the unique mild solution of (1.74) described in Theorem 1.144. Let $F\colon [0, T] \times H \to \mathbb{R}$ be such that F and its derivatives F_t, DF, D^2F are continuous in $[0, T] \times H$. Suppose that $DF\colon [0, T] \times H \to D(A^*)$, that A^*DF is continuous in $[0, T] \times H$, that $D^2F : [0, T] \times H \to \mathcal{L}_1(H)$ is continuous, and that there exist $C > 0$ and $N \ge 1$ such that

$$|F(s, x)| + |DF(s, x)| + |F_t(s, x)| + \|D^2F(s, x)\|_{\mathcal{L}_1(H)} + |A^*DF(s, x)| \le C(1 + |x|)^N. \tag{1.112}$$

Then, for any $s \in [0, T]$,

$$\begin{aligned} \mathbb{E}\left[F(s, X(s))\right] = F(0, x) + \mathbb{E}\int_0^s F_t(r, X(r))dr \\ + \mathbb{E}\int_0^s \langle A^*DF(r, X(r)), X(r)\rangle\, dr \\ + \mathbb{E}\int_0^s \left\langle (\lambda I - A^*)DF(r, X(r)), (\lambda I - A)^{-1}b(r, X(r), a(r))\right\rangle dr \\ + \frac{1}{2}\mathbb{E}\int_0^s \mathrm{Tr}\left[\sigma(r, X(r), a(r))\sigma(r, X(r), a(r))^* D^2F(r, X(r))\right] dr. \end{aligned} \tag{1.113}$$

PROOF. We will use the notation $b(r, x) := b(r, x, a(r))$, $\sigma(r, x) := \sigma(r, x, a(r))$. We approximate the process $X(\cdot)$ by processes $X^k(\cdot)$ from Lemma 1.146. As observed in Remark 1.147, the approximating problems satisfy the hypotheses of Proposition 1.156 so we have

$$\begin{aligned} \mathbb{E}\left[F(s, X^k(s))\right] = F(0, x) + \int_0^s \mathbb{E}\, F_t(r, X^k(r))dr \\ + \int_0^s \mathbb{E}\left\langle A^*DF(r, X^k(r)), X^k(r)\right\rangle dr \\ + \int_0^s \mathbb{E}\left\langle DF(r, X^k(r)), e^{\frac{1}{k}A}b(r, X^k(r))\right\rangle dr \\ + \frac{1}{2}\int_0^s \mathbb{E}\,\mathrm{Tr}\left[e^{\frac{1}{k}A}\sigma(r, X^k(r))P_k(e^{\frac{1}{k}A}\sigma(r, X^k(r))P_k)^* D^2F(r, X^k(r))\right] dr. \end{aligned} \tag{1.114}$$

The claim will follow if we can pass to the limit as $k \to +\infty$ in each term of (1.114).

Using (1.90), (1.89), (1.81) and the the dominated converge theorem it is easy to see that

$$\lim_{k\to\infty} |X(\cdot) - X^k(\cdot)|_{M^2_\mu(0,T;H)} = 0.$$

Therefore we can find a subsequence, still denoted by $X^k(\cdot)$, that converges to $X(\cdot)$ $dt \otimes d\mathbb{P}$ a.e..

We will only show how to prove the convergence of the last two terms in (1.114) since the arguments for other terms are similar. Using the assumptions it is obvious that

$$\begin{aligned}
&\left\langle DF(r, X^k(r)), e^{\frac{1}{k}A} b(r, X^k(r)) \right\rangle \\
&\qquad = \left\langle (\lambda I - A^*) DF(r, X^k(r)), e^{\frac{1}{k}A} (\lambda I - A)^{-1} b(r, X^k(r)) \right\rangle \\
&\qquad\qquad \to \left\langle (\lambda I - A^*) DF(r, X(r)), (\lambda I - A)^{-1} b(r, X(r)) \right\rangle
\end{aligned}$$

as $k \to +\infty$. Moreover, thanks to (1.111), (1.112) and (1.89),

$$\begin{aligned}
&\int_0^s \mathbb{E} \left| \left\langle (\lambda I - A^*) DF(r, X^k(r)), e^{\frac{1}{k}A} (\lambda I - A)^{-1} b(r, X^k(r)) \right\rangle \right|^2 dr \\
&\qquad\qquad \le C_1 \int_0^s \mathbb{E}(1 + |X^k(r)|)^{2(N+1)} dr \le C_2
\end{aligned}$$

for some C_2 independent of k. Similarly we obtain

$$\int_0^s \mathbb{E} \left| \left\langle (\lambda I - A^*) DF(r, X(r)), (\lambda I - A)^{-1} b(r, X(r)) \right\rangle \right|^2 dr \le C_3$$

for some C_3. Therefore it follows from Lemma 1.50 that

$$\begin{aligned}
&\lim_{k\to+\infty} \int_0^s \mathbb{E} \left\langle DF(r, X^k(r)), e^{\frac{1}{k}A} b(r, X^k(r)) \right\rangle dr \\
&\qquad = \int_0^s \mathbb{E} \left\langle (\lambda I - A^*) DF(r, X(r)), (\lambda I - A)^{-1} b(r, X(r)) \right\rangle dr.
\end{aligned}$$

Regarding the last term in (1.114),

$$\begin{aligned}
&\mathrm{Tr}\left[e^{\frac{1}{k}A} \sigma(r, X^k(r)) P_k (e^{\frac{1}{k}A} \sigma(r, X^k(r)) P_k)^* D^2 F(r, X^k(r)) \right] \\
&\qquad - \mathrm{Tr}\left[\sigma(r, X(r)) \sigma(r, X(r))^* D^2 F(r, X(r)) \right] = I_1 + I_2 := \\
&\mathrm{Tr}\left[e^{\frac{1}{k}A} \sigma(r, X^k(r)) P_k (e^{\frac{1}{k}A} \sigma(r, X^k(r)) P_k)^* (D^2 F(r, X^k(r)) - D^2 F(r, X(r))) \right] \\
&+\mathrm{Tr}\left[\left[e^{\frac{1}{k}A} \sigma(r, X^k(r)) P_k (e^{\frac{1}{k}A} \sigma(r, X^k(r)) P_k)^* - \sigma(r, X(r)) \sigma(r, X(r))^* \right] D^2 F(r, X(r))) \right].
\end{aligned}$$

By Proposition B.26 and (1.111) we have

$$|I_1| \le C_1 (1 + |X^k(r)|)^2 \| D^2 F(r, X^k(r)) - D^2 F(r, X(r)) \|_{\mathcal{L}_1(H)} \to 0 \quad \text{as} \ \ k \to +\infty$$

$dt \otimes d\mathbb{P}$ a.e.. Let $\{e_1, e_2, ...\}$ be an orthonormal basis of eigenvectors of $D^2 F(r, X(r))$ and $\lambda_1, \lambda_2, ...$ be the corresponding eigenvalues. Then

$$\begin{aligned}
&\mathrm{Tr}\left[e^{\frac{1}{k}A} \sigma(r, X^k(r)) P_k (e^{\frac{1}{k}A} \sigma(r, X^k(r)) P_k)^* D^2 F(r, X(r))) \right] \\
&\qquad = \sum_{n=1}^\infty \lambda_n \left| P_k \sigma(r, X^k(r))^* e^{\frac{1}{k}A^*} e_n \right|_\Xi^2 \to \sum_{n=1}^\infty \lambda_n \left| \sigma(r, X(r))^* e_n \right|_\Xi^2 \\
&\qquad\qquad = \mathrm{Tr}\left[\sigma(r, X(r)) \sigma(r, X(r))^* D^2 F(r, X(r))) \right] \quad \text{as} \ \ k \to +\infty
\end{aligned}$$

$dt \otimes d\mathbb{P}$ a.e. Therefore $\lim_{k\to+\infty}(I_1 + I_2) = 0$ $dt \otimes d\mathbb{P}$ a.e. Since we also have

$$\int_0^s \mathbb{E}|I_1 + I_2|^2 dr \leq C_2$$

for some constant C_2 independent of k, the convergence of the last term in (1.114) now follows from Lemma 1.50. □

REMARK 1.160 It is obvious from the proof of Proposition 1.159 that the assumption $\sigma : [0,T] \times H \times \Lambda \to \mathcal{L}(\Xi, H)$ and is continuous in the x variable can be weakened by the requirement that $\sigma : [0,T] \times H \times \Lambda \to \mathcal{L}(\Xi, H)$ and $\sigma^* v$ is continuous in the x variable for every $v \in H$. ■

PROPOSITION 1.161 *Let Hypotheses 1.137 and 1.139 be satisfied with $f(s) = Ls^{-\gamma_1}$ for some $L > 0$, $\gamma_1 \in [0,1)$. Assume there exists $\gamma > 0$ such that (1.67) is satisfied. Suppose moreover that there exist $\lambda \in \mathbb{R}$ such that $(\lambda I - A)^{-1} a_2(\cdot) : [0,T] \times \Omega \to H$ is in $M^2_\mu(0,T;H)$. Let X be the unique mild solution of (1.68) described in Proposition 1.141. Let $F\colon [0,T] \times H \to \mathbb{R}$ be such that F and its derivatives F_t, DF, D^2F are continuous in $[0,T] \times H$. Suppose that $DF\colon [0,T] \times H \to D(A^*)$, that A^*DF is continuous in $[0,T]\times H$, that $D^2F : [0,T]\times H \to \mathcal{L}_1(H)$ is continuous, and that there exists $C > 0$ such that (1.112) holds with $N = 0$. Then, for any $s \in [0,T]$,*

$$\begin{aligned}
\mathbb{E}\left[F(s, X(s))\right] &= F(0,x) + \mathbb{E}\int_0^s F_t(r, X(r))dr \\
&+ \mathbb{E}\int_0^s \langle A^*DF(r, X(r)), X(r)\rangle\, dr + \mathbb{E}\int_0^s \langle DF(r, X(r)), b_0(r, X(r), a_1(r))\rangle\, dr \\
&+ \mathbb{E}\int_0^s \big\langle (\lambda I - A^*)DF(r, X(r)), (\lambda I - A)^{-1} a_2(r)\big\rangle\, dr \\
&+ \frac{1}{2}\mathbb{E}\int_0^s \operatorname{Tr}\left[\sigma Q \sigma^* D^2F(r, X(r))\right] dr.
\end{aligned}$$

PROOF. The proof is similar to the proof of Proposition 1.159 so we simply comment how to adapt the latter. Without loss of generality we can assume that $Q = I$. First, as we did in Lemma 1.146 for the solutions of (1.74), we need to approximate the solution X. For example we can use the sequence X^k of the solutions of the problems

$$\begin{cases} dX^k(s) = \left[AX^k(s) + b^k(s, X^k(s))\right] ds + \sigma P^k dW(s), \quad s > 0, \\ X(0) = x, \end{cases} \tag{1.115}$$

where, for sufficiently big $k \in \mathbb{N}$,

- $P^k\colon \Xi \to \Xi$ is the projection on the span of the first k vectors of an orthonormal basis $\{e_i\}_{i\in\mathbb{N}}$ of Ξ.
- $b^k(s,x) = b_0^k(s,x,a_1(s)) + (\lambda I - A_k)(\lambda I - A)^{-1} a_2(s) \mathbf{1}_{B_k}(s,\omega)$, where $b_0^k(s,x,a_1(s)) = b_0(s,x,a_1(s))\phi(ks)$, where $\phi\colon [0,+\infty) \to [0,1]$ is a smooth, increasing function equal to 0 on the interval $[0,1]$ and equal to 1 on the interval $[2,+\infty)$, A_k is the Yosida approximation of A, and $B_k = \{(s,\omega) : |(\lambda I - A)^{-1} a_2(s,\omega)| \leq k\}$.

Using arguments similar to these in the proof of Lemma 1.146 for the solutions of (1.74) one can prove analogues of (1.89) for $p = q, p \geq 2$ (see Hypothesis 1.139), (1.90) and (1.91) and then argue as in the proof Proposition 1.159. We leave the details to the reader. □

REMARK 1.162 Proposition 1.161 is also true if (1.112) holds with some $N \geq 1$ however then one has to assume that q in Hypothesis 1.139 is large enough (depending on N and β). ■

1.8. Bibliographical notes

Section 1.1 contains elements of basic probability and measure theory. Classical references are for example [12, 42, 44, 199, 281, 374, 407]. We refer in particular to [42, 44, 199, 281] for general theory of measure and probability (Section 1.1.1) and to [42, 44, 199, 407] for results on measurability (Section 1.1.2 and Section 1.1.3). For the Bochner integral and the integration of Banach-valued functions (Section 1.1.3) the reader can consult [137, 138, 140, 306]; some results, useful from the stochastic calculus perspective are contained in [130]. For Sections 1.1.4 and 1.1.5 conditional expectation for Banach-valued random variables the reader can see [130, 267, 281, 374, 446]. Gaussian measure in Hilbert spaces (Section 1.1.6) and Fourier transform are nicely introduced in [130, 110], a more extended study of the subject is contained in [300].

Generalities about stochastic processes, martingales and stopping times in Section 1.2 can be found in many different books, e.g. [267, 283, 294, 344, 345, 346, 395, 398, 446], while for Hilbert-valued martingales (Section 1.2.2) the reader may consult [130, 220, 380]. For standard Wiener and Q-Wiener processes and related results we refer to [93, 130, 220, 283, 345, 344, 349]. The material of Section 1.2.4 is based on [130]. Definition 1.87 is not contained in the standard literature, it is introduced here because it is useful to study stochastic control problems. The presentation of Lemma 1.89 is based on [402] and [283]. The material of Section 1.2.5 is loosely based on [130, 220, 283].

The material of Section 1.3 is based on [130, 127, 220] (see also [93, 385]). These books present the theory in Hilbert spaces while [345, 344] (see also [139]) present the Banach space case.

The presentation of Section 1.4 about solutions of stochastic differential equations in Hilbert spaces is also based on [130, 220]. In particular [130] is a standard reference in the theory. Other references on strong and mild solutions are for example in [93, 127] while good introduction to variational solutions is in [93, 297, 385, 406]. The reader is also referred to [130] for more on weak mild solutions. Section 1.4.4 containing some results about uniqueness in law uses the approach of [368]. For a different approach to weak uniqueness based on the theorem of Yamada-Watanabe we refer the reader to [385], Appendix E.

Section 1.5 contains existence and uniqueness results for stochastic differential equations with special unbounded terms and cylindrical additive noise. They are more or less common knowledge however we presented proofs since no complete references seem to be available in the literature.

Classical results on transition semigroups (Section 1.6) can be found in [130]. The statements here they are a little modified and extended in order to be used in our applications to optimal control, mainly in Chapter 4.

Section 1.7 contains various versions of Itô's and Dynkin's formulae (Propositions 1.155, 1.156, 1.157) in connection with mild solutions for functions that have properties of test functions used in the definition of viscosity solution (Definition 3.32). Such results have been known and used in the viscosity solution literature, however complete proofs are available only in [286]. The statements here are slightly more general than these in [286] and we presented proofs for the reader's convenience. The last two results of Section 1.7 (Propositions 1.159 and 1.161) are used to prove the verification theorems of Sections 4.8 and 5.5. They have been used in the literature (e.g in [231]) but without complete proofs, hence we provide them for completeness. We finally recall that Itô's formula related to variational solutions of linear stochastic parabolic equations is proved in [364].

CHAPTER 2

Optimal control problems and examples

In this chapter we discuss the connection between the study of infinite dimensional stochastic optimal control problems and that of second order Hamilton-Jacobi-Bellman (HJB) equations in Hilbert spaces. This so called "dynamic programming approach" to optimal control problems is based on two main results:

- The *dynamic programming principle*, which is a functional equation for the value function of the control problem, and whose differential form is the HJB equation. It is the core result in the dynamic programming approach.
- The *verification theorem*, which gives a sufficient (and sometimes necessary) condition for optimality. Verification theorems rely on the HJB equation and open the way to the so-called *optimal synthesis* i.e. the expression of the optimal control strategy as function of the current state trajectory (the *feedback form*).

To carry out this dynamic programming approach one needs suitable existence, uniqueness and regularity results for the solutions of the HJB equation. With this in mind we organize the chapter as follows.

In Section 2.1 we describe a general stochastic infinite dimensional optimal control problem (in both strong and weak formulations).

Sections 2.2 and 2.3 contain the dynamic programming principle (with a complete proof) and the equivalence between weak and strong formulations when the underlying "information structure" of the problem is based on a less general notion of a reference probability space, see Definition 2.7. The problem and the statement of the dynamic programming principle are formulated in an abstract form so that they can be used in many cases when the solutions of the state equation (which is an infinite dimensional SDE) are interpreted in various ways (strong, mild, variational, etc.). Since this increases the level of technicalities, we recommend that the readers assume on first reading that the state equation in the control problem is the one described in Section 2.2.3 with solutions defined in the mild sense, as this is the most common case in this book and the theory then applies more straightforwardly. Section 2.4 is devoted to the infinite horizon problem.

In Section 2.5 we present classical verification theorems and the optimal synthesis when the value function is regular, in both the finite and the infinite horizon cases.

Finally, in Section 2.6 we discuss various examples of stochastic infinite dimensional optimal control problems, which arise in applications, and which can be studied in the framework of the theory presented in this book.

The material on the dynamic programming principle and the examples are presented for optimal control problems defined on the whole space. We do not discuss in the book problems in bounded domains with more general cost functionals including cost of exiting through the boundary, problems with optimal stopping, state constraints problems, singular control problems, risk sensitive control problems, ergodic control problems, stochastic differential games. Some references to results for such problems are scattered throughout other sections.

2.1. Stochastic optimal control problem: general formulation

2.1.1. Strong formulation. We start with a description of a general stochastic optimal control problem in an infinite dimensional Hilbert space. We will be using the convention of Notation 1.65.

We make the following assumptions:

HYPOTHESIS 2.1
- (i) *The* state space H *and the* noise space Ξ *are real separable Hilbert spaces.*
- (ii) *The* control space Λ *is a Polish space.*
- (iii) *The horizon of the problem is* $T \in (0, +\infty) \cup \{+\infty\}$*, and the initial time is* $t \in [0, T)$.
- (iv) $\mu := \left(\Omega^\mu, \mathscr{F}^\mu, \left\{\mathscr{F}^t_{\mu,s}\right\}_{s\in[t,T]}, \mathbb{P}^\mu, W^\mu_Q\right)$ *is a generalized reference probability space from Definition 1.95 with* $W_Q(t) = 0$*,* $\mathbb{P}$ *a.s..*

We introduce the set of *admissible controls*

$$\mathcal{U}^\mu_t := \left\{a(\cdot)\colon [t,T]\times\Omega \to \Lambda \;:\; a(\cdot) \text{ is } \mathscr{F}^t_{\mu,s} - \text{progressively measurable}\right\}. \tag{2.1}$$

The notation $\mathcal{U}^\mu_t$ emphasizes the dependence on the generalized reference probability space. Sometimes additional conditions (e.g. state constraints) are imposed on the admissible controls.

In a general infinite dimensional stochastic optimal control problem, we consider, for every $a^\mu(\cdot) \in \mathcal{U}^\mu_t$, a system driven by an abstract stochastic differential equation in H

$$\begin{cases} dX(s) = \beta(s, X(s), a^\mu(s))ds + \sigma(s, X(s), a^\mu(s))dW^\mu_Q(s), \quad s \in [t,T] \\ X(t) = x \in H, \end{cases} \tag{2.2}$$

where β, σ are appropriate functions for which the above equation is well posed (in a sense to be made precise, see Remark 2.2) for every admissible control. Such equation is called the *state equation* and we denote by $X(\cdot; t, x, a^\mu(\cdot)) : [t,T] \to H$ (or simply by $X(\cdot)$ when its meaning is clear) its unique solution. This is the *state trajectory* of the system. The couple $(a^\mu(\cdot), X(\cdot; t, x, a^\mu(\cdot)))$ will be called an *admissible couple* (or *admissible pair*).

The goal is to *minimize*, over all $a^\mu(\cdot) \in \mathcal{U}^\mu_t$, the *cost functional*

$$\begin{aligned} J^\mu(t,x;a^\mu(\cdot)) = \mathbb{E}^\mu\Bigg[\int_t^T & e^{-\int_t^s c(X(\tau;t,x,a^\mu(\cdot)))d\tau} l(s, X(s;t,x,a^\mu(\cdot)), a^\mu(s))ds \\ & + e^{-\int_t^T c(X(\tau;t,x,a^\mu(\cdot)))d\tau} g(X(T;t,x,a^\mu(\cdot)))\Bigg], \end{aligned} \tag{2.3}$$

where $l\colon [t,T]\times H \times \Lambda \to \mathbb{R}$, $c, g\colon H \to \mathbb{R}$ are Borel measurable functions, and c is bounded from below. The function l is the so called running cost, g is the terminal cost, and c is a function responsible for discounting. When $T = +\infty$ the standing convention will be that $g = 0$, i.e. the cost functional only depends on the running cost and discounting. When T is finite the problem is called a *finite horizon problem*, and when $T = +\infty$ it is called an *infinite horizon problem.* The expectation $\mathbb{E}^\mu$ is computed with respect to the probability measure $\mathbb{P}^\mu$ so it depends on the generalized reference probability space. When the generalized reference probability space is clear we will often drop the superscript μ in our notation. We will refer to the above problem as the *strong formulation* of the stochastic optimal control problem (2.2)-(2.3) on $[t,T]$. *Strong here means that the generalized reference probability space is fixed.*

The discounting function c may also depend on the control variable. The results of this chapter can be easily extended to cover such case. However we chose not to include this dependence in order not to overcomplicate the presentation which is already very technical.

REMARK 2.2 In the infinite dimensional case the state equation (2.2) can have different forms which may call for various definitions of solutions (strong solutions, mild solutions, variational solutions, etc.) and various approaches to solve them. For this reason in our general formulation we do not specify the concept of solution of (2.2) and we do not specify the required assumptions. Later, in Section 2.2, we will formulate and prove the dynamic programming principle (DPP) in a general form so that it can be applied in these different situations. Thus we will make a series of rather abstract assumptions (see Hypotheses 2.11-2.12) about (2.2) that are verified in various cases for different concepts of solutions and which are sufficient to prove the DPP.

However the reader should keep in mind that our primary guiding examples are the control problems of the type (2.2)-(2.3) where the state equation is a stochastic evolution equation with solutions interpreted in the mild sense. In such cases we have $\beta = A + b$, where A is the generator of a C_0 semigroup on H, and b, σ are functions satisfying suitable Lipschitz conditions. This case requires a less general formulation to prove the DPP and will be discussed separately in Subsection 2.2.3. The cases which do not use mild solutions include optimal control problems for the Duncan-Mortensen-Zakai, Burgers, Navier-Stokes, and reaction diffusion equations. ■

The value function for problem (2.2)-(2.3) in the strong formulation with initial time t is defined as

$$V_t^\mu(x) = \inf_{a(\cdot)\in\mathcal{U}_t^\mu} J^\mu(t,x;a^\mu(\cdot)). \tag{2.4}$$

Notice however, that in this strong formulation the generalized reference probability space changes when we change t and so does the control set $\mathcal{U}_t^\mu$.

DEFINITION 2.3 (Optimal control/couple) *If, for given initial data (t,x), $a^*(\cdot) \in \mathcal{U}_t^\mu$ minimizes (2.3), i.e. if $J^\mu(t,x;a^*(\cdot)) = V_t^\mu(x)$, we say that $a^*(\cdot)$ is an* optimal control *at (t,x). The associated state trajectory $X^*(\cdot) := X(\cdot;t,x,a^*(\cdot))$ (i.e. the solution of (2.2) with control $a^*(\cdot)$) is an* optimal state *at (t,x). The pair $(a^*(\cdot), X^*(\cdot))$ is called an* optimal couple *(or* optimal pair*) at (t,x).*

To perform the dynamic programming approach in the strong formulation we need to consider a family of problems (2.2)-(2.3) parameterized by the initial time t which are defined on a common generalized reference probability space, and introduce a value function defined on $[0,T]\times H$. To do this we take a generalized reference probability space $\mu = \left(\Omega^\mu, \mathscr{F}^\mu, \left\{\mathscr{F}^0_{\mu,s}\right\}_{s\in[0,T]}, \mathbb{P}^\mu, W_Q^\mu\right)$ on $[0,T]$ with $W_Q(0) = 0$ (i.e. μ satisfies Hypothesis 2.1 with initial time $t=0$). We then define the value function

$$V^\mu(t,x) = \inf_{a(\cdot)\in\mathcal{U}_0^\mu} J^\mu(t,x;a^\mu(\cdot)), \tag{2.5}$$

where $J^\mu(t,x;a^\mu(\cdot))$ is defined by (2.3) with $X(\cdot;t,x,a^\mu(\cdot)))$ solving (2.2). We notice that for μ as above, the generalized reference probability spaces $\mu_t := \left(\Omega^\mu, \mathscr{F}^\mu, \left\{\mathscr{F}^0_{\mu,s}\right\}_{s\in[t,T]}, \mathbb{P}^\mu, W_Q^\mu(\cdot) - W_Q^\mu(t)\right)$ satisfy Hypothesis 2.1 with initial time t. Thus it is reasonable to expect that $V^\mu(t,x)$ should be equal to $V_t^{\mu_t}(x)$ for $(t,x)\in[0,T]\times H$. This is indeed the case for control problems considered in this book and it is a simple consequence of the properties of the stochastic integral (see e.g. (2.16)). Thus the requirement that $W_Q(t) = 0$ in Hypothesis 2.1-(iv) can be dropped for all practical purposes.

2.1.2. Weak formulation. In the strong formulation of the optimal control problem (2.2)-(2.3), the generalized reference probability space $\mu := (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W_Q)$ is fixed. However it is often more convenient or necessary to include the generalized reference probability space as part of the control. This approach is used to prove the dynamic programming principle and to construct optimal feedback controls (see Sections 2.2 and 2.5). This leads us to the weak formulation of the stochastic optimal control problem.

In the weak formulation the state equation and the cost functional are the same, however, for each fixed $t \in [0, T]$, any generalized reference probability space μ is allowed and so the class of admissible controls is enlarged. We define

$$\overline{\mathcal{U}}_t := \bigcup_{\mu} \mathcal{U}_t^{\mu}, \tag{2.6}$$

where the union is taken over all generalized reference probability spaces μ satisfying Hypothesis 2.1-(iv). We say that $a(\cdot)$ is an admissible control if $a(\cdot) \in \overline{\mathcal{U}}_t$, i.e. if there exist a generalized reference probability space $\mu = \left(\Omega^{\mu}, \mathscr{F}^{\mu}, \mathscr{F}_s^{\mu,t}, \mathbb{P}^{\mu}, W_Q^{\mu}\right)$ satisfying Hypothesis 2.1-(iv) such that $a(\cdot)\colon [t, T] \times \Omega^{\mu} \to \Lambda$ is $\mathscr{F}_s^{\mu,t}$-progressively measurable. We will often write $a^{\mu}(\cdot)$ to indicate the dependence of $a(\cdot)$ on the generalized reference probability space. This way, choosing an admissible control also means choosing a generalized reference space μ, so with a slight abuse of notation, we will often write, $a(\cdot) = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W_Q, a(\cdot))$.

Given a control $a^{\mu}(\cdot) \in \overline{\mathcal{U}}_t$ and the related trajectory $X(\cdot; t, x, a^{\mu}(\cdot))$[1], we call the couple $(a^{\mu}(\cdot), X(s; t, x, a^{\mu}(\cdot)))$ an *admissible couple* (or *admissible pair*) in the weak sense.

REMARK 2.4 To avoid misunderstandings we clarify that the use of the term “weak" in the “weak formulation" of our stochastic control problem, is referred only to the fact that the generalized reference probability spaces vary with the controls and not to the concept of solution of the state equation. Indeed, in our framework, once the control $a^{\mu}(\cdot)$ is fixed (and with it also the generalized reference space), the solution is taken in the same generalized reference space (i.e. in the so-called strong probabilistic sense). Solutions of the state equation in the weak probabilistic sense will be used only to treat the closed loop equations in some cases, see Remark 2.6 ■

In the weak formulation the goal is to minimize the same cost functional (2.3), however now over all controls $a^{\mu}(\cdot) \in \overline{\mathcal{U}}_t$. Consequently, the value function in the weak formulation is now defined by

$$\overline{V}(t, x) = \inf_{a^{\mu}(\cdot) \in \overline{\mathcal{U}}_t} J^{\mu}(t, x; a^{\mu}(\cdot)), \quad (t, x) \in [0, T) \times H, \tag{2.7}$$

and we set $\overline{V}(T, x) := g(x)$ for $x \in H$ if $T < +\infty$. From the above definition we clearly have

$$\overline{V}(t, x) = \inf_{\mu} \inf_{a(\cdot) \in \mathcal{U}_t^{\mu}} J^{\mu}(t, x; a(\cdot)) = \inf_{\mu} V_t^{\mu}(x). \tag{2.8}$$

[1]To know that such trajectory exists and is unique we need to assume that the state equation (2.2) is well posed for every admissible control $a^{\mu}(\cdot)$, so in particular for every generalized reference space.

REMARK 2.5 For the optimal control problem we could as well have required that controls in $\mathcal{U}_t^\mu$ be measurable and adapted instead of progressively measurable, since, by Lemma 1.69, every adapted $a(\cdot)$ has a progressively measurable modification $\tilde{a}(\cdot)$. We chose to deal with progressively measurable controls to avoid unnecessary technical issues. In light of Lemma 1.94, we could have chosen to work with predictable controls as well.

Moreover, in the definition of $\mathcal{U}_t^\mu$ we did not specify possible further restrictions on the control and on the state (state constraints, integrability conditions on the controls, etc.) which commonly arise in examples, see Section 2.6. Such kinds of restrictions usually lead to more complicated problems, however in principle they can be treated in the same framework. ■

REMARK 2.6 To study problems where neither existence, nor uniqueness of solutions of the state equation is guaranteed for arbitrary control process $a(\cdot)$ (in particular to study the existence of optimal feedback controls) it is useful to extend the formulation of an optimal control problem. In such cases the *extended* formulation of the control problem can be given as follows. Given a generalized reference probability space μ, we call $(a(\cdot), X(\cdot))$ an *admissible control pair* if $a(\cdot)$ is an $\mathscr{F}_s^t$-progressively measurable process with values in Λ and $X(\cdot)$ is a (not necessarily unique) solution of (2.2) corresponding to $a(\cdot)$. To every admissible control pair we associate the cost (2.3). The optimal control problem in the *extended* strong formulation consists in minimizing the functional $J^\mu(t, x; a(\cdot), X(\cdot))$ over all admissible control pairs $(a(\cdot), X(\cdot))$, and in characterizing the value function (where we use the same notation for simplicity)

$$V_t^\mu(x) = \inf_{(a(\cdot),X(\cdot))} J^\mu(t, x; a(\cdot), X(\cdot)).$$

The optimal control problem in the *extended* weak formulation consists in further minimizing with respect to all generalized reference probability spaces, i.e. in characterizing the value function (where again we use the same notation for simplicity)

$$\overline{V}(t, x) = \inf_{\mu} V_t^\mu(x).$$

Such formulation is often much more suitable for construction of optimal feedback controls, see Corollary 2.37, and is employed for instance in Chapter 6 (and partly also in Chapters 4 and 5). ■

2.2. Dynamic Programming Principle: setup and assumptions

In this section we introduce the Dynamic Programming Principle (DPP). It is one of the fundamental results of stochastic optimal control, whose formulation and proof are very technical, here even more so since we deal with the infinite dimensional case. We first present the stochastic setup and the main assumptions, and then follow with the statement of the DPP and the proof.

2.2.1. The setup.

DEFINITION 2.7 *A* reference probability space *is a generalized reference probability space $\nu := (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W_Q)$ (see Definition 1.95), where $W_Q(t) = 0$, $\mathbb{P}$ a.s., and $\mathscr{F}_s^t = \sigma(\mathscr{F}_s^{t,0}, \mathcal{N})$, where $\mathscr{F}_s^{t,0} = \sigma(W_Q(\tau) : t \le \tau \le s)$ is the filtration generated by W_Q, and $\mathcal{N}$ is the collection of the $\mathbb{P}$-null sets in $\mathscr{F}$.*

DEFINITION 2.8 *We will say that a reference probability space ν is* standard *if there exists a σ-field $\mathscr{F}'$ such that $\mathscr{F}_T^{t,0} \subset \mathscr{F}' \subset \mathscr{F}$, $\mathscr{F}$ is the completion of $\mathscr{F}'$, and $(\Omega, \mathscr{F}')$ is a standard measurable space (see Section 1.1).*

We will consider control problem (2.2)-(2.3) in the weak formulation in which *generalized reference probability spaces* are replaced by *reference probability spaces* This means that we are restricting the set of admissible controls. The set of all admissible controls is now defined by

$$\mathcal{U}_t := \bigcup_{\nu} \mathcal{U}_t^{\nu}, \tag{2.9}$$

where the union is taken over all reference probability spaces ν. Obviously $\mathcal{U}_t \subseteq \overline{\mathcal{U}}_t$, where $\overline{\mathcal{U}}_t$ is defined by (2.6). Thus $a(\cdot)$ is an admissible control now if there exist a reference probability space $\nu = \left(\Omega^\nu, \mathscr{F}^\nu, \mathscr{F}_s^{\nu,t}, \mathbb{P}^\nu, W_Q^\nu\right)$ such that $a(\cdot)\colon [t,T]\times\Omega^\nu \to \Lambda$ is $\mathscr{F}_s^{\nu,t}$-progressively measurable. As before we will often write $a^\nu(\cdot)$ to indicate the dependence of $a(\cdot)$ on the reference probability space.

With this definition the value function is now defined by

$$V(t,x) = \inf_{a^\nu(\cdot)\in\mathcal{U}_t} J^\nu(t,x;a^\nu(\cdot)). \tag{2.10}$$

(with the same convention that $V(T,x) := g(x)$ if $T < +\infty$) and, clearly, if $\overline{V}$ is the value function defined in (2.7),

$$\overline{V}(t,x) \le V(t,x) \le V_t^\nu(x), \qquad \text{for every reference probability space } \nu. \tag{2.11}$$

We will later see (Theorem 2.22) that the last inequality is indeed an equality under our assumptions. When solutions of the HJB equations are regular enough to allow construction of optimal feedbacks we will also see in Chapters 4, 5 and 6 that both inequalities become equalities (see e.g. Theorem 4.140). We do not study this issue here but the reader may check [364] for an argument that for control problems considered in [364] the first inequality is an equality. It is possible that the approach from [364] can be applied to the control problems in this book.

2.2.2. The general assumptions. We make the following general assumption.

Hypothesis 2.9

(i) *The* state space H *and the* noise space Ξ *are real separable Hilbert spaces.*

(ii) *The* control space Λ *is a Polish space.*

(iii) *The horizon of the problem is* $T \in (0,+\infty)\cup\{+\infty\}$ *and the initial time is* $t \in [0,T)$.

(iv) $Q \in \mathcal{L}_1^+(\Xi)$

Remark 2.10 We assume here that $Q \in \mathcal{L}_1^+(\Xi)$ (i.e. $\mathrm{Tr}\,(Q) < +\infty$), which implies that the processes W_Q in the reference probability spaces are Ξ-valued Q-Wiener processes. We do this for technical reasons, because in our proof of the DPP it is important that the Q-Wiener processes always have values in some (fixed) space Ξ. However in many examples discussed in this chapter we will encounter Q-Wiener processes for which $\mathrm{Tr}\,(Q) = +\infty$. Recalling the definition of a Q-Wiener process (see Definition 1.83) we then have to choose and fix a space Ξ_1 such that W_Q is a Ξ_1-valued Q_1-Wiener process. (Since the space Ξ_1 is often not important, abusing notation, we will still call such process a W_Q-Wiener process, see Remark 1.84.) It puts us in the framework developed in this chapter and this is how the reader should understand such control problems, as it will not be repeated in the future when we discuss the examples in this chapter unless it is essential. However, as mentioned in Remark 1.86, if $W_Q(s) = \sum_{k=1}^\infty g_k\beta_k(s)$ for some orthonormal basis $\{g_k\}_{k\in\mathbb{N}}$ of Ξ_0 (see Definition 1.83), then regardless of the choice of Ξ_1, $\mathscr{F}_s^{t,0} = \sigma\{\beta_k(r) : t \le r \le s, k \in \mathbb{N}\}$. Thus the filtration does not depend on the choice of Ξ_1, and then also the class of integrable processes is independent of Ξ_1. Therefore the control problems discussed in the examples are independent of the choice of Ξ_1 and the theory can be applied to optimal control problems for which $Q \in \mathcal{L}^+(\Xi)$. ■

The following comment is important. The Q-Wiener processes in the reference probability spaces in general have trajectories which are only $\mathbb{P}$ a.e. continuous. However we can always modify them on a set of measure zero so that the trajectories are continuous everywhere. Moreover it is obvious that such modified Q-Wiener process generates the same filtration $\mathscr{F}_s^t$ as the original one so the set of admissible controls does not change. Moreover the solutions of the stochastic differential equations for control problems considered in this book are indistinguishable after this modification of the Q-Wiener processes so the cost functional will be the same (see assumption $(A1)$ of Hypothesis 2.12). Therefore, unless specified otherwise, without loss of generality, **we will always assume that the Q-Wiener processes in the reference probability spaces have everywhere continuous paths** however we will point it out explicitly if it is important to avoid any misunderstandings.

The assumptions about existence and uniqueness of solutions of the state equation are the following.

HYPOTHESIS 2.11 *Assume that, for every $0 \le t < T$, reference probability space $\nu := (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W_Q)$, $a(\cdot) \in \mathcal{U}_t^\nu$ and an H-valued $\mathscr{F}_t^t$-measurable random variable ζ (i.e, $\zeta = x$, $\mathbb{P}$ a.s. for some $x \in H$), we have a unique, up to a modification, solution (in a certain sense) $X(\cdot) = X(\cdot; t, \zeta, a(\cdot))$ on $[t, T]$ of the abstract stochastic differential equation*

$$\begin{cases} dX(s) = \beta(s, X(s), a(s))ds + \sigma(s, X(s), a(s))dW_Q(s), \\ X(t_1) = \zeta. \end{cases} \tag{2.12}$$

The solution $X(\cdot; t, \zeta, a(\cdot))$ is $\mathscr{F}_s^t$-progressively measurable, has continuous trajectories in H and $X(t; t, \zeta, a(\cdot)) = \zeta$, $\mathbb{P}$ a.e..

The above hypothesis in particular implies that any modification of a solution is still a solution of the same equation as long as it has continuous trajectories. To emphasize the dependence of the solution on the reference probability space we will sometimes use the notation $X^\nu(\cdot; t, \zeta, a(\cdot))$.

Hypothesis 2.12 collects assumptions about the properties of the family of solutions of the state equation. To simplify notation we will write W instead of W_Q in Hypothesis 2.12 and in other places when notation becomes cumbersome and when the meaning of it is clear.

HYPOTHESIS 2.12 *Assume that Hypothesis 2.11 holds. For every $0 \le t \le \eta < T$, $x \in H$, reference probability space $\nu = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W)$, $a(\cdot) \in \mathcal{U}_t^\mu$, and an H-valued $\mathscr{F}_t^t$-measurable random variable ζ such that $\zeta = x$, $\mathbb{P}$ a.s., we have the following properties:*

(A0)

$$X(\cdot; t, \zeta, a(\cdot)) = X(\cdot; t, x, a(\cdot)) \text{ on } [t, T] \ \mathbb{P} \ a.e..$$

(A1) *If $\nu_1 = \big(\Omega_1, \mathscr{F}_1, \mathscr{F}_{1,s}^t, \mathbb{P}_1, W_1\big)$, $\nu_2 = \big(\Omega_2, \mathscr{F}_2, \mathscr{F}_{2,s}^t, \mathbb{P}_2, W_2\big)$ are two reference probability spaces, $a_1(\cdot) \in \mathcal{U}_t^{\nu_1}, a_2(\cdot) \in \mathcal{U}_t^{\nu_2}$, and $\mathcal{L}_{\mathbb{P}_1}(a_1(\cdot), W_1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(a_2(\cdot), W_2(\cdot))$ (see Definition 1.127), then*

$$\mathcal{L}_{\mathbb{P}_1}(X(\cdot; t, x, a_1(\cdot)), a_1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(X(\cdot; t, x, a_2(\cdot)), a_2(\cdot)).$$

(A2) *If $a_1(\cdot)$, $a_2(\cdot) \in \mathcal{U}_t^\nu$ are such that $a_1(\cdot) = a_2(\cdot)$, $dt \otimes \mathbb{P}$ a.e. on $[t, \eta] \times \Omega$, then*

$$X(\cdot; t, x, a_1(\cdot)) = X(\cdot; t, x, a_2(\cdot)) \text{ on } [t, \eta], \ \mathbb{P} \text{ a.e..}$$

(A3) *Let $\nu = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W)$ be a standard reference probability space (Definition 2.8) with W having everywhere continuous trajectories. Let $\nu_{\omega_0} = \left(\Omega, \mathscr{F}_{\omega_0}, \mathscr{F}_{\omega_0,s}^\eta, \mathbb{P}_{\omega_0}, W_\eta\right)$, where $\mathbb{P}_{\omega_0} = \mathbb{P}(\cdot|\mathscr{F}_\eta^{t,0})(\omega_0)$ is the regular conditional probability, $\mathscr{F}_{\omega_0}$ is the augmentation of $\mathscr{F}'$ by the $\mathbb{P}_{\omega_0}$ null sets, and $\mathscr{F}_{\omega_0,s}^\eta$ is the augmented filtration generated by W_η[2]. Let $a(\cdot) \in \mathcal{U}_t^\nu$ and $a_{|[\eta,T]}(\cdot) \in \mathcal{U}_\eta^{\nu_{\omega_0}}$ for $\mathbb{P}$ a.e. ω_0. Then the process $X^\nu(s; t, x, a(\cdot))$ has an indistinguishable version such that, for $\mathbb{P}$ a.e. ω_0, $X^{\nu_{\omega_0}}(\cdot; \eta, X^\nu(\eta), a(\cdot)) = X^\nu(\cdot; t, x, a(\cdot))$ on $[\eta, T]$ $\mathbb{P}_{\omega_0}$ a.s..*

REMARK 2.13 It is possible to relax and slightly simplify Hypothesis 2.12 by combining conditions $(A0) - (A1)$ into one condition

(A1') If $\nu_1 = \left(\Omega_1, \mathscr{F}_1, \mathscr{F}_{1,s}^t, \mathbb{P}_1, W_1\right)$, $\nu_2 = \left(\Omega_2, \mathscr{F}_2, \mathscr{F}_{2,s}^t, \mathbb{P}_2, W_2\right)$ are two reference probability spaces, $x \in H$, ζ is an H-valued $\mathscr{F}_{1,t}^t$-measurable random variable such that $\zeta = x$, $\mathbb{P}_1$ a.s., $a_1(\cdot) \in \mathcal{U}_t^{\nu_1}, a_2(\cdot) \in \mathcal{U}_t^{\nu_2}$, and $\mathcal{L}_{\mathbb{P}_1}(a_1(\cdot), W_1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(a_2(\cdot), W_2(\cdot))$, then

$$\mathcal{L}_{\mathbb{P}_1}(X(\cdot; t, \zeta, a_1(\cdot)), a_1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(X(\cdot; t, x, a_2(\cdot)), a_2(\cdot)).$$

With this change the proof of the dynamic programming principle is virtually unchanged. However since condition $(A0)$ is standard and is satisfied by control problems considered in this book, we opted to keep it in the formulation of Hypothesis 2.12 hoping that this way the proof of the dynamic programming principle will be slightly easier to follow. ■

REMARK 2.14 We point out that since the trajectories of the solutions are continuous, $(A1)$ implies in particular that

$$\mathcal{L}_{\mathbb{P}_1}(X(\cdot; t, x, a_1(\cdot))) = \mathcal{L}_{\mathbb{P}_2}(X(\cdot; t, x, a_2(\cdot))) \quad \text{on } [t, T].$$

■

Let us briefly explain the nature of the abstract assumptions of Hypothesis 2.12. Condition $(A0)$ guarantees pathwise uniqueness for our solutions with almost deterministic initial conditions, while condition $(A1)$ is a statement about uniqueness in law which guarantees that the joint law of $(X(\cdot; t, x, a(\cdot)), a(\cdot))$ only depends on the joint law of $(a(\cdot), W(\cdot))$. Condition $(A2)$ is a requirement that if the controls are "almost the same" then the solutions do not change. Finally, the most complicated condition $(A3)$ is a technical assumption which is needed since we do not define precisely what we mean by a solution. It guarantees, roughly speaking, that if we have a solution X in one reference probability space, then, for $\mathbb{P}$ a.e. ω_0, X is still a solution in reference probability spaces equipped with measures $\mathbb{P}_{\omega_0}$ provided certain conditions are satisfied. We remark that for $\mathbb{P}$ a.e. ω_0, $X^\nu(\eta)$ is $\mathbb{P}_{\omega_0}$ a.e. constant and is equal to $X^\nu(\eta)(\omega_0)$. We remark that condition $(A3)$ in particular implies that the version of $X^\nu(s; t, x, a(\cdot))$ is $\mathscr{F}_{\omega_0,s}^\eta$-progressively measurable on $[\eta, T]$ for $\mathbb{P}$ a.e. ω_0, and has continuous trajectories $\mathbb{P}_{\omega_0}$ a.e. for $\mathbb{P}$ a.e. ω_0. These two properties can be proved for every solution X^ν satisfying our Hypothesis 2.11. We required that ν is a standard reference probability space to guarantee the existence of the regular conditional probability $\mathbb{P}_{\omega_0}$. The requirement that $a_{|[\eta,T]}(\cdot) \in \mathcal{U}_\eta^{\nu_{\omega_0}}$ can be always assumed since we will see in Lemma 2.26 that

[2]We remark that ν_{ω_0} in $(A3)$ is a reference probability space for $\mathbb{P}$ a.e. ω_0 by Lemma 2.25.

for every $a(\cdot) \in \mathcal{U}_t^\nu$ there is $a_1(\cdot) \in \mathcal{U}_t^\nu$ such that $a(\cdot) = a_1(\cdot)$, $\mathbb{P} \otimes dt$ a.e. and $a_{1|[\eta,T]}(\cdot) \in \mathcal{U}_\eta^{\nu_{\omega_0}}$ for $\mathbb{P}$ a.e. ω_0.

REMARK 2.15 This is a very important remark regarding optimal control problems with additional conditions on the set of admissible controls. In the proof of the dynamic programming principle we will use the following property of admissible controls.

(A4) If ν is a standard reference probability space as in $(A3)$ and $a(\cdot) \in \mathcal{U}_t^\nu$, then there exists $a_1(\cdot) \in \mathcal{U}_t^\nu$ such that $a(\cdot) = a_1(\cdot)$ $\mathbb{P} \otimes dt$ a.e. and $a_{|[\eta,T]}(\cdot) \in \mathcal{U}_\eta^{\nu_{\omega_0}}$ for $\mathbb{P}$ a.e. ω_0, where ν_{ω_0} is as in $(A3)$.

This property is always true for our abstract optimal control problem and it is shown in Lemma 2.26 whose proof is only based on considerations of measurability. However, if the set of admissible controls is characterized by additional conditions, for instance some integrability conditions, property $(A4)$ must be established in each particular case, see for instance Remark 2.27. Therefore the reader should be very careful adapting the abstract proof of the dynamic programming principle to such cases. ■

We close with important remarks about standard reference probability spaces and regular conditional probabilities. Let $\nu = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W_Q)$ be a standard reference probability space. Regular conditional probabilities will be denoted by $\mathbb{P}_\omega$, and to indicate that $\mathbb{P}_{\omega_0}$ is the regular conditional probability given a sigma field $\mathscr{F}_s^{t,0}$ we will write $\mathbb{P}_{\omega_0} = \mathbb{P}(\cdot|\mathscr{F}_s^{t,0})(\omega_0)$ even though this is a slight abuse of notation. The expectation with respect to $\mathbb{P}_{\omega_0}$ will be denoted by $\mathbb{E}_{\omega_0}$.

For every $\Omega_1 \in \mathscr{F}$ such that $\mathbb{P}(\Omega_1) = 1$ there exists $\Omega_2 \subset \Omega_1, \Omega_2 \in \mathscr{F}'$ (from Definition 2.8) such that $\mathbb{P}(\Omega_2) = 1$. Therefore

$$1 = \mathbb{P}(\Omega_2) = \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}_{\Omega_2}|\mathscr{F}_s^{t,0}\right](\omega_0)\right] = \mathbb{E}\left[\mathbb{P}_{\omega_0}(\Omega_2)\right].$$

Thus we obtain that $\Omega_1 \in \mathscr{F}_{\omega_0}$ for $\mathbb{P}$ a.e. ω_0 and

$$1 = \mathbb{P}_{\omega_0}(\Omega_2) = \mathbb{P}_{\omega_0}(\Omega_1).$$

Now suppose that $Y \in L^1(\Omega, \mathscr{F}, \mathbb{P})$. Let $Y' \in L^1(\Omega, \mathscr{F}', \mathbb{P})$ be such that $Y = Y'$, $\mathbb{P}$ a.s. Hence also $Y = Y'$ $\mathbb{P}_{\omega_0}$ a.s. for $\mathbb{P}$ a.s. ω_0, which implies that Y is $\mathscr{F}_{\omega_0}$ measurable for $\mathbb{P}$ a.s. ω_0. Thus for $\mathbb{P}$ a.s. ω_0

$$\mathbb{E}\left[Y|\mathscr{F}_s^t\right](\omega_0) = \mathbb{E}\left[Y'|\mathscr{F}_s^{t,0}\right](\omega_0) = \int Y'(\omega)d\mathbb{P}_{\omega_0}(\omega) = \int Y(\omega)d\mathbb{P}_{\omega_0}(\omega) = \mathbb{E}_{\omega_0}[Y].$$

Therefore $\mathbb{E}_{\omega_0}[Y]$ (as a function of ω_0) is in $L^1(\Omega, \mathscr{F}, \mathbb{P})$ and

$$\mathbb{E}[Y] = \mathbb{E}\left[\mathbb{E}\left[Y|\mathscr{F}_s^t\right]\right] = \mathbb{E}\left[\mathbb{E}_{\omega_0}[Y]\right].$$

This fact will be used frequently in the following chapters without repeating the technical details.

2.2.3. The assumptions in the case of control problem for mild solutions. In this subsection we briefly illustrate the abstract setup for the case which is the most frequent among problems treated in the book, namely optimal control problem driven by general stochastic evolution equation (with Lipschitz coefficients) with solutions interpreted in the mild sense, and explain how Hypotheses 2.11 and 2.12 are satisfied. In this case the state equation (2.12) is of type (1.31) i.e.

$$\begin{cases} dX(s) = (AX(s) + b(s, X(s), a(s)))\,ds + \sigma(s, X(s), a(s))dW_Q(s) \\ X(t) = \zeta. \end{cases} \tag{2.13}$$

where A, b and σ satisfy Hypothesis 1.119 and its solution is understood in the mild sense of Definition 1.113, i.e. we have

$$X(s) = e^{(s-t)A}\zeta + \int_t^s e^{(s-r)A} b(r, X(r), a(r))dr + \int_t^s e^{(s-r)A}\sigma(r, X(r), a(r))dW_Q(r) \tag{2.14}$$

on $[t, T]$, $\mathbb{P}$ a.e..

PROPOSITION 2.16 *Consider equation (2.13) under Hypotheses 1.119 and 2.9. Then Hypotheses 2.11 and 2.12 are satisfied for its mild solutions.*

PROOF. Hypothesis 2.11 follows from Theorem 1.121. Regarding Hypothesis 2.12, condition $(A0)$ follows from the fact that $e^{sA}\zeta = e^{sA}x$ $\mathbb{P}$-a.e. for all $s \in [t, T]$ and from (1.40). Condition $(A1)$ follows from Proposition 1.131 and $(A2)$ follows from Corollary 1.122.

To show $(A3)$, we will first show that $X^\nu(s) := X^\nu(s; t, x, a(\cdot))$ has a modification X_1^ν, which is everywhere continuous on and for $\mathbb{P}$ a.e. ω_0 is $\mathscr{F}^\eta_{\omega_0,s}$-progressively measurable on $[\eta, T]$. In general X^ν is only $\mathscr{F}^t_s$-progressively measurable. Let Ω_0 be such that $\mathbb{P}(\Omega_0) = 1$, and for $\omega \in \Omega_0$, $X^\nu(\cdot, \omega)$ is continuous on $[t, T]$. Let $\{s_k\}, k \geq 1, s_1 = t$, be a countable dense set in $[t, T]$, and let $A_k \in \mathcal{F}^{t,0}_{s_k}, k \geq 1$ be sets such that $\mathbb{P}(A_k) = 1$, and $X^\nu(s_k) = \xi_k$ on A_k for some $\mathcal{F}^{t,0}_{s_k}$-measurable random variable ξ_k. Set $\Omega_1 = \Omega_0 \cap \bigcap_{k+1}^\infty A_k$. Then $\mathbb{P}(\Omega_1) = 1$ and thus for $\mathbb{P}$ a.e. ω_0, $\mathbb{P}_{\omega_0}(\Omega_1) = 1$, which implies that $\Omega_1 \in \mathscr{F}^\eta_{\omega_0,s}$ for $\mathbb{P}$ a.e. ω_0. We now define $X_1^\nu(s) = X^\nu(s)$ for $s \in [t, T], \omega \in \Omega_1$ and $X_1^\nu(s) = 0$ for $s \in [t, T], \omega \in \Omega \setminus \Omega_1$. The process X^ν has continuous trajectories. Since for $\omega \in \Omega_1$, $X_1^\nu(s) = \lim_{s_k \to s, s_k \leq s} \xi_k$, X_1^ν is $\sigma(\mathscr{F}^{t,0}_s, \Omega_1)$-adapted. However, thanks to Lemma 2.26, $\mathscr{F}^{t,0}_s \subset \mathscr{F}^\eta_{\omega_0,s}$ for $\mathbb{P}$ a.s. ω_0, and so it follows that X_1^ν is $\mathscr{F}^\eta_{\omega_0,s}$-adapted, which, noticing that it has continuous trajectories implies by Lemma 1.69 that it is $\mathscr{F}^\eta_{\omega_0,s}$-progressively measurable for $\mathbb{P}$ a.s. ω_0. From now on we will write X^ν for X_1^ν.

We notice that $X^\nu(\cdot) \in M^p_\nu(t, T; H), p > 2$, so in particular

$$\mathbb{E}\left[\mathbb{E}\left[\int_\eta^T |X^\nu(s)|^p ds | \mathscr{F}^{t,0}_\eta\right]\right] = \mathbb{E}\left[\int_\eta^T |X^\nu(s)|^p ds\right] < +\infty, \tag{2.15}$$

so for $\mathbb{P}$ a.e. ω_0, $X^\nu(\cdot) \in M^p_{\nu_{\omega_0}}(\eta, T; H)$. Thus by uniqueness of mild solutions given by Theorem 1.121 we will be done provided we know that, for $\mathbb{P}$ a.e. ω_0, $X^\nu(\cdot)$ is a mild solution, in the interval $[\eta, T]$ in the reference probability space ν_{ω_0}.

We have the flow property

$$\begin{aligned}
&X^\nu(s) = e^{(s-\eta)A}\left[\zeta + \int_t^\eta e^{(\eta-r)A} b(r, X^\nu(r), a(r))dr\right. \\
&\left.+ \int_t^\eta e^{(\eta-r)A}\sigma(r, X^\nu(r), a(r))dW(r)\right] + \int_\eta^s e^{(s-r)A} b(r, X^\nu(r), a(r))dr \\
&+ \int_\eta^s e^{(s-r)A}\sigma(r, X^\nu(r), a(r))dW(r) = e^{(s-\eta)A}X^\nu(\eta) \\
&+ \int_\eta^s e^{(s-r)A} b(r, X^\nu(r), a(r))dr + \int_\eta^s e^{(s-r)A}\sigma(r, X^\nu(r), a(r))dW(r) \\
&= X^\nu(s; \eta, X(\eta; t, \zeta, a(\cdot)), a(\cdot))
\end{aligned}$$

Since $\mathbb{P}$ a.s.

$$\int_\eta^s e^{(s-r)A}\sigma(r, X^\nu(r), a(r))dW(r) = \int_\eta^s e^{(s-r)A}\sigma(r, X^\nu(r), a(r))dW_\eta(r) \tag{2.16}$$

on $[\eta, s]$, and since for every set Ω_1 such that $\mathbb{P}(\Omega_1) = 1$ we have $\mathbb{P}_{\omega_0}(\Omega_1) = 1$ for $\mathbb{P}$ a.e. ω_0, the equality

$$X^\nu(s) = e^{(s-\eta)A} X^\nu(\eta) + \int_\eta^s e^{(s-r)A} b(r, X^\nu(r), a(r)) dr \\ + \int_\eta^s e^{(s-r)A} \sigma(r, X^\nu(r), a(r)) dW_\eta(r).$$

is satisfied $\mathbb{P}_{\omega_0}$ a.s. for $\mathbb{P}$ a.e. ω_0. This equality is exactly the same for the mild solution in the interval $[\eta, T]$ in the reference probability space ν_{ω_0} except for the fact that there the stochastic integral is taken in the reference probability space ν_{ω_0} instead of ν. So to conclude it is enough to show that for $\mathbb{P}$ a.e. ω_0

$$I_\nu(s) := \int_\eta^s e^{(s-r)A} \sigma(r, X^\nu(r), a(r)) dW_\eta(r)$$

is $\mathbb{P}_{\omega_0}$ a.e. equal on $[\eta, T]$ to the stochastic integral in the reference probability space ν_{ω_0}, which we denote by $I_{\nu_{\omega_0}}(s)$.[3] By continuity of the paths of the stochastic convolution (see Proposition 1.107) it is enough to show it for a single s. Denote $\Phi(r) = e^{(s-r)A} \sigma(r, X^\nu(r), a(r))$. By Lemma 1.93 and the proof of Lemma 1.94 there exist a sequence of elementary and $\mathscr{F}_s^{t,0}$-progressively measurable processes Φ_n with values in $\mathcal{L}(\Xi; H)$, such that $\|\Phi - \Phi_n\|_{\mathcal{N}^2_{Q,\nu}(\eta,T;H)} \to 0$, where we indicated the dependence on the reference probability space in the notation for the norm. By Lemma 2.26, Φ_n are also $\mathscr{F}^\eta_{\omega_0,s}$-progressively measurable. Since $\mathbb{E}^\nu[\int_\eta^s [\Phi_n(r) - \Phi(r)] dW_\eta(r)]^2 \to 0$, passing to a subsequence if necessary, we can assume that

$$\int_\eta^s \Phi_n(r) dW_\eta(r) \to I_\nu(s), \quad \mathbb{P} \text{ a.e.}, \tag{2.17}$$

say on a set Ω_2, where $\mathbb{P}(\Omega_2) = 1$ and we can assume that $\mathbb{P}_{\omega_0}(\Omega_2) = 1$ for $\mathbb{P}$ a.e. ω_0.

On the other hand, again by using conditional expectation as in (2.15), we know that, up to a subsequence, for $\mathbb{P}$ a.e. ω_0 we have $\|\Phi - \Phi_n\|_{\mathcal{N}^2_{Q,\nu_{\omega_0}}(\eta,T;H)} \to 0$. So, for $\mathbb{P}$ a.e. ω_0, we have that there exists a subsequence of Φ_n such that

$$\int_\eta^s \Phi_n(r) dW_\eta(r) \to I_{\nu_{\omega_0}}(s), \quad \mathbb{P}_{\omega_0} \text{ a.e.}. \tag{2.18}$$

Since $\mathbb{P}_{\omega_0}(\Omega_2) = 1$ for $\mathbb{P}$ a.e. ω_0, (2.17) and (2.18), imply that, for $\mathbb{P}$ a.e. ω_0, $I_\nu(s) = I_{\nu_{\omega_0}}(s)$, $\mathbb{P}_{\omega_0}$ a.e..

A different approach to proving $(A3)$ can be found in [427]. □

REMARK 2.17 Another two examples of systems satisfying Hypotheses 2.11 and 2.12 are given by the boundary control system described in Section 1.5.1 (Theorem 1.135) and by the semilinear system with non-nuclear covariance described in Section 1.5.2 (Proposition 1.141). We briefly sketch how one can show that Hypotheses 2.11 and 2.12 hold in these two cases. However we point out that these cases do not fully conform to our general abstract control problem as additional integrability conditions on the controls must be assumed to guarantee the existence and uniqueness of a unique mild solution. Thus the formulation of the control problem must be slightly adjusted in an obvious way.

Concerning the case of Section 1.5.1, suppose that the assumptions of Theorem 1.135 including (1.57) are satisfied. Then Hypothesis 2.11 follows from Theorem 1.135. Regarding Hypothesis 2.12, $(A0)$ and $(A2)$ follow from part (i) of Proposition

[3] Observe that we can compute the integral $I_{\nu_{\omega_0}}(s)$ since the control and the trajectory $X^\nu(\cdot)$ are $\mathscr{F}^\eta_{\omega_0,s}$-progressively measurable on $[\eta, T]$ for $\mathbb{P}$ a.e. ω_0.

1.136, $(A1)$ follows from part (ii) of Proposition 1.136 while for $(A3)$ one can argue as in Proposition 2.16, using $(A4)$ which holds by Remark 2.27.

As regards the case of Section 1.5.2, suppose that the assumptions of Propositions 1.141 and 1.142 are satisfied and $a_2(\cdot)$ is as in Remark 1.140. Hypothesis 2.11 follows from Propositions 1.141 For Hypothesis 2.12, $(A0)$ and $(A2)$ follow from part (i) of Proposition 1.142, $(A1)$ follows from part (ii) of Proposition 1.142, while for $(A3)$ one can again argue as in Proposition 2.16 using $(A4)$. ■

2.3. Dynamic Programming Principle: statement and proof

This section is devoted to the formulation and the proof of the Dynamic Programming Principle. Throughout the whole section we always assume that Hypothesis 2.9 is satisfied. We begin with a technical subsection.

2.3.1. Pullback to the canonical reference probability space.

Fix $t \in [0,T]$. The canonical reference probability space is the 5-tuple $\nu_W := (\mathbf{W}, \mathcal{F}_*, \mathbb{P}_*, \mathcal{B}^{\mathbf{t}}_{\mathbf{s}}, \mathcal{W})$, where $\mathbf{W} := \{\omega \in C([t,T];\Xi) : \omega(t) = 0\}$ equipped with the usual sup-norm, $\mathbb{P}_*$ is the Wiener measure on $(\mathbf{W}, \mathcal{B}(\mathbf{W}))$ (where $\mathcal{B}(\mathbf{W})$ is the Borel σ-field), i.e. the unique probability measure on $\mathbf{W}$ that makes the mapping

$$\begin{cases} \mathcal{W}\colon [t,T] \times \mathbf{W} \to \Xi \\ \mathcal{W}(s,\omega) = \omega(s) \end{cases} \tag{2.19}$$

a Q-Wiener process in Ξ (see [300]), $\mathcal{F}_*$ is the completion of $\mathcal{B}(\mathbf{W})$, and for $s \in [t,T]$, $\mathcal{B}^{t,0}_s = \sigma(\mathcal{W}(\tau) : t \le \tau \le s)$, $\mathcal{B}^t_s = \sigma\left(\mathcal{B}^{t,0}_s, \mathcal{N}^*\right)$, where $\mathcal{N}^*$ are the $\mathbb{P}_*$-null sets. $\mathbf{W}$ is a Polish space.

It is easy to see that $\mathcal{B}(\mathbf{W})$ is generated by the one dimensional cylinder sets $C = \{\omega : \omega(s) \in A\}$, where $s \in [t,T]$, A is open in Ξ, and that $\mathcal{B}^{t,0}_T = \mathcal{B}(\mathbf{W})$ (Lemma 2.18). Theorem 1.12 thus guarantee that ν_W is a standard reference probability space.

The canonical reference probability space on $[t,+\infty)$ is defined the same except that now $\mathbf{W} := \{\omega \in C([t,+\infty);\Xi) : \omega(t) = 0\}$ is equipped with the metric

$$\rho(w_1,w_2) = \sum_{n=1}^{\infty} 2^{-n}(\|w_1 - w_2\|_{C([t,t+n];\Xi)} \wedge 1),$$

which makes it a Polish space.

Lemma 2.18 *Let for $s \in [t,T]$ the map $\varphi_s : \mathbf{W} \to \mathbf{W}$ be defined by $\varphi_s(\omega)(\tau) = \omega(\tau \wedge s)$. Then*

$$\mathcal{B}^{t,0}_s = \varphi_s^{-1}(\mathcal{B}(\mathbf{W})).$$

In particular $\mathcal{B}^{t,0}_T = \mathcal{B}(\mathbf{W})$.

Proof. Observe that for a one dimensional cylinder $C = \{\omega : \omega(r) \in A\}$, where $r \in [t,T]$, A is open in Ξ, we have

$$\varphi_s^{-1}(C) = \{\omega \in \mathbf{W} \ : \ \varphi_s(\omega)(r) \in C\} = \{\omega \in \mathbf{W} \ : \ \omega(r \wedge s) \in C\} \in \mathcal{B}^{t,0}_s.$$

Since the cylinder sets C generate $\mathcal{B}(\mathbf{W})$, we thus obtain $\varphi_s^{-1}(\mathcal{B}(\mathbf{W})) \subset \mathcal{B}^{t,0}_s$.

For the opposite inclusion, since $\mathcal{B}^{t,0}_s$ is generated by sets of the form $B = \mathcal{W}^{-1}(r,\cdot)(V)$, where $r \in [t,s]$, V is open in Ξ, we have

$$B = \{\omega \in \mathbf{W} \ : \ \omega(r) \in V\} = \{\omega \in \mathbf{W} \ : \ \omega(r \wedge s) \in V\} = \varphi_s^{-1}(\{\omega \in \mathbf{W} : \omega(r) \in V\}).$$

Thus $\mathcal{B}^{t,0}_s \subset \varphi_s^{-1}(\mathcal{B}(\mathbf{W}))$. □

LEMMA 2.19 *Let $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W)$ be a reference probability space (i.e. it satisfies Definition 2.7), and let the paths of the Q-Wiener process $W(\cdot,\omega)$ be continuous for every $\omega \in \Omega$. Then, for $s \in [t,T]$,*

$$\mathscr{F}_s^{t,0} = W(\cdot \wedge s)^{-1}(\mathcal{B}(\mathbf{W})).$$

PROOF. The proof is similar to that of Lemma 2.18. □

We denote by $\mathcal{P}^{\mathbf{W}}_{[t,T]}$ to be the sigma field of $\mathcal{B}_s^{t,0}$-predictable sets, i.e. the sigma field generated by the sets of the form $(s,r] \times A, t \le s < r \le T, A \in \mathcal{B}_s^{t,0}$ and $\{t\} \times A, A \in \mathcal{B}_t^{t,0}$. For a reference probability space $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W_Q)$ we denote by $\mathcal{P}^{\Omega}_{[t,T]}$ to be the sigma field of $\mathscr{F}_s^{t,0}$-predictable sets.

We will use the following simple representation lemma from [427].

LEMMA 2.20 *Let $a(\cdot) = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W, a(\cdot)) \in \mathcal{U}_t$ (defined by (2.9)) be $\mathscr{F}_s^{t,0}$-predictable, and let the paths of the Q-Wiener process $W(\cdot,\omega)$ be continuous for every $\omega \in \Omega$. Then there exists a $\mathcal{P}^{\mathbf{W}}_{[t,T]}/\mathcal{B}(\Lambda)$-measurable function $f : [t,T] \times \mathbf{W} \to \Lambda$ such that*

$$a(s,\omega) = f(s, W(\cdot,\omega)), \quad \textit{for } \omega \in \Omega,\ s \in [t,T]. \tag{2.20}$$

PROOF. Define the process

$$\begin{cases} \beta \colon [t,T] \times \Omega \to [t,T] \times \mathbf{W} \\ \beta(\tau,\omega) = (\tau, W(\cdot,\omega)). \end{cases}$$

The sets of the form $A_1 = (s,r] \times \{\omega \in \Omega :\ W(\eta,\omega) \in B\}, t \le \eta \le s < r \le T, B \in \mathcal{B}(\Xi)$, and $A_2 = \{t\} \times \{\omega \in \Omega :\ W(t,\omega) \in B\}, B \in \mathcal{B}(\Xi)$, generate $\mathcal{P}^{\Omega}_{[t,T]}$. But $(\tau,\omega) \in A_1$ if and only if $\tau \in (s,r]$ and $W(\cdot,\omega) \in \tilde{B}_1 = \{\xi \in \mathbf{W} : \xi(\eta) \in B\} \in \mathcal{B}_s^{t,0}$, and $(t,\omega) \in A_2$ if and only if $W(\cdot,\omega) \in \tilde{B}_2 = \{\xi \in \mathbf{W} : \xi(t) \in B\} \in \mathcal{B}_t^{t,0}$. Therefore, $A_1 = \beta^{-1}((s,r] \times \tilde{B}_1)$, $A_2 = \beta^{-1}(\{t\} \times \tilde{B}_2)$. Since the sets of the form $(s,r] \times \{\xi \in \mathbf{W} :\ \xi(\eta) \in B\}, t \le \eta \le s < r \le T, B \in \mathcal{B}(\Xi)$, and $\{t\} \times \{\xi \in \mathbf{W} : \xi(t) \in B\}, B \in \mathcal{B}(\Xi)$, generate $\mathcal{P}^{\mathbf{W}}_{[t,T]}$, we have $\mathcal{P}^{\Omega}_{[t,T]} = \beta^{-1}(\mathcal{P}^{\mathbf{W}}_{[t,T]})$. Therefore, by Theorem 1.9, there exists a $\mathcal{P}^{\mathbf{W}}_{[t,T]}/\mathcal{B}(\Lambda)$-measurable function $f : [t,T] \times \mathbf{W} \to \Lambda$ such that (2.20) is satisfied. □

COROLLARY 2.21 *Let $a(\cdot) = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W, a(\cdot)) \in \mathcal{U}_t$ be $\mathscr{F}_s^{t,0}$-predictable. Let $\left(\Omega_1, \mathscr{F}_1, \mathscr{F}_{1,s}^t, \mathbb{P}_1, W_1\right)$ be another reference probability space. Suppose that W and W_1 have everywhere continuous trajectories. Let $f : [t,T] \times \mathbf{W} \to \Lambda$ be the function from Lemma 2.20 satisfying (2.20). Then the process*

$$\tilde{a}(s,\omega_1) = f(s, W_1(\cdot,\omega_1)) \tag{2.21}$$

is $\mathscr{F}_{1,s}^{t,0}$-predictable and hence $\mathscr{F}_{1,s}^{t,0}$-progressively measurable on $[t,T] \times \Omega_1$, and

$$\mathcal{L}_{\mathbb{P}}(a(\cdot), W(\cdot)) = \mathcal{L}_{\mathbb{P}_1}(\tilde{a}(\cdot), W_1(\cdot)). \tag{2.22}$$

2.3.2. Independence of reference probability spaces. To prove the Dynamic Programming Principle we have formulated our optimal control problem in a special weak form in which we only use reference probability spaces. Here we show that the control problem does not depend on the choice of a reference probability space ν and thus the strong and weak formulations are equivalent.

We will formulate the result only for the case $T < +\infty$ however the reader can easily modify the assumptions so that the result also holds for $T = +\infty$.

THEOREM 2.22 (Independence of the reference probability space) *Let $T \in (0,+\infty)$. Let Hypotheses 2.9, 2.11, 2.12 be satisfied. Let the functions $l : [0,T] \times H \times \Lambda \to \mathbb{R}$, $g : H \to \mathbb{R}$, $c : H \to \mathbb{R}$ be Borel measurable, c be bounded from* *below, and let, for every* $0 \le t < T, x \in H$*, every reference probability space* $\nu = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W)$ *and* $a(\cdot) \in \mathcal{U}_t^\nu$*,*

$$l(\cdot, X(\cdot; t, x, a(\cdot)), a(\cdot)) \in M_\nu^1(t, T; \mathbb{R}), \; g(X(T; t, x, a(\cdot))) \in L^1(\Omega, \mathscr{F}, \mathbb{P}),$$

Then, for every $0 \le t < T, x \in H$*, every two reference probability spaces* $\nu_1 = \left(\Omega_1, \mathscr{F}_1, \mathscr{F}_{1,s}^t, \mathbb{P}_1, W_1\right), \nu_2 = \left(\Omega_2, \mathscr{F}_2, \mathscr{F}_{2,s}^t, \mathbb{P}_2, W_2\right)$*, and* $a(\cdot) \in \mathcal{U}_t^{\nu_1}$*, there exists* $a_2(\cdot) \in \mathcal{U}_t^{\nu_2}$ *such that*

$$\mathcal{L}_{\mathbb{P}_1}(X^{\nu_1}(\cdot; t, x, a(\cdot)), a(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(X^{\nu_2}(\cdot; t, x, a_2(\cdot)), a_2(\cdot)).$$

In particular, for every reference probability space ν*,*

$$V_t^\nu(x) = V(t, x).$$

Proof. Let $a(\cdot) \in \mathcal{U}_t^{\nu_1}$. Let $a_1(\cdot)$ be the $\mathscr{F}_s^{t,0}$-predictable processes from Lemma 1.94 such that $a_1(\cdot) = a(\cdot)$, $\mathbb{P}_1 \otimes dt$ a.e.. Let $\tilde{a}_1(\cdot) \in \mathcal{U}_t^{\nu_2}$ be the process from Corollary 2.21. (Without loss of generality we can assume that W_1, W_2 have everywhere continuous trajectories.) Since $\mathcal{L}_{\mathbb{P}_1}(a_1(\cdot), W_1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(\tilde{a}_1(\cdot), W_2(\cdot))$, it follows from $(A1)$, $(A2)$, and Theorem 1.128 that, denoting $X^{\nu_1}(\cdot) = X^{\nu_1}(\cdot; t, x, a_1(\cdot)), X^{\nu_2}(\cdot) = X^{\nu_2}(\cdot; t, x, \tilde{a}_1(\cdot))$,

$$f_1(s) = e^{-\int_\eta^s c(X^{\nu_1}(\tau))d\tau}, \quad f_2(s) = e^{-\int_\eta^s c(X^{\nu_2}(\tau))d\tau},$$

we have

$$\mathcal{L}_{\mathbb{P}_1}(f_1(\cdot), X^{\nu_1}(\cdot), a(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(f_2(\cdot), X^{\nu_2}(\cdot), \tilde{a}_1(\cdot)). \tag{2.23}$$

This shows the first claim.

Using (2.23) we thus obtain

$$J^{\nu_1}(t, x; a(\cdot)) = J^{\nu_1}(t, x; a_1(\cdot)) = J^{\nu_2}(t, x; \tilde{a}_1(\cdot)),$$

which implies

$$\inf_{a(\cdot) \in \mathcal{U}_t^{\nu_1}} J^{\nu_1}(t, x; a(\cdot)) \ge \inf_{a(\cdot) \in \mathcal{U}_t^{\nu_2}} J^{\nu_2}(t, x; a(\cdot)).$$

The opposite inequality is obtained in the same way and thus it follows that

$$\inf_{a(\cdot) \in \mathcal{U}_t^{\nu_1}} J^{\nu_1}(t, x; a(\cdot)) = \inf_{a(\cdot) \in \mathcal{U}_t^{\nu_2}} J^{\nu_2}(t, x; a(\cdot)). \tag{2.24}$$

This completes the proof. □

2.3.3. The proof of the abstract principle of optimality. We now state and prove the Dynamic Programming Principle (DPP) in an abstract formulation. We will do it only for the finite horizon problem. However the same proof applies to the infinite horizon case if Hypothesis 2.23 is slightly changed to accommodate for $T = +\infty$. A special infinite horizon case is discussed in more details in Section 2.4.

Hypothesis 2.23 *Let* $T \in (0, +\infty)$*. The functions* $l : [0, T] \times H \times \Lambda \to \mathbb{R}$*,* $g : H \to \mathbb{R}$*,* $c : H \to \mathbb{R}$ *are Borel measurable, and* c *is bounded from below. For every* $0 \le t \le \eta < T, x \in H$*, every reference probability space* $\nu = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W)$ *and* $a(\cdot) \in \mathcal{U}_t^\nu$

$$l(\cdot, X(\cdot; t, x, a(\cdot)), a(\cdot)) \in M_\nu^1(t, T; \mathbb{R}), \; g(X(T; t, x, a(\cdot))) \in L^1(\Omega, \mathscr{F}, \mathbb{P}),$$

$$V(\eta, X(\eta; t, x, a(\cdot))) \in L^1(\Omega, \mathscr{F}, \mathbb{P}).$$

Moreover the functional $J(t, y; a(\cdot))$ *is uniformly continuous in the variable* y *on bounded sets of* H*, uniformly for* $a(\cdot) \in \mathcal{U}_t$*.*

Hypothesis 2.23 in particular ensures that the value function V is finite.

THEOREM 2.24 (Dynamic Programming Principle) *Assume that Hypotheses 2.9, 2.11, 2.12, and 2.23 are satisfied. Let $0 \le t < \eta < T, x \in H$. Then*

$$V(t,x) = \inf_{a(\cdot)\in\mathcal{U}_t} \mathbb{E}\left[\int_t^\eta e^{-\int_t^s c(X(\tau))d\tau} l(s, X(s), a(s))\, ds + e^{-\int_t^\eta c(X(\tau))d\tau} V(\eta, X(\eta))\right]. \tag{2.25}$$

The proof is very technical so we will proceed slowly. We begin with two simple lemmas.

LEMMA 2.25 *Let $0 \le t \le \eta < T$. Let $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W)$ be a standard reference probability space, and let W have everywhere continuous trajectories. Define, for $s \in [\eta, T]$, $W_\eta(s) := W(s) - W(\eta)$. Then for $\mathbb{P}$ a.e. $\omega_0 \in \Omega$, W_η is a Q-Wiener process on $\left(\Omega, \mathscr{F}_{\omega_0}, \mathscr{F}^\eta_{\omega_0,s}, \mathbb{P}_{\omega_0}\right)$, where $\mathbb{P}_{\omega_0} = \mathbb{P}(\cdot|\mathscr{F}_\eta^{t,0})(\omega_0)$ is the regular conditional probability, $\mathscr{F}_{\omega_0}$ is the augmentation of $\mathscr{F}'$ (see Definition 2.8) by the $\mathbb{P}_{\omega_0}$ null sets, and $\mathscr{F}^\eta_{\omega_0,s}$ is the augmented filtration generated by W_η.*

PROOF. We notice that for $\eta \le s \le T$

$$\mathscr{F}^{\eta,0}_{\omega_0,s} = \sigma\left(W_\eta(\tau) : \eta \le \tau \le s\right) \subset \mathscr{F}_s^{t,0}$$

(observe that $\mathscr{F}^{\eta,0}_{\omega_0,t_1}$ is independent of ω_0) and, by Lemma 2.26-(i), for $\mathbb{P}$ a.e. ω_0, $\mathscr{F}_s^{t,0} \subset \mathscr{F}^\eta_{\omega_0,s}$ for every $\eta \le s \le T$. Thus for $\mathbb{P}$ a.e. ω_0, $\mathscr{F}^\eta_{\omega_0,s}$ is the augmentation of $\mathscr{F}_s^{t,0}$ by the $\mathbb{P}_{\omega_0}$ null sets for every $\eta \le s \le T$.

We fix $\eta \le t_1 < t_2, y \in \Xi$. We want to apply Lemma 1.88 (with $\Xi_1 = \Xi$ and $Q_1 = Q$) so we need to compute for $\mathbb{P}$ a.s ω_0,

$$h := \mathbb{E}_{\omega_0}\left[e^{i\langle y, W_\eta(t_2)-W_\eta(t_1)\rangle_\Xi} | \mathscr{F}^\eta_{\omega_0,t_1}\right] = \mathbb{E}_{\omega_0}\left[e^{i\langle y, W_\eta(t_2)-W_\eta(t_1)\rangle_\Xi} | \mathscr{F}^{t,0}_{t_1}\right] \quad \mathbb{P}_{\omega_0}\ a.s.$$

Thus we can assume that h is $\mathscr{F}^{t,0}_{t_1}$-measurable. We have

$$\int_A h(\omega) d\mathbb{P}_{\omega_0}(\omega) = \int_A e^{i\langle y, W_\eta(t_2)-W_\eta(t_1)\rangle_\Xi}(\omega) d\mathbb{P}_{\omega_0}(\omega) \qquad \forall A \in \mathscr{F}^{t,0}_{t_1},$$

which (by the definition of $\mathbb{P}_{\omega_0}$) means that for $\mathbb{P}$ a.s ω_0

$$\begin{aligned}\int_A h(\omega) d\mathbb{P}_{\omega_0}(\omega) &= \mathbb{E}\left[e^{i\langle y, W_\eta(t_2)-W_\eta(t_1)\rangle_\Xi}\mathbf{1}_A | \mathscr{F}^{t,0}_\eta\right](\omega_0) \\ &= \mathbb{E}\left[\mathbf{1}_A \mathbb{E}\left[e^{i\langle y, W(t_2)-W(t_1)\rangle_\Xi} | \mathscr{F}^t_{t_1}\right] | \mathscr{F}^{t,0}_\eta\right](\omega_0) \\ &= \mathbb{E}\left[e^{-\frac{\langle Qy,y\rangle_\Xi}{2}(t_2-t_1)}\mathbf{1}_A | \mathscr{F}^{t,0}_\eta\right](\omega_0) = e^{-\frac{\langle Qy,y\rangle_\Xi}{2}(t_2-t_1)}\mathbb{P}_{\omega_0}(A). \end{aligned} \tag{2.26}$$

Therefore, since $\mathscr{F}^{t,0}_{t_1}$ is countably generated, it follows that $h = e^{-\frac{\langle Qy,y\rangle_\Xi}{2}(t_2-t_1)}$ for $\mathbb{P}$ a.s. ω_0. Thus by the separability of Ξ, for $\mathbb{P}$ a.s. ω_0 we have $h = e^{-\frac{\langle Qy,y\rangle_\Xi}{2}(t_2-t_1)}$ for all $y \in \Xi$. Consider now all pairs $(t_1^k, t_2^k), k = 1, 2, ...,$ where $t_1^k = \eta$ or t_1^k is rational, t_2^k is rational and $\eta \le t_1^k < t_2^k \le T$. We can conclude from the above that there is a set Ω_0 such that $\mathbb{P}(\Omega_0) = 1$ and such that for every $\omega_0 \in \Omega_0$, $y \in \Xi$ and $k = 1, 2, ...$

$$\mathbb{E}_{\omega_0}\left[e^{i\left\langle y, W_\eta(t_2^k)-W_\eta(t_1^k)\right\rangle_\Xi} | \mathscr{F}^{t,0}_{t_1^k}\right] = e^{-\frac{\langle Qy,y\rangle_\Xi}{2}(t_2^k-t_1^k)}. \tag{2.27}$$

It remains to prove that if $\omega_0 \in \Omega_0$, $y \in \Xi$ and $\eta \le t_1 < t_2 \le T$, then

$$\mathbb{E}_{\omega_0}\left[e^{i\langle y, W_\eta(t_2)-W_\eta(t_1)\rangle_\Xi} | \mathscr{F}^{t,0}_{t_1}\right] = e^{-\frac{\langle Qy,y\rangle_\Xi}{2}(t_2-t_1)}. \tag{2.28}$$

So let $\omega_0 \in \Omega_0$, $y \in \Xi$ and $\eta \le t_1 < t_2 \le T$. We will assume that $t_1 \ne \eta$ and t_1, t_2 are not rational since in such cases the argument is similar and easier. Then for some subsequence of our sequence of pairs, which we will still denote by (t_1^k, t_2^k), we have $t_1^k \to t_1, t_2^k \to t_2$ and $t_1^k < t_1, t_2^k < t_2$. We claim that $\mathbb{P}_{\omega_0}$ a.s.

$$\lim_{k\to+\infty} \mathbb{E}_{\omega_0}\left[e^{i\left\langle y, W_\eta(t_2^k)-W_\eta(t_1^k)\right\rangle_\Xi}|\mathscr{F}^{t,0}_{t_1^k}\right] = \mathbb{E}_{\omega_0}\left[e^{i\langle y, W_\eta(t_2)-W_\eta(t_1)\rangle_\Xi}|\mathscr{F}^{t,0}_{t_1}\right] \tag{2.29}$$

which, together with (2.27), will establish (2.28). First we notice that, since the filtration $\mathscr{F}^{t,0}_s$ is left continuous, by Proposition 1.40-(viii)

$$\lim_{k\to+\infty} \mathbb{E}_{\omega_0}\left[e^{i\langle y, W_\eta(t_2)-W_\eta(t_1)\rangle_\Xi}|\mathscr{F}^{t,0}_{t_1^k}\right] = \mathbb{E}_{\omega_0}\left[e^{i\langle y, W_\eta(t_2)-W_\eta(t_1)\rangle_\Xi}|\mathscr{F}^{t,0}_{t_1}\right] \ \ \mathbb{P}_{\omega_0}\ a.s.$$

Then we observe that by Proposition 1.40-(vi)

$$\begin{aligned}&\mathbb{E}_{\omega_0}\left|\mathbb{E}_{\omega_0}\left[e^{i\left\langle y, W_\eta(t_2^k)-W_\eta(t_1^k)\right\rangle_\Xi}|\mathscr{F}^{t,0}_{t_1^k}\right] - \mathbb{E}_{\omega_0}\left[e^{i\langle y, W_\eta(t_2)-W_\eta(t_1)\rangle_\Xi}|\mathscr{F}^{t,0}_{t_1^k}\right]\right| \\ &\le \sqrt{2}\mathbb{E}_{\omega_0}\left|e^{i\left\langle y, W_\eta(t_2^k)-W_\eta(t_1^k)\right\rangle_\Xi} - e^{i\langle y, W_\eta(t_2)-W_\eta(t_1)\rangle_\Xi}\right| \to 0 \ \text{ as } k\to+\infty.\end{aligned}$$

Thus for some subsequence, $\mathbb{P}_{\omega_0}$ a.s.

$$\lim_{k\to+\infty}\left|\mathbb{E}_{\omega_0}\left[e^{i\left\langle y, W_\eta(t_2^k)-W_\eta(t_1^k)\right\rangle_\Xi}|\mathscr{F}^{t,0}_{t_1^k}\right] - \mathbb{E}_{\omega_0}\left[e^{i\langle y, W_\eta(t_2)-W_\eta(t_1)\rangle_\Xi}|\mathscr{F}^{t,0}_{t_1^k}\right]\right| = 0.$$

These two convergences show (2.29). □

The reader can consult [427] for a different argument to prove Lemma 2.25.

LEMMA 2.26 *Let $0 \le t \le \eta < T$. Let $\nu = (\Omega, \mathscr{F}, \mathscr{F}^t_s, \mathbb{P}, W)$ be a standard reference probability space and let W have everywhere continuous trajectories. Let $a(\cdot) \in \mathcal{U}^\nu_t$, and let $a_1(\cdot)$ be from Lemma 1.94. Then*

(i) *For $\mathbb{P}$ a.e. $\omega_0 \in \Omega$, $\mathscr{F}^{t,0}_s \subset \mathscr{F}^\eta_{\omega_0,s}$ for every $\eta \le s \le T$.*

(ii) *For $\mathbb{P}$ a.e. $\omega_0 \in \Omega$, $a^{\omega_0}(\cdot) := \left(\Omega, \mathscr{F}_{\omega_0}, \mathscr{F}^\eta_{\omega_0,s}, \mathbb{P}_{\omega_0}, W_\eta, a_1|_{[\eta,T]}(\cdot)\right) \in \mathcal{U}_\eta$.*

PROOF. To show Part *(i)*, we take a countable generating family $\{A_k\}$ of $\mathcal{B}(\Xi)$ and a countable dense subset $\{s_m\}$ in $[t,T]$. We will show that for a.e. $\omega_0 \in \Omega$, $W(s_m)^{-1}(A_k) \in \mathscr{F}^\eta_{\omega_0,s}$ for all $k \ge 1, s_m \le s$. If $s_m \le \eta$, since $\mathscr{F}^{t,0}_\eta$ is countably generated, we obtain by Theorem 1.44 that $W(s_m)(\omega) = W(s_m)(\omega_0)$, $\mathbb{P}_{\omega_0}$ a.e., for $\mathbb{P}$ a.e. $\omega_0 \in \Omega$. Thus, up to a set of $\mathbb{P}_{\omega_0}$ measure 0, $W(s_m)^{-1}(A_k)$ is either empty or is equal to Ω and so it is in $\mathscr{F}^\eta_{\omega_0,s}$. If $s_m > \eta$ then again up to a set of $\mathbb{P}_{\omega_0}$ measure 0, $W(s_m)^{-1}(A_k) = W_\eta(s_m)^{-1}(A_k - W_\eta(\omega_0))$ and so it is in $\mathscr{F}^\eta_{\omega_0,s}$ for $\mathbb{P}$ a.e. $\omega_0 \in \Omega$. This implies that $\sigma\{W(s_m) : s_m \le s\} \subset \mathscr{F}^\eta_{\omega_0,s}$ for $\mathbb{P}$ a.e. $\omega_0 \in \Omega$. It remains to notice that $\mathscr{F}^{t,0}_s = \sigma\{W(s_m) : s_m \le s\}$.

Part *(ii)*: In view of Lemma 2.25 it is enough to show that for $\mathbb{P}$ a.e. $\omega_0 \in \Omega$, $a_1(\cdot)$ is $\mathscr{F}^\eta_{\omega_0,s}$-progressively measurable on $[\eta, T]$. This fact follows from Part *(i)*. □

REMARK 2.27 If $a(\cdot)$ in Lemma 2.26 is such that $a(\cdot) \in M^p_\nu(t,T;E), p \ge 1$, for some Banach space E, it is easy to see that we also have $a_1|_{[\eta,T]}(\cdot) \in M^p_{\nu_{\omega_0}}(\eta,T;E)$, for $\mathbb{P}$ a.e. ω_0, where $\nu_{\omega_0} = \left(\Omega, \mathscr{F}_{\omega_0}, \mathscr{F}^\eta_{\omega_0,s}, \mathbb{P}_{\omega_0}, W_\eta\right)$ is as in Lemma 2.26. ■

PROOF OF THEOREM 2.24. We first do the following reduction. Denote (similarly to (2.9))

$$\tilde{\mathcal{U}}_t = \left\{\bigcup_\nu \mathcal{U}^\nu_t : \nu \text{ is a standard reference probability space}\right\}.$$

The set $\tilde{\mathcal{U}}_t$ is nonempty since for instance the canonical reference probability space ν_W is in it. It is clear from Theorem 2.22 (and the same argument used to justify

(2.23) there) that (2.25) will follow if we can prove it with $\mathcal{U}_t$ replaced by $\tilde{\mathcal{U}}_t$. Therefore it remains to show that

$$V(t,x) = \inf_{a(\cdot)\in\tilde{\mathcal{U}}_t} \mathbb{E}\Bigg[\int_t^\eta e^{-\int_t^s c(X(\tau))d\tau} l(s, X(s), a(s))\, ds \\ + e^{-\int_t^\eta c(X(\tau))d\tau} V(\eta, X(\eta))\Bigg]. \quad (2.30)$$

(In fact, using Theorem 2.22 it would be enough to replace $\mathcal{U}_t$ by $\mathcal{U}_t^{\nu_W}$ and do everything on canonical reference probability spaces, however we will do the proof in the more general setup since the arguments and technicalities are similar.) We remind that we assume that all Q-Wiener processes in the reference probability spaces have everywhere continuous trajectories.

Part 1. (inequality $\geq$ in (2.30)): Let $a(\cdot) \in \tilde{\mathcal{U}}_t$. Denoting $X(s) = X(s;t,x,a(\cdot)))$ we have

$$J(t,x;a(\cdot)) = \mathbb{E}\left[\int_t^\eta e^{-\int_t^s c(X(\tau))d\tau} l(s, X(s), a(s))\, ds\right] \\ +\mathbb{E}\left[\int_\eta^T e^{-\int_t^s c(X(\tau))d\tau} l(s, X(s), a(s))\, \mathrm{d}s + e^{-\int_t^T c(X(\tau))d\tau} g(X(T))\right] \quad (2.31)$$

Let $a_1(\cdot)$ be from Lemma 1.94, and for $\mathbb{P}$ a.e. $\omega_0 \in \Omega$, $a^{\omega_0}(\cdot)$ be the control from Lemma 2.26. Let $\mathbb{P}_{\omega_0} = \mathbb{P}\left(\cdot|\mathscr{F}_\eta^{t,0}\right)(\omega_0)$ and $\mathbb{E}_{\omega_0}$ be the expectation with respect to $\mathbb{P}_{\omega_0}$.

Let $X_1(s) = X(s;t,x,a_1(\cdot))$. By $(A2)$, X_1 and X are indistinguishable. Thus we obtain by $(A3)$ that (up to an indistinguishable modification) $X_1(s) = X(s;\eta,X(\eta),a^{\omega_0}(\cdot))$ in $\left(\Omega,\mathscr{F}_{\omega_0},\mathscr{F}^\eta_{\omega_0,s},\mathbb{P}_{\omega_0},W_\eta\right)$ for $\mathbb{P}$ a.s. ω_0.

Therefore, using this, $(A0)$ and the fact that for $\mathbb{P}$ a.s. ω_0, $\mathbb{P}_{\omega_0}(\{\omega : X_1(\eta,\omega) = X_1(\eta,\omega_0)\}) = 1$, we have

$$\mathbb{E}\left[\int_\eta^T e^{-\int_t^s c(X(\tau))d\tau} l(s, X(s), a(s))\, \mathrm{d}s + e^{-\int_t^T c(X(\tau))d\tau} g(X(T))\right] \\ = \mathbb{E}\left[\int_\eta^T e^{-\int_t^s c(X_1(\tau))d\tau} l(s, X_1(s), a_1(s))\, \mathrm{d}s + e^{-\int_t^T c(X_1(\tau))d\tau} g(X_1(T))\right] \\ = \mathbb{E}\Bigg[e^{-\int_t^\eta c(X(\tau))d\tau}\mathbb{E}\Bigg[\int_\eta^T e^{-\int_\eta^s c(X_1(\tau))d\tau} l(s, X_1(s), a_1(s))\, \mathrm{d}s \\ + e^{-\int_\eta^T c(X_1(\tau))d\tau} g(X_1(T))\,|\mathscr{F}_\eta^t\Bigg]\Bigg] \\ = \mathbb{E}\left[e^{-\int_t^\eta c(X(\tau))d\tau} J(\eta, X_1(\eta,\omega_0); a^{\omega_0}(\cdot))\right] \geq \mathbb{E}\left[e^{-\int_t^\eta c(X(\tau))d\tau} V(\eta, X_1(\eta,\omega_0))\right] \\ = \mathbb{E}\left[e^{-\int_t^\eta c(X(\tau))d\tau} V(\eta, X(\eta))\right], \quad (2.32)$$

where we used the remarks at the end of Section 2.2.2. Thus, using (2.31), we obtain

$$J(t,x;a(\cdot)) \geq \mathbb{E}\left[\int_t^\eta e^{-\int_t^s c(X(\tau))d\tau} l(s, X(s), a(s))\, ds + e^{-\int_t^\eta c(X(\tau))d\tau} V(\eta, X(\eta))\right]$$

and the claim follows by taking the infimum over all $a(\cdot) \in \tilde{\mathcal{U}}_t$.

Part 2. (inequality $\leq$ in (2.30)): Let $t \leq \eta \leq T$. We fix $a(\cdot) = (\Omega,\mathscr{F},\mathscr{F}_s^t,\mathbb{P},W,a(\cdot)) \in \tilde{\mathcal{U}}_t$.

We can choose $\delta_1 > 0$ be such that, for $|x - \bar{x}| < \delta_1$ and $|x|, |\bar{x}| \leq 1$ we have, for each $\tilde{a}(\cdot) \in \mathcal{U}_\eta$

$$|J(\eta, x; \tilde{a}(\cdot)) - J(\eta, \bar{x}; \tilde{a}(\cdot))| + |V(\eta, x) - V(\eta, \bar{x})| < \epsilon$$

Since H is separable we can choose a partition $\left\{D_j^R\right\}_{j\in\mathbb{N}}$ of $B_H(0,1)$ into countable disjoint Borel subsets with $diam(D_j^1) < \delta_1$. Similarly we can choose a (possibly smaller) $\delta_2 > 0$ such that, for $|x - \bar{x}| < \delta_2$ and $|x|, |\bar{x}| \leq 2$ we have, for each $\tilde{a}(\cdot) \in \mathcal{U}_\eta$

$$|J(\eta, x; \tilde{a}(\cdot)) - J(\eta, \bar{x}; \tilde{a}(\cdot))| + |V(\eta, x) - V(\eta, \bar{x})| < \epsilon$$

and we can choose a partition $\left\{D_j^2\right\}_{j\in\mathbb{N}}$ of $B_H(0,2)\backslash B_H(0,1)$ into countable disjoint Borel subsets with $diam(D_j^2) < \delta_2$.

Iterating the argument we can find a partition $\{D_j\}_{j\in\mathbb{N}}$ of H into countable disjoint Borel subsets with the following property: for all D_j and all $x, \bar{x} \in D_j$ we have, for each $\tilde{a}(\cdot) \in \mathcal{U}_\eta$,

$$|J(\eta, x; \tilde{a}(\cdot)) - J(\eta, \bar{x}; \tilde{a}(\cdot))| + |V(\eta, x) - V(\eta, \bar{x})| < \epsilon$$

For each $j \in \mathbb{N}$ we choose $x_j \in D_j$ and $a_j(\cdot) = \left(\Omega_j, \mathscr{F}_j, \mathscr{F}^\eta_{j,s}, \mathbb{P}_j, W_j, a_j(\cdot)\right) \in \mathcal{U}_\eta^{\nu_j}$ such that

$$J(\eta, x_j; a_j(\cdot)) < V(\eta, x_j) + \epsilon. \tag{2.33}$$

We define a new control $a^\eta(\cdot) \in \tilde{\mathcal{U}}_t$ on the probability space $(\Omega, \mathscr{F}, \mathscr{F}^t_s, \mathbb{P}, W)$ as follows. Let $a_{j,1}(\cdot)$ be the $\mathscr{F}^{\eta,0}_{j,s}$-predictable processes from Lemma 1.94 such that $a_{j,1}(\cdot) = a_j(\cdot)$, $\mathbb{P}_j \otimes dt$ a.e.. Let $f_j : [\eta, T] \times C([\eta, T]; \Xi) \to \Lambda$ be the functions from Lemma 2.20 such that

$$f_j(s, W_j(\cdot, \omega)) = a_{j,1}(s, \omega), \qquad \text{for } \omega \in \Omega_j,\ s \in [\eta, T].$$

We now set $\tilde{a}_j(s, \omega) = f_j(s, W_\eta(\cdot, \omega))$. By Corollary 2.21 and Lemma 2.26 the process $\tilde{a}_j(\cdot)$ is $\mathscr{F}^{t,0}_s$-progressively measurable and, for $\mathbb{P}$-a.e. ω_0, is $\mathscr{F}^\eta_{\omega_0,s}$-progressively measurable in the reference probability spaces $\nu_{\omega_0} := \left(\Omega, \mathscr{F}_{\omega_0}, \mathscr{F}^\eta_{\omega_0,s}\mathbb{P}_{\omega_0}, W_\eta\right)$ defined in Lemma 2.25. Moreover $\mathcal{L}_{\mathbb{P}_{\omega_0}}(\tilde{a}_j(\cdot), W_\eta(\cdot)) = \mathcal{L}_{\mathbb{P}_j}(a_{j,1}(\cdot), W_j(\cdot))$. We define

$$a^\eta(s, \omega) = a(s, \omega)\,\mathbf{1}_{\{t\leq s<\eta\}} + \mathbf{1}_{\{s\geq\eta\}} \sum_{j\in\mathbb{N}} \tilde{a}_j(s, \omega)\,\mathbf{1}_{\{X(\eta;t,x,a(\cdot))\in D_j\}}. \tag{2.34}$$

Obviously $(\Omega, \mathscr{F}, \mathscr{F}^t_s, \mathbb{P}, W, a^\eta(\cdot)) \in \tilde{\mathcal{U}}_t$.

Let $X(s) = X(s; t, x, a^\eta(\cdot))$. Notice that $X(s; t, x, a^\eta(\cdot)) = X(s; t, x, a(\cdot))$ on $[t, \eta]$, $\mathbb{P}$ a.e..

Denote $O_j := \{\omega : X(\eta; t, x, a(\cdot)) \in D_j\}$. Since for $\mathbb{P}$-a.s. ω_0, $\mathbb{P}_{\omega_0}(\{\omega : X(\eta, \omega) = X(\eta, \omega_0)\}) = 1$, if $\omega_0 \in O_j$, then $\mathbb{P}_{\omega_0}(\Omega \setminus O_j) = 0$, which implies that in this case $\tilde{a}_j(\cdot) = a^\eta(\cdot)$ on $[\eta, T]$, $\mathbb{P}_{\omega_0}$ a.s., and thus, for $\mathbb{P}$ a.s. ω_0, $a^\eta|_{[\eta,T]} \in \mathcal{U}_\eta^{\nu_{\omega_0}}$, and

$$\mathcal{L}_{\mathbb{P}_{\omega_0}}(a^\eta(\cdot), W_\eta(\cdot)) = \mathcal{L}_{\mathbb{P}_j}(a_{j,1}(\cdot), W_j(\cdot)), \quad j \in \mathbb{N}. \tag{2.35}$$

Moreover we can assume, by $(A3)$, that for $\mathbb{P}$ a.e. ω_0, $X(s) = X^{\nu_{\omega_0}}(\cdot; \eta, X(\eta), a^\eta(\cdot))$ on $[\eta, T]$ $\mathbb{P}_{\omega_0}$ a.s..

By the definition of V,

$$\begin{aligned}
V(t,x) &\leq \mathbb{E}\left[\int_t^T e^{-\int_t^s c(X(\tau))d\tau} l(s, X(s), a^\eta(s))\,ds + e^{-\int_t^T c(X(\tau))d\tau} g(X(T))\right] \\
&= \mathbb{E}\left[\int_t^\eta e^{-\int_t^s c(X(\tau))d\tau} l(s, X(s), a(s))\,ds\right] \\
&+\mathbb{E}\left[\int_\eta^T e^{-\int_t^s c(X(\tau))d\tau} l(s, X(s), a^\eta(s))\,ds + e^{-\int_t^T c(X(\tau))d\tau} g(X(T))\right].
\end{aligned} \tag{2.36}$$

We have

$$
\begin{aligned}
&\mathbb{E}\left[\int_\eta^T e^{-\int_t^s c(X(\tau))d\tau} l\left(s, X\left(s\right), a^\eta\left(s\right)\right) ds + e^{-\int_t^T c(X(\tau))d\tau} g\left(X\left(T\right)\right)\right] \\
&= \mathbb{E}\bigg[e^{-\int_t^\eta c(X(\tau))d\tau} \mathbb{E}\bigg[\int_\eta^T e^{-\int_\eta^s c(X(\tau))d\tau} l\left(s, X\left(s\right), a^\eta\left(s\right)\right) ds \\
&\qquad\qquad + e^{-\int_\eta^T c(X(\tau))d\tau} g\left(X\left(T\right)\right) | \mathscr{F}_\eta^t\bigg]\bigg] \\
&= \sum_{j\in\mathbb{N}} \int_{O_j} e^{-\int_t^\eta c(X(\tau))d\tau} \mathbb{E}_{\omega_0}\bigg[\int_\eta^T e^{-\int_\eta^s c(X(\tau))d\tau} l\left(s, X\left(s\right), a^\eta\left(s\right)\right) ds \\
&\qquad\qquad + e^{-\int_\eta^T c(X(\tau))d\tau} g\left(X\left(T\right)\right)\bigg] d\mathbb{P}(\omega_0)
\end{aligned}
$$

By (2.35) and $(A0) - (A1)$ we obtain

$$
\mathcal{L}_{\mathbb{P}_{\omega_0}}(X(\cdot), a^\eta(\cdot)) = \mathcal{L}_{\mathbb{P}_j}(X^{\nu_j}(\cdot), a_{j,1}(\cdot)), \quad j \in \mathbb{N},
$$

where $X^{\nu_j}(s) = X^{\nu_j}(s; \eta, X(\eta; t, x, a(\cdot))(\omega_0), a_{j,1}(\cdot))$. Thus, it follows from Theorem 1.128 that, denoting

$$
f(s) = e^{-\int_\eta^s c(X(\tau))d\tau}, \quad f_j(s) = e^{-\int_\eta^s c(X^{\nu_j}(\tau))d\tau},
$$

$$
\mathcal{L}_{\mathbb{P}_{\omega_0}}(f(\cdot), X(\cdot), a^\eta(\cdot)) = \mathcal{L}_{\mathbb{P}_j}(f_j(\cdot), X^{\nu_j}(\cdot), a_{j,1}(\cdot)), \quad j \in \mathbb{N}.
$$

Therefore,

$$
\begin{aligned}
&\mathbb{E}\left[\int_\eta^T e^{-\int_t^s c(X(\tau))d\tau} l\left(s, X\left(s\right), a^\eta\left(s\right)\right) ds + e^{-\int_t^T c(X(\tau))d\tau} g\left(X\left(T\right)\right)\right] \\
&= \sum_{j\in\mathbb{N}} \int_{O_j} e^{-\int_t^\eta c(X(\tau))d\tau} J_{\mathbb{P}_{\omega_0}}\left(\eta, X(\eta; t, x, a(\cdot))(\omega_0); a^\eta(\cdot)\right) d\mathbb{P}(\omega_0) \\
&= \sum_{j\in\mathbb{N}} \int_{O_j} e^{-\int_t^\eta c(X(\tau))d\tau} J_{\mathbb{P}_j}\left(\eta, X(\eta; t, x, a(\cdot))(\omega_0); a_{j,1}(\cdot)\right) d\mathbb{P}(\omega_0).
\end{aligned}
$$

Moreover, using (2.33), we get for a.s. $\omega_0 \in O_j$

$$
\begin{aligned}
J_{\mathbb{P}_j}\left(\eta, X\left(\eta; t, x, a(\cdot)\right)\left(\omega_0\right); a_j(\cdot)\right) &\le J_{\mathbb{P}_j}(\eta, x_j; a_j(\cdot)) + \epsilon \\
&\le V(\eta, x_j) + 2\epsilon \le V\left(\eta, X\left(\eta; t, x, a(\cdot)\right)\left(\omega_0\right)\right) + 3\epsilon,
\end{aligned}
$$

so we finally obtain

$$
\begin{aligned}
&\mathbb{E}\left[\int_\eta^T e^{-\int_t^s c(X(\tau))d\tau} l\left(s, X\left(s\right), a^\eta\left(s\right)\right) ds + e^{-\int_t^T c(X(\tau))d\tau} g\left(X\left(T\right)\right)\right] \\
&\le \mathbb{E}\left[e^{-\int_t^\eta c(X(\tau))d\tau} V\left(\eta, X\left(\eta; t, x, a(\cdot)\right)\right)\right] + C\epsilon.
\end{aligned}
$$

Therefore, by (2.36) and the arbitrariness of $a(\cdot)$,

$$
\begin{aligned}
V\left(t, x\right) \le \inf_{a(\cdot)\in\tilde{\mathcal{U}}_t} \mathbb{E}\bigg[\int_t^\eta & e^{-\int_t^s c(X(\tau))d\tau} l\left(s, X\left(s\right), a\left(s\right)\right) ds \\
&+ e^{-\int_t^\eta c(X(\tau))d\tau} V\left(\eta, X\left(\eta\right)\right)\bigg] + C\epsilon
\end{aligned}
$$

and the claim follows by letting $\epsilon \to 0$. □

If we know more information about the value function and the control problem, in particular that the value function is continuous in both variables, the dynamic programming principle can be strengthened to include stopping times. We do not do it here in the abstract case. We explain how to obtain such formulation of the dynamic programming principle for a control problem for mild solutions in Section 3.6, Theorem 3.70.

2.4. Infinite horizon problems

In this section we consider a special infinite horizon problem described by an evolution equation

$$\begin{cases} dX(s) = \beta(X(s), a(s))ds + \sigma(X(s), a(s))dW_Q(s) \\ X(t) = x, \end{cases} \tag{2.37}$$

with a cost functional of the form

$$J(t, x; a(\cdot)) = \mathbb{E}\left[\int_t^{+\infty} e^{-\int_t^s c(X(\tau;t,x,a(\cdot)))d\tau} l(X(s; t, x, a(\cdot)), a(s))ds\right], \tag{2.38}$$

where $c \geq \lambda > 0$. We are really only interested in the case $t = 0$ but we will keep the dependence on t for a while.

This is a very important class of problems which are semi-"autonomous" in a sense that the coefficients β, σ and the cost l do not depend explicitly on time. In this case the value function does not depend on time and the DPP takes on a simpler form.

We define the value function for $t \geq 0$ as

$$V(t, x) = \inf_{a(\cdot)\in\mathcal{U}_t} \mathbb{E}\left[\int_t^{+\infty} e^{-\int_t^s c(X(\tau;t,x,a(\cdot)))d\tau} l(X(s; t, x, a(\cdot)), a(s))ds\right] \tag{2.39}$$

and set

$$J(x; a(\cdot)) := J(0, x; a(\cdot)), \quad V(x) := V(0, x). \tag{2.40}$$

We assume now that Hypotheses 2.9, 2.11 and 2.12 are satisfied with $T = +\infty$ (i.e. reference probability spaces and solutions are defined on $[t, +\infty)$). We also replace Hypothesis 2.23 by the following one.

HYPOTHESIS 2.28 *The functions $l : H \times \Lambda \to \mathbb{R}$, $c : H \to \mathbb{R}$ are Borel measurable and there exists $\lambda > 0$ such that $c(x) \geq \lambda$ for every $x \in H$. Moreover for every $0 \leq \eta < +\infty, x \in H$, reference probability space ν, $a(\cdot) \in \mathcal{U}_0^\nu$*

$$e^{-\int_0^{\cdot} c(X(\cdot;0,x,a(\cdot)))d\tau} l(X(\cdot; 0, x, a(\cdot)), a(\cdot)) \in M^1_\nu(0, +\infty; \mathbb{R}),$$
$$V(X(\eta; 0, x, a(\cdot))) \in L^1(\Omega, \mathscr{F}, \mathbb{P}).$$

Finally $J(\cdot; a(\cdot))$ is uniformly continuous on bounded sets of H, uniformly for $a(\cdot) \in \mathcal{U}_0$.

Since we are dealing with an abstract state equation we have to add another hypothesis which reflects the "autonomous" nature of the system. First we notice that

$$\left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s\geq t}, \mathbb{P}, W_Q(\cdot), a(\cdot)\right) \in \mathcal{U}_t$$
$$\iff \left(\Omega, \mathscr{F}, \{\mathscr{F}_{s+t}^t\}_{s\geq 0}, \mathbb{P}, W_Q(t+\cdot), a(t+\cdot)\right) \in \mathcal{U}_0.$$

HYPOTHESIS 2.29 *Assume that the family of the solutions $X(\cdot; t, x, a(\cdot))$ of (2.37) satisfies the following property. For every $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W_Q(\cdot), a(\cdot)) \in \mathcal{U}_t$*

(A5) *$\mathcal{L}_{\mathbb{P}}(X(t+\cdot; t, x, a(\cdot)), a(t+\cdot)) = \mathcal{L}_{\mathbb{P}}(X(\cdot; 0, x, a(t+\cdot)), a(t+\cdot))$ on $[0, +\infty)$, where $X(\cdot; 0, x, a(t+\cdot))$ is the solution of (2.37) with $W_Q(\cdot)$ replaced by $W_Q(t+\cdot)$.*

REMARK 2.30 An example of a state equation satisfying Hypothesis 2.29 is given by the mild solution of an SDE

$$\begin{cases} dX(s) = AX(s)ds + b(X(s), a(s))ds + \sigma(X(s), a(s))dW_Q(s) \\ X(t) = x, \end{cases} \tag{2.41}$$

where A, b and σ satisfy the assumptions described in Hypothesis 1.119. ■

Using Hypothesis 2.29, by a change of variable and $(A5)$, we observe that

$$\begin{aligned} V(t,x) &= \inf_{a(\cdot)\in\mathcal{U}_t} \mathbb{E}\left[\int_t^{+\infty} e^{-\int_t^s c(X(\tau;t,x,a(\cdot)))d\tau} l(X\left(s;t,x,a(\cdot)\right), a(s))ds\right] \\ &= \inf_{a(\cdot)\in\mathcal{U}_t} \mathbb{E}\left[\int_0^{+\infty} e^{-\int_0^s c(X(t+\tau;t,x,a(\cdot)))d\tau} l(X\left(t+s;t,x,a(\cdot)\right), a(t+s))ds\right] \\ &= \inf_{a(\cdot)\in\mathcal{U}_t} \mathbb{E}\left[\int_0^{+\infty} e^{-\int_0^s c(X(\tau;0,x,a(t+\cdot)))d\tau} l(X\left(s;0,x,a(t+\cdot)\right), a(t+s))ds\right] \\ &= \inf_{a(\cdot)\in\mathcal{U}_0} \mathbb{E}\left[\int_0^{+\infty} e^{-\int_0^s c(X(\tau;0,x,a(\cdot)))d\tau} l(X\left(s;0,x,a(\cdot)\right) a(s))ds\right] = V(x). \end{aligned}$$

We thus have the following theorem whose proof is obtained by a simple modification of the proofs of Theorems 2.22 and 2.24.

THEOREM 2.31 (DPP, infinite horizon case) *Assume that Hypotheses 2.9, 2.11 and 2.12 for $T = +\infty$ hold, and that Hypotheses 2.28 and 2.29 are satisfied. Then the value function V satisfies the Dynamic Programming Principle: For every $\eta > 0$, $x \in H$,*

$$V(x) = \inf_{a(\cdot)\in\mathcal{U}_0} \mathbb{E}\left[\int_0^{\eta} e^{-\int_0^s c(X(\tau))d\tau} l\left(X(s), a(s)\right) ds + e^{-\int_0^{\eta} c(X(\tau))d\tau} V\left(X(\eta)\right)\right]. \tag{2.42}$$

Moreover

$$V(x) = V^{\nu}(x) := \inf_{a(\cdot)\in\mathcal{U}_0^{\nu}} \mathbb{E}\left[\int_0^{+\infty} e^{-\int_0^s c(X(\tau))d\tau} l\left(X(s), a(s)\right) ds\right]$$

for every reference probability space ν.

2.5. HJB equation and optimal synthesis in the smooth case

Once we know that the dynamic programming principle (DPP) holds, we want to use it to solve the control problem i.e. to find optimal couples and, possibly, to study their properties. In the dynamic programming approach the path to do this consists in:

- Writing a differential form of the DPP (the HJB equation);
- Finding a solution v of the HJB equation (which we do not know ex ante to be the value function);
- Using such solution v to prove a verification theorem i.e. sufficient, and possibly necessary, conditions for optimality which express optimal controls as functions of the current state (feedback controls);
- Performing the optimal synthesis, i.e. using the optimality conditions of the previous step to find optimal feedback controls: this will also imply that v is indeed the value function.

Such a program can be performed if we know in advance that the HJB equation has a smooth solution or if we know that the value function is sufficiently regular, both of which may not be true even in finite dimensions. However it is still useful to present how the program works in the smooth case to understand the machinery of the dynamic programming approach. We do it for our model problem, when the state equation admits a solution in the mild sense, assuming that the value function is smooth. We prove the following three results (in both finite and infinite horizon cases):

- The value function solves the HJB equation.
- The verification theorem (necessary and sufficient conditions for optimality).
- The existence of optimal couples in feedback form.

One of the main goals of the theory presented in this book is to obtain some of these results under more realistic assumptions.

It is important to note that, if one finds a sufficiently smooth solution of the HJB equation, then the verification theorem and the existence of optimal feedbacks can be done without using the DPP and this is done in Chapters 4-6.

We will present everything for a control problem in the weak formulation of Section 2.1.2, i.e. when the set of admissible controls is equal to $\overline{\mathcal{U}}_t$, as this setup is more convenient when discussing optimal feedback controls. However the same results are also true for control problems in the weak formulation of Section 2.2 used to prove the DPP, with $\overline{\mathcal{U}}_t$ replaced by $\mathcal{U}_t$, or in the strong formulation of Section 2.1.1.

2.5.1. Finite horizon problem: Parabolic HJB equation. Let Hypothesis 2.1 hold. Consider an optimal control problem of minimizing the cost functional (2.3) for the system governed by (2.13), where for simplicity we do not have discounting in (2.3), i.e. we set $c = 0$. We rewrite it here for the reader's convenience. The state equation is

$$\begin{cases} dX(s) = (AX(s) + b(s, X(s), a^\mu(s)))\, ds + \sigma(s, X(s), a^\mu(s))dW_Q(s) \\ gX(t) = x, \end{cases} \tag{2.43}$$

where $a^\mu(\cdot) \in \mathcal{U}_t^\mu$ for some generalized reference probability space μ satisfying Hypothesis 2.1, and the cost functional

$$J(t,x;a^\mu(\cdot)) = \mathbb{E}^\mu\left[\int_t^T l(s, X(s;t,x;a^\mu(\cdot)), a^\mu(s))ds + g(X(T;t,x,a^\mu(\cdot)))\right]. \tag{2.44}$$

We consider the control problem in the weak formulation of Section 2.1.2, and assume that Hypothesis 1.119 is satisfied. The HJB equation associated with this problem is

$$\begin{cases} v_t + \langle Dv, Ax\rangle + \inf_{a\in\Lambda}\left\{\frac{1}{2}\mathrm{Tr}\left[\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^* D^2 v\right]\right. \\ \qquad\qquad \left. + \langle Dv, b(t,x,a)\rangle + l(t,x,a)\right\} = 0, \\ v(T,x) = g(x). \end{cases} \tag{2.45}$$

In the above equation Dv, D^2v are the Fréchet derivatives of v with respect to x, which are identified respectively with elements of H and $S(H)$, the set of bounded, self-adjoint operators in the Hilbert space H. For $(t,x,p,S,a) \in [0,T]\times H\times H\times S(H)\times\Lambda$, the term

$$F_{CV}(t,x,p,S,a) := \frac{1}{2}\mathrm{Tr}\left[\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^* S\right] + \langle p, b(t,x,a)\rangle + l(t,x,a) \tag{2.46}$$

will be called the *current value Hamiltonian* of the system and its infimum over $a \in \Lambda$

$$F(t,x,p,S) := \inf_{a\in\Lambda} \left\{ \frac{1}{2} Tr \left[\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right) \left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^* S \right] \right. \\ \left. + \langle p, b(t,x,a)\rangle + l(t,x,a) \right\} \quad (2.47)$$

will be called the *Hamiltonian.*[4] Using this notation, the HJB equation (2.45) can be rewritten as

$$\begin{cases} v_t + \langle Dv, Ax\rangle + F(t,x,Dv,D^2v) = 0, \\ v(T,x) = g(x). \end{cases} \quad (2.48)$$

The HJB equation (2.45) can be viewed as a differential form of the DPP.

DEFINITION 2.32 (Classical solution, parabolic case) *A function $v\colon (0,T] \times H \to \mathbb{R}$ is a* classical solution *of (2.45) if $v \in C^{1,2}((0,T) \times H) \cap C((0,T] \times H)$, $Dv\colon (0,T) \times H \to D(A^*)$, $A^*Dv \in C((0,T) \times H; H)$ and v satisfies*

$$\begin{cases} v_t + \langle A^*Dv, x\rangle + \inf_{a\in\Lambda} \left\{ \frac{1}{2}\mathrm{Tr} \left[\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right) \left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^* D^2 v \right] \right. \\ \qquad \left. + \langle Dv, b(t,x,a)\rangle + l(t,x,a) \right\} = 0, \qquad (t,x) \in (0,T) \times H, \\ v(T,x) = g(x), \qquad x \in H. \end{cases}$$

pointwise.

We will use the following assumption.

HYPOTHESIS 2.33

(i) *The functions $\sigma(t,x,a)$, $b(t,x,a)$ and $l(t,x,a)$ are uniformly continuous in t on $[0,T]$, uniformly for $(x,a) \in B(0,R) \times \Lambda$ for every $R > 0$.*

(ii) *There exist $C, N > 0$ such that*

$$|l(t,x,a)| \le C(1+|x|)^N \quad (2.49)$$

for all $(t,x,a) \in [0,T] \times H \times \Lambda$.

(iii) *The function $v\colon [0,T] \times H \to \mathbb{R}$ is uniformly continuous on bounded subsets of $[0,T] \times H$, and its derivatives Dv, D^2v, v_t are uniformly continuous on bounded subsets of $(0,T) \times H$. Moreover $Dv\colon (0,T) \times H \to D(A^*)$ and A^*Dv is uniformly continuous on bounded subsets of $(0,T) \times H$. Finally there exist $C, N > 0$ such that*

$$|v(t,x)| + |Dv(t,x)| + |v_t(t,x)| + \|D^2v(t,x)\| + |A^*Dv(t,x)| \le C(1+|x|)^N \quad (2.50)$$

for all $(t,x) \in (0,T) \times H$.

THEOREM 2.34 *Let Hypotheses 1.119, 2.1 and 2.33 be satisfied, $v(T,x) = g(x)$ for every $x \in H$, and let the function v satisfy the DPP, i.e. for every $0 < t < \eta < T, x \in H$,*

$$v(t,x) = \inf_{a(\cdot)\in\overline{\mathcal{U}}_t} \mathbb{E}\left[\int_t^\eta l(s,X(s),a(s))ds + v(\eta, X(\eta))\right]. \quad (2.51)$$

Then v is a classical solution of (2.45).

PROOF. To prove that equation (2.45) is satisfied we show separately the two inequalities. We will not present all the details here as the proof follows the lines of the proof of Theorem 3.66 where it is showed that the value function is a viscosity solution applying Dynkin's formula to a suitable family of test functions.

[4] Sometimes it is called the *minimum value Hamiltonian.*

Part 1. (Supersolution inequality). We fix $(t,x) \in (0,T) \times H$. By (2.51), for every $\epsilon \in (0, T-t)$ we can choose a control $a^{\mu_\epsilon}(\cdot) \in \mathcal{U}_t^{\mu_\epsilon}$, such that,

$$v(t,x) + \epsilon^2 \geq \mathbb{E}^{\mu_\epsilon}\left[\int_t^{t+\epsilon} l(s, X^{\mu_\epsilon}(s), a^{\mu_\epsilon}(s))ds + v(t+\epsilon, X^{\mu_\epsilon}(t+\epsilon))\right],$$

where $X^{\mu_\epsilon}(\cdot)$ is the trajectory starting at (t,x) driven by $a^{\mu_\epsilon}(\cdot)$. Dividing the above by ϵ we have

$$\epsilon \geq \mathbb{E}^{\mu_\epsilon}\frac{v(t+\epsilon, X^{\mu_\epsilon}(t+\epsilon)) - v^{\mu_\epsilon}(t,x)}{\epsilon} + \frac{1}{\epsilon}\mathbb{E}^{\mu_\epsilon}\int_t^{t+\epsilon} l(s, X^{\mu_\epsilon}(s), a^{\mu_\epsilon}(s))ds$$

and, using Dynkin's formula from Proposition 1.156,

$$\begin{aligned}\epsilon \geq \frac{1}{\epsilon}\mathbb{E}^{\mu_\epsilon}\Bigg[\int_t^{t+\epsilon}\Bigg[&v_t(s, X^{\mu_\epsilon}(s)) + \langle A^* Dv(s, X^{\mu_\epsilon}(s)), X^{\mu_\epsilon}(s)\rangle \\ &+ \langle Dv(s, X^{\mu_\epsilon}(s)), b(s, X^{\mu_\epsilon}(s), a^{\mu_\epsilon}(s))\rangle \\ + \frac{1}{2}\mathrm{Tr}\left[\left(\sigma(s, X^{\mu_\epsilon}(s), a^{\mu_\epsilon}(s))Q^{1/2}\right)\right.&\left.\left(\sigma(s, X^{\mu_\epsilon}(s), a^{\mu_\epsilon}(s))Q^{1/2}\right)^* D^2 v(s, X^{\mu_\epsilon}(s))\right] \\ &+ l\left(s, X^{\mu_\epsilon}(s), a^{\mu_\epsilon}(s)\right)\Bigg]ds\Bigg] = \frac{1}{\epsilon}\mathbb{E}^{\mu_\epsilon}\int_t^{t+\epsilon}\Psi(s, X^{\mu_\epsilon}(s), a^{\mu_\epsilon}(s))ds, \end{aligned} \quad (2.52)$$

where

$$\begin{aligned}\Psi(s,y,a) := {}& v_t(s,y) + \langle A^* Dv(s,y), y\rangle + \langle Dv(s,y), b(s,y,a)\rangle \\ &+ \frac{1}{2}\mathrm{Tr}\left[\left(\sigma(s,y,a)Q^{1/2}\right)\left(\sigma(s,y,a)Q^{1/2}\right)^* D^2 v(s,y)\right] + l(s,y,a).\end{aligned}$$

By our assumptions we have, for some $h > 0$ and modulus ρ, depending on t, x,

$$|\Psi(s,y,a) - \Psi(t,x,a)| \leq \rho(|s-t| + |y-x|) \quad \text{for all } (s,y) \in [t,t+h] \times B_1(x), a \in \Lambda \quad (2.53)$$

and, for some C and $M \geq 0$,

$$|\Psi(s, X^{\mu_\epsilon}(s), a^{\mu_\epsilon}(s))| \leq C(1 + |X^{\mu_\epsilon}(s)|^M). \quad (2.54)$$

Moreover it follows from (1.39) that there is $r_\epsilon > 0$, independent of a^{μ_ϵ}, such that $r_\epsilon \to 0$ as $\epsilon \to 0$, and denoting

$$\Omega_1^\epsilon = \{\omega \in \Omega^{\mu_\epsilon} : \sup_{s \in [t,t+\epsilon]} |X^{\mu_\epsilon}(s) - x| \leq r_\epsilon\},$$

then

$$\mathbb{P}^{\mu_\epsilon}(\Omega_1^\epsilon) \geq \gamma(\epsilon) \to 1 \quad \text{as } \epsilon \to 0. \quad (2.55)$$

Thus, using (1.38), (2.52), (2.53), (2.54), and (2.55) we obtain (see the proof of Theorem 3.66 for more details) that there exists a modulus $\rho_1(\epsilon)$, depending on t and x, such that

$$\begin{aligned}\rho_1(\epsilon) &\geq \frac{1}{\epsilon}\mathbb{E}^{\mu_\epsilon}\int_t^{t+\epsilon}\Psi(t,x,a^{\mu_\epsilon}(s))ds \\ &\geq v_t(t,x) + \langle A^* Dv(t,x), x\rangle + \frac{1}{\epsilon}\mathbb{E}^{\mu_\epsilon}\Bigg[\int_t^{t+\epsilon}\inf_{a\in\Lambda}\Bigg\{\frac{1}{2}\mathrm{Tr}\Big[\left(\sigma(t,x,a)Q^{1/2}\right) \\ &\qquad \times \left(\sigma(t,x,a)Q^{1/2}\right)^* D^2 v(t,x)\Big] + \langle Dv(t,x), b(t,x,a)\rangle + l(t,x,a)\Bigg\}ds\Bigg] \\ = v_t(t,x) &+ \langle A^* Dv(t,x), x\rangle + \inf_{a\in\Lambda}\Bigg\{\frac{1}{2}\mathrm{Tr}\Big[\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^* D^2 v(t,x)\Big] \\ &\qquad + \langle Dv(t,x), b(t,x,a)\rangle + l(t,x,a)\Bigg\}. \end{aligned} \quad (2.56)$$

The inequality follows letting $\epsilon \to 0$.

Part 2. (Subsolution inequality). Choose $a \in \Lambda$ and consider the constant control $\bar{a}(\cdot) \equiv a \in \Lambda$ for some generalized reference probability space μ. Denote by $X(s)$ the trajectory starting from (t,x) driven by the control $\bar{a}(\cdot)$. From (2.51) we have for $\epsilon \in (0, T-t)$

$$v(t,x) \le \mathbb{E}^\mu \left[\int_t^{t+\epsilon} l\left(s, X\left(s\right), a\right) ds + v\left(t+\epsilon, X\left(t+\epsilon\right)\right) \right].$$

Using again Dynkin's formula from Proposition 1.156, we thus obtain

$$\begin{aligned} 0 \le \frac{\mathbb{E}^\mu [v(t+\epsilon, X(t+\epsilon)) - v(t,x)]}{\epsilon} + \frac{1}{\epsilon}\mathbb{E}^\mu \int_t^{t+\epsilon} l\left(s, X\left(s\right), a\right) ds \\ = \frac{1}{\epsilon}\mathbb{E}^\mu \Bigg[\int_t^{t+h} \Big[v_t(s, X(s)) + \langle A^* Dv(s, X(s)), X(s) \rangle \\ + \langle Dv(s, X(s)), b(s, X(s), a) \rangle \\ + \frac{1}{2}\mathrm{Tr}\left[\left(\sigma(s, X(s), a) Q^{1/2}\right) \left(\sigma(s, X(s), a) Q^{1/2}\right)^* D^2 v(s, X(s)) \right] \\ + l\left(s, X\left(s\right), a\right) \Big] ds \Bigg]. \end{aligned} \tag{2.57}$$

We can now pass to the limit as $\epsilon \to 0$ above like in Part 1 to obtain that for every $a \in \Lambda$

$$\begin{aligned} 0 \le v_t(t,x) + \langle A^* Dv(t,x), x \rangle + \frac{1}{2}\mathrm{Tr}\left[\left(\sigma(t,x,a) Q^{\frac{1}{2}}\right) \left(\sigma(t,x,a) Q^{\frac{1}{2}}\right)^* D^2 v(t,x) \right] \\ + \langle Dv(t,x), b(t,x,a) \rangle + l(t,x,a),\ t \in (0,T),\ x \in H. \end{aligned} \tag{2.58}$$

The inequality follows by taking the infimum over $a \in \Lambda$ above. □

We now show how to use the HJB equation to characterize optimal controls. First we prove the so-called verification theorem.

THEOREM 2.35 (Smooth Verification Theorem, Sufficient Condition) *Let $v\colon [0,T] \times H \to \mathbb{R}$ be a classical solution of (2.45) as defined in Definition 2.32. Let Hypotheses 1.119, 2.1 and 2.33-(ii)(iii) be satisfied. Then:*

(i) *We have*

$$v(t,x) \le \overline{V}(t,x) \qquad \textit{for all } (t,x) \in [0,T] \times H. \tag{2.59}$$

(ii) *Let $(a^*(\cdot), X^*(\cdot))$ be an admissible pair at (t,x) such that*

$$a^*(s) \in \arg\min_{a \in \Lambda} F_{CV}(s, X^*(s), Dv(s, X^*(s)), D^2 v(s, X^*(s)), a), \tag{2.60}$$

for almost every $s \in [t,T]$ and $\mathbb{P}$-almost surely. Then the pair $(a^*(\cdot), X^*(\cdot))$ is optimal at (t,x), and $v(t,x) = \overline{V}(t,x)$.

PROOF. We prove first the following identity[5]: for every $a(\cdot) \in \overline{\mathcal{U}}_t$

$$\begin{aligned} v(t,x) = J(t,x; a(\cdot)) \\ - \mathbb{E}\int_t^T \Big[F_{CV}\left(r, X(r), Dv(r, X(r)), D^2 v(r, X(r)), a(r)\right) \\ - F\left(r, X(r), Dv(r, X(r)), D^2 v(r, X(r))\right) \Big] dr. \end{aligned} \tag{2.61}$$

[5]This is often called the *fundamental identity* for the optimal control problem.

Indeed, consider $a(\cdot) \in \overline{\mathcal{U}}_t$ and the corresponding trajectory $X(\cdot)$ starting at x at time t. We apply Proposition 1.156, to the process $v(s, X(s))$, $s \in [t, T]$ obtaining

$$\begin{aligned}
\mathbb{E}v(T, X(T)) &= v(t,x) + \mathbb{E}\int_t^T v_t(r, X(r))dr \\
&+ \mathbb{E}\int_t^T \langle A^* Dv(r, X(r)), X(r)\rangle\, dr + \mathbb{E}\int_t^T \langle Dv(r, X(r)), b(r, X(r), a(r))\rangle\, dr \\
&+ \frac{1}{2}\mathbb{E}\int_t^T \mathrm{Tr}\left[\left(\sigma(r, X(r), a(r))Q^{1/2}\right)\left(\sigma(r, X(r), a(r))Q^{1/2}\right)^* D^2 v(r, X(r))\right] dr.
\end{aligned} \tag{2.62}$$

We now use that $v(T, \cdot) = g$, rearrange the terms, and we add and subtract $\mathbb{E}\int_t^T l(r, X(r), a(r))dr$ obtaining, by the definition of the current value Hamiltonian F_{CV} in (2.46),

$$\begin{aligned}
v(t,x) = \mathbb{E}g(X(T)) &+ \mathbb{E}\int_t^T l(r, X(r), a(r))dr \\
&- \mathbb{E}\int_t^T v_t(r, X(r)) + \langle A^* Dv(r, X(r)), X(r)\rangle\, dr \\
&- \mathbb{E}\int_t^T F_{CV}(r, X(r), Dv(r, X(r)), D^2 v(r, X(r)), a(r))dr.
\end{aligned} \tag{2.63}$$

Equality (2.61) is now a consequence of the definition of the functional J and the fact that v is a classical solution of the HJB equation (2.45).

Therefore (i) follows by observing that, by definition, $F_{CV} - F \geq 0$ everywhere, and by taking the infimum over $a(\cdot) \in \overline{\mathcal{U}}_t$ in the right hand side of (2.61).

Regarding (ii), let $(a^*(\cdot), X^*(\cdot))$ be an admissible pair at (t,x) satisfying (2.60) for almost every $s \in [t, T]$ and $\mathbb{P}$-almost surely. We then have

$$\begin{aligned}
\mathbb{E}\int_t^T \Big[F_{CV}&\left(r, X(r), Dv(r, X(r)), D^2 v(r, X(r)), a(r)\right) \\
&- F\left(r, X(r), Dv(r, X(r)), D^2 v(r, X(r))\right)\Big] dr = 0.
\end{aligned} \tag{2.64}$$

Thus, by (2.61), we get

$$v(t,x) = J(t, x, a^*(\cdot)), \tag{2.65}$$

which, together with (i), implies that $(a^*(\cdot), X^*(\cdot)$ is optimal at (t,x) and $v(t,x) = V(t,x)$. □

Note that part (i) of above theorem remains true if v is any classical subsolution of the HJB equation (2.45)[6] with the required regularity.

If we know from the beginning that the solution v in Theorem 2.35 is the value function $\overline{V}$ then (2.60) becomes also a necessary condition for optimality.

Corollary 2.36 (Smooth Verification Theorem, Necessary Condition) *Let the assumptions of Theorem 2.35 hold for $v = \overline{V}$. Let $(a^*(\cdot), X^*(\cdot))$ be an optimal pair at (t,x). Then we must have*

$$a^*(s) \in \arg\min_{a \in \Lambda} F_{CV}(s, X^*(s), D\overline{V}(s, X^*(s)), D^2\overline{V}(s, X^*(s)), a), \tag{2.66}$$

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

[6]in the sense that $v(T,x) \leq g(x)$ and it satisfies (2.45) with the inequality $\geq$.

PROOF. Now the function $v = \overline{V}$ satisfies (2.61). Since $(a^*(\cdot), X^*(\cdot))$ is an optimal pair at (t, x), we have $\overline{V}(t, x) = J(t, x; a^*(\cdot))$. Therefore, (2.61) for $\overline{V}$ implies that the integrand of the last term of (2.61) is zero $dt \otimes \mathbb{P}$-a.e. and the claim follows. □

Assume now that we have a classical solution v of the HJB equation (2.45). Define the multivalued function

$$\begin{cases} \Phi\colon (0,T) \times H \to \mathcal{P}(\Lambda) \\ \Phi\colon (t,x) \mapsto \arg\min_{a \in \Lambda} F_{CV}(t, x, Dv(t,x), D^2v(t,x), a). \end{cases} \tag{2.67}$$

The *Closed Loop Equation* (CLE) associated with our problem and v is then formally defined as

$$\begin{cases} dX(s) \in AX(s)dt + b(s, X(s), \Phi(s, X(s)))ds + \sigma(s, X(s), \Phi(s, X(s)))dW_Q(s) \\ X(t) = x. \end{cases} \tag{2.68}$$

If we can find a solution (in a suitable sense) $X_\Phi(\cdot)$ of such stochastic differential inclusion, we expect that, if $a_\Phi(\cdot)$ is a suitable measurable selection of $\Phi(\cdot, X(\cdot))$, then the couple $(a_\Phi(\cdot), X_\Phi(\cdot))$ is optimal at (t, x). This is indeed the statement of the next corollary.

COROLLARY 2.37 (Optimal Feedback Controls) *Let the assumptions of Theorem 2.35 hold. Assume moreover that the feedback map Φ defined in (2.67) admits a measurable selection $\phi : (0,T) \times H \to \Lambda$ such that the Closed Loop Equation*

$$\begin{cases} dX(s) = AX(s)ds + b(s, X(s), \phi(s, X(s)))ds + \sigma(s, X(s), \phi(s, X(s)))dW_Q(s) \\ X(t) = x, \end{cases} \tag{2.69}$$

has a mild weak solution (see Definition 1.115) $X_\phi(\cdot)$ in some generalized reference probability space μ satisfying Hypothesis 2.1-(iv). Then the couple $(a_\phi(\cdot), X_\phi(\cdot))$, where the control $a_\phi(\cdot)$ is defined by the feedback law $a_\phi(s) = \phi(s, X_\phi(s))$, is admissible and it is optimal at (t, x).

PROOF. By construction the couple $(a_\phi(\cdot), X_\phi(\cdot))$ satisfies (2.60). Then, by Theorem 2.35-(ii) we obtain that such couple is optimal. Observe that since the assumptions of Theorem 2.35 are satisfied, $X_\phi(\cdot)$ is the unique mild solution (in the strong probabilistic sense) of the state equation associated to the control $a_\phi(\cdot)$ in generalized reference probability space μ. □

In the above corollary we assumed that the closed loop equation has a weak mild solution to obtain existence of an optimal feedback control in the weak formulation. If we consider the control problem (2.43)-(2.44) in the strong formulation with the value function defined by (2.5), all the results above remain true except for Corollary 2.37. Indeed the weak mild solution $X_\phi(\cdot)$, and so the control $a_\phi(\cdot)$ may be defined in a different reference probability space than the starting one. To get existence of an optimal feedback control in the strong formulation one needs to have existence of solutions of the closed loop equation (2.69) in the strong probabilistic sense (i.e. in the mild or strong sense of Definitions 1.112 and 1.113).

To conclude let us reiterate the three step process to do the so-called synthesis of optimal control for the problem (2.43)-(2.44) once we have a classical solution v of our HJB equation.

(1) Define the function Φ as in (2.67) .
(2) Look for a solution $X^*(\cdot)$ of the closed loop equation (2.68).
(3) Define $a^*(s) := \phi(s, X^*(s))$. It is optimal thanks to Theorem 2.35.

Of course, given an optimal control problem like (2.43)-(2.44), it may not be possible to perform the above steps as they are. However, even if the HJB equation does not have a classical solution, we may still be able to synthesize optimal controls. This will be explained in later chapters for some special cases. The general synthesis of optimal controls is still a largely open problem.

As it was explained in Remark 2.6, the extended weak formulation may be more suitable for Corollary 2.37 (and also Corollary 2.43 in the infinite horizon case). This is done in Chapter 6 (see also Chapters 4 and 5).

2.5.2. Infinite horizon problem: Elliptic HJB equation. Consider the optimal control problem of minimizing the infinite horizon functional (2.38) for the system governed by the state equation (2.41) with $t = 0$ and the set of admissible controls equal to $\overline{\mathcal{U}}_0$. For simplicity we will assume that $c(\cdot) \equiv \lambda > 0$. This is a typical infinite horizon problem with constant discounting.

The current value Hamiltonian is now defined by

$$F_{CV}(x,p,S,a) := \frac{1}{2}\mathrm{Tr}\left[\left(\sigma(x,a)Q^{\frac{1}{2}}\right)\left(\sigma(x,a)Q^{\frac{1}{2}}\right)^* S\right] + \langle p, b(x,a)\rangle + l(x,a), \quad (2.70)$$

the Hamiltonian is given by

$$F(x,p,S) := \inf_{a\in\Lambda}\left\{\frac{1}{2}\mathrm{Tr}\left[\left(\sigma(x,a)Q^{\frac{1}{2}}\right)\left(\sigma(x,a)Q^{\frac{1}{2}}\right)^* S\right] + \langle p, b(x,a)\rangle + l(x,a)\right\}, \quad (2.71)$$

and the HJB equation associated to our infinite horizon optimal control problem is

$$\lambda v - \langle Dv, Ax\rangle - F(x, Dv, D^2 v) = 0 \quad (2.72)$$

for the unknown function $v : H \to \mathbb{R}$.

We present here the infinite horizon versions of the results of the previous subsection, adapted to the infinite horizon case.

DEFINITION 2.38 (Classical solution, elliptic case) *A function $v\colon H \to \mathbb{R}$ is a* classical solution *of (2.72) if $v \in C^2(H)$, $A^*Dv \in C(H;H)$ and v satisfies*

$$\lambda v - \langle A^*Dv, x\rangle - F(x, Dv, D^2 v) = 0.$$

pointwise.

Similarly to the previous section we will need the following assumption.

HYPOTHESIS 2.39

(i) *There exist $C, N > 0$ such that*

$$|l(x,a)| \le C(1+|x|)^N \quad (2.73)$$

for all $(x,a) \in H \times \Lambda$.

(ii) *The function $v\colon H \to \mathbb{R}$ and its derivatives Dv, D^2v, v_t are uniformly continuous on bounded subsets of H. Moreover $Dv\colon H \to D(A^*)$ and A^*Dv are uniformly continuous on bounded subsets of H.*

$$|v(x)| + |Dv(x)| + \|D^2 v(x)\| + |A^*Dv(x)| \le C(1+|x|)^N \quad (2.74)$$

for all $x \in H$.

THEOREM 2.40 *Let Hypotheses 1.119, 2.1 for $T = +\infty$, and 2.39 be satisfied. Assume that the function v satisfy the DPP, i.e. for every $0 < \eta < +\infty, x \in H$,*

$$v(x) = \inf_{a(\cdot)\in\overline{\mathcal{U}}_0} \mathbb{E}\left[\int_0^\eta e^{-\lambda s} l(s, X(s), a(s))ds + e^{-\lambda\eta} v(X(\eta))\right]. \quad (2.75)$$

Then, the function v is a classical solution of (2.72).

PROOF. The proof follows the lines of the proof of Theorem 2.34. □

We now pass to the verification theorem, the necessary conditions and the closed loop equation. In these results we may encounter integrability problems. To avoid technical complications, here we consider the case where the discount factor λ is sufficiently big.

THEOREM 2.41 (Smooth Verification, Sufficient Condition, Infinite Horizon) *Let $v\colon H \to \mathbb{R}$ be a classical solution of (2.72) as defined in Definition 2.38. Let Hypotheses 1.119, 2.1 for $T = +\infty$, and 2.39 be satisfied, and let $\lambda > \overline{\lambda} = 2(N+2)(C + (N+1)C^2)$, where C is the constant from (1.34) and (1.35) (see Proposition 3.24 for $m = N+2$). Then:*

(i) *For all $x \in H$*

$$v(x) \le \overline{V}(x) \quad \text{for all } x \in H. \tag{2.76}$$

(ii) *Let $(a^*(\cdot), X^*(\cdot))$ be an admissible pair at x such that*

$$a^*(s) \in \arg\min_{a\in\Lambda} F_{CV}(X^*(s), Dv(s, X^*(s)), D^2v(s, X^*(s)), a), \tag{2.77}$$

for almost every $s \in [0, +\infty)$ and $\mathbb{P}$-almost surely. Then the pair $(a^*(\cdot), X^*(\cdot))$ is optimal at x, and $v(x) = \overline{V}(x)$.

PROOF. The proof is similar to that of Theorem 2.35 except for the fact that we now have to take the limit as $T \to +\infty$, in (2.61). Indeed, arguing as in the proof of Theorem 2.35, we obtain that for every $a(\cdot) \in \overline{\mathcal{U}}_0$, and every $T > 0$,

$$\begin{aligned} v(x) = e^{-\lambda T}\mathbb{E}v(X(T)) &+ \int_0^T e^{-\lambda r} l(X(r), a(r))dr \\ &- \mathbb{E}\int_0^T e^{-\lambda r}\Big[F_{CV}\left(X(r), Dv(r, X(r)), D^2v(r, X(r)); a(r)\right) \\ &\qquad - F\left(X(r), Dv(r, X(r)), D^2v(r, X(r))\right)\Big]dr. \end{aligned} \tag{2.78}$$

The condition $\lambda > \overline{\lambda}$ guarantees, due to estimate (3.34), that we can pass to the limit as $T \to +\infty$ above, obtaining the *fundamental identity*:

$$\begin{aligned} v(x) = &\int_0^{+\infty} e^{-\lambda r} l(X(r), a(r))dr \\ &- \mathbb{E}\int_0^{+\infty} e^{-\lambda r}\Big[F_{CV}\left(X(r), Dv(r, X(r)), D^2v(r, X(r)); a(r)\right) \\ &\qquad - F\left(X(r), Dv(r, X(r)), D^2v(r, X(r))\right)\Big]dr. \end{aligned} \tag{2.79}$$

The claims now follow as in the proof of Theorem 2.35. □

COROLLARY 2.42 (Smooth Verification, Necessary Cond., Infinite Horizon) *Let the assumptions of Theorem 2.41 hold for $v = \overline{V}$. Let $(a^*(\cdot), X^*(\cdot))$ be an optimal pair at x. Then we must have*

$$a^*(s) \in \arg\min_{a\in\Lambda} F_{CV}(X^*(s), D\overline{V}(s, X^*(s)), D^2\overline{V}(s, X^*(s)), a), \tag{2.80}$$

for almost every $s \in [0, +\infty)$ and $\mathbb{P}$-almost surely.

PROOF. The same as Corollary 2.36 using (2.79). □

As in the finite horizon case we assume that we have a classical solution v of the HJB equation (2.72). We define the multivalued function

$$\begin{cases} \Phi\colon H \to \mathcal{P}(\Lambda) \\ \Phi\colon x \mapsto \arg\min_{a\in\Lambda} F_{CV}(x, Dv(t,x), D^2v(t,x); a) \end{cases} \tag{2.81}$$

The *Closed Loop Equation* (CLE) associated with our problem and v is then formally defined as

$$\begin{cases} dX(s) \in AX(s)dt + b(X(s), \Phi(X(s)))ds + \sigma(X(s), \Phi(X(s)))dW_Q(s) \\ X(0) = x. \end{cases} \tag{2.82}$$

Again, if a solution $X_\Phi(\cdot)$ of this stochastic differential inclusion can be found, and we can find $a_\Phi(\cdot)$, a suitable measurable selection of $\Phi(\cdot, X(\cdot))$, we would expect the couple $(a_\Phi(\cdot), X_\Phi(\cdot))$ to be optimal at x.

COROLLARY 2.43 (Optimal Feedback Controls, Infinite Horizon) *Let the assumptions of Theorem 2.41 hold. Assume moreover that the feedback map Φ defined in (2.81) admits a measurable selection $\phi : H \to \Lambda$ such that the Closed Loop Equation*

$$\begin{cases} dX(s) = AX(s)dt + b(X(s), \phi(X(s)))ds + \sigma(X(s), \phi(X(s)))dW(s) \\ X(0) = x, \end{cases} \tag{2.83}$$

has a mild weak solution (see Definition 1.115) $X_\phi(\cdot)$ in some generalized reference probability space satisfying Hypothesis 2.1-(iv). Then the couple $(a_\phi(\cdot), X_\phi(\cdot))$, where the control $a_\phi(\cdot)$ is defined by the feedback law $a_\phi(s) = \phi(s, X_\phi(s))$, is admissible and it is optimal at (t, x).

PROOF. The proof is the same as that of Corollary 2.37. □

The optimal synthesis is performed in the same way as in the finite horizon case.

REMARK 2.44 In Section 2.4 and in this subsection we have considered infinite horizon problems satisfying Hypothesis 2.29 which substantially means that the data b, σ, l and c must be time independent. We did this restriction partly for simplicity of exposition and partly because such cases are very common in applied models. However with little effort, it is possible to apply the dynamic programming approach to infinite horizon problems with "non-autonomous" data (so without Hypothesis 2.29). In such cases the value function would be a function of (t, x), the DPP would have the form (2.25), and the HJB equation would be a parabolic equation on $(0, +\infty) \times H$ like (2.48) but with a zeroth order term coming from the discount factor, and without a terminal condition:

$$v_t - \lambda v + \langle Dv, Ax \rangle + F(t, x, Dv, D^2 v) = 0. \tag{2.84}$$

Such problems are more difficult but are still interesting e.g. in some financial applications (see [107, 174, 218]). ■

2.6. Some motivating examples

In this section we describe several examples that motivate the study of stochastic optimal control problems in infinite dimensions. Our goal here is to show how various applied problems, arising in different areas of science and engineering, are naturally modeled within the framework of the infinite dimensional stochastic analysis. The first five examples are concerned with the control of various kinds of stochastic PDE, while the last deals with the control of stochastic delay equations. Despite similarities, these examples are very different from each other and it is difficult to find a general theory that includes all of them. This is an unpleasant feature of the infinite dimensional optimal control which will force us to apply different approaches and adaptations of the main general theory.

We have chosen our examples, among many others, since they are representative of interesting applied models and since they motivate the four different approaches described in this book (viscosity solutions, strong solutions, L^2_μ solutions, solutions via BSDE). In all examples the criteria to maximize/minimize are given as the expectation of a Bolza-type functionals (in finite or infinite horizon cases). We leave aside other types of criteria (mean-variance, risk sensitive, ergodic, etc) for which we will refer the reader to the existing literature. For each example we provide a short motivation, the main mathematical framework (*state equation, objective functional and constraints*), and show how to translate the problem into the abstract framework of infinite dimensional stochastic optimal control introduced before. Finally we discuss the issue of the Dynamic Programming Principle, and we present the associated HJB equations with references to further material in the book and in the literature. For purposes of the DPP we always take the optimal control problems in the weak formulation of Section 2.2, however the control problems can be studied with other formulations.

We remark that the existing theory is far from providing a satisfactory treatment of all problems: many challenging questions remain open and call for further research. An important issue in this respect is the one of constraints: to obtain more realistic control problems it is often necessary to impose suitable constraints on the state and on the control variables. Such constraints strongly depend on the particular problems. Since the addition of state constraints makes the dynamic programming approach much harder and not much is known at the present stage, we will mention how state constraints arise in specific problems, however we will not deal with state constraint problems.

Finally we mention a few things about the notation.

- To be consistent with the general setting introduced before, the initial time is always $t \geq 0$.
- In all examples we start first with the finite dimensional notation. Thus the state equation is first written as a PDE (or a functional equation) in finite dimensions with solutions defined informally, then in a subsequent section the state equation is rewritten as an evolution equation in an infinite dimensional space. To distinguish the two cases we write y to denote the finite dimensional state and α for the finite dimensional control, while the infinite dimensional state and control will be denoted, as before, by X and a, respectively.
- Following the usual convention, even if all variables depend on the "scenario" ω, we will always drop such dependence unless needed in the context.

WARNING: The HJB equations that appear in the examples in this section are formal and are all written using the convention adapted from the natural way the equation was written in (2.45)-(2.47) in Section 2.5.1. This form is preferable from the PDE point of view as all the terms (when they are defined) use only the reference Hilbert space H of the independent variable x. We want to focus on the examples and leave the details of how the equations are interpreted and solved to later chapters. In some cases the Hamiltonians appearing here are always well defined and will not need any interpretation. In some cases some terms may not make sense the way they are written here and they will need special interpretation which would take too long to explain here. This is especially true of equations discussed in Sections 2.6.2 and 2.6.3. The formal versions are enough to point out the main difficulties posed by the equations, however the reader should be careful as the equations may not be what they appear.

2.6.1. Stochastic controlled heat equation: distributed control. Our first example concerns the problem of controlling a nonlinear stochastic heat equation in a given space region $\mathcal{O} \subset \mathbb{R}^N$. This is a very popular example and here, under "reasonable" assumptions, the theory applies quite well giving rise to results about the HJB equation and the synthesis of optimal controls.

The optimal control problems of this type arise in various applied contexts. We recall some of them.

- Optimal control of the heat distribution of a given body (the region $\mathcal{O}$). The deterministic case is described for instance in the monograph [312] (pages 3-5). The presence of the stochastic additive term in the equation can be justified by the presence of (small) random perturbations in the system. (It may be of interest to see what happens when such term goes to zero, see e.g. [63]).
- Optimal control of stochastic reaction diffusion equations, where the white noise term describes the internal fluctuation of the system due to its many-particle nature (see e.g. page 8 of [130], and [10] for the model without control and [79, 81] for the model with control).
- Optimal control of the motion of an elastic string in a random viscous environment (see e.g. page 4 of [130] and [214] for the model without control).
- Optimal control of the stochastic cable equation (arising also in neurosciences, see page 9 of [130] and [444] for the stochastic model, and e.g [53] for the optimal control problem in the deterministic case).
- Optimal advertising problems arising in economics (see [331]).

2.6.1.1. *Setting of the problem.* We are given an open, connected and bounded set $\mathcal{O} \subset \mathbb{R}^N$ with C^1 boundary $\partial\mathcal{O} \subset \mathbb{R}^N$ and a reference probability space $(\Omega, \mathscr{F}, (\mathscr{F}_s)_{s\in[t,T]}, \mathbb{P}, W_Q)$. We consider a controlled dynamical system driven by the following stochastic PDE in the time interval $[t, T]$, for $0 \le t \le T < +\infty$

$$\begin{cases} dy(s,\xi) = [\Delta_\xi y(s,\xi) + f(y(s,\xi)) + \alpha(s,\xi)]\, ds + dW_Q(s)(\xi), \quad s \in (t,T],\ \xi \in \mathcal{O} \\ y(s,\xi) = 0 \quad (s,\xi) \in (t,T] \times \partial\mathcal{O} \\ y(t,\xi) = x(\xi) \in L^2(\mathcal{O}),\ \xi \in \mathcal{O}, \end{cases} \tag{2.85}$$

where:

- the function $y : [t,T] \times \mathcal{O} \times \Omega \to \mathbb{R}$, $(s,\xi,\omega) \mapsto y(s,\xi,\omega)$ is a stochastic process that describes e.g. the evolution of the temperature distribution and is the *state trajectory* of the system;
- the function $\alpha : [t,T] \times \mathcal{O} \times \Omega \to \mathbb{R}$, $(s,\xi,\omega) \mapsto \alpha(s,\xi,\omega)$ is a stochastic process giving e.g. the dynamics of the external source of heat acting at every interior point of $\mathcal{O}$ and is the *control strategy* of the system.

We will omit the variable ω writing simply $y(s,\xi)$ and $\alpha(s,\xi)$. Moreover:

- Δ_ξ is the Laplace operator. We consider the Dirichlet boundary conditions, however the problem can be studied similarly with the Neumann boundary conditions. Conditions of mixed type are also possible, see on this e.g. Chapters 3 and 5 of [323];
- $f \in C^0(\mathbb{R})$ is a nonlinear function of the state (which may represent a "reaction" term);

- W_Q is a Q-Wiener process with $Q \in \mathcal{L}^+(L^2(\mathcal{O}))$ and $(\mathscr{F}_s)_{s\in[t,T]}$ is the augmented filtration generated by W_Q[7] (see Remark 2.10 if $\mathrm{Tr}(Q) = +\infty$);
- $x(\cdot) \in L^2(\mathcal{O})$ is the initial state (e.g. temperature distribution) in the region $\mathcal{O}$.

The solution[8] of (2.85) will be denoted by $y^{\alpha,t,x}$ to underline the dependence of the state y on the control α and on the initial data t, x. Having in mind the control of the temperature distribution, a reasonable objective of the controller here can be the one of getting such distribution $y^{\alpha,t,x}$ to be close to a required one $\bar{y}$ (for each time $s \in [t,T]$ or only at the final time T) while spending the fewest amount of energy doing this. In such case a reasonable cost functional may be of the form (for suitable constants $c_0, c_1, c_2 \in \mathbb{R}$)

$$I_1(t,x;\alpha) = \mathbb{E}\Bigg\{ \int_t^T \int_{\mathcal{O}} \left[c_0|y^{\alpha,t,x}(s,\xi) - \bar{y}(s,\xi)|^2 + c_1|\alpha(s,\xi)|^2\right] d\xi ds + \int_{\mathcal{O}} c_2|y^{\alpha,t,x}(T,\xi) - \bar{y}(T,\xi)|^2 d\xi \Bigg\}, \tag{2.86}$$

and the objective would be to minimize the functional I_1 above over all control strategies α, progressively measurable with respect to the filtration generated by W_Q, and satisfying suitable constraints and integrability conditions (e.g. such that the state equation and above integrals make sense). More generally one could consider a cost functional

$$I_2(t,x;\alpha) = \mathbb{E}\Bigg\{ \int_t^T \int_{\mathcal{O}} \beta(y^{\alpha,t,x}(s,\xi), \alpha(s,\xi)) d\xi ds + \int_{\mathcal{O}} \gamma(y^{\alpha,t,x}(T,\xi)) d\xi \Bigg\} \tag{2.87}$$

where $\beta \in C^0(\mathbb{R}^2)$ an $\gamma \in C^0(\mathbb{R})$ are given functions depending on the objective of the controller.

Finally, the constraints: If the state is the absolute temperature it is natural to require the positivity of $y^{\alpha,t,x}$; moreover it is reasonable to assume bounds on the control strategies depending on the physical device used to control the system (e.g. $\alpha(s,\xi) \in [m,M]$ for given $m < M$). The constraints depend on a particular problem.

2.6.1.2. *The infinite dimensional setting and the HJB equation.* Take $H = \Xi = L^2(\mathcal{O})$ and let Λ be a closed, bounded subset of $L^2(\mathcal{O})$. For instance, if $\alpha(s,\xi) \in [m,M]$ for every $(s,\xi) \in [t,T] \times \mathcal{O}$ then we would take $\Lambda = \big\{ f \in L^2(\mathcal{O}) : \ f(\xi) \in [m,M], \ \forall \xi \in \mathcal{O} \big\}$. Consider the Laplace operator with Dirichlet boundary conditions defined as (see e.g. [432], Section 5.2 page 180):

$$\begin{cases} D(A) = H^2(\mathcal{O}) \cap H^1_0(\mathcal{O}), \\ Ax = \Delta x, \ \text{for} \ \ x \in D(A) \end{cases} \tag{2.88}$$

that generates an analytic semigroup of compact operators $\{e^{tA}\}_{t\geq 0}$. Moreover define the Nemytskii operator $b : H \to H$ as

$$b(x)(\xi) = f(x(\xi)). \tag{2.89}$$

[7]Indeed also stochastic PDE's with more general types of noise can be treated, see e.g. [380] but this is beyond the scope of this book.

[8]For the concept of solution and the assumptions on the data f, Q and on the control strategy α that guarantee the existence and uniqueness of it see next section.

Defining $X(s) := y(s,\cdot) \in L^2(\mathcal{O})$ and $a(s) := \alpha(s,\cdot) \in \Lambda$, the state equation (2.85) can be rewritten as an SDE in H as follows

$$\begin{cases} dX(s) = [AX(s) + b(X(s)) + a(s)]\, ds + dW_Q(s) \\ X(t) = x \in H. \end{cases} \tag{2.90}$$

From Proposition 1.141 we know that, when f is Lipschitz[9] (and so is b) and $Q_r := \int_0^T e^{\tau A} Q e^{\tau A^*} d\tau$ is trace class for all $r > 0$, the above equation admits a unique mild solution denoted by $X(s;t,x,a)$ (or simply $X(s)$ when no confusion is possible). If also (1.67) holds, then such solution has continuous trajectories.[10] If $\mathrm{Tr}(Q) < +\infty$ then, thanks to Proposition 2.16, Hypotheses 2.11 and 2.12 are satisfied. If $\mathrm{Tr}(Q) = +\infty$ the claim is still true as outlined in Remark 2.17.

Defining $l \colon H \times \Lambda \to \mathbb{R}$

$$l(x,a) = \int_{\mathcal{O}} \beta(x(\xi), a(\xi)) d\xi,$$

and $g \colon H \to \mathbb{R}$ as

$$g(x) = \int_{\mathcal{O}} \gamma(x(\xi)) d\xi,$$

the functional I_2 of (2.87) can be rewritten in the Hilbert space setting as

$$J_2(t,x;a(\cdot)) = \mathbb{E}\left\{ \int_t^T l(X(s),a(s))ds + g(x(T)) \right\}. \tag{2.91}$$

Suppose that β and γ satisfy the right conditions so that Hypothesis 2.23 holds. This is done for instance in Section 3.6, Propositions 3.61 and 3.62. Then, all the assumptions of Theorem 2.24 are satisfied and hence the dynamic programming principle holds.

The Hamilton-Jacobi-Bellman equation associated with problem (2.90)-(2.91) is the following:

$$\begin{cases} v_t + \dfrac{1}{2}\,\mathrm{Tr}\,[QD^2 v] + \langle Ax + b(x), Dv\rangle + \inf_{a\in\Lambda} \{\langle a, Dv\rangle + l(x,a)\} = 0, \\ v(T,x) = g(x). \end{cases} \tag{2.92}$$

This problem falls into the classes studied for instance in [286, 421, 422][11] by the viscosity solution approach, in [23, 63, 64, 79, 136, 231, 232, 333, 335, 339][12] by the strong solutions approach, in [94, 223] by the L^2 approach, in [211, 337][13] by the BSDE approach. The theory of such HJB equations is described in Chapters 3-6. We also refer to Section 4.8.3 for a specific example and to Section 4.10.1 where an explicit solution of the HJB equation (2.92) is found in the case of a quadratic Hamiltonian.

---

[9] In the case studied in [79, 81], b is not Lipschitz, see on this Section 4.9

[10] Such assumptions are true e.g. when $N = 1$ and Q is the identity, or when $N = 2$ and $Q = A^{-\alpha}$ for some $\alpha > 0$.

[11] These papers treat the fully nonlinear case.

[12] [79] deals with nonlipschitz b, [333, 335, 339] also treat the case of multiplicative noise, and [335] treats a Banach space case.

[13] These papers also treat the case of multiplicative noise and [337] considers it in a Banach space.

2.6.1.3. *The infinite horizon case.* For the infinite horizon case we again rewrite the state equation (2.85) as (2.90), starting at time 0 at the point $x \in H$. The cost functional

$$I_3(x;a) = \mathbb{E}\left\{ \int_0^{+\infty} e^{-\rho s} \int_{\mathcal{O}} \beta(y^{\alpha,0,x}(s,\xi), \alpha(s,\xi)) d\xi ds \right\} \tag{2.93}$$

is then expressed as

$$J_3(x;a) = \mathbb{E}\left\{ \int_0^{+\infty} e^{-\rho s} l(X(s;0,x,a(\cdot)), a(s)) ds \right\}. \tag{2.94}$$

Hypotheses 2.11 and 2.12 are satisfied as in the finite horizon part. Hypothesis 2.29 holds thanks to Remark 2.30. Hypothesis 2.28 holds if β satisfies proper conditions. This is discussed in Section 3.6, Propositions 3.73 and 3.74. The Hamilton-Jacobi-Bellman equation associated with the problem is now

$$\rho v - \frac{1}{2} \operatorname{Tr}\, [QD^2 v] - \langle Ax + b(x), Dv \rangle - \inf_{a \in \Lambda} \{\langle a, Dv \rangle + l(x,a)\} = 0. \tag{2.95}$$

As regards the literature, we refer for to [286, 421, 422] for the viscosity solution approach in the fully nonlinear case and with multiplicative noise, to [81, 241, 334][14] for the strong solutions approach, to [212] for the BSDE approach in the case of multiplicative noise. In [226], an ergodic control problem is studied, using the results for the infinite horizon problem.

2.6.2. Stochastic controlled heat equation: boundary control. The second example is also concerned with the control of a nonlinear stochastic heat equation in a given space region $\mathcal{O}$ but perhaps in a more realistic case, when the control can be exercised only at the boundary of $\mathcal{O}$ or in a subset of $\mathcal{O}$. We present only the case of the control at the boundary, remarking that the case of the control on a subdomain of $\mathcal{O}$ (that may even reduce to a point) gives rise to very similar mathematical difficulties that are treated e.g. in [206, 336]. We consider two cases that are the most standard and commonly used: the first when the control at the boundary enters through the Dirichlet boundary condition, and the second when one controls the flow, i.e. the Neumann boundary condition.

2.6.2.1. *Setting of the problem: Dirichlet case.* As in the previous example, assume $\mathcal{O}$ to be an open, connected, bounded subset of $\mathbb{R}^N$ with smooth boundary $\partial\mathcal{O}$. We consider the controlled dynamical system driven by the following stochastic PDE on the time interval $[t,T]$, for $0 \le t \le T < +\infty$,

$$\begin{cases} dy(s,\xi) = [\Delta_\xi y(s,\xi) + f\,(y(s,\xi))] + dW_Q(s)(\xi) & \text{in } (t,T] \times \mathcal{O} \\ y(t,\xi) = x(\xi) & \text{on } \mathcal{O} \\ y(s,\xi) = \alpha(s,\xi) & \text{on } (t,T] \times \partial\mathcal{O}, \end{cases} \tag{2.96}$$

where Δ_ξ, f, W_Q, x, y are as in equation (2.85). The difference with respect to equation (2.85) is that here the control is no longer in the drift term of the state equation but it influences the system through its values at the boundary (the so-called Dirichlet boundary condition). So here the *control strategy* of the system is the function $\alpha : [t,T] \times \partial\mathcal{O} \times \Omega \to \mathbb{R}$ which may be interpreted as the dynamics of an external source of heat acting at every boundary point of $\mathcal{O}$.

Following the notation of Section 2.6.1.1, we denote the unique solution (whenever it exists, see next subsection for more precise setting) of (2.96) by $y^{\alpha,t,x}$ to

[14] The paper [334] also treats the case of a multiplicative noise.

underline the dependence of the state y, on the control α and on the initial data t, x.

Similarly to the distributed control case, a reasonable objective of the controller can be the one of minimizing a functional

$$I(t,x;\alpha) = \mathbb{E}\left\{\int_t^T\int_{\mathcal{O}} \beta_1(y^{\alpha,t,x}(s,\xi))d\xi + \int_{\partial\mathcal{O}} \beta_2(\alpha(s,\xi))d\xi ds + \int_{\mathcal{O}} \gamma(y^{\alpha,t,x}(T,\xi))d\xi\right\} \tag{2.97}$$

where $\beta_1, \beta_2, \gamma \in C^0(\mathbb{R})$ are given functions depending on the objective of the controller. Observe that the difference with respect to the cost functional I_2 in (2.87) is that here we take the integral on $\partial\mathcal{O}$ when the control α is involved. The goal of the controller here would be to minimize the functional I above, over all control strategies α which are progressively measurable with respect to the augmented filtration generated by W_Q, and such that the above integrals make sense. The constraints can be the same as in the distributed control case in Section 2.6.1.1.

2.6.2.2. *Setting of the problem: Neumann case.* In this case the boundary condition in (2.96) is replaced by

$$\frac{\partial y(s,\xi)}{\partial n} = \alpha(s,\xi) \text{ on } (t,T] \times \partial\mathcal{O}, \tag{2.98}$$

where n is the outward unit normal vector to $\partial\mathcal{O}$. This means that one controls the heat flow across the boundary. The goal again is to minimize a cost functional of type (2.97) over all admissible controls α.

2.6.2.3. *The infinite dimensional setting and the HJB equation.* To rewrite the state equation (2.96) (with either Dirichlet or Neumann boundary condition) in an infinite dimensional setting we take $H = L^2(\mathcal{O})$ and Λ to be a closed subset of $L^2(\partial\mathcal{O})$ depending on the control constraints.

We first consider the Dirichlet case. Let A be the operator defined in (2.88) and $b : H \to H$ be the Nemytskii operator defined in (2.89). Let D be the Dirichlet operator defined in (C.2). We denote, as before, $X(s) := y(s,\cdot) \in L^2(\mathcal{O})$, and $a(s) := \alpha(s,\cdot) \in \Lambda$. We assume in addition that Λ is bounded in $L^2(\partial\mathcal{O})$. Then, as explained in Appendix C, Section C.2 (Notation C.14), the state equation (2.96) can be formally rewritten as

$$\begin{cases} dX(s) = [AX(s) + b(X(s)) - ADa(s)]\, ds + dW_Q(s), & t < s \le T \\ X(t) = x, \quad x \in H. \end{cases} \tag{2.99}$$

Now, thanks to (C.5), if we write the term $-AD$ as $(-A)^\beta B$ for $\beta \in (3/4, 1)$, where $B := A^{1-\beta}D$, the operator B is bounded from $L^2(\partial\mathcal{O})$ to H. Thus, passing to the integral form (see again Notation C.14), we can write (2.99) as follows

$$X(s) = e^{(s-t)A}x + \int_t^s e^{(s-r)A}b(X(r))dr + \int_t^s (-A)^\beta e^{(s-r)A}Ba(r)dr + \int_t^s e^{(s-r)A}dW_Q(s). \tag{2.100}$$

The Neumann boundary control case is handled similarly. Here we take $\Lambda = L^2(\partial\mathcal{O})$ and $\mathcal{U}_t^\nu = M^2_\nu(t,T;L^2(\partial\mathcal{O}))$. Let A be the Laplace operator with Neumann boundary conditions (see e.g. [432], Section 5.2, page 180):

$$\begin{cases} D(A) = \left\{x \in H^2(\mathcal{O}) : \frac{\partial x}{\partial n} = 0\right\}, \\ Ax = \Delta x, \text{ for } x \in D(A). \end{cases} \tag{2.101}$$

It generates an analytic semigroup of compact operators $\{e^{tA}\}_{t\geq 0}$ in H. We consider, for fixed $\lambda > 0$ the Neumann operator N_λ defined in (C.7). Similarly to the Dirichlet boundary control case, as explained in Appendix C, Section C.3 (Notation C.17), the state equation can be formally expressed as an evolution equation as follows:

$$\begin{cases} dX(s) = [AX(s) + b(X(s)) + (\lambda I - A)N_\lambda a(s)]\, ds + dW_Q(s), \quad t < s \leq T \\ \\ X(t) = x, \quad x \in H. \end{cases} \tag{2.102}$$

Now, thanks to (C.10), if we write the term $(\lambda I - A)N_\lambda$ as $(\lambda I - A)^\beta B_\lambda$, for $\beta \in (1/4, 1/2)$ and $B_\lambda := (\lambda I - A)^{1-\beta} N_\lambda$, the operator B_λ is bounded from $L^2(\partial\mathcal{O})$ to H. Then, passing to the integral form (see again Notation C.17), we can rewrite (2.102) as

$$\begin{aligned} X(s) = e^{(s-t)A}x + \int_t^s e^{(s-r)A} b(X(r))dr \\ + \int_t^s (\lambda I - A)^\beta e^{(s-r)A} B_\lambda a(r)dr + \int_t^s e^{(s-r)A} dW_Q(s). \end{aligned} \tag{2.103}$$

If f (and thus b) is Lipschitz[15] and (1.67) holds (if $\mathrm{Tr}(Q) = +\infty$) both integral equations (2.100) and (2.103) have unique mild solutions (see Theorem 1.135 and Proposition 1.141) with continuous trajectories, which we denote by $X(s\,; t, x, a(\cdot))$ (or simply $X(s)$ if its meaning is clear).

Thus it follows from the discussion in Remark 2.17 that Hypotheses 2.11 and 2.12, and condition $(A4)$ in the Neumann case, needed for the dynamic programming principle hold for both problems.

We now define $l_1 \colon H \to \mathbb{R}$ by

$$l_1(x) = \int_{\mathcal{O}} \beta_1(x(\xi))d\xi,$$

$l_2 \colon \Lambda \to \mathbb{R}$ by

$$l_2(a) = \int_{\partial\mathcal{O}} \beta_2(a(\xi))d\xi,$$

and $g \colon H \to \mathbb{R}$ by

$$g(x) = \int_{\mathcal{O}} \gamma(x(\xi))d\xi.$$

The functional I in (2.97) can be rewritten in the Hilbert space setting as

$$J(t, x; a(\cdot)) = \mathbb{E}\left\{ \int_t^T [l_1(X(s)) + l_2(a(s))]\, ds + g(x(T)) \right\}. \tag{2.104}$$

Thus, if β_1, β_2, γ satisfy proper continuity and growth conditions that guarantee Hypothesis 2.23, then the hypotheses of Theorem 2.24 are satisfied, and thus the dynamic programming principle stated there holds.

The associated HJB equation in both cases can be written as

$$\begin{cases} v_t + \frac{1}{2}\,\mathrm{Tr}\,[QD^2 v] + \langle Ax + b(x), Dv\rangle \\ \qquad + \inf_{a\in\Lambda} \left\{ \left\langle (\lambda I - A)^\beta B_\lambda a, Dv \right\rangle + l_2(a) \right\} + l_1(x) = 0, \\ \\ v(T, x) = g(x), \end{cases} \tag{2.105}$$

[15] More general assumptions on f could be used, like e.g. in [79, 81] in the distributed control case.

where in the Dirichlet case we take $\lambda = 0$ and $\beta \in (3/4, 1)$, while in the Neumann case we take $\lambda > 0$ and $\beta \in (1/4, 1/2)$.

Observe that the term $\langle (\lambda I - A)^\beta B_\lambda a, Dv \rangle$ caused by the presence of the boundary control term in the state equation does not make sense in general. However if his term is interpreted as $\langle B_\lambda a, (\lambda I - A)^\beta Dv \rangle$ then (writing $l(x, a) = l_1(x) + l_2(a)$), the Hamiltonian F_0 defined by

$$F_0(x, p) = \inf_{a \in \Lambda} \left\{ \langle B_\lambda a, (\lambda I - A)^\beta p \rangle + l(x, a) \right\} \tag{2.106}$$

is well defined on $H \times D\left((\lambda I - A)^\beta\right)$. Such unboundedness of F_0 is difficult to treat and typically requires better regularity properties of the solution, e.g. that the $Dv(t, x)$ belongs to the narrower space $D\left((\lambda I - A)^\beta\right)$. Since $D\left((\lambda I - A)^\beta\right)$ is larger in the Neumann case, the regularity needed for the value function to solve the HJB equation in the Neumann case is weaker than in the Dirichlet case. Thus the Neumann case can be studied under weaker assumptions and/or with better results. In the framework of viscosity solutions, the unboundedness of F_0 may require additional conditions on test functions. Overall this problem is much more difficult to study than the one of the previous section.

The theory of viscosity solutions has been developed for such equations in [242].[16] It is presented in Section 3.12, where existence and uniqueness of viscosity solutions for stationary HJB equations is proved in large generality, covering Dirichlet boundary conditions and very general drift and diffusion coefficients allowing for possibly fully nonlinear Hamiltonians. However no results about feedback controls exist with this approach. A Cauchy problem was also studied in [439] using the techniques of [242]. A related finite horizon problem has been studied partly with a viscosity solution approach in [451] in a case with boundary control and boundary noise: uniqueness of solutions is not proved.

Regarding the strong solution approach, only a special one dimensional Neumann case has been investigated in [136, 235] when the term b is zero or regular (see also Chapter 4 and, in particular, the examples of Section 4.8). The existence and uniqueness of a regular solution of the HJB equation, and the existence of feedback controls were obtained there.[17] The L^2 approach presented in Chapter 5 and the BSDE approach of Chapter 6 have not yet been applied to such equations, however there are results when the boundary control comes together with the boundary noise, see Section 2.6.3.

2.6.3. Stochastic controlled heat equation: boundary control and boundary noise. The third example still concerns the control of the stochastic heat equation in a given space region $\mathcal{O}$. In this case we assume that both the noise and the control act only at the boundary of that region. The problem is very hard since the presence of the noise at the boundary introduces a strong unboundedness in the model. For this reason, up to now, only one dimensional cases have been studied, and so we present here a one-dimensional example with a Neumann boundary condition taken from [131]. We will also mention what happens in a more difficult case of Dirichlet boundary condition (see [165, 338]).

2.6.3.1. *Setting of the problem.* We consider an optimal control problem for a state equation of parabolic type on a bounded interval, which, for convenience, we take to be $[0, \pi]$. We consider a Neumann boundary condition in which the derivative of the unknown function is equal to the sum of the control and of a white

[16]See also, for the deterministic case, the papers [67, 70, 71, 163, 164].

[17]See also, for the deterministic case, [167, 168, 171].

noise in time, namely:

$$\begin{cases} \dfrac{\partial y}{\partial s}(s,\xi) = \Delta_\xi y(s,\xi) + f\left(y(s,\xi)\right) & \text{in } (t,T] \times (0,\pi), \\[2ex] y(t,\xi) = x(\xi) & \text{on } (0,\pi), \\[2ex] \dfrac{\partial y(s,0)}{\partial n} = a_1(s) + \dot{W}_1(s), \quad \dfrac{\partial y(s,\pi)}{\partial n} = a_2(s) + \dot{W}_2(s) & \text{on } (t,T], \end{cases} \tag{2.107}$$

In the above equation, $\{W_i(t),\, t \geq 0\}$, $i = 1,2$, are independent standard real Wiener processes; the unknown $y(s;\xi,\omega)$, representing the state of the system, is a real-valued process; the control is modeled by the real-valued processes $a_i(s,\omega)$, $i = 1,2$ acting, respectively, at $\xi = 0$ and $\xi = \pi$; x is in $L^2(0,\pi)$, which are progressively measurable with respect to the augmented filtration generated by $W = (W_1, W_2)$. The function f belongs to $C_b(\mathbb{R})$ and is globally Lipschitz continuous.

The functional to minimize is

$$I(t,x;a_1(\cdot),a_2(\cdot)) =$$

$$\mathbb{E}\left[\int_t^T \left(\int_0^\pi \beta_1(\xi, y(s,\xi))d\xi + \beta_2(a_1(s),a_2(s))\right)ds + \int_0^\pi \gamma(\xi, y(T,\xi))ds\right] \tag{2.108}$$

2.6.3.2. *The infinite dimensional setting.* To rewrite the problem in an infinite dimensional setting we take $H = L^2(0,\pi)$, $\Lambda = \Xi = \mathbb{R}^2$, Q is the identity operator on Ξ, $W_Q = W$, and $a(\cdot) = (a_1(\cdot), a_2(\cdot))$. As in the previous example, using the results of Section 1.4 and Appendix C Section C.3, we get, formally, the following infinite dimensional state equation for the variable $X(s) = y(s,\cdot)$:

$$\begin{cases} dX(s) = \left[AX(s) + b(X(s)) + (\lambda I - A)N_\lambda a(s)\right]ds + (\lambda I - A)N_\lambda dW_Q(s), \;\; s \in (t,T] \\[2ex] X(t) = x, \quad x \in H, \end{cases} \tag{2.109}$$

where b and N_λ are as in the Section 2.6.2.3. This equation is interpreted in the mild form as

$$\begin{aligned} X(s) = e^{(s-t)A}x + \int_t^s e^{(s-r)A}[b(X(s)) + (\lambda I - A)N_\lambda a(s)]dr \\ + \int_t^s (\lambda I - A)^\beta e^{(s-r)A} B_\lambda dW_Q(r), \end{aligned} \tag{2.110}$$

where $\beta \in (1/4, 1/2)$ and $B_\lambda := (\lambda I - A)^{1-\beta} N_\lambda$ is a bounded operator. We take $\mathcal{U}_t$ to be the set of processes $a(\cdot)$ belonging to $M^2_\nu(t,T;\mathbb{R}^2)$ for a given reference probability space.

Theorem 1.135 guarantees the existence and uniqueness of a mild solution $X(s) := X(s;t,x,a(\cdot))$ of (2.110) with continuous trajectories. The validity of Hypotheses 2.11 and 2.12, and $(A4)$ needed for the dynamic programming principle are discussed in Remark 2.17.

We now define $l\colon H \to \mathbb{R}$ by

$$l_1(x) = \int_0^\pi \beta_1(\xi, x(\xi))d\xi,$$

$l_2\colon \Lambda \to \mathbb{R}$ by

$$l_2(a) = \beta_2(a_1, a_2),$$

and $g\colon H \to \mathbb{R}$ by

$$g(x) = \int_0^\pi \gamma(\xi, x(\xi))d\xi.$$

The functional I in (2.97) can thus be rewritten as

$$J(t,x;a(\cdot)) = \mathbb{E}\left\{\int_t^T [l_1(X(s)) + l_2(a(s))]\,ds + g(x(T))\right\}. \tag{2.111}$$

Again, β_1, β_2, γ must satisfy the right continuity and growth assumptions to guarantee Hypothesis 2.23, so that we can claim that the dynamic programming principle be satisfied.

2.6.3.3. *The HJB equation.* The HJB equation associated with the problem (2.110)-(2.111) is

$$\begin{cases} v_t + \dfrac{1}{2}\operatorname{Tr}\left[(\lambda I - A)N_\lambda \left[(\lambda I - A)N_\lambda\right]^* D^2 v\right] + \langle Ax, Dv\rangle + F_0(Dv) + l_2(x) = 0, \\ \\ v(T,x) = g(x),\ x \in H, \end{cases}$$

where the Hamiltonian F_0 is given by

$$F_0(p) = \inf_{a\in\mathbb{R}} \left\{\langle (\lambda I - A)N_\lambda a, p\rangle + l_2(a)\right\}.$$

Similarly to the boundary control case, here the Hamiltonian makes sense when one rewrites the term $\langle (\lambda I - A)N_\lambda a, p\rangle$ as $\langle B_\lambda a, (\lambda I - A)^\beta p\rangle$, and then F_0 is unbounded with respect to the variable p as it is only defined if $p \in D((\lambda I - A)^\beta)$. However, the extra difficulty arises due to the second order term $\operatorname{Tr}[(\lambda I - A)N_\lambda \left[(\lambda I - A)N_\lambda\right]^* D^2 v]$ which is written here in a formal way and needs to be given special interpretation. We notice that the same "operator" $(\lambda I - A)N_\lambda$ acts on the control and on the Wiener process in (2.109), and thus the control acts on the solution in the same way as the noise. This allows to use the BSDE approach to mild solutions (see [131] and, later, [448, 464, 465]). The L^2 approach is in principle applicable to this problem but, up to now, it has not been developed. Concerning a viscosity solution approach we mention the paper [451] where the authors show that the value function is a viscosity solution of the HJB equation but without proving uniqueness. At the present stage, the perturbation approach to mild solution presented in Chapter 4 does not seem applicable here.

REMARK 2.45 A one dimensional control problem in the half-line $[0, +\infty)$ with boundary control and noise in the Dirichlet case (i.e. with the boundary condition of the type $y(s,0) = a(s) + \dot{W}(s)$ for $s \in (t,T]$) has been studied in [165] and [338].

However, the choice of the infinite dimensional setting in this case presents a problem, since, choosing as the state space $H = L^2(0,+\infty)$, the continuity of the trajectories in $L^2(0,+\infty)$ is not ensured (see for example [125] Proposition 3.1, page 176). We can have the continuity only in some spaces of distributions extending $L^2(0,+\infty)$, or in L^2 with a suitable weight (see [165] Proposition 2.2, Lemma 2.2 and Theorem 2.7).

In [165] the linear quadratic case is studied while in [338] a more general case is studied by the BSDE approach. The problem has not been studied yet by other methods.

We remark that, similarly to the case of boundary control, the HJB equation for the Dirichlet boundary noise case is more difficult than the one for the Neumann boundary noise (as the unbounded operators arising in the first and second order terms contain "higher powers of A"). ■

2.6.4. Optimal control of the stochastic Burgers equation. Our fourth example concerns optimal control of the stochastic Burgers equation. Deterministic Burgers equation has been introduced by J. M. Burgers (see e.g. [61, 62]) as a model in fluid mechanics and has been later used in various areas of applied mathematics such as acoustics, dispersive water waves, gas dynamics, traffic flow, heat conduction, etc. As explained in [126] (page 255) the deterministic Burgers equation is not a good model for turbulence since it does not display any chaotic phenomena; even when a force is added to the right hand side, all solutions converge to a unique stationary solution as time goes to infinity. The situation is different when the force is random. Several authors have indeed suggested using the stochastic Burgers equation as a simple model for turbulence, as [83, 88, 279]. In [284] it is used to model the growth of a one-dimensional interface. Among other papers on the subject we mention [119], [146], [118].

Here we present a simple optimal control problem for the one dimensional stochastic Burgers equation motivated by a model of the control of turbulence formulated in [88] and studied in [112, 113].

2.6.4.1. *Setting of the problem.* The state equation is the following stochastic controlled viscous Burgers equation

$$\begin{cases} dy(s,\xi) = \left[\dfrac{\partial^2 y(s,\xi)}{\partial \xi^2} + \dfrac{1}{2}\dfrac{\partial}{\partial \xi} y^2(s,\xi) + \sqrt{Q}\alpha(s,\cdot)(\xi)\right] ds + dW_Q(s)(\xi), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad s \in (t,T], \xi \in (0,1), \\ y(t,\xi) = x(\xi), \quad \xi \in [0,1], \\ y(s,0) = y(s,1) = 0, \quad s \in [t,T]. \end{cases} \tag{2.112}$$

Here:

- the function $y : [t,T] \times [0,1] \times \Omega \to \mathbb{R}$, $(s,\xi,\omega) \mapsto y(s,\xi,\omega)$ describes the evolution of the velocity field of the fluid;
- the control $\alpha : [t,T] \times (0,1) \times \Omega \to \mathbb{R}$, $(s,\xi) \mapsto \alpha(s,\xi,\omega)$ gives the dynamics of the external force acting at every point of $(0,1)$;
- W_Q is a Q-Wiener process with $Q \in \mathcal{L}_1^+(L^2(0,1))$ and $(\mathscr{F}_s^t)_{s\in[t,T]}$ is the augmented filtration generated by W_Q;
- $x(\cdot) \in L^2(0,1)$ gives the distribution of the initial velocity field.

As before, the solution of (2.112) is denoted by $y^{\alpha,t,x}$. A possible objective of the controller (used in [88, 112, 113]) is to minimize a functional

$$I(t,x;\alpha(\cdot)) = \mathbb{E}\left\{ \int_t^T \int_0^1 \left[\left| \frac{\partial y^{\alpha,t,x}(s,\xi)}{\partial \xi} \right|^2 + \frac{1}{2}|\alpha(s,\xi)|^2 \right] d\xi ds \right.$$
$$\left. + \int_0^1 \frac{1}{2} |y^{\alpha,t,x}(T,\xi) - \bar{y}(\xi)|^2 d\xi \right\}. \tag{2.113}$$

where $\bar{y}$ is a given “desired” velocity profile. The main idea beyond this form of the cost functional is that we try to get the final velocity field to be close to $\bar{y}$ while minimizing the “vorticity" of the flow (measured here by the integral of the space derivative) and the energy spent controlling the system.

We then minimize the functional I over all control strategies α which are progressively measurable with respect to $\mathscr{F}_s^t$, and such that $\mathbb{E}\int_t^T \int_0^1 |a(s,\xi)|^2 d\xi ds < +\infty$. Sometimes we may require some additional bounds on the control strategies (e.g. $\alpha(s,\xi) \in [m,M]$ for some $m < M$).

Note that the operator $\sqrt{Q}$ acting on the control is the square root of the covariance operator of the Wiener process. This can be interpreted as "the noise acting on the control."

2.6.4.2. *The infinite dimensional setting and the HJB equation.* We take $H = \Lambda = L^2(0,1)$. The state equation (2.112) and the functional (2.113) can be rewritten as an abstract evolution equation in H using the operator

$$\begin{cases} D(A) = H^2(0,1) \cap H_0^1(0,1), \\ Ax = \frac{\partial^2}{\partial \xi^2} x, \text{ for } x \in D(A), \end{cases} \tag{2.114}$$

and the nonlinear operator

$$\begin{cases} D(B) = H^1(0,1) \\ B(x)(\xi) = x(\xi) \frac{\partial}{\partial \xi} x(\xi), \text{ for } x \in D(B). \end{cases} \tag{2.115}$$

Indeed, once we set $X(s) = y(s,\cdot) \in L^2(0,1)$, $a(s) = \alpha(s,\cdot) \in L^2(0,1)$, the state equation (2.112) becomes

$$\begin{cases} dX(s) = \left(AX(s) + B\left(X(s)\right) + \sqrt{Q}a(s)\right) ds + dW_Q(s) \\ X(t) = x, \end{cases} \tag{2.116}$$

and (2.113) is equivalent to

$$J(t,x;a(\cdot)) = \mathbb{E}\left\{ \int_t^T \left[\left| (-A)^{1/2} x(s) \right|_H^2 + \frac{1}{2} |a(s)|_H^2 \right] ds + \frac{1}{2} \left| X(T) - \bar{y} \right|_H^2 \right\}. \tag{2.117}$$

Differently from the previous examples, the standard mild solution approach does not work for equation (2.116). Therefore, the existence and uniqueness results require a different framework, see [119] and [127] Chapter 14 (the result is also stated in Section 4.9). An unpleasant consequence of this is the fact that it is not obvious to conclude that Hypotheses 2.11 and 2.12 needed for the dynamic programming principle are satisfied. We will not deal explicitly with this problem in the book, but we will see in Chapter 3 how to show the DPP for a much more difficult problem, namely the optimal control of the 2-D stochastic Navier-Stokes equations (see next section).

The HJB equation related to our control problem is (see [113] equation (2.4))

$$\begin{cases} v_t(t,x) + \frac{1}{2} \mathrm{Tr}\left[Q D^2 v(t,x) \right] + \langle Dv(t,x), Ax + B(x) \rangle \\ \qquad\qquad - \frac{1}{2} \left| \sqrt{Q} Dv(t,x) \right|^2 + \left| (-A)^{1/2} x \right|^2 = 0, \\ v(T,x) = \frac{1}{2} |x - \bar{y}|^2. \end{cases} \tag{2.118}$$

This equation is difficult to investigate due to the presence of the non-linear unbounded term $\langle Dv(t,x), B(x) \rangle$, coming from the state equation, and the term $\left| (-A)^{-1/2} x \right|^2$, coming from the objective functional. It was studied in [113] by a Hopf-type change of variable and in [112, 114] using a variant the mild solution approach. In these papers the authors use finite dimensional approximations of the state equation and are able to obtain existence and uniqueness of regular solutions to the HJB equation and to find optimal control in feedback form (see Subsection 4.9.1.1).

It is interesting to note that this technique has been extended to the case of control of stochastic Navier-Stokes equations in dimensions 2 and 3 (see next section). Equation (2.118) can also be investigated using the viscosity solution approach even though there are no explicit results. However, we refer the readers to Chapter 3 and [244]).

2.6.5. Optimal control of the stochastic Navier-Stokes equations. The stochastic Navier-Stokes equations are used to model turbulent flows. We refer the reader to the books [299, 443], the survey article [187], Chapter 15 of [126], and the paper [39] for more on this.

The optimal control of the Navier-Stokes equations, in the deterministic and stochastic cases, is a very challenging problem, both from the theoretical and applied points of view. For a survey on this subject we refer to the book [215] for the deterministic case and the paper [419] for the stochastic case. The dynamic programming approach to the optimal control of stochastic Navier Stokes equations has been investigated in the papers [115, 244, 328].

We consider a model problem for a control of turbulent flow governed by the stochastic two-dimensional Navier-Stokes equations for incompressible fluids. We follow mainly the paper [115] with some changes borrowed e.g. from [419] and [244]. It has been partly generalized to the three-dimensional case in [328].

2.6.5.1. *Setting of the problem.* We are given an open domain $\mathcal{O} \subset \mathbb{R}^2$ with locally Lipschitz boundary: it includes e.g. the case where $\mathcal{O}$ is a rectangle (which is quite common in the literature, see e.g. [116, 436, 244]). Given $\xi = (\xi_1, \dots, \xi_n), \eta = (\eta_1, \dots, \eta_n) \in \mathbb{R}^n$ we use the notation $\xi \cdot \eta := \sum_{i=1}^n \xi_i \eta_i$.

We take any reference probability space $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W_Q)$ satisfying the usual conditions, where W_Q is an $L^2(\mathcal{O};\mathbb{R}^2))$-valued Q-Wiener process with $Q \in \mathcal{L}_1^+(L^2(\mathcal{O};\mathbb{R}^2))$.

The control variable is an external force $\alpha(s,\xi)$ acting at every point ξ of $\mathcal{O}$; for models with control on subdomains one can see e.g. [262] and [419] (page 3). The controls stochastic processes progressively measurable with respect to the filtration $\mathscr{F}_s^t$ for a given reference probability space, and such that $|\alpha(s,\cdot)|_{L^2(\mathcal{O};\mathbb{R}^2))} \le R$ for some fixed $R > 0$, for all $(s,\omega) \in [t,T] \times \Omega$. The unknowns are the velocity vector field $y(s,\xi) = (y_1(s,\xi), y_2(s,\xi))$ and the pressure $p(s,\xi)$: they satisfy the system

$$\begin{cases} dy(s) + [(y(s,\xi)\cdot\nabla)y(s,\xi) + \nabla p(s,\xi)]\,ds \\ \qquad\qquad = [\nu\Delta y(s,\xi) + \alpha(s,\xi)]\,ds + dW_Q(s)(\xi) \text{ in } (t,T]\times\mathcal{O} \\ \operatorname{div}(y(s,\xi)) = 0 \text{ in } [t,T]\times\mathcal{O} \\ y(s,\xi) = 0 \text{ on } [t,T]\times\partial\mathcal{O} \\ y(0,\xi) = x(\xi) \text{ on } \mathcal{O}, \end{cases} \tag{2.119}$$

(∇ denotes $(\partial_{\xi_1}, \partial_{\xi_2})$ and $y \cdot \nabla$ denotes $y_1\partial_{\xi_1} + y_2\partial_{\xi_2}$). The positive constant ν represents the kinematic viscosity. We remark that distributed control can be approximately realized for electrically conducting fluids (like salt water, liquid metals, etc.) by a suitable Lorentz force distribution. The boundary control, which is not present in this example, is typically implemented by blowing and suction at the boundary.

Suppose, as in the previous section, that we want to achieve the desired profile $\bar{y}$ of the flow while minimizing the turbulence of the flow and the amount of energy used to control it. This is is the most common case in engineering applications. We recall that we can measure how turbulent a flow is by evaluating the time averaged enstrophy, which is defined by

$$\int_{\mathcal{O}} |\operatorname{curl} y(s,\xi)|^2 \, d\xi,$$

where the rotational operator curl in dimension 2 is defined as

$$\mathrm{curl}\,(y_1, y_2) = \frac{\partial y_1}{\partial \xi_2} - \frac{\partial y_2}{\partial \xi_1}. \tag{2.120}$$

Thus we consider the problem of minimizing the following functional over all control strategies α:

$$I(t,x;\alpha(\cdot)) = \mathbb{E}\left[\int_t^T \int_{\mathcal{O}} \left[|\mathrm{curl}\; y(s,\xi)|^2 + \frac{1}{2}|\alpha(s,\xi)|^2\right] d\xi ds + \int_{\mathcal{O}} |y(T,\xi) - \bar{y}(\xi)|^2\, d\xi\right]. \tag{2.121}$$

In areas like combustion the goal may be to maximize mixing (and hence turbulence) of the flow. As remarked in [419] (page 3) in some flow control problems and in data assimilation problems in meteorology one may also minimize the functional

$$I_1(t,x;\alpha(\cdot)) = \mathbb{E}\left[\int_t^T \int_{\mathcal{O}} \left[|\mathrm{curl}\;(y(s,\xi) - \bar{y}_d(s,\xi))|^2 + \frac{1}{2}|\alpha(s,\xi)|^2\right] d\xi ds\right] \tag{2.122}$$

for a given velocity field $y_d(s,\xi)$. (See also [215], page 167, formula (1.15), for a similar type of functional, in the deterministic case.)

2.6.5.2. *The infinite dimensional setting and the HJB equation.* Denote

$$\mathscr{V} := \left\{ f \in C_0^\infty(\mathcal{O};\mathbb{R}^2) \;:\; \mathrm{div}(f) = 0 \right\}, \tag{2.123}$$

$$H := \text{the closure of } \mathscr{V} \text{ in } L^2(\mathcal{O};\mathbb{R}^2), \tag{2.124}$$

and

$$V := \text{the closure of } \mathscr{V} \text{ in } H^1(\mathcal{O};\mathbb{R}^2). \tag{2.125}$$

Recall that we have an orthogonal decomposition

$$L^2(\mathcal{O};\mathbb{R}^2) = H \times H^\perp,$$

where $H^\perp = \{f = \nabla p :\ \text{for some } p \in H^1(\mathcal{O})\}$.

We define the unbounded operator in H

$$\begin{cases} D(A) := H^2(\mathcal{O};\mathbb{R}^2) \cap V \subset H \\ A := P\Delta \end{cases}$$

where P is the orthogonal projection in $L^2(\mathcal{O};\mathbb{R}^2)$ onto H. The operator A is self-adjoint and strictly negative (see [437]), generates a C_0-semigroup on H, and moreover $V = D((-A)^{1/2})$. We also define the bilinear operator

$$\begin{cases} B\colon V \times V \to H \\ B(x,y) = P(x\cdot\nabla)y \end{cases}$$

and set $B(x) := B(x,x)$.

Applying the projection P to equation (2.119) and setting $X(s) = y(s,\cdot) \in H$, $a(s) = P\alpha(s,\cdot) \in H$, we obtain

$$\begin{cases} dX(s) = (\nu AX(s) - B\,(X(s)) + a(s))\,ds + PdW_Q(s) \\ X(t) = x. \end{cases} \tag{2.126}$$

Since $|a(s)|_H = |a(s)|_{L^2(\mathcal{O};\mathbb{R}^2)} \le |\alpha(s,\cdot)|_{L^2(\mathcal{O};\mathbb{R}^2)}$, we can obviously restrict the set of controls to these with values in H. Moreover, for $x \in V$, $|\mathrm{curl}\; x|_{L^2(\mathcal{O})} = |\nabla x|_{L^2(\mathcal{O};\mathbb{R}^4)} = |(-A)^{1/2}x|_H$. Thus the minimization of the functional I in (2.121) is equivalent to the minimization of

$$J(t,x;a(\cdot)) = \mathbb{E}\left[\int_t^T \left[\left|(-A)^{1/2}X(s)\right|_H^2 + \frac{1}{2}|a(s)|_H^2\right] ds + |X(T) - \bar{y}|_H^2\right] \tag{2.127}$$

over all $a(\cdot) \in \mathcal{U}_t$, which is defined as in Section 2.2.1 with $\Lambda := B_H(0,R)$.

There are various ways to define solutions of (2.126) and for existence and uniqueness results we refer for instance to [93, 127, 342, 443] and to Chapter 3 and Section 4.9 here. Unfortunately we cannot apply the definition of mild solution, and thus showing that Hypotheses 2.11 and 2.12 are satisfied requires a different argument. We explain it in Chapter 3.

The Hamiltonian for the control problem is

$$F(p) := \begin{cases} -\frac{1}{2}|p|^2 & \text{if } |p| \le R \\ -|p|R + \frac{1}{2}R^2 & \text{if } |p| > R, \end{cases}$$

and the Hamilton-Jacobi-Bellman equation for the system becomes

$$\begin{cases} v_t + \dfrac{1}{2}\mathrm{Tr}\left[PQP^* D^2 v\right] + \langle Dv, \nu Ax - B(x)\rangle \\ \qquad\qquad\qquad\qquad + F(Dv) + \left|(-A)^{1/2}x\right|^2 = 0 \\ v(T,x) = |x - \bar{y}|^2. \end{cases} \tag{2.128}$$

We note that one can also associate a different control problem with (2.128) by considering PW_Q to be a $\tilde{Q}$-Wiener process in H with $\tilde{Q} = PQP^*$ and taking controls to be the progressively measurable processes with respect to the augmented filtration generated by PW_Q.

Similarly to the case of the control of the stochastic Burgers equation discussed in the previous section, the difficulty of the HJB equation (2.128) comes from the presence of the unbounded terms $\langle Dv, b(x)\rangle$ and $\left|(-A)^{1/2}x\right|^2$. The unboundedness caused by the operator b is much worse now. However the approach used for the Burgers case by Da Prato and Debussche still allows to obtain satisfactory results on existence and uniqueness of regular solutions [115] under some conditions on $\tilde{Q}$. In [328] existence of regular solutions is proved for the three dimensional case. These results are described in Subsections 4.9.1.2, 4.9.1.3. The viscosity solution approach can be applied to the two dimensional case with more general cost functional and noise and it yields existence and uniqueness of viscosity solutions (see [244] and Chapter 3). The viscosity solution approach for the three dimensional case is still open. Some results in the deterministic case are in [412].

2.6.6. Optimal control of the Duncan-Mortensen-Zakai equation. This example concerns a class of finite dimensional stochastic optimal control problems with partial observation and correlated noises. We present the problem and we briefly show its connection with the so-called "separated" problem (see e.g. [34, 156, 157, 193, 370]) which is a fully observable infinite dimensional stochastic optimal control problem. The setting of the partially observed control system we describe here is the same as in [245] and is borrowed from [370, 468, 469], (see also [257, 258, 259, 320]). Duncan-Mortensen-Zakai (DMZ) equation, separated problem and optimal control of the DMZ equation are also discussed in details in [364]. The presentation in [364] relies on [406] which also discusses filtering problems.

2.6.6.1. *An optimal control problem with partial observation.* Consider, in the interval $[t,T]$ a random state process y in $\mathbb{R}^d$ and a random observation process y_1 in $\mathbb{R}^m$. The state-observation equation is

$$\begin{cases} dy(s) = b^1(y(s),a(s))ds + \sigma^1(y(s),a(s))dW^1(s) + \sigma^2(y(s),a(s))dW^2(s), \\ y(t) = \eta, \\ dy_1(s) = h(y(s))ds + dW^2(s), \\ y_1(t) = 0, \end{cases} \tag{2.129}$$

where W^1 and W^2 are two independent Brownian motions in $\mathbb{R}^d$ and $\mathbb{R}^m$ respectively on some stochastic basis $\left(\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s\in[t,T]}, \mathbb{P}\right)$ which is a complete probability space with the filtration satisfying the usual conditions. The initial condition η is assumed to be $\mathscr{F}_t$-measurable and square integrable. The control set $\Lambda \subset \mathbb{R}^n$, and admissible controls are the processes $a(\cdot) : [t,T]\times\Omega \to \Lambda$ that are progressively measurable with respect to the filtration $\{\mathscr{F}_s^{y_1} : s \in [t,T]\}$, which is the augmented filtration of the filtration $\left\{\mathscr{F}_s^{y_1,0} : s \in [t,T]\right\}$ generated by the observation process y_1. We assume the following.

HYPOTHESIS 2.46 *The set Λ is a closed subset of $\mathbb{R}^n$. The functions*

$$b^1 : \mathbb{R}^d \times \Lambda \to \mathbb{R}^d, \qquad h : \mathbb{R}^d \to \mathbb{R}^m$$

are uniformly continuous and $b^1(\cdot, a), h$ have their $C^2(\mathbb{R}^d)$ norms bounded, uniformly for $a \in \Lambda$. Moreover the functions

$$\sigma^1 : \mathbb{R}^d \times \Lambda \to \mathcal{L}\left(\mathbb{R}^d, \mathbb{R}^d\right), \qquad \sigma^2 : \mathbb{R}^d \times \Lambda \to \mathcal{L}\left(\mathbb{R}^m, \mathbb{R}^d\right)$$

are uniformly continuous and $\sigma^1(\cdot, a), \sigma^2(\cdot, a)$ have their $C^3(\mathbb{R}^d)$ norms bounded, uniformly for $a \in \Lambda$, and

$$\sigma^1(\xi, a)\left[\sigma^1(\xi, a)\right]^T \geq \lambda I$$

for some $\lambda > 0$ and all $\xi \in \mathbb{R}^d, a \in \Lambda$.

This assumption in particular guarantees the existence of a unique strong solution of the state equation (2.129), see e.g. Theorem 1.121. We denote its solution at time s by $(y(s; t, \eta, a(\cdot)), y_1(s; t, a(\cdot)))$ or simply by $(y(s), y_1(s))$.

We now consider the problem of minimizing the cost functional

$$I(t, \eta; a(\cdot)) = \mathbb{E}\left\{\int_t^T l_1(y(s), a(s))ds + g_1(y(T))\right\} \tag{2.130}$$

over all admissible controls, where the cost functions

$$l_1 : \mathbb{R}^d \times \Lambda \to \mathbb{R}; \qquad g_1 : \mathbb{R}^d \to \mathbb{R}$$

are suitable continuous functions, say with at most polynomial growth at infinity in the variable y, uniformly with respect to the variable a. A control strategy $a^*(\cdot)$ minimizing the cost I in (2.130) is called an *optimal control in the strict sense.*

2.6.6.2. *The separated problem.* The optimal control of partially observed diffusions is a very difficult problem with still many open questions (e.g the existence of optimal controls in the strict sense, see e.g. [193], p. 261). One way of dealing with it is through the so called "separated" problem where one looks at the associated problem of controlling the unnormalized conditional probability density $Y(\cdot) : [t,T] \to L^1\left(\mathbb{R}^d\right)$ of the state process y given the observation y_1. This idea, that arises from well known results in nonlinear filtering (see e.g. [143, 350, 456]), has been introduced first in [193] to prove existence of optimal controls in a suitable weak sense. Here, following mainly [370], we briefly and informally explain the separated problem and how it arises.

To introduce the new state Y and to compute the equation for it we consider for each $s \in [t,T]$ the conditional law Π_s of the random variable $y(s)$ given the path of y_1 up to time s i.e., in our setting, given the σ-field $\mathscr{F}_s^{y_1}$, and look at its density with respect to the Lebesgue measure in $\mathbb{R}^d$. This density, up to a normalizing factor, will be the new state $Y(s)$ at time s. The conditional law Π_s is a measure valued process such that for every $f \in C_b(\mathbb{R}^d)$, the conditional expectation

$$\mathbb{E}\left[f(y(s))|\mathscr{F}_s^{y_1}\right] = \int_{\mathbb{R}^d} f(y)d\Pi_s(y) \quad \mathbb{P}\ a.s..$$

Using the notation of [370], the above expression will be denoted by $\Pi_s(f)$. The process Π_s exists if there exists a regular conditional probability given $\mathscr{F}_s^{y_1,0}$. To compute it, it is more convenient (as explained e.g. in [370] at the end of Section I.5) to change the probability measure. We define the new probability measure $\overline{\mathbb{P}}$ by

$$d\overline{\mathbb{P}} = \kappa^{-1}(T)d\mathbb{P},$$

where

$$\kappa(s) = \exp\left[\int_t^s \langle h(y(r)), dy_1(r)\rangle_{\mathbb{R}^m} - \frac{1}{2}\int_t^s |h(y(r))|^2_{\mathbb{R}^m} dr\right].$$

Since $\kappa(s)$ is a martingale we have

$$d\overline{\mathbb{P}} = \kappa^{-1}(s)d\mathbb{P}, \quad \text{on } \mathscr{F}_s. \tag{2.131}$$

It follows from the Girsanov Theorem (see e.g. [283], Section 3.5) that the processes W^1 and y_1 become two independent Brownian motions respectively in $\mathbb{R}^d$ and $\mathbb{R}^m$ in the new probability space $\left(\Omega, \mathscr{F}, (\mathscr{F}_s)_{s\in[t,T]}, \overline{\mathbb{P}}\right)$. In this space the equation for the process $y(\cdot)$ becomes

$$\begin{cases} dy(s) = \left[b^1(y(s), a(s)) - \sigma^2(y(s), a(s))h(y(s))\right] ds \\ \qquad\qquad +\sigma^1(y(s), a(s))dW^1(s) + \sigma^2(y(s), a(s))dy_1(s), \\ y(t) = \eta. \end{cases} \tag{2.132}$$

It follows from Bayes formula (see [406], page 225 or [364], Proposition 1.3, page 18), see Lemma 3.1 in [370], that for $s \in [t, T]$ and $f \in C_b(\mathbb{R}^d)$,

$$\Pi_s(f) = \mathbb{E}\left[f(y(s))|\mathscr{F}_s^{y_1}\right] = \frac{\overline{\mathbb{E}}\left[f(y(s))\kappa(s)|\mathscr{F}_s^{y_1}\right]}{\overline{\mathbb{E}}\left[\kappa(s)|\mathscr{F}_s^{y_1}\right]}.$$

So, if we are able to compute, for every $s \in [t, T]$ and for every $f \in C_b(\mathbb{R}^d)$, the quantity $\overline{\mathbb{E}}\left[f(y(s))\kappa(s)|\mathscr{F}_s^{y_1}\right]$, then we can also find $\Pi_s(f)$ for every such f.

By Itô's formula we have, for $f \in C_b^2(\mathbb{R}^d)$, $s \in [t, T]$,

$$\begin{aligned} f(y(s))\kappa(s) &= f(\eta) + \int_t^s \kappa(r)L_{a(r)}f(y(r))dr \\ &+ \int_t^s \kappa(r)\langle \nabla f(y(r)), \sigma^1(y(r), a(r))dW^1(r)\rangle_{\mathbb{R}^d} \\ &+ \int_t^s \kappa(r)\langle \nabla f(y(r)), \sigma^2(y(r), a(r))dy_1(r)\rangle_{\mathbb{R}^d} + \int_t^s \kappa(r)f(y(r))\langle h(y(r)), dy_1(r)\rangle_{\mathbb{R}^m}, \end{aligned}$$

where for every $a \in \Lambda$, $L_a : C_b^2(\mathbb{R}^d) \to C_b(\mathbb{R}^d)$ is given by

$$\begin{aligned} (L_a f)(\xi) &= \langle b^1(\xi, a), \nabla f(\xi)\rangle_{\mathbb{R}^d} \\ &+ \tfrac{1}{2}\mathrm{Tr}\left[\left(\sigma^1(\xi, a)[\sigma^1(\xi, a)]^T + \sigma^2(\xi, a)[\sigma^2(\xi, a)]^T\right) D^2 f(\xi)\right]. \end{aligned}$$

Now, computing the conditional expectation (see [370], Section I.4) we have, for $s \in [t, T]$,

$$\begin{aligned} \overline{\mathbb{E}}\left[f(y(s))\kappa(s)|\mathscr{F}_s^{y_1}\right] &= \overline{\mathbb{E}}[f(\eta)] + \int_t^s \overline{\mathbb{E}}\left[\kappa(r)L_{a(r)}f(y(r))|\mathscr{F}_r^{y_1}\right] dr \\ &+ \int_t^s \langle \overline{\mathbb{E}}\left[\kappa(r)[\sigma^2(y(r), a(r))]^T \nabla f(y(r)) + \kappa(r)f(y(r))h(y(r))|\mathscr{F}_r^{y_1}\right], dy_1(r)\rangle_{\mathbb{R}^m}. \end{aligned}$$

If $\overline{\Pi}_s$ is a measure-valued process such that, for every $f \in C_b(\mathbb{R}^d)$, $\overline{\Pi}_s(f) = \overline{\mathbb{E}}\left[\kappa(s)f(y(s))|\mathscr{F}_s^{y_1}\right]$, (which exists if Π_s exists) then the equation above implies that $\overline{\Pi}_s$ must satisfy the equation

$$\overline{\Pi}_s(f) = \overline{\Pi}_t(f) + \int_t^s \overline{\Pi}_r(L_{a(r)}f)dr + \int_t^s \langle \overline{\Pi}_r(B_{a(r)}f), dy_1(r)\rangle_{\mathbb{R}^m}, \qquad s \in [t, T],$$

where for every $a \in \Lambda$, $B_a : C_b^1(\mathbb{R}^d) \to C_b(\mathbb{R}^d; \mathbb{R}^m)$ is given by

$$(B_a f)(\xi) = [\sigma^2(\xi, a)]^T \nabla f(\xi) + f(\xi) h(\xi).$$

(To justify that $\overline{\mathbb{E}}\left[\kappa(r) F(y(r), a(r)) | \mathscr{F}_r^{y_1}\right] = \overline{\Pi}_t(F(\cdot, a(r)), \overline{\mathbb{P}}$ a.s. for a bounded and continuous function $F(\xi, a)$, one can use approximation by step functions, Proposition 1.40-(vii) and the Lebesgue dominated convergence theorem, since the equality is true for functions $F(\xi, a) = \mathbf{1}_A(a) \mathbf{1}_B(\xi)$, where A, B are Borel subsets of $\Lambda, \mathbb{R}^d$ respectively, as $a(r)$ is $\mathscr{F}_r^{y_1}$ measurable.)

If $\overline{\Pi}_s$ has a density $Y(s) \in L^1(\mathbb{R}^d)$ with respect to the Lebesgue measure, then the process $Y(\cdot)$ should satisfy, at least in a weak sense, the so-called Duncan-Mortensen-Zakai (DMZ) equation (introduced in [143, 350, 456])

$$dY(s) = L_{a(s)}^* Y(s) ds + \langle B_{a(s)}^* Y(s), dy_1(s) \rangle, \qquad Y(t) = x, \tag{2.133}$$

where x is the density of the law of the initial datum η of equation (2.129). The process $Y(\cdot)$ is called the unnormalized conditional density of the state with respect of the observation process. If one can prove that equation (2.133) has a solution, it is the density with respect to the Lebesgue measure of $\overline{\Pi}_s$, see [370], Section I.4 for more on this.

Now it is possible to rewrite the functional I in (2.130) in terms of the new probability space and infinite dimensional state $Y(\cdot)$. Indeed, assuming that the process $Y(\cdot)$ takes values in $L^2(\mathbb{R}^d)$, using (2.131) we have

$$\begin{aligned} I(t, \eta; a(\cdot)) &= \int_t^T \overline{\mathbb{E}}\left[\kappa(s) l_1(y(s), a(s))\right] ds + \overline{\mathbb{E}}\left[\kappa((T) g_1(y(T))\right] \\ &= \overline{\mathbb{E}}\left[\int_t^T \overline{\mathbb{E}}\left[\kappa(s) l_1(y(s), a(s)) | \mathscr{F}_s^{y_1}\right] ds\right] + \overline{\mathbb{E}}\left[\overline{\mathbb{E}}\left[\kappa(T) g_1(y(T)) | \mathscr{F}_s^{y_1}\right]\right] \\ &= \overline{\mathbb{E}}\left\{\int_t^T \langle l_1(\cdot, a(s)), Y(s) \rangle_{L^2} ds + \langle g_1(\cdot), Y(T) \rangle_{L^2}\right\} =: J(t, x; a(\cdot)) \end{aligned} \tag{2.134}$$

Computing the adjoint operators, L_a^*, B_a^*, we can rewrite (2.133) in an explicit and more familiar form

$$dY(s) = A_{a(s)} Y(s) ds + \sum_{k=1}^m S_{a(s)}^k Y(s) dy_{1,k}(s), \qquad Y(t) = x, \tag{2.135}$$

where for every $a \in \Lambda$, A_a and S_a^k $(k = 1, \dots, m)$ are the following differential operators

$$(A_a x)(\xi) = \sum_{i,j=1}^d \partial_i \left[a_{i,j}(\xi, a) \partial_j x(\xi)\right] + \sum_{i=1}^d \partial_i \left[b_i(\xi, a) x(\xi)\right], \tag{2.136}$$

and

$$\left(S_a^k x\right)(\xi) = \sum_{i=1}^d d_{ik}(\xi, a) \partial_i x(\xi) + e_k(\xi, a) x(\xi); \qquad k = 1, ..., m, \tag{2.137}$$

where

$$\begin{gathered} a(\xi, a) = \sigma^1(\xi, a) \left[\sigma^1(\xi, a)\right]^T + \sigma^2(\xi, a) \left[\sigma^2(\xi, a)\right]^T, \\ b_i(\xi, a) = -b_i^1(\xi, a) + \partial_j a_{i,j}(\xi, a); \qquad i = 1, ...d, \\ d(\xi, a) = -\sigma^2(\xi, a), \\ e_k(\xi, a) = h_k(\xi) - \partial_i \sigma_{ik}^2(\xi, a); \qquad k = 1, ...m. \end{gathered}$$

The *separated problem* is thus the problem of maximizing the functional J over all admissible controls $a(\cdot)$, with state equation (2.135). It is an infinite dimensional optimal control problem which can be studied within the framework of this book in the state space $H = L^2(\mathbb{R}^d)$ or other spaces. In Section 3.11 we will investigate it in suitable weighted spaces. It is worth noticing that even though the original control problem was nonlinear, the DMZ equation (2.135) is linear and the cost functional J is also linear in the state variable x.

The mild solution approach cannot be applied to (2.135). Existence and uniqueness of solutions in a variational sense was proved in [297] (see also [369]). We discuss variational solutions in Section 3.11 where we show that under suitable assumptions equation (2.135) is well posed in $L^2(\mathbb{R}^d)$ and in weighted versions of it. We also explain how to show that Hypotheses 2.11 and 2.12 hold.

The HJB equation for the infinite dimensional problem has the form

$$\begin{cases} v_t + \inf_{a\in\Lambda} \left\{ \frac{1}{2} \sum_{k=1}^m \left\langle D^2 v S_a^k x, S_a^k x \right\rangle + \langle A_a x, Dv\rangle + f(x,a) \right\} = 0, \\ v(T,x) = g(x). \end{cases} \tag{2.138}$$

This is a fully nonlinear equation with unbounded first and second order terms and up to know only a viscosity solution approach has given some results on existence and uniqueness of solutions [15, 245, 259, 320]. In [320] the equation was studied in a standard L^2 space when the operators S_a^k were bounded multiplication operators. In [259] it was shown that the value function is a viscosity solution in a very weak sense when the HJB equation was considered in the space of measures (see also [261], [190]). The results of [245] are presented in Section 3.11, where (2.138) will be studied in a weighted L^2 space. The optimal control problem for the DMZ equation and the HJB equation (2.138) are also discussed in [364].

2.6.7. Super-hedging of forward rates. We now present a stochastic optimal control problem arising in finance, in pricing derivatives. When the market is incomplete there is not a unique way to price a derivative product. It is then useful, in some cases, to find the range of all possible prices, i.e. the maximum and the minimum of possible prices, called the super-hedging and the sub-hedging price. Such prices are defined as value functions of suitable optimal control problems and the finite dimensional theory of this problem has been widely studied: see e.g. [14, 151] for the one dimensional case and [248, 249, 324, 403], [87, 415, 416, 417, 418] in the multidimensional case; see also [440] for a first idea of the method in the infinite dimensional case.

When the underlying asset is a forward rate the natural model for it is the so-called Musiela model introduced in [352] that describes the dynamics of forward rates in terms of evolution of an infinite dimensional diffusion process. Consequently, the super-hedging problem in such case is naturally formulated as a stochastic optimal control problem in infinite dimensions. We present now such problem, taken from the paper [287].

The Musiela model of interest rates [352] is a reparametrization of the Heath-Jarrow-Morton (HJM) model. In this model the forward rate process $\{r(t,\sigma)\}_{\sigma,t\geq 0}$ evolves according to a stochastic differential equation

$$dr(t,\sigma) = \left(\frac{\partial}{\partial\sigma} r(t,\sigma) + \sum_{i=1}^d \tau_i(t,\sigma) \int_0^\sigma \tau_i(t,\mu)d\mu \right) dt + \sum_{i=1}^d \tau_i(t,\sigma) dw(t)^i,$$

where $W = (w^1,...,w^d)$ is a standard d-dimensional Brownian motion, and τ_i are certain functions. Using the notation $A = \frac{d}{d\sigma}$, $\tau(t)(\sigma) = (\tau_1(t,\sigma),...,\tau_d(t,\sigma))$,

$$b(\tau(t))(\sigma) = \sum_{i=1}^d \tau_i(t,\sigma) \int_0^\sigma \tau_i(t,\mu)d\mu,$$

the above equation can be written as an abstract infinite dimensional stochastic differential equation

$$dr(t) = (Ar(t) + b(\tau(t)))dt + \tau(t) \cdot dW(t), \quad r(0) \in H, \tag{2.139}$$

where H is some separable Hilbert space of functions on $\mathbb{R}^+$ (for instance $H = L^2(\mathbb{R}^+)$, $H = H^1(\mathbb{R}^+)$ or their weighted versions), and $\cdot$ is the inner product in $\mathbb{R}^d$ (see [228, 352, 440, 441]). We call equation (2.139) the Heath-Jarrow-Morton-Musiela (HJMM) equation. Given the right choice of the space H and proper assumptions on τ, the equation has a unique mild solution, see Section 3.10. Using the process $r(t,\sigma)$, the price at time t of a zero-coupon bond (see [353]) with maturity T is

$$B_T(t) = e^{-\int_0^{T-t} r(t,\sigma)d\sigma}.$$

This model can be used to price swaptions, caps, and other interest rates and currency derivatives.

Consider first a case of European options. Given a contingent claim with the payoff function $g : H \to \mathbb{R}$ and an initial curve at time t, $r(t)(\sigma) = x(\sigma)$, the rational price of the option maturing at time T is

$$V(t,x) = \mathbb{E}\left(e^{-\int_t^T r(s,0)ds} g(r(T)) : r(t)(\sigma) = x(\sigma)\right). \tag{2.140}$$

For instance for a European swaption on a swap with cash-flows $C_i, i = 1, ..., n$, at times $T < T_1 < ... < T_n$,

$$g(z) = \left(K - \sum_{i=1}^n C_i e^{\int_0^{T_i - T} z(\sigma)d\sigma}\right)^+ \tag{2.141}$$

for some $K > 0$. The function V given by the (Feynman-Kac) formula (2.140) should satisfy the partial differential equation

$$\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{1}{2}\displaystyle\sum_{i=1}^d \left\langle D^2 u\, \tau_i(t), \tau_i(t)\right\rangle + \langle b(\tau(t)) + Ax, Du\rangle - x(0)u = 0 \\ \\ u(T,x) = g(x), \qquad\qquad\qquad\qquad (t,x) \in (0,T) \times H, \end{cases} \tag{2.142}$$

where $\langle \cdot, \cdot \rangle$ is the inner product in H. The above equation is called an infinite dimensional Black-Scholes equation. It was analyzed in [227] in the space $H = L^2((0,+\infty))$ where existence of smooth solutions was proved for smooth g and τ independent of t but possibly depending on the state variable x. The existence of solutions was also shown for some non-smooth g when τ was a constant by an argument that allowed a parallel between (2.142) and a finite dimensional Black-Scholes equation (see also [221, 440]).

The problem of pricing of American options in the framework of the Musiela model can be rephrased as an optimal stopping problem for the above infinite dimensional diffusion process and is connected to an obstacle problem

$$\max\left\{\frac{\partial u}{\partial t} + \frac{1}{2}\sum_{i=1}^d \left\langle D^2 u\, \tau_i(t), \tau_i(t)\right\rangle + \langle b(\tau(t)) + Ax, Du\rangle - x(0)u, u - \varphi\right\} = 0 \tag{2.143}$$

for some function φ. This equation was studied in [454] from the point of view of Bellman's inclusions and a similar obstacle problem for a related model was also investigated in [219].

One of the drawbacks of the Musiela model is that it does not guarantee the positivity of rates and in some cases it is almost certain that they are not positive (see [440]). To avoid such possibilities the term $x(0)$ was replaced by $x^+(0)$ in [454].

We do the same here and throughout the section we always take the positive part of the rates.

Let us explain now the super-hedging problem. Suppose that the dynamics of the forward rates are given by equation (2.139), however we are not able to determine precisely the process τ that describes the volatility of the market. We only know that it takes values in some set $\Lambda \subset H^d$. We consider an agent who wants to price and hedge a European contingent claim with payoff $g(r(T))$ that depends on the value of the forward rate curve at the maturity time T (note that in cases of interest in finance the payoff function g is not even C^1).

To find the super-hedging price, given an initial condition $r(t) \in H$, we try to maximize the payoff

$$\mathbb{E}\left(e^{-\int_t^T r^+(s,0)ds} g(r(T))\right) \tag{2.144}$$

with respect to all progressively measurable stochastic processes τ taking values in Λ. The processes τ become controls and the maximization of (2.144) gives the value function V which should provide the super-hedging price at time t as $C(t) = V(t, r(t))$. Following the standard finite dimensional theory for such problems (see e.g. [151, 324, 403]), a super-hedging strategy in such context (i.e. an investment strategy that replicates the super-hedging price) "should" then be given by the process $\pi(t) := DV(t, r(t))$, so in terms of the space-like derivative (i.e. with respect to r) of the value function V. In the infinite dimensional case similar results have not been proved yet as the method of proof requires strong regularity properties of the value function V which are not known. However the problem provides a strong motivation for studying the following optimal control problem: maximize (2.144) over all processes $\tau \in \mathcal{U}_t$, where the state equation is given by (2.139). For this optimal control problem Hypotheses 2.11 and 2.12 are satisfied thanks to Proposition 2.16. Moreover, if g is locally uniformly continuous and has at most polynomial growth, it can be proved that also Hypothesis 2.23 holds and so the dynamic programming principle holds. This is explained in Section 3.10. The associated HJB equation is

$$\begin{cases} \dfrac{\partial u}{\partial t} + \langle Ax, Du\rangle + F(x, u, Du, D^2u) = 0 \quad \text{in } (0,T) \times H \\ \\ u(T, x) = g(x) \quad \text{in } H, \end{cases} \tag{2.145}$$

where for $x \in H, s \in \mathbb{R}, p \in H$ and $X \in S(H)$,

$$F(x, s, p, X) = \sup_{\tau \in \Lambda} \left\{ \frac{1}{2} \sum_{i=1}^{d} \langle X \tau_i, \tau_i\rangle + \langle b(\tau), p\rangle - x^+(0)s \right\}.$$

Equation (2.145) is called an infinite dimensional Black-Scholes-Barenblatt (BSB) equation associated to the contingent claim g. In cases of interest in finance the payoff function g is not even C^1 and a notion of a generalized solution is needed. It was studied in [287]) using viscosity solutions. In this context (2.145) has a unique viscosity solution that coincides with the value function provided by the maximization of (2.144). This is discussed in Section 3.10. The results are shown in the space $H = H^1(\mathbb{R}^+)$ which makes the term $x^+(0)$ continuous. One can also investigate the problem in weighted versions of $H^1(\mathbb{R}^+)$.

The problem of pricing derivatives in the HJMM model when the Gaussian noise is replace by a Lévy noise, and the analysis of the associated non-local BSB equation is studied in [427].

2.6.8. Optimal control of stochastic delay equations. In this last example we consider finite dimensional stochastic controlled systems with delay in the state and/or in the control variables. Such control systems arise in many applications (for example in optimal advertising theory, see [238, 239, 330], optimal portfolio management of pension funds, see e.g. [172]) and can be rephrased, using a well known procedure (see e.g. [34] for the deterministic case and [89, 238] for the stochastic case), as infinite dimensional controlled systems without delay. We present two cases: the first is a system with pointwise delay only in the state variable (taken from [223, 172], see also a special case in [238, 239]), while the second one displays delays (pointwise or distributed) both in the state and in the control variable (taken from [238, 239]). We separate the two cases, since they give rise to a different settings with different mathematical difficulties.

2.6.8.1. *Delay in the state variable only.* Let us consider a simple controlled one dimensional linear stochastic differential equation with a delay $r > 0$ in the state variable:

$$\begin{cases} dy(s) = (\beta_0 y(s) + \beta_1 y(s-r) + \alpha(s))\, ds + \sigma dW_0(s), \\ y(t) = x_0, \\ y(t+\theta) = x_1(\theta),\, \theta \in [-r, 0), \end{cases} \tag{2.146}$$

where $\sigma > 0$, $\beta_0, \beta_1 \in \mathbb{R}$ are given constants, W_0 is a one-dimensional standard Brownian motion defined on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and $\mathscr{F}_s^t$ is the augmented filtration generated by W_0. The control $\alpha(\cdot)$ is an $\mathscr{F}_s^t$-progressively measurable process with values in an interval $[0, R]$ for some $R > 0$. We assume that $x_1(\cdot) \in L^2(-r, 0)$. This type of equations is used e.g. in optimal portfolio management of pension funds (see e.g. [172], where the state variable is the wealth of the fund and the control variable is the investment strategy), and in optimal advertising (see e.g. [330, 238, 239], where the state variable is the "goodwill" of a given product and the control is the investment in advertising). Such equations also seem to be relevant for some models arising in studying economic growth in a stochastic environment (see [18, 19] in the deterministic case).

Given three real functions $\varphi_0, h_0, g_0 : \mathbb{R} \to \mathbb{R}$, we consider the problem of minimizing a functional

$$I(t, y_0, y_1; \alpha(\cdot)) = \mathbb{E}\left\{ \int_t^T \left[\varphi_0\left(y\left(s\right)\right) + h_0\left(\alpha\left(s\right)\right)\right] ds + g_0(y(T)) \right\} \tag{2.147}$$

over all control strategies $\alpha(\cdot) \in \mathcal{U}_t$. For example, in the optimal advertising problem the function h_0 represents the cost of advertising while $-\varphi_0$, and $-g_0$ represent the profit coming from the so-called "goodwill" associated to a given product. In such an applied problem it is reasonable to assume the positivity of the state variable (the "goodwill" $y(\cdot) \geq 0$) and of the control variable (the investment $\alpha(\cdot) \geq 0$), and a constraint on the control space (for example $\sup_{s\in[t,T]} \alpha(s) \leq R$ for some $R > 0$ as we have done for other examples).

Existence, uniqueness and properties of solutions of delay equations like (2.146) can be studied either directly (see e.g. [274] Section 5, [410], or the survey [275]) or by introducing an equivalent infinite dimensional formulation. If we follow the former direction, the dynamic programming approach can be used only for special problems where the HJB equation reduces to a finite dimensional differential equation (see [307], one can find similar ideas in [158, 179]), while rephrasing the state equation and therefore the whole optimization problem in infinite dimensions allows to study a larger class of problems. There are different possibilities to rewrite stochastic delay differential equations in the form (2.146) as evolution equations in Hilbert or Banach spaces. Here we present the approach of [89] which allows to rewrite equation (2.146) in the Hilbert space $\mathbb{R} \times L^2(-r, 0)$. Regarding other choices of state spaces we refer for example to [347, 348], where the state space is

$C([-r,0])$, or to the recent paper [189] (see in particular Theorem 2.2), where more general spaces are used.

The setting of [89] is the following. Denote by H the space $\mathbb{R}\times L^2(-r,0)$ and consider the linear operator A_1 on H defined by:

$$\begin{cases} D(A_1)=\left\{\begin{pmatrix} x_0 \\ x_1(\cdot)\end{pmatrix}\in\mathbb{R}\times W^{1,2}(-r,0;\mathbb{R}),\ x_0=x_1(0)\right\} \\ A_1\begin{pmatrix} x_0 \\ x_1(\cdot)\end{pmatrix}=\begin{pmatrix}\beta_0x_0+\beta_1x_1(-r) \\ x_1'(\cdot)\end{pmatrix}.\end{cases}$$

The operator A_1 generates a strongly continuous semigroup $S_1(t)$ on H and, for $z=(z_0,z_1(\cdot))\in H$, $S_1(t)\,z$ can be written in terms of the solution of the linear deterministic delay equation

$$\begin{cases} \dot{y}(t)=\beta_0y(t)+\beta_1y(t-r), \\ y(0)=z_0,\ y(\theta)=z_1(\theta),\ \theta\in[-r,0), \end{cases} \tag{2.148}$$

as follows:

$$S_1(t)\,z=\begin{pmatrix} y(t) \\ y(t+\cdot)\end{pmatrix}\in H,\quad t\ge 0$$

(see [89, 127] and also [32]). Set now $\Xi=\mathbb{R},\Lambda=[0,R]$ for a suitable $R>0$, and define $Q:\Xi\mapsto\Xi$ and $B:\Lambda\to H$, $G:\Xi\to H$ by

$$Qw_0=w_0,\qquad B_1w_0=\begin{pmatrix} w_0 \\ 0\end{pmatrix},\qquad Gw_0=\begin{pmatrix}\sigma w_0 \\ 0\end{pmatrix}.$$

Then, setting $X(s)=(y(s),y(s+\cdot)),a(s)=\alpha(s)$, and $W_Q=W_0$, the controlled stochastic delay equation (2.146) can be rewritten (see again [89, 127]) as the following linear evolution equation in H:

$$\begin{cases} dX(s)=[A_1X(s)+B_1a(s)]\,dt+GdW_Q(s), \\ X(t)=\begin{pmatrix} x_0 \\ x_1\end{pmatrix}:=\begin{pmatrix} y_0 \\ y_1\end{pmatrix}\in H.\end{cases} \tag{2.149}$$

Thanks to Theorem 1.121 the state equation (2.149) admits a unique mild solution (denoted by $X(s;t,x,a(\cdot))$ or simply by $X(s)$), and thus, thanks to Proposition 2.16, Hypotheses 2.11 and 2.12 are satisfied.

Moreover the functional (2.147) can be rewritten as follows. Set

$$\begin{cases} \varphi(x_0,x_1):=\varphi_0(x_0) \\ g(x_0,x_1):=g_0(x_0)\end{cases}$$

so that, for a given initial datum $x\in H$, the functional I becomes

$$J(t,x;a):=\mathbb{E}\left\{\int_t^T[\varphi(X(s))+h_0(a(s))]\,ds+g(X(T))\right\}. \tag{2.150}$$

If φ_0, h_0 and g_0 satisfy proper continuity and growth conditions that ensure Hypothesis 2.23, then the dynamic programming principle holds (see Section 3.6 on this).

The associated Hamilton-Jacobi-Bellman equation is

$$\begin{cases} v_t+\frac{1}{2}\mathrm{Tr}(GG^*D^2v)+\langle Dv,A_1x\rangle+F_0(Dv)+\varphi(x)=0, \\ v(T,x)=g(x),\end{cases} \tag{2.151}$$

where

$$F_0(p) := \inf_{0\le a\le R} \{h_0(a) + \langle p, B_1 a\rangle\} = \inf_{0\le a\le R} \{h_0(a) + p_0 a\}\,. \tag{2.152}$$

Since the second component of B_1 is always zero, the Hamiltonian F_0 only depends on the one dimensional component p_0 (i.e. on the first derivative $D_0 v$ of v with respect to the "present" component). Similarly in the second order term, the fact that the second component of G is zero implies that this term only depends on the second derivative D^2_{00} of v with respect to the "present" component. Thus we can just write $\frac{1}{2}\mathrm{Tr}(GG^* D^2 v) = \frac{1}{2}\gamma^2 D^2_{00} v$.

This kind of equations was studied in [223] by the L^2 approach, where existence of weakly differentiable solutions in Sobolev spaces with respect to a suitable invariant measure μ was proved (see Chapter 5 and, in particular, Section 5.6). Also the BSDE approach, which produces Gâteaux differentiable solutions can be applied here, since the so-called *structure condition* $Im\, B_1 \subset Im\, G$ holds. It is developed, also for ore general equations, first in [209, 337], then in [466, 467], and finally in [464, 465] including also boundary control/noise. See on this Section 6.5. Some results for the viscosity solution approach have been obtained, also for more general cases, in [172, 173, 462, 463], see also the bibliographical notes in Section 3.14.

We also mention, that in the deterministic case, using the fact that the Hamiltonian F_0 only depends on the derivative with respect to the "present", in [175] a regularity result was proved for the viscosity solution of a first order HJB equation of type (2.151) for a case with nonlinear state equation and with state constraints.

2.6.8.2. *Delay in the state and control.* We now consider a stochastic optimal control problem whose state equation has delay in both the state and the control. Such equations are used, for example, to model the evolution of the goodwill stock in advertising models (see e.g. [238, 239]). Suppose we have a controlled stochastic delay differential equation

$$\begin{cases} dy(s) = \left[\beta_0 y(s) + \displaystyle\int_{-r}^0 \beta_1(\xi) y(s+\xi)\,d\xi + \gamma_0\alpha(s) + \int_{-r}^0 \gamma_1(\xi)\alpha(s+\xi)\,d\xi\right] ds \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad +\sigma\, dW_0(s), \quad t\le s\le T, \\ y(t) = y_0, \\ y(t+\xi) = y_1(\xi),\ \alpha(t+\xi) = \delta(\xi),\ \ \xi\in[-r,0], \end{cases} \tag{2.153}$$

where $\sigma > 0$, $\beta_0, \gamma_0 \in \mathbb{R}$ are given real numbers, $\beta_1(\cdot), \gamma_1(\cdot) \in L^2(-r,0)$, and $W_0, \alpha(\cdot)$ are as in the previous subsection. Since $\beta_1(\cdot), \gamma_1(\cdot)$ are functions, we rule out the case of pointwise delay. In fact also the pointwise delay case can be studied, however it gives rise to an unbounded control operator B_2 in the state equation (2.154) below. For the moment we do not consider this case: we will say more about it in the comments after the HJB equation.

The initial data (y_0, y_1, δ) are taken in $\mathbb{R}\times L^2(-r,0)\times L^2(-r,0)$. We again try to minimize the functional I defined by (2.147), over all controls $\alpha(\cdot)\in\mathcal{U}_t$.

The problem can be rewritten in an infinite dimensional setting using a technique which is slightly different from the one of the previous subsection; the results we use are proved in [238] and they generalize these proved in the deterministic setting in [442].

We take, as before $H := \mathbb{R} \times L^2(-r,0)$, $\Xi = R, W_Q = W_0$ and $\Lambda = [0,R]$ for a suitable $R > 0$. We define the operator $A_2 : D(A_2) \subset H \to H$ as follows:

$$\begin{cases} D(A_2) := \left\{x \in H \ : \ x_1 \in W^{1,2}(-r,0), \ x_1(-r) = 0\right\} \\ A_2 : (x_0, x_1(\cdot)) \mapsto \left(\beta_0 x_0 + x_1(0), \beta_1(\cdot)x_0 - \frac{dx_1(\cdot)}{d\xi}\right). \end{cases}$$

Moreover, we define the bounded linear control operator B_2 by

$$\begin{cases} B_2 : \mathbb{R} \to H \\ B_2 : a \mapsto \left(\gamma_0 a, \gamma_1(\cdot)a\right), \end{cases}$$

and the operator $G : \mathbb{R} \to H$ by $G : w_0 \mapsto (\sigma w_0, 0)$, as in the case of delay only in the state. The control variable will remain the same in the new system, so $a(s) := \alpha(s)$. The state variable is called the *structural state* and is defined using the following proposition proved in [238].

PROPOSITION 2.47 *Let $X(\cdot)$ be the mild solution of the abstract evolution equation*

$$\begin{cases} dX(s) = (AX(s) + B_2 a(s))\, dt + G\, dW_Q(s) \\ X(t) = x \in H, \end{cases} \tag{2.154}$$

with arbitrary initial datum $x \in H$ and control $a(\cdot) \in M^2_\mu(t,T;\mathbb{R})$. Then, for $s \geq t$, one has, $\mathbb{P}$-a.s.,

$$X(s) = M(X_0(s), X_0(s+\cdot), a(s+\cdot)),$$

where

$$\begin{cases} M\colon H \times L^2(-r,0) \to H \\ M\colon (x_0, x_1(\cdot), v(\cdot)) \mapsto (x_0, m(\cdot)), \end{cases}$$

($X_0(s)$ is the first component of $X(s)$) and

$$m(\xi) := \int_{-r}^{\xi} \beta_1(\zeta) x_1(\zeta - \xi)\, d\zeta + \int_{-r}^{\xi} \gamma_1(\zeta) v(\zeta - \xi)\, d\zeta.$$

Moreover, let $\{y(s),\ s \geq t\}$ be a continuous solution of the stochastic delay differential equation (2.153), and $X(\cdot)$ be the mild solution of the abstract evolution equation (2.154) with initial condition

$$x = M(y_0, y_1, \delta(\cdot)).$$

Then, for $s \geq t$, one has, $\mathbb{P}$-a.s.,

$$X(s) = M(y(s), y(s+\cdot), a(s+\cdot)),$$

hence $y(s) = X_0(s)$, $\mathbb{P}$-a.s., for all $s \geq 0$.

Using this equivalence result, we can now give a reformulation of our problem in the Hilbert space H. The state equation is (2.154) with initial condition $x := M(y_0, y_1, \delta(\cdot))$ and we denote its mild solution (which exists and is unique thanks to Theorem 1.121) by $X(s) := X(s; t, x, a(\cdot))$. The objective functional to minimize is the same J given by (2.150), where g and φ have the same meaning. Therefore, in this setup, Hypotheses 2.11 and 2.12 are satisfied thanks to Proposition 2.16. Thus, again, if φ_0, h_0 and g_0 satisfy proper continuity and growth conditions, we can ensure that the dynamic programming principle holds.

The Hamilton-Jacobi-Bellman equation in the infinite-dimensional setting is

$$\begin{cases} v_t + \frac{1}{2}\mathrm{Tr}(GG^* D^2 v) + \langle Dv, A_2 x\rangle + \inf_{0\leq a\leq R}\{h_0(a) + \langle Dv, B_2 a\rangle\} = 0, \\ v(T,x) = g(x). \end{cases} \tag{2.155}$$

This kind of HJB equations is more difficult than (2.151) since the so-called structure condition ($Im\, B_2 \subset Im\, G$) is no longer true, and thus it is impossible to use the BSDE approach of [209, 337] and the approach of strong solutions in Sobolev spaces is used in [223]. However in a special case with no delay in the state, a clever variant of the perturbation approach to mild solutions can be applied, see [240].

Concerning a viscosity solution approach we are not aware of any results in the stochastic case. For the deterministic case, please see [178] where also regularity results for viscosity solutions are proved (see also [175] for such results).

Finally, we remark that (as it can be seen e.g. in [238]) in the case of pointwise delay (i.e. when β_1 is the Dirac delta at $-r$), the operator B_2 above is unbounded. This unboundedness is similar to the one arising in boundary control problems, and up to now HJB equations of this kind have been investigated only in a special case in [240].

2.7. Bibliographical notes

The stochastic optimal control problem introduced in Section 2.1 is an abstract infinite dimensional version of problems studied in the literature. We refer to [38, 47, 149, 194, 195, 294, 317, 318, 349, 357, 364, 382, 449] for the finite dimensional theory.

For deterministic optimal control problems and their connection with HJB equations the reader may consult [30, 40, 41, 69, 95, 96, 316, 455] and the books [23, 312] for the infinite dimensional case. Some aspects of the theory of stochastic optimal control in infinite dimensions and second-order HJB equations can be found in the books [129, 364].

We present the optimal control problem in its weak (Section 2.1.2) and strong (Section 2.1.1) formulations. The two distinct forms appeared already in the sixties, in the early days of the studies of finite-dimensional stochastic optimal control problems (see e.g. [191] and [303]); we follow the terminology of [449]. We recall that for us the "weak" in "the weak formulation" is referred only to the fact that the generalized reference probability spaces vary with the controls and not to the concept of solution that in this context is always strong in the probabilistic sense (see Remark 2.4). The notion of weak formulation is also different from that used in [197] where the word "weak" is meant in the sense of the convex duality.

In Section 2.2 and more precisely in Subsection 2.2.1 we introduce a third formulation that we use to prove the DPP. We can call this third setup *weak DPP* formulation. In this framework, as in the weak formulation, we allow the probability spaces and Q-Wiener processes W_Q to vary but we only consider the (augmented) filtration generated by the Q-Wiener processes. Thus the difference is that we pass from generalized reference probability spaces (Definition 1.95) to reference probability spaces (Definition 2.7). Other formulations of stochastic optimal control problems have been proposed in the literature with various notions of control processes. Markov (feedback) controls, i.e. controls of the form $a(t, X(t))$, where $X(t)$ is the state of the system at time t, have been considered e.g. in [47, 194, 195, 294, 304]. So called *natural strategies* i.e. controls that can be expressed at time t as functions of the state trajectory up to time t are considered in [294] where it is also shown that, under suitable hypotheses, the value function of the problem for natural strategies equals the value function of the problem in the strong formulation (Theorem 7, page 132 of [294]). Relaxed controls have been considered (see e.g. [47, 153, 254, 301]) mostly to prove existence of optimal controls. Our formulations of control problems follow most closely these of [449]. In Theorem 2.22 we show that the weak DPP formulation and the strong formulation if a reference probability space is used are equivalent in the sense that the problems in the two forms have the same value function. For similar results in the finite-dimensional case one can see [155, 195, 294]. Nisio in [364] proves that the weak formulations using the reference and the generalized reference probability spaces give the same value functions assuming that the control set is a convex subset of $\mathbb{R}^q$.

The DPP proved in Section 2.3 (Theorem 2.24) is very abstract and general. We follow to a large extent the strategy from [449]. The proof uses the continuity of the value function in the spatial variable, however this assumption can be relaxed with very little change in the proof (see e.g. [217]). Theorem 3.70, Section 3.6 (next chapter) contains a version of the DPP for mild solutions in the formulation with stopping times when the value function is continuous. The DPP is often considered a standard result, however we included complete proofs since even in finite dimensions it is very technical and many of the proofs available in the literature miss a lot of details.

Several other approaches to the proof of the DPP are available in the literature. Krylov [294] uses approximation of controls by step controls. A PDE based proof is in Fleming and Soner [195] where the DPP is first proved for a uniformly parabolic case where the HJB equation has a smooth solution, and then the value function is approximated by smooth value functions solving uniformly parabolic HJB equations. The proof in [47] uses Markov controls. Nisio [357] uses approximations with switching controls at binary times and a reduction to the canonical reference probability space, while the proof in [356] uses so called non-anticipative controls and approximations by controls with continuous trajectories. A proof based on discrete time dynamic programming principle and approximation by switching controls and is presented in [364]. In [364] the DPP is proved for the weak formulation of control problem from Section 2.1.2. The proof in [349] is based on a reduction to the canonical reference probability space. In [382] a sketch of the proof is given for a measurable value function, which however omits delicate measurability issues. Soner and Touzi [416] use deep measurable selection theorems to show the DPP for stochastic target problems without continuity assumptions on the value function. In a recent paper [155] El Karoui and Tan prove the DPP under general assumptions. Other proofs (see e.g. [153, 254, 301]) use relaxed controls and compactness of the set of admissible controls. In [244] the authors adapt to the infinite dimensional case the arguments used in [196], where the DPP was shown for a two player, zero sum stochastic differential game in finite dimensions, to prove the DPP for a control problem for stochastic Navier-Stokes equations in the canonical reference probability space.

A different approach to the DPP has been introduced, for the finite dimensional case, in [51]. In that paper the authors introduce the notion of *weak dynamic programming*: roughly speaking, instead of proving a result similar to (2.25) they prove the following, weaker, fact: for any couple of continuous test functions ϕ and ψ such that $\phi \le V \le \psi$,

$$\begin{aligned}\inf_{a(\cdot)\in\mathcal{U}_t^{\phi}} \mathbb{E}\left[\int_t^{\eta} e^{-\int_t^s c(X(\tau))d\tau} l\left(s, X\left(s\right), a\left(s\right)\right) ds + e^{-\int_t^{\eta} c(X(\tau))d\tau}\phi\left(\eta, X\left(\eta\right)\right)\right]& \\ \le V\left(t,x\right)& \\ \le \inf_{a(\cdot)\in\mathcal{U}_t^{\psi}} \mathbb{E}\left[\int_t^{\eta} e^{-\int_t^s c(X(\tau))d\tau} l\left(s, X\left(s\right), a\left(s\right)\right) ds + e^{-\int_t^{\eta} c(X(\tau))d\tau}\psi\left(\eta, X\left(\eta\right)\right)\right].&\end{aligned} \tag{2.156}$$

In this way the difficulties due to the possible lack of continuity of the value function V are avoided because the condition deals with test functions that are continuous. This formulation is of course tailored to the study of viscosity solutions of HJB equations which are defined in terms of regular test functions (see Chapter 3). The weak dynamic programming approach introduced in [51] has been generalized to the case of expectation constraints and state constraints in [50], where an abstract dynamic programming result was stated. The weak dynamic programming was also used for a class of finite-dimensional impulsive problems in [49].

The verification theorem and construction of optimal feedback controls for a smooth value function presented in Section 2.5 follow similar standard results for the finite-dimensional case which can be found for instance in Chapter 4 of [194], Chapter 3 of [195] or in Chapter 5 of [449]. In infinite dimensional Hilbert spaces, for the case of quadratic Hamiltonian, the reader is referred to Chapter 13 of [129]. When the value function does not satisfy the strong regularity conditions of Section 2.5 ($C^{1,2}$ regularity and the derivative in the domain of A^*), only a few specific results are available:

- There are no results in infinite dimensions for viscosity solutions that are only continuous. A finite dimensional verification theorem can be found in [246, 247, 449]. In the deterministic case, some verification results for a Hilbert space case can be found in [66, 166, 312].
- For mild solutions in spaces of continuous functions there are several contributions [79, 81, 112, 113, 115, 231, 232, 235, 238, 239, 241, 333, 334], some of which will be discussed in Chapter 4.
- For mild solutions in the space of L^2 functions we refer to [3, 4, 223, 226], see also Chapter 5.
- For optimal synthesis obtained via backward stochastic differential equations the reader is referred to [205, 209, 210, 211, 212] and to Chapter 6.

A different approach to optimal control problems that is not developed in the book is the use of maximum principle; it is closely related to the study of backward stochastic differential equations (BSDEs). A general result for the finite-dimensional case is given in [376], see also [449]. A generalization to problems with noises with jumps is addressed in [433, 434], where the authors first characterize the adjoint process of the second variation as the solution of a BSDE in the Hilbert space of Hilbert Schmidt operators.

In infinite dimensions the problem was initially studied in [33] and [263] for the case of diffusion independent of the controls, and in [469] for a problem with linear state equation and cost functional. Recently, thanks to the developments in the study of backward stochastic differential equations in infinite dimensions, new results on maximum principle for stochastic infinite dimensional problems appeared. In [141, 142, 207] the second variation is characterized as a certain stochastic bilinear form defined on $L^4(\Omega; H)$, while in [322] a general case when the coefficients are Fréchet-differentiable (twice for non-convex control domain) is treated and the second variation is characterized as a solution of a BSDE "in the sense of transposition". The approach of [141, 142, 207] is used in [208] where regularity conditions on the coefficients are weakened to study a large class of optimal control problems driven by stochastic PDE of parabolic type on a bounded open set of $\mathbb{R}^n$. Other results for specific classes of equations include [253], for a one-dimensional heat equation with noise and control on the boundary, and [367], for a class of problems with delay state equation (both with distributed and discrete delay). The papers [142] and [367] also include, respectively, an unbounded diffusion term and Lévy noise. In general the maximum principle approach only gives necessary conditions for optimality, however under suitable convexity assumptions also sufficiency can be proved. Such results for finite dimensional systems can be found in [16, 449, 470]. A sufficiency result for a class of infinite dimensional systems is proved in [343], while a sufficient condition for certain delay systems with diffusion independent of the controls is characterized in [266].

CHAPTER 3

Viscosity solutions

This chapter is devoted to the theory of viscosity solutions of Hamilton-Jacobi-Bellman equations in Hilbert spaces. At its core is the notion of the so called B-continuous viscosity solution which was introduced for first order equations by M. G. Crandall and P. L. Lions in [103, 104] and later extended to second order equations in [422]. The theory applies to fully nonlinear equations with various unbounded terms. This is its main advantage over the notions of mild and strong solutions discussed in Chapter 4, mild solutions in L^2 spaces discussed in Chapter 5 and the BSDE techniques of Chapter 6. After the introduction of the core theory we discuss several special cases which require various adjustments in the definition of viscosity solution. The material of the chapter is arranged in the following way:

- In Section 3.1 we introduce the notion of B-continuity, the spaces $H_{-\alpha}$ defined by a strictly positive self-adjoint operator B, and we present several estimates involving $|\cdot|_{-1}$ norms for solutions of deterministic and stochastic evolution equations. We also discuss a smooth perturbed optimization principle in Hilbert spaces.
- In Section 3.2 we present a maximum principle for B-upper semi-continuous functions in Hilbert spaces. This is a key technical result needed in the proofs of uniqueness of viscosity solutions.
- In Section 3.3 we introduce the definition of viscosity solution and in Section 3.4 we discuss basic convergence properties of viscosity solutions.
- Section 3.5 is devoted to uniqueness of viscosity solutions. We prove several comparison theorems for degenerate parabolic and elliptic equations.
- In Sections 3.6 and 3.7 we present results on existence of viscosity solutions. In Section 3.6 we study properties of value functions of stochastic optimal control problems and prove that they are viscosity solutions of the associated HJB equations. In Section 3.7 we discuss how to obtain existence of viscosity solutions for more general equations, for instance of Isaacs type, by the method of finite dimensional approximations.
- In Section 3.9 another method to prove existence of viscosity solutions, Perron's method, is presented. In this section we also explain how in certain cases the method of half-relaxed limits of Barles-Perthame can be adapted to viscosity solutions in Hilbert spaces. A classical limiting problem of singular perturbations is discussed in Section 3.8.
- In Section 3.10 we explain how the theory redof viscosity solutions is applied to the infinite dimensional Black-Scholes-Barenblatt equation originating in the theory of bond markets.
- Sections 3.11, 3.12 and 3.13 discuss three special cases, the HJB equation related to the optimal control of Duncan-Mortensen-Zakai equation, the HJB equation for a boundary optimal control problem and the HJB equation for optimal control of stochastic Navier-Stokes equations. These cases require modifications of the definition of viscosity

solution. We explain how the basic theory of Sections 3.5 and 3.6 can be extended and adapted to equations containing special unbounded terms.

Throughout this chapter H is a real, separable Hilbert space with inner product $\langle\cdot,\cdot\rangle$ and norm $|\cdot|$. We remind that we identify H with its dual. We denote by $S(H)$ the set of bounded, self-adjoint operators on H.

3.1. Preliminary results

3.1.1. B-continuity and weak and strong B-conditions.

DEFINITION 3.1 *Given a strictly positive $B \in S(H)$ and $\alpha > 0$, we define the space $H_{-\alpha}$ as the completion of H with respect to the norm*

$$|x|^2_{-\alpha} := \langle B^\alpha x, x\rangle .$$

The strict positivity of B ensures that the operator $B^{\alpha/2}$ extends to an isometry of $H_{-\alpha}$ onto H that we denote again by $B^{\alpha/2}$. $H_{-\alpha}$ is a Hilbert space when endowed with the inner product induced by $B^{\alpha/2}$:

$$\langle x, y\rangle_{-\alpha} := \left\langle B^{\alpha/2}x, B^{\alpha/2}y\right\rangle .$$

DEFINITION 3.2 *If B, α are as in Definition 3.1, we denote by H_α the space $H_\alpha := B^{\alpha/2}(H)$ endowed with the Hilbert space structure characterized by the following inner product:*

$$\langle x, y\rangle_{\alpha} := \left\langle B^{-\alpha/2}x, B^{-\alpha/2}y\right\rangle .$$

Thanks to the strict positivity of B, $B^{-\alpha/2}\colon H_\alpha \to H$ is an isometry onto H; H_α can be identified with the dual of $H_{-\alpha}$.

Of course, even if not explicitly emphasized by the notation, the spaces H_α depend on the choice of B.

We will often use a notion of continuity, called B-continuity, which is stronger than the usual continuity and weaker than weak sequential continuity.

DEFINITION 3.3 (B-upper/lower semicontinuity) *Let $B \in S(H)$ be a strictly positive operator on H. Given $I \subseteq \mathbb{R}$ and $U \subseteq H$, we say that a function $u\colon I\times U \to \mathbb{R}\cup\{\pm\infty\}$ is B-upper semicontinuous (respectively, B-lower semicontinuous) if, for any $\{t_n\}_{n\in\mathbb{N}} \subseteq I$ and $\{x_n\}_{n\in\mathbb{N}} \subseteq U$, such that $t_n \to t \in I$, $x_n \rightharpoonup x \in U$ and $Bx_n \to Bx$ as $n \to \infty$, we have*

$$\limsup_{n\to\infty} u(t_n, x_n) \le u(t,x) \qquad (\textit{respectively,} \liminf_{n\to\infty} u(t_n,x_n) \ge u(t,x)\,).$$

DEFINITION 3.4 (B-continuity) *Given B, I and U as in Definition 3.3, we say that a function $u\colon I\times U \to \mathbb{R}$ is B-continuous if it is both B-upper semicontinuous and B-lower semicontinuous.*

REMARK 3.5 It is easy to see that one gets the same definition of B-upper/lower-semicontinuity if the condition $x_n \rightharpoonup x \in U$ in Definition 3.3 is replaced by the requirement that $\{x_n\}_{n\in\mathbb{N}}$ is bounded and $x \in U$. ■

LEMMA 3.6 *Let B be as in Definition 3.3. Then:*

(i) *If B is compact then u is B-upper semi continuous (respectively, B-lower semi continuous, B-continuous) if and only if u is weakly sequentially upper semi continuous (respectively, weakly sequentially lower semi continuous, weakly sequentially continuous).*

(ii) *Let $\alpha > 0$. Then u is B-upper semi continuous (respectively, B-lower semi continuous, B-continuous) if and only if u is B^α-upper semi continuous (respectively, B^α-lower semi continuous, B^α-continuous).*

(iii) *Let U be weakly sequentially closed, and $\alpha > 0$. Then u is B-continuous on $I \times U$ if and only if u is continuous in the $|\cdot|\times|\cdot|_{-\alpha}$ norm on bounded subsets of $I \times U$. If B is compact and $I = [a, b]$, then u is B-continuous on $[a, b] \times U$ if and only if u is uniformly continuous in the $|\cdot| \times |\cdot|_{-\alpha}$ norm on $[a, b] \times (U \cap B_R)$ for every $R > 0$. Finally, if u is weakly sequentially continuous on $[a, b] \times U$ then u is uniformly continuous in the $|\cdot| \times |\cdot|_{-\alpha}$ norm on $[a, b] \times (U \cap B_R)$ for every $R > 0$.*

(iv) *Let $B_1, B_2 \in S(H)$ be two strictly positive operators on H such that $B_1(H) = B_2(H)$ and $\{x_n\}_{n\in\mathbb{N}} \subseteq H$. Then $B_1x_n \to B_1x$ if and only if $B_2x_n \to B_2x$. In particular the notions of B_1-continuity and B_2-continuity are equivalent.*

Proof. Part (i) is obvious.

(ii) We show that, for any $\alpha \geq \beta > 0$, u is B^α-continuous (resp. B^α-lower semi continuous, B^α-upper semi continuous) if and only if it is B^β-continuous (resp. B^β-lower semi continuous, B^β-upper semi continuous). To show this fact it is enough to prove that for a given weakly convergent sequence $x_n \rightharpoonup x \in H$ we have that $|B^\alpha(x_n - x)| \to 0$ if and only if $|B^\beta(x_n - x)| \to 0$. Since $\alpha \geq \beta$ the "if" part is obvious. For the "only if" part, assume that $|B^\alpha(x_n - x)| \to 0$ and observe that, since x_n is weakly convergent and hence bounded,

$$|B^{\alpha/2}(x_n - x)|^2 = \langle x_n, B^\alpha(x_n - x)\rangle - \langle x, B^\alpha(x_n - x)\rangle \to 0.$$

So, if $\alpha/2 \leq \beta$, this fact and the "if" part allow to conclude the proof, otherwise one can conclude iterating the argument.

(iii) The first statement follows from (ii). The only nontrivial statement of the second claim is the "only if" part of it. So let B be compact. From (ii) we can assume $\alpha = 2$. Assume by contradiction that, for some $\varepsilon > 0$, there exist two sequences (t_n, x_n) and (s_n, y_n) in $[a, b] \times (U \cap B_R)$ s.t.

$$|t_n - s_n| + |x_n - y_n|_{-2} \to 0 \text{ and } |u(t_n, x_n) - u(s_n, y_n)| > \epsilon. \tag{3.1}$$

Since $[a, b] \times (U \cap B_R)$ is weakly sequentially compact we can assume that $x_n \rightharpoonup x$ and $y_n \rightharpoonup y$ for some $x, y \in U \cap B_R$ and that $t_n, s_n \to s$ for some $s \in [a, b]$. So we have $B(x_n - y_n) \to 0$ and $(x_n - y_n) \rightharpoonup (x - y)$ and thus (since the graph of a continuous operator is weakly closed), $B(x - y) = 0$ which implies that $x = y$. Since B is compact we also have $Bx_n \to Bx$ and $By_n \to By = Bx$. So, since u is B-continuous, $u(t_n, x_n) \to u(s, x)$ and $u(s_n, y_n) \to u(s, x)$ and this contradicts (3.1). If the third claim is not true then again there must exist sequences (t_n, x_n) and (s_n, y_n) s.t. (3.1) holds. But then again, up to a subsequence, $t_n, s_n \to s$ for some $s \in [a, b]$ and $x_n, y_n \rightharpoonup x$ for some $x \in U \cap B_R$ and this, together with the weak sequential continuity of u, contradicts (3.1).

(iv). Let $B_1x_n \to B_1x$ as $n \to \infty$. It follows easily from the closed graph theorem theorem that $B_1^{-1}B_2$ is bounded. Thus $B_2B_1^{-1} = (B_1^{-1}B_2)^*$ on $B_1(H)$ and $(B_1^{-1}B_2)^*$ is a bounded operator. Therefore

$$B_2x_n = B_2B_1^{-1}B_1x_n = (B_1^{-1}B_2)^*B_1x_n \to (B_1^{-1}B_2)^*B_1x = B_2B_1^{-1}B_1x = B_2x.$$

The other implication is proved similarly. □

Definition 3.7 (B-closed set) *We will say that a set $U \subset H$ is B-closed if whenever $x_n \in U, x_n \rightharpoonup x, Bx_n \to Bx$ then $x \in U$.*

Remark 3.8 Every weakly sequentially closed subset of H is B-closed, in particular every convex closed subset of H is B-closed. ■

The following weak and strong B-conditions were introduced in [103, 104].

DEFINITION 3.9 (Weak B-condition) *Let A be a linear, densely defined, closed operator in H. We say that an operator $B \in \mathcal{L}(H)$ satisfies the* weak B-condition *for A if B is strictly positive, self-adjoint, $A^*B \in \mathcal{L}(H)$, and*

$$-A^*B + c_0 B \geq 0 \quad \textit{for some } c_0 \geq 0. \tag{3.2}$$

DEFINITION 3.10 (Strong B-condition) *Let A be a linear, densely defined, closed operator in H. We say that an operator $B \in \mathcal{L}(H)$ satisfies the* strong B-condition *for A if B is strictly positive, self-adjoint, $A^*B \in \mathcal{L}(H)$, and*

$$-A^*B + c_0 B \geq I \quad \textit{for some } c_0 \geq 0. \tag{3.3}$$

It is well known that if A is a densely defined closed operator in H then the operator $B = (I + AA^*)^{-1/2}$ is bounded, strictly positive, self-adjoint and $A^*B \in \mathcal{L}(H)$. The strong and weak B-conditions require a little more. We will apply them when the operator A is maximal dissipative. The following result has been shown in [396].

THEOREM 3.11 *If A is a linear, densely defined maximal dissipative operator in H then the weak B condition is satisfied with $B = ((\mu I - A)(\mu I - A)^*)^{-1/2}$ and $c_0 = \mu$, where $\mu \geq 0$ is any constant such that $\mu I - A^* \geq \delta I$ for some $\delta > 0$.*

PROOF. Denote $C = (\mu I - A)(\mu I - A)^*$. By our assumptions, C^{-1} exists and $C^{-1} \in \mathcal{L}(H)$. It is also easy to see that $C = C^* > 0$. We set $B = C^{-1/2}$. Then $B = B^* > 0$, and we have, for $x \in H$,

$$|Bx|^2 = \langle \left((\mu I - A)^{-1}\right)^* (\mu I - A)^{-1} x, x\rangle = \left|(\mu I - A)^{-1} x\right|^2.$$

Therefore, by Proposition B.2-(i) (see also [130], Proposition B.1, p. 429 or [455], Theorem 2.2, p. 208), it follows that $R(B) = R\left(\left((\mu I - A)^{-1}\right)^*\right) = R\left((\mu I - A^*)^{-1}\right) = D(A^*)$.

Denote $S = (\mu I - A)^* B$. Then $S \in \mathcal{L}(H)$ and it is unitary. In fact SB^{-1} is the polar decomposition of $\mu I - A^*$. It remains to show that $S \geq 0$.

To this end we complexify the space and the operators. Let $H_c = \{\tilde{x} = x + iy : x, y \in H\}$ with standard operations $(x + iy) + (z + iw) = (x + z) + i(y + w)$, $(a + ib)(x + iy) = (ax - by) + i(bx + ay)$ and the inner product $\langle (x + iy), (z + iw)\rangle_c = \langle x, z\rangle + \langle y, w\rangle + i\langle y, z\rangle - i\langle x, w\rangle$. An operator T in H is complexified by setting $T_c(x + iy) = Tx + iTy$ and then $(T_c)^* = (T^*)_c$. It is easy to see that we still have $\mu I_c - A_c^* \geq 0$ in the sense that $\mathrm{Re}\langle(\mu I_c - A_c^*)\tilde{x}, \tilde{x}\rangle_c \geq 0$, $B_c = B_c^* > 0$ and S_c is unitary (and thus normal). It is enough to show that $S_c \geq 0$.

Suppose that S_c is not nonnegative. Since S_c is normal it then follows from the spectral representation theorem that there is a nontrivial closed subspace K of H_c which is invariant for S_c and S_c is strongly dissipative on K, i.e. $\mathrm{Re}\langle S_c \tilde{x}, \tilde{x}\rangle_c \leq -\nu |\tilde{x}|_c^2$ for some $\nu > 0$. Let P_K be the orthogonal projection onto K. Then the operator $P_K B_c : K \to K$ is self adjoint and strictly positive. We choose $\lambda > 0$ in the spectrum of $P_K B_c$. Then there exists $\tilde{y} \neq 0$ in K such that

$$|P_K B_c \tilde{y} - \lambda \tilde{y}|_c \leq \frac{\lambda \nu}{2\|S_c\|} |\tilde{y}|_c.$$

We set $\tilde{x} = B_c \tilde{y}$. Then

$$\begin{aligned}\langle \mu I_c - A_c^* \tilde{x}, \tilde{x}\rangle_c &= \langle S_c \tilde{y}, \tilde{x}\rangle_c = \langle S_c \tilde{y}, P_K \tilde{x}\rangle_c = \langle S_c \tilde{y}, P_K B_c \tilde{y}\rangle_c \\ &= \lambda \langle S_c \tilde{y}, \tilde{y}\rangle_c + \langle S_c \tilde{y}, P_K B_c \tilde{y} - \lambda \tilde{y}\rangle_c\end{aligned}$$

Taking the real part of the above relation and using that $\mathrm{Re}\langle(\mu I_c - A_c^*)\tilde{x}, \tilde{x}\rangle_c \geq 0$ we obtain

$$0 \leq -\lambda\nu|\tilde{y}|_c^2 + \|S_c\||\tilde{y}|_c \frac{\lambda\nu}{2\|S_c\|}|\tilde{y}|_c = -\frac{\lambda\nu}{2}|\tilde{y}|_c^2$$

which is a contradiction. Therefore $S_c \geq 0$ and thus $S \geq 0$. In fact one can show (see [396]) that $S > 0$. □

The following are two concrete examples of operators satisfying the weak B-condition:

EXAMPLE 3.12 If the operator A is maximal dissipative and skew-adjoint, i.e. $A^* = -A$, the above implies that we can take $B = (\mu I - A^2)^{-1/2}$ for every $\mu > 0$. However, in such case a compact B cannot satisfy the strong B-condition, since if it did, then for every eigenvalue λ of B with an eigenvector e we would have $|e|^2 \leq \langle(-A^*B + c_0B)e, e\rangle = \lambda\langle(A + c_0I)e, e\rangle \leq \lambda c_0|e|^2$ which is impossible since the eigenvalues accumulate at zero. ■

EXAMPLE 3.13 (Operators coming from hyperbolic equations) Let A be a maximal dissipative, self-adjoint operator in a Hilbert space H with a bounded inverse. It is then well known (see e.g. A.5.4 in [130]) that the operator

$$\mathcal{D}(\mathcal{A}) = \begin{pmatrix} D(A) \\ \times \\ D((-A)^{1/2}) \end{pmatrix}, \qquad \mathcal{A} = \begin{pmatrix} 0 & I \\ A & 0 \end{pmatrix},$$

is maximal dissipative in the Hilbert space $\mathcal{H} = \begin{pmatrix} D((-A)^{1/2}) \\ \times \\ H \end{pmatrix}$, equipped with the following "energy" type inner product

$$\left\langle \begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} \right\rangle_{\mathcal{H}} = \langle(-A)^{1/2}u, (-A)^{1/2}\bar{u}\rangle_H + \langle v, \bar{v}\rangle_H, \qquad \begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} \bar{u} \\ \bar{v} \end{pmatrix} \in \mathcal{H}.$$

Moreover, $\mathcal{A}^* = -\mathcal{A}$.

It is easy to check that the operator

$$\mathcal{B} = \begin{pmatrix} (-A)^{-1/2} & 0 \\ 0 & (-A)^{-1/2} \end{pmatrix}$$

is bounded, positive, self-adjoint on $\mathcal{H}$, and such that $\mathcal{A}^*\mathcal{B}$ is bounded and the weak B-condition holds with constant $c_0 = 0$. In fact

$$\left\langle -\mathcal{A}^*\mathcal{B}\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} u \\ v \end{pmatrix} \right\rangle_{\mathcal{H}} = 0.$$

Moreover we have

$$\left| \begin{pmatrix} u \\ v \end{pmatrix} \right|_{-1} = \left(|(-A)^{1/4}u|^2 + |(-A)^{-1/4}v|^2 \right)^{1/2}.$$

■

Let us now examine the strong B-condition in some examples.

EXAMPLE 3.14 If A is maximal dissipative and self-adjoint in H, it satisfies the strong B-condition with $B = (I - A)^{-1}$ and $c_0 = 1$. ■

EXAMPLE 3.15 Suppose now that A_0 be a densely defined, closed operator in H which satisfies the strong B-condition for some operator B_0 and constant c_0. Let

A_1 be another densely defined, closed operator in H such that $A_1^* B_0$ is bounded and

$$-A_1^* B_0 + c_1 B_0 \geq -\nu I \tag{3.4}$$

for some $\nu \in (0,1)$ and some constant c_1. It is then clear that $A = A_0 + A_1$ satisfies the strong B-condition with $B = (1/(1-\nu))B_0$ and the new constant $c := c_0 + c_1$. Obviously (3.4) holds if $\|A_1^* B_0\| < 1$. Also rather standard arguments show that (3.4) is satisfied for every $\nu \in (0,1)$ and some constant c_1 if $A_1^* B_0$ is compact. To see this let $\{e_1, e_2, ...\}$ be an orthonormal basis of H. For $N \geq 1$ we denote $H_N = \text{span}\{e_1, ..., e_N\}$, P_N be the orthogonal projection onto H_N, and $Q_N := I - P_N$. For $x \in H$ we will write $x_N := P_N x$, $x_N^\perp := Q_N x$. Since $A_1^* B_0$ is compact, there is $N_1 \geq 1$ such that $\|A_1^* B_0 - P_{N_1} A_1^* B_0 P_{N_1}\| \leq \nu/2$. Therefore it is enough to prove that there is c_1 such that $-P_{N_1} A_1^* B_0 P_{N_1} + c_1 B \geq -\nu/2I$ which, since $\langle P_{N_1} A_1^* B_0 P_{N_1} x, x\rangle \leq C|x_{N_1}|^2$ will be true if

$$C|x_{N_1}|^2 \leq c_1 \langle B_0 x, x\rangle + \nu|x|^2/2,$$

i.e. if

$$C \leq c_1 \langle B_0 x, x\rangle + \nu|x|^2/2 \quad \text{for any } x \text{ such that } |x_{N_1}| = 1.$$

The above is certainly satisfied if $|x_{N_1}^\perp| \geq (2C/\nu)^{1/2}$. Moreover, it is easy to see that

$$\inf_{\{x : |x_{N_1}| = 1, |x_{N_1}^\perp| \leq (2C/\nu)^{1/2}\}} \langle B_0 x, x\rangle = \delta_{N_1} > 0.$$

Thus it is enough to take $c_1 = C/\delta_{N_1}$. ■

EXAMPLE 3.16 (*Operators coming from elliptic equations*) Let $\mathcal{O}$ be a bounded (regular enough) domain in $\mathbb{R}^n$. Let

$$\begin{cases} A := \sum_{i,j}^n \partial_i (a_{ij} \partial_j) + \sum_i^n b_i \partial_i + c \\ D(A) := H_0^1(\mathcal{O}) \cap H^2(\mathcal{O}), \end{cases}$$

where $a_{ij} = a_{ji}, b_i, c \in L^\infty(\mathcal{O})$ for $i, j \in \{1, .., n\}$, and there exists $\theta > 0$ such that

$$\sum_{i,j}^n a_{ij} \xi_i \xi_j \geq \theta |\xi|^2 \quad \forall \xi \in \mathbb{R}^n$$

a.e. in $\mathcal{O}$. We observe that if A_0 is the operator A with $c = b_i = 0$, $i = 1, ..., n$, then A_0 is maximal dissipative and self-adjoint in $H = L^2(\mathcal{O})$ and the strong B-condition holds for A_0 with $B_0 = (I - A_0)^{-1}$ and $c_0 = 1$. Moreover B_0 is compact as an operator from $L^2(\mathcal{O})$ to $H_0^1(\mathcal{O})$ and thus, if $A_1 = A - A_0$, it follows that $A_1^* B_0$ is compact. Thus the strong B-condition is satisfied for A with $B = \lambda B_0$ for some constant λ.

If in addition $a_{ij} \in W^{1,\infty}(\mathcal{O}), b_i = 0, i, j = 1, ..., n$ one can also take $B_0 = \lambda(\hat{A})^{-1}$ above, where

$$\begin{cases} \hat{A} f := -\Delta f \\ D(\hat{A}) := H_0^1(\mathcal{O}) \cap H^2(\mathcal{O}) \end{cases}$$

for λ big enough. This follows from an application of the Sobolevskii inequality (see for instance Theorem 1.1 in [315], see also [305, 414]). ■

LEMMA 3.17 *Let $B \in S(H)$ be a strictly positive operator on H and A be a linear, densely defined, maximal dissipative operator. Then:*

(i) *If $D(A^*) = D(B^{-1})$, then the operator $S = -A^* B + c_0 B$ is invertible for any $c_0 > 0$, and $S^{-1} \in \mathcal{L}(H)$.*

(ii) *If B satisfies the strong B-condition for A, then $D(A^*) = D(B^{-1})$.*

PROOF. (i) The statement is obvious since $B^{-1}(-A^* + c_0 I)^{-1}$ is bounded and it is the inverse of S.

(ii) Let S be defined as in part (i) but with c_0 being the constant from the strong B-condition for A. Since S is bounded and $S \geq I$, S^{-1} exists and it is bounded. Moreover we have $B = (-A^* + c_0 I)^{-1} S$ which, by the invertibility of S implies that $D(B^{-1}) = R(B) = R(-A^* + c_0 I)^{-1} = D(A^*)$. □

We refer the reader to [396] for an abstract condition involving interpolation spaces which ensures that the strong B-condition is satisfied and to [103, 104] for other comments about B-continuity and strong and weak B-conditions.

3.1.2. Estimates for solutions of stochastic differential equations. Let $T > 0$. Let A be a linear, densely defined, maximal dissipative operator in H, and $Q \in \mathcal{L}^+(\Xi)$. Let $(\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s\in[0,T]}, \mathbb{P}, W_Q)$ be a generalized reference probability space. Let Λ be a Polish space. Let $b\colon [0,T] \times H \times \Lambda \to H$ be $\mathcal{B}([0,T]) \otimes \mathcal{B}(H) \otimes \mathcal{B}(\Lambda)/\mathcal{B}(H)$ measurable, and $\sigma\colon [0,T] \times H \times \Lambda \to \mathcal{L}_2(\Xi_0, H)$ be $\mathcal{B}([0,T]) \otimes \mathcal{B}(H) \otimes \mathcal{B}(\Lambda)/\mathcal{B}(\mathcal{L}_2(\Xi_0, H))$ measurable. Let $a(\cdot) : [0,T] \times \Omega \to \Lambda$ be $\mathscr{F}_s$-progressively measurable. For $x \in H$ we consider the following SDE

$$\begin{cases} dX(s) = (AX(s) + b(s, X(s), a(s)))\, dt + \sigma(s, X(s), a(s)) dW_Q(s) \\ X(0) = x \end{cases} \tag{3.5}$$

and its approximation

$$\begin{cases} dX^n(s) = (A_n X^n(s) + b(s, X^n(s), a(s)))\, dt + \sigma(s, X^n(s), a(s)) dW_Q(s) \\ X^n(0) = x, \end{cases} \tag{3.6}$$

where A_n is the Yosida approximation of A defined in (B.8). The approximating equation (3.6) was already introduced in Chapter 1. Here we discuss some more specific results that will be needed in later chapters.

Let $C \geq 0$ and $\gamma \in [0,1]$. We will make use of the following assumptions.

$$|b(s,x,a) - b(s,y,a)| \leq C|x-y| \qquad \forall x,y \in H,\ s \in [0,T],\ a \in \Lambda, \tag{3.7}$$

$$\|\sigma(s,x,a) - \sigma(s,y,a)\|_{\mathcal{L}_2(U_0;H)} \leq C|x-y| \qquad \forall x,y \in H,\ s \in [0,T],\ a \in \Lambda, \tag{3.8}$$

$$|b(s,x,a)| \leq C(1+|x|) \qquad \forall x \in H,\ s \in [0,T],\ a \in \Lambda, \tag{3.9}$$

$$\|\sigma(s,x,a)\|_{\mathcal{L}_2(U_0;H)} \leq C(1+|x|^\gamma) \qquad \forall x \in H,\ s \in [0,T],\ a \in \Lambda. \tag{3.10}$$

Recall that, thanks to Theorem 1.121, assumptions (3.7), (3.8), (3.9) and (3.10), ensure the existence of unique mild solutions $X(\cdot)$ and $X^n(\cdot)$ of (3.5) and (3.6).

PROPOSITION 3.18 *Let $T > 0$ and $\gamma \in [0,1]$. Assume that (3.7), (3.8), (3.9) and (3.10) hold. Let $X(\cdot)$ be the mild solution of (3.5). Then there exist constants $c_1 > 0, c_2 > 0$ (depending only on T, C, γ) such that*

$$\mathbb{E}\left(\sup_{0\leq s\leq T} e^{c_1(1+|X(s)|^2)^{(1-\gamma)}} \right) \leq c_2 e^{(1+|x|^2)^{(1-\gamma)}} \qquad \text{if } \gamma \in [0,1) \tag{3.11}$$

and

$$\mathbb{E}\left(\sup_{0\leq s\leq T} e^{c_1(\log(2+|X(s)|^2))^2} \right) \leq c_2 e^{(\log(2+|x|^2))^2} \qquad \text{if } \gamma = 1. \tag{3.12}$$

PROOF. We first consider the case $0 \leq \gamma < 1$. Let X^n be the mild solution of the approximating equation (3.6) and τ_k be the minimum of T and the first exit time of X^n from the set $\{|z| \leq k\}$. Let $\beta > 0$ and $\alpha > 0$ be numbers to be specified later. Since A_n is bounded, X^n solves the integral equation

$$\begin{aligned} X^n(s) = x + \int_0^s A_n X^n(r) + b(r, X^n(r), a(r))dr \\ + \int_0^s \sigma(r, X^n(r), a(r))dW_Q(r). \end{aligned} \tag{3.13}$$

Thus we can apply Ito's formula (see Theorem 1.154) to the function

$$\begin{cases} \Phi\colon [0,T] \times H \to \mathbb{R} \\ \Phi(s,x) = e^{\beta e^{-\alpha s}(1+|x|^2)^{1-\gamma}} \end{cases}$$

and obtain,

$$\begin{aligned}
&e^{\beta e^{-\alpha(s\wedge\tau_k)}(1+|X^n(s\wedge\tau_k)|^2)^{1-\gamma}} \\
&\quad = e^{\beta(1+|x|^2)^{1-\gamma}} - \int_0^{s\wedge\tau_k} \alpha\beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma} e^{\beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}} dr \\
&\quad + \int_0^{s\wedge\tau_k} 2(1-\gamma)\beta e^{-\alpha r} e^{\beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}} (1+|X^n(r)|^2)^{-\gamma} \\
&\qquad\qquad \times \langle A_n X^n(r) + b(r, X^n(r), a(r)), X^n(r)\rangle dr \\
&\quad + \int_0^{s\wedge\tau_k} 2(1-\gamma)\beta e^{-\alpha r} e^{\beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}} (1+|X^n(r)|^2)^{-\gamma} \\
&\qquad\qquad \times \langle X^n(r), \sigma(r, X^n(r), a(r))dW_Q(r)\rangle \\
&\quad + \frac{1}{2}\int_0^{s\wedge\tau_k} e^{\beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}} \mathrm{Tr}\Big(\Big(\sigma(r, X^n(r), a(r))Q^{\frac12}\Big) \Big(\sigma(r, X^n(r), a(r))Q^{\frac12}\Big)^* \\
&\qquad\qquad \times 2\Big[2\beta^2 e^{-2\alpha r}(1+|X^n(r)|^2)^{-2\gamma}(1-\gamma)^2 X^n(r)\otimes X^n(r) \\
&\qquad\qquad - 2\beta\gamma e^{-\alpha r}(1+|X^n(r)|^2)^{-\gamma-1}(1-\gamma) X^n(r)\otimes X^n(r) \\
&\qquad\qquad\qquad + \beta e^{-\alpha r}(1+|X^n(r)|^2)^{-\gamma}(1-\gamma) I\Big]\Big) dr
\end{aligned}$$

$$\begin{aligned}
&\le e^{\beta(1+|x|^2)^{1-\gamma}} + \int_0^{s\wedge\tau_k} (1+|X^n(r)|^2)^{1-\gamma} e^{\beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}} \\
&\qquad\qquad \times (-\alpha + C(\beta))\beta e^{-\alpha r} dr \\
&\quad + 2\int_0^s \mathbf{1}_{[0,\tau_k]}(1-\gamma)\beta e^{-\alpha r} e^{\beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}} (1+|X^n(r)|^2)^{-\gamma} \\
&\qquad\qquad \times \langle X^n(r), \sigma(r, X^n(r), a(r))dW_Q(r)\rangle
\end{aligned} \tag{3.14}$$

for some absolute constant $C(\beta)$, nondecreasing in β and also depending on C, γ, where we used Lemma 1.105 in the last line of (3.14) .

Therefore, choosing $\alpha = C(\beta) + 1$ in (3.14) we obtain

$$\begin{aligned}
&e^{\beta e^{-\alpha(s\wedge\tau_k)}(1+|X^n(s\wedge\tau_k)|^2)^{1-\gamma}} + \int_0^{s\wedge\tau_k} \beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma} e^{\beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}} dr \\
&\le e^{\beta(1+|x|^2)^{1-\gamma}} + 2\int_0^s \mathbf{1}_{[0,\tau_k]}(1-\gamma)\beta e^{-\alpha r} e^{\beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}} (1+|X^n(r)|^2)^{-\gamma} \\
&\qquad\qquad \times \langle X^n(r), \sigma(r, X^n(r), a(r))dW_Q(r)\rangle.
\end{aligned} \tag{3.15}$$

Therefore, taking expectation in (3.15), yields

$$\mathbb{E}e^{\beta e^{-\alpha(s\wedge\tau_k)}(1+|X^n(s\wedge\tau_k)|^2)^{1-\gamma}} + \mathbb{E}\int_0^{s\wedge\tau_k} \beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma} e^{\beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}} dr \le e^{\beta(1+|x|^2)^{1-\gamma}}. \quad (3.16)$$

Now we choose $\alpha = C(2)+1$ in (3.14) so that (3.15) and (3.16) are satisfied for $\beta = 2$. Moreover we can observe that, since $C(\beta)$ is an increasing function of β and since the term $\beta e^{-\alpha r} e^{\beta e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}}(1+|X^n(r)|^2)^{1-\gamma}$ is always positive, (3.15) and (3.16) are satisfied also when we choose $\beta = 1$ and $\alpha = C(2)+1$. Using (3.15) with this last choice of α and β and observing that the integral in the left hand side of (3.15) is positive, we get

$$\sup_{0\le u\le s} e^{e^{-\alpha(u\wedge\tau_k)}(1+|X^n(u\wedge\tau_k)|^2)^{1-\gamma}} \le e^{(1+|x|^2)^{1-\gamma}} + \sup_{0\le u\le s}\left| \int_0^u 2e^{-\alpha r} e^{e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}} \mathbf{1}_{[0,\tau_k]} \times (1-\gamma)(1+|X^n(r)|^2)^{-\gamma}\langle X^n(r), \sigma(r, X^n(r), a(r)) dW_Q(r)\rangle \right|,$$

and therefore, using Burkholder-Davis-Gundy inequality (see Theorem 1.106), we have

$$\begin{aligned}\mathbb{E}\sup_{0\le u\le s} & e^{e^{-\alpha(u\wedge\tau_k)}(1+|X^n(u\wedge\tau_k)|^2)^{1-\gamma}} \le e^{(1+|x|^2)^{1-\gamma}} \\ &+ \left(\mathbb{E}\sup_{0\le u\le s}\left| \int_0^u 2e^{-\alpha r} e^{e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}} \mathbf{1}_{[0,\tau_k]} \times (1-\gamma)(1+|X^n(r)|^2)^{-\gamma}\langle X^n(r), \sigma(r, X^n(r), a(r)) dW_Q(r)\rangle \right|\right) \\ &\le e^{(1+|x|^2)^{1-\gamma}} \\ &+ \left(\mathbb{E}\int_0^s C_1 e^{-2\alpha r} e^{2e^{-\alpha r}(1+|X^n(r)|^2)^{1-\gamma}}(1+|X^n(r)|^2)^{1-\gamma}\mathbf{1}_{[0,\tau_k]} dr\right)^{\frac{1}{2}}.\end{aligned} \quad (3.17)$$

Using (3.16) with $\beta = 2$ we see that the last two lines of (3.17) are less than or equal to

$$C_2 e^{(1+|x|^2)^{1-\gamma}}.$$

Above, the constants C_i, $i = 1, 2$, only depend on C, γ, T. Thus we have obtained

$$\mathbb{E}\sup_{0\le s\le T} e^{e^{-\alpha T}(1+|X^n(s\wedge\tau_k)|^2)^{1-\gamma}} \le C_2 e^{(1+|x|^2)^{1-\gamma}}. \quad (3.18)$$

Since $\lim_{k\to+\infty}\tau_k = T$ a.s., letting $k \to +\infty$ in (3.18) and using Fatou's lemma, we obtain

$$\mathbb{E}\sup_{0\le s\le T} e^{e^{-\alpha T}(1+|X^n(s)|^2)^{1-\gamma}} \le C_2 e^{(1+|x|^2)^{1-\gamma}}.$$

It now remains to use (see Theorem 1.125) that

$$\lim_{n\to\infty}\mathbb{E}\big(\sup_{0\le s\le T}|X^n(s) - X(s)|^2\big) = 0.$$

It implies the existence of a subsequence X^{n_k} satisfying $\lim_{n\to\infty}\sup_{0\le s\le T}|X^{n_k}(s) - X(s)|^2 = 0$ almost surely and then it ensures that, a.s.,

$$\lim_{n_k\to\infty}\sup_{0\le s\le T} e^{e^{-\alpha T}(1+|X^{n_k}(s)|^2)^{1-\gamma}} = \sup_{0\le s\le T} e^{e^{-\alpha T}(1+|X(s)|^2)^{1-\gamma}}.$$

We can then apply Fatou's lemma again to obtain the claim.

For $\gamma = 1$ we can repeat the same arguments applied to the function

$$e^{\beta e^{-\alpha s}(\log(2+|x|^2))^2}.$$

□

Lemma 3.19 *Let A be a linear, densely defined maximal dissipative operator in H and B an operator satisfying the weak B-condition for A for some constant $c_0 > 0$. Then:*

(i) *For any $R > 0$ there exists a constant $C(R)$ such that, for $x \in H, |x| \le R$ and $t \ge 0$,*

$$|e^{tA}x - x|_{-1} \le C(R)\sqrt{t}. \tag{3.19}$$

(ii) *If B satisfies the strong B-condition for A with constant c_0 then, for $x \in H$ and $t \ge 0$,*

$$|e^{tA}x|_{-1}^2 + 2t|e^{tA}x|^2 \le e^{2c_0 t}|x|_{-1}^2. \tag{3.20}$$

Proof. *(i)* Denote $Z(t) = e^{tA}x$. If $x \in D(A)$, using that A is maximal dissipative, Theorem B.45, and (3.2) we have

$$\begin{aligned}
|Z(t) - x|_{-1}^2 &= \int_0^t \langle 2B(Z(s) - x), AZ(s)\rangle ds \\
&\le 2\int_0^t \langle A^*BZ(s), Z(s)\rangle ds + 2\|A^*B\||x|^2 t \\
&\le 2c_0 \int_0^t |Z(s)|_{-1}^2 ds + 2\|A^*B\||x|^2 t \le (2c_0\|B\||x|^2 + 2\|A^*B\||x|^2)t.
\end{aligned}$$

The estimate now follows by density of $D(A)$.

(ii) Again it is enough to show the estimate for $x \in D(A)$. We then have by (3.3)

$$\frac{d}{ds}|Z(s)|_{-1}^2 = 2\langle A^*BZ(s), Z(s)\rangle \le 2c_0|Z(s)|_{-1}^2 - 2|Z(s)|^2,$$

and thus

$$\begin{aligned}
\frac{d}{ds}\left(e^{-2c_0 s}|Z(s)|_{-1}^2\right) &= -2c_0 e^{-2c_0 s}|Z(s)|_{-1}^2 \\
&\quad + e^{-2c_0 s}\frac{d}{ds}|Z(s)|_{-1}^2 \le -2e^{-2c_0 s}|Z(s)|^2.
\end{aligned} \tag{3.21}$$

Integrating we obtain

$$e^{-2c_0 t}|Z(t)|_{-1}^2 + 2\int_0^t e^{-2c_0 s}|Z(s)|^2 ds \le |x|_{-1}^2.$$

The inequality now follows upon noticing that $e^{-2c_0 t}|Z(t)|^2 \le e^{-2c_0 s}|Z(s)|^2$ for $0 \le s \le t$ since e^{sA} is a semigroup of contractions. □

Lemma 3.20 *Let A be a linear, densely defined maximal dissipative operator in H and B an operator satisfying the weak B-condition for A for some $c_0 \ge 0$. Let (3.7), (3.9) and (3.10) with $\gamma = 1$ hold, and let*

$$\langle b(s,x,a) - b(s,y,a), B(x-y)\rangle \le C|x-y|_{-1}^2 \tag{3.22}$$

$$\|\sigma(s,x,a) - \sigma(s,y,a)\|_{\mathcal{L}_2(U_0,H)} \le C|x-y|_{-1}, \tag{3.23}$$

for all $x, y \in H$, $s \in [0,T]$ and $a \in \Lambda$. If $X(\cdot)$ and $Y(\cdot)$ are the mild solutions of (3.5) with initial conditions $X(0) = x$ and $Y(0) = y$ respectively, driven by the process $a : [0,T] \times \Omega \to \Lambda$, then

$$\sup_{s\in[0,T]} \mathbb{E}\left[|X(s) - Y(s)|_{-1}^2\right] \le C(T)|x-y|_{-1}^2 \tag{3.24}$$

where $C(T)$ is a constant depending only on $T, C, c_0, \|B\|$.

PROOF. We define the function

$$\begin{cases} F\colon H \to \mathbb{R} \\ F(z) = |z|_{-1}^2 = \langle Bz, z\rangle \,. \end{cases}$$

We notice that $DF(z) = 2Bz$ and $D^2F(z) = 2B$. We will apply Ito's formula to F along the trajectories of the process $Z(s) := X(s) - Y(s)$, which is a mild solution of $dZ(s) = (AZ(s) + f(s))\,dt + \Phi(s)dW_Q(s)$, where $f(s) := b(s, X(s), a(s)) - b(s, Y(s), a(s))$ and $\Phi(s) := \sigma(s, X(s), a(s)) - \sigma(s, Y(s), a(s))$. Thanks to (3.7), (3.9), (3.10) and (3.23), the assumptions of Theorem 1.124 are satisfied and thus we have (1.38) and the hypotheses of Proposition 1.155 are satisfied. Therefore we have

$$\mathbb{E}\left[|Z(s)|_{-1}^2\right] = |x-y|_{-1}^2 + \int_0^s \mathbb{E}\left[\langle 2A^*BZ(r), Z(r)\rangle + \langle 2BZ(r), f(r)\rangle\right] ds \\ + \int_0^s \mathbb{E}\left[\mathrm{Tr}\Big(\left(\Phi(r)Q^{\frac{1}{2}}\right) \left(\Phi(r)Q^{\frac{1}{2}}\right)^* B\Big)\right] dr, \quad (3.25)$$

and using (3.2), (3.22) and (3.23), we find

$$\mathbb{E}\left[|Z(s)|_{-1}^2\right] \le |x-y|_{-1}^2 + \int_0^s \mathbb{E}\left[2c_0|Z(r)|_{-1}^2 + 2C|Z(r)|_{-1}^2\right] dr \\ + \int_0^s \mathbb{E}\left[\|B\|C^2|Z(r)|_{-1}^2\right] dr. \quad (3.26)$$

Applying Gronwall's lemma we obtain (3.24). □

REMARK 3.21 Condition (3.22) is obviously satisfied if

$$|b(s,x,a) - b(s,y,a)|_{-1} \le C|x-y|_{-1}$$

for all $x, y \in H$, $s \in [0,T]$ and $a \in \Lambda$. ■

LEMMA 3.22 *Let A be a linear, densely defined maximal dissipative operator in H. Let B be a bounded, strictly positive, self-adjoint operator on H such that A^*B is bounded. Let (3.7), (3.8), (3.9) and (3.10) with $\gamma = 1$ hold. If $X(\cdot)$ is the mild solution of (3.5) with initial condition $X(0) = x$ driven by the process $a : [0,T] \times \Omega \to \Lambda$, then*

$$\mathbb{E}\left[|X(s) - x|_{-1}^2\right] \le C(|x|, T)s \quad (3.27)$$

where $C(|x|, T)$ is a constant depending only on $T, |x|, C, \|B\|, \|A^*B\|$.

PROOF. We define the function

$$\begin{cases} F\colon H \to \mathbb{R} \\ F(z) = |z-x|_{-1} = \langle B(z-x), z-x\rangle \,. \end{cases}$$

We have $DF(z) = 2B(z-x)$ and $D^2F(z) = 2B$, and applying Proposition 1.156 yields

$$\mathbb{E}\left[|X(s) - x|_{-1}^2\right] \\ = 2\int_0^s \mathbb{E}\left[\langle X(r), A^*B(X(r)-x)\rangle + \langle b(r, X(r), a(r)), B(X(r)-x)\rangle\right] dr \\ + \int_0^s \mathbb{E}\left[\mathrm{Tr}\Big(\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right) \left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)^* B\Big)\right] dr. \quad (3.28)$$

Using (3.9), (3.10), the boundedness of A^*B and (1.38), we easily deduce using Cauchy-Schwarz inequality, that the absolute values of the integrands in the right hand side of (3.28) remain bounded by some constant $C(T,|x|)$ depending only on $T,|x|,C,\|B\|,\|A^*B\|$. This concludes the proof of (3.27). □

LEMMA 3.23 *Let A be a linear, densely defined maximal dissipative operator in H and B an operator satisfying the strong B-condition for A for some $c_0 \geq 0$. Let (3.7), (3.9), (3.10) and (3.23) hold. If $X(\cdot)$ and $Y(\cdot)$ are the mild solutions of (3.5) with initial conditions $X(0)=x$ and $Y(0)=y$ respectively, driven by the process $a:[0,T]\times\Omega\to\Lambda$, then*

$$\sup_{s\in[0,T]}\left(\mathbb{E}\left[|X(s)-Y(s)|_{-1}^2\right]+\mathbb{E}\int_0^s|X(r)-Y(r)|^2dr\right)\leq C(T)|x-y|_{-1}^2 \quad (3.29)$$

and, for any $s\in(0,T)$,

$$\mathbb{E}\left[|X(s)-Y(s)|^2\right]\leq\frac{C(T)}{s}|x-y|_{-1}^2, \quad (3.30)$$

where $C(T)$ is a constant depending only on $T,C,c_0,\|B\|$.

PROOF. Following the proof of Lemma 3.20 if $Z(s)=X(s)-Y(s)$, $f(s)=b(s,X(s),a(s))-b(s,Y(s),a(s))$ and $\Phi(s)=\sigma(s,X(s),a(s))-\sigma(s,Y(s),a(s))$ we have (as in (3.25)):

$$\begin{aligned}\mathbb{E}\left[|Z(s)|_{-1}^2\right]=|x-y|_{-1}^2+\int_0^s\mathbb{E}\left[\langle 2A^*BZ(r),Z(r)\rangle+\langle 2BZ(r),f(r)\rangle\right]ds\\+\int_0^s\mathbb{E}\left[\mathrm{Tr}\left(\left(\Phi(r)Q^{\frac12}\right)\left(\Phi(r)Q^{\frac12}\right)^*B\right)\right]dr. \quad (3.31)\end{aligned}$$

We observe that (3.3) implies

$$\langle 2A^*BZ(r),Z(r)\rangle+\langle 2Z(r),Z(r)\rangle\leq\langle 2c_0BZ(r),Z(r)\rangle,$$

which, together with (3.7), gives

$$\begin{aligned}&\langle 2A^*BZ(r),Z(r)\rangle+\langle 2BZ(r),f(r)\rangle\\&\quad\leq 2c_0|Z(r)|_{-1}^2-2|Z(r)|^2+2C\|B\|^{1/2}|Z(r)|_{-1}|Z(r)|\leq c_1|Z(r)|_{-1}^2-|Z(r)|^2 \quad (3.32)\end{aligned}$$

where c_1 depends on c_0, $\|B\|$ and C. Using (3.23) we have

$$\mathrm{Tr}\left(\left(\Phi(r)Q^{\frac12}\right)\left(\Phi(r)Q^{\frac12}\right)^*B\right)\leq\|B\|C^2|Z(r)|_{-1}^2=c_2|Z(r)|_{-1}^2. \quad (3.33)$$

It thus follows from (3.31), (3.32), (3.33) that

$$\mathbb{E}\left[|Z(s)|_{-1}^2\right]+\int_0^s\mathbb{E}\left[|Z(r)|^2\right]dr\leq|x-y|_{-1}^2+(c_1+c_2)\int_0^s\mathbb{E}\left[|Z(r)|_{-1}^2\right]dr.$$

Such an inequality holds, of course, also dropping the positive term $\int_0^s\mathbb{E}\left[|Z(r)|^2\right]dr$ and then (3.29) follows easily from Gronwall's lemma. Regarding (3.30), using the definition of mild solution, (3.20), (3.29), and elementary computations, we have

$$\mathbb{E}\left[|X(s)-Y(s)|^2\right]\leq C_1(|e^{sA}(x-y)|^2+|x-y|_{-1}^2)\leq\frac{C_2}{s}|x-y|_{-1}^2,$$

where C_1 and C_2 only depend on $T,C,c_0,\|B\|$. □

PROPOSITION 3.24 *Let $m>0$. Let (3.7) and (3.8), (3.9) and (3.10) with $\gamma=1$ hold. Let $X(\cdot)$ be the mild solution of (3.5), driven by the process $a(\cdot):[0,T]\times\Omega\to\Lambda$ and let C be the constant appearing in (3.9) and (3.10). Denote*

$$\bar\lambda=Cm+\frac12C^2m(m-1)\quad\text{if } m\geq 2,$$

and

$$\bar{\lambda} = Cm + \frac{1}{2}C^2 m \quad \text{if } 0 < m < 2.$$

Then for every $\lambda > \bar{\lambda}$ *there exists a constant* C_λ *such that*

$$\mathbb{E}\left[\left(C_\lambda + |X(s)|^2\right)^{\frac{m}{2}}\right] \le (C_\lambda + |x|^2)^{\frac{m}{2}} e^{\lambda s} \quad \text{for all } s \ge 0. \tag{3.34}$$

PROOF. Let $\lambda > \bar{\lambda}$ and let $\epsilon = \epsilon(\lambda) > 0$ be such that $\bar{\lambda}(1+\epsilon) = \lambda$. We set $C_\lambda > 0$ to be a number such that $2r \le C_\lambda - 1 + \epsilon r^2$ for all $r \ge 0$. It is then easy to see that

$$Cmr(1+r) + \frac{1}{2}C^2 m(m-1)(1+r)^2 \le \lambda(K_\epsilon + r^2) \quad \text{for all } r \ge 0.$$

Define $F(z) = (C_\lambda + |z|^2)^{\frac{m}{2}}$. Then $DF(z) = m(C_\lambda + |z|^2)^{\frac{m-2}{2}} z$ and $D^2F(z) = m(m-2)(C_\lambda + |z|^2)^{\frac{m-4}{2}} z \otimes z + m(1 + |z|^2)^{\frac{m-2}{2}} I$.

Assume first that $m \ge 2$. Using Proposition 1.157 and (3.9), (3.10) we then have

$$\begin{aligned} \mathbb{E}\left[\left(C_\lambda + |X(s)|^2\right)^{\frac{m}{2}}\right] &\le (C_\lambda + |x|^2)^{\frac{m}{2}} \\ &\quad + \int_0^s \mathbb{E}\Bigg[m(C_\lambda + |X(r)|^2)^{\frac{m-2}{2}} \langle X(r), b(r, X(r), a(r))\rangle \\ &\quad + \frac{1}{2}\mathrm{Tr}\Bigg(\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)^* \\ \times \Big(m(m-2)&(C_\lambda + |X(r)|^2)^{\frac{m-4}{2}} X(r) \otimes X(r) + m(C_\lambda + |X(r)|^2)^{\frac{m-2}{2}} I \Big) \Bigg)\Bigg] dr \\ &\le (C_\lambda + |x|^2)^{\frac{m}{2}} + \lambda \int_0^s \mathbb{E}\left[\left(1 + |X(r)|^2\right)^{\frac{m}{2}}\right] dr \end{aligned} \tag{3.35}$$

and we conclude applying Gronwall's lemma.

For $0 < m < 2$ the first term in the fourth line of (3.35) can be dropped and we argue as before since

$$Cmr(1+r) + \frac{1}{2}C^2 m(1+r)^2 \le \lambda(K_\epsilon + r^2) \quad \text{for all } r \ge 0.$$

□

3.1.3. Perturbed optimization. The following is a classical result of Ekeland and Lebourg [147], see also [420] and [312], Lemma 4.2 page 245, for a more general formulation.

THEOREM 3.25 *Let D be a bounded closed subset of a real Hilbert space K and $f\colon D \to \mathbb{R} \cup \{-\infty\}$ be upper semicontinuous and such that* $\mathrm{dom}\,(f) := \{x \in D : f(x) \in \mathbb{R}\} \neq \emptyset$. *Suppose that f is bounded from above. Then, for any $\delta > 0$, there exist $y \in K, \hat{x} \in D$ such that $|y|_K < \delta$ and the function*

$$x \mapsto f(x) + \langle y, x\rangle_K$$

has a strict maximum over D at $\hat{x}$.

COROLLARY 3.26 *Let H be a real, separable Hilbert space with inner product $\langle\cdot,\cdot,\rangle$, let B be a strictly positive operator in $S(H)$ and $D \subseteq H$ be a bounded, B-closed subset of H. Let $f\colon D \to \mathbb{R} \cup \{-\infty\}$ be a B-upper semicontinuous function, bounded from above. Then, for any $\delta > 0$, there exist $p \in H, \hat{x} \in D$ such that $|p| < \delta$, and the function*

$$x \to f(x) + \langle Bp, x\rangle$$

attains a maximum over D at $\hat{x}$, which is strict in the topology of H_{-2}.

PROOF. We want to apply Theorem 3.25 to D endowed with the topology induced by H_{-2}.

D is obviously bounded in H_{-2} and it is easy to see that D is closed in H_{-2}. To show it, let $(x_n)_{n\in\mathbb{N}}$ be a sequence in D such that $x_n \xrightarrow[H_{-2}]{n\to\infty} x \in H_{-2}$, i.e. $Bx_n \to z$ for some $z \in H$. Since D is bounded in H, there is a subsequence, still denoted by x_n, such that $x_n \overset{H}{\rightharpoonup} \tilde{x} \in H$ for some $\tilde{x} \in D$. But the graph of B is weakly sequentially closed, so we obtain $B\tilde{x} = z$, which implies that $|x_n - \tilde{x}|_{-2} \to 0$, and thus $x = \tilde{x}$. Since D is B-closed, we thus have $x \in D$.

In particular we showed that if (x_n) is a sequence in D such that $x_n \xrightarrow[H_{-2}]{n\to\infty} x$, then $x \in D$ and $x_n \rightharpoonup x$. Since f is B-upper semicontinuous, this shows that f is upper semicontinuous on D considered as a subset H_{-2}.

We can now apply Theorem 3.25 to obtain that for all $\delta > 0$ there exists $y \in H_{-2}$ with $|y|_{H_{-2}} < \delta$ such that $x \mapsto f(x) + \langle y, x\rangle_{H_{-2}}$ attains a strict maximum (in the topology of H_{-2}) on D at some point $\hat{x}$. Denote $p := By \in H$. Since $B\colon H_{-2} \to H$ is an isometry, we have that $|p| = |y|_{-2} < \delta$. Therefore for $x \in H$, $\langle y, x\rangle_{H_{-2}} = \left\langle B^2 y, x\right\rangle_H = \langle Bp, x\rangle_H$, which completes the proof. □

3.2. Maximum principle

From now on, throughout the rest of Chapter 3, unless stated otherwise, A is a linear, densely defined, maximal dissipative operator in H.

In this section B is any strictly positive operator in $S(H)$. Let $\{e_1, e_2, ...\}$ be an orthonormal basis in H_{-1} (see Definition 3.1) made of elements of H. For $N > 2$ we denote $H_N = \operatorname{span}\{e_1, ..., e_N\}$. Let $P_N : H_{-1} \to H_{-1}$ be the orthogonal projection onto H_N. It is clear that P_N is also a bounded operator on H and therefore so is $Q_N := I - P_N$, i.e. $P_N, Q_N \in \mathcal{L}(H)$. It is also easy to see that $BP_N = P_N^* BP_N = P_N^* B, BQ_N = Q_N^* BQ_N = Q_N^* B$, where P_N^*, Q_N^* are adjoints of P_N, Q_N as operators in $\mathcal{L}(H)$. For $x \in H$ we will write $x_N := P_N x$, $x_N^\perp := Q_N x$.

We remark that if B is compact then $\|B^\gamma Q_N\| \to 0$ as $N \to +\infty$ for every $\gamma > 0$. Also in this case a natural choice for the basis $\{e_1, e_2, ...\}$ is to take $e_i = B^{-\frac{1}{2}} f_i$, where $\{f_1, f_2, ...\}$ is an orthonormal basis of H composed of eigenvectors of B. This choice of basis has a property that it is orthogonal in H_{-1} and H.

For a function $w \in C^2(H_{-1})$ we will write $D_{H_{-1}} w$, $D^2_{H_{-1}} w$ to denote the Fréchet derivatives of w as a function in $C^2(H_{-1})$ whereas $Dw, D^2 w$ mean the Fréchet derivatives of w as a function in $C^2(H)$. We remark that the spaces H_1, H_2 in Theorem 3.27 are the spaces introduced in Section 3.1.1, not one of the spaces H_N defined above. This is why we put the restriction $N > 2$.

THEOREM 3.27 (Maximum Principle) *Let $B \in S(H)$ be strictly positive and let $N > 2$, $\kappa > 0$. Let $u, v : H \to \mathbb{R} \cup \{-\infty\}$ be B-upper semi-continuous, bounded from above functions and such that*

$$\limsup_{|x|\to+\infty} \frac{u(x)}{|x|} < 0 \quad \textit{and} \quad \limsup_{|x|\to+\infty} \frac{v(x)}{|x|} < 0. \tag{3.36}$$

Let $\Phi \in C^2(H_N \times H_N)$ be such that

$$u(x_N + x_N^\perp) + v(y_N + y_N^\perp) - \Phi(x_N, y_N)$$

has a strict global maximum over $H \times H$ at a point $(\bar{x}, \bar{y})$. Then there exist functions $\varphi_k, \psi_k \in C^2(H)$ for $k = 1, 2, ...$ such that $\varphi_k, B^{-1}D\varphi_k, D^2\varphi_k, \psi_k, B^{-1}D\psi_k, D^2\psi_k$ are bounded and uniformly continuous, and such that

$$u(x) - \varphi_k(x)$$

has a global maximum at some point x_k,

$$v(y) - \psi_k(y)$$

has a global maximum at some point y_k, *and*

$$\big(x_k, u(x_k), D\varphi_k(x_k), D^2\varphi_k(x_k)\big) \xrightarrow{k\to+\infty} \big(\bar{x}, u(\bar{x}), D_x\Phi(\bar{x}_N, \bar{y}_N), X_N\big) \quad \text{in } H \times \mathbb{R} \times H_2 \times \mathcal{L}(H_{-1}; H_1), \tag{3.37}$$

$$\big(y_k, v(y_k), D\psi_k(y_k), D^2\psi_k(y_k)\big) \xrightarrow{k\to+\infty} \big(\bar{y}, v(\bar{y}), D_y\Phi(\bar{x}_N, \bar{y}_N), Y_N\big) \quad \text{in } H \times \mathbb{R} \times H_2 \times \mathcal{L}(H_{-1}; H_1), \tag{3.38}$$

where $X_N, Y_N \in S(H)$, $X_N = P_N^* X_N P_N, Y_N = P_N^* Y_N P_N$,

$$-\left(\frac{1}{\kappa} + \|C\|_{\mathcal{L}(H_{-1}\times H_{-1})}\right)\begin{pmatrix} BP_N & 0 \\ 0 & BP_N \end{pmatrix} \le \begin{pmatrix} X_N & 0 \\ 0 & Y_N \end{pmatrix} \le \begin{pmatrix} B & 0 \\ 0 & B \end{pmatrix}\big(C + \kappa C^2\big) \tag{3.39}$$

and $C = D^2_{H_{-1}\times H_{-1}}\Phi(\bar{x}_N, \bar{y}_N)$.

We remark that in fact $\varphi_k, \psi_k \in C^2(H_{-1})$.

PROOF. Define

$$\tilde{u}(x_N) := \sup_{x_N^\perp \in Q_N H} u(x_N + x_N^\perp),$$

$$\tilde{v}(y_N) := \sup_{y_N^\perp \in Q_N H} v(y_N + y_N^\perp),$$

the partial sup-convolutions of u and v respectively, and let $\tilde{u}^*$ and $\tilde{v}^*$ be their upper semi-continuous envelopes (see Definition D.9). We remark that $\tilde{u}, \tilde{v}$ do not need to be upper semi-continuous (see [102]). Since $u + v - \Phi$ had a strict global maximum at $(\bar{x}, \bar{y})$ it easily follows that

$$\tilde{u}^*(x_N) + \tilde{v}^*(y_N) - \Phi(x_N, y_N) \tag{3.40}$$

has a strict global maximum over $H_N \times H_N$ at $(\bar{x}_N, \bar{y}_N)$. Moreover we have $\tilde{u}^*(\bar{x}_N) = u(\bar{x})$, $\tilde{v}^*(\bar{y}_N) = v(\bar{y})$.

We can now apply the finite dimensional maximum principle (see Theorem E.10, that is a particular case of Theorem 3.2 in [101]) when we consider H_N as a subspace of H_{-1}. (We remind that in H_N the topology of H_{-1} is equivalent to the topology of H). Denote H_N with this topology by $\tilde{H}_N$. We also note that $\Phi \in C^2(\tilde{H}_N \times \tilde{H}_N)$ and thus we can consider it as a function in $C^2(H_{-1} \times H_{-1})$ by setting $\Phi(x, y) := \Phi(P_N x, P_N y)$.

Therefore there exist bounded functions $\varphi_k, \psi_k \in C^2(\tilde{H}_N)$ with bounded and uniformly continuous derivatives (which we can consider as functions in $C^2(H_{-1})$ by setting $\varphi_k(x) := \varphi_k(P_N x)$ and $\psi_k(y) := \psi_k(P_N y)$) such that $\tilde{u}^*(x_N) - \varphi_k(x_N)$ has a strict global maximum at some point x_N^k, $\tilde{v}^*(y_N) - \psi_k(y_N)$ has a strict global minimum at some point y_N^k, and such that

$$\big(x_N^k, \tilde{u}^*(x_N^k), D_{H_{-1}}\varphi_k(x_N^k), D^2_{H_{-1}}\varphi_k(x_N^k)\big) \xrightarrow{k\to\infty} \big(\bar{x}_N, u(\bar{x}), D_{H_{-1},x}\Phi(\bar{x}_N, \bar{y}_N), \tilde{X}_N\big), \tag{3.41}$$

$$\big(y_N^k, \tilde{v}^*(y_N^k), D_{H_{-1}}\psi_k(y_N^k), D^2_{H_{-1}}\psi_k(y_N^k)\big) \xrightarrow{k\to\infty} \big(\bar{y}_N, v(\bar{y}), D_{H_{-1},y}\Phi(\bar{x}_N, \bar{y}_N), \tilde{Y}_N\big), \tag{3.42}$$

and

$$-\left(\frac{1}{\kappa}+\|C\|_{\mathcal{L}(H_{-1}\times H_{-1})}\right)\begin{pmatrix} P_N & 0 \\ 0 & P_N \end{pmatrix} \\ \le \begin{pmatrix} \tilde{X}_N & 0 \\ 0 & \tilde{Y}_N \end{pmatrix} \le C+\kappa C^2 \quad \text{in } H_{-1}\times H_{-1} \tag{3.43}$$

for some $\tilde{X}_N, \tilde{Y}_N \in S(H_{-1})$ that satisfy $\tilde{X}_N = P_N\tilde{X}_N P_N$, $\tilde{Y}_N = P_N\tilde{Y}_N P_N$ as operators in $\mathcal{L}(H_{-1})$ and, since

$$D_{H_{-1}}\varphi_k(x_N^k) = P_N D_{H_{-1}}\varphi_k(x_N^k),\ D^2_{H_{-1}}\varphi_k(x_N^k) = P_N D^2_{H_{-1}}\varphi_k(x_N^k)P_N,$$
$$D_{H_{-1}}\psi_k(x_N^k) = P_N D_{H_{-1}}\psi_k(y_N^k),\ D^2_{H_{-1}}\psi_k(y_N^k) = P_N D^2_{H_{-1}}\psi_k(y_N^k)P_N,$$

and in H_N the topology of H_{-1} is equivalent to the topology of H, the convergences (3.41), (3.42) hold in $H\times\mathbb{R}\times H\times\mathcal{L}(H_{-1})$.

It is easy to see that

$$D\varphi_k(x) = BD_{H_{-1}}\varphi_k(x), \quad D^2\varphi_k(x) = BD^2_{H_{-1}}\varphi_k(x),$$
$$D\psi_k(y) = BD_{H_{-1}}\psi_k(y), \quad D^2\psi_k(y) = BD^2_{H_{-1}}\psi_k(y).$$

(Note that if $X\in S(H_{-1})$ then $BX\in S(H)$ since for $x,y\in H$ we have $\langle BXx,y\rangle = \langle Xx,y\rangle_{-1} = \langle x,Xy\rangle_{-1} = \langle x,BXy\rangle$.) Therefore, setting $X_N = B\tilde{X}_N, Y_N = B\tilde{Y}_N$, we obtain from (3.41) and (3.42) that

$$\left(x_N^k, \tilde{u}^*(x_N^k), D\varphi_k(x_N^k), D^2\varphi_k(x_N^k)\right) \xrightarrow{k\to\infty} \left(\bar{x}_N, u(\bar{x}), D_x\Phi(\bar{x}_N,\bar{y}_N), X_N\right) \\ \text{in } H\times\mathbb{R}\times H_2\times\mathcal{L}(H_{-1};H_1), \tag{3.44}$$

$$\left(y_N^k, \tilde{v}^*(y_N^k), D\psi_k(y_N^k), D^2\psi_k(y_N^k)\right) \xrightarrow{k\to\infty} \left(\bar{y}_N, v(\bar{y}), D_y\Phi(\bar{x}_N,\bar{y}_N), Y_N\right) \\ \text{in } H\times\mathbb{R}\times H_2\times\mathcal{L}(H_{-1};H_1), \tag{3.45}$$

$X_N = P_N^* X_N P_N, Y_N = P_N^* Y_N P_N$, and (3.39) is satisfied.

Now, using Corollary 3.26 and (3.36), for every k and j big enough we can find $p_j^k, q_j^k \in H$, such that $|p_j^k|+|q_j^k|\le 1/j$, and

$$u(x)-\varphi_k(x)-\langle Bp_j^k, x\rangle \quad \text{has a global maximum at some point } x_j^k, \tag{3.46}$$

and

$$v(y)-\psi_k(y)-\langle Bq_j^k, y\rangle \quad \text{has a global maximum at some point } y_j^k, \tag{3.47}$$

where, because of (3.36), $|x_j^k|+|y_j^k|\le R_k$ for some $R_k>0$. Then if $|x|,|y|\le R$ for $R\ge R_k$

$$\begin{aligned} &\tilde{u}^*((x_j^k)_N)+\tilde{v}^*((y_j^k)_N)-\varphi_k(x_j^k)-\psi_k(y_j^k) \\ &\quad\ge u(x_j^k)+v(y_j^k)-\varphi_k(x_j^k)-\psi_k(y_j^k) \\ &\quad\ge u(x)+v(y)-\varphi_k(x)-\psi_k(y)-\langle Bp_j^k, x-x_j^k\rangle-\langle Bq_j^k, y-y_j^k\rangle \\ &\quad\ge u(x)+v(y)-\varphi_k(x)-\psi_k(y)-\frac{4R\|B\|}{j}. \end{aligned} \tag{3.48}$$

Since by (3.36) if j is big enough, $u(x)-\varphi_k(x)-\langle Bp_j^k, x-x_j^k\rangle\to-\infty$ as $|x|\to+\infty$, and $v(y)-\psi_k(y)-\langle Bq_j^k, y-y_j^k\rangle\to-\infty$ as $|y|\to+\infty$, choosing R big enough and taking suprema over $x_N^\perp, y_N^\perp$ in (3.48) and then envelopes at x_N^k, y_N^k we obtain for sufficiently big j that

$$\begin{aligned}\tilde{u}^*((x_j^k)_N) + \tilde{v}^*((y_j^k)_N) - \varphi_k(x_j^k) - \psi_k(y_j^k)\\ \geq u(x_j^k) + v(y_j^k) - \varphi_k(x_j^k) - \psi_k(y_j^k)\\ \geq \tilde{u}^*(x_N^k) + \tilde{v}^*(y_N^k) - \varphi_k(x^k) - \psi_k(y^k) - \frac{4R\|B\|}{j}.\end{aligned} \tag{3.49}$$

Since $\tilde{u}^*(x_N) + \tilde{v}^*(y_N) - \varphi_k(x_N) - \psi_k(y_N)$ has a strict global maximum at (x_N^k, y_N^k), we deduce from (3.49) that

$$(x_j^k)_N \to x_N^k, (y_j^k)_N \to y_N^k, \quad \tilde{u}^*((x_j^k)_N) \to \tilde{u}^*(x_N^k), \quad \tilde{v}^*((y_j^k)_N) \to \tilde{v}^*(y_N^k) \tag{3.50}$$

as $j \to +\infty$ and then also

$$u(x_j^k) \to \tilde{u}^*(x_N^k), \quad v(y_j^k) \to \tilde{v}^*(y_N^k) \quad \text{as } j \to +\infty. \tag{3.51}$$

Using these and (3.44)-(3.45) we can therefore select a subsequence j_k such that

$$\big((x_{j_k}^k)_N, u(x_{j_k}^k), D\varphi_k(x_{j_k}^k), D^2\varphi_k(x_{j_k}^k)\big) \xrightarrow{k\to\infty} \big(\bar{x}_N, u(\bar{x}), D_x\Phi(\bar{x}_N, \bar{y}_N), X_N\big)$$
$$\text{in } H \times \mathbb{R} \times H_2 \times \mathcal{L}(H_{-1}; H_1),$$

$$\big((y_{j_k}^k)_N, v(y_{j_k}^k), D\psi_k(y_{j_k}^k), D^2\psi_k(y_{j_k}^k)\big) \xrightarrow{k\to\infty} \big(\bar{y}_N, v(\bar{y}), D_y\Phi(\bar{x}_N, \bar{y}_N), Y_N\big)$$
$$\text{in } H \times \mathbb{R} \times H_2 \times \mathcal{L}(H_{-1}; H_1).$$

It remains to show that $x_{j_k}^k \to \bar{x}$ and $y_{j_k}^k \to \bar{y}$. This is however now obvious since

$$u(x_{j_k}^k) + v(x_{j_k}^k) - \Phi((x_{j_k}^k)_N, (y_{j_k}^k)_N) \to u(\bar{x}) + v(\bar{y}) - \Phi(\bar{x}_N, \bar{y}_N)$$

and by assumption this function has a strict global maximum at $(\bar{x}, \bar{y})$. Therefore the lemma holds with $\varphi_k(x) := \varphi_k(x) + \langle Bp_{j_k}^k, x\rangle$, $\psi_k(y) := \psi_k(y) + \langle Bq_{j_k}^k, y\rangle$ and $x_k := x_{j_k}^k, y_k := y_{j_k}^k$. □

Theorem 3.27 applied to $\Phi(x_N, y_N) = \frac{1}{2\epsilon}|x_N - y_N|_{-1}^2$ for $\epsilon > 0$ yields the following result which will be used in the proofs of comparison theorems.

COROLLARY 3.28 *Let $B \in S(H)$ be strictly positive and let $N \geq 1, \epsilon > 0$. Let $u, -v : H \to \mathbb{R} \cup \{-\infty\}$ be B-upper semi-continuous, bounded from above functions satisfying (3.36). Let*

$$u(x_N + x_N^\perp) - v(y_N + y_N^\perp) - \frac{|x_N - y_N|_{-1}^2}{2\epsilon}$$

have a strict global maximum over $H \times H$ at a point $(\bar{x}, \bar{y})$. Then there exist functions $\varphi_k, \psi_k \in C^2(H)$ for $k = 1, 2, ...$ such that $\varphi_k, B^{-1}D\varphi_k, D^2\varphi_k, \psi_k, B^{-1}D\psi_k, D^2\psi_k$ are bounded and uniformly continuous, and such that

$$u(x) - \varphi_k(x)$$

has a global maximum at some point x_k,

$$v(y) - \psi_k(y)$$

has a global minimum at some point y_k, and

$$\big(x_k, u(x_k), D\varphi_k(x_k), D^2\varphi_k(x_k)\big) \xrightarrow{k\to\infty} \left(\bar{x}, u(\bar{x}), \frac{B(\bar{x}_N - \bar{y}_N)}{\epsilon}, X_N\right)$$
$$\text{in } H \times \mathbb{R} \times H_2 \times \mathcal{L}(H_{-1}; H_1), \tag{3.52}$$

$$\big(y_k, v(y_k), D\psi_k(y_k), D^2\psi_k(y_k)\big) \xrightarrow{k\to\infty} \left(\bar{y}, v(\bar{y}), \frac{B(\bar{x}_N - \bar{y}_N)}{\epsilon}, Y_N\right)$$
$$\textit{in } H \times \mathbb{R} \times H_2 \times \mathcal{L}(H_{-1}; H_1), \quad (3.53)$$

where $X_N = P_N^* X_N P_N, Y_N = P_N^* Y_N P_N$,

$$-\frac{3}{\epsilon}\begin{pmatrix} BP_N & 0 \\ 0 & BP_N \end{pmatrix} \le \begin{pmatrix} X_N & 0 \\ 0 & -Y_N \end{pmatrix} \le \frac{3}{\epsilon}\begin{pmatrix} BP_N & -BP_N \\ -BP_N & BP_N \end{pmatrix}. \quad (3.54)$$

PROOF. Observe that if $\Phi(x_N, y_N) = \frac{1}{2\epsilon}|x_N - y_N|_{-1}^2$ then

$$C = D^2_{H_{-1}\times H_{-1}}\Phi(x_N, y_N) = \frac{1}{\epsilon}\begin{pmatrix} P_N & -P_N \\ -P_N & P_N \end{pmatrix}$$

and thus $\kappa C^2 = \frac{2\kappa}{\epsilon}C$ and $\|C\|_{\mathcal{L}(H_{-1}\times H_{-1})} = \frac{2}{\epsilon}$. Then (3.54) follows from (3.39) choosing $\kappa = \epsilon$. □

We remark that the convergence in $\mathcal{L}(H_{-1}; H_1)$ in particular implies convergence in $\mathcal{L}(H)$.

The time dependent analogue of Corollary 3.28 is the following.

COROLLARY 3.29 *Let $B \in S(H)$ be strictly positive and let $N \ge 1, \epsilon, \beta > 0$. Let $u, -v : (0,T) \times H \to \mathbb{R} \cup \{-\infty\}$ be B-upper semi-continuous, bounded from above functions satisfying*

$$\limsup_{|x|\to+\infty} \sup_{t\in(0,T)} \frac{u(t,x)}{|x|} < 0 \quad \textit{and} \quad \limsup_{|x|\to+\infty} \sup_{t\in(0,T)} \frac{-v(t,x)}{|x|} < 0. \quad (3.55)$$

Let

$$u(t, x_N + x_N^\perp) - v(s, y_N + y_N^\perp) - \frac{|x_N - y_N|_{-1}^2}{2\epsilon} - \frac{(t-s)^2}{2\beta}$$

have a strict global maximum over $(0,T) \times H \times (0,T) \times H$ at a point $(\bar{t}, \bar{x}, \bar{s}, \bar{y})$. Then there exist functions $\varphi_k, \psi_k \in C^2((0,T) \times H)$ for $k = 1, 2, ...$ such that $\varphi_k, (\varphi_k)_t, B^{-1}D\varphi_k, D^2\varphi_k, \psi_k, (\psi_k)_t, B^{-1}D\psi_k, D^2\psi_k$ are bounded and uniformly continuous, and such that

$$u(t,x) - \varphi_k(t,x)$$

has a global maximum at some point (t_k, x_k),

$$v(s,y) - \psi_k(s,y)$$

has a global minimum at some point (s_k, y_k), and

$$\big(t_k, x_k, u(t_k,x_k), (\varphi_k)_t(t_k,x_k), D\varphi_k(t_k,x_k), D^2\varphi_k(t_k,x_k)\big)$$
$$\xrightarrow[\mathbb{R}\times H\times\mathbb{R}\times H_2\times\mathcal{L}(H_{-1};H_1)]{k\to\infty} \left(\bar{t}, \bar{x}, u(\bar{t},\bar{x}), \frac{\bar{t}-\bar{s}}{\beta}, \frac{B(\bar{x}_N - \bar{y}_N)}{\epsilon}, X_N\right) \quad (3.56)$$

$$\big(s_k, y_k, v(s_k,y_k), (\psi_k)_t(s_k,y_k), D\psi_k(s_k,y_k), D^2\psi_k(s_k,y_k)\big)$$
$$\xrightarrow[\mathbb{R}\times H\times\mathbb{R}\times H_2\times\mathcal{L}(H_{-1};H_1)]{k\to\infty} \left(\bar{s}, \bar{y}, v(\bar{s},\bar{y}), \frac{\bar{t}-\bar{s}}{\beta}, \frac{B(\bar{x}_N - \bar{y}_N)}{\epsilon}, Y_N\right) \quad (3.57)$$

where $X_N = P_N^* X_N P_N, Y_N = P_N^* Y_N P_N$ and they satisfy (3.54).

PROOF. We can obviously extend u, v to $\mathbb{R} \times H$ preserving all the properties of the functions and the strict global maximum at $(\bar{t}, \bar{x}, \bar{s}, \bar{y})$. We consider the space $\tilde{H} = \mathbb{R} \times H$ and the operator $\tilde{B} := I_{\mathbb{R}} \times B$. Writing (t,x) for elements of this extended space we now consider the function $\Phi(t,x_N,s,y_N) = \frac{1}{2\epsilon}|x_N - y_N|_{-1}^2 + \frac{1}{2\beta}(t-s)^2$. We now rescale time by setting

$$\tilde{u}(t,x) = u\left(\left(\frac{\beta}{\epsilon}\right)^{\frac{1}{2}} t, x\right), \quad \tilde{v}(s,y) = u\left(\left(\frac{\beta}{\epsilon}\right)^{\frac{1}{2}} s, y\right),$$

$$\tilde{\Phi}(t,x_N,s,y_N) = \Phi\left(\left(\frac{\beta}{\epsilon}\right)^{\frac{1}{2}} t, x_N, \left(\frac{\beta}{\epsilon}\right)^{\frac{1}{2}} s, y_N\right) = \frac{|x_N - y_N|_{-1}^2}{2\epsilon} + \frac{(t-s)^2}{2\epsilon}.$$

Then

$$\tilde{u}(t,x) - \tilde{v}(s,y) - \tilde{\Phi}(t,x_N,s,y_N)$$

has a strict global maximum over $\tilde{H} \times \tilde{H}$ at the point $((\frac{\epsilon}{\beta})^{\frac{1}{2}}\bar{t}, \bar{x}, (\frac{\epsilon}{\beta})^{\frac{1}{2}}\bar{s}, \bar{y})$. We can now apply Corollary 3.28 to produce the required functions φ_k, ψ_k. We now have have operators $\tilde{X}_N, \tilde{Y}_N$ satisfying a version of (3.54) on $\tilde{H} \times \tilde{H}$. However it is easy to see that its restriction to $\{0\} \times H \times \{0\} \times H$ gives (3.54). The claim follows after rescaling time back to the original variables which will only change the time derivatives of φ_k, ψ_k. □

REMARK 3.30 A different type of time dependent maximum principle can be obtained which relies on the finite dimensional parabolic maximum principle presented in Theorem E.11. Such a result can be found in [102], Theorem 3.2, however it is stated there in a version which uses second order parabolic jets and is applicable to equations with bounded terms (see Section 3.3.1). Theorem 3.2 in [102] also imposes an additional condition on the functions, which is satisfied when they are viscosity sub- and supersolutions of bounded time depended second order equations in a Hilbert space. The maximum principle stated in Corollary 3.29 does not impose extra conditions and thus it is more universal and can be applied more easily which is why we prefer it here. However the other maximum principle has certain advantages. For instance one can use it to prove comparison principles for bounded time dependent second order equations in a Hilbert space without the assumption that the viscosity subsolutions and supersolutions attain the initial/terminal values locally uniformly. Such results can be found in [290]. This type of maximum principle was also implicitly used in [421, 422]. ■

3.3. Viscosity solutions

Throughout this section U is an open subset of H and the operator B satisfies the following assumption (see Section 3.1.1).

HYPOTHESIS 3.31 *$B \in S(H)$ is a strictly positive operator such that A^*B is bounded.*

Contrary to the finite dimensional case, in infinite dimensions there is no one universal definition of viscosity solution. The basic idea of using pointwise maximum principle and replacing nonexistent derivatives of a solution by derivatives of test functions is still the same. However, because of the presence of unbounded terms and operators, the choice of test functions and the interpretation of unbounded terms must be often adjusted for different types of equations. In this section we present a generic definition of viscosity solution for a general class of stationary equations and time dependent Cauchy problems. The solutions defined below are called B-continuous viscosity solutions.

We consider the following boundary and terminal boundary value problems:

$$\begin{cases} -\langle Ax, Du\rangle + F(x,u,Du,D^2u) = 0 \quad \text{in } U \\ u(x) = f(x) \quad \text{on } \partial U \end{cases} \tag{3.58}$$

and

$$\begin{cases} u_t - \langle Ax, Du\rangle + F(t,x,u,Du,D^2u) = 0 \quad \text{in } (0,T)\times U \\ u(0,x) = g(x) \quad \text{for } x \in U, \\ u(t,x) = f_1(t,x) \quad \text{for } (t,x) \in (0,T)\times \partial U, \end{cases} \tag{3.59}$$

where $F : (0,T)\times U \times \mathbb{R} \times H \times S(H) \to \mathbb{R}$, $g : U \to \mathbb{R}$, $f : \partial U \to \mathbb{R}$, and $f_1 : (0,T)\times\partial U \to \mathbb{R}$ are continuous.

DEFINITION 3.32 *A function ψ is a test function if $\psi = \varphi + h(t,|x|)$, where:*

(i) *$\varphi \in C^{1,2}((0,T)\times U)$ is locally bounded, and is such that φ is B-lower semicontinuous, and $\varphi_t,\ A^*D\varphi,\ D\varphi,\ D^2\varphi$ are uniformly continuous on $(0,T)\times U$.*

(ii) *$h \in C^{1,2}((0,T)\times\mathbb{R})$ and is such that for every $t\in(0,T)$, $h(t,\cdot)$ is even and $h(t,\cdot)$ is non-decreasing on $[0,+\infty)$.*

For stationary equations φ and h are independent of t.

We remark that even though $|x|$ is not differentiable at 0, the function $h(t,|x|) \in C^{1,2}((0,T)\times H)$. The requirement that $\varphi_t, A^*D\varphi, D\varphi, D^2\varphi$ are uniformly continuous (and hence grow at most linearly at infinity) is a little arbitrary. It can be replaced by a requirement that they are locally uniformly continuous and have some prescribed growth at infinity, for instance at most polynomial. The growth restriction can also be removed, however it is useful in applications to stochastic optimal control since it is not clear if one can modify a test function φ outside a fixed set so that the modification has a required growth at infinity, while preserving the property that $A^*D\varphi$ is continuous. Thus the radial part h of test functions, which can be modified at will, plays the role of a cut-off function which takes care of the growth at infinity. We can thus require φ to be as nice as we want as long as our choice gives us enough test functions which are needed to build a good theory. The requirement that $A^*D\varphi$ is uniformly continuous can also be replaced by a requirement that $B^{-1}D\varphi$ is continuous. This however would make the definition more dependent on the choice of B. The reader can experiment with various modifications of the above definition and we will later see how the choice of test functions must be adjusted to particular cases.

DEFINITION 3.33 *A locally bounded B-upper semi-continuous (see Definition 3.3) function u on $\overline{U}$ is a viscosity subsolution of (3.58) if $u \le f$ on ∂U and whenever $u - \psi$ has a local maximum at a point x for a test function $\psi = \varphi + h(|x|)$ then*

$$-\langle x, A^*D\varphi(x)\rangle + F(x,u(x),D\psi(x),D^2\psi(x)) \le 0. \tag{3.60}$$

A locally bounded B-lower semi-continuous function u on $\overline{U}$ is a viscosity supersolution of (3.58) if $u \ge f$ on ∂U and whenever $u+\psi$ has a local minimum at a point x for a test function $\psi = \varphi + h(|x|)$ then

$$\langle x, A^*D\varphi(x)\rangle + F(x,u(x),-D\psi(x),-D^2\psi(x)) \ge 0. \tag{3.61}$$

A viscosity solution of (3.58) is a function which is both a viscosity subsolution and a viscosity supersolution of (3.58).

DEFINITION 3.34 *A locally bounded B-upper semi-continuous function u on $[0,T)\times\overline{U}$ is a viscosity subsolution of (3.59) if $u(0,y) \le g(y)$ for $y\in U$, $u \le f_1$ on $(0,T)\times\partial U$ and whenever $u-\psi$ has a local maximum at a point $(t,x)\in(0,T)\times U$ for a test function $\psi(s,y) = \varphi(s,y) + h(s,|y|)$ then*

$$\psi_t(t,x) - \langle x, A^*D\varphi(t,x)\rangle + F(t,x,u(t,x),D\psi(t,x),D^2\psi(t,x)) \le 0. \tag{3.62}$$

A locally bounded B-lower semi-continuous function u on $[0,T)\times\overline{U}$ is a viscosity supersolution of (3.59) if $u(0,y) \ge g(y)$ for $y\in U$, $u \ge f_1$ on $(0,T)\times\partial U$ and *whenever $u + \psi$ has a local minimum at a point $(t, x) \in (0, T) \times U$ for a test function $\psi(s, y) = \varphi(s, y) + h(s, |y|)$ then*

$$-\psi_t(t, x) + \langle x, A^* D\varphi(t, x)\rangle + F(t, x, u(t, x), -D\psi(t, x), -D^2\psi(t, x)) \geq 0. \quad (3.63)$$

A viscosity solution of (3.59) is a function which is both a viscosity subsolution and a viscosity supersolution of (3.59).

The main idea behind this definition of solution is the following. Test functions are split into two categories. Good test functions φ provide enough functions to apply the doubling argument in the proof of comparison and produce maxima and minima using perturbed optimization by functions in this class. Radial functions h are needed as cut-off functions to be able to produce local/global maxima and minima and to confine the region of their possible locations. As always in the theory of viscosity solutions, non-existing derivatives of u are replaced by existing derivatives of test functions. The term $\langle Ax, D\varphi(t, x)\rangle$ is interpreted as $\langle x, A^* D\varphi(t, x)\rangle$. We cannot do the same with the term $\langle Ax, Dh(t, |x|)\rangle$ since $Dh(t, |x|) = h_r(t, |x|)\frac{x}{|x|}$ (where h_r is the partial derivative of h with respect to the second variable) and we cannot hope in general that $x \in D(A^*)$ (nor that $x \in D(A)$). Therefore this term is dropped. This can be done effectively since the term $\frac{h_r(t,|x|)}{|x|}\langle Ax, x\rangle$ (or $\frac{h_r(t,|x|)}{|x|}\langle A^* x, x\rangle$) would be non-positive if it were well defined. Thus the definition is consistent with what the definition of viscosity solution should be under ideal conditions.

In applications to control problems it is more natural to work with terminal value problems instead of the initial value problems. A terminal value problem can be converted into an initial value problem by a change of variable $\tilde{t} := T - t$. Thus a terminal boundary value problem corresponding to (3.59) is

$$\begin{cases} u_t + \langle Ax, Du\rangle - F(t, x, u, Du, D^2u) = 0 \quad \text{in } (0, T) \times U \\ u(T, x) = g(x) \quad \text{for } x \in U, \\ u(t, x) = f(t, x) \quad \text{for } (t, x) \in (0, T) \times \partial U, \end{cases} \quad (3.64)$$

where $f(t, x) = f_1(T - t, x)$. Since we will be working with terminal value problems we state below the definition of viscosity solution adapted to this case, which is a consequence of Definition 3.34. (We keep the minus sign in front of the Hamiltonian F since we will formulate the conditions for F that will apply to both stationary and time dependent terminal value problems.)

Definition 3.35 *A locally bounded B-upper semi-continuous function u on $(0, T] \times \overline{U}$ is a viscosity subsolution of (3.64) if $u(T, y) \leq g(y)$ for $y \in U$, $u \leq f$ on $(0, T) \times \partial U$ and whenever $u - \psi$ has a local maximum at a point $(t, x) \in (0, T) \times U$ for a test function $\psi(s, y) = \varphi(s, y) + h(s, |y|)$ then*

$$\psi_t(t, x) + \langle x, A^* D\varphi(t, x)\rangle - F(t, x, u(t, x), D\psi(t, x), D^2\psi(t, x)) \geq 0. \quad (3.65)$$

A locally bounded B-lower semi-continuous function u on $(0, T] \times \overline{U}$ is a viscosity supersolution of (3.64) if $u(T, y) \geq g(y)$ for $y \in U$, $u \geq f$ on $(0, T) \times \partial U$ and whenever $u + \psi$ has a local minimum at a point $(t, x) \in (0, T) \times U$ for a test function $\psi(s, y) = \varphi(s, y) + h(s, |y|)$ then

$$-\psi_t(t, x) - \langle x, A^* D\varphi(t, x)\rangle - F(t, x, u(t, x), -D\psi(t, x), -D^2\psi(t, x)) \leq 0. \quad (3.66)$$

A viscosity solution of (3.64) is a function which is both a viscosity subsolution and a viscosity supersolution of (3.64).

REMARK 3.36 It is easy to see that if u is a viscosity subsolution (respectively, supersolution) of (3.64) on $(0,T]\times\overline{U}$ then it is a viscosity subsolution (respectively, supersolution) of (3.64) on $(T_1,T]\times\overline{U}$ for every $0<T_1<T$. ■

LEMMA 3.37 *Without loss of generality the maxima and minima in Definitions 3.33, 3.34, and 3.35 can be assumed to be global and strict.*

PROOF. We will only show it for the case of subsolution in Definition 3.33 as the other cases are similar. Let

$$u(x)-\varphi(x)-h(|x|)\geq u(y)-\varphi(y)-h(|y|) \quad \text{for } y\in B_R(x)\subset U$$

for some $R>0$.

We will show that there exist test functions $\tilde{\varphi}$ and $\tilde{h}(|\cdot|)$ such that $D\tilde{\varphi}(x)=D\varphi(x)$, $D^2\tilde{\varphi}(x)=D^2\varphi(x)$, $D\tilde{h}(|x|)=Dh(|x|)$, $D^2\tilde{h}(|x|)=D^2h(|x|)$, and $u-\tilde{\varphi}-\tilde{h}(|\cdot|)$ has a strict global maximum at x. Let $\eta\in C^2([0,\infty))$ be an increasing function such that

$$r+\sup_{|y|\leq r,\, y\in U}\{|u(y)|+|\varphi(y)|\}\leq \eta(r).$$

Let $g_1\in C^2((0,\infty))$ be a function such that

$$g_1(r)=\begin{cases} 0 & \text{if } r\leq |x| \\ (r-|x|)^4 & \text{if } |x|<r<|x|+1 \\ \text{increasing} & \text{if } |x|+1\leq r\leq |x|+2 \\ \eta(r) & \text{if } r>|x|+2. \end{cases}$$

Let $\varphi_1\in C^2([0,\infty))$ be defined by

$$\varphi_1(r)=\begin{cases} r^4 & r\leq 1, \\ \text{increasing} & 1<r<2, \\ 2 & r\geq 2. \end{cases}$$

Now for $n\geq 1$ consider the function

$$\Phi_n(y)=u(y)-\varphi(y)-n\varphi_1(|x-y|_{-1})-h(|y|)-g_1(|y|).$$

Obviously we have

$$\Phi_n(x)=u(x)-\varphi(x)-h(|x|).$$

Suppose there is a subsequence $n_k\to+\infty$ and $y_{n_k}\in U$, $y_{n_k}\neq x$ such that $\Phi_{n_k}(y_{n_k})\geq\Phi_{n_k}(x)$. Then we must have $|x-y_{n_k}|_{-1}\to 0$ as $k\to\infty$ and $|y_{n_k}|\leq C_1$ for some $C_1>0$, i.e. $y_{n_k}\rightharpoonup x$ and $By_{n_k}\to Bx$. Since $u,-\varphi$ are B-upper semicontinuous, and $h+g_1$ is increasing on $[|x|,+\infty)$, this implies that $|y_n|\to|x|$, and therefore $y_{n_k}\to x$ in H. But then $y_{n_k}\in B_R(x)$ for big k and so we get

$$\Phi_{n_k}(y_{n_k})<u(y_{n_k})-\varphi(y_{n_k})-h(|y_{n_k}|)\leq u(x)-\varphi(x)-h(|x|)$$

which is a contradiction. Therefore there must exist $\bar{n}$ such that $\Phi_{\bar{n}}(y)<\Phi_{\bar{n}}(x)$ for all $y\in U, y\neq x$. It then easily follows that $\Phi_{\bar{n}+1}$ has a strict global maximum at x. Therefore the conclusion follows with $\tilde{\varphi}(y)=\varphi(y)+(\bar{n}+1)\varphi_1(|x-y|_{-1})$ and $\tilde{h}(|y|)=h(|y|)+g_1(|y|)$. □

It follows from the proof of Lemma 3.37 that if we know a priori that u has certain growth at ∞ (at least quadratic) and $U=H$, we can then obtain the same growth for $\tilde{h}$. (We notice that if $U=H$ then φ has at most quadratic growth at infinity.) For instance if u has a polynomial growth at ∞ we can have $\tilde{h}$ which is a polynomial of some special form for big $|x|$. This can be important in applications to stochastic optimal control where we may want to impose additional conditions on test functions to be able to apply stochastic calculus. In these applications it may also be useful to assume that $h'(r)/r$ is globally bounded away from 0 for the radial test functions h. To avoid technical difficulties it may then be more convenient to choose h belonging to one particular class of functions, say certain polynomials with growth depending on the growth of sub- and super-solutions we are dealing with. However when using such narrow classes of radial test functions one may be forced to require that the maxima and minima in the definition of viscosity solution be global as the definitions using global and local maxima and minima may no longer be equivalent.

REMARK 3.38 We assumed in Definitions 3.33, 3.34, and 3.35 that A was a linear, densely defined, maximal dissipative operator in H, i.e. that it generated a C_0-semigroup of contractions e^{tA}. The definitions can be used to cover the case when $A - \omega I$ is maximal dissipative for some $\omega > 0$, i.e. if

$$\|e^{tA}\| \leq e^{\omega t} \quad \text{for all } t \geq 0. \tag{3.67}$$

One way to do it is to replace A by $\tilde{A} = A - \omega I$ and F by $\tilde{F}(t, x, r, p, X) = F(t, x, r, p, X) - \omega\langle x, p\rangle$. Another way is by making a change of variables $\tilde{u}(t, x) = u(t, e^{\omega t}x)$ in the equation which reduces equation (3.64) to an equation with A replaced by $A - \omega I$ and F replaced by $\tilde{F}(t, x, r, p, X) = F(t, e^{\omega t}x, e^{-\omega t}p, e^{-2\omega t}X)$. ■

LEMMA 3.39 *Let $F : [0, T) \times U \times \mathbb{R} \times H \times S(H) \to \mathbb{R}$ be continuous. Definition 3.35 is equivalent to the definition in which we only require that the maxima and minima be one-sided (i.e. that $u \mp \psi$ has a local maximum/minimum at a point (t, x) restricted to $[t, T) \times U$), if we also require that the subsolutions and supersolutions are continuous. In particular the equation is also satisfied at $t = 0$ if we in addition require in the case when a one-sided maximum/minimum is attained at $(0, x)$ that the test functions $\varphi \in C^{1,2}([0, T) \times U)$ and $h \in C^{1,2}([0, T) \times \mathbb{R})$.*

PROOF. Suppose that u is a viscosity subsolution of (3.64) in the sense of Definition 3.35 and let $u(s, y) - \varphi(s, y) - h(s, |y|) = u(s, y) - \psi(s, y)$ has a one sided local maximum at (t, x) over $[t, t + \epsilon) \times B_\epsilon(x)$. Arguing as in the proof of Lemma 3.37, we can assume that the one-sided local maximum is strict. Then for big n there exist by Corollary 3.26 (which we can apply because any closed convex subset of H is B-closed, see Remark 3.8) $a_n \in \mathbb{R}, p_n \in H$, $|a_n| + |p_n| \leq 1/n$ such that the function

$$u(s, y) - \psi(s, y) - \frac{1}{n(s - t)} - a_n s - \langle Bp_n, y\rangle$$

has a local (two-sided) maximum at $(s_n, y_n) \in (t, t + \epsilon) \times B_\epsilon(x)$. Since the initial local maximum was strict we have that $(s_n, y_n) \to (t, x)$. Without loss of generality we can assume that is $\frac{1}{n(s-t)}$ is a test function by modifying it around $s = t$ and then extending it to $(0, T)$. Therefore we obtain using Definition 3.35 that

$$\begin{aligned} \psi_t(s_n, y_n) + a_n - \frac{1}{n(s_n - t)^2} + \langle y_n, A^*(D\varphi(s_n, y_n) + Bp_n)\rangle \\ - F(s_n, y_n, u(s_n, y_n), D\psi(s_n, y_n) + Bp_n, D^2\psi(s_n, y_n)) \geq 0 \end{aligned}$$

which gives us

$$\psi_t(t, x) + \langle x, A^* D\varphi(t, x)\rangle - F(t, x, u(t, x), D\psi(t, x), D^2\psi(t, x)) \geq 0.$$

after letting $n \to +\infty$. □

3.3.1. Bounded equations. If $A = 0$ there is no need to use the notion of B-continuity. Viscosity solutions can then be defined in the same way as for finite dimensional problems. We present the definition for the time independent problem

$$\begin{cases} F(x,u,Du,D^2u) = 0 & \text{in } U \\ u(x) = f(x) & \text{on } \partial U. \end{cases} \tag{3.68}$$

The definition for time dependent problems is similar. We call such equations "bounded" since they do not contain any unbounded terms.

DEFINITION 3.40 *A locally bounded upper semi-continuous function u on $\overline{U}$ is a viscosity subsolution of (3.68) if $u \le f$ on ∂U and whenever $u - \varphi$ has a local maximum at a point x for a test function $\varphi \in C^2(U)$ then*

$$F(x, u(x), D\varphi(x), D^2\varphi(x)) \le 0.$$

A locally bounded lower semi-continuous function u on $\overline{U}$ is a viscosity supersolution of (3.68) if $u \ge f$ on ∂U and whenever $u - \varphi$ has a local minimum at a point x for a test function $\varphi \in C^2(U)$ then

$$F(x, u(x), D\varphi(x), D^2\varphi(x)) \ge 0.$$

A viscosity solution of (3.68) is a function which is both a viscosity subsolution and a viscosity supersolution of (3.68).

Equation (3.68) and its parabolic version were studied, together with the associated control problems, in [319, 321]. In [319] a stronger definition of viscosity solution was introduced, allowing for more general test functions which are not necessarily twice Fréchet differentiable. One can also replace Definition 3.40 with a definition using second-order jets (see [101, 321]). Regularity results for bounded equations and their obstacle problems have been obtained in [319, 426]. Existence and uniqueness results for such equations can also be found in [290].

3.4. Consistency of viscosity solutions

Consistency property of viscosity solutions, i.e. the ability to pass to limits in the equations under minimal assumptions on the solutions, is one of the greatest strengths of the notion of viscosity solution.

Let B be an operator satisfying Hypothesis 3.31. Let $A_n, n = 1, 2, ...,$ be linear, densely defined, maximal dissipative operators in H such that $D(A^*) \subset D(A_n^*)$. We consider equations

$$u_t + \langle A_n x, Du \rangle - F_n(t, x, u, Du, D^2u) = 0 \quad \text{in } (0,T) \times U, \tag{3.69}$$

where U is an open subset of H. We assume that viscosity sub- and super-solutions of (3.69) are B-upper (respectively, lower) semicontinuous with the same fixed B. We note that if φ is a test function in Definition 3.32-(i), then

$$A_n^* D\varphi = A_n^*(I - A^*)^{-1}(I - A^*)D\varphi,$$

and thus, since $A_n^*(I - A^*)^{-1} \in \mathcal{L}(H)$, φ is a test function of type (i) for equation (3.69).

We have the following result.

THEOREM 3.41 *Let the above assumptions about $A_n, n = 1, 2, ...$ be satisfied. Let $u_n, n = 1, 2, ...$ be a viscosity subsolutions (respectively, supersolutions) of (3.69) (with some terminal and boundary conditions which are not essential here). Suppose that $F_n : (0,T) \times U \times \mathbb{R} \times H \times S(H) \to \mathbb{R}, n = 1, 2, ...$ are continuous, and*

$$\text{whenever } x, x_n \in D(A^*),\ x_n \to x, \ \text{ and } A^* x_n \to A^* x, \ \text{ then } A_n^* x_n \to A^* x. \tag{3.70}$$

Let u_n converge locally uniformly to a function u on $(0,T) \times U$. Then u is a viscosity subsolution of

$$u_t + \langle Ax, Du \rangle - F_-(t, x, u, Du, D^2u) = 0 \quad \text{in } (0,T) \times U$$

(respectively, supersolution of

$$u_t + \langle Ax, Du\rangle - F^+(t,x,u,Du,D^2u) = 0 \quad \textit{in } (0,T)\times U,)$$

where

$$F_-(t,x,r,p,X) = \lim_{i\to+\infty} \inf \Big\{ F_n(\tau,y,s,q,Y) : n \geq i,$$
$$|t-\tau| + |x-y| + |r-s| + |p-q| + \|X-Y\| \leq \frac{1}{i} \Big\}, \quad (3.71)$$

$$F^+(t,x,r,p,X) = \lim_{i\to+\infty} \sup \Big\{ F_n(\tau,y,s,q,Y) : n \geq i,$$
$$|t-\tau| + |x-y| + |r-s| + |p-q| + \|X-Y\| \leq \frac{1}{i} \Big\}. \quad (3.72)$$

NOTATION 3.42 We denote the right-hand side of (3.71) and (3.72) respectively by

$$\liminf_{n\to+\infty}{}_* F_n(t,x,r,p,X),$$

and

$$\limsup_{n\to+\infty}{}^* F_n(t,x,r,p,X).$$

Obviously

$$\limsup_{n\to+\infty}{}^*(-F_n(t,x,r,p,X)) = -\liminf_{n\to+\infty}{}_* F_n(t,x,r,p,X).$$

■

PROOF. We will only do the proof for the subsolution case. The function u is obviously locally bounded and B-upper semi continuous. Suppose that $u(s,y) - \psi(s,y) = u(s,y) - \varphi(s,y) - h(s,|y|)$ has a local maximum at a point (t,x). By Lemma 3.37 the maximum can be assumed to be strict. Let $D = \{(s,y) : |t-s| \leq \delta, |x-y| \leq \delta\}$ for some $\delta > 0$. Applying Corollary 3.26 on D we obtain, for every n, $a_n \in \mathbb{R}, p_n \in H$ such that $|a_n| + |p_n| \leq \frac{1}{n}$ and such that

$$u_n(s,y) - (\psi(s,y) + a_n s + \langle Bp_n, y\rangle)$$

has maximum over D at some point (t_n,x_n). Since the original maximum at (t,x) was strict and the u_n converge uniformly on D to u, it is easy to see that we must have $(t_n,x_n) \to (t,x)$ as $n \to +\infty$. Since we have

$$\psi_t(t_n,x_n) + a_n + \langle x_n, A_n^*(D\varphi(t_n,x_n) + Bp_n)\rangle$$
$$-F_n(t_n,x_n,u_n(t_n,x_n), D\psi(t_n,x_n) + Bp_n, D^2\psi(t_n,x_n)) \geq 0,$$

the claim follows passing to the $\limsup_{n\to+\infty}{}^*$ in the above inequality. □

The result for time independent equations is similar. In finite dimensional spaces one can pass to weaker limits with viscosity solutions. In particular, the method of half-relaxed limits of Barles-Perthame (see [30], [101] and [195]) allows to conclude that for a family of subsolutions (respectively, supersolutions) u_n, the function $u^+ = \limsup_{n\to+\infty}{}^* u_n$ is a subsolution and $u_- = \liminf_{n\to+\infty}{}_* u_n$ is a supersolution. Unfortunately this is no longer true in infinite dimensions due to lack of local compactness. The following simple example from [424] illustrates this phenomenon. Half-relaxed limits in a special case are discussed in Section 3.9.

EXAMPLE 3.43 Let H be the real l^2 space. Let $F(p) = 1 - |p|$, and $u_n(x) = x_n$, where $x = (x_1, ..., x_n, ...)$. Then the functions u_n are classical (and thus viscosity) solutions of

$$F(Du_n) = 0.$$

However $u^+ \equiv 0$ and therefore $F(Du^+) = 1$, i.e. u^+ is not a subsolution of $F(Du^+) = 0$. To see that $u^+ \equiv 0$ we notice that if $n \geq i$ and $|x - y| \leq 1/i$ then

$$|u_n(y)| = |y_n| \leq |x_n| + \frac{1}{i} \to 0 \quad \text{as } n, i \to \infty.$$

■

3.5. Comparison theorems

In this section we present comparison results for viscosity solutions. They are proved under either the weak or the strong B-condition for A (see Definitions 3.9 and 3.10).

We use the notation from Section 3.2. In particular, we recall that $\{e_1, e_2, ...\}$ is an orthonormal basis in H_{-1} made of elements of H, and for $N > 2$, $H_N = \text{span}\{e_1, ..., e_N\}$, P_N is the orthogonal projection in H_{-1} onto H_N, and $Q_N := I - P_N$.

We will make the following assumptions about the function $F : (0, T) \times U \times \mathbb{R} \times H \times S(H) \to \mathbb{R}$.

HYPOTHESIS 3.44 *F is uniformly continuous on bounded subsets of* $(0, T) \times U \times \mathbb{R} \times H \times S(H)$.

HYPOTHESIS 3.45 *There exists $\nu \geq 0$ such that for every $(t, x, p, X) \in (0, T) \times U \times H \times S(H)$*

$$F(t, x, r, p, X) - F(t, x, s, p, X) \geq \nu(r - s) \quad \text{when } r \geq s.$$

HYPOTHESIS 3.46 *For every $(t, x, r, p) \in (0, T) \times U \times \mathbb{R} \times H$*

$$F(t, x, r, p, X) \geq F(t, x, r, p, Y) \quad \text{when } X \leq Y.$$

HYPOTHESIS 3.47 *For all $t \in (0, T)$, $r \in \mathbb{R}$, $x \in U$, $p \in H$, $R > 0$,*

$$\sup \Big\{ |F(t, x, p, X + \lambda B Q_N) - F(t, x, p, X)| : \\ \|X\|, |\lambda| \leq R, X = P_N^* X P_N \Big\} \xrightarrow{N \to +\infty} 0. \quad (3.73)$$

HYPOTHESIS 3.48 *For every $R > 0$ there exists a modulus ω_R such that, for all $(t, x, y, r) \in (0, T) \times U \times U \times \mathbb{R}$ such that $|r|$, $|x|$, $|y| \leq R$, for any $\epsilon > 0$, for all $X, Y \in S(H)$ such that $X = P_N^* X P_N, Y = P_N^* Y P_N$ for some N and satisfying (3.54), we have*

$$F\left(t, x, r, \frac{B(x - y)}{\epsilon}, X\right) - F\left(t, y, r, \frac{B(x - y)}{\epsilon}, Y\right) \\ \geq -\omega_R\left(|x - y|_{-1}\left(1 + \frac{|x - y|_{-1}}{\epsilon}\right)\right).$$

HYPOTHESIS 3.49 *There exist $\gamma \in [0, 1]$ and a constant $M_F \geq 0$ such that*

$$|F(t, x, r, p + q, X + Y) - F(t, x, r, p, X)| \\ \leq M_F\left((1 + |x|)|q| + (1 + |x|^\gamma)^2 \|Y\|\right)$$

for all $(t, x, r) \in (0, T) \times U \times \mathbb{R}$, $p, q \in H$, $X, Y \in S(H)$.

Hypothesis 3.45 guarantees that F is nondecreasing in the zeroth order variable. If $\nu > 0$ we say that F is proper. Hypothesis 3.46 ensures that F is monotone in the second order variable. When it is satisfied we say that F (and therefore the equation) is degenerate elliptic/parabolic.

3.5.1. Degenerate parabolic equations. In Theorem 3.50, the boundary and terminal value functions f and g are not explicitly mentioned since they are not relevant. However the subsolution function u and the supersolution function v are defined on $(0,T] \times \overline{U}$ and conditions (3.74) and (3.75) describe their joint behavior along the boundary ∂U and the terminal value T.

THEOREM 3.50 (Comparison under weak B-condition) *Let $U \subset H$ be open and $\overline{U}$ be B-closed (see Definition 3.7). Let (3.2) hold and let F satisfy Hypotheses 3.44, 3.46, 3.47, 3.48 and 3.49 and 3.45 with $\nu = 0$. Let u be a viscosity subsolution of (3.64), and v be a viscosity supersolution of (3.64). Suppose that for every $R > 0$ there exists a modulus $\tilde{\omega}_R$ such that*

$$(u(t,x) - v(s,y))_+ + (u(t,y) - v(s,x))_+ \le \tilde{\omega}_R(|t-s| + |x-y|_{-1}) \tag{3.74}$$

for $t,s \in (0,T), x \in \partial U, y \in \overline{U}, |x|, |y| \le R$, and that

$$\lim_{R\to+\infty} \lim_{r\to 0} \lim_{\eta\to 0} \sup \Big\{ u(t,x) - v(s,y) : \ |x-y|_{-1} < r \\ x,y \in \overline{U} \cap B_R,\ T-\eta \le t,\ s \le T \Big\} \le 0. \tag{3.75}$$

Moreover suppose that there exist constants $C, a > 0$ such that

$$u, -v \le Ce^{a|x|^{2-2\gamma}} \quad (t,x) \in (0,T) \times H, \ \text{ if } \gamma \in [0,1), \tag{3.76}$$

and

$$u, -v \le Ce^{a(\log(1+|x|))^2} \quad (t,x) \in (0,T) \times H, \ \text{ if } \gamma = 1. \tag{3.77}$$

Then for every $\kappa > 0$

$$\lim_{R\to+\infty} \lim_{r\to 0} \lim_{\eta\to 0} \sup \Big\{ u(t,x) - v(s,y) : \ |x-y|_{-1} < r, |t-s| < \eta \\ x,y \in \overline{U} \cap B_R, \kappa < t, s \le T \Big\} \le 0. \tag{3.78}$$

In particular $u \le v$.

REMARK 3.51 It is easy to see that (3.78) implies that for every $\kappa > 0$ and $R > 0$ there exists a modulus $\tilde{\omega}_R$ such that

$$u(t,x) - v(s,y) \le \tilde{\omega}_{\kappa,R}(|x-y|_{-1} + |t-s|) \quad \text{for } x,y \in \overline{U} \cap B_R, \kappa < t, s \le T. \tag{3.79}$$

■

PROOF OF THEOREM 3.50. Let $0 < \tau < 1$ be such that $a < 1/\sqrt{\tau}$. Additional condition on τ will be given later. Set $T_1 = T - \tau$. The proof will be done in several steps. We will first show (3.78) for $T_1 \le s, t \le T$ and then reapply the procedure to intervals $[T - 3\tau/2, T - \tau/2], [T - 4\tau/2, T - 2\tau/2]$,.... We will first do the proof for the case $\gamma = 1$.

We argue by contradiction and assume that (3.78) is not true. Then there is $\kappa > 0$ such that

$$m = \lim_{R\to+\infty} \lim_{r\to 0} \lim_{\eta\to 0} \sup \Big\{ u(t,x) - v(s,y) : \ |x-y|_{-1} < r, |t-s| < \eta \\ x,y \in \overline{U} \cap B_R,\ T_1 + \kappa \le t, s \le T \Big\} > 0.$$

We note that m can be $+\infty$. Denote

$$m_\delta := \lim_{r\to 0}\lim_{\eta\to 0}\sup\Big\{u(t,x)-v(s,y)-\delta e^{\frac{(\log(2+|x|^2))^2}{\sqrt{t-T_1}}}-\delta e^{\frac{(\log(2+|y|^2))^2}{\sqrt{s-T_1}}}:$$
$$|x-y|_{-1}<r,\ |t-s|<\eta,\ x,y\in\overline{U},\ T_1\le t,s\le T\Big\},$$

$$m_{\delta,\epsilon} := \lim_{\eta\to 0}\sup\Big\{u(t,x)-v(s,y)-\delta e^{\frac{(\log(2+|x|^2))^2}{\sqrt{t-T_1}}}-\delta e^{\frac{(\log(2+|y|^2))^2}{\sqrt{s-T_1}}}$$
$$-\frac{|x-y|_{-1}^2}{2\epsilon}:\ |t-s|<\eta,\ x,y\in\overline{U},\ T_1\le t,s\le T\Big\},$$

$$m_{\delta,\epsilon,\beta} := \sup\Big\{u(t,x)-v(s,y)-\delta e^{\frac{(\log(2+|x|^2))^2}{\sqrt{t-T_1}}}-\delta e^{\frac{(\log(2+|y|^2))^2}{\sqrt{s-T_1}}}$$
$$-\frac{|x-y|_{-1}^2}{2\epsilon}-\frac{(t-s)^2}{2\beta}:\ x,y\in\overline{U},\ T_1\le t,s\le T\Big\}.$$

It is very easy to see that

$$m \le \lim_{\delta\to 0} m_\delta, \tag{3.80}$$
$$m_\delta = \lim_{\epsilon\to 0} m_{\delta,\epsilon}, \tag{3.81}$$
$$m_{\delta,\epsilon} = \lim_{\beta\to 0} m_{\delta,\epsilon,\beta}. \tag{3.82}$$

Setting $u(t,x)=-\infty$ if $x\notin\overline{U}$ and $v(t,x)=+\infty$ if $x\notin\overline{U}$ we can consider u and v to be defined on $(T_1,T]\times H$. Since $\overline{U}$ is B-closed such extended u is B-upper semi-continuous on $(T_1,T]\times H$ and v is B-lower semi-continuous on $(T_1,T]\times H$.

Define

$$\Psi(t,s,x,y)=u(t,x)-v(s,y)-\delta e^{\frac{(\log(2+|x|^2))^2}{\sqrt{t-T_1}}}-\delta e^{\frac{(\log(2+|y|^2))^2}{\sqrt{s-T_1}}}-\frac{|x-y|_{-1}^2}{2\epsilon}-\frac{(t-s)^2}{2\beta}.$$

We notice that by (3.77) we obtain for instance

$$\Psi(t,s,x,y)\le -(|x|^2+|y|^2)\quad \text{if } |x|+|y|\ge K_\delta \tag{3.83}$$

for some $K_\delta>0$. Therefore, using Corollary 3.26, for every $n\ge 1$ we can find $a_n,b_n\in\mathbb{R}$, and $p_n,q_n\in H$ such that $|a_n|+|b_n|+|p_n|+|q_n|\le\frac{1}{n}$ such that

$$\Psi(t,s,x,y)+a_nt+b_ns+\langle Bp_n,x\rangle+\langle Bq_n,y\rangle$$

achieves a strict global maximum at some point $(\bar t,\bar s,\bar x,\bar y)\in[T_1,T]\times[T_1,T]\times H\times H$. (The maximum is initially strict in the $|\cdot|_{-2}$ norm but since the radial functions are strictly increasing the maximum is in fact strict in the $|\cdot|$ norm.) Moreover for a fixed δ

$$|\bar x|,\ |\bar y|,\ |u(\bar t,\bar x)|,\ |v(\bar s,\bar y)|\le R_\delta \tag{3.84}$$

for some R_δ independently of ϵ,β,n. Obviously $(\bar t,\bar s,\bar x,\bar y)\in(T_1,T]\times(T_1,T]\times\overline{U}\times\overline{U}$. It follows from (3.83) that

$$m_{\delta,\epsilon,\beta}\le\Psi(\bar t,\bar s,\bar x,\bar y)+\frac{C_\delta}{n} \tag{3.85}$$

for some constant $C_\delta>0$. Therefore it follows that

$$m_{\delta,\epsilon,\beta}+\frac{|\bar t-\bar s|^2}{4\beta}\le\Psi(\bar t,\bar s,\bar x,\bar y)+\frac{|\bar t-\bar s|^2}{4\beta}+\frac{C_\delta}{n}\le m_{\delta,\epsilon,2\beta}+\frac{C_\delta}{n} \tag{3.86}$$

and

$$m_{\delta,\epsilon,\beta}+\frac{|\bar x-\bar y|_{-1}^2}{4\epsilon}+\frac{|\bar t-\bar s|^2}{4\beta}\le m_{\delta,2\epsilon,2\beta}+\frac{C_\delta}{n}. \tag{3.87}$$

Inequalities (3.86) and (3.82) imply

$$\lim_{\beta\to 0}\limsup_{n\to+\infty}\frac{|\bar t-\bar s|^2}{\beta}=0\quad\text{for every }\delta,\epsilon>0,\tag{3.88}$$

and then (3.88), (3.87) and (3.81) imply

$$\lim_{\epsilon\to 0}\limsup_{\beta\to 0}\limsup_{n\to\infty}\frac{|\bar x-\bar y|_{-1}^2}{\epsilon}=0\quad\text{for every }\delta>0.\tag{3.89}$$

In particular it follows from (3.80), (3.81), (3.82), (3.85), (3.88) and (3.89) that there exists $\delta_0>0$ such that for all $\delta<\delta_0$

$$\liminf_{\epsilon\to 0}\liminf_{\beta\to 0}\liminf_{n\to\infty}(u(\bar t,\bar x)-v(\bar s,\bar y))\geq\bar m=\min\left(\frac{m}{2},1\right).\tag{3.90}$$

Conditions (3.74), (3.75), together with (3.88) and (3.89), imply that if δ,ϵ,β are small enough and n is sufficiently big we must have $(\bar t,\bar s,\bar x,\bar y)\in(T_1,T)\times(T_1,T)\times U\times U$.

We now have for $N>2$

$$|x-y|_{-1}^2=|P_N(x-y)|_{-1}^2+|Q_N(x-y)|_{-1}^2$$

and

$$\begin{aligned}|Q_N(x-y)|_{-1}^2\leq 2\langle BQ_N(\bar x-\bar y),x-y\rangle+2|Q_N(x-\bar x)|_{-1}^2\\+2|Q_N(y-\bar y)|_{-1}^2-|Q_N(\bar x-\bar y)|_{-1}^2\end{aligned}$$

with equality at $\bar x,\bar y$. Therefore, defining

$$\begin{aligned}u_1(t,x)=u(t,x)-\delta e^{\frac{(\log(2+|x|^2))^2}{\sqrt{t-T_1}}}-\frac{\langle BQ_N(\bar x-\bar y),x\rangle}{\epsilon}-\frac{|Q_N(x-\bar x)|_{-1}^2}{\epsilon}\\+\frac{|Q_N(\bar x-\bar y)|_{-1}^2}{2\epsilon}+a_nt+\langle Bp_n,x\rangle\end{aligned}$$

and

$$\begin{aligned}v_1(s,y)=v(s,y)+\delta e^{\frac{(\log(2+|y|^2))^2}{\sqrt{s-T_1}}}-\frac{\langle BQ_N(\bar x-\bar y),y\rangle}{\epsilon}+\frac{|Q_N(y-\bar y)|_{-1}^2}{\epsilon}\\-b_ns-\langle Bq_n,y\rangle.\end{aligned}$$

we see that

$$u_1(t,x)-v_1(s,y)-\frac{1}{2\epsilon}|P_N(x-y)|_{-1}^2-\frac{1}{2\beta}|t-s|^2$$

has a strict global maximum at $(\bar t,\bar s,\bar x,\bar y)$ over $[T_1,T]\times[T_1,T]\times H\times H$. We can therefore apply Corollary 3.29 to obtain test functions φ_k,ψ_k and points $(t_k,x_k),(s_k,y_k)$ such that $u_1(t,x)-\varphi_k(t,x)$ has a maximum at (t_k,x_k), $v_1(s,y)-\psi_k(s,y)$ has a minimum at (s_k,y_k), and such that (3.56), (3.57) are satisfied for u_1,v_1 respectively. In particular $(t_k,x_k),(s_k,y_k)\in(T_1,T)\times U$ for big k.

Define

$$\begin{aligned}\varphi(t,x)=\varphi_k(t,x)+\frac{\langle BQ_N(\bar x-\bar y),x\rangle}{\epsilon}+\frac{|Q_N(x-\bar x)|_{-1}^2}{\epsilon}\\-\frac{|Q_N(\bar x-\bar y)|_{-1}^2}{2\epsilon}-a_nt-\langle Bp_n,x\rangle,\end{aligned}\tag{3.91}$$

and

$$h(t,|x|)=\delta e^{\frac{(\log(2+|x|^2))^2}{\sqrt{t-T_1}}}.$$

Since u is a viscosity subsolution of (3.64) on $(T_1, T] \times \overline{U}$, using the definition of viscosity subsolution we have

$$\begin{aligned}&\varphi_t(t_k, x_k) + h_t(t_k, |x_k|) + \langle x_k, A^* D\varphi(t_k, x_k)\rangle \\ &- F(t_k, x_k, u(t_k, x_k), D\varphi(t_k, x_k) + Dh(t_k, |x_k|), D^2\varphi(t_k, x_k) + D^2 h(t_k, |x_k|)) \geq 0.\end{aligned} \tag{3.92}$$

Letting $k \to +\infty$ in (3.92) and using (3.56) yields

$$\begin{aligned}&\frac{\bar{t} - \bar{s}}{\beta} - a_n + h_t(\bar{t}, |\bar{x}|) + \left\langle \bar{x}, A^* \left(\frac{B(\bar{x} - \bar{y})}{\epsilon} - Bp_n \right) \right\rangle \\ &-F\left(\bar{t}, \bar{x}, u(\bar{t}, \bar{x}), \frac{B(\bar{x} - \bar{y})}{\epsilon} - Bp_n + Dh(\bar{t}, |\bar{x}|), X_N + \frac{2BQ_N}{\epsilon} + D^2 h(\bar{t}, |\bar{x}|) \right) \geq 0.\end{aligned} \tag{3.93}$$

We now compute

$$h_t(t, |x|) = -\frac{\delta(\log(2 + |x|^2))^2}{2(t - T_1)^{\frac{3}{2}}} e^{\frac{(\log(2+|x|^2))^2}{\sqrt{t-T_1}}},$$

$$Dh(t, |x|) = e^{\frac{(\log(2+|x|^2))^2}{\sqrt{t-T_1}}} \frac{4\delta \log(2 + |x|^2)}{\sqrt{t - T_1}} \frac{x}{2 + |x|^2},$$

and

$$\begin{aligned}D^2 h(t, |x|) = \frac{4\delta}{\sqrt{t - T_1}} e^{\frac{(\log(2+|x|^2))^2}{\sqrt{t-T_1}}} &\left[\left(\frac{4(\log(2 + |x|^2))^2}{\sqrt{t - T_1}(2 + |x|^2)^2} \right.\right. \\ &\left.\left. + \frac{2}{(2 + |x|^2)^2} - \frac{2\log(2 + |x|^2)}{(2 + |x|^2)^2} \right) x \otimes x + \frac{\log(2 + |x|^2)}{2 + |x|^2} I \right].\end{aligned}$$

We have

$$\frac{|Dh(t, |x|)|}{1 + |x|} + \|D^2 h(t, |x|)\| \leq \frac{C_1 \delta(\log(2 + |x|^2))^2}{(t - T_1)(1 + |x|)^2} e^{\frac{(\log(2+|x|^2))^2}{\sqrt{t-T_1}}}$$

for some absolute constant C_1. Therefore we obtain from Hypothesis 3.49

$$\begin{aligned}&\left| F\left(\bar{t}, \bar{x}, u(\bar{t}, \bar{x}), \frac{B(\bar{x} - \bar{y})}{\epsilon} - Bp_n + Dh(\bar{t}, |\bar{x}|), X_N + \frac{2BQ_N}{\epsilon} + D^2 h(\bar{t}, |\bar{x}|) \right) \right. \\ &\quad \left. - F\left(\bar{t}, \bar{x}, u(\bar{t}, \bar{x}), \frac{B(\bar{x} - \bar{y})}{\epsilon} - Bp_n, X_N + \frac{2BQ_N}{\epsilon} \right) \right| \\ &\quad \leq M_F((1 + |\bar{x}|)|Dh(\bar{t}, |\bar{x}|)| + (1 + |\bar{x}|)^2 \|D^2 h(\bar{t}, |\bar{x}|)\|) \\ &\quad \quad \leq \frac{M_F C_1 \delta(\log(2 + |\bar{x}|^2))^2}{\bar{t} - T_1} e^{\frac{(\log(2+|\bar{x}|^2))^2}{\sqrt{\bar{t}-T_1}}} \leq -\frac{1}{2} h_t(\bar{t}, |\bar{x}|)\end{aligned}$$

if

$$\tau \leq \frac{1}{(4M_F C_1)^2}. \tag{3.94}$$

Hence if (3.94) is satisfied, using that $\frac{1}{2} h_t(\bar{t}, |\bar{x}|) \leq -C_\tau \delta$ for some $C_\tau > 0$, it follows from (3.93) that

$$\begin{aligned}&- C_\tau \delta + \frac{\bar{t} - \bar{s}}{\beta} - a_n + \left\langle \bar{x}, A^* \left(\frac{B(\bar{x} - \bar{y})}{\epsilon} - Bp_n \right) \right\rangle \\ &\quad - F\left(\bar{t}, \bar{x}, u(\bar{t}, \bar{x}), \frac{B(\bar{x} - \bar{y})}{\epsilon} - Bp_n, X_N + \frac{2BQ_N}{\epsilon} \right) \geq 0.\end{aligned} \tag{3.95}$$

Arguing similarly we obtain from the fact that $v_1(s,y) - \psi_k(s,y)$ has a minimum at (s_k, y_k) that

$$\frac{\bar{t}-\bar{s}}{\beta} + b_n + \left\langle \bar{y}, A^*\left(\frac{B(\bar{x}-\bar{y})}{\epsilon} + Bq_n\right)\right\rangle - F\left(\bar{s}, \bar{y}, v(\bar{s},\bar{y}), \frac{B(\bar{x}-\bar{y})}{\epsilon} + Bq_n, Y_N - \frac{2BQ_N}{\epsilon}\right) \le 0. \quad (3.96)$$

Therefore subtracting (3.95) from (3.96) and using (3.84), Hypothesis 3.47 yield

$$C_\tau \delta - \left\langle \bar{x}-\bar{y}, \frac{A^*B(\bar{x}-\bar{y})}{\epsilon}\right\rangle + F\left(\bar{t}, \bar{x}, u(\bar{t},\bar{x}), \frac{B(\bar{x}-\bar{y})}{\epsilon}, X_N\right) - F\left(\bar{s}, \bar{y}, v(\bar{s},\bar{y}), \frac{B(\bar{x}-\bar{y})}{\epsilon}, Y_N\right) \le \omega_1(\delta, \epsilon, \beta; n, N), \quad (3.97)$$

where $\lim_{n\to+\infty} \lim_{N\to+\infty} \omega_1(\delta,\epsilon,\beta;n,N) = 0$ for fixed δ, ϵ, β. Now Hypothesis 3.45, (3.88), (3.90) and (3.97) imply

$$C_\tau \delta - \left\langle \bar{x}-\bar{y}, \frac{A^*B(\bar{x}-\bar{y})}{\epsilon}\right\rangle + F\left(\bar{t}, \bar{x}, u(\bar{t},\bar{x}), \frac{B(\bar{x}-\bar{y})}{\epsilon}, X_N\right) - F\left(\bar{t}, \bar{y}, u(\bar{t},\bar{x}), \frac{B(\bar{x}-\bar{y})}{\epsilon}, Y_N\right) \le \omega_2(\delta; \epsilon, \beta, n, N), \quad (3.98)$$

where $\limsup_{\epsilon\to 0} \limsup_{\beta\to 0} \limsup_{n\to+\infty} \limsup_{N\to+\infty} \omega_2(\delta;\epsilon,\beta,n,N) = 0$ for sufficiently small δ. We recall that X_N, Y_N satisfy (3.54). We can now use (3.2), Hypothesis 3.48, (3.84) and then invoke (3.89) to get

$$C_\tau \delta \le c_0 \frac{|\bar{x}-\bar{y}|_{-1}^2}{\epsilon} + \omega_{R_\delta}\left(|\bar{x}-\bar{y}|_{-1}\left(1 + \frac{|\bar{x}-\bar{y}|_{-1}}{\epsilon}\right)\right) + \omega_2(\delta;\epsilon,\beta,n,N) \le \omega_3(\delta;\epsilon,\beta,n,N),$$

where $\limsup_{\epsilon\to 0} \limsup_{\beta\to 0} \limsup_{n\to+\infty} \limsup_{N\to+\infty} \omega_3(\delta;\epsilon,\beta,n,N) = 0$ for sufficiently small δ. This yields a contradiction for small δ.

Thus we obtain that $m \le 0$ and this allows us to reapply the procedure to intervals $[T-3\tau/2, T-\tau/2], [T-4\tau/2, T-2\tau/2], ..., [0, T-k\tau/2]$, where k is such that $T - k\tau/2 > 0 \ge T-(k+2)\tau/2$.

For $\gamma \in [0,1)$ the proof is the same but we have to replace the functions

$$\delta e^{\frac{(\log(2+|x|^2))^2}{\sqrt{t-T_1}}} \quad \text{and} \quad \delta e^{\frac{(\log(2+|y|^2))^2}{\sqrt{s-T_1}}}$$

by

$$\delta e^{\frac{(1+|x|^2)^{1-\gamma}}{\sqrt{t-T_1}}} \quad \text{and} \quad \delta e^{\frac{(1+|y|^2)^{1-\gamma}}{\sqrt{s-T_1}}},$$

respectively. □

The assumptions of the comparison theorems can be weakened if we replace the weak B condition by the strong B-condition i.e. if we replace (3.2) by (3.3). In this case we will use the following assumption instead of Hypothesis 3.48.

HYPOTHESIS 3.52 *For every $R > 0$ there exists a modulus ω_R such that, for all $(t,x,y,r) \in (0,T) \times U \times U \times \mathbb{R}$ such that $|r|$, $|x|$, $|y| \le R$, for any $\epsilon > 0$, for all*

$X, Y \in S(H)$ such that $X = P_N^* X P_N, Y = P_N^* Y P_N$ for some N and satisfying (3.54), we have

$$F\left(t, x, r, \frac{B(x-y)}{\epsilon}, X\right) - F\left(t, y, r, \frac{B(x-y)}{\epsilon}, Y\right)$$
$$\geq -\omega_R\left(|x-y|\left(1 + \frac{|x-y|_{-1}}{\epsilon}\right)\right).$$

HYPOTHESIS 3.53 *For every $R > 0$ there exists a modulus ω_R such that*

$$|g(x) - g(y)| \leq \omega_R(|x-y|) \quad \text{if } x, y \in H, |x|, |y| \leq R.$$

THEOREM 3.54 (Comparison under B-strong condition) *Let $U = H$. Let (3.3) hold and let F satisfy Hypotheses 3.44, 3.46, 3.47, 3.52, 3.49 and 3.45 with $\nu = 0$. Let g satisfy Hypothesis 3.53. Let u be a viscosity subsolution of (3.64) in $(0, T] \times H$, and v be a viscosity supersolution of (3.64) in $(0, T] \times H$. Suppose that*

$$\lim_{t \to T}\left[(u(t,x) - g(e^{(T-t)A}x))_+ + (v(t,x) - g(e^{(T-t)A}x))_-\right] = 0 \tag{3.99}$$

uniformly on bounded subsets of H and that either of (3.76) or (3.77) is satisfied. Then for every $0 < \mu < T$

$$m_\mu = \lim_{R \to +\infty} \lim_{r \to 0} \lim_{\eta \to 0} \sup\Big\{u(t,x) - v(s,y) : \ |x-y|_{-1} < r, \ |t-s| < \eta$$
$$x, y \in B_R, \ \mu < t, s \leq T - \mu\Big\} \leq 0. \tag{3.100}$$

In particular $u \leq v$.

PROOF. Let us again assume that $\gamma = 1$ and that (3.76) is satisfied. As in the proof of Theorem 3.50 we take $0 < \tau \leq \min(1, 1/\sqrt{a}, 1/(4M_F C_1)^2$, where C_1 is the constant appearing in that proof, and we set $T_1 = T - \tau$. Define for $0 < \mu < \tau$

$$m_\mu = \lim_{R \to +\infty} \lim_{r \to 0} \lim_{\eta \to 0} \sup\Big\{u(t,x) - v(s,y) : \ |x-y|_{-1} < r, \ |t-s| < \eta$$
$$x, y \in B_R, \ T_1 + \mu \leq t, s \leq T - \mu\Big\}.$$

If there is μ_0 such that $m_{\mu_0} > \tilde{m} > 0$ then $m_\mu > \tilde{m} > 0$ for all $0 < \mu < \mu_0$. Defining

$$\Psi(t,s,x,y) = u(t,x) - v(s,y) - \delta e^{\frac{(\log(2+|x|^2))^2}{\sqrt{t-T_1}}} - \delta e^{\frac{(\log(2+|y|^2))^2}{\sqrt{s-T_1}}} - \frac{|x-y|_{-1}^2}{2\epsilon} - \frac{(t-s)^2}{2\beta}$$

we again have that for every $n \geq 1$ there exist $a_n, b_n \in \mathbb{R}, p_n, q_n \in H$ such that $|a_n| + |b_n| + |p_n| + |q_n| \leq \frac{1}{n}$ and that

$$\Psi(t,s,x,y) + a_n t + b_n s + \langle B p_n, x\rangle + \langle B q_n, y\rangle$$

achieves a strict global maximum over $[T_1, T-\mu] \times [T_1, T-\mu] \times H \times H$ at some point $(\bar{t}_\mu, \bar{s}_\mu, \bar{x}_\mu, \bar{y}_\mu) \in (T_1, T-\mu] \times (T_1, T-\mu] \times H \times H$, and that (3.84), (3.88), (3.89) hold (note that the constant R_δ in (3.84) is independent of μ). Moreover, since for $\mu < \mu_0$, $\Psi(\bar{t}_\mu, \bar{s}_\mu, \bar{x}_\mu, \bar{y}_\mu) \geq \Psi(\bar{t}_{\mu_0}, \bar{s}_{\mu_0}, \bar{x}_{\mu_0}, \bar{y}_{\mu_0}) - \frac{C_\delta}{n}$ for some $C_\delta > 0$ independent of μ it is easy to see that for every $0 < \mu < \mu_0$

$$\begin{aligned}\liminf_{\epsilon \to 0} \liminf_{\beta \to 0} \liminf_{n \to +\infty} & (u(\bar{t}_\mu, \bar{x}_\mu) - v(\bar{s}_\mu, \bar{y}_\mu)) \\ &= \liminf_{\epsilon \to 0} \liminf_{\beta \to 0} \liminf_{n \to +\infty} \Psi(\bar{t}_\mu, \bar{s}_\mu, \bar{x}_\mu, \bar{y}_\mu) \\ &\geq \liminf_{\epsilon \to 0} \liminf_{\beta \to 0} \liminf_{n \to +\infty} \Psi(\bar{t}_{\mu_0}, \bar{s}_{\mu_0}, \bar{x}_{\mu_0}, \bar{y}_{\mu_0}) \\ &= \liminf_{\epsilon \to 0} \liminf_{\beta \to 0} \liminf_{n \to +\infty} (u(\bar{t}_{\mu_0}, \bar{x}_{\mu_0}) - v(\bar{s}_{\mu_0}, \bar{y}_{\mu_0})) \geq \tilde{m}\end{aligned} \tag{3.101}$$

if $\delta < \delta_0$ for some $\delta_0 > 0$ (depending only on μ_0).

If for all $\delta, \epsilon, \beta, n$ we have $(\bar{t}_\mu, \bar{s}_\mu, \bar{x}_\mu, \bar{y}_\mu) \in (T_1, T-\mu) \times (T_1, T-\mu) \times H \times H$ then as in the proof of Theorem 3.50 and using the notation there we arrive at (3.98), i.e. that

$$
\begin{aligned}
&C_\tau \delta - \left\langle \bar{x}_\mu - \bar{y}_\mu, \frac{A^* B(\bar{x}_\mu - \bar{y}_\mu)}{\epsilon} \right\rangle \\
&+ F\left(\bar{t}_\mu, \bar{x}_\mu, u(\bar{t}_\mu, \bar{x}_\mu), \frac{B(\bar{x}_\mu - \bar{y}_\mu)}{\epsilon}, X_N\right) - F\left(\bar{t}_\mu, \bar{y}_\mu, u(\bar{t}_\mu, \bar{x}_\mu), \frac{B(\bar{x}_\mu - \bar{y}_\mu)}{\epsilon}, Y_N\right) \\
&\qquad \leq \omega_2(\delta; \epsilon, \beta, n, N),
\end{aligned}
$$

where $\limsup_{\epsilon\to 0} \limsup_{\beta\to 0} \limsup_{n\to+\infty} \limsup_{N\to+\infty} \omega_2(\delta; \epsilon, \beta, n, N) = 0$ for sufficiently small δ. We then use (3.2), Hypothesis 3.52, (3.84) and (3.89) to get

$$
\begin{aligned}
C_\tau \delta \leq c_0 \frac{|\bar{x}_\mu - \bar{y}_\mu|_{-1}^2}{\epsilon} - \frac{|\bar{x}_\mu - \bar{y}_\mu|^2}{\epsilon} + \omega_{R_\delta}\left(|\bar{x}_\mu - \bar{y}_\mu|\left(1 + \frac{|\bar{x}_\mu - \bar{y}_\mu|_{-1}}{\epsilon}\right)\right) \\
+ \omega_2(\delta; \epsilon, \beta, n, N).
\end{aligned}
$$

Let K_δ be a constant such that $\omega_{R_\delta}(r) \leq C_\tau\delta/4 + K_\delta r$. Then

$$
\begin{aligned}
\omega_{R_\delta}\left(|\bar{x}_\mu - \bar{y}_\mu|\left(1 + \frac{|\bar{x}_\mu - \bar{y}_\mu|_{-1}}{\epsilon}\right)\right) &\leq C_\tau\delta/4 + K_\delta |\bar{x}_\mu - \bar{y}_\mu|\left(1 + \frac{|\bar{x}_\mu - \bar{y}_\mu|_{-1}}{\epsilon}\right) \\
&\leq C_\tau\delta/2 + \frac{|\bar{x}_\mu - \bar{y}_\mu|^2}{\epsilon} + \tilde{K}_\delta \frac{|\bar{x}_\mu - \bar{y}_\mu|_{-1}^2}{\epsilon}
\end{aligned}
$$

for some $\tilde{K}_\delta > 0$ and small enough ϵ. Therefore we obtain that

$$
\frac{C_\tau \delta}{2} \leq (c_0 + \tilde{K}_\delta) \frac{|\bar{x}_\mu - \bar{y}_\mu|_{-1}^2}{\epsilon} + \omega_2(\delta; \epsilon, \beta, n, N)
$$

and this yields a contradiction in light of (3.89).

Therefore for small δ, ϵ, β and large n we must have $\bar{t} = T - \mu$ or $\bar{s} = T - \mu$. Without loss of generality suppose that $\bar{s}_\mu = T - \mu$. Recalling that $|\bar{x}_\mu|, |\bar{y}_\mu| \leq R_\delta$ for some $R_\delta > 0$ and using (3.99) we have

$$
\begin{aligned}
&u(\bar{t}_\mu, \bar{x}_\mu) - v(\bar{s}_\mu, \bar{y}_\mu) = (u(\bar{t}_\mu, \bar{x}_\mu) - g(e^{(T-\bar{t}_\mu)A}\bar{x}_\mu))_+ \\
&+ (g(e^{(T-\bar{t}_\mu)A}\bar{x}_\mu) - g(e^{(T-\bar{s}_\mu)A}\bar{y}_\mu)) + (g(e^{(T-\bar{s}_\mu)A}\bar{y}_\mu) - v(\bar{s}_\mu, \bar{y}_\mu))_+ \\
&\leq \tilde{\omega}_{R_\delta}(\mu + |\bar{t}_\mu - \bar{s}_\mu|) + |g(e^{(T-\bar{t}_\mu)A}\bar{x}_\mu) - g(e^{(T-\bar{s}_\mu)A}\bar{y}_\mu)|,
\end{aligned}
$$

where $\tilde{\omega}_{R_\delta}$ is a modulus for every δ, depending on (3.99). Then by (3.19), (3.20) and Hypothesis 3.53

$$
\begin{aligned}
&u(\bar{t}_\mu, \bar{x}_\mu) - v(\bar{s}_\mu, \bar{y}_\mu) \leq \tilde{\omega}_{R_\delta}(\mu + |\bar{t}_\mu - \bar{s}_\mu|) + \omega_{R_\delta}(|e^{(T-\bar{t}_\mu)A}\bar{x}_\mu - e^{(T-\bar{s}_\mu)A}\bar{y}_\mu|) \\
&\leq \tilde{\omega}_{R_\delta}(\mu + |\bar{t}_\mu - \bar{s}_\mu|) + \omega_{R_\delta}\left(\frac{e^{c_0\mu}}{(2\mu)^{\frac{1}{2}}} |e^{(\bar{s}_\mu - \bar{t}_\mu)A}\bar{x}_\mu - \bar{y}_\mu|_{-1}\right) \\
&\leq \tilde{\omega}_{R_\delta}(\mu + |\bar{t}_\mu - \bar{s}_\mu|) + \omega_{R_\delta}\left(\frac{e^{c_0\mu}}{(2\mu)^{\frac{1}{2}}} (C(R_\delta)|\bar{s}_\mu - \bar{t}_\mu|^{\frac{1}{2}} + |\bar{x}_\mu - \bar{y}_\mu|_{-1})\right).
\end{aligned}
$$

Therefore it follows from this, (3.88) and (3.89) that for $\delta < \delta_0$, and $\mu < \mu_0$ such that $\tilde{\omega}_{R_\delta}(\mu) \leq \tilde{m}/2$

$$
\limsup_{\epsilon\to 0} \limsup_{\beta\to 0} \limsup_{n\to+\infty} (u(\bar{t}, \bar{x}) - v(\bar{s}, \bar{y})) \leq \frac{\tilde{m}}{2}.
$$

This is impossible in light of (3.101).

Thus we obtain that $m_\mu \leq 0$ for all $\mu < \tau$ and this allows us to reapply the procedure to intervals $[T-3\tau/2, T-\tau/2], [T-4\tau/2, T-2\tau/2]$,... directly as in the proof of Theorem 3.50 since we now have

$$\lim_{R\to+\infty} \lim_{r\to 0} \limsup_{\eta\to 0} \Big\{ u(t,x) - v(s,y) : \ |x-y|_{-1} < r \\ x, y \in B_R, \ T-\tau/2-\eta \leq t, s \leq T-\tau/2 \Big\} \leq 0. \quad (3.102)$$

For $\gamma \in [0,1)$ the proof again uses the same modifications as indicated in the proof of Theorem 3.50. □

3.5.2. Degenerate elliptic equations. In this subsection we consider the degenerate elliptic case. We first introduce a slightly different version of Hypothesis 3.49.

HYPOTHESIS 3.55 *For $\gamma \in [0,1]$ there exist $M_F, N_F \geq 0$ such that*

$$|F(x,r,p+q,X+Y) - F(x,r,p,X)| \\ \leq M_F(1+|x|^\gamma)|q| + N_F(1+|x|^\gamma)^2\|Y\|$$

for all $(x,r) \in U\times\mathbb{R}$, $p,q \in H$, $X,Y \in S(H)$.

THEOREM 3.56 (Comparison under B-weak condition) *Let $U \subset H$ be open and $\overline{U}$ be B-closed. Let (3.2) hold and let F satisfy Hypotheses 3.44, 3.46, 3.47, 3.48, 3.55 and 3.45 with $\nu > 0$. Let u and v be, respectively, a viscosity subsolution and a viscosity supersolution of (3.58). Suppose that for every $R > 0$ there exists a modulus $\tilde{\omega}_R$ such that*

$$(u(x)-v(y))_+ + (u(y)-v(x))_+ \leq \tilde{\omega}_R(|x-y|_{-1}) \quad (3.103)$$

for $x \in \partial U$, $y \in \overline{U}$, $|x|,|y| \leq R$. Moreover suppose that there exist constants $C, a > 0$ such that one of the following conditions is satisfied

1. $\gamma \in (0,1)$, $\exists\ \bar{k} \geq 0$ *s.t.* $u, -v \leq C(1+|x|^{\bar{k}})\ \forall x \in H$, (3.104)
2. $\gamma = 0$, $2M_F a + 4N_F(a+a^2) < \nu$, *and* $u, -v \leq Ce^{a|x|}\ \forall x \in H$, (3.105)
3. $\gamma = 1$, $\exists \bar{k} \geq 0$ *s.t.* $M_F\bar{k} + N_F\bar{k}(\bar{k}-1) < \nu$ *if* $\bar{k} \geq 2$,
   $\bar{k}(M_F+N_F) < \nu$ *if* $\bar{k} < 2$, *and* $u, -v \leq C(1+|x|^{\bar{k}})\ \forall x \in H$. (3.106)

Then

$$m = \lim_{R\to+\infty} \lim_{r\to 0} \sup \Big\{ u(x) - v(y) : \ |x-y|_{-1} < r, \quad x, y \in \overline{U}\cap B_R \Big\} \leq 0. \quad (3.107)$$

In particular $u \leq v$.

PROOF. We will first do the proof in the case $\gamma \in (0,1)$. We argue by contradiction and assume that $m > 0$. Let $k > \bar{k}, k \geq 2$. Denote

$$m_\delta := \lim_{r\to 0} \sup \Big\{ u(x) - v(y) - \delta|x|^k - \delta|y|^k \ : \ |x-y|_{-1} < r, x, y \in \overline{U} \Big\},$$

$$m_{\delta,\epsilon} := \sup \left\{ u(x) - v(y) - \delta|x|^k - \delta|y|^k - \frac{|x-y|_{-1}^2}{2\epsilon} \ : \ x, y \in \overline{U} \right\}.$$

As in the proof of Theorem 3.50 it is easy to see that

$$m = \lim_{\delta\to 0} m_\delta, \quad (3.108)$$

$$m_\delta = \lim_{\epsilon\to 0} m_{\delta,\epsilon}. \quad (3.109)$$

Again, setting $u(x) = -\infty$ if $x \notin \overline{U}$ and $v(x) = +\infty$ if $x \notin \overline{U}$ we can consider u and v to be defined on H. Since $\overline{U}$ is B-closed such extended u is B-upper semi-continuous on $(0,T]\times H$ and v is B-lower semi-continuous on $(0,T]\times H$.

Define

$$\Psi(x,y) = u(x) - v(y) - \delta|x|^k - \delta|y|^k - \frac{|x-y|_{-1}^2}{2\epsilon}.$$

By (3.104) we can apply Corollary 3.26 to produce for every $n \geq 1$ elements $p_n, q_n \in H$ such that $|p_n| + |q_n| \leq \frac{1}{n}$ and such that

$$\Psi(x,y) + \langle Bp_n, x\rangle + \langle Bq_n, y\rangle$$

achieves a strict global maximum over $H \times H$ at some point $(\bar{x}, \bar{y}) \in \overline{U} \times \overline{U}$. Moreover we have

$$m_{\delta,\epsilon} \leq \Psi(\bar{x},\bar{y}) + \frac{C_\delta}{n}$$

for some constant $C_\delta > 0$. Therefore it follows that

$$m_{\delta,\epsilon} + \frac{|\bar{x}-\bar{y}|^2}{4\epsilon} \leq m_{\delta,2\epsilon} + \frac{C_\delta}{n}. \tag{3.110}$$

Inequalities (3.110) and (3.109) imply

$$\lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{|\bar{x}-\bar{y}|_{-1}^2}{\epsilon} = 0 \quad \text{for every } \delta > 0. \tag{3.111}$$

Condition (3.103), together with (3.111), now imply that if δ, ϵ are small enough and n is sufficiently big we must have $(\bar{x},\bar{y}) \in U \times U$.

Similarly to the proof of Theorem 3.56 we now have for $N > 2$ that if we define

$$u_1(x) = u(x) - \delta|x|^k - \frac{\langle BQ_N(\bar{x}-\bar{y}), x\rangle}{\epsilon} - \frac{|Q_N(x-\bar{x})|_{-1}^2}{\epsilon} + \frac{|Q_N(\bar{x}-\bar{y})|_{-1}^2}{2\epsilon} + \langle Bp_n, x\rangle$$

and

$$v_1(y) = v(y) + \delta|y|^k - \frac{\langle BQ_N(\bar{x}-\bar{y}), y\rangle}{\epsilon} + \frac{|Q_N(y-\bar{y})|_{-1}^2}{\epsilon} - \langle Bq_n, y\rangle.$$

then

$$u_1(x) - v_1(y) - \frac{1}{2\epsilon}|P_N(x-y)|_{-1}^2$$

has a strict global maximum at $(\bar{x},\bar{y})$ over $H \times H$. We can therefore apply Corollary 3.28 to obtain test functions φ_k, ψ_k and points x_k, y_k such that $u_1(x) - \varphi_k(x)$ has a maximum at x_k, $v_1(y) - \psi_k(y)$ has a minimum at y_k, and such that (3.52), (3.53) are satisfied for u_1, v_1 respectively. In particular $x_k, y_k \in U$ for big k.

Therefore, since u is a viscosity subsolution of (3.58) in U, using the definition of viscosity subsolution, letting $k \to +\infty$ and using (3.52) we obtain

$$-\left\langle \bar{x}, A^*\left(\frac{B(\bar{x}-\bar{y})}{\epsilon} - Bp_n\right)\right\rangle + F(\bar{x}, u(\bar{x}), p_{n,\delta,\epsilon}, X_{n,\delta,\epsilon}) \leq 0 \tag{3.112}$$

where we denoted

$$p_{n,\delta,\epsilon} = \frac{B(\bar{x}-\bar{y})}{\epsilon} - Bp_n + \delta k|\bar{x}|^{k-2}x,$$

and

$$X_{n,\delta,\epsilon} = X_N + \frac{2BQ_N}{\epsilon} + \delta k|\bar{x}|^{k-2}((k-2)\frac{\bar{x}\otimes\bar{x}}{|\bar{x}|^2} + I)).$$

Hence we obtain from Hypothesis 3.55 that

$$-\left\langle \bar{x}, A^*\left(\frac{B(\bar{x}-\bar{y})}{\epsilon} - Bp_n\right)\right\rangle + F\left(\bar{x}, u(\bar{x}), \frac{B(\bar{x}-\bar{y})}{\epsilon} - Bp_n, X_N + \frac{2BQ_N}{\epsilon}\right) - \delta M_F\left((1+|\bar{x}|^\gamma)k|\bar{x}|^{k-1} + (1+|\bar{x}|^\gamma)^2 k(k-1)|\bar{x}|^{k-2}\right) \leq 0. \tag{3.113}$$

Arguing similarly we obtain that

$$-\left\langle \bar{y}, A^*\left(\frac{B(\bar{x}-\bar{y})}{\epsilon}+Bq_n\right)\right\rangle + F\left(\bar{y}, v(\bar{y}), \frac{B(\bar{x}-\bar{y})}{\epsilon}+Bq_n, X_N - \frac{2BQ_N}{\epsilon}\right)$$
$$+\delta M_F((1+|\bar{y}|^\gamma)k|\bar{y}|^{k-1}+(1+|\bar{y}|^\gamma)^2k(k-1)|\bar{y}|^{k-2}) \geq 0. \quad (3.114)$$

Therefore subtracting (3.114) from (3.113) and using Hypothesis 3.47 yield

$$-\left\langle \bar{x}-\bar{y}, \frac{A^*B(\bar{x}-\bar{y})}{\epsilon}\right\rangle$$
$$+F\left(\bar{x}, u(\bar{x}), \frac{B(\bar{x}-\bar{y})}{\epsilon}, X_N\right) - F\left(\bar{y}, v(\bar{y}), \frac{B(\bar{x}-\bar{y})}{\epsilon}, Y_N\right)$$
$$-C\delta\left(1+|\bar{x}|^{k-1+\gamma}+|\bar{y}|^{k-1+\gamma}\right) \leq \omega_1(\delta,\epsilon;n,N), \quad (3.115)$$

for some $C = C(M_F, k, \gamma)$, where $\lim_{n\to+\infty}\lim_{N\to+\infty}\omega_1(\delta,\epsilon;n,N) = 0$ for fixed δ, ϵ.

It now follows from (3.108), (3.109), (3.110) that

$$\liminf_{\delta\to 0}\liminf_{\epsilon\to 0}\liminf_{n\to+\infty}(u(\bar{x})-v(\bar{y})-\delta|\bar{x}|^k-\delta|\bar{y}|^k) > \bar{m} = \min\left(\frac{m}{2},1\right) > 0. \quad (3.116)$$

This, Hypothesis 3.45 and (3.115) imply

$$-\left\langle \bar{x}-\bar{y}, \frac{A^*B(\bar{x}-\bar{y})}{\epsilon}\right\rangle + \nu(u(\bar{x})-v(\bar{y}))$$
$$+F\left(\bar{x}, v(\bar{y}), \frac{B(\bar{x}-\bar{y})}{\epsilon}, X_N\right) - F\left(\bar{y}, v(\bar{y}), \frac{B(\bar{x}-\bar{y})}{\epsilon}, Y_N\right)$$
$$-C\delta\left(1+|\bar{x}|^{k-1+\gamma}+|\bar{y}|^{k-1+\gamma}\right) \leq \omega_1(\delta,\epsilon;n,N), \quad (3.117)$$

We recall that X_N, Y_N satisfy (3.54). We can now use (3.2), Hypothesis 3.48, and the fact that $|\bar{x}|, |\bar{y}|, |u(\bar{x})|, |v(\bar{y})| \leq R_\delta$ for some R_δ independent of ϵ, n to get

$$\begin{aligned}
&\nu(u(\bar{x})-v(\bar{y})) - C\delta(1+|\bar{x}|^{k-1+\gamma}+|\bar{y}|^{k-1+\gamma})\\
&\leq c_0\frac{|\bar{x}-\bar{y}|_{-1}^2}{\epsilon} + \omega_{R_\delta}\left(|\bar{x}-\bar{y}|_{-1}\left(1+\frac{|\bar{x}-\bar{y}|_{-1}}{\epsilon}\right)\right) + \omega_1(\delta,\epsilon;n,N) \qquad (3.118)\\
&\leq \omega_2(\delta;\epsilon,n,N),
\end{aligned}$$

where $\limsup_{\epsilon\to 0}\limsup_{n\to+\infty}\limsup_{N\to+\infty}\omega_2(\delta;\epsilon,n,N) = 0$ for sufficiently small δ. Therefore we have from (3.116) and (3.118) that

$$\tilde{m} \leq -\nu\delta(|\bar{x}|^k+|\bar{y}|^k) + C\delta(1+|\bar{x}|^{k-1+\gamma}+|\bar{y}|^{k-1+\gamma}) + \omega_3(\delta,\epsilon,n,N), \quad (3.119)$$

where $\limsup_{\delta\to 0}\limsup_{\epsilon\to 0}\limsup_{n\to+\infty}\limsup_{N\to+\infty}\omega_3(\delta,\epsilon,n,N) = 0$. Since

$$\max_{r\geq 0}\left(-\nu\delta r^k + C\delta r^{k-1+\gamma}\right) \leq C_1\delta,$$

taking $\limsup_{\delta\to 0}\limsup_{\epsilon\to 0}\limsup_{n\to+\infty}\limsup_{N\to+\infty}$ in (3.119), we conclude

$$\nu\bar{m} \leq 0,$$

which is a contradiction unless $m \leq 0$.

For $\gamma = 0$ the proof is almost the same. We replace the functions

$$\delta|x|^k \quad \text{and} \quad \delta|x|^k$$

by

$$\delta e^{b\sqrt{1+|x|^2}} \quad \text{and} \quad \delta e^{b\sqrt{1+|y|^2}},$$

respectively, where $b > a$ is such that $\nu > 2M_F b + 4N_F(b + b^2)$. We then obtain, in place of (3.118),

$$\nu(u(\bar{x}) - v(\bar{y})) - \delta(2M_F b + 4N_F(b + b^2))\left(e^{b\sqrt{1+|\bar{x}|^2}} + e^{b\sqrt{1+|\bar{y}|^2}}\right) \le \omega_2(\delta; \epsilon, n, N)$$

which, using $\nu > 2M_F b + 4N_F(b + b^2)$ and the fact that now

$$\liminf_{\delta\to 0}\ \liminf_{\epsilon\to 0}\ \liminf_{n\to+\infty}\left(u(\bar{x}) - v(\bar{y}) - \delta e^{b\sqrt{1+|\bar{x}|^2}} - \delta e^{b\sqrt{1+|\bar{y}|^2}}\right) > \bar{m} > 0,$$

produces again

$$\nu\bar{m} \le \omega_4(\delta, \epsilon, n, N), \tag{3.120}$$

where $\limsup_{\delta\to 0}\limsup_{\epsilon\to 0}\limsup_{n\to+\infty}\limsup_{N\to+\infty}\omega_4(\delta, \epsilon, n, N) = 0$.

For $\gamma = 1$ the proof is also very similar. Let $\bar{k} \ge 2$. We take $k_1 > k > \bar{k}$ such that $\nu > M_F k_1 + N_F k_1(k_1 - 1)$ and replace the functions $\delta|x|^k$ and $\delta|x|^k$ in the definition of m_δ respectively by $h(x) = \delta(1 + |x|^2)^{\frac{k}{2}}$ and $h(y) = \delta(1 + |y|^2)^{\frac{k}{2}}$. It is easy to check that

$$|Dh(x)| \le \delta k(1 + |x|^2)^{\frac{k}{2}-1}|x|, \quad |D^2h(x)| \le \delta k(k-1)(1 + |x|^2)^{\frac{k}{2}-1}$$

and so there exists $r > 0$ such that

$$(1 + |x|)|Dh(x)| \le \delta k_1(1 + |x|^2)^{\frac{k}{2}} \quad \text{if } |x| \ge r,$$
$$(1 + |x|)^2\|D^2h(x)\| \le \delta k_1(k_1 - 1)(1 + |x|^2)^{\frac{k}{2}} \quad \text{if } |x| \ge r,$$

If we now repeat the arguments of the proof and use the above estimates we obtain, in place of (3.118),

$$\nu(u(\bar{x}) - v(\bar{y})) - \delta(M_F k_1 + N_F k_1(k_1 - 1))\left((1 + |\bar{x}|^2)^{\frac{k}{2}} + (1 + |\bar{y}|^2)^{\frac{k}{2}}\right)$$
$$\le \omega_2(\delta; \epsilon, n, N) + \omega_3(\delta) \tag{3.121}$$

for some modulus ω_3 which depends on r. The result now follows upon noticing that

$$\liminf_{\delta\to 0}\ \liminf_{\epsilon\to 0}\ \liminf_{n\to+\infty}(u(\bar{x}) - v(\bar{y}) - \delta(1 + |\bar{x}|^2)^{\frac{k}{2}} + \delta(1 + |\bar{y}|^2)^{\frac{k}{2}}) > \bar{m} > 0$$

and using $\nu > M_F k_1 + N_F k_1(k_1 - 1)$.

If $\bar{k} < 2$ we proceed the same as for $\bar{k} \ge 2$. We take $k_1 > k > \bar{k}$ such that $\nu > M_F k_1 + N_F k_1$ and as before take $h(x) = \delta(1 + |x|^2)^{\frac{k}{2}}$ and $h(y) = \delta(1 + |y|^2)^{\frac{k}{2}}$. However now

$$D^2h(x) = \delta k(k-2)(1 + |x|^2)^{\frac{k}{2}-2}x \otimes x + \delta k(1 + |x|^2)^{\frac{k}{2}-1}I \le \delta k(1 + |x|^2)^{\frac{k}{2}-1}I$$

Thus when we plug the derivatives of h into the equation in the proof of comparison, using Hypothesis 3.46 we can replace $D^2h(\bar{x})$ by $\delta k(1 + |\bar{x}|^2)^{\frac{k}{2}-1}I$ and also do similarly for $D^2h(\bar{y})$. The rest of the arguments are the same. □

Remark 3.57 The conditions in (3.104)-(3.106) may not be optimal for some equations due to a rather general assumption Hypothesis 3.55 and the way it is written. However they are optimal in some cases. Consider a simple first order equation $u - xu' = 0$ in $\mathbb{R}$ which has two obvious classical solutions $u_1 \equiv 0$ and $u_2(x) = x$, and the second order equation $2u - x^2u'' = 0$ which has solutions $u_1 \equiv 0$ and $u_2(x) = x^2$. For $u - xu' = 0$, (3.106) produces $\bar{k} < 1$, and for $2u - x^2u'' = 0$ we obtain $\bar{k} < 2$. Equation $u - \mu u' = 0, \mu > 0$, has two classical solutions $u_1 \equiv 0$ and $u_2(x) = e^{x/\mu}$. Notice that here $M_F = \mu/2$ and (3.105) gives $a < 1/\mu$. ■

Theorem 3.58 (Comparison under B-strong condition) *The conclusions of Theorem 3.56 hold if (3.2) is replaced by (3.3) and Hypothesis 3.48 is replaced by Hypothesis 3.52.*

PROOF. The proof is exactly the same as the proof of Theorem 3.56 with one modification. Using the notation of this proof, instead of (3.118) (for $\gamma \in (0,1)$), by (3.3) and Hypothesis 3.52 we now have

$$\nu(u(\bar{x}) - v(\bar{y})) - C\delta(1 + |\bar{x}|^{k-1+\gamma} + |\bar{y}|^{k-1+\gamma}) \le c_0 \frac{|\bar{x}-\bar{y}|_{-1}^2}{\epsilon} - \frac{|\bar{x}-\bar{y}|^2}{\epsilon} + \omega_{R_\delta}\left(|\bar{x}-\bar{y}|\left(1 + \frac{|\bar{x}-\bar{y}|_{-1}}{\epsilon}\right)\right) + \omega_1(\delta,\epsilon;n,N). \tag{3.122}$$

If $\omega_{R_\delta}(r) \le \nu\bar{m}/4 + K_\delta s$ for some $K_\delta > 0$ we obtain

$$\omega_{R_\delta}\left(|\bar{x}-\bar{y}|\left(1 + \frac{|\bar{x}-\bar{y}|_{-1}}{\epsilon}\right)\right) \le \frac{\nu\bar{m}}{4} + K_\delta|\bar{x}-\bar{y}|\left(1 + \frac{|\bar{x}-\bar{y}|_{-1}}{\epsilon}\right) \le \frac{\nu\bar{m}}{2} + \frac{|\bar{x}-\bar{y}|^2}{\epsilon} + \tilde{K}_\delta \frac{|\bar{x}-\bar{y}|_{-1}^2}{\epsilon} \tag{3.123}$$

for some $\tilde{K}_\delta > 0$ and small enough ϵ. Therefore putting this in (3.122) and applying (3.111) yields

$$\nu(u(\bar{x}) - v(\bar{y})) - C\delta(1 + |\bar{x}|^{k-1+\gamma} + |\bar{y}|^{k-1+\gamma}) \le \frac{\nu\bar{m}}{2} + \omega_2(\delta;\epsilon,n,N), \tag{3.124}$$

where $\limsup_{\epsilon\to 0} \limsup_{n\to+\infty} \limsup_{N\to+\infty} \omega_2(\delta;\epsilon,n,N) = 0$ for sufficiently small δ. This allows us to continue and conclude the proof using the same arguments as those used in the proof of Theorem 3.56. Other cases are done similarly. □

REMARK 3.59 We remark that all results of this section extend to equations of the form

$$u_t + \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \left\{ \langle A_{\alpha,\beta}x, Du\rangle - F_{\alpha,\beta}(t,x,u,Du,D^2u) \right\} = 0$$

and

$$\inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \left\{ -\langle A_{\alpha,\beta}x, Du\rangle + F_{\alpha,\beta}(t,x,u,Du,D^2u) \right\} = 0,$$

where $\mathcal{A}, \mathcal{B}$ are arbitrary sets, provided that all assumptions are satisfied by $A_{\alpha,\beta}, F_{\alpha,\beta}$, uniformly in α and β. Since the definition of viscosity solution depended on the operators A and B, we have to assume that there exist a linear, densely, defined maximal dissipative operator A in H such that $D(A^*) \subset D(A^*_{\alpha,\beta})$ for all α, β, and a bounded, strictly positive, self-adjoint operator B on H such that A^*B is bounded. To ensure the uniformity of test functions, they are now defined by Definition 3.32 for A, and the notion of B-continuity is defined using our fixed B which works for all $A_{\alpha,\beta}$. All operators $A_{\alpha,\beta}$ must then satisfy either the weak or strong B-condition with this B and a constant c_0 independent of α and β. ■

3.6. Existence of solutions: Value function

In this section we investigate existence of viscosity solutions for Hamilton-Jacobi-Bellman equations associated with stochastic optimal control problems. In such cases the Hamiltonians F in equations (3.58) and (3.64) are convex/concave in u, Du, D^2u. We show that, under suitable hypotheses, the unique viscosity solution of (3.64) (respectively, (3.58)) is the value function of the associated finite horizon (respectively, infinite horizon) optimal control problem. A key ingredient in the proof will be the use of the Dynamic Programming Principle (Theorem 2.24). We recall briefly the weak formulation of a stochastic optimal control problem that has been introduced in Chapter 2.

We fix a final time $0 < T \le +\infty$, a Polish space Λ (the control space), a real, separable Hilbert space Ξ (the space of the noise) and $Q \in \mathcal{L}_1^+(\Xi)$.

Following Definition 2.7, for $t \in [0,T)$, we say that the 5-tuple $\nu := \left(\Omega^\nu, \mathscr{F}^\nu, \mathscr{F}_s^{\nu,t}, \mathbb{P}^\nu, W_Q^\nu\right)$ is a *reference probability space* if:

(i) $(\Omega^\nu, \mathscr{F}^\nu, \mathbb{P}^\nu)$ is a complete probability space.

(ii) $W_Q^\nu = \{W_Q^\nu(s) : t \le s \le T\}$ is a Ξ-valued Q-Wiener process on $(\Omega^\nu, \mathscr{F}^\nu, \mathbb{P}^\nu)$ (with $W_Q^\nu(t) = 0$, $\mathbb{P}^\nu$ a.s.).

(iii) The filtration $\mathscr{F}_s^{\nu,t} = \sigma\{\mathscr{F}_s^{\nu,t,0}, \mathcal{N}\}$, where $\mathscr{F}_s^{\nu,t,0} = \sigma\{W_Q^\nu(\tau) : t \le \tau \le s\}$ and $\mathcal{N}$ are the $\mathbb{P}^\nu$-null sets in $\mathscr{F}^\nu$.

We say that a process $a(\cdot)$ is an admissible control on $[t,T]$ (respectively on $[t,+\infty)$ if $T = +\infty$) if there exists a reference probability space $\nu = \left(\Omega^\nu, \mathscr{F}^\nu, \mathscr{F}_s^{\nu,t}, \mathbb{P}^\nu, W_Q^\nu\right)$ such that $a(\cdot) \colon [t,T]\times\Omega^\nu \to \Lambda$ (respectively $a(\cdot) \colon [t,+\infty)\times\Omega^\nu \to \Lambda$) is $\mathscr{F}_s^{\nu,t}$-progressively measurable. To indicate the dependence of $a(\cdot)$ on the reference probability space we will write $a^\nu(\cdot)$ and, with a slight abuse the notation, we will often write $a^\nu(\cdot)$ to denote the whole 6-tuple $\left(\Omega^\nu, \mathscr{F}^\nu, \mathscr{F}_s^{\nu,t}, \mathbb{P}^\nu, W_Q^\nu, a^\nu(\cdot)\right)$. We denote the set of all admissible controls $a^\nu(\cdot)$ by $\mathcal{U}_t$.

The finite horizon problem: Let $T < +\infty$. For any $a^\nu(\cdot) \in \mathcal{U}_t$ we consider the system evolving according to the following state equation

$$\begin{cases} dX(s) = (AX(s) + b(s, X(s), a^\nu(s)))\, ds + \sigma(s, X(s), a^\nu(s)) dW_Q^\nu(s) \\ X(t) = x, \end{cases} \tag{3.125}$$

where A is a linear, densely defined, maximal dissipative operator in H generating a C_0-semigroup of contractions e^{tA}. The functions b and σ satisfy conditions that will be specified below. Our hypotheses will guarantee that (3.125) admits, for any $a^\nu(\cdot) \in \mathcal{U}_t$, a unique mild solution (see Definition 1.113) denoted by $X(s;t,x;a^\nu(\cdot))$. We consider the problem of minimizing a cost functional

$$J(t,x;a^\nu(\cdot)) = \mathbb{E}^\nu \Bigg[\int_t^T e^{-\int_t^s c(X(\tau;t,x,a^\nu(\cdot)))d\tau} l(s, X(s;t,x;a^\nu(\cdot)), a^\nu(s)) ds + e^{-\int_t^T c(X(\tau;t,x,a^\nu(\cdot)))d\tau} g(X(T;t,x,a^\nu)) \Bigg] \tag{3.126}$$

over all $a^\nu(\cdot) \in \mathcal{U}_t$. The value function of this minimization problem is defined as follows:

$$V(t,x) := \inf_{a^\nu(\cdot)\in\mathcal{U}_t} J(t,x;a(\cdot)), \tag{3.127}$$

while the associated Hamilton-Jacobi-Bellman equation is given by

$$\begin{cases} v_t + \langle Ax, Dv\rangle - F(t,x,v,Dv,D^2v) = 0 \\ v(T,x) = g(x), \end{cases} \tag{3.128}$$

where

$$F(t,x,r,p,X) := \sup_{a\in\Lambda} \Bigg\{ -\frac{1}{2}\mathrm{Tr}\left(\left(\sigma(t,x,a)Q^{\frac12}\right) \left(\sigma(t,x,a)Q^{\frac12}\right)^* X \right) - \langle p, b(t,x,a)\rangle + c(x)r - l(t,x,a) \Bigg\}. \tag{3.129}$$

Remark 3.60 We point out that if $\sigma(t,x,a) \in \mathcal{L}(\Xi, H)$, then the term

$$\left(\sigma(t,x,a)Q^{\frac12}\right) \left(\sigma(t,x,a)Q^{\frac12}\right)^*$$

can be written in a more common and convenient form

$$\sigma(t,x,a)Q\sigma(t,x,a)^*.$$

■

The infinite horizon problem: We only study the case of constant discounting, i.e. we assume that $c = \lambda > 0$. However, under suitable assumptions, the results we prove could be adapted to a more general case of non-constant c. For any $a^\nu(\cdot) \in \mathcal{U}_0$ we consider a system described by a stochastic differential equation

$$\begin{cases} dX(s) = (AX(s) + b(X(s), a^\nu(s)))\, ds + \sigma(X(s), a^\nu(s)) dW_Q^\nu(s) \\ X(0) = x. \end{cases} \tag{3.130}$$

The mild solution of (3.130) will be denoted by $X(s; 0, x; a^\nu(\cdot))$. The infinite horizon problem consists in minimizing a cost functional

$$J(x; a^\nu(\cdot)) = \mathbb{E}^\nu \left[\int_0^{+\infty} e^{-\lambda t} l(X(s; 0, x; a^\nu), a^\nu(s)) ds \right]. \tag{3.131}$$

over all controls $a^\nu(\cdot) \in \mathcal{U}_0$. The value function is given by

$$V(x) := \inf_{a^\nu(\cdot) \in \mathcal{U}_0} J(x; a(\cdot)), \tag{3.132}$$

and the corresponding Hamilton-Jacobi-Bellman equation is

$$\lambda v(x) - \langle Ax, Dv \rangle + F(x, v, Dv, D^2 v) = 0, \tag{3.133}$$

where

$$F(x, r, p, X) := \sup_{a \in \Lambda} \left\{ -\frac{1}{2} \mathrm{Tr} \left(\left(\sigma(x, a) Q^{\frac{1}{2}} \right) \left(\sigma(x, a) Q^{\frac{1}{2}} \right)^* X \right) - \langle p, b(x, a) \rangle - l(x, a) \right\}. \tag{3.134}$$

3.6.1. Finite horizon problem. In this subsection we prove that, under suitable hypotheses, the value function (3.127) of the finite horizon problem is the unique viscosity solution of (3.128). We obtain the results under two sets of hypotheses: in the first the generator A satisfies the weak B-condition for some operator B (see Definition 3.9), in the second it satisfies the strong B-condition (Definition 3.10) which allows to put milder assumptions on the coefficients of the state equation and of the cost functional. Note that the words *strong* and *weak* used to describe the B-conditions have nothing to do with the strong and weak formulation of the optimal control problem (see Sections 2.1.2 and 2.1.2).

To avoid cumbersome notation we will drop the index "ν" whenever it does not cause and confusion.

PROPOSITION 3.61 (Regularity of V under weak B-condition) *Let B satisfy the weak B-condition for A (Definition 3.9) and $b\colon [0,T] \times H \times \Lambda \to H$, $\sigma\colon [0,T] \times H \times \Lambda \to \mathcal{L}_2(\Xi_0, H)$, $l\colon [0,T] \times H \times \Lambda \to \mathbb{R}$ be continuous. Assume that b and σ satisfy (3.7), (3.9), (3.10) with $\gamma = 1$, (3.22) and (3.23), and let c be bounded from below. Suppose that there exist local moduli $\omega_l(\cdot, \cdot)$ and $\omega(\cdot, \cdot)$ such that*

$$|l(t, x, a) - l(s, y, a)| \le \omega_l(|x - y|_{-1} + |s - t|, R) \quad \text{for all } x, y \in B(0, R),\ a \in \Lambda,\ s, t \in [0, T] \tag{3.135}$$

and

$$|g(x) - g(y)|, |c(x) - c(y)| \le \omega(|x - y|_{-1}, R) \qquad \text{for all } x, y \in B(0, R). \tag{3.136}$$

Moreover assume that there exist two nonnegative constants C, m such that

$$|c(x)|, |g(x)|, |l(t, x, a)| \le C(1 + |x|^m) \tag{3.137}$$

for all $x \in H$, $a \in \Lambda$ and $t \in [0, T]$. Then:

(i) *There exists a local modulus $\sigma_1(\cdot, \cdot)$ such that*

$$|J(t, x; a(\cdot)) - J(t, y; a(\cdot))| \le \sigma_1(|x - y|_{-1}, R) \tag{3.138}$$

for all $x, y \in B(0, R)$, $t \in [0, T]$ and $a(\cdot) \in \mathcal{U}_t$.

(ii) *There exists a nonnegative constant $\tilde{C}$, and a local modulus $\sigma_2(\cdot,\cdot)$ such that*

$$|J(t,x;a(\cdot))|, |V(t,x)| \leq \tilde{C}(1+|x|^m), \tag{3.139}$$

for all $(t,x) \in [0,T] \times H$ and $a(\cdot) \in \mathcal{U}_t$. and

$$|V(t,x) - V(s,y)| \leq \sigma_2(|t-s| + |x-y|_{-1}, R) \tag{3.140}$$

for all $x, y \in B(0,R)$, $t, s \in [0,T]$.

PROOF. Part (i): Let L be a constant such that $c(x) \geq L$ for all $x \in H$. We will assume that $L < 0$. Choose $x, y \in B(0,R)$, $a(\cdot) \in \mathcal{U}_t$ and denote $X(\cdot;t,x;a(\cdot))$ and $X(\cdot;t,y;a(\cdot))$ respectively by $X(\cdot)$ and $Y(\cdot)$. We have

$$\begin{aligned}
&|J(t,y;a(\cdot)) - J(t,x;a(\cdot))| \leq I_1 + I_2 \\
&:= \left(\int_t^T \mathbb{E}\left| e^{-\int_t^r c(X(\tau))d\tau} l(r,X(r),a(r)) - e^{-\int_t^r c(Y(\tau))d\tau} l(r,Y(r),a(r)) \right| dr \right) \\
&\qquad + \left(\mathbb{E}|e^{-\int_t^T c(X(\tau))d\tau} g(X(T)) - e^{-\int_t^T c(Y(\tau))d\tau} g(Y(T))| \right).
\end{aligned} \tag{3.141}$$

We first consider I_1.

$$\begin{aligned}
I_1 &\leq I_{11} + I_{12} \\
&:= \int_t^T \mathbb{E}\left[e^{-\int_t^r c(X(\tau))d\tau} |l(r,X(r),a(r)) - l(r,Y(r),a(r))| \right] dr \\
&\quad + \int_t^T \mathbb{E}\left[|l(r,Y(r),a(r))| \left| e^{-\int_t^r c(X(\tau))d\tau} - e^{-\int_t^r c(Y(\tau))d\tau} \right| \right] dr.
\end{aligned} \tag{3.142}$$

In the following we will denote M any absolute constant independent of R and of the control. Given $\epsilon > 0$ we can find, thanks to (D.1), a positive constant K_ϵ (non-increasing in ϵ) such that, for any $s > 0$, $\omega_l(s, \frac{1}{\epsilon}) \leq \epsilon + K_\epsilon s$. Using (3.135) and (3.137), we obtain

$$\begin{aligned}
I_{11} &\leq e^{-TL} \int_t^T \mathbb{E}|l(r,X(r),a(r)) - l(r,Y(r),a(r))| dr \\
&\leq e^{-TL} \int_t^T \int_{\{|X(r)|<\frac{1}{\epsilon} \text{ and } |Y(r)|<\frac{1}{\epsilon}\}} \omega_l\left(|X(r)-Y(r)|_{-1}, \frac{1}{\epsilon} \right) d\mathbb{P} \\
&+ e^{-TL} \int_t^T \int_{\{|X(r)|\geq\frac{1}{\epsilon} \text{ or } |Y(r)|\geq\frac{1}{\epsilon}\}} M(2+|X(r)|^m + |Y(r)|^m) d\mathbb{P} dr. \\
&\leq M \int_t^T (\epsilon + K_\epsilon \mathbb{E}|X(r)-Y(r)|_{-1}) dr \\
&+ M \int_t^T \left(\mathbb{E}(1+|X(r)|^{2m} + |Y(r)|^{2m}) \right)^{\frac{1}{2}} \left(\mathbb{P}\left(|X(r)| \geq \frac{1}{\epsilon} \right) + \mathbb{P}\left(|Y(r)| \geq \frac{1}{\epsilon} \right) \right)^{\frac{1}{2}} dr.
\end{aligned}$$

It follows from (1.38) that we have

$$\mathbb{E}\big(\sup_{t\leq r\leq T} |X(r)|^{2m} + \sup_{t\leq r\leq T} |Y(r)|^{2m} \big) \leq C_R, \tag{3.143}$$

where C_R is a constant independent of the control but depending on R. In particular, this implies by Chebychev's inequality, that

$$\left(\mathbb{P}\left(\sup_{t\leq r\leq T} |X(r)| \geq \frac{1}{\epsilon} \right) + \mathbb{P}\left(\sup_{t\leq r\leq T} |Y(r)| \geq \frac{1}{\epsilon} \right) \right)^{\frac{1}{2}} \leq \gamma(\epsilon, R), \tag{3.144}$$

for some local modulus γ. Thus, using (3.143), (3.144) and (3.24), we obtain

$$I_{11} \leq M(\epsilon + K_\epsilon |x-y|_{-1}) + \gamma_1(\epsilon, R) \tag{3.145}$$

for some local modulus γ_1. Taking the infimum of the right hand side of (3.145) over $\epsilon > 0$ produces a local modulus $\varrho(\cdot, R)$ such that

$$I_{11} \le \varrho(|x-y|_{-1}, R). \tag{3.146}$$

To estimate I_{12} observe that

$$\begin{aligned}
\int_t^T \mathbb{E}\left[|l(r,Y(r),a(r))| \left|e^{-\int_t^s c(X(\tau))d\tau} - e^{-\int_t^s c(Y(\tau))d\tau}\right|\right] dr \\
\le \left(\int_t^T \mathbb{E}\left[|l(r,Y(r),a(r))|^2\right] dr\right)^{\frac{1}{2}} \\
\times \left(\int_t^T \mathbb{E}\left[\left|e^{-\int_t^s c(X(\tau))d\tau} - e^{-\int_t^s c(Y(\tau))d\tau}\right|^2\right] dr\right)^{\frac{1}{2}}.
\end{aligned} \tag{3.147}$$

We observe that for $a, b \in \mathbb{R}, a, b \ge L$ one has $\left|e^{-a} - e^{-b}\right| \le e^{-L}|a-b|$. Therefore, using (3.143), it follows similarly as before that, for some $C_R > 0$ depending on R but independent of the choice of the control,

$$\begin{aligned}
I_{12} \le C_R\Bigg[\int_t^T \mathbb{E}\left(\int_t^T \omega\left(|X(r)-Y(r)|_{-1}, \frac{1}{\epsilon}\right) dr\right)^2 \\
+ \mathbb{P}\left(\sup_{t\le r\le T} |X(r)| \ge \frac{1}{\epsilon}\right) + \mathbb{P}\left(\sup_{t\le r\le T} |Y(r)| \ge \frac{1}{\epsilon}\right) dr\Bigg]^{\frac{1}{2}}.
\end{aligned}$$

We now use again (3.144), (3.24), and argue as for I_{11} to find that there exists a local modulus $\varrho(\cdot, R)$ such that

$$I_{12} \le \varrho(|x-y|_{-1}, R). \tag{3.148}$$

The term I_2 in (3.141) can be estimated similarly. Thus we obtain claim (i).

Part (ii): Estimate (3.139) follows directly from (3.137) and (1.38). Moreover, by (3.138), we have

$$|V(t,x) - V(t,y)| \le \sigma_1(|x-y|_{-1}, R) \qquad \forall x, y \in B(0,R), t \in [0,T]. \tag{3.149}$$

It remains to prove that

$$|V(t,x) - V(s,x)| \le \tilde{\sigma}(|t-s|, R) \qquad \forall x \in B(0,R), t, s \in [0,T] \tag{3.150}$$

for some local modulus $\tilde{\sigma}$.

We notice that it follows from our assumptions, (3.138), (3.139), and Proposition 2.16, that the assumptions of Theorem 2.24 are satisfied and thus the dynamic programming principle (2.25) holds. Let now (3.150), let $0 \le t < s \le T$ and $x \in B(0,R)$. Let $X(\cdot)$ be the solution of (3.125).

Using (2.25) we have, for some constant C_R depending on R,

$$\begin{aligned}
|V(s,x) - V(t,x)| &\le \sup_{a(\cdot)\in\mathcal{U}_t} e^{-LT}\mathbb{E}\int_t^s |l(r,X(r),a(r))|dr \\
&\quad + \sup_{a(\cdot)\in\mathcal{U}_t} \mathbb{E}\left|V(s,x) - V(s,X(s))e^{-\int_t^s c(X(r))dr}\right| \\
&\le C_R|t-s| + e^{-LT}\sup_{a(\cdot)\in\mathcal{U}_t} \mathbb{E}\left|V(s,x) - V(s,X(s))\right| \\
&\quad + \sup_{a(\cdot)\in\mathcal{U}_t} \mathbb{E}\left|V(s,x)\left(e^{-\int_t^s c(X(r))dr} - 1\right)\right| =: C_R|t-s| + D_1 + D_2,
\end{aligned} \tag{3.151}$$

where we have used (1.38), (3.137).

Since V satisfies (3.139) and (3.149), arguing as in the estimates for I_{11} and using (3.27) and (3.144), we obtain for every $\epsilon > 0$, $a(\cdot) \in \mathcal{U}_t$

$$\begin{aligned}\mathbb{E}\,|V(s,x) - V(s,X(s))| &\le \mathbb{E}\sigma_1(|x - X(s)|_{-1}, \frac{1}{\epsilon}) + C_R\left(\mathbb{P}(\sup_{t\le r\le T}|X(r)| \ge \frac{1}{\epsilon})\right)^{\frac{1}{2}} \\ &\le \epsilon + C_{R,\epsilon}|t-s|^{\frac{1}{2}} + \gamma(\epsilon, R), \quad (3.152)\end{aligned}$$

thus the same estimate holds for D_1.

To estimate D_2, let $C_\epsilon \ge 0$ be a constant such that $c(y) \le C_\epsilon$ when $|y| \le \frac{1}{\epsilon}$. Then, for every $\epsilon > 0$, $a(\cdot) \in \mathcal{U}_t$,

$$\begin{aligned}&\mathbb{E}\Big(|V(s,x)|\left|e^{-\int_t^s c(X(r))dr} - 1\right|\Big) \\ &\quad\le C_R \max\left(e^{-|t-s|L} - 1, \left(1 - e^{-C_\epsilon|t-s|}\right) + \mathbb{P}(\sup_{t\le r\le T}|X(r)| \ge \frac{1}{\epsilon})\right) \\ &\qquad\le C_R \max\left(e^{-|t-s|L} - 1, \left(1 - e^{-C_\epsilon|t-s|}\right) + \gamma(\epsilon, R)\right), \quad (3.153)\end{aligned}$$

and D_2 satisfies the same estimate. Plugging (3.152) and (3.153) into (3.151) and taking infimum over $\epsilon > 0$ provides (3.150) □

PROPOSITION 3.62 (Regularity of V under strong B-condition) *Let B satisfy the strong B-condition for A (Definition 3.10). Let $b\colon [0,T]\times H\times\Lambda \to H$, $\sigma\colon [0,T]\times H\times\Lambda \to \mathcal{L}_2(\Xi_0, H)$, $l\colon [0,T]\times H\times\Lambda \to \mathbb{R}$ be continuous, let b and σ satisfy (3.7), (3.9), (3.10) with $\gamma = 1$ and (3.23), and let c be bounded from below. Suppose that there exist local moduli $\omega_l(\cdot,\cdot)$ and $\omega(\cdot,\cdot)$ such that*

$$|l(t,x,a) - l(s,y,a)| \le \omega_l(|x-y| + |s-t|, R), \tag{3.154}$$

for all $x, y \in B(0,R)$, $a \in \Lambda$, $s,t \in [0,T]$ and

$$|g(x) - g(y)|, |c(x) - c(y)| \le \omega(|x-y|, R), \tag{3.155}$$

for all $x, y \in B(0,R)$, and that (3.137) holds.

Then:

(i) *The functions J and V satisfy (3.139) and there exists a local modulus $\sigma(\cdot,\cdot)$ such that*

$$|J(t,x;a(\cdot)) - J(t,y;a(\cdot))| \le \sigma(|x-y|, R) \tag{3.156}$$

for all $x, y \in B(0,R)$, $t \in [0,T]$, $a(\cdot) \in \mathcal{U}_t$.

(ii) *For any $\tau \in (0,T)$, there exists a local modulus $\sigma_\tau(\cdot,\cdot)$ such that*

$$|V(t,x) - V(s,y)| \le \sigma_\tau(|x-y|_{-1}, R) \tag{3.157}$$

for all $x, y \in B(0,R)$, $t \in [0,\tau]$.

(iii) *There exists a local modulus $\rho(\cdot,\cdot)$ such that*

$$|V(t,x) - V(s, e^{(s-t)A}x)| \le \rho(s-t, R) \tag{3.158}$$

for all $x \in B(0,R)$, $0 \le t \le s \le T$.

PROOF. Obviously J and V satisfy (3.139) as in Proposition 3.61. Also (3.156) is proved exactly as (3.138) in Proposition 3.61. The only difference is that, since l, g, c are now continuous in the usual norm of H instead of the $|\cdot|_{-1}$ norm, we have to replace (3.24) by (1.40).

To show (3.157) we begin as in (3.141). The term I_1 is estimated in exactly the same way as in the proof of Proposition 3.61 using (3.29) instead of (3.24). For the term I_2 we have

$$I_2 \le I_{21} + I_{22} := \mathbb{E}\left[e^{-\int_t^T c(X(r))dr} |g(X(T)) - g(Y(T))| \right] + \mathbb{E}\left[|g(Y(T))| \left| e^{-\int_t^T c(X(r))dr} - e^{-\int_t^T c(Y(r))dr} \right| \right].$$

The term I_{22} is again standard if we use (3.29). If g satisfied (3.136) we could also proceed as before with the term I_{21} to obtain (3.160) (see Remark 3.63). Since g only satisfies (3.155) we have to proceed slightly differently. We have, by (3.137), (3.155), (3.143), (3.144), (3.30)

$$\begin{aligned} I_{21} &\le e^{-TL}\mathbb{E}\,|g(X(T)) - g(Y(T))| \le \mathbb{E}\,\omega\left(|X(T) - Y(T)|, \frac{1}{\epsilon} \right) \\ &+ e^{-TL} \int_{\{|X(T)|\ge \frac{1}{\epsilon} \text{ or } |Y(T)|\ge\frac{1}{\epsilon}\}} C(2 + |X(T)|^m + |Y(T)|^m) d\mathbb{P}. \\ &\le e^{-TL}(\epsilon + K_\epsilon \mathbb{E}|X(T) - Y(T)|) + \gamma(\epsilon, R) \\ &\le e^{-TL}\epsilon + e^{-TL}\left(\frac{C(T)}{\tau}\right)^{\frac{1}{2}} |x-y|_{-1} + \gamma(\epsilon, R), \end{aligned}$$

where $C(T)$ is the constant from (3.30) and γ is a local modulus. It remains to take the infimum over all $\epsilon > 0$

The proof of (3.158) is also very similar to the proof of (3.150). We can now claim that the dynamic programming principle is satisfied and thus, as in (3.159), if $x \in B(0,R)$ and $0 \le t \le s \le T$, we have

$$\begin{aligned} |V(t,x) - V(s, e^{(s-t)A}x)| &\le \sup_{a(\cdot)\in\mathcal{U}_t} e^{-LT}\mathbb{E}\int_t^s |l(r, X(r), a(r))| dr \\ &+ e^{-LT} \sup_{a(\cdot)\in\mathcal{U}_t} \mathbb{E}\left| V(s, e^{(s-t)A}x) - V(s, X(s)) \right| \\ &+ \sup_{a(\cdot)\in\mathcal{U}_t} \mathbb{E}\left| V(s, e^{(s-t)A}x)\left(e^{-\int_t^s c(X(r))dr} - 1 \right) \right|, \end{aligned} \tag{3.159}$$

where $X(r)$ is the solution of (3.125). The first and the third term above are estimated as in (3.151) and (3.153). For the middle term we first notice that there exists some constant C_R depending on R but independent of the control such that

$$\mathbb{E}|X(s) - e^{(s-t)A}x|^2 \le C_R(s-t).$$

Therefore, using (3.156),

$$\begin{aligned} \mathbb{E}\left| V(s, e^{(s-t)A}x) - V(s, X(s)) \right| &\le \mathbb{E}\,\sigma\left(|X(s) - e^{(s-t)A}x|, \frac{1}{\epsilon} \right) \\ &+ C_R\left(\mathbb{P}(\sup_{t\le r\le T} |X(r)| \ge \frac{1}{\epsilon}) \right)^{\frac{1}{2}} \le \epsilon + C_{R,\epsilon}|t-s|^{\frac{1}{2}} + \gamma(\epsilon, R), \end{aligned}$$

which implies the claim as all the constants and the local modulus γ are independent of t and the controls. □

REMARK 3.63 It follows easily from the above proof that if g satisfies (3.136) instead of (3.155), then

$$|V(t,x) - V(s,y)| \le \omega(|t-s| + |x-y|_{-1}, R) \qquad \forall x, y \in B(0,R),\ t, s \in [0,T] \tag{3.160}$$

for some local modulus $\omega(\cdot,\cdot)$. ■

REMARK 3.64 It is clear from the proof that (3.156), and the same estimate for V, still holds if (3.23) is replaced by (3.8) and the strong B-condition for A is replaced by a standard requirement that A generates a C_0-semigroup. ■

In the next lemma we provide Itô like formulae for test functions $\psi = \varphi + h(t,|x|)$ introduced in Definition 3.32. As we remarked after this definition, even though $|x|$ is not differentiable at 0, the function $h_0(t,x) := h(t,|x|) \in C^{1,2}((0,T)\times H)$, so with a slight abuse of notation, in the following we will write $h(t,x)$ instead of $h(t,|x|)$, $Dh(t,x)$ instead of $Dh_0(t,x) = \frac{x}{|x|}\frac{d}{dr}h(t,r)|_{r=|x|}$ (which is 0 when $x=0$), and $D^2h(t,x)$ instead of $D^2h_0(t,x)$.

LEMMA 3.65 *Let b and σ be continuous, satisfy (3.7), (3.8), (3.9), (3.10), and let $c: H \to \mathbb{R}$ be continuous, bounded from below and satisfy (3.137). Consider a test function (in the sense of Definition 3.32) $\psi = \varphi + h$. Suppose that h satisfies (1.105) and consider the solution $X(\cdot)$ of (3.125) for a given control $a(\cdot) \in \mathcal{U}_t$. Then, for any $s \in [t,T]$,*

$$\begin{aligned}\mathbb{E}\left[e^{-\int_t^s c(X(\tau))d\tau}\varphi(s,X(s))\right] = \varphi(t,x)&\\ \mathbb{E}\left[\int_t^s e^{-\int_t^r c(X(\tau))d\tau}\Big(\varphi_t(r,X(r)) + \langle X(r), A^*D\varphi(r,X(r))\rangle\right.&\\ + \langle b(r,X(r),a(r)), D\varphi(r,X(r))\rangle - c(X(r))\varphi(r,X(r))&\\ \left.+\frac{1}{2}\mathrm{Tr}\left[\left(\sigma(r,X(r),a(r))Q^{\frac{1}{2}}\right)\left(\sigma(r,X(r),a(r))Q^{\frac{1}{2}}\right)^* D^2\varphi(r,X(r))\right]\Big)dr\right]&\end{aligned} \tag{3.161}$$

and

$$\begin{aligned}\mathbb{E}\left[e^{-\int_t^s c(X(\tau))d\tau}h(s,X(s))\right] \le h(t,x) + \mathbb{E}\left[\int_t^s e^{-\int_t^r c(X(\tau))d\tau}\Big(h_t(r,X(r))\right.&\\ + \langle b(r,X(r),a(r)), Dh(r,X(r))\rangle - c(X(r))h(r,X(r))&\\ \left.+\frac{1}{2}\mathrm{Tr}\left[\left(\sigma(r,X(r),a(r))Q^{\frac{1}{2}}\right)\left(\sigma(r,X(r),a(r))Q^{\frac{1}{2}}\right)^* D^2h(r,X(r))\right]\Big)dr\right].&\end{aligned} \tag{3.162}$$

PROOF. We define $c_n(y) := \min(c(y),n)$ and observe that, $\eta(r) := e^{-\int_t^r c_n(X(\tau))d\tau}$ is the unique solution of $d\eta_n(r) = b_n(r)dr$ (and $\eta_n(t)=1$), where $b_n(r) = -\eta_n(r)c_n(X(\tau))$ is bounded. We can thus use Propositions 1.156 and 1.157 and send $n \to +\infty$ to obtain the claim. □

THEOREM 3.66 (Existence under weak B-condition) *Let the assumptions of Proposition 3.61 be satisfied, and let in addition $b(\cdot,x,a)$ and $\sigma(\cdot,x,a)$ be uniformly continuous on $[0,T]$, uniformly in $(x,a) \in B(0,R)\times\Lambda$ for every $R>0$. Suppose also that, for every (t,x),*

$$\lim_{N\to+\infty}\sup_{a\in\Lambda}\mathrm{Tr}\left[\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^* BQ_N\right] = 0. \tag{3.163}$$

Then the value function $V(t,x)$, defined in (3.127), is the unique viscosity solution of (3.128) among functions in the set

$$\begin{aligned}S := \Big\{u\colon [0,T]\times H \to \mathbb{R} \;:\; |u(t,x)| \le C(1+|x|^k) \text{ for some } k \ge 0,&\\ \lim_{t\to T}|u(t,x)-g(x)| = 0 \text{ uniformly on bounded subsets of } H\Big\}.&\end{aligned}$$

PROOF. Without loss of generality we can assume that c is positive since if $c \ge L$ for $L<0$ then V is a viscosity solution of (3.128) if and only if $\tilde{V} = e^{L(T-t)}V$ is a viscosity solution of (3.128) with c replaced by $\tilde{c} = c - L$ and l replaced by $e^{L(T-t)}l$.

Existence: Proposition 3.61 ensures that V is B-continuous and that it belongs to S. We first prove that V is a viscosity supersolution of (3.128). Let $V+\psi$ have a local minimum at $(t,x) \in (0,T) \times H$ for a test function $\psi = \varphi + h$ (in sense of Definition 3.32). We can assume that h and its derivatives Dh, D^2h, h_t have polynomial growth (see on this the discussion following Lemma 3.37), that the minimum is global (see Lemma 3.37), and that $V(t,x) + \psi(t,x) = 0$, so for all (s,y) we have $V(s,y) \geq -\psi(s,y)$.

By Proposition 2.16, Theorem 2.24, and Proposition 3.61, the dynamic programming principle (2.25) is satisfied. Thus for $\epsilon > 0$ there exists a control $a^{\nu_\epsilon}(\cdot) \in \mathcal{U}_t$ such that, denoting $X(\cdot) := X(\cdot; t, x; a^{\nu_\epsilon}(\cdot))$,

$$
\begin{aligned}
V(t,x) + \epsilon^2 \geq \mathbb{E}^{\nu_\epsilon} \bigg[\int_t^{t+\epsilon} & e^{-\int_t^r c(X(\tau))d\tau} l(r, X(r), a^{\nu_\epsilon}(r)) dr \\
& + e^{-\int_t^{t+\epsilon} c(X(\tau))d\tau} V(t+\epsilon, X(t+\epsilon)) \bigg].
\end{aligned}
$$

This implies that

$$
\begin{aligned}
\epsilon^2 - \varphi(t,x) - h(t,x) \geq \mathbb{E}^{\nu_\epsilon} \bigg[\int_t^{t+\epsilon} & e^{-\int_t^r c(X(\tau))d\tau} l(r, X(r), a^{\nu_\epsilon}(r)) dr \\
- e^{-\int_t^{t+\epsilon} c(X(\tau))d\tau} \varphi(t+\epsilon, X(t+\epsilon)) & - e^{-\int_t^{t+\epsilon} c(X(\tau))d\tau} h(t+\epsilon, X(t+\epsilon)) \bigg]. \quad (3.164)
\end{aligned}
$$

Using (3.161), (3.162) and (3.164) we find

$$
\begin{aligned}
0 \leq \epsilon + \frac{1}{\epsilon} \mathbb{E}^{\nu_\epsilon} \bigg[\int_t^{t+\epsilon} & e^{-\int_t^r c(X(\tau))d\tau} \bigg(- l(r, X(r), a^{\nu_\epsilon}(r)) \\
& + \psi_t(r, X(r)) + \langle b(r, X(r), a^{\nu_\epsilon}(r)), D\psi(r, X(r)) \rangle \\
+ \frac{1}{2} \mathrm{Tr} \Big[\Big(\sigma(r, X(r), a^{\nu_\epsilon}(r)) Q^{\frac{1}{2}} \Big) & \Big(\sigma(r, X(r), a^{\nu_\epsilon}(r)) Q^{\frac{1}{2}} \Big)^* D^2\psi(r, X(r)) \Big] \\
& - c(X(r))\psi(r, X(r)) + \langle X(r), A^* D\varphi(r, X(r)) \rangle \bigg) dr \bigg]. \quad (3.165)
\end{aligned}
$$

Now we observe that, thanks to (1.39), we can find a constant $r_\epsilon > 0$, depending on $\epsilon > 0$ but independent of the control $a^{\nu_\epsilon}(\cdot)$, such that $r_\epsilon \xrightarrow{\epsilon \to 0^+} 0$, and the set

$$
\Omega_1^\epsilon = \left\{ \omega \in \Omega^{\nu_\epsilon} : \sup_{r \in [t, t+\epsilon]} |X(r) - x| \leq r_\epsilon \right\},
$$

satisfies

$$
\mathbb{P}^{\nu_\epsilon}(\Omega_1^\epsilon) \to 1 \quad \text{as } \epsilon \to 0. \tag{3.166}
$$

We set $\Omega_2^\epsilon = \Omega^{\nu_\epsilon} \backslash \Omega_1^\epsilon$. If we denote by $\Psi(r)$ the integrand in (3.165), the assumptions and properties of test functions imply,

$$
|\Psi(r)| \leq C(1 + |X(r)|^N) \tag{3.167}
$$

for some $N \geq 0$. Thus, by (1.38), (3.166), (3.167), and the continuity of the functions in the integrand , we obtain

$$
\begin{aligned}
0 \leq \epsilon + \frac{1}{\epsilon}\mathbb{E}^{\nu_\epsilon}\Big[\int_t^{t+\epsilon} \Big(& - l(t,x,a^{\nu_\epsilon}(r)) + \psi_t(t,x) + \langle b(t,x,a^{\nu_\epsilon}(r)), D\psi(r,X(r))\rangle \\
& + \frac{1}{2}\mathrm{Tr}\left[\left(\sigma(t,x,a^{\nu_\epsilon}(r))Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a^{\nu_\epsilon}(r))Q^{\frac{1}{2}}\right)^* D^2\psi(t,x)\right] \\
& - c(x)\psi(t,x) + \langle x, A^* D\varphi(t,x)\rangle \Big)\mathbf{1}_{\Omega_1^\epsilon} dr\Big] \\
& + C\frac{1}{\epsilon}\int_t^{t+\epsilon} (\mathbb{P}(\Omega_2^\epsilon)^{\frac{1}{2}}(\mathbb{E}[1+|X(r)|^N]^2)^{\frac{1}{2}} + \gamma_1(\epsilon) \\
\leq \frac{1}{\epsilon}\mathbb{E}^{\nu_\epsilon}\Big[\int_t^{t+\epsilon} \Big(& - l(t,x,a^{\nu_\epsilon}(r)) + \psi_t(t,x) + \langle b(t,x,a^{\nu_\epsilon}(r)), D\psi(r,X(r))\rangle \\
& + \frac{1}{2}\mathrm{Tr}\left[\left(\sigma(t,x,a^{\nu_\epsilon}(r))Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a^{\nu_\epsilon}(r))Q^{\frac{1}{2}}\right)^* D^2\psi(t,x)\right] \\
& + c(x)V(t,x) + \langle x, A^* D\varphi(t,x)\rangle \Big) dr\Big] + \gamma_2(\epsilon) \\
& \leq \frac{1}{\epsilon}\int_t^{t+\epsilon} \mathbb{E}^{\nu_\epsilon}\big[\psi_t(t,x) + \langle x, A^* D\varphi(t,x)\rangle \\
& + F(t,x,V(t,x),-D\psi(t,x),-D^2\psi(t,x))\big] dr + \gamma_2(\epsilon) \\
= \psi_t(t,x) + \langle x, A^* D\varphi(t,x)\rangle & + F(t,x,V(t,x),-D\psi(t,x),-D^2\psi(t,x)) + \gamma_2(\epsilon),
\end{aligned}
$$

where γ_1, γ_2 above are such that $\lim_{\epsilon\to 0}\gamma_i(\epsilon) = 0, i = 1,2$, and are independent of the control $a^{\nu_\epsilon}(r)$ and of the reference probability space ν_ϵ. The claim follows after we let $\epsilon \to 0$.

To show the subsolution property, let $V - \psi$ has a global maximum at (t,x), where h and its derivatives Dh, D^2h, h_t have polynomial growth, and $V(t,x) = \psi(t,x)$. We choose $a \in \Lambda$ and take a constant control $a(s) \equiv a$ defined on some reference probability space, and we denote again $X(\cdot) := X(\cdot;t,x;a)$. Using dynamic programming principle (2.25) we have

$$
\begin{aligned}
& \psi(t,x) = V(t,x) \\
& \leq \mathbb{E}\left[\int_t^{t+\epsilon} e^{-\int_t^r c(X(\tau))d\tau} l(r,X(r),a)dr + e^{-\int_t^{t+\epsilon} c(X(\tau))d\tau} V(t+\epsilon, X(t+\epsilon))\right] \\
& \leq \mathbb{E}\left[\int_t^{t+\epsilon} e^{-\int_t^r c(X(\tau))d\tau} l(r,X(r),a)dr + e^{-\int_t^{t+\epsilon} c(X(\tau))d\tau} \psi(t+\epsilon, X(t+\epsilon))\right]
\end{aligned}
$$

and then as before we get

$$
\begin{aligned}
\frac{1}{\epsilon}\mathbb{E}\Big[\int_t^{t+\epsilon} e^{-\int_t^r c(X(\tau))d\tau}\Big(& l(r,X(r),a) + \psi_t(r,X(r)) + \langle X(r), A^* D\varphi(r,X(r))\rangle \\
& - c(X(r))\psi(r,X(r)) + \langle b(r,X(r),a), D\varphi(r,X(r))\rangle \\
+ \frac{1}{2}\mathrm{Tr}&\left[\left(\sigma(r,X(r),a)Q^{\frac{1}{2}}\right)\left(\sigma(r,X(r),a)Q^{\frac{1}{2}}\right)^* D^2\psi(r,X(r))\right]\Big)dr\Big] \geq 0.
\end{aligned}
$$

The same argument as in the proof of the supersolution part now yields

$$
\begin{aligned}
& l(t,x,a) + \psi_t(t,x) + \langle x, A^* D\varphi(t,x)\rangle - c(x)V(t,x) \\
& \quad + \langle b(t,x,a), D\varphi(t,x)\rangle + \frac{1}{2}\mathrm{Tr}\left[\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^* D^2\psi(t,x)\right] \geq 0
\end{aligned}
\tag{3.168}
$$

and the claim follows after we take the $\inf_{a\in\Lambda}$ in (3.168).

Uniqueness: To prove the uniqueness of the solution we need to show that the hypotheses of Theorem 3.50 are satisfied with the set $U = H$.

Hypothesis 3.44 follows from the local uniform continuity of $b(\cdot,\cdot,a), \sigma(\cdot,\cdot,a), l(\cdot,\cdot,a), c(\cdot)$, uniform in $a \in \Lambda$, (3.9), (3.10), and (3.137). Hypothesis 3.45 follows from the positivity of c. For Hypothesis 3.46 we can argue as follows: since $\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^*$ is a positive, self-adjoint, trace class operator, it is obvious that, for $X, Y \in S(H)$ with $X \le Y$,

$$-\mathrm{Tr}\left(\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^* X\right) \ge -\mathrm{Tr}\left(\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^* Y\right),$$

and then taking the supremum over $a \in \Lambda$ we see that Hypothesis 3.46 is satisfied. Hypothesis 3.47 follows from (3.163) (see further comments about it after the end of the proof).

To show that Hypothesis 3.48 holds observe that, using (3.22), (3.136) and (3.135), we have, for $|r|, |x|, |y| \le R$,

$$\begin{aligned}
&F\left(t,x,r,\frac{B(x-y)}{\epsilon},X\right) - F\left(t,y,r,\frac{B(x-y)}{\epsilon},Y\right)\\
&\ge -\sup_{a\in\Lambda}\Bigg(l(t,x,a)-l(t,y,a)+\left\langle\frac{B(x-y)}{\epsilon},b(t,x,a)\right\rangle-\left\langle\frac{B(x-y)}{\epsilon},b(t,y,a)\right\rangle\\
&\qquad -r(c(x)-c(y))+\frac{1}{2}\mathrm{Tr}\left(\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^*X\right)\\
&\qquad\qquad -\frac{1}{2}\mathrm{Tr}\left(\left(\sigma(t,y,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,y,a)Q^{\frac{1}{2}}\right)^*Y\right)\Bigg)\\
&\ge -\sup_{a\in\Lambda}|l(t,x,a)-l(t,y,a)| - R\sup_{a\in\Lambda}|c(x)-c(y)|\\
&\qquad -\sup_{a\in\Lambda}\left\langle\frac{B(x-y)}{\epsilon},b(t,x,a)-b(t,y,a)\right\rangle\\
&\qquad -\sup_{a\in\Lambda}\Bigg(\frac{1}{2}\mathrm{Tr}\left(\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^*X\right)\\
&\qquad\qquad -\frac{1}{2}\mathrm{Tr}\left(\left(\sigma(t,y,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,y,a)Q^{\frac{1}{2}}\right)^*Y\right)\Bigg)\\
&\ge -\omega_l(|x-y|_{-1},R) - C\frac{|x-y|_{-1}^2}{\epsilon} - R\omega_c(|x-y|_{-1},R)\\
&\qquad -\sup_{a\in\Lambda}\Bigg(\frac{1}{2}\mathrm{Tr}\left(\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^*X\right)\\
&\qquad\qquad -\frac{1}{2}\mathrm{Tr}\left(\left(\sigma(t,y,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,y,a)Q^{\frac{1}{2}}\right)^*Y\right)\Bigg).
\end{aligned}$$

To estimate the last term we use that X and Y satisfy (3.54). In particular we have

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \le \frac{3}{\epsilon}\begin{pmatrix} B & -B \\ -B & B \end{pmatrix}.$$

Multiplying both sides of this inequality by the operator

$$Z = \begin{pmatrix} \left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^* & \left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,y,a)Q^{\frac{1}{2}}\right)^* \\ \left(\sigma(t,y,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^* & \left(\sigma(t,y,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,y,a)Q^{\frac{1}{2}}\right)^* \end{pmatrix}$$

and taking the trace preserves the inequality. This can be seen by evaluating the trace on the basis of eigenvectors of Z as it is a compact, self-adjoint, and positive operator. Therefore, thanks to (3.23),

$$\begin{aligned}\operatorname{Tr}\left(\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^{*}X\right)-\operatorname{Tr}\left(\left(\sigma(t,y,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,y,a)Q^{\frac{1}{2}}\right)^{*}Y\right)\\ \leq\frac{3}{\epsilon}\operatorname{Tr}\Big[\left(\left(\sigma(t,x,a)-\sigma(t,y,a)\right)Q^{\frac{1}{2}}\right)\left(\left(\sigma(t,x,a)-\sigma(t,y,a)\right)Q^{\frac{1}{2}}\right)^{*}B\Big]\\ \leq C\frac{|x-y|_{-1}^{2}}{\epsilon},\end{aligned}\tag{3.169}$$

for all $a\in\Lambda$ for some C. We thus conclude that

$$\begin{aligned}F\left(t,x,r,\frac{B(x-y)}{\epsilon},X\right)-F\left(t,y,r,\frac{B(x-y)}{\epsilon},Y\right)\\ \geq-\omega_{l}(|x-y|_{-1},R)-R\omega_{c}(|x-y|_{-1},R)-C\frac{|x-y|_{-1}^{2}}{\epsilon}\end{aligned}\tag{3.170}$$

for some constant C, and so Hypothesis 3.48 is satisfied. Hypothesis 3.49 with $\gamma=2$ follows from (3.9) and (3.10). This concludes the proof of the uniqueness. □

Let us analyze condition (3.163). Let $\{u_1,u_2,...\}$ be any orthonormal basis of Ξ. Let $\{e_1,e_2,...\}$ be an orthonormal basis in H_{-1} made of elements of H as in Section 3.2. Then $\{f_1,f_2,...\}$, where $f_i=B^{\frac{1}{2}}e_i$ is an orthonormal basis of H.

We have

$$\begin{aligned}&\operatorname{Tr}\left[\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^{*}BQ_N\right]\\ &=\operatorname{Tr}\left[(\sigma(t,x,a)Q^{\frac{1}{2}})^{*}BQ_N(\sigma(t,x,a)Q^{\frac{1}{2}})\right]\\ &=\operatorname{Tr}\left[(\sigma(t,x,a)Q^{\frac{1}{2}})^{*}Q_N^{*}BQ_N(\sigma(t,x,a)Q^{\frac{1}{2}})\right]\\ &=\sum_{i=1}^{\infty}\left\langle BQ_N\sigma(t,x,a)Q^{\frac{1}{2}}u_i,Q_N\sigma(t,x,a)Q^{\frac{1}{2}}u_i\right\rangle=\sum_{i=1}^{\infty}|Q_N\sigma(t,x,a)Q^{\frac{1}{2}}u_i|_{-1}^{2}\\ &=\sum_{i=1}^{\infty}|B^{\frac{1}{2}}Q_N\sigma(t,x,a)Q^{\frac{1}{2}}u_i|^{2}=\sum_{i=1}^{\infty}|(\sigma(t,x,a)Q^{\frac{1}{2}})^{*}Q_N^{*}B^{\frac{1}{2}}f_i|_{\Xi}^{2}\\ &=\sum_{i=1}^{\infty}|(\sigma(t,x,a)Q^{\frac{1}{2}})^{*}Q_N^{*}Be_i|_{\Xi}^{2}=\sum_{i=N+1}^{\infty}|(\sigma(t,x,a)Q^{\frac{1}{2}})^{*}Be_i|_{\Xi}^{2}\\ &=\sum_{i=N+1}^{\infty}|(\sigma(t,x,a)Q^{\frac{1}{2}})^{*}B^{\frac{1}{2}}f_i|_{\Xi}^{2}=\sum_{i=N+1}^{\infty}|(B^{\frac{1}{2}}\sigma(t,x,a)Q^{\frac{1}{2}})^{*}f_i|_{\Xi}^{2}\end{aligned}$$

(see also [286], page 33). Therefore, we have

$$\begin{aligned}f_N(a):=\operatorname{Tr}\left[\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^{*}BQ_N\right]=\textstyle\sum_{i=1}^{\infty}|Q_N\sigma(t,x,a)Q^{\frac{1}{2}}u_i|_{-1}^{2}\\ =\textstyle\sum_{i=N+1}^{\infty}|(B^{\frac{1}{2}}\sigma(t,x,a)Q^{\frac{1}{2}})^{*}f_i|_{\Xi}^{2}.\end{aligned}\tag{3.171}$$

The functions $f_N:\Lambda\to\mathbb{R}$, $N\geq 1$, are continuous, nonnegative, and since by (3.10),

$$\sum_{i=1}^{\infty}|(B^{\frac{1}{2}}\sigma(t,x,a)Q^{\frac{1}{2}})^{*}f_i|_{\Xi}^{2}=\operatorname{Tr}\left[\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)\left(\sigma(t,x,a)Q^{\frac{1}{2}}\right)^{*}B\right]\leq C_1,$$

it follows from (3.171) that for every $a \in \Lambda$, $f_N(a) \downarrow 0$ as $N \to +\infty$. Thus, if Λ is compact, we must have $f_N(a) \to 0$ uniformly on Λ as $N \to +\infty$, which means that (3.163) is satisfied.

Another case when (3.163) is satisfied is when B is compact. This is an obvious consequence of the fact that in this case $\|BQ_N\| \to 0$ as $N \to +\infty$.

One can use (3.171) to obtain other criteria for (3.163) to hold. For instance it will be satisfied if

$$\sum_{i=1}^{\infty} a_i < +\infty,$$

where

$$a_i := \sup_{a \in \Lambda} |\sigma(t,x,a) Q^{\frac{1}{2}} u_i|_{-1}^2,$$

and if for every i

$$\lim_{N \to +\infty} \sup_{a \in \Lambda} |Q_N \sigma(t,x,a) Q^{\frac{1}{2}} u_i|_{-1} = 0.$$

THEOREM 3.67 (Existence under strong B-condition) *Let the assumptions of Proposition 3.62 and (3.163) be satisfied, and let in addition $b(\cdot,x,a)$, $\sigma(\cdot,x,a)$ be uniformly continuous on $[0,T]$, uniformly for $(x,a) \in B(0,R) \times \Lambda$ for every $R > 0$. Then the value function $V(t,x)$, defined in (3.127) is the unique viscosity solution of (3.128) among functions in the set*

$$S := \Big\{ u \colon [0,T] \times H \to \mathbb{R} \ : \ |u(t,x)| \le C(1+|x|^k) \ \textit{for some}\ k \ge 0,$$
$$\lim_{t \to T} |u(t,x) - g(e^{(T-t)A}x)| = 0 \ \textit{uniformly on bounded subsets of}\ H \Big\}.$$

PROOF. The proof follows the lines of the proof for the weak case. To prove uniqueness we now use Theorem 3.54 instead of Theorem 3.50 so we need to verify Hypothesis 3.52 instead of Hypothesis 3.48. This can be done arguing as before using (3.7) instead of (3.22). □

REMARK 3.68 B-continuity is built into the definition of viscosity solution, however it is clear from the proof of existence that B-continuity of the value function is not needed to show that it satisfies the sub- and supersolution conditions required by the definition of viscosity solution. Thus, if we disregard the requirement of B-continuity, we can still prove that the value function is a "viscosity solution" under much weaker sets of assumptions than these of Theorems 3.66 and 3.67. ■

EXAMPLE 3.69 (Controlled stochastic wave equation) Consider a control problem for the stochastic wave equation

$$\begin{cases} \frac{\partial^2 y}{\partial s^2}(s,\xi) = \Delta y(s,\xi) + f(\xi, y(s,\xi), a(s)) & \\ \qquad\qquad + h(\xi, y(s,\xi), a(s)) \frac{\partial}{\partial s} \tilde{W}_{\tilde{Q}}(s,\xi), & s > t, \ \xi \in \mathcal{O}, \\ y(s,\xi) = 0, & s > t, \ \xi \in \partial\mathcal{O}, \\ y(t,\xi) = y_0(\xi), & \xi \in \mathcal{O}, \\ \frac{\partial y}{\partial t}(t,\xi) = z_0(\xi), & \xi \in \mathcal{O}, \end{cases} \tag{3.172}$$

where $\mathcal{O}$ is a bounded regular domain in $\mathbb{R}^d$, $y_0 \in H_0^1(\mathcal{O})$, $z_0 \in L^2(\mathcal{O})$, $\tilde{Q}$ is an operator in $\mathcal{L}_1^+(L^2(\mathcal{O}))$ and $\tilde{W}_{\tilde{Q}}$ is a $\tilde{Q}$-Wiener process, Λ is a Polish space and $a(\cdot) \in \mathcal{U}_t$. In addition $f, h : \mathcal{O} \times \mathbb{R} \times \Lambda \to \mathbb{R}$. Suppose we want to minimize the cost functional

$$I(t, y_0, z_0; a(\cdot)) = \mathbb{E}\left[\int_t^T \int_{\mathcal{O}} \beta(s, y(s,\xi), a(s)) d\xi ds + \int_{\mathcal{O}} \gamma(y(T,\xi)) d\xi \right]$$

over all $a(\cdot) \in \mathcal{U}_t$, where $\beta : [0,T] \times \mathbb{R} \times \Lambda \to \mathbb{R}, \gamma : \mathbb{R} \to \mathbb{R}$.

Let Δ_ξ be the Laplace operator with the domain $D(\Delta_\xi) = H^2(\mathcal{O}) \cap H^1_0(\mathcal{O})$. Then (see Section C.1 and in particular (C.11)) $D((-\Delta_\xi)^{\frac{1}{2}}) = H^1_0(\mathcal{O})$. We set[1]

$$H = \begin{pmatrix} H^1_0(\mathcal{O}) \\ \times \\ L^2(\mathcal{O}) \end{pmatrix}$$

equipped with the inner product

$$\left\langle \begin{pmatrix} y \\ z \end{pmatrix}, \begin{pmatrix} \bar{y} \\ \bar{z} \end{pmatrix} \right\rangle_H = \langle (-\Delta_\xi)^{1/2} y, (-\Delta_\xi)^{1/2} \bar{y} \rangle_{L^2(\mathcal{O})} + \langle z, \bar{z} \rangle_{L^2(\mathcal{O})}, \qquad \begin{pmatrix} y \\ z \end{pmatrix}, \begin{pmatrix} \bar{y} \\ \bar{z} \end{pmatrix} \in H.$$

The operator

$$D(A) = \begin{pmatrix} H^2(\mathcal{O}) \cap H^1_0(\mathcal{O}) \\ \times \\ H^1_0(\mathcal{O}) \end{pmatrix}, \qquad A = \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix},$$

is maximal dissipative in H and $A^* = -A$. Equation (3.172) can then be rewritten as the following evolution equation

$$dX(s) = (AX(s) + b(X(s), a(s)))\, dt + \sigma(X(s), a(s)) dW_Q(s), \quad X(t) = x := \begin{pmatrix} y_0 \\ z_0 \end{pmatrix}, \tag{3.173}$$

in H, where

$$b\left(\begin{pmatrix} y \\ z \end{pmatrix}, a \right) = \begin{pmatrix} 0 \\ f(\cdot, y(\cdot), a) \end{pmatrix}, \quad \sigma\left(\begin{pmatrix} y \\ z \end{pmatrix}, a \right) \begin{pmatrix} \bar{y} \\ \bar{z} \end{pmatrix} = \begin{pmatrix} 0 \\ h(\cdot, y(\cdot), a)\bar{z} \end{pmatrix}, \tag{3.174}$$

$$W_Q = \begin{pmatrix} 0 \\ \tilde{W}_{\tilde{Q}} \end{pmatrix}, \quad Q \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ \tilde{Q} z \end{pmatrix}.$$

We consider the operator

$$B = \begin{pmatrix} (-\Delta_\xi)^{-1/2} & 0 \\ 0 & (-\Delta_\xi)^{-1/2} \end{pmatrix}.$$

It is bounded, positive, self-adjoint on H, A^*B is bounded and

$$\left\langle A^* B \begin{pmatrix} y \\ z \end{pmatrix}, \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle_H = 0.$$

Moreover we have

$$\left| \begin{pmatrix} y \\ z \end{pmatrix} \right|_{-1} = \left(|(-\Delta_\xi)^{1/4} y|^2 + |(-\Delta_\xi)^{-1/4} z|^2 \right)^{1/2}.$$

Assume that f, h are continuous in all variables, $f(\xi, \cdot, a), h(\xi, \cdot, a)$ are Lipschitz continuous with Lipschitz constant L independent of ξ, a, and $f(\cdot, 0, \cdot), h(\cdot, 0, \cdot)$ are bounded. Then

$$\begin{aligned} &\left| b\left(\begin{pmatrix} y \\ z \end{pmatrix}, a \right) - b\left(\begin{pmatrix} \tilde{y} \\ \tilde{z} \end{pmatrix}, a \right) \right|_H = |f(\cdot, y(\cdot), a) - f(\cdot, \tilde{y}(\cdot), a)|_{L^2(\mathcal{O})} \\ &\le L \left(\int_{\mathcal{O}} |y(\xi) - \tilde{y}(\xi)|^2 d\xi \right)^{\frac{1}{2}} = L |y - \tilde{y}|_{L^2(\mathcal{O})} \\ &\le C |(-\Delta_\xi)^{1/4} (y - \tilde{y})|_{L^2(\mathcal{O})} \le C \left| \begin{pmatrix} y \\ z \end{pmatrix} - \begin{pmatrix} \tilde{y} \\ \tilde{z} \end{pmatrix} \right|_{-1} \end{aligned}$$

[1]We remark that if $\mathrm{Tr}(\tilde{Q}) = +\infty$ then the right choice of the state space for the stochastic wave equation is $L^2(\mathcal{O}) \times D((-\Delta_\xi)^{-\frac{1}{2}})$, at least for additive noise, see [130] Example 5.8 page 127.

for some constant C which follows from embeddings $D((-\Delta_\xi)^{1/4}) \hookrightarrow H^{1/2}(\mathcal{O}) \hookrightarrow L^2(\mathcal{O})$ (see Section C.1). Thus the function b satisfies (3.9) and (3.22). Unfortunately we need more in order for σ to satisfy (3.23) (or even (3.8)). We present one sufficient condition. Obviously other conditions are possible. Denote $\Xi_0 = Q^{1/2}H$. Suppose that $d > 1$ and $(-\Delta_\xi)^{(d-1)/4}\tilde{Q}^{1/2} \in \mathcal{L}_2(L^2(\mathcal{O}))$. Then we have by Proposition B.26

$$\begin{aligned}
&\left\| \sigma\left(\begin{pmatrix} y \\ z \end{pmatrix}, a\right) - \sigma\left(\begin{pmatrix} \tilde{y} \\ \tilde{z} \end{pmatrix}, a\right) \right\|_{\mathcal{L}_2(\Xi_0, H)} \\
&\qquad = \left\| (h(\cdot, y(\cdot), a) - h(\cdot, \tilde{y}(\cdot), a))\tilde{Q}^{1/2} \right\|_{\mathcal{L}_2(L^2(\mathcal{O}))} \\
&\le \left\| (h(\cdot, y(\cdot), a) - h(\cdot, \tilde{y}(\cdot), a))(-\Delta_\xi)^{(1-d)/4} \right\|_{\mathcal{L}(L^2(\mathcal{O}))} \left\| (-\Delta_\xi)^{(d-1)/4}\tilde{Q}^{1/2} \right\|_{\mathcal{L}_2(L^2(\mathcal{O}))} \\
&\le L \sup_{|z|_{L^2(\mathcal{O})} \le 1} \left(\int_{\mathcal{O}} |y(\xi) - \tilde{y}(\xi)|^2 |(-\Delta_\xi)^{(1-d)/4} z(\xi)|^2 d\xi \right)^{\frac{1}{2}} \left\| (-\Delta_\xi)^{(d-1)/4}\tilde{Q}^{1/2} \right\|_{\mathcal{L}_2(L^2(\mathcal{O}))} \\
&\le L|y - \tilde{y}|_{L^{2d/(d-1)}(\mathcal{O})} \sup_{|z|_{L^2(\mathcal{O})} \le 1} |(-\Delta_\xi)^{(1-d)/4} z|_{L^{2d}(\mathcal{O})} \left\| (-\Delta_\xi)^{(d-1)/4}\tilde{Q}^{1/2} \right\|_{\mathcal{L}_2(L^2(\mathcal{O}))} \\
&\qquad \le C|(-\Delta_\xi)^{1/4}(y - \tilde{y})|_{L^2(\mathcal{O})} \left\| (-\Delta_\xi)^{(d-1)/4}\tilde{Q}^{1/2} \right\|_{\mathcal{L}_2(L^2(\mathcal{O}))} \\
&\qquad \le C \left| \begin{pmatrix} y \\ z \end{pmatrix} - \begin{pmatrix} \tilde{y} \\ \tilde{z} \end{pmatrix} \right|_{-1} \left\| (-\Delta_\xi)^{(d-1)/4}\tilde{Q}^{1/2} \right\|_{\mathcal{L}_2(L^2(\mathcal{O}))}, \qquad (3.175)
\end{aligned}$$

where we have used Sobolev embeddings $H^{1/2}(\mathcal{O}) \hookrightarrow L^{2d/(d-1)}(\mathcal{O})$ and $|(-\Delta_\xi)^{(1-d)/4} z|_{L^{2d}(\mathcal{O})} \le C_1 |z|_{L^2(\mathcal{O})}$ if $d > 1$ (see Section C.1). Thus σ satisfies (3.23). The same calculation shows that for $d = 1$ it is enough that $(-\Delta_\xi)^\alpha \tilde{Q}^{1/2} \in \mathcal{L}_2(L^2(\mathcal{O}))$ for some $\alpha > 0$. It is now easy to check that (3.10) is true with $\gamma = 1$.

We now define

$$l\left(s, \begin{pmatrix} y \\ z \end{pmatrix}, a\right) = \begin{pmatrix} 0 \\ \int_{\mathcal{O}} \beta(s, y(\xi), a) d\xi \end{pmatrix}, \quad g\begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ \int_{\mathcal{O}} \gamma(y(\xi)) d\xi \end{pmatrix},$$

and rewrite the cost functional I as

$$J(t, x; a(\cdot)) = \mathbb{E}\left[\int_t^T l(s, X(s), a(s)) ds + g(X(T)) \right].$$

It is easy to see by calculations similar to these for b that if β is continuous and $\beta(\cdot, \cdot, a), \gamma$ are uniformly continuous with a modulus of continuity independent of a then

$$|l(s, x, a) - l(t, y, a)| + |g(x) - g(y)| \le \omega(|x - y|_{-1} + |s - t|), \quad s, t \in [0, T], x, y \in H, a \in \Lambda$$

for some modulus ω. If $\beta(0, 0, \cdot)$ is bounded then l satisfies (3.137) with $m = 1$. ■

3.6.2. Improved version of dynamic programming principle. Once we know that the value function is continuous in both variables, a stronger version of the dynamic programming principle involving stopping times can be proved. We will only do it for the finite horizon problem discussed in the previous section, however the same result would be true for other optimal control problems, including these on infinite horizon.

We define the set $\mathcal{V}_t$ in the following way. For every $a(\cdot) \in \mathcal{U}_t^\nu$ for some reference probability space $\nu = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W_Q)$, we choose an $\mathscr{F}_s^t$-stopping time $t \le \tau_{a(\cdot)} \le T$. The set $\mathcal{V}_t$ is the set of all such pairs $\big(a(\cdot), \tau_{a(\cdot)}\big)$. We also define $\tilde{\mathcal{V}}_t$ to be the set of these pairs in $\mathcal{V}_t$ for which the underlying reference probability space is standard, i.e. $\tilde{\mathcal{V}}_t = \big(a(\cdot), \tau_{a(\cdot)}\big)$, where $a(\cdot) \in \tilde{\mathcal{U}}_t^\nu$. To simplify notation we will just write $(a(\cdot), \tau)$ instead of $\big(a(\cdot), \tau_{a(\cdot)}\big)$.

THEOREM 3.70 *Let $b\colon [0,T] \times H \times \Lambda \to H$, $\sigma\colon [0,T] \times H \times \Lambda \to \mathcal{L}_2(\Xi_0, H)$, $l\colon [0,T] \times H \times \Lambda \to \mathbb{R}$ be continuous, let b and σ satisfy (3.7), (3.8), (3.9), and (3.10) with $\gamma = 1$. Let c be bounded from below, and let l, g, c satisfy (3.137), (3.154), (3.155). Then*

$$V(t,x) = \inf_{(a(\cdot),\tau)\in\mathcal{V}_t} \mathbb{E}\left[\int_t^\tau e^{-\int_t^s c(X(r))dr} l\left(s, X\left(s\right), a\left(s\right)\right) ds + e^{-\int_t^\tau c(X(r))dr} V\left(\tau, X\left(\tau\right)\right)\right]. \tag{3.176}$$

PROOF. Without loss of generality we always assume that the Q-Wiener processes in the reference probability spaces have everywhere continuous paths. We recall that, by Proposition 3.62 and Remark 3.64, J satisfies (3.156) and V satisfies

$$|V(t,x) - V(s,y)| \le \omega(|t-s| + |x-y|, R) \qquad \forall\, x, y \in B(0,R), t, s \in [0,T]$$

for some local modulus ω.

Let $a(\cdot) \in \tilde{\mathcal{U}}_t^\nu$ for some standard reference probability space $\nu = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W_Q)$ and let τ_n be an $\mathscr{F}_s^t$-stopping time which has finite number of values, i.e.

$$\tau_n = \sum_{i=1}^k \mathbf{1}_{A_i} t_i$$

for some pairwise disjoint sets $A_1, ..., A_k$ such that $\bigcup_{i=1}^n A_i = \Omega$ and $A_i \in \mathscr{F}_{t_i}^t, i = 1, ..., k$. It then follows from the proof of the first part of Theorem 2.24 (see (2.32)) that

$$\begin{aligned}
J(t,x;a\left(\cdot\right)) &= \sum_{i=1}^k \mathbb{E}\left[\mathbf{1}_{A_i} \int_t^{t_i} e^{-\int_t^s c(X(r))dr} l\left(s, X\left(s\right), a\left(s\right)\right) ds\right.\\
&\left.+\mathbf{1}_{A_i}\mathbb{E}\left[\int_{t_i}^T e^{-\int_t^s c(X(r))dr} l\left(s, X\left(s\right), a\left(s\right)\right) ds + e^{-\int_t^T c(X(r))dr} g\left(X\left(T\right)\right) | \mathscr{F}_{t_i}^t\right]\right]\\
&\ge \mathbb{E}\left[\sum_{i=1}^k \mathbf{1}_{A_i}\left[\int_t^{t_i} e^{-\int_t^s c(X(r))dr} l\left(s, X\left(s\right), a\left(s\right)\right) ds + e^{-\int_t^{t_i} c(X(r))dr} V\left(t_i, X\left(t_i\right)\right)\right]\right]\\
&= \mathbb{E}\left[\int_t^{\tau_n} e^{-\int_t^s c(X(r))dr} l\left(s, X\left(s\right), a\left(s\right)\right) ds + e^{-\int_t^{\tau_n} c(X(r))dr} V\left(\tau_n, X\left(\tau_n\right)\right)\right].
\end{aligned} \tag{3.177}$$

By Proposition 1.77, every stopping time can be approximated by stopping times τ_n with finite number of values. Therefore, thanks to (1.38), (3.137), (3.139), (3.154) and (3.155) we can apply the dominated convergence theorem to obtain in the limit that (3.177) is satisfied for every $(a(\cdot), \tau) \in \tilde{\mathcal{V}}_t$. Since $\tilde{\mathcal{V}}_t \subset \mathcal{V}_t$, it thus follows that

$$\begin{aligned}
V(t,x) &= \inf_{a(\cdot)\in\mathcal{U}_t} J(t,x;a\left(\cdot\right)) = \inf_{a(\cdot)\in\tilde{\mathcal{U}}_t} J(t,x;a\left(\cdot\right))\\
&\ge \inf_{(a(\cdot),\tau)\in\mathcal{V}_t} \mathbb{E}\left[\int_t^\tau e^{-\int_t^s c(X(r))dr} l\left(s, X\left(s\right), a\left(s\right)\right) ds + e^{-\int_t^\tau c(X(r))dr} V\left(\tau, X\left(\tau\right)\right)\right],
\end{aligned}$$

where we used Theorem 2.22 to obtain the second equality.

To show the reverse inequality, let $t \le \eta \le T$, $a(\cdot) \in \tilde{\mathcal{U}}_t^\nu$ for some standard reference probability space $\nu = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W_Q)$, and let $X(s) = X(s; t, x, a(\cdot))$.

Let, for $\omega_0 \in \Omega$, $\nu_{\omega_0} = \left(\Omega, \mathscr{F}_{\omega_0}, \mathscr{F}^t_{\omega_0,s}, \mathbb{P}_{\omega_0}, W_{Q,\eta}\right)$, where $W_{Q,\eta}(s) = W_Q(s) - W_Q(\eta)$, and $a^{\omega_0}(\cdot) = a_1(\cdot)$ (on $[\eta, T]$) be from Lemma 2.26-(ii).[2] We have $X(\cdot) = X(\cdot; \eta, X(\eta), a_1(\cdot))$, on $[\eta, T], \mathbb{P}$ a.e., and (see the argument in the proof of $(A3)$ in Proposition 2.16) that it is indistinguishable from a process, still denoted by $X(s)$, such that, for $\mathbb{P}$ a.e. ω_0,

$$X(\cdot) = X^{\nu_{\omega_0}}(\cdot; \eta, X(\eta)(\omega_0), a^{\omega_0}(\cdot)), \quad \text{on } [\eta, T], \mathbb{P}_{\omega_0}\ a.e..$$

Therefore, as a consequence of Theorem 2.24, applied to the reference probability space ν_{ω_0}, we have that, for $\mathbb{P}$ a.e. ω_0,

$$V(\eta, X(\eta)(\omega_0)) \le \mathbb{E}_{\omega_0}\left[\int_\eta^s e^{-\int_\eta^r c(X(\theta))d\theta} l\left(r, X\left(r\right), a^{\omega_0}(\cdot)\left(r\right)\right) dr \right.$$
$$\left. + e^{-\int_\eta^s c(X(\theta))d\theta} V\left(s, X\left(s\right)\right)\right]. \quad (3.178)$$

Set

$$M(s) = \int_t^s e^{-\int_t^r c(X(\theta))d\theta} l\left(r, X\left(r\right), a\left(r\right)\right) dr + e^{-\int_t^s c(X(\theta))d\theta} V\left(s, X\left(s\right)\right).$$

Then, by (3.178) and the definition of $M(s)$ (see also remarks at the end of Section 2.2.2), for $\mathbb{P}$ a.e. ω_0

$$\begin{aligned}
M(\eta)(\omega_0) \le & \left(\int_t^\eta e^{-\int_t^r c(X(\theta))d\theta} l\left(r, X\left(r\right), a\left(r\right)\right) dr\right)(\omega_0) \\
& + \left(e^{-\int_t^\eta c(X(\theta))d\theta}\right)(\omega_0)\, \mathbb{E}_{\omega_0}\left[\int_\eta^s e^{-\int_\eta^r c(X(\theta))d\theta} l\left(r, X\left(r\right), a^{\omega_0}\left(r\right)\right) dr \right. \\
& \left. + e^{-\int_\eta^s c(X(\theta))d\theta} V\left(s, X\left(s\right)\right)\right] \\
= & \left(\int_t^\eta e^{-\int_t^r c(X(\theta))d\theta} l\left(r, X\left(r\right), a\left(r\right)\right) dr\right)(\omega_0) \\
& + \left(e^{-\int_t^\eta c(X(\theta))d\theta}\right)(\omega_0)\, \mathbb{E}\left[\int_\eta^s e^{-\int_\eta^r c(X(\theta))d\theta} l\left(r, X\left(r\right), a\left(r\right)\right) dr \right. \\
& \left. + e^{-\int_\eta^s c(X(\theta))d\theta} V\left(s, X\left(s\right)\right) \middle| \mathscr{F}^t_\eta\right](\omega_0) = \mathbb{E}\left[M(s)|\mathscr{F}^t_\eta\right](\omega_0) \quad (3.179)
\end{aligned}$$

for every $s \in [\eta, T]$. Therefore, M is a submartingale, and thus, by the Optional Sampling Theorem (Theorem 1.79), if τ is an $\mathscr{F}^t_s$-stopping time,

$$\begin{aligned}
V(t,x) &= M(t) \le \mathbb{E}\left[M(\tau)|\mathscr{F}^t_t\right] \\
&= \mathbb{E}\left[\int_t^\tau e^{-\int_t^s c(X(r))dr} l\left(s, X\left(s\right), a\left(s\right)\right) ds + e^{-\int_t^\tau c(X(r))dr} V\left(\tau, X\left(\tau\right)\right) \middle| \mathscr{F}^t_t\right].
\end{aligned}$$

Taking the expectation above (or noticing that $\mathscr{F}^t_t$ is trivial), we thus obtain

$$V(t,x) \le \mathbb{E}\left[\int_t^\tau e^{-\int_t^s c(X(r))dr} l\left(s, X\left(s\right), a\left(s\right)\right) ds + e^{-\int_t^\tau c(X(r))dr} V\left(\tau, X\left(\tau\right)\right)\right] \quad (3.180)$$

for every $a(\cdot) \in \tilde{\mathcal{U}}^\nu_t$ for any standard reference probability space $\nu = (\Omega, \mathscr{F}, \mathscr{F}^t_s, \mathbb{P}, W_Q)$, and every $\mathscr{F}^t_s$-stopping time τ.

It remains to justify that (3.180) is true for every $(a(\cdot), \tau) \in \mathcal{V}_t$. We sketch the argument. Let $a(\cdot) \in \mathcal{U}^{\nu_1}_t$, where $\nu_1 = \left(\Omega_1, \mathscr{F}_1, \mathscr{F}^t_{1,s}, \mathbb{P}_1, W^1_Q\right)$, and let τ_n

[2] In Lemma 2.26-(ii) $a^{\omega_0}(\cdot)$ denotes the 6-uple $\left(\Omega, \mathscr{F}_{\omega_0}, \mathscr{F}^\eta_{\omega_0,s}, \mathbb{P}_{\omega_0}, W_\eta, a_1|_{[\eta,T]}(\cdot)\right)$ while here we use $a^{\omega_0}(\cdot)$ only to indicate the process.

be $\mathscr{F}^t_{1,s}$-stopping times with finite number of values approximating τ. Let $\nu_2 = \left(\Omega_2, \mathscr{F}_2, \mathscr{F}^t_{2,s}, \mathbb{P}_2, W^2_Q\right)$ be a standard reference probability space. We proceed as in the proof of Theorem 2.22. Let $a_1(\cdot), \tilde{a}_1(\cdot)$ be as in the proof of Theorem 2.22. We denote $X^{\nu_1}(\cdot) = X^{\nu_1}(\cdot; t, x, a(\cdot)) = X^{\nu_1}(\cdot; t, x, a_1(\cdot)), X^{\nu_2}(\cdot) = X^{\nu_2}(\cdot; t, x, \tilde{a}_1(\cdot))$. Since τ_n has finite number of values, we can assume that

$$\tau_n = \sum_{i=1}^k \mathbf{1}_{A_i} t_i$$

for some pairwise disjoint sets $A_1, ..., A_k$ such that $\bigcup_{i=1}^n A_i = \Omega$ and $A_i \in \mathscr{F}^{t,0}_{1,t_i}, i = 1, ..., k$. Let $B_1, ..., B_k \in \mathcal{B}(\mathbf{W})$ be such that $W^1_Q(\cdot \wedge t_i)^{-1}(B_i) = A_i, i = 1, ..., k$ (see Lemma 2.19). We set

$$\tilde{\tau}_n = \sum_{i=1}^k \mathbf{1}_{W^2_Q(\cdot\wedge t_i)^{-1}(B_i)} t_i.$$

Then $\tilde{\tau}_n$ is an $\mathscr{F}^t_{2,s}$-stopping time with finite number of values, and it follows that

$$\mathcal{L}_{\mathbb{P}_1}(\tau_n \wedge \cdot, X^{\nu_1}(\cdot), a(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(\tilde{\tau}_n \wedge \cdot, X^{\nu_2}(\cdot), \tilde{a}_1(\cdot)).$$

One can then conclude that

$$\begin{aligned} &\mathbb{E}^{\nu_1}\left[\int_t^{\tau_n} e^{-\int_t^s c(X^{\nu_1}(r))dr} l\left(s, X^{\nu_1}(s), a(s)\right) ds + e^{-\int_t^{\tau_n} c(X^{\nu_1}(r))dr} V\left(\tau_n, X^{\nu_1}(\tau_n)\right)\right] \\ &= \mathbb{E}^{\nu_2}\left[\int_t^{\tilde{\tau}_n} e^{-\int_t^s c(X^{\nu_2}(r))dr} l\left(s, X^{\nu_2}(s), \tilde{a}_1(s)\right) ds + e^{-\int_t^{\tilde{\tau}_n} c(X^{\nu_2}(r))dr} V\left(\tilde{\tau}_n, X^{\nu_2}(\tilde{\tau}_n)\right)\right]. \end{aligned} \tag{3.181}$$

Combining (3.180) with (3.181), we thus obtain that

$$V(t,x) \le \mathbb{E}^{\nu_1}\left[\int_t^{\tau_n} e^{-\int_t^s c(X^{\nu_1}(r))dr} l\left(s, X^{\nu_1}(s), a(s)\right) ds + e^{-\int_t^{\tau} c(X^{\nu_1}(r))dr} V\left(\tau_n, X^{\nu_1}(\tau_n)\right)\right]$$

It remains to let $n \to +\infty$ above to conclude that (3.180) holds for every pair $(a(\cdot), \tau) \in \mathcal{V}_t$. □

Remark 3.71 The proof of existence in Theorem 3.66 works almost exactly the same (and is in fact easier) if we use Theorem 3.70 and replace $t+\epsilon$ by $\min(t+\epsilon, \tau)$, where τ is the exit time of $X(\cdot)$ from some ball $B_{r_0}(x)$ for some $r_0 > 0$ (or from $B_{r_\epsilon}(x)$ for some $r_\epsilon > 0$). In this way one can always work with local maxima and minima in the definition of viscosity solution and avoid the requirements about global uniform continuity (and hence growth at infinity) of test functions and their derivatives. We do not pursue this here and leave the details of such a version of viscosity solution to the interested readers. ■

3.6.3. Infinite horizon problem. In this subsection we characterize the value function of the infinite horizon optimal control problem (3.132) as the unique solution of the associated HJB equation (3.133). We consider the following set of assumptions for $b : H \times \Lambda \to H, \sigma : H \times \Lambda \to \mathcal{L}_2(\Xi_0, H), l : H \times \Lambda \to \mathbb{R}$.

There exist constants $C, m \geq 0$, and a local modulus ω_l such that:

$$
\begin{aligned}
&|b(x,a) - b(y,a)| \leq C|x-y| && \forall x,y \in H,\ a \in \Lambda, && (3.182)\\
&\|\sigma(x,a) - \sigma(y,a)\|_{\mathcal{L}_2(\Xi_0;H)} \leq C|x-y| && \forall x,y \in H,\ a \in \Lambda, && (3.183)\\
&|b(x,a)| \leq C(1+|x|) && \forall x,y \in H,\ a \in \Lambda, && (3.184)\\
&\|\sigma(x,a)\|_{\mathcal{L}_2(\Xi_0;H)} \leq C(1+|x|) && \forall x,y \in H,\ a \in \Lambda, && (3.185)\\
&\langle b(x,a) - b(y,a), B(x-y)\rangle \leq C|x-y|_{-1}^2 && \forall\ x,y \in H,\ a \in \Lambda, && (3.186)\\
&\|\sigma(x,a) - \sigma(y,a)\|_{\mathcal{L}_2(\Xi_0;H)} \leq C|x-y|_{-1} && \forall\ x,y \in H,\ a \in \Lambda, && (3.187)\\
&|l(x,a) - l(y,a)| \leq \omega_l(|x-y|, R) && \forall x,y \in B(0,R),\ a \in \Lambda, && (3.188)\\
&|l(x,a) - l(y,a)| \leq \omega_l(|x-y|_{-1}, R) && \forall x,y \in B(0,R),\ a \in \Lambda, && (3.189)\\
&|l(x,a)| \leq C(1+|x|^m) && \forall x \in H,\ a \in \Lambda. && (3.190)
\end{aligned}
$$

Proposition 3.24 suggests that in order for the value function to be well defined we need λ to be sufficiently big. We thus impose the following hypothesis.

HYPOTHESIS 3.72 *If $m > 0$, the discount constant λ in the functional (3.131) satisfies $\lambda > \bar{\lambda}$, where $\bar{\lambda}$ is the constant from Proposition 3.24, where C is the constant from (3.184) and (3.185) and m is the constant appearing in (3.190). If $m = 0$, we have $\lambda > 0$.*

PROPOSITION 3.73 (Regularity of V under weak B-condition) *Suppose that (3.2) hold, that b and σ are continuous, and that b σ and l satisfy (3.182), (3.184), (3.185), (3.186), (3.187), (3.189) and (3.190). Assume that Hypotheses 2.28 and 3.72 hold. Then there exists a local modulus ω such that:*

(i) *The cost functional (3.131) satisfies*

$$
|J(x,a(\cdot)) - J(y,a(\cdot))| \leq \omega(|x-y|_{-1}, R), \tag{3.191}
$$

for all $x,y \in B(0,R)$, $a(\cdot) \in \mathcal{U}_0$.

(ii) *There exists a constant $\tilde{C}$ such that*

$$
|J(x;a(\cdot))|, |V(x)| \leq \tilde{C}(1+|x|^m) \tag{3.192}
$$

for all $x \in H, a(\cdot) \in \mathcal{U}_0$.

(iii) *The value function V defined in (3.132) satisfies*

$$
|V(x) - V(y)| \leq \omega(|x-y|_{-1}, R) \qquad \forall\ x,y \in B(0,R). \tag{3.193}
$$

PROOF. Part (i): Let $R > 0$, $x,y \in B(0,R)$, and $a(\cdot) \in \mathcal{U}_0$. Set $X(\cdot) := X(\cdot;0,x;a^\nu(\cdot)), Y(\cdot) := X(\cdot;0,y;a^\nu(\cdot))$.

Choose $\epsilon > 0$. Thanks to (3.34) and Hypothesis 3.72, there exists T_ϵ, also depending on C, m, λ, R but independent of $a(\cdot)$, such that

$$
\mathbb{E}\int_{T_\epsilon}^{\infty} e^{-\lambda r}|l(X(r),a(r)) - l(Y(r),a(r))|dr \leq \epsilon. \tag{3.194}
$$

We now proceed as in the proof of Proposition 3.61.

$$\begin{aligned}
&\int_0^{T_\epsilon} e^{-\lambda r} \int_{\{|X(r)|<\frac{1}{\epsilon} \text{ and } |Y(r)|<\frac{1}{\epsilon}\}} |l(X(r),a(r)) - l(Y(r),a(r))| d\mathbb{P}dr \\
&\quad + \int_0^{T_\epsilon} e^{-\lambda r} \int_{\{|X(r)|\geq\frac{1}{\epsilon} \text{ or } |Y(r)|\geq\frac{1}{\epsilon}\}} |l(X(r),a(r)) - l(Y(r),a(r))| d\mathbb{P}dr \\
&\quad \leq \int_0^{T_\epsilon} e^{-\lambda r} \int_{\{|X(r)|<\frac{1}{\epsilon} \text{ and } |Y(r)|<\frac{1}{\epsilon}\}} \omega_l(|X(r)-Y(r)|_{-1}, \frac{1}{\epsilon}) d\mathbb{P}dr \\
&\quad + \int_0^{T_\epsilon} e^{-\lambda r} \int_{\{|X(r)|\geq\frac{1}{\epsilon} \text{ or } |Y(r)|\geq\frac{1}{\epsilon}\}} C(2 + |X(r)|^m + |Y(r)|^m) d\mathbb{P}dr \\
&\quad = J_1 + J_2. \qquad (3.195)
\end{aligned}$$

Thanks to (1.38), arguing as in the proof of Proposition 3.61, we have

$$J_2 \leq \gamma_1(\epsilon, R) \tag{3.196}$$

for some local modulus γ_1, independent of $a(\cdot)$.

Let K_ϵ be such $\omega_l(s, \frac{1}{\epsilon}) \leq \epsilon + K_\epsilon s$. Using (3.24) we obtain

$$J_1 \leq \frac{\epsilon}{\lambda} + K_\epsilon \int_0^{T_\epsilon} e^{-\lambda r} \mathbb{E}|X(r) - Y(r)|_{-1} dr \leq \frac{\epsilon}{\lambda} + C_\epsilon |x-y|_{-1} \tag{3.197}$$

for some C_ϵ independent of $a(\cdot)$.

Therefore, (3.194), (3.195), (3.196) and (3.197), yield

$$|J(x,a(\cdot)) - J(y,a(\cdot))| \leq \epsilon + \frac{\epsilon}{\lambda} + C_\epsilon |x-y|_{-1} + \gamma_1(\epsilon, R), \tag{3.198}$$

and (3.191) follows by taking the infimum above over $\epsilon > 0$.

Estimate (3.192) follows from (3.34) and Hypothesis 3.72 and (3.193) is an obvious consequence of (3.191).

□

PROPOSITION 3.74 (Regularity of V under strong B-condition) *Let (3.3) hold, let b and σ be continuous, satisfy (3.182), (3.184), (3.185) and (3.187), and let l be continuous and satisfy (3.188), (3.190). Assume that Hypotheses 2.28 and 3.72 hold. Then there exist a local modulus ω and a constant $\tilde{C}$ such that (i)-(iii) of Proposition 3.73 are satisfied.*

PROOF. The proof is exactly the same as the proof of Proposition 3.73. We just have to replace the term $|X(r) - Y(r)|_{-1}$ by $|X(r) - Y(r)|$ in (3.195) and (3.197), and then use (3.29) instead of (3.24). □

THEOREM 3.75 (Existence under weak B-condition) *Let the assumptions of Proposition 3.73 be satisfied, and let, for every x,*

$$\lim_{N\to+\infty} \sup_{a\in\Lambda} \mathrm{Tr}\left[\left(\sigma(x,a)Q^{\frac{1}{2}}\right)\left(\sigma(x,a)Q^{\frac{1}{2}}\right)^* BQ_N\right] = 0. \tag{3.199}$$

Then the value function V defined in (3.132) is the unique viscosity solution of (3.133) among functions in the set

$$\begin{aligned}
S := \{u \colon H \to \mathbb{R} \ : \ &|u(x)| \leq C_1(1 + |x|^k) \\
&\textit{for some } C_1 \geq 0 \textit{ and } k \geq 0 \textit{ satisfying } (3.201)\}. \qquad (3.200)
\end{aligned}$$

$$\begin{cases} k < \frac{\lambda}{C+\frac{1}{2}C^2} \quad \textit{if} \quad \frac{\lambda}{C+\frac{1}{2}C^2} \leq 2, \\ Ck + \frac{1}{2}C^2 k(k-1) < \lambda \quad \textit{if} \quad \frac{\lambda}{C+\frac{1}{2}C^2} > 2, \\ k \textit{ can be any positive number if } C = 0. \end{cases} \tag{3.201}$$

(C is the constant from (3.184).)

PROOF. The proof follows the lines of the proof of Theorem 3.66. Proposition 3.73 and Hypothesis 3.72 guarantee that V is B-continuous and that it belongs to S. By Proposition 2.16, Theorem 2.31, and Proposition 3.73 (which ensures that Hypothesis 2.28 holds), the dynamic programming principle (2.42) is satisfied.

To show that V is a viscosity supersolution of (3.128), suppose that there exist a test function $\psi = \varphi + h$ and a point $x \in H$ such that $V + \psi$ has a local minimum at x. Without loss of generality we can assume that h, Dh, D^2h have at most polynomial growth at infinity, and that the minimum is global. We can also require that $V(x) + \psi(x) = 0$, so for all y we have $V(y) \geq -\psi(y)$.

For $\epsilon > 0$, by the dynamic programming principle, we can find $a^{\nu_\epsilon}(\cdot) \in \mathcal{U}_0$ such that, denoting $X(\cdot) := X(\cdot; 0, x; a^{\nu_\epsilon}(\cdot))$,

$$V(x) + \epsilon^2 \geq \mathbb{E}^{\nu_\epsilon} \left[\int_0^\epsilon e^{-\lambda s} l(X(s), a^{\nu_\epsilon}(s)) ds + e^{-\lambda \epsilon} V(X(\epsilon)) \right].$$

Therefore we have

$$\epsilon^2 - \varphi(x) - h(x) \geq \mathbb{E}^{\nu_\epsilon} \left[\int_0^\epsilon e^{-\lambda r} l(X(r), a^{\nu_\epsilon}(r)) dr - e^{-\lambda \epsilon} (\varphi(X(\epsilon)) + h(X(\epsilon))) \right],$$

which, upon using (3.161), (3.162), yields

$$\begin{aligned}\epsilon + \frac{1}{\epsilon} \mathbb{E}^{\nu_\epsilon} \Bigg[\int_0^\epsilon e^{-\lambda r} \Big(& - l(X(r), a^{\nu_\epsilon}(r)) + \psi_t(X(r)) + \langle b(X(r), a^{\nu_\epsilon}(r)), D\psi(X(r)) \rangle \\ & + \frac{1}{2} \mathrm{Tr} \left[\left(\sigma(X(r), a^{\nu_\epsilon}(r)) Q^{\frac{1}{2}} \right) \left(\sigma(X(r), a^{\nu_\epsilon}(r)) Q^{\frac{1}{2}} \right)^* D^2 \psi(X(r)) \right] \\ & \qquad - \lambda \psi(X(r)) + \langle X(r), A^* D\varphi(X(r)) \rangle \Big) dr \Bigg] \geq 0.\end{aligned}$$

Using exactly the same arguments as these in the proof of Theorem 3.66, it follows that there exists a modulus $\tilde{\rho}$, independent of the control $a^\nu(\cdot)$, such that:

$$\begin{aligned}& \frac{1}{\epsilon} \mathbb{E}^{\nu_\epsilon} \Bigg[\int_0^\epsilon \lambda V(x) - l(x, a^{\nu_\epsilon}(r)) + \langle b(x, a^{\nu_\epsilon}(r)), D\psi(x) \rangle \\ & + \frac{1}{2} \mathrm{Tr} \left[\left(\sigma(x, a^{\nu_\epsilon}(r)) Q^{\frac{1}{2}} \right) \left(\sigma(x, a^{\nu_\epsilon}(r)) Q^{\frac{1}{2}} \right)^* D^2 \psi(x) \right] + \langle x, A^* D\varphi(x) \rangle \, dr \Bigg] \geq -\tilde{\rho}(\epsilon).\end{aligned}$$

Therefore, taking the supremum over $a \in \Lambda$ inside the integral and then letting $\epsilon \to 0$ we obtain

$$\begin{aligned}\lambda V(x) + \langle x, A^* D\varphi(x) \rangle + \sup_{a \in \Lambda} \Big\{ & - l(x, a) + \langle b(x, a), D\psi(x) \rangle \\ & + \frac{1}{2} \mathrm{Tr} \left[\left(\sigma(x, a) Q^{\frac{1}{2}} \right) \left(\sigma(x, a) Q^{\frac{1}{2}} \right)^* D^2 \psi(x) \right] \Big\} \geq 0.\end{aligned}$$

This shows that V is a viscosity supersolution of (3.133).

To show that V is a viscosity subsolution we take a constant control, apply DPP, and again argue the same as in the proof of Theorem 3.66. We leave it to the readers.

To prove that V is the unique viscosity solution among functions in S we need to show that the hypotheses of Theorem 3.56 are satisfied. This was already done in the proof of Theorem 3.66 apart from Hypotheses 3.45 and 3.55 for $\gamma = 1$, and condition (3.106). Hypothesis 3.45 is obviously true with $\nu = \lambda$. As regards Hypothesis 3.55 for $\gamma = 1$, by (3.184) and (3.185), we obtain for for all $(x, r) \in$

$H \times \mathbb{R}, p, q \in H, X, Y \in S(H)$.

$$
\begin{aligned}
&|F(x, r, p+q, X+Y) - F(x, r, p, X)| \\
&\qquad \leq C(1+|x|)|q| + \frac{1}{2}C^2(1+|x|)^2\|Y\|,
\end{aligned}
$$

i.e. Hypothesis 3.55 holds with $M_F = C, N_F = \frac{1}{2}C^2$. Condition (3.106) thus follows from the definition of the set S. □

THEOREM 3.76 (Existence under strong B-condition) *Let the assumptions of Proposition 3.74 and (3.199) be satisfied. Then the value function V defined in (3.132) is the unique viscosity solution of (3.133) among functions in S defined in Theorem 3.75.*

PROOF. The only difference with respect to the proof for the weak case is that to show uniqueness we now use Theorem 3.54 instead of Theorem 3.58. The fact that Hypothesis 3.52 is satisfied was already noticed in the proof of Theorem 3.67. □

When conditions (3.184) and (3.185) are replaced by $\|\sigma(x,a)\|_{\mathcal{L}_2(\Xi_0;H)} \leq C(1+|x|^\gamma)$ and $|b(x,a)| \leq C(1+|x|^\gamma)$ for some $\gamma \in [0,1)$, conditions which must be imposed on a set of functions to guarantee that the value function is the unique viscosity solution among them, can be easily deduced from (3.104) and (3.105).

3.7. Existence of solutions: Finite dimensional approximations

We have shown in Section 3.6 that value functions of stochastic optimal control problems are viscosity solutions of their associated HJB equations. This gives a direct method of establishing existence of viscosity solutions for a large class of equations where we have an explicit representation formula for a solution. However many interesting equations cannot be linked to a stochastic optimal control problem. The best examples are Isaacs equations which are associated to zero-sum, two-player, stochastic differential games. For Isaacs equations, one way or showing existence of viscosity solutions is by proving directly that the associated (upper or lower) value of the game is the solution. Such results can be found in [192, 361, 363]. This method however runs into technical difficulties as the proof of the dynamic programming principle is very complicated. In this section we will present a more general method of showing existence of viscosity solutions based on finite dimensional approximations. This method can be thought of as a Galerkin type approximation for PDE in infinitely many variables. It was first introduced in [103] for first order equations and later generalized to second order equations in [421, 422]. We will present the proofs for the initial value problems .

Let A be a linear, densely defined, maximal dissipative operator in H. Let B be a bounded, strictly positive, self-adjoint, compact operator on H such that A^*B is bounded. For $N > 1$ let H_N be the finite dimensional space spanned by the eigenvectors of B corresponding to the eigenvalues which are greater than or equal to $1/N$. Let P_N, Q_N be defined as in Section 3.2. We notice that B commutes with P_N and Q_N, and P_N, Q_N are now also orthogonal projections in H.

We need to change slightly the structure conditions on the Hamiltonian F.

HYPOTHESIS 3.77 *There exists a modulus ω such that*

$$
\begin{aligned}
&F\left(t, x, \frac{B(x-y)}{\epsilon}, X\right) - F\left(t, y, \frac{B(x-y)}{\epsilon}, Y\right) \\
&\qquad \geq -\omega\left(|x-y|\left(1 + \frac{|x-y|_{-1}}{\epsilon}\right)\right)
\end{aligned}
$$

for all $(t,x,y) \in (0,T) \times H \times H, \epsilon > 0$, *and* $X, Y \in S(H)$, $X = P_N X P_N, Y = P_N Y P_N$ *for some* N *and such that (3.54) holds.*

For a bounded, strictly positive, self-adjoint, operator C on H we will use the notation

$$|x|_C := |C^{1/2}x|.$$

HYPOTHESIS 3.78 *Let* C *be a bounded, strictly positive, self-adjoint, operator on* H . *We say that Hypothesis 3.78-C is satisfied if there exists a modulus* ω_1 *such that*

$$F(t,x,c_1C(x-y),X) - F(t,y,c_1C(x-y),Y)$$
$$\geq -\omega_1 \left(|x-y|_C \left(1 + (c_1+c_2+c_3)|x-y|_C\right)\right)$$

for all $(t,x,y) \in (0,T) \times H \times H$, *and* $X, Y \in S(H)$, $X = P_N X P_N, Y = P_N Y P_N$ *for some* N *and such that*

$$-c_2 \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \leq \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq c_3 \begin{pmatrix} C & -C \\ -C & C \end{pmatrix}, \tag{3.202}$$

for some $c_1, c_2, c_3 \geq 0$.

HYPOTHESIS 3.79 *There exists* $h \in C^2(H)$, *radial, nondecreasing, nonnegative,* $h(x) \to \infty$ *as* $|x| \to \infty, Dh$, D^2h *are bounded, and*

$$F(t,x,p+\alpha Dh(x), X + \alpha D^2h(x)) \geq F(t,x,p,X) - \sigma(\alpha, \|p\| + \|X\|) \tag{3.203}$$

$\forall x, p, X$, $\forall \alpha \geq 0$, *where* σ *is a local modulus.*

Hypotheses 3.77, 3.78, and 3.79 will be sometimes applied to Hamiltonians F defined on finite dimensional spaces, i.e. when $F : (0,T) \times H_{N_0} \times H_{N_0} \times S(H_{N_0}) \to \mathbb{R}$ for some N_0. In such cases it will be understood that N in Hypotheses 3.77, 3.78 will always be equal to N_0 and that every $X \in S(H_{N_0})$ is naturally extended to an operator in $S(H)$ by taking $P_{N_0} X P_{N_0}$.

We first show continuity estimates for viscosity solutions of finite dimensional problems.

LEMMA 3.80 *Let* $\delta > 0, l > 0$, *and let* ω *be a modulus. Then there exist a nondecreasing, concave,* C^2 *function* φ_δ *on* $[0,+\infty)$ *such that* $\varphi_\delta(0) < \delta$ *and*

$$\omega(|\varphi_\delta''(r)|r^2 + \varphi_\delta'(r)r + r) \leq \varphi_\delta(r) \qquad \text{for } 0 \leq r \leq l. \tag{3.204}$$

PROOF. For $\epsilon \in (0,l]$, $r \geq 0$, $0 < \gamma \leq 1$, thanks to the subadditivity of ω, we have

$$\omega(r) \leq \omega(\epsilon) + \frac{\omega(\epsilon)}{\epsilon} r \leq \omega(\epsilon) + \frac{\omega(\epsilon)}{\epsilon}(1+l)^{1-\gamma} r^\gamma. \tag{3.205}$$

Let ϵ be such that $\omega(\epsilon) < \delta/2$. Define

$$g_\gamma(r) = \omega(\epsilon) + 2\frac{\omega(\epsilon)}{\epsilon}(1+l)^{1-\gamma} r^\gamma.$$

An elementary calculation and (3.205) give

$$g_\gamma(r) - \omega(|g_\gamma''(r)|r^2 + g_\gamma'(r)r + r) \geq \frac{\omega(\epsilon)}{\epsilon}(1+l)^{1-\gamma} r^\gamma \left(1 - 2\gamma(2-\gamma)\frac{\omega(\epsilon)}{\epsilon}(1+l)\right) \geq 0$$

if γ is small enough. We choose such γ_0 and set

$$\varphi_\delta(r) = g_{\gamma_0}(r + r_0),$$

where $0 < r_0 < 1$ is such that $g_{\gamma_0}(r_0) < \delta$. The function φ_δ has the required properties. $\square$

PROPOSITION 3.81 *Let C be a bounded, strictly positive, self-adjoint, operator on $\mathbb{R}^k$. Let $u \in USC([0,T) \times \mathbb{R}^k), v \in LSC([0,T) \times \mathbb{R}^k)$ be respectively a viscosity subsolution and a viscosity supersolution of*

$$\begin{cases} u_t + F(t,x,Du,D^2u) = 0 \quad \text{for } t \in (0,T), x \in \mathbb{R}^k, \\ u(0,x) = \psi(x) \quad \text{for } x \in \mathbb{R}^k, \end{cases} \tag{3.206}$$

where $F : (0,T)\times\mathbb{R}^k\times\mathbb{R}^k\times S(\mathbb{R}^k) \to \mathbb{R}$ is continuous, degenerate elliptic (Hypothesis 3.46) and satisfies Hypotheses 3.78 (with $H = H_N = \mathbb{R}^k$) and 3.79, and $\psi \in UC_b(\mathbb{R}^k)$.

(i) If $u\,, -v \le M$, then there is a modulus of continuity m, depending only on M, T, ω_1, and a modulus of continuity of ψ in the $|\cdot|_C$ norm, such that

$$u(t,x) - v(t,y) \le m(|x-y|_C) \tag{3.207}$$

for all $t \in [0,T)$ and $x, y \in \mathbb{R}^k$.

(ii) If

$$\sup_{x\in\mathbb{R}^k, t\in(0,T)} |F(t,x,0,0)| = K < +\infty, \tag{3.208}$$

then there exists a unique bounded viscosity solution u of (3.206). The norm $\|u\|_0$ only depends on $\|\psi\|_0$ and K.

PROOF. (i) Let m_1 be a modulus of continuity of ψ in the $|\cdot|_C$ norm. Given $\mu > 0$, set

$$u_1(t,x) = u(t,x) - \frac{\mu}{T-t} \tag{3.209}$$

$$v_1(t,x) = v(t,x) + \frac{\mu}{T-t}. \tag{3.210}$$

Let $\kappa = 3(T+1)(1+2\|C\|)$. Lemma 3.80 applied with the modulus $m_2(r) = m_1(r) + \kappa\omega(r) + (2M+1)r$ and $l = 2$ gives us for every $\delta > 0$ a function $\varphi_\delta \in C^2([0,\infty))$, nondecreasing, concave, such that

$$\varphi_\delta(0) < \delta, \quad \varphi_\delta(1) \ge 2M+1 \tag{3.211}$$

and

$$\varphi_\delta(r) - m_2(|\varphi_\delta''(r)|r^2 + \varphi_\delta'(r)r + r) \ge 0 \tag{3.212}$$

for $0 \le r \le 2$.
We are going to show that for every $\delta > 0$

$$u_1(t,x) - v_1(t,y) \le \varphi_\delta(|x-y|_C)(1+t)$$

for $t \in [0,T]$ and $\{|x-y|_C < 1\} = \Delta$. Let for $\gamma > 0$

$$\varphi(t,x,y) = \varphi_\delta(|x-y|_C^2 + \gamma)^{\frac{1}{2}}(1+t).$$

Suppose that

$$\sup_{(x,y)\in\Delta\,,\,t\in[0,T]} (u_1(t,x) - v_1(t,y) - \varphi(t,x,y)) > 0$$

(if not we are done). Then, for small $\alpha > 0$, using h from Hypothesis 3.79,

$$\sup_{(x,y)\in\Delta, t\in[0,T]} (u_1(t,x) - v_1(t,y) - \varphi(t,x,y) - \alpha h(x) - \alpha h(y)) > 0$$

and is attained at a point $(\bar{t}, \bar{x}, \bar{y})$. Moreover (3.211) and (3.212) imply that $(\bar{x}, \bar{y}) \in \Delta$ and $0 < \bar{t} < T$.

We compute

$$D_x\varphi(\bar{t},\bar{x},\bar{y}) = \varphi_\delta{}'\left((|\bar{x}-\bar{y}|_C^2 + \gamma)^{\frac{1}{2}}\right) \frac{C(\bar{x}-\bar{y})}{(|\bar{x}-\bar{y}|_C^2 + \gamma)^{\frac{1}{2}}}(\bar{t}+1), \tag{3.213}$$

$$\begin{aligned}
D^2_{xx}\varphi(\bar t,\bar x,\bar y) &= \varphi_\delta{}''((|\bar x-\bar y|_C^2+\gamma)^{\frac12})\frac{C(\bar x-\bar y)\otimes C(\bar x-\bar y)}{|\bar x-\bar y|_C^2+\gamma}(\bar t+1)\\
&+ \varphi_\delta{}'((|\bar x-\bar y|_C^2+\gamma)^{\frac12})\frac{C}{(|\bar x-\bar y|_C^2+\gamma)^{\frac12}}(\bar t+1) \qquad (3.214)\\
&- \varphi_\delta{}'((|\bar x-\bar y|_C^2+\gamma)^{\frac12})\frac{C(\bar x-\bar y)\otimes C(\bar x-\bar y)}{(|\bar x-\bar y|_C^2+\gamma)^{\frac32}}(\bar t+1).
\end{aligned}$$

We may rewrite (3.214) as $D^2_{xx}\varphi(\bar x,\bar y)=B_1+B_2+B_3$, where B_1,B_2,B_3 are the three terms appearing in (3.214). Since φ_δ is nondecreasing and concave, $B_2\ge 0$ and $B_1,B_3\le 0$. Using this notation we have

$$D^2\varphi(\bar t,\bar x,\bar y)=\begin{pmatrix} B_2 & -B_2\\ -B_2 & B_2\end{pmatrix}+\begin{pmatrix} B_1+B_3 & -B_1-B_3\\ -B_1-B_3 & B_1+B_3\end{pmatrix}. \qquad (3.215)$$

If we denote the two matrices in (3.215) by D_1 and $-D_2$ respectively, we obtain $D^2\varphi(\bar t,\bar x,\bar y)=D=D_1-D_2$, where $D_1,D_2\ge 0$.

Applying Theorem E.11 with $\epsilon=1/(\|D_1\|+\|D_2\|)$, there exist $b_1,b_2\in\mathbb R$ and matrices $X,Y\in S(\mathbb R^k)$ such that

$$\begin{aligned}
&\left(b_1,\varphi_\delta{}'\left((|\bar x-\bar y|_C^2+\gamma)^{\frac12}\right)\frac{C(\bar x-\bar y)}{(|\bar x-\bar y|_C^2+\gamma)^{\frac12}}(1+\bar t),X\right)\in\bar{\mathcal P}^{2,+}(u_1-\alpha h)(\bar t,\bar x),\\
&\left(b_2,\varphi_\delta{}'\left((|\bar x-\bar y|_C^2+\gamma)^{\frac12}\right)\frac{C(\bar x-\bar y)}{(|\bar x-\bar y|_C^2+\gamma)^{\frac12}}(1+\bar t),Y\right)\\
&\qquad\qquad\in\bar{\mathcal P}^{2,-}(v_1+\alpha h)(\bar t,\bar y), \qquad (3.216)
\end{aligned}$$

$$b_1-b_2=\varphi_\delta\left((|\bar x-\bar y|_C^2+\gamma)^{\frac12}\right), \qquad (3.217)$$

and

$$\begin{aligned}
-2(\|D_1\|+\|D_2\|)\begin{pmatrix} I & 0\\ 0 & I\end{pmatrix} &\le \begin{pmatrix} X & 0\\ 0 & -Y\end{pmatrix}\\
&\le D+\frac{1}{\|D_1\|+\|D_2\|}D^2\le 2D_1,
\end{aligned}$$

where in the last line we used $D^2\le(\|D_1\|+\|D_2\|)(D_1+D_2)$. Computing the norms, we thus have obtained

$$\begin{aligned}
&-2\|C\|(1+T)\left[2\left|\varphi_\delta{}''\left((|\bar x-\bar y|_C^2+\gamma)^{\frac12}\right)\right|+\frac{3\varphi_\delta{}'\left((|\bar x-\bar y|_C^2+\gamma)^{\frac12}\right)}{(|\bar x-\bar y|_C^2+\gamma)^{\frac12}}\right]\begin{pmatrix} I & 0\\ 0 & I\end{pmatrix}\\
&\le\begin{pmatrix} X & 0\\ 0 & -Y\end{pmatrix}\le\frac{2\varphi_\delta{}'\left((|\bar x-\bar y|_C^2+\gamma)^{\frac12}\right)}{(|\bar x-\bar y|_C^2+\gamma)^{\frac12}}(1+T)\begin{pmatrix} C & -C\\ -C & C\end{pmatrix} \qquad (3.218)
\end{aligned}$$

We set $\bar r=(|\hat x-\hat y|_C^2+\gamma)^{\frac12}$ and

$$d=\sup\{|\varphi_\delta''(r)|+\varphi_\delta'(r)\ :0\le r\le 2\}.$$

Using the equation, (3.216), (3.217), (3.218), and Hypothesis 3.79 we now have for $\gamma<1$ and small α

$$\begin{aligned}
&\varphi_\delta(\bar r)+\frac{2\mu}{T^2}\le F\left(\bar t,\bar y,\frac{(1+\bar t)\varphi_\delta{}'(\bar r)}{\bar r}C(\bar x-\bar y),Y\right)\\
&-F\left(\bar t,\bar x,\frac{(1+\bar t)\varphi_\delta{}'(\bar r)}{\bar r}C(\bar x-\bar y),X\right)+2\sigma\left(\alpha,\frac{6d(T+1)\|C\|}{\gamma^{\frac12}}+d\|C\|^{\frac12}(T+1)+1\right).
\end{aligned}$$

It thus follows from Hypothesis 3.78 that

$$\begin{aligned}
&\varphi_\delta(\bar r) + \frac{2\mu}{T^2} \\
&\le \omega_1 \left(|\hat x - \hat y|_C \left(1 + 3(T+1)(1+2\|C\|)\frac{\varphi'_\delta(\bar r)}{\bar r} + 4(T+1)\|C\||\varphi_\delta{}''(\bar r)| \right) |\hat x - \hat y|_C \right) \\
&\quad +2\sigma\left(\alpha, \frac{6d(T+1)\|C\|}{\gamma^{\frac12}} + d\|C\|^{\frac12}(T+1)+1\right)
\end{aligned}$$

for some local modulus σ_1. Thus, since ω_1 is concave, we get

$$\begin{aligned}
\varphi_\delta(\bar r) + \frac{2\mu}{T^2} &\le 3(T+1)(1+2\|C\|)\omega_1\left(|\varphi_\delta{}''(\bar r)|\bar r^2 + \varphi_\delta{}'(\bar r)\bar r + \bar r\right) \\
&+2\sigma\left(\alpha, \frac{6d(T+1)\|C\|}{\gamma^{\frac12}} + d\|C\|^{\frac12}(T+1)+1\right)
\end{aligned}$$

Therefore we obtain a contradiction if we let $\alpha \to 0$. This implies

$$u_1(t,x) - v_1(t,y) \le \varphi_\delta(|x-y|_C)(1+T) + 2M|x-y|_C$$

for all $x, y \in \mathbb{R}^k$ and $t \in [0,T)$. The claim now follows by letting $\mu \to 0$.

(ii) We remark that part (i) in particular guarantees that comparison principle holds for equation (3.206). It is standard to notice that under our assumptions one can construct a bounded viscosity subsolution $\underline{u}$ and a bounded viscosity supersolution $\overline{u}$ such that $\underline{u}(0,x) = \psi(x) = \overline{u}(0,x)$ and $\underline{u} \le \overline{u}$ (see Proposition 3.94 for a similar construction). We can thus use Perron's method (see Theorem E.12) to obtain a bounded viscosity solution which is unique by (i). □

The above existence and uniqueness result for finite dimensional HJB equations will be an important tool in constructing viscosity solutions of HJB equations in Hilbert spaces by finite dimensional approximations. We begin with the case when the strong B-condition for A is satisfied.

PROPOSITION 3.82 *Let B be compact and satisfy the strong B-condition for A as in Definition 3.10. Let u, v be respectively a viscosity subsolution and a viscosity supersolution of*

$$\begin{cases} u_t - \langle Ax, Du\rangle + F(t,x,Du,D^2u) = 0 \quad \text{for } t\in(0,T), x\in H, \\ u(0,x) = \psi(x) \quad \text{for } x \in H, \end{cases} \tag{3.219}$$

where $F : (0,T)\times H\times H\times S(H) \to \mathbb{R}$ satisfies (for $U = H$) Hypotheses 3.44, 3.46, 3.47, 3.77, and 3.79. Let $\psi \in UC_b(H)$. Let $u, -v \le M$ and be such that

$$|u(t,x) - u(t,y)| + |v(t,x) - v(t,y)| \le m(|x-y|) \tag{3.220}$$

for all $t \in [0,T)$ and $x, y \in H$, for some modulus m. Assume moreover that

$$\lim_{t\to 0} \tilde\rho(t) = 0, \tag{3.221}$$

where

$$\tilde\rho(t) = \sup_{x\in H}\left[(u(t,x) - \psi(e^{tA}x))_+ + (v(t,x) - \psi(e^{tA}x))_-\right].$$

Then for every $0 < \tau < T$ there exists a modulus m_τ, depending only on $\tau, m, \omega, \tilde\rho, T, M$, the constant c_0 in Definition 3.10 and the modulus of continuity of ψ, such that

$$u(t,x) - v(t,y) \le m_\tau(|x-y|) \quad \text{for all } x, y \in H, t \in [\tau, T). \tag{3.222}$$

PROOF. We will first show that there exists constants $C_\epsilon > 0$, depending only on ϵ, m, c_0, ω such that
$$\lim_{\epsilon\to 0} C_\epsilon = 0, \tag{3.223}$$
and for every $0<\tau<T$,
$$\sup_{t\in[\tau,T)} a_{\epsilon,C_\epsilon}(t) = a_{\epsilon,C_\epsilon}(\tau), \tag{3.224}$$
where
$$a_{\epsilon,C}(t) = \sup_{x,y\in H}\left\{u(t,x)-v(t,y)-\frac{|x-y|_{-1}^2}{2\epsilon}-Ct\right\}.$$
For $\mu>0,\alpha>0,\beta>0$ we consider the function
$$\Psi(t,s,x,y) = u(t,x)-\frac{\mu}{T-t}-v(s,y)-\frac{\mu}{T-s}-\frac{|x-y|_{-1}^2}{2\epsilon} - \alpha h(x)-\alpha h(y)-\frac{(t-s)^2}{2\beta}-Ct.$$
where h is the function from Hypothesis 3.79. Since B is compact, B-upper semi continuity is equivalent to weak sequential upper semicontinuity, so Ψ attains a maximum at some point $(\bar t,\bar s,\bar x,\bar y)$. Moreover as always we have
$$\lim_{\beta\to 0}\frac{(\bar t-\bar s)^2}{2\beta}=0 \quad \text{for fixed } \epsilon,\alpha. \tag{3.225}$$
Therefore, using weak sequential upper semicontinuity of the above function, it is easy to see that if $\sup_{t\in[\tau,T)} a_{\epsilon,C}(t) > a_{\epsilon,C}(\tau)$, then for small $\mu>0,\alpha>0,\beta>0$, we must have $\tau<\bar t,\bar s<T$.

We can now argue as in the proof of Theorem 3.50 (from (3.90) to (3.93)) to obtain that for $N>2$ there exist $X_N, Y_N\in S(H)$ satisfying (3.54) and such that
$$\frac{\bar t-\bar s}{\beta}+\frac{\mu}{(T-\bar t)^2}+C-\left\langle \bar x, A^*\left(\frac{B(\bar x-\bar y)}{\epsilon}\right)\right\rangle + F\left(\bar t,\bar x,\frac{B(\bar x-\bar y)}{\epsilon}+\alpha Dh(\bar x), X_N+\frac{2BQ_N}{\epsilon}+\alpha D^2h(\bar x)\right)\le 0 \tag{3.226}$$
and
$$\frac{\bar t-\bar s}{\beta}-\frac{\mu}{(T-\bar s)^2}+\left\langle \bar y, A^*\left(\frac{B(\bar x-\bar y)}{\epsilon}\right)\right\rangle + F\left(\bar s,\bar y,\frac{B(\bar x-\bar y)}{\epsilon}-\alpha Dh(\bar x), Y_N-\frac{2BQ_N}{\epsilon}-\alpha D^2h(\bar x)\right)\ge 0. \tag{3.227}$$
Since $u,-v$ are bounded from below it is obvious that
$$\frac{|B(\bar x-\bar y)|}{\epsilon}\le R_\epsilon \tag{3.228}$$
for some R_ϵ, possibly depending[3] on $u,-v$. Also, since $\Psi(\bar t,\bar s,\bar x,\bar x)+\Psi(\bar t,\bar s,\bar y,\bar y)\le 2\Psi(\bar t,\bar s,\bar x,\bar y)$, we get
$$\frac{|\bar x-\bar y|_{-1}^2}{2\epsilon}\le m(|\bar x-\bar y|). \tag{3.229}$$

[3]In fact, using uniform continuity of u, since for every $w\in H, |w|=1$ we have $\Psi(\bar t,\bar s,\bar x+w,\bar y)\le\Psi(\bar t,\bar s,\bar x,\bar y)$, we can obtain for $\alpha<1$
$$\frac{|B(\bar x-\bar y)|}{\epsilon}=\sup_{|w|=1}\frac{\langle B(\bar x-\bar y),w\rangle}{\epsilon}\le\frac{\langle Bw,w\rangle}{2\epsilon}+u(\bar t,\bar x)-u(\bar t,\bar x-w) + \alpha(h(\bar x-w)-h(\bar x))\le\frac{\|B\|}{2\epsilon}+m(1)+L,$$
where L is the Lipschitz constant of h.

Thus, subtracting (3.227) from (3.226), and using Hypotheses 3.44, 3.47, 3.77, 3.79, and (3.3), (3.225), (3.228), yields

$$C+\frac{2\mu}{T^2} \le c_0\frac{|\bar{x}-\bar{y}|_{-1}^2}{\epsilon}-\frac{|\bar{x}-\bar{y}|^2}{\epsilon}+\omega\left(|\bar{x}-\bar{y}|\left(1+\frac{|\bar{x}-\bar{y}|_{-1}}{\epsilon}\right)\right)+\sigma_1(\epsilon;\alpha,\beta,N), \tag{3.230}$$

where $\lim_{\alpha\to 0}\limsup_{\beta\to 0}\limsup_{N\to+\infty}\sigma_1(\epsilon;\alpha,\beta,N)=0$. Denote

$$\gamma(\epsilon)=\sup_{x,y\in H}\left\{c_0\frac{|\bar{x}-\bar{y}|_{-1}^2}{\epsilon}-\frac{|\bar{x}-\bar{y}|^2}{\epsilon}+\omega\left(|\bar{x}-\bar{y}|\left(1+\frac{|\bar{x}-\bar{y}|_{-1}}{\epsilon}\right)\right)\right\}.$$

using (3.229) we have

$$\gamma(\epsilon)\le\sup_{r\ge 0}\left\{2c_0m(r)-\frac{r^2}{\epsilon}+\omega\left(r\left(1+\frac{(2m(r))^{\frac{1}{2}}}{\epsilon^{\frac{1}{2}}}\right)\right)\right\}. \tag{3.231}$$

This expression can be estimated from above by

$$C_1\left(1+r+\frac{r}{\epsilon^{\frac{1}{2}}}+\frac{r^{\frac{3}{2}}}{\epsilon^{\frac{1}{2}}}\right)-\frac{r^2}{\epsilon}, \tag{3.232}$$

where C_1 only depends on ω, m and c_0. It is easily seen that (3.232) is positive only if $r\le C_2\epsilon^{\frac{1}{2}}$ for $\epsilon\le 1$, where C_2 only depends on C_1. But then it easily follows from (3.231) that $\lim_{\epsilon\to 0}\gamma(\epsilon)=0$. Thus, if $C=C_\epsilon:=2\gamma(\epsilon)$, we obtain a contradiction after we take $\limsup_{\alpha\to 0}\limsup_{\beta\to 0}\limsup_{N\to+\infty}$ in (3.230). Hence (3.224) must be true with this choice of C_ϵ.

A consequence of (3.224) is that for all $t\in[\tau,T), x,y\in H$

$$u(t,x)-v(t,y)-\frac{|x-y|_{-1}^2}{2\epsilon}-C_\epsilon t\le\sup_{x,y\in H}\left\{u(\tau,x)-v(\tau,y)-\frac{|x-y|_{-1}^2}{2\epsilon}\right\}. \tag{3.233}$$

Let m_ψ be the modulus of continuity of ψ. Then, by (3.20), (3.223) and (3.221), we obtain from (3.233)

$$\begin{aligned}u(t,x)-v(t,y)&\le\frac{|x-y|_{-1}^2}{2\epsilon}+C_\epsilon t+\sup_{x,y\in H}\left\{u(\tau,x)-v(\tau,y)-\frac{|x-y|_{-1}^2}{2\epsilon}\right\}\\&\le\frac{|x-y|_{-1}^2}{2\epsilon}+C_\epsilon T+2\tilde{\rho}(\tau)+\sup_{x,y\in H}\left\{|\psi(e^{\tau A}x)-\psi(e^{\tau A}y)|-\frac{|x-y|_{-1}^2}{2\epsilon}\right\}\\&\le\frac{|x-y|_{-1}^2}{2\epsilon}+C_\epsilon T+2\tilde{\rho}(\tau)+\sup_{x,y\in H}\left\{m_\psi\left(\frac{e^{c_0T}|x-y|_{-1}}{2\tau^{\frac{1}{2}}}\right)-\frac{|x-y|_{-1}^2}{2\epsilon}\right\}\\&\le\frac{|x-y|_{-1}^2}{2\epsilon}+\rho_\tau(\epsilon)\end{aligned} \tag{3.234}$$

where ρ_τ depends only on $C_\epsilon, T, \tau, m_\psi, \tilde{\rho}$, and $\lim_{\epsilon\to 0}\rho_\tau(\epsilon)=0$. Thus for every $\epsilon>0$

$$u(t,x)-v(t,y)\le\min\left\{2M|x-y|_{-1},\frac{|x-y|_{-1}^2}{2\epsilon}+\rho_\tau(\epsilon)\right\}$$

which implies (3.222). □

Proposition 3.82 in particular implies comparison principle for bounded and uniformly continuous viscosity sub- and super-solutions of (3.219). However we want to mention that comparison also holds without the requirement of uniform continuity of u and v with almost the same proof.

We will be using the following operators to approximate the operator A. For $N\ge 1$ we define

$$A_N=(P_NA^*P_N)^*.$$

The A_N are bounded, dissipative, operators in H and it is easy to see that

$$A_N P_N = A_N = P_N A_N, \tag{3.235}$$

and thus it follows that

$$e^{tA_N} P_N = P_N e^{tA_N}. \tag{3.236}$$

Moreover we have

$$-A_N^* B + c_0 B \geq P_N. \tag{3.237}$$

We alert the readers that in the lemma below we will use x_N to denote a sequence in H, not $P_N x$ as we have done in previous sections.

LEMMA 3.83 *Let B be a positive, self-adjoint, compact operator satisfying the strong B-condition as in Definition 3.10. Then:*

(i) *Let $x, x_N \in D(A^*), x_N \to x$, and $A^* x_N \to A^* x$. Then $A_N^* x_N \to A^* x$.*

(ii) *For every $x \in H, T > 0$*

$$e^{tA_N} x \to e^{tA} x \tag{3.238}$$

uniformly on $[0, T]$.

PROOF. (i) We know from Lemma 3.17(ii) that the operator $S = -A^* B + c_0 B$ is invertible, $S^{-1} \in \mathcal{L}(H)$, and $D(A^*) = D(B^{-1})$. We have

$$A^* = -SB^{-1} + c_0 I, \quad A_N^* = -P_N S B^{-1} P_N + c_0 P_N.$$

Since $B^{-1} = S^{-1}(-A^* + c_0 I)$ we thus obtain

$$B^{-1} x_N \to B^{-1} x.$$

Therefore

$$A^* x_N - A_N^* x = -Q_N S B^{-1} x_N - P_N S Q_N B^{-1} x_N + c_0 Q_N x_N \to 0$$

since P_N converges strongly to I and Q_N converges strongly to 0. This proves the claim.

(ii) We notice that $e^{tA_N^*}$ and e^{tA_N} are semigroups of contractions. Using (3.3) we have

$$\begin{aligned} &|e^{tA} x|_{-1}^2 + 2 \int_0^t \langle e^{sA} x, S e^{sA} x \rangle ds - 2c_0 \int_0^t |e^{sA} x|_{-1}^2 ds = |x|_{-1}^2, \\ &|e^{tA_N} x|_{-1}^2 + 2 \int_0^t \langle e^{sA_N} x, S_N e^{sA_N} x \rangle ds - 2c_0 \int_0^t |e^{sA_N} x|_{-1}^2 ds = |x|_{-1}^2, \end{aligned} \tag{3.239}$$

where

$$S_N = -A_N^* B + c_0 B = -P_N A^* P_N B + c_0 B = P_N(-A^* B + c_0 B) P_N + c_0 Q_N B.$$

By Trotter-Kato theorem (see Theorem B.46), for every $x \in H$, $e^{tA_N^*} x \to e^{tA^*} x$ uniformly on $[0, T]$. Thus, taking adjoints, it follows that

$$e^{tA_N} x \rightharpoonup e^{tA} x \qquad \text{for every } x \in H, t \geq 0. \tag{3.240}$$

Since B is compact, this implies

$$e^{tA_N} x \to e^{tA} x \qquad \text{in } H_{-1}. \tag{3.241}$$

Thus, passing to the limit as $N \to +\infty$ in (3.239) and using (3.241), we obtain

$$\int_0^t \langle e^{sA_N} x, S_N e^{sA_N} x \rangle ds \to \int_0^t \langle e^{sA} x, S e^{sA} x \rangle ds,$$

which, upon observing that $\|Q_N B\| \to 0$ and $P_N x \to x$, yields

$$\int_0^t \langle e^{sA_N} x, S e^{sA_N} x \rangle ds \to \int_0^t \langle e^{sA} x, S e^{sA} x \rangle ds. \tag{3.242}$$

Denote $y = e^{sA}x, y_N = e^{sA_N}x$. Then

$$0 \le |y - y_N|^2 \le \langle y - y_N, S(y - y_N)\rangle = \langle y, Sy\rangle - \langle y_N, Sy\rangle - \langle y, Sy_N\rangle + \langle y_N, Sy_N\rangle. \tag{3.243}$$

Using (3.240) it thus follows that

$$0 \le \liminf_{N\to+\infty} \langle y_N, Sy_N\rangle - \langle y, Sy\rangle.$$

This, together with Fatou's lemma and (3.242), implies that

$$\lim_{N\to+\infty} \langle e^{sA_N}x, Se^{sA_N}x\rangle = \langle e^{sA}x, Se^{sA}x\rangle$$

for a.e. s. We then get from (3.243) that

$$\lim_{N\to+\infty} |e^{sA}x - e^{sA_N}x|^2 = 0 \qquad \text{for a.e. } s. \tag{3.244}$$

The uniform convergence on $[0, T]$ follows from standard arguments. □

THEOREM 3.84 *Let B be compact and satisfy the strong B-condition for A as in Definition 3.10. Let $F : (0, T)\times H\times H\times S(H) \to \mathbb{R}$ satisfy (for $U = H$) Hypotheses 3.44, 3.46, Hypothesis 3.47 with $B = I$, and Hypotheses 3.77, 3.78-I, and 3.79. Let $\psi \in UC_b(H)$ and let for every $R > 0$*

$$F_R := \sup\{|F(t, x, p, X)| : t \in (0, T), x \in H, |p| + \|X\| \le R\} < +\infty. \tag{3.245}$$

Then there exists a unique bounded viscosity solution $u \in UC_b^x([0, T) \times H) \cap UC_b^x([\tau, T) \times H_{-1})$ for $0 < \tau < T$, of

$$\begin{cases} u_t - \langle Ax, Du\rangle + F(t, x, Du, D^2u) = 0 \quad \text{for } t \in (0, T), x \in H, \\ u(0, x) = \psi(x) \quad \text{for } x \in H, \end{cases} \tag{3.246}$$

satisfying

$$\lim_{t\to0} \sup_{x\in H} |u(t, x) - \psi(e^{tA}x)| = 0. \tag{3.247}$$

Moreover, there is a modulus ρ such that

$$|u(t, x) - u(s, e^{(t-s)A}x)| \le \rho(t - s) \quad \text{for all } 0 \le s \le t < T, x \in H. \tag{3.248}$$

PROOF. We consider two approximating equations.

$$\begin{cases} (u_N)_t - \langle A_N x, Du_N\rangle + F(t, P_N x, P_N Du_N, P_N D^2u_N P_N) = 0 \quad \text{in } (0, T) \times H \\ u_N(0, x) = \psi(P_N x) \quad \text{in } H, \end{cases} \tag{3.249}$$

and

$$\begin{cases} (v_N)_t - \langle A_N x, Dv_N\rangle + F(t, x, Dv_N, P_N D^2v_N P_N) = 0 \quad \text{in } (0, T) \times H_N \\ v_N(0, x) = \psi(x) \quad \text{in } H_N. \end{cases} \tag{3.250}$$

We notice that, since A_N is dissipative, the Hamiltonian $\tilde{F}_N : (0, T) \times H_N \times H_N \times S(H_N) \to \mathbb{R}$ defined by

$$\tilde{F}_N(t, x, p, X) = \langle A_N x, p\rangle + F(t, x, p, P_N X P_N)$$

satisfies all the assumptions of Proposition 3.81 with $C = I$, uniformly in N. Therefore, by Proposition 3.81, there is a unique bounded viscosity solution v_N of (3.250), $M \ge 0$, and a modulus m such that for all N

$$\begin{cases} \|v_N\|_0 \le M \\ |v_N(t, x) - v_N(t, y)| \le m(|x - y|) \quad \text{for all } t \in [0, T),\ x, y \in H_N \end{cases} \tag{3.251}$$

and so (3.251) is also satisfied by u_N on H.

We remark that the monotonicity of A_N guarantees that v_N is also a viscosity solution in the sense of Definition 3.34 on the finite dimensional Hilbert space H_N.

We now extend v_N to H by setting $u_N(x) = v_N(P_N x)$. We claim that u_N is a viscosity solution of (3.249). Again, since all the terms are bounded and A_N is monotone it is enough to show it in the classical sense of (the parabolic counterpart of) Definition 3.40. To prove that u_N is a viscosity subsolution, suppose that $u_N(t,x) - \varphi(t,x)$ has a maximum at $(\hat{t},\hat{x})$ for a smooth test function φ. Then $v_N(t,z) - \varphi(t, z + Q_N\hat{x})$ has a maximum at $(\hat{t}, P_N\hat{x})$ in $(0,T)\times H_N$. Therefore, using the fact that v_N is a subsolution of (3.250), we get

$$\varphi_t(\hat{t},\hat{x}) - \langle A_N P_N \hat{x}, P_N D\varphi(\hat{t},\hat{x})\rangle + F(\hat{t}, P_N\hat{x}, P_N D\varphi(\hat{t},\hat{x}), P_N D^2\varphi(\hat{t},\hat{x})P_N) \le 0$$

and the claim follows by (3.235). The supersolution case is done similarly.

We now show that there is a modulus ρ, depending only on m and the function F_R, such that

$$|v_N(t,x) - v_N(s, e^{-(t-s)A_N}x)| \le \rho(t-s) \tag{3.252}$$

for $x \in H_N, 0 \le s \le t < T$. Because of (3.251) it is enough to show (3.252) for $s = 0$ since the estimate can be reapplied at any later time. To do this we begin with $\psi \in C_b^{1,1}(H)$. We denote the Lipschitz constant of $D\psi$ by $L_{D\psi}$. We use the fact that $w(t,x) = \psi(e^{tA_N}x)$ is a classical (and viscosity) solution of

$$w_t - \langle A_N x, Dw\rangle = 0 \quad \text{in } (0,T)\times H_N, \quad u(0,x) = \psi(x) \text{ in } H_N,$$

which implies that

$$w + tF_{L_{D\psi} + \|D\psi\|_0}, \; w - tF_{L_{D\psi}+\|D\psi\|_0}$$

are respectively a viscosity super- and a subsolution of (3.250). Comparison then gives

$$|v_N(t,x) - \psi(e^{tA_N}x)| \le tF_{L_{D\psi}+\|D\psi\|_0} \tag{3.253}$$

For $\psi \in UC_b(H)$ we can approximate it by its inf-sup convolutions $\overline{\psi}_\epsilon \in C_b^{1,1}(H)$ (see Proposition D.21). This approximation is such that $c_\epsilon = \|\psi - \overline{\psi}_\epsilon\|_0$ and $K_\epsilon = L_{D\overline{\psi}_\epsilon} + \|D\overline{\psi}_\epsilon\|_0$ only depend on the modulus of continuity of ψ, and moreover $\lim_{\epsilon\to 0} c_\epsilon = 0$. Let v_N^ϵ be the viscosity solution of (3.250) with initial condition $\overline{\psi}_\epsilon$. It follows from comparison guaranteed by Proposition 3.81 that

$$\|v_N - v_N^\epsilon\|_0 \le \|\psi - \overline{\psi}_\epsilon\|_0 = c_\epsilon.$$

Using this and (3.253) we thus have

$$\begin{aligned} &|v_N(t,x) - \psi(e^{tA_N}x)| \le |v_N(t,x) - v_N^\epsilon(t,x)| + |v_N^\epsilon(t,x) - \overline{\psi}_\epsilon(e^{tA_N}x)| \\ &+ |\overline{\psi}_\epsilon(e^{tA_N}x) - \psi(e^{tA_N}x)| \le 2\|\psi - \overline{\psi}_\epsilon\|_0 + tF_{L_{D\overline{\psi}_\epsilon}+\|D\overline{\psi}_\epsilon\|_0} = 2c_\epsilon + tK_\epsilon. \end{aligned} \tag{3.254}$$

Therefore

$$|v_N(t,x) - \psi(e^{tA_N}x)| \le \rho(t) = \inf_{\epsilon>0}\{2c_\epsilon + tK_\epsilon\}$$

which completes the proof of (3.252). We also conclude, by (3.236), that for $0 \le s \le t < T, x \in H$,

$$\begin{aligned} |u_N(t,x) - u_N(s, e^{(t-s)A_N}x)| &= |v_N(t, P_N x) - v_N(s, P_N e^{(t-s)A_N}x)| \\ &= |v_N(t, P_N x) - v_N(s, e^{(t-s)A_N}P_N x)| \le \rho(t). \end{aligned} \tag{3.255}$$

We will now to show that for every $0 < \tau < T$ there exists a modulus m_τ, such that

$$|u_N(t,x) - u_N(t,y)| \le m_\tau(|x-y|_{-1}) \quad \text{for all } x,y \in H, t \in [\tau, T). \tag{3.256}$$

We notice that (3.237) implies that B restricted to H_N (i.e. $B_N = BP_N$) satisfies the strong condition for A_N on H_N with the same constant c_0. Therefore (3.256) follows from (3.251), (3.252) and Proposition 3.82 applied on spaces $H = H_N$, since all assumptions are independent of N. (In fact we do not need the full force of Proposition 3.82 since we deal with bounded equations on finite dimensional spaces.)

Since B also satisfies the weak B-condition for A_N with constant c_0, we notice that, by (3.19), for every N, $|e^{tA_N}x - x|_{-1} \le C(R)\sqrt{t}$ for $|x| \le R$, where $C(R)$ is independent of N. Thus for $0 < \tau \le s < t < T, R > 0$, using (3.255), (3.256), we obtain for $|x| < R$

$$|u_N(t,x) - u_N(s,x)| \le |u_N(t,x) - u_N(s, e^{(t-s)A_N}x)| + |u_N(s, e^{(t-s)A_N}x) - u_N(s,x)|$$
$$\le \rho(|t-s|) + m_\tau(|e^{(t-s)A_N}x - x| \le \rho(|t-s|) + m_\tau(C(R)\sqrt{|t-s|}) =: \rho_{\tau,R}(|t-s|).$$

Combining it with (3.256) we have

$$|u_N(t,x) - u_N(s,x)| \le m_\tau(|x-y|_{-1}) + \rho_{\tau,R}(|t-s|), \quad N \ge 1, \tau \le t, s < T, |x|, |y| \le R. \tag{3.257}$$

Therefore (extending u_N to $t = T$), the family $\{u_N\}$ is equicontinuous in the topology of $\mathbb{R} \times H_{-1}$ on sets $[\tau, T] \times \{|x| \le R\}$ for $\tau > 0$. But since B is compact such sets are compact in $\mathbb{R} \times H_{-1}$. Therefore, by the Arzela-Ascoli theorem there is a subsequence of u_N, still denoted by u_N, and a function u, such that $u_N \to u$ uniformly on bounded subsets of $[\tau, T] \times H$ for $\tau > 0$. Obviously u satisfies (3.248), (3.251), (3.256) and (3.257).The conclusion that u is a viscosity solution of (3.246) will follow from Theorem 3.41 (reformulated for the initial value problem), Lemma 3.83(i) and Lemma 3.85 below applied with $\tilde{F}(X) := F(t,x,p,X)$ for some fixed $(t,x,p) \in (0,T) \times H \times H$. Uniqueness is a consequence of Proposition 3.82. □

LEMMA 3.85 *If $\tilde{F} : S(H) \to \mathbb{R}$ is locally uniformly continuous and satisfies Hypothesis 3.46 and Hypothesis 3.47 with $B = I$, then for every $X \in S(H)$*

$$\tilde{F}(P_N X P_N) \to \tilde{F}(X) \quad \textit{as } N \to \infty.$$

PROOF. For every $\epsilon > 0$ we have

$$P_N(X - \epsilon X^2)P_N - \left(\|X\| + \frac{1}{\epsilon}\right) Q_N \le X \le P_N(X + \epsilon X^2)P_N + \left(\|X\| + \frac{1}{\epsilon}\right) Q_N.$$

Therefore, Hypotheses 3.46 and 3.47 imply

$$\tilde{F}(P_N(X + \epsilon X^2)P_N) - \sigma_1(N,\epsilon) \le \tilde{F}(X) \le \tilde{F}(P_N(X - \epsilon X^2)P_N) + \sigma_1(N,\epsilon),$$

where σ_1 is a local modulus. Using uniform continuity of $\tilde{F}$ we thus obtain

$$\tilde{F}(P_N X P_N) - \sigma_1(N,\epsilon) - \sigma_2(\epsilon) \le \tilde{F}(X) \le \tilde{F}(P_N X P_N) + \sigma_1(N,\epsilon) + \sigma_2(\epsilon),$$

for some modulus σ_2. Thus

$$|\tilde{F}(X) - \tilde{F}(P_N X P_N)| \le \{\sigma_1(N,\epsilon) + \sigma_2(\epsilon)\} \to \sigma_2(\epsilon) \quad \text{as } N \to +\infty.$$

and the claim follows thanks to the arbitrariness of ϵ. □

We now study the case when B satisfies the weak B-condition for A, i.e. when $-A^*B + c_0 B \ge 0$. In this case we do not have an analogue of Lemma 3.83 so we will have to add another layer of approximations of A. We will first replace A by its Yosida approximation A_λ and then approximate A_λ by $A_{\lambda,N} = P_N A_\lambda P_N$. The operators A_λ and $A_{\lambda,N}$ are bounded and dissipative. We first notice that B also satisfies a weak B-condition for A_λ and $A_{\lambda,N}$ with a different constant. Indeed, since for every $y \in D(A)$,

$$(1 - \lambda c_0)\langle By, y\rangle \le \langle B(I - \lambda A)y, y\rangle,$$

taking $y = (I - \lambda A)^{-1}x$, we get

$$(1 - \lambda c_0)|B^{\frac{1}{2}}(I - \lambda A)^{-1}x|^2 \le |B^{\frac{1}{2}}x||B^{\frac{1}{2}}(I - \lambda A)^{-1}x|,$$

which yields

$$|B^{\frac{1}{2}}(I-\lambda A)^{-1}x| \le \frac{|B^{\frac{1}{2}}x|}{1-\lambda c_0}.$$

It thus follows that for every $x \in H$,

$$\langle B(I-\lambda A)^{-1}x, x\rangle \le \frac{1}{1-\lambda c_0}\langle Bx, x\rangle.$$

Therefore,

$$-BA_\lambda + \frac{c_0}{1-\lambda c_0}B = \frac{1}{\lambda}\left(\frac{1}{1-\lambda c_0}B - B(I-\lambda A)^{-1}\right) \ge 0$$

and we conclude that

$$A_\lambda^* B + \frac{c_0}{1-\lambda c_0}B \ge 0. \tag{3.258}$$

Thus B satisfies the weak B-condition for A_λ with constant $2c_0$ for $\lambda < 1/(2c_0)$. Obviously (3.258) is also satisfied if A_λ is replaced by $A_{\lambda,N}$.

THEOREM 3.86 *Let B be compact and satisfy the weak B-condition for A as in Definition 3.9. Let $F : (0,T)\times H\times H\times S(H) \to \mathbb{R}$ satisfy (for $U = H$) Hypotheses 3.44, 3.46, Hypothesis 3.47 with $B = I$, and Hypotheses 3.78-B, and 3.79. Let $\psi \in UC_b(H_{-1})$ and let*

$$\sup\{|F(t,x,0,0)| : t \in (0,T), x \in H\} = K < +\infty. \tag{3.259}$$

Then there exists a unique bounded viscosity solution $u \in UC_b^x([0,T)\times H_{-1})]$ of (3.246). Moreover, for every $R > 0$, there is a modulus ρ_R such that

$$|u(t,x) - u(s,x)| \le \rho_R(|t-s|) \quad \textit{for all } 0 \le s,\ t < T,\ |x| \le R. \tag{3.260}$$

PROOF. We first solve, for $N > 2, 0 < \lambda < 1/(2c_0)$, the equations

$$\begin{cases} (v_{\lambda,N})_t + \langle A_{\lambda,N}x, Dv_{\lambda,N}\rangle + F(t,x,Dv_{\lambda,N}, P_N D^2 v_{\lambda,N} P_N) = 0 & \text{in } (0,T)\times H_N \\ v_{\lambda,N}(0,x) = \psi(x) \quad \text{in } H_N. \end{cases} \tag{3.261}$$

To do this we notice that, since $A_{\lambda,N}$ is dissipative and the weak B condition holds with constant $1/(2c_0)$, the Hamiltonian $\tilde{F}_{\lambda,N} : (0,T)\times H_N\times H_N\times S(H_N) \to \mathbb{R}$ defined by

$$\tilde{F}_{\lambda,N}(t,x,p,X) = \langle A_{\lambda,N}x, p\rangle + F(t,x,p,P_N X P_N)$$

satisfies all the assumptions of Proposition 3.81 with $C = B$, uniformly in N. (Again we identify $X \in S(H_N)$ with $P_N X P_N \in S(H)$.) In fact we have

$$\begin{aligned} &\tilde{F}_{\lambda,N}\left(t,x,c_1B(x-y),X\right) - \tilde{F}_{\lambda,N}\left(t,y,c_1B(x-y),Y\right) \\ &\qquad \ge -\omega_1\left(|x-y|_{-1}\left(1+(c_1+c_2+c_3)|x-y|_{-1}\right)\right) - \frac{c_1}{2c_0}|x-y|_{-1}^2 \end{aligned}$$

in Hypothesis 3.78-B now. Therefore, there exists a unique viscosity solution of $v_{\lambda,N}$ of (3.261), $M \ge 0$, and a modulus m such that for all λ, N

$$\begin{cases} \|v_{\lambda,N}\|_0 \le M \\ |v_{\lambda,N}(t,x) - v_{\lambda,N}(t,y)| \le m(|x-y|_{-1}) \quad \text{for all } t\in[0,T), x,y\in H_N \end{cases} \tag{3.262}$$

As in the proof of Theorem 3.84, the functions $u_{\lambda,N}(t,x) = v_{\lambda,N}(t,P_N x)$ are viscosity solutions of

$$\begin{cases} (u_{\lambda,N})_t - \langle A_{\lambda,N}x, Du_{\lambda,N}\rangle + F(t,P_N x,P_N Du_{\lambda,N}, P_N D^2 u_{\lambda,N} P_N) = 0 \\ u_{\lambda,N}(0,x) = \psi(P_N x) \quad \text{in } H, \end{cases} \tag{3.263}$$

and they also satisfy (3.262).

We will now show that for every $R > 0$ there exists a modulus ρ_R such that for all $N > 2$, $0 < \lambda < 1/(2c_0)$,

$$|v_{\lambda,N}(t,x) - v_{\lambda,N}(s,x)| \le \rho_R(|t-s|) \quad \text{for } 0 \le t, s < T, |x| \le R. \tag{3.264}$$

It is obvious that the function

$$\overline{w}(t,x) = Kt + M$$

is a classical and viscosity supersolution of (3.261) for every λ, N. For every $\epsilon > 0, x \in H$ there exists C_ϵ, depending only on m, such that

$$\psi(y) \le \psi(x) + \epsilon + C_\epsilon |x-y|_{-1}^2.$$

Set

$$R_\epsilon = \left(\frac{KT + 2M}{\epsilon}\right)^{\frac{1}{2}}.$$

We notice that

$$\psi_{x,\epsilon}(y) = \psi(x) + \epsilon + C_\epsilon|x-y|_{-1}^2 + \epsilon|y|^2 > KT + M \quad \text{for } |y| \ge R_\epsilon.$$

Now let $|x| \le R$. Since $\|A_{\lambda,N}^* B\| \le \|A^* B\|$, we have

$$|\langle A_{\lambda,N} y, 2C_\epsilon B(y-x)\rangle| \le 2C_\epsilon \|A^*B\| R_\epsilon (R_\epsilon + R) \quad \text{for } |y| \le R_\epsilon.$$

Thus if we set

$$\begin{aligned} F_{R,\epsilon} = \sup\{&|F(t,x,p,X| : t \in (0,T), |y| \le R_\epsilon, \\ &|p| + \|X\| \le 2[\epsilon(R_\epsilon + 1) + C_\epsilon \|B\|(R + R_\epsilon + 1)]\} + 2C_\epsilon \|A^*B\| R_\epsilon (R_\epsilon + R), \end{aligned}$$

it is easy to see that for every ϵ, N the function

$$\eta_{x,\epsilon}(t,y) = F_{R,\epsilon} t + \psi_{x,\epsilon}(y)$$

is a viscosity supersolution of (3.261) in $[0,T) \times \{|y| < R_\epsilon\}$. Therefore, for every ϵ, λ, N, the function

$$\overline{w}_{x,\epsilon} = \min\{\overline{w}, \eta_{x,\epsilon}\}$$

is a bounded viscosity supersolution of (3.261) in $[0,T) \times H_N$. By comparison we have $v_{\lambda,N} \le \overline{w}_{x,\epsilon}$. In particular

$$v_{\lambda,N}(t,x) - \psi(x) \le \overline{w}_{x,\epsilon}(t,x) - \psi(x) \le \epsilon + \epsilon R^2 + F_{R,\epsilon} t.$$

Taking infimum over $\epsilon > 0$ we obtain a modulus ρ_R such that

$$v_{\lambda,N}(t,x) - \psi(x) \le \rho_R(t) \quad \text{for } t \ge 0, |x| \le R.$$

Similar construction for subsolutions provides the same bound from below. Since the construction only depended on M and m in (3.262) it can be applied for any starting point $0 \le s < T$ which yields (3.264) which is obviously also true for $u_{\lambda,N}$.

We can finish as in the proof of Theorem 3.84. By (3.262) and (3.264), the family $u_{\lambda,N}$ is equibounded and equicontinuous in the topology of $\mathbb{R} \times H_{-1}$ on bounded sets of $[0,T] \times H$ and thus by the Arzela-Ascoli theorem, for every λ there is a subsequence of $u_{\lambda,N}$, still denoted by $u_{\lambda,N}$, and a function u_λ, such that $u_{\lambda,N} \to u_\lambda$ uniformly on bounded subsets of $[0,T] \times H$. Obviously u_λ satisfies (3.262) and (3.264). The fact that u_λ is a viscosity solution of (3.246) with A replaced by A_λ is standard and follows from Theorem 3.41 and Lemma 3.85 since all the terms are bounded. We then again us Arzela-Ascoli theorem to obtain that, up to a subsequence, u_λ converges uniformly on bounded subsets of $[0,T] \times H$ to a function u which satisfies (3.262) and (3.264). Using again Theorem 3.41, Lemma 3.85, and a well known analogue of Lemma 3.83(i) for Yosida approximations we finally conclude that u is a viscosity solution of (3.246).

The proof of uniqueness is similar to the proof of Proposition 3.82 and it will be omitted. Alternatively it can be deduced from the proof of Theorem 3.50 where we now have to first let $\delta \to 0$ and then $\epsilon \to 0$ there. $\square$

3.8. Singular perturbations

Passing to limits with viscosity solutions for equations in infinite dimensional spaces was discussed in Section 3.4. Despite its ease, some finite dimensional techniques cannot be applied due to the lack of local compactness, and we need to know a priori that solutions converge locally uniformly. When A is more coercive a version of the method of half-relaxed limits will be discussed in Section 3.9. In this section we look at a classical "vanishing viscosity" limit in which one tries to establish convergence of viscosity solutions of singularly perturbed equations. Such problems arise for instance in large deviation considerations and we will focus on equations having such origins.

Suppose we have a sequence of SDE

$$\begin{cases} dX_n(s) = (AX_n(s) + b(s, X_n(s)))ds + \frac{1}{\sqrt{n}}\sigma(s, X_n(s))dW_Q(s) \quad \text{for } s > t, \\ \\ X(t) = x \in H \end{cases} \tag{3.265}$$

in a real, separable Hilbert space H, where A is a linear, densely defined, maximal dissipative operator in H, $Q \in \mathcal{L}^+(H)$ and W_Q is a Q-Wiener process defined on some reference probability space. We want to investigate large deviation principle for the processes X_n. One of the key components in the study of large deviations is establishing the existence of the so called Laplace limit, i.e.

$$\lim_{n\to+\infty} \frac{1}{n} \log \mathbb{E}\left[e^{-ng(X_n(T))} : X_n(t) = x\right]$$

for a given continuous and bounded function g, where $T > t$. Defining

$$v_n(t, x) := -\frac{1}{n} \log \mathbb{E}\left[e^{-ng(X_n(T))}\right],$$

by formally applying Ito's formula the function v_n should be a viscosity solution of the second order equation

$$\begin{cases} (v_n)_t + \frac{1}{2n}\mathrm{Tr}\left((\sigma(t,x)Q^{1/2})(\sigma(t,x)Q^{1/2})^* D^2 v_n\right) - \frac{1}{2}|(\sigma(t,x)Q^{\frac{1}{2}})^* Dv_n|^2 \\ \qquad\qquad\qquad\qquad\qquad\qquad + \langle Ax + b(t,x), Dv_n\rangle = 0, \\ \\ v_n(T,x) = g(x) \qquad \text{in } (0,T)\times H. \end{cases} \tag{3.266}$$

Sending $n \to +\infty$ in (3.266) we obtain the limiting first order PDE

$$\begin{cases} v_t + \langle Ax + b(t,x), Dv\rangle - \frac{1}{2}|(\sigma(t,x)Q^{\frac{1}{2}})^* Dv|^2 = 0, \\ \\ v(T,x) = g(x) \qquad \text{in } (0,T)\times H. \end{cases} \tag{3.267}$$

This is the HJB equation associated to the deterministic optimal control problem characterized by the state equation

$$\begin{cases} \frac{d}{ds}X(s) = AX(s) + b(s, X(s)) + \sigma(s, X(s))Q^{\frac{1}{2}}z(s) \quad s > t, \\ \\ X(t) = x, \end{cases} \tag{3.268}$$

where we minimize the cost functional

$$J(t, x; z(\cdot)) = \int_t^T \frac{1}{2}|z(s)|^2 ds + g(X(T)) \tag{3.269}$$

over all controls $z(\cdot) \in L^2(t, T; H)$. The value function of the problem should be the unique viscosity solution of (3.267). Thus we can show the existence of the Laplace limit and identify it if we can prove that solutions v_n of the PDE (3.266) converge to the viscosity solution v of the limiting PDE (3.267). This is a classical singular perturbation limit problem which can be solved using the theory of viscosity solutions presented in this book. The details of the above program (which is based on a general PDE approach to large deviations developed in [184], see also [180, 181, 182]) and further study of this large deviation problem are in [425]. Here we will only show how the convergence of the v_n can be established using the techniques from the proof of comparison principle. We also point out that equations (3.266) and (3.267) have a quadratic gradient term which makes them more difficult. In particular they do not satisfy the assumptions of Section 3.5.

Let $T > 0$. Let B be an operator satisfying the weak B-condition (3.2) for A. We make the following assumptions.

HYPOTHESIS 3.87 *The functions* $b : [0, T] \times H \to H$, $\sigma : [0, T] \times H \to \mathcal{L}_2(\Xi_0; H)$ *are uniformly continuous on bounded sets and there exist constants* L, M *such that*

$$|b(t, x) - b(t, y)| \leq L|x - y|, \quad t \in [0, T],\ x, y \in H, \tag{3.270}$$

$$\langle b(t, x) - b(t, y), B(x - y) \rangle \leq L|x - y|_{-1}^2, \quad t \in [0, T],\ x, y \in H, \tag{3.271}$$

$$\|\sigma(t, x) - \sigma(t, y)\|_{\mathcal{L}_2(\Xi_0; H)} \leq L|x - y|_{-1}, \quad t \in [0, T],\ x, y \in H, \tag{3.272}$$

$$\|\sigma(t, x)\|_{\mathcal{L}_2(\Xi_0; H)} \leq M, \quad t \in [0, T],\ x \in H. \tag{3.273}$$

The function $g : H \to \mathbb{R}$ *is bounded and*

$$|g(x) - g(y)| \leq L|x - y|_{-1}, \quad x, y \in H. \tag{3.274}$$

It was shown in [425] that the functions v_n are unique viscosity solutions of (3.266). The assumptions in [425] were slightly different from Hypothesis 3.87 and some additional restrictions on test functions were placed to deal with exponential moments, however the proof of existence follows the standard arguments and the test function restrictions can be circumvented by localization using Itô's formulas with stopping times and (for instance) Theorem 3.70. The uniqueness part is more difficult. Moreover it was shown in [425] that the value function v of the deterministic control problem satisfies

$$|v(t, x) - v(s, y)| \leq C_1|x - y|_{-1} + C_2(\max\{|x|, |y|\})|t - s|^{\frac{1}{2}}$$

for all $x, y \in H$ and $t, s \in [0, T]$.

The following theorem addresses the convergence problem. It is a general statement about a singular perturbation problem. The theorem could be stated for more general HJB equations, however the main difficulty here is the quadratic gradient term.

THEOREM 3.88 *Let Hypothesis 3.87 hold. Let* v_n *be a bounded viscosity solution of (3.266), and* v *be a bounded viscosity solution of (3.267) such that*

$$\lim_{t \to T} \{|v_n(t, x) - g(x)| + |v(t, x) - g(x)|\} = 0, \ \textit{uniformly on bounded sets} \tag{3.275}$$

and

$$|v(t, x) - v(t, y)| \leq L|x - y|_{-1}, \quad 0 \leq t \leq T, x, y \in H. \tag{3.276}$$

Then there exists a constant C *independent of* n *such that*

$$\|v_n - v\|_0 \leq \frac{C}{\sqrt{n}}. \tag{3.277}$$

PROOF. Set

$$u_n := v + \frac{C}{\sqrt{n}}(T - t + 1).$$

Then u_n is a viscosity solution of

$$(u_n)_t + \langle Ax + b(t,x), Du_n\rangle - \frac{1}{2}|(\sigma(t,x)Q^{\frac{1}{2}})^* Du_n|^2 = -\frac{C}{\sqrt{n}}. \tag{3.278}$$

We will show that there exists C independent of n such that $v_n \le u_n$. If $v_n \not\le u_n$, then for $\mu, \delta, \beta > 0, m \in \mathbb{N}$, there exist $p_m, q_m \in H, a_m, b_m \in \mathbb{R}$ such that $|p_m|, |q_m|, |a_m|, |b_m| \le 1/m$, and

$$\begin{aligned}\Psi(t,s,x,y) &:= v_n(t,x) - u_n(s,y) - \frac{\mu}{t} - \frac{\mu}{s} - \frac{\sqrt{n}}{2}|x-y|_{-1}^2 - \delta(|x|^2 + |y|^2)\\ &-\frac{(t-s)^2}{2\beta} + \langle Bp_m, x\rangle + \langle Bq_m, y\rangle + a_m t + b_m s\end{aligned} \tag{3.279}$$

has a global maximum over $(0,T] \times H \times (0,T] \times H$ at some points $\bar{t}, \bar{s}, \bar{x}, \bar{y}$, where $\Psi(\bar{t}, \bar{s}, \bar{x}, \bar{y}) \ge \eta_n > 0$ for small μ, δ and large m. Similarly to the proof of Theorem 3.50 we have

$$\limsup_{\beta\to 0}\limsup_{m\to\infty} \frac{(\bar{t}-\bar{s})^2}{2\beta} = 0 \quad \text{for fixed } \mu, \epsilon, \delta, \tag{3.280}$$

$$\limsup_{\delta\to 0}\limsup_{\beta\to 0}\limsup_{m\to\infty} \delta(|\bar{x}|^2 + |\bar{y}|^2) = 0 \text{ for fixed } \mu. \tag{3.281}$$

Since $\Psi(\bar{t}, \bar{s}, \bar{x}, \bar{x}) \le \Psi(\bar{t}, \bar{s}, \bar{x}, \bar{y})$, it follows from (3.276)

$$\begin{aligned}\frac{\sqrt{n}}{2}|\bar{x}-\bar{y}|_{-1}^2 &\le u_n(\bar{s}, \bar{x}) - u_n(\bar{s}, \bar{y}) + \delta|x|^2 + \langle Bq_m, \bar{y} - \bar{x}\rangle\\ &\le \left(L + \frac{\|B^{1/2}\|}{m}\right)|\bar{x}-\bar{y}|_{-1} + \delta|x|^2.\end{aligned}$$

Therefore

$$\limsup_{\delta\to 0}\limsup_{\beta\to 0}\limsup_{m\to\infty} |\bar{x}-\bar{y}|_{-1} \le \frac{2L}{\sqrt{n}}. \tag{3.282}$$

If either $\bar{s}$ or $\bar{t}$ is equal to T, we thus obtain from (3.275), (3.276), (3.280), (3.281) and (3.282)

$$\begin{aligned}\eta_n &\le \limsup_{\delta\to 0}\limsup_{\beta\to 0}\limsup_{m\to\infty} \Psi(\bar{t}, \bar{s}, \bar{x}, \bar{y})\\ &\le \limsup_{\delta\to 0}\limsup_{\beta\to 0}\limsup_{m\to\infty}\left(L|\bar{x}-\bar{y}|_{-1} - \frac{C}{\sqrt{n}}\right) \le \frac{2L^2 - C}{\sqrt{n}}.\end{aligned}$$

Thus if $C \ge 2L^2$ we must have $0 < \bar{t}, \bar{s} < T$.

We now use that v_n is a viscosity subsolution of (3.266) to obtain

$$\begin{aligned}-\frac{\mu}{\bar{t}^2} - a_m + \frac{\bar{t}-\bar{s}}{\beta} + \frac{1}{2n}\mathrm{Tr}\left((\sigma(\bar{t},\bar{x})Q^{1/2})(\sigma(\bar{t},\bar{x})Q^{1/2})^*(\sqrt{n}B + 2\delta I)\right)\\ - \frac{1}{2}|(\sigma(\bar{t},\bar{x})Q^{\frac{1}{2}})^*(\sqrt{n}B(\bar{x}-\bar{y}) + 2\delta\bar{x} - Bp_m)|^2\\ + \langle \bar{x}, A^*[\sqrt{n}B(\bar{x}-\bar{y}) - Bp_m]\rangle + \langle b(\bar{t},\bar{x}), \sqrt{n}B(\bar{x}-\bar{y}) + 2\delta\bar{x} - Bp_m\rangle \ge 0.\end{aligned} \tag{3.283}$$

Moreover, since u_n is a viscosity supersolution of (3.278), we get

$$\frac{\mu}{\bar{s}^2} + b_m + \frac{\bar{t}-\bar{s}}{\beta} - \frac{1}{2}|(\sigma(\bar{s},\bar{y})Q^{\frac{1}{2}})^*(\sqrt{n}B(\bar{x}-\bar{y}) - 2\delta\bar{y} + Bq_m)|^2$$
$$+ \langle \bar{y}, A^*[\sqrt{n}B(\bar{x}-\bar{y}) + Bq_m]\rangle + \langle b(\bar{s},\bar{y}), \sqrt{n}B(\bar{x}-\bar{y}) - 2\delta\bar{y} + Bq_m\rangle$$
$$\leq -\frac{C}{\sqrt{n}}. \quad (3.284)$$

Subtracting (3.284) from (3.283) and using (3.2), (3.270), (3.271), (3.272), (3.273), (3.280), (3.281) and (3.282) give us

$$2\frac{\mu}{T^2} \leq \frac{n}{2}|(\sigma(\bar{t},\bar{y})Q^{\frac{1}{2}})^*B(\bar{x}-\bar{y})|^2 - \frac{n}{2}|(\sigma(\bar{t},\bar{x})Q^{\frac{1}{2}})^*B(\bar{x}-\bar{y})|^2$$
$$+\frac{1}{2\sqrt{n}}M^2\|B\| + c_1\sqrt{n}|\bar{x}-\bar{y}|_{-1}^2 - \frac{C}{\sqrt{n}} + \gamma(\delta,\beta,m), \quad (3.285)$$

where c_1 is some constant depending only on L and c_0 in (3.2), and γ is a function such that $\limsup_{\delta\to 0}\limsup_{\beta\to 0}\limsup_{m\to\infty}\gamma(\delta,\beta,m) = 0$. Now

$$|(\sigma(\bar{t},\bar{y})Q^{\frac{1}{2}})^*B(\bar{x}-\bar{y})|^2 - |(\sigma(\bar{t},\bar{x})Q^{\frac{1}{2}})^*B(\bar{x}-\bar{y})|^2$$
$$= \mathrm{Tr}\left((\sigma(\bar{t},\bar{y})Q^{1/2})(\sigma(\bar{t},\bar{y})Q^{1/2})^*B(\bar{x}-\bar{y})\otimes B(\bar{x}-\bar{y})\right)$$
$$- \mathrm{Tr}\left((\sigma(\bar{t},\bar{x})Q^{1/2})(\sigma(\bar{t},\bar{x})Q^{1/2})^*B(\bar{x}-\bar{y})\otimes B(\bar{x}-\bar{y})\right)$$
$$\leq c_2|\bar{x}-\bar{y}|_{-1}^3,$$

where c_2 is some constant depending only on $L, M, \|B^{1/2}\|$. Plugging this inequality into (3.285) and invoking (3.282) we thus obtain

$$2\frac{\mu}{T^2} \leq \limsup_{\delta\to 0}\limsup_{\beta\to 0}\limsup_{m\to\infty}\left(c_1\sqrt{n}|\bar{x}-\bar{y}|_{-1}^2 + c_2 n|\bar{x}-\bar{y}|_{-1}^3\right) + \frac{1}{2\sqrt{n}}M^2\|B\| - \frac{C}{\sqrt{n}}$$
$$\leq \frac{1}{\sqrt{n}}\left(4L^2c_1 + 8L^3c_2 + \frac{1}{2}M^2\|B\|\right) - \frac{C}{\sqrt{n}}.$$

This yields a contradiction if $C \geq 4L^2c_1 + 8L^3c_2 + \frac{1}{2}M^2\|B\|$. Thus we must have

$$v_n \leq v + \frac{C}{\sqrt{n}}(T-t+1) \leq v + \frac{C(T+1)}{\sqrt{n}}.$$

Similar arguments give us

$$v - \frac{C}{\sqrt{n}}(T-t+1) \leq v_n$$

and thus the result follows. □

The rate of convergence provided by Theorem 3.88 is the same as the rate for finite dimensional problems.

Remark 3.89 It is obvious from the proof that Theorem 3.88 remains the same if the term $\langle b(t,x), Du\rangle$ in (3.266) and (3.267) is replaced by a general Hamiltonian $F(t,x,Du)$, where $F : [0,T]\times H\times H \to \mathbb{R}$ is uniformly continuous on bounded sets and for instance satisfies

$$|F(t,x,p) - F(t,x,q)| \leq C|p-q|(1+|x|),$$
$$F\left(t,x,\frac{B(x-y)}{\epsilon}\right) - F\left(t,y,\frac{B(x-y)}{\epsilon}\right) \leq C|x-y|_{-1}^2,$$

for all $t\in[0,T], x,y,p,q\in H$. Such equations arise in risk sensitive optimal control problems. We refer to [84, 85, 276, 359, 360, 362, 363, 424] for such problems in infinite dimensional spaces and to [145, 195] for more on risk sensitive control problems. A result similar to Theorem 3.88 has been proved in [363] for a risk sensitive control problem using representation formulas and probabilistic methods. ■

3.9. Perron's method and half-relaxed limits

Perron's method is one of the main techniques for producing viscosity solutions of PDE in finite dimensional spaces (see [101, 269] and Appendix E.4). It is based on the principle that the supremum of the family of all viscosity subsolutions which are less than or equal to a viscosity supersolution of an equation is a (possibly discontinuous) viscosity solution. Thus to construct a viscosity solution, all we need is to produce one subsolution u_0 and one supersolution v_0 that both satisfy the boundary and initial conditions and such that $u_0 \leq v_0$. If we have a comparison theorem, the viscosity solution produced by Perron's method can then be proved to be continuous. Perron's method has a rather trivial extension to infinite dimensional bounded equations (3.68), see [321]. Perron's method was also used to prove existence of viscosity solutions using Ishii's definitions of viscosity solutions [271, 272]. However it is not known if a version of Perron's method can be implemented for B-continuous viscosity solutions of (3.58) and (3.64), even if the equations are of first order. The reason for this is that B-continuous viscosity sub-/super-solutions are semi-continious in a weaker topology and this makes the problem difficult. However it was shown in [288] how to adapt Perron's method to B-continuous viscosity solutions of (3.58) and (3.64) when the operator A is more coercive. We only discuss the initial value problems

$$\begin{cases} u_t - \langle Ax, Du\rangle + F(t,x,u,Du,D^2u) = 0 \quad (t,x) \in (0,T)\times H \\ u(0,x) = g(x), \end{cases} \tag{3.286}$$

where H is a real, separable Hilbert space and A is a linear, densely defined, maximal dissipative operator in H. The presentation here is based on [288] and we refer to this paper for further results and more details.

In order to develop Perron's method we need to introduce a notion of a discontinuous viscosity solution. Let B be an operator satisfying (3.2). For a function u we will write $u^{*,-1}$ and $u_{*,-1}$ to denote the upper- and lower-semicontinuous envelopes of u in the $|\cdot|\times|\cdot|_{-1}$ norm, i.e.

$$u^{*,-1}(t,x) = \limsup\{u(s,y) : s\to t, |y-x|_{-1}\to 0\},$$

$$u_{*,-1}(t,x) = \liminf\{u(s,y) : s\to t, |y-x|_{-1}\to 0\}.$$

Observe that $u^{*,-1}$ is upper semicontinuous in the $|\cdot|\times|\cdot|_{-1}$ norm and thus, thanks to Lemma 3.6(ii), it is B-upper semicontinuous.

We assume that $F:(0,T)\times H\times\mathbb{R}\times H\times S(H)\to\mathbb{R}$ satisfies Hypotheses 3.44, 3.45 and 3.46. We also impose the following coercivity condition on A.

$$-\langle A^*x, x\rangle \geq \lambda |x|_1^2, \qquad \text{for } x\in D(A^*) \tag{3.287}$$

for some $\lambda > 0$.

The above implies in particular that $D(A^*)\subset H_1$. Assumption (3.287) is satisfied for instance for self-adjoint invertible operators A if $B=(-A)^{-1}$.

Definition 3.90 *A locally bounded function u is a discontinuous viscosity subsolution of (3.286) if $u(0,y)\leq g(y)$ on H, and whenever $(u-h)^{*,-1}-\varphi$ has a local maximum in the topology of $|\cdot|\times|\cdot|_{-1}$ at a point (t,x) for a test function $\psi(s,y)=\varphi(s,y)+h(s,|y|)$ from Definition 3.32 such that φ is B-continuous and*

$$u(s,y)-h(s,|y|)\to-\infty \quad \textit{as } |y|\to\infty \quad \textit{locally uniformly in } s \tag{3.288}$$

then

$$x \in H_1$$

and

$$\psi_t(t,x) + \lambda |x|_1^2 \frac{h_r(t,|x|)}{|x|} - \langle x, A^* D\varphi(t,x)\rangle$$
$$+ F(t,x,(u-h)^{*,-1}(t,x) + h(t,|x|), D\psi(t,x), D^2\psi(t,x)) \le 0,$$

where h_r is the partial derivative of h with respect to the second variable.
A locally bounded function u is a discontinuous viscosity supersolution of (3.286) if $u(0,y) \ge g(y)$ on H, and whenever $(u+h)_{*,-1} + \varphi$ has a local minimum in the topology of $|\cdot| \times |\cdot|_{-1}$ at a point (t,x) for a test function $\psi(s,y) = \varphi(s,y) + h(s,|y|)$ from Definition 3.32 such that φ is B-continuous and

$$u(s,y) + h(s,|y|) \to +\infty \quad \text{as } |y| \to \infty \quad \text{locally uniformly in } s \tag{3.289}$$

then

$$x \in H_1$$

and

$$-\psi_t(t,x) - \lambda |x|_1^2 \frac{h_r(|x|)}{|x|} + \langle x, A^* D\varphi(t,x)\rangle$$
$$+ F(t,x,(u+h)_{*,-1}(t,x) - h(t,|x|), -D\psi(t,x), -D^2\psi(t,x)) \ge 0.$$

A discontinuous viscosity solution of (3.286) is a function which is both a discontinuous viscosity subsolution and a discontinuous viscosity supersolution.

The maxima and minima in Definition 3.90 can be assumed to be global and strict in the $|\cdot| \times |\cdot|_{-1}$ norm. Compared to Definition 3.34, apart from discontinuity of sub/super-solutions, the main difference here is that we require that $x \in H_1$ and the term $\langle x, A^* Dh(t,x)\rangle$ is not dropped entirely. We notice that if $x \in A^*$ then $-\langle x, A^* Dh(t,x)\rangle = -\frac{h_r(t,|x|)}{|x|}\langle x, A^* x\rangle \ge \lambda |x|_1^2 \frac{h_r(t,|x|)}{|x|}$ and this term is well defined and is left in the definition. If $x = 0$ the term $|x|_1^2 \frac{h_r(t,|x|)}{|x|}$ by definition is equal to 0. We also remark that if u is B-upper semicontinuous then $(u-h)^{*,-1} = u - h$ and if u is B-lower semicontinuous then $(u+h)_{*,-1} = u + h$. Definitions of discontinuous viscosity solutions have been first used in [271, 272]. Definitions requiring that points where maxima/minima occur belong to better spaces appeared in [71, 106] and have been successfully employed for some second order equations which are discussed in this book in Sections 3.11, 3.12, 3.13 (see also [242, 244, 245]).

For simplicity we restrict ourselves to F which does not depend on u. We will often say that u is a viscosity sub-/supersolution in an open set V. This will mean that we disregard the initial condition and the conditions of Definition 3.90 must be satisfied only if $(t,x) \in V$. However all functions involved must be defined on $(0,T) \times H$ and the maxima/minima in Definition 3.90 are local in the topology of $|\cdot| \times |\cdot|_{-1}$ in the whole $(0,T) \times H$.

LEMMA 3.91 *Let Hypotheses 3.44, 3.46 and condition (3.287) hold and let V be an open subset of $(0,T) \times H$. Let $\psi = \varphi + h$ be a test function such that $w = -\psi$ (respectively, $w = \psi$) satisfies*

$$w_t(t,x) - \langle x, A^* Dw(t,x)\rangle + F(t,x,Dw(t,x),D^2 w(t,x)) \le 0 \quad x \in D(A^*) \cap V, t \in (0,T)$$

(respectively,

$$w_t(t,x) - \langle x, A^* Dw(t,x)\rangle + F(t,x,Dw(t,x),D^2 w(t,x)) \ge 0 \quad x \in D(A^*) \cap V, t \in (0,T).)$$

Then the function w is a viscosity subsolution (respectively, supersolution) of (3.286).

PROOF. We will only show the proof in the subsolution case. We notice that $w = -\psi$ is B-upper semicontinuous. Suppose that $w(s,y) - \tilde{\varphi}(s,y) - \tilde{h}(s,|y|)$ has a local maximum at (t,x) for a test function $\tilde{\psi} = \tilde{\varphi} + \tilde{h}$. Then

$$w_t(t,x) = \tilde{\psi}_t(t,x), \quad -D\varphi(t,x) - \frac{h_r(t,|x|)}{|x|}x - D\tilde{\varphi}(t,x) - \frac{\tilde{h}_r(t,|x|)}{|x|}x = 0$$

and

$$D^2 w(t,x) \le D^2(\tilde{\varphi} + \tilde{h})(t,x).$$

Therefore either $x = 0$ or

$$\left(\frac{h_r(t,|x|)}{|x|} + \frac{\tilde{h}_r(t,|x|)}{|x|}\right) x = -D\varphi(t,x) - D\tilde{\varphi}(t,x) \in D(A^*),$$

i.e. $x \in D(A^*)$. Thus, using (3.287) and Hypothesis 3.46, we obtain

$$\begin{aligned}
&\tilde{\psi}_t(t,x) + \lambda |x|_1^2 \tfrac{\tilde{h}_r(t,|x|)}{|x|} - \langle x, A^* D\tilde{\varphi}(t,x)\rangle + F(t,x,D\psi(t,x),D^2\psi(t,x)) \\
&\le w_t(t,x) - \tfrac{\tilde{h}_r(t,|x|)}{|x|}\langle x, A^* x\rangle - \langle x, A^* D\tilde{\varphi}(t,x)\rangle + F(t,x,Dw(t,x),D^2 w(t,x)) \\
&= w_t(t,x) - \langle x, A^* Dw(t,x)\rangle + F(t,x,Dw(t,x),D^2 w(t,x)) \le 0
\end{aligned}$$

and the claim is proved. □

PROPOSITION 3.92 *Let Hypotheses 3.44, 3.46 and condition (3.287) be satisfied. Let $\mathcal{A}$ be a family of viscosity subsolutions of (3.286) in the sense of Definition 3.90. Suppose that the function*

$$u(x) = \sup\{w(x) : w \in \mathcal{A}\} \tag{3.290}$$

is locally bounded. Then u is a viscosity subsolution of (3.286) in the sense of Definition 3.90.

PROOF. Suppose that $(u-h)^{*,-1} - \varphi$ has a strict in $|\cdot| \times |\cdot|_{-1}$ norm global maximum at a point (t,x) for a test functions $\psi = \varphi + h$. (We can assume that $(u-h)^{*,-1}(s,y) - \varphi(s,y) \le -|y|$ as $|y| \to \infty$.) Perturbed optimization (see Corollary 3.26) and Definition 3.90 yield that there exist $w_n \in \mathcal{A}$, $x_n \in H_1$, t_n, and $a_n \in \mathbb{R}$, $p_n \in H$, $|a_n| + |p_n| \le 1/n$ such that

$$t_n \to t, \; B^{\frac{1}{2}} x_n \to B^{\frac{1}{2}} x, \; x_n \rightharpoonup x \text{ in } H \text{ as } n \to \infty, \tag{3.291}$$

$$(w_n - h)^{*,-1}(s,y) - \varphi(s,y) + \langle Bp_n, y\rangle + a_n s$$

has a strict in $|\cdot| \times |\cdot|_{-1}$ norm global maximum at (t_n, x_n), and

$$(w_n - h)^{*,-1}(t_n,x_n) \to (u-h)^{*,-1}(t,x) \text{ as } n \to \infty. \tag{3.292}$$

Therefore,

$$\begin{aligned}
\psi_t(t_n,x_n) - a_n + \lambda |x_n|_1^2 \frac{h_r(t_n,|x_n|)}{|x_n|} &- \langle x_n, A^*(D\varphi(x_n) - Bp_n)\rangle \\
&+ F(t_n, x_n, D\psi(t_n,x_n) - Bp_n, D^2\psi(t_n,x_n)) \le 0.
\end{aligned} \tag{3.293}$$

Since the x_n are bounded, using the local boundedness of F we thus obtain that either $x_n \to 0 = x$ or, up to a subsequence, $|x_n| > c > 0$ which leads to

$$|x_n|_1^2 \le C$$

for some constant C which, together with (3.291), implies that $x \in H_1$, and $B^{-\frac{1}{2}} x_n \rightharpoonup B^{-\frac{1}{2}} x$ as $n \to \infty$. Therefore, by (3.291),

$$|x_n - x|^2 = \langle B^{-\frac{1}{2}}(x_n - x), B^{\frac{1}{2}}(x_n - x)\rangle \to 0 \text{ as } n \to \infty,$$

i.e. $x_n \to x$ in H. Using this, the continuity of F, and the lower semicontinuouity of $|\cdot|_1$ in H, we can now pass to the lim inf as $n \to \infty$ in (3.293) to obtain

$$\psi_t(t,x) + \lambda |x|_1^2 \frac{h_r(t,|x|)}{|x|} - \langle x, A^* D\varphi(t,x)\rangle + F(t,x,D\psi(t,x),D^2\psi(t,x)) \le 0$$

which completes the proof. □

THEOREM 3.93 *Let Hypotheses 3.44, 3.46 and condition (3.287) be satisfied. Let u_0, v_0 be respectively a viscosity subsolution and a viscosity supersolution of (3.286) in the sense of Definition 3.90 such that $u_0 \le v_0$ and $u_0(0,x) = v_0(0,x) = g(x)$, $x \in H$. Then the function*

$$u(t,x) = \sup\{v(t,x) : u_0 \le v \le v_0, v \text{ is a viscosity subsolution of (3.286) in the sense of Definition 3.90}\} \tag{3.294}$$

is a viscosity solution of (3.286) in the sense of Definition 3.90.

PROOF. It follows from Proposition 3.92 that u is a viscosity subsolution. Suppose now that $(u+h)_{*,-1} + \varphi$ has a strict in $|\cdot| \times |\cdot|_{-1}$ norm global minimum at a point (t,x) for a test function $\psi = \varphi + h$ satisfying (3.289). First we notice that if

$$(u+h)_{*,-1}(t,x) = (v_0+h)_{*,-1}(t,x)$$

then $(v_0+h)_{*,-1} + \varphi$ has a global minimum at (t,x) and so we are done since v_0 is a viscosity supersolution. Therefore we only need to consider the case

$$(u+h)_{*,-1}(t,x) < (v_0+h)_{*,-1}(t,x).$$

It then follows from the above inequality, the B-continuity of φ and the weak sequential lower semi-continuity of $|\cdot|$ that there is $\epsilon_0 > 0$ such that for every $R > 0$ there exist $\eta_0 > 0$ such that

$$\begin{aligned}\epsilon + (u+h)_{*,-1}(t,x) + \varphi(t,x) - \varphi(s,y) - h(s,|y|)& \\ < (v_0+h)_{*,-1}(s,y) - h(s,|y|) \le v_0(s,y)&\end{aligned} \tag{3.295}$$

for $(s,y) \in (t-\eta_0, t+\eta_0) \times (B_{H_{-1}}(x,\eta_0) \cap B(x,R))$, $0 < \epsilon < \epsilon_0$. Denote

$$w(y) = \epsilon + (u+h)_{*,-1}(t,x) + \varphi(t,x) - \varphi(s,y) - h(s,|y|). \tag{3.296}$$

By further modifying h for large values of $|y|$ and $s \notin (t-\eta_0, t+\eta_0)$ if necessary, we can also assume that there is $R_0 > 0$ such that

$$w(s,y) \le u(s,y) - 1 \quad y \notin B(x,R_0),\ s \in (0,T). \tag{3.297}$$

Moreover, if R_0 is big enough there exist $s_n \to t, y_n \in B(x,R_0), y_n \to x$ in H_{-1} such that

$$u(s_n,y_n) + h(s_n,|y_n|) + \varphi(s_n,y_n) \to (u+h)_{*,-1}(t,x) + \varphi(t,x)$$

which means that for every $\eta > 0$ there exist points $(s,y) \in (t-\eta, t+\eta) \times (B_{H_{-1}}(x,\eta) \cap B(x,R_0))$ for which

$$u(s,y) < w(s,y). \tag{3.298}$$

If the condition for u being a viscosity supersolution of (3.286) is violated at (t,x) for the test function ψ then one of the following must hold:

(i) $x \notin H_1$.

(ii) $x \in H_1$ but

$$\begin{aligned}-\psi_t(t,x) - \lambda |x|_1^2 \frac{h_r(t,|x|)}{|x|} + \langle x, A^* D\varphi(t,x)\rangle& \\ + F(t,x,-D\psi(t,x),-D^2\psi(t,x)) < -\nu < 0&\end{aligned} \tag{3.299}$$

for some $\nu > 0$.

If (i) is satisfied then we must have

$$\liminf_{\substack{y \to x \text{ in } H_{-1} \\ y \in H_1}} |y|_1 = +\infty. \tag{3.300}$$

Otherwise we would have a sequence y_n such that $B^{\frac{1}{2}} y_n \to B^{\frac{1}{2}} x$ and $|B^{-\frac{1}{2}} y_n| \le C$. Then for some subsequence (still denoted by y_n) $B^{-\frac{1}{2}} y_n \rightharpoonup z$ for some $z \in H$ which would imply $x \in H_1$ and $z = B^{-\frac{1}{2}} x$. Using the local boundedness of F, condition (3.300) now implies that for every $R > 0$

$$w_t(s,y) - \lambda |y|_1^2 \frac{h_r(s,|y|)}{|y|} + \langle y, A^* D\varphi(s,y)\rangle + F(s,y,Dw(s,y),D^2 w(s,y)) < -\frac{\nu}{2}, \tag{3.301}$$

for $(s,y) \in (t-\eta_1, t+\eta_1) \times (B_{H_{-1}}(x,\eta_1) \cap B(x,R) \cap H_1)$ for some $\eta_1 > 0$.

Suppose that (ii) is true. We will show that for every $R > 0$ (3.301) holds for $(s,y) \in (t-\eta_1, t+\eta_1) \times (B_{H_{-1}}(x,\eta_1) \cap B(x,R) \cap H_1)$ for some $\eta_1 > 0$. If not there exist sequences $t_n \to t, x_n \to x$ in H_{-1}, $|x_n| \le R$ such that

$$\begin{aligned} w_t(t_n,x_n) - \lambda |x_n|_1^2 \frac{h_r(t_n,|x_n|)}{|x_n|} &+ \langle x_n, A^* D\varphi(t_n,x_n)\rangle \\ &+ F(t_n,x_n,Dw(t_n,x_n),D^2 w(t_n,x_n)) \ge -\frac{\nu}{2}. \end{aligned} \tag{3.302}$$

If $x_n \to 0 = x$ then letting $n \to +\infty$ in (3.302) would contradict (3.299). If $x_n \not\to 0$ then for some subsequence (still denoted by x_n) we would have $h_r(t_n,|x_n|)/|x_n| \ge \gamma > 0$ for some γ and this would imply $|x_n|_1 \le C$ for some constant C as otherwise (3.302) would be violated. But then we must have $B^{-\frac{1}{2}} x_n \rightharpoonup B^{-\frac{1}{2}} x$ in H and thus we obtain $x_n \to x$. However then (3.299), (3.302) and the lower semi-continuity of $\|\cdot\|_1$ again imply

$$\begin{aligned} -\frac{\nu}{2} \le \limsup_{n \to \infty} \Bigg(& w_t(t_n,x_n) - \lambda |x_n|_1^2 \frac{h_r(t_n,|x_n|)}{|x_n|} + \langle x_n, A^* D\varphi(t_n,x_n)\rangle \\ &+ F(t_n,x_n,Dw(t_n,x_n),D^2 w(t_n,x_n)) \Bigg) < -\nu \end{aligned}$$

which gives a contradiction.

Thus we have proved that in both cases (i) and (ii), for every $R > 0$ (3.301) holds for $(s,y) \in (t-\eta_1, t+\eta_1) \times (B_{H_{-1}}(x,\eta_1) \cap B(x,R) \cap H_1)$ for some $\eta_1 > 0$.

Recall now the definition of w given in (3.296). Since $(u+h)_{*,-1} + \varphi$ has a global minimum at (t,x), strict in $|\cdot| \times |\cdot|_{-1}$ norm, given $\eta > 0$ and ϵ small enough (depending on η) there exists a constant $\mu_\eta > 0$ such that

$$w(s,y) < (u+h)_{*,-1}(s,y) - h(s,|y|) - \mu_\eta \le u(s,y) - \mu_\eta \tag{3.303}$$

for $y \notin B_{H_{-1}}(x,\eta)$, $s \in (t-\eta, t+\eta)$.

Using (3.295), (3.297), (3.303), and (3.301) we can therefore conclude that there exist numbers $R, \eta, \epsilon, \mu > 0$ such that

$$w \le v_0 \quad \text{in } [0,T) \times H, \tag{3.304}$$

$$w(s,y) < u(s,y) - \mu \quad \text{for } (s,y) \notin (t-\eta, t+\eta) \times (B_{H_{-1}}(x,\eta) \cap B(x,R)), \tag{3.305}$$

and such that (3.301) is satisfied for $(s,y) \in (t-2\eta, t+2\eta) \times (B_{H_{-1}}(x,2\eta) \cap B(x,2R) \cap H_1)$.

We now claim that the function w is a viscosity subsolution of (3.286) in the interior of $(t-2\eta,t+2\eta)\times(B_{-1}(x,2\eta)\cap B(x,2R))$. This follows from Lemma 3.91 upon noticing that by (3.301) and (3.287) we have

$$\begin{aligned}&w_t(s,y)-\langle y,A^*Dw(s,y)\rangle+F(s,y,Dw(s,y),D^2w(s,y))\\&\le w_t(s,y)-\lambda|y|_1^2\tfrac{h_r(s,|y|)}{|y|}+\langle y,A^*D\varphi(s,y)\rangle+F(s,y,Dw(s,y),D^2w(s,y))<0\end{aligned}$$

for $(s,y)\in(t-2\eta,t+2\eta)\times(B_{H_{-1}}(x,2\eta)\cap B(x,2R)\cap D(A^*))$.

It remains to show that the function

$$u_1=\max(w,u) \tag{3.306}$$

is a viscosity subsolution in the sense of Definition 3.90. It follows from the definition that u_1 is a viscosity subsolution in the interior of $(t-2\eta,t+2\eta)\times(B_{-1}(x,2\eta)\cap B(x,2R))$. If $(s,y)\notin(t-3/2\eta,t+3/2\eta)\times(B_{-1}(x,3/2\eta)\cap B(x,3/2R))$ and $(u_1-\tilde h)^{*,-1}-\tilde\varphi$ has a maximum at (s,y), and

$$(u_1-\tilde h)^{*,-1}(s,y)=\lim_{n\to+\infty}(u_1(s_n,y_n)-\tilde h(s_n,y_n)),$$

where $|s_n-s|+|y_n-y|_{-1}\to0$ and $y_n\rightharpoonup y$, then since $|y|\le\liminf_{n\to+\infty}|y_n|$, we obtain $(s_n,y_n)\notin(t-\eta,t+\eta)\times(B_{-1}(x,\eta)\cap B(x,R))$ for large n. Thus by (3.305) $u_1(s_n,y_n)=u(s_n,y_n)$ which implies $(u_1-\tilde h)^{*,-1}(s,y)=(u-\tilde h)^{*,-1}(s,y)$. Therefore the subsolution condition is satisfied for u_1 at (s,y) for the test function $\tilde\psi=\tilde h+\tilde\varphi$, and hence u_1 is a discontinuous viscosity subsolution of (3.286).

By the definition of u_1 and (3.304) we know that $u_0\le u_1\le v_0$. Thus, by (3.294), we should have $u_1\le u$ but this contradicts (3.298). $\square$

Comparison theorem in the whole space can be proved under the same assumptions as these of Theorem 3.50. The proof is almost exactly the same. The reader can also check the proof of Theorem 4.1 in [288] for a proof in a simpler time independent case. Comparison theorem in particular implies that a discontinuous viscosity solution (if it exists) is in fact B-continuous. Thus, in particular, if comparison theorem holds, a viscosity solution in the sense of Definition 3.90 is the usual viscosity solution in the sense of Definition 3.34 (with additional requirement that test functions φ are B-continuous). However we now have a very convenient way to prove the existence of solution by Perron's method. The remaining question is how to construct a sub- and a supersolution u_0 and v_0 as in Theorem 3.93 that in addition attain the initial condition locally uniformly so that we can later use comparison theorem.

PROPOSITION 3.94 *Let Hypotheses 3.44, 3.46 and condition (3.287) hold and let g be locally uniformly B-continuous and such that $|g(x)|\le\mu(1+|x|)$ for $x\in H$ for some constant μ. Then there are a viscosity subsolution u_0 and viscosity supersolution v_0 of equation (3.286) in the sense of Definition 3.90 such that*

$$\lim_{t\downarrow0}(|u_0(t,x)-g(x)|+|v_0(t,x)-g(x)|)=0$$

uniformly on bounded sets of H.

PROOF. We will only show how to construct v_0.
Define

$$C(r)=\sup\{|F(t,x,p,X)|:x\in H,t\in[0,T],|p|\le r,\|X\|\le r\}.$$

Let $v(t,x)=\alpha t+2\mu\sqrt{1+|x|^2}$. Notice that $v(0,x)\ge g(x)$, $x\in H$. By Lemma 3.91, v is a viscosity supersolution of (3.286) if

$$\alpha+F\big(t,x,Dv(t,x),D^2v(t,x)\big)\ge0$$

for all $(t,x) \in (0,T) \times H$. Since $Dv(t,x)$ and $D^2v(t,x)$ are bounded we can therefore select α, depending only on μ, such that the above condition is satisfied.

Let $z \in H, \epsilon > 0$. We first choose a constant $R = R(|z|) \geq |z|$ such that $(|x| - |z|)_+^4 \geq 2v(t,x)$ for $|x| \geq R, t \in (0,T)$. We then find $M = M(|z|, \epsilon)$ such that

$$\bar{w}_{z,\epsilon}(x) := g(z) + \epsilon + M|x - z|_{-1}^2 + (|x| - |z|)_+^4 \geq g(x)$$

for $|x| \leq R$. Let now $\gamma = \sup\{|D\bar{w}_{z,\epsilon}(x)| + \|D^2\bar{w}_{z,\epsilon}(x)\| : |x| \leq R\}$. Using again Lemma 3.91, in order for $w_{z,\epsilon}(t,x) := \beta t + \bar{w}_{z,\epsilon}(x)$ to be a viscosity supersolution of (3.286) in the interior of $(0,T) \times B(0,R)$ we need

$$\beta + 2M\langle x, A^*B(x-z)\rangle + F(t,x,Dw_{z,\epsilon}(t,x), D^2w_{z,\epsilon}(t,x)) \geq 0$$

in this set. This can be achieved by taking $\beta = 2RM(R + |z|)\|A^*B\| + C(\gamma)$.

Since $w_{z,\epsilon}(t,x) > v(t,x)$ if $t \in (0,T), |x| \geq R$, it thus follows that

$$\hat{\omega}_{z,\epsilon}(t,x) := \min\{w_{z,\epsilon}(t,x), v(t,x)\}$$

is a B-lower semicontinuous viscosity supersolution of (3.286) in $[0,T) \times H$. It is now clear from the construction of the $\hat{\omega}_{z,\epsilon}$ and Proposition 3.92 for supersolutions that the function $v_0(t,x) := \inf_{z,\epsilon} \hat{\omega}_{z,\epsilon}(t,x)$ is a viscosity supersolution of (3.286) in the sense of Definition 3.90 such that $\lim_{t\downarrow 0} |v_0(t,x) - g(x)| = 0$ uniformly on bounded sets of H. □

In the last part of this section we show how the method of half-relaxed limits of Barles-Perthame (see [101]) can be generalized to infinite dimensional spaces. This method improves the general consistency result of Section 3.4. Suppose that we have equations

$$u_t - \langle A_n x, Du\rangle + F_n(t,x,u,Du,D^2u) = 0 \quad (t,x) \in (0,T) \times H, \tag{3.307}$$

where $F_n : [0,T] \times H \times \mathbb{R} \times H \times \mathcal{S}(H) \to \mathbb{R}$, and $A_n, n = 1,2,...$ are linear, densely defined maximal dissipative operators in H such that $D(A^*) \subset D(A_n^*)$. Let F^+, F_- be defined as in Theorem 3.41. We define

$$u^+(x) = \lim_{i\to\infty} \sup\left\{u_n(y) : n \geq i, |x-y| \leq \frac{1}{i}\right\},$$

$$u_-(x) = \lim_{i\to\infty} \inf\left\{u_n(y) : n \geq i, |x-y| \leq \frac{1}{i}\right\}.$$

THEOREM 3.95 *Let the operator B satisfying (3.2) be compact. Let A_n be as above, let A, A_n, $n = 1,2,...$ satisfy (3.287), let (3.70) hold, and let for every test function φ, the family $A_n^*D\varphi, n = 1,2,...$ be locally uniformly bounded. Suppose that F_n, $n = 1,2,...$ are continuous, locally bounded uniformly in n, and satisfy Hypotheses 3.45 and 3.46. Let u_n be locally bounded, uniformly in n, B-upper semicontinuous (respectively, B-lower semicontinuous) viscosity subsolutions, (respectively, supersolutions) of*

$$(u_n)_t - \langle A_n x, Du_n\rangle + F_n(t,x,u_n,Du_n,D^2u_n) = 0 \quad in\ (0,T) \times H \tag{3.308}$$

in the sense of Definition 3.90. Then the function u^+ (respectively, u_-) is a viscosity subsolution (respectively, supersolution) of

$$(u^+)_t - \langle Ax, Du^+\rangle + F_-(t,x,u^+,Du^+,D^2u^+) = 0 \quad in\ (0,T) \times H$$

(respectively,

$$(u_-)_t - \langle Ax, Du_-\rangle + F^+(t,x,u_-,Du_-,D^2u_-) = 0 \quad in\ (0,T) \times H)$$

in the sense of Definition 3.90.

PROOF. Let $(u^+ - h)^{*,-1} - \varphi$ have a local maximum (equal to 0) at (t, x) for some test function $\psi = \varphi + h$. In light of local uniform boundedness of the u_n we can assume that the maximum is global, strict in the $|\cdot| \times |\cdot|_{-1}$ norm, and such that

$$u^+(y) - h(s, |y|) \to -\infty, \quad (u^+ - h)^{*,-1}(y) - \varphi(s, y) \to -\infty,$$

and

$$u_n(s, y) - h(s, |y|) - \varphi(s, y) \to -\infty$$

as $|y| \to +\infty$, uniformly in n and $s \in (0, T)$, and as $s \to 0$ and $s \to T$, uniformly in n and y in bounded sets. Then there must exist a sequences t_n, x_n such that $|t_n - t| + |x_n - x|_{-1} \to 0$, $|x_n| \le C$, and

$$u^+(t_n, x_n) - h(t_n, |x_n|) - \varphi(t_n, x_n) \ge -\frac{1}{n}.$$

Therefore there exist τ_n, y_n and i_n such that

$$u_{i_n}(\tau_n, y_n) - h(\tau_n, |y_n|) - \varphi(\tau_n, y_n) \ge -\frac{2}{n}. \tag{3.309}$$

Let (s_n, z_n) be a global maximum of

$$u_{i_n}(s, y) - h(s, |y|) - \varphi(s, y).$$

It exists because of the decay of this function at infinity and around $0, T$, and the fact that, because B is compact, B-upper semicontinuity is equivalent to weak sequential upper semicontinuity. Obviously $|z_n| \le C_1$ and we also have

$$\begin{aligned}\psi_t(s_n, z_n) + \lambda |z_n|_1^2 \frac{h_r(s_n, |z_n|)}{|z_n|} - \langle z_n, A_{i_n}^* D\varphi(s_n, z_n)\rangle \\ + F_{i_n}(s_n, z_n, u_{i_n}(s_n, z_n), D\psi(s_n, z_n), D^2\psi(s_n, z_n)) \le 0.\end{aligned} \tag{3.310}$$

We can assume that $s_n \to s$. Now either $z_n \to 0$ or for a subsequence (still denoted by z_n) $|z_n| \ge c_1 > 0, n = 1, 2, ...$, which implies $h_r(s_n, |z_n|)/|z_n| > c_2 > 0$, $n = 1, 2, ...$ It then follows from the local uniform boundedness of the F_n and $A_{i_n}^* D\varphi$, that $|z_n|_1 \le C_2$ which implies $z_n \rightharpoonup z$ in H_1 for some $z \in H_1$ and thus, since B is compact, $z_n \to z$ in H.

Therefore $u^+(s, z) \ge \limsup_{n\to\infty} u_{i_n}(s_n, z_n)$ which, together with (3.309), gives

$$\begin{aligned}0 &\ge (u^+ - h)^{*,-1}(s, z) - \varphi(s, z) \ge u^+(s, z) - h(s, z) - \varphi(s, z) \\ &\ge \limsup_{n\to\infty} (u_{i_n}(s_n, z_n) - h(s_n, |z_n|) - \varphi(s_n, z_n)) \ge 0.\end{aligned}$$

Thus $(s, z) = (t, x)$ and moreover

$$(u^+ - h)^{*,-1}(t, x) + h(t, x) = \limsup_{n\to\infty} u_{i_n}(s_n, z_n).$$

It now remains to pass to $\liminf_{n\to+\infty}$ in (3.310) and use (3.70) to conclude the proof. □

If $F^+ = F_-$ and comparison holds for the limiting equation one can obtain the convergence of the u_n to the unique viscosity solution of the limiting equation. Moreover, the limiting Hamiltonians F^+ and F^- may be of first order so the above theorem can be applied to singular perturbation problems discussed in Section 3.8. Other applications related to the convergence of finite dimensional approximations (like these in Section 3.7) when condition (3.287) is satisfied by the operators A_n only on a family of finite dimensional spaces can be found in [288].

3.10. Infinite dimensional Black-Scholes-Barenblatt equation

In this section we show how the theory of viscosity solutions and the results of previous sections can be used to deal with the infinite dimensional Black-Scholes-Barenblatt equation (2.145) introduced in Section 2.6.7. We refer the reader to Section 2.6.7 for details about the financial meaning of the equation and the associated optimal control the problem.

Let H be the Sobolev space $H^1([0,+\infty))$ and let A be the maximal dissipative operator

$$\begin{cases} D(A) := H^2([0,+\infty)) \\ A(x)(\sigma) := \dfrac{dx}{d\sigma}(\sigma) \end{cases}$$

The operator A generates, by Theorem B.45, a C_0-semigroup of contractions e^{tA} in H. Let B be a bounded, self-adjoint, strictly positive operator satisfying (3.2). We introduce the space

$$\mathfrak{V} = \left\{ x \in H \ : \ \sigma \mapsto \sqrt{\sigma} x(\sigma), \ \sigma \mapsto \sqrt{\sigma}\frac{dx}{d\sigma}(\sigma) \in \mathrm{L}^2(0,\infty) \right\},$$

equipped with the norm

$$|x|^2_{\mathfrak{V}} = \int_0^\infty (1+\sigma)\left(x^2(\sigma) + \left(\frac{dx}{d\sigma}(\sigma)\right)^2\right) d\sigma.$$

and we denote by Λ a fixed bounded and closed subset of $\mathfrak{V}^d$. The space H will be the state space and Λ will be the control space.

The set of admissible controls $\mathcal{U}_t$ is defined as in Section 2.1.2, where W in reference probability spaces ν there are d-dimensional standard Brownian motions.

LEMMA 3.96 *The function $b : \mathfrak{V}^d \to H$ defined by*

$$b(x)(\sigma) = \sum_{k=1}^d x_k(\sigma) \int_0^\sigma x_k(\mu) d\mu,$$

is locally Lipschitz. Above $x = (x_1, ..., x_k)$.

PROOF. Let $R > 0$ and $x, y \in \mathfrak{V}^d$ be such that $|x_k|_{\mathfrak{V}}, |y_k|_{\mathfrak{V}} \le R$, for $k = 1, ..., d$. Then

$$|b(x)(\sigma) - b(y)(\sigma)|^2 \le \sum_{k=1}^d 2d\bigg(\sigma |x_k(\sigma) - y_k(\sigma)|^2 \int_0^\sigma x_k^2(\mu) d\mu \\ + \sigma y_k^2(\sigma) \int_0^\sigma |x_k(\mu) - y_k(\mu)|^2 d\mu \bigg).$$

Integrating we have

$$\int_0^{+\infty} |b(x)(\sigma) - b(y)(\sigma)|^2 d\sigma \le \sum_{k=1}^d 2dR^2 \int_0^{+\infty} (1+\sigma)|(x_k - y_k)(\sigma)|^2 d\sigma.$$

Similarly, we obtain

$$|(b(x))'(\sigma) - (b(y))'(\sigma)|^2 \le 3d \sum_{k=1}^d \bigg(4R^2 |x_k(\sigma) - y_k(\sigma)|^2 \\ + \sigma \left((y_k)'\right)^2(\sigma) \int_0^\sigma |x_k(\mu) - y_k(\mu)|^2 d\mu \\ + \sigma \left|\left((x_k)'\right)(\sigma) - \left((y_k)'\right)(\sigma)\right|^2 \int_0^\sigma x_k^2(\mu) d\mu \bigg)$$

which after integration yields

$$\int_0^{+\infty} |(b(x))'(\sigma) - (b(y))'(\sigma)|^2 d\sigma \leq 3d(4R^2+1) \sum_{k=1}^{d} \left(\int_0^{+\infty} |(x_k - y_k)(\sigma)|^2 d\sigma \right.$$
$$\left. + \int_0^{+\infty} \sigma \left| ((x_k)')(\sigma) - ((y_k)')(\sigma) \right|^2 d\sigma \right).$$

The claim now follows easily. □

The previous lemma implies in particular that for $\tau \in \mathcal{U}_t$, the process $b(\tau(s))$ is progressively measurable and bounded. Therefore the state equation for the problem

$$\begin{cases} dr(s) &= (Ar(s) + b(\tau(s)))ds + \tau(s) \cdot dW(s), \quad s \in (t,T] \\ r(t) &= x, \end{cases} \tag{3.311}$$

is well posed in H for any reference probability space $\nu = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W)$ and any $\tau(\cdot) \in \mathcal{U}_t$ (see Theorem 1.121). We denote its unique mild solution by $r(\cdot)$.

Our control problem consists in maximizing the cost functional

$$\mathbb{E}\left(e^{-\int_t^T r^+(s,0)ds} g(r(T)) \right) \tag{3.312}$$

over all controls $\tau(\cdot) \in \mathcal{U}_t$. (We used $r^+(s,0)$ to denote $r^+(s)(0)$.) This defines the value function

$$V(t,x) := \sup_{\tau(\cdot)\in\mathcal{U}_t} \mathbb{E}\left(e^{-\int_t^T r^+(s,0)ds} g(r_T) \right).$$

We assume the function g satisfies the following hypothesis.

HYPOTHESIS 3.97 *The function g is locally uniformly B-continuous and*

$$|g(x)| \leq C(1+|x|^m) \quad \textit{for all } x \in H,$$

for some $C, m \geq 0$.

We can now apply the results of the previous sections to the HJB of the problem (2.145). Observe first that, if we define

$$c(x) = x^+(0),$$

for x in H, then c is weakly sequentially continuous on H and so it is uniformly continuous in the $|\cdot|_{-1}$ norm on bounded sets of H. Moreover it is easy to see that c has at most linear growth at infinity and for instance

$$x^+(0) \leq 2|x| \tag{3.313}$$

so the hypotheses needed to prove the "existence" part of of Theorem 3.66 are satisfied. It guarantees the existence of a local modulus ω such that

$$|V(t,x) - V(s,y)| \leq \omega\big(|t-s| + |x-y|_{-1}; R\big) \tag{3.314}$$

for all $0 \leq t, s \leq T$, $x, y \in B(0,R)$. It also ensures that V is a viscosity solution of the Hamilton-Jacobi-Bellman equation (the BSB equation) (2.145).

As regards the uniqueness of viscosity solutions of the BSB equation (2.145) we notice that Hypotheses 3.44-3.46, 3.48, 3.49 with $\gamma = 0$ are satisfied. To guarantee Hypothesis 3.47 we need an additional assumption.

We suppose that Λ is a compact subset of H_{-1}^d. It is then obvious that

$$\sup_{\tau\in\Lambda} \sum_{i=1}^{d} |Q_N \tau_i|_{-1}^2 \to 0 \quad \text{as } N \to \infty,$$

where Q_N is defined as in Section 3.5. This implies that Hypothesis 3.47 holds. Therefore, by Theorem 3.50, comparison holds for (2.145) and thus we have the following result.

THEOREM 3.98 *Let Hypothesis 3.97 hold and let Λ be a bounded and closed subset of $\mathfrak{V}^d$ which is also a compact subset of H^d_{-1}. Then the value function V satisfies (3.314) and is the unique viscosity solution of the BSB equation (2.145) among functions satisfying (3.76) with $\gamma = 0$ and*

$$\lim_{t\to T} |u(t,x) - g(x))| = 0 \tag{3.315}$$

uniformly on bounded sets.

If g is bounded and weakly sequentially continuous, it can be shown that the value function can be approximated by viscosity solutions of finite dimensional approximations of the BSB equation (2.145). This assumption holds in many interesting cases, for example if g is given by (2.141). We refer to [287] for details of this.

3.11. HJB equation for control of the Duncan-Mortensen-Zakai equation

This section is a continuation of Section 2.6.6 and the reader should be familiar with it. We have seen in Section 2.6.6 how the Duncan-Mortensen-Zakai (DMZ) equation arises in control problems with partial observation. In the so called "separated" problem the DMZ equation is the state equation for the unnormalized conditional probability density of the state process with respect to the observation process. This gives rise to an optimal control problem for the DMZ equation which is fully observable. In this section we discuss how the HJB techniques can be applied to this problem. We first present basic results about variational solutions of SPDE.

3.11.1. Variational solutions. In this section we make the following assumptions. Let V, H be real separable Hilbert spaces. We identify H with its dual H'. Suppose that V is continuously and densely embedded in H. We then have the continuous and dense embeddings

$$V \subset H \subset V'$$

and V' is also separable. We denote the norms in V, H, V' by $|\cdot|_V$, $|\cdot|$, $|\cdot|_{V'}$ respectively. The inner product in H is denoted by $\langle\cdot,\cdot\rangle$. The duality pairing between V' and V are denoted by $\langle\cdot,\cdot\rangle_{\langle V',V\rangle}$. The duality pairing agrees with the inner product on H, i.e. for every $x \in H, v \in V, \langle x,v\rangle = \langle x,v\rangle_{\langle V',V\rangle}$. The triple (V,H,V') with the above properties is called a Gelfand triple.

Let H, Ξ be real separable Hilbert spaces, Λ be a Polish space, $Q \in \mathcal{L}_1^+(\Xi)$ and $T \in (0,+\infty)$. Let $\mu = \left(\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s\in[0,T]}, \mathbb{P}, W_Q\right)$ be a generalized reference probability space. Let $a(\cdot) \in \mathcal{U}^\mu := \mathcal{U}_0^\mu$ (see (2.1)). We assume the following hypothesis.

HYPOTHESIS 3.99 *The following conditions are satisfied:*

(i) *The linear operators $A(t,a) : V \to V'$ are closed with a common domain $D(A)$ for $(t,a) \in [0,T]\times\Lambda$, and for every $t \in [0,T]$, the map $\tilde{A} : [0,T]\times\Lambda\times V \to V'$, $\tilde{A}(t,a,v) = A(t,a)v$, restricted to $[0,t]\times\Lambda\times V$, is $\mathcal{B}([0,t])\otimes\mathcal{B}(\Lambda)\otimes\mathcal{B}(V)/\mathcal{B}(V')$ measurable. Moreover there exist C,γ, and $\beta > 0$ such that for all $u,v \in V$,*

$$|\langle A(s,a)u,v\rangle_{\langle V',V\rangle}| \le C|u|_V|v|_V,\ (s,a) \in [0,T]\times\Lambda, \tag{3.316}$$

$$\langle A(s,a)v,v\rangle_{\langle V',V\rangle} \le -\beta|v|_V^2 + \gamma|v|^2, \ (s,a)\in[0,T]\times\Lambda. \qquad (3.317)$$

(ii) *The functions $b:[0,T]\times V\times\Lambda\to H$ and $\sigma:[0,T]\times V\times\Lambda\to\mathcal{L}_2(\Xi_0,H)$ are such that for every $t\in[0,T]$ their restrictions to $[0,T]\times V\times\Lambda$ are respectively $\mathcal{B}([0,t])\otimes\mathcal{B}(V)\otimes\mathcal{B}(\Lambda)/\mathcal{B}(H)$ and $\mathcal{B}([0,t])\otimes\mathcal{B}(V)\otimes\mathcal{B}(\Lambda)/\mathcal{B}(\mathcal{L}_2(\Xi_0,H))$ measurable.*

(iii) *There exists C such that for all $u,v\in V$*

$$|b(s,v,a)| + \|\sigma(s,v,a)\|_{\mathcal{L}_2(\Xi_0,H)} \le C(1+|v|_V), \qquad (3.318)$$

$$|b(s,u,a)-b(s,v,a)| + \|\sigma(s,u,a)-\sigma(s,v,a)\|_{\mathcal{L}_2(\Xi_0,H)} \le C|u-v|_V, \qquad (3.319)$$

for all $(s,a)\in[0,T]\times\Lambda$.

(iv) *There exist C,γ_1, and $\beta_1>0$ such that for all $v\in V$,*

$$\langle A(s,a)v,v\rangle_{\langle V',V\rangle} + \|\sigma(s,v,a)\|^2_{\mathcal{L}_2(\Xi_0,H)} \le -\beta_1|v|_V^2 + \gamma_1|v|^2 + C, \qquad (3.320)$$

for all $(s,a)\in[0,T]\times\Lambda$.

(v) *There exists δ such that for all $u,v\in V$,*

$$2\langle A(s,a)(u-v),u-v\rangle_{\langle V',V\rangle} + 2\langle b(s,u,a)-b(s,v,a),u-v\rangle + \|\sigma(s,u,a)-\sigma(s,v,a)\|^2_{\mathcal{L}_2(\Xi_0,H)} \le \delta|u-v|^2, \quad (3.321)$$

for all $(s,a)\in[0,T]\times\Lambda$.

It is now easy to see that the maps $A(s,a(s))v$, $b(s,v,a(s))$, $\sigma(s,v,a(s))$ defined on $[0,T]\times\Omega\times V$ are such that for every $t\in[0,T]$ their restrictions to $[0,t]\times\Omega\times V$ are respectively $\mathcal{B}([0,t])\otimes\mathscr{F}_t\otimes\mathcal{B}(V)/\mathcal{B}(V')$, $\mathcal{B}([0,t])\otimes\mathscr{F}_t\otimes\mathcal{B}(V)/\mathcal{B}(H)$ and $\mathcal{B}([0,t])\otimes\mathscr{F}_t\otimes\mathcal{B}(V)/\mathcal{B}(\mathcal{L}_2(\Xi_0,H))$ measurable, and they satisfy (3.316)-(3.321) (with a replaced by $a(s)$) for a.e. $(s,\omega)\in[0,T]\times\Omega$.

We consider the following stochastic PDE

$$\begin{cases} dX(s) = (A(s,a(s))X(s) + b(s,X(s),a(s)))ds + \sigma(s,X(s),a(s)))dW_Q(s) \\ X(0)=\xi. \end{cases} \qquad (3.322)$$

Definition 3.100 (Variational solution of (3.322)) *A process $X(\cdot)\in M^2_\mu(0,T;H)$ is called a* variational solution *of (3.322) if*

$$\mathbb{E}\left[\int_0^T |X(r)|_V^2 dr\right] < +\infty$$

and for every $\phi\in V$ we have

$$\langle X(s),\phi\rangle = \langle\xi,\phi\rangle + \int_0^s \langle A(r,a(r))X(r),\phi\rangle_{\langle V',V\rangle}dr + \int_0^s \langle b(r,X(r),a(r)),\phi\rangle dr + \int_0^s \langle\sigma(r,X(r),a(r))dW_Q(r),\phi\rangle \quad \textit{for each } s\in[0,T]\text{, } \mathbb{P}\textit{-a.e.}. \qquad (3.323)$$

We remark that the integrand $A(r,a(r))X(r)$ above is evaluated at a V valued progressively measurable equivalent version of $X(\cdot)$, and the process $\mathbf{1}_{X(s)\in V}X(s)$ is equivalent to the process $X(s)$ and, by Lemma 1.17-(iii), belongs to $M^2_\mu(0,T;V)$. Moreover, (3.323) is equivalent to the equality

$$X(s) = \xi + \int_0^s (A(r,a(r))X(r) + b(r,X(r),a(r)))dr + \int_0^s \sigma(r,X(r),a(r))dW_Q(r)$$

as elements of V'.

The following result is taken from [296], Theorem I.3.1, in the version from [220], Theorem 4.3, page 165, see also [385], Theorem 4.2.5, and [406].

THEOREM 3.101 *Let μ be a generalized reference probability space, ξ be $\mathscr{F}_0$ measurable H valued random variable such that $\mathbb{E}^\mu[|\xi|^2] < +\infty$, let $Y(\cdot) \in M^2_\mu(0,T;V')$, $Z(\cdot) \in \mathcal{N}^2_Q(0,T;H)$. We define the continuous V' valued process*

$$X(s) = \xi + \int_0^s Y(r)dr + \int_0^s Z(r)dW_Q(r), \quad s \in [0,T].$$

If $X(\cdot)$ has an equivalent version $\tilde{X}(\cdot) \in M^2_\mu(0,T;V)$, then $X(\cdot) \in M^2_\mu(0,T;H) \cap L^2(\Omega; C([0,T];H))$,

$$\mathbb{E}\left[\sup_{0\le s\le T} |X(s)|^2\right] \le +\infty,$$

and the following Ito's formula holds $\mathbb{P}$-a.e.

$$\begin{aligned}|X(s)|^2 = |\xi|^2 + \int_0^s \Big(2\langle Y(r), \tilde{X}(r)\rangle_{\langle V',V\rangle} + \|Z(r)\|^2_{\mathcal{L}_2(U_0,H)}\Big)\, dr \\ +2\int_0^s \langle Z(r)dW_Q(r), X(r)\rangle \quad s \in [0,T].\end{aligned} \tag{3.324}$$

THEOREM 3.102 *Let μ be a generalized reference probability space, ξ be $\mathscr{F}_0$ measurable H valued random variable such that $\mathbb{E}^\mu[|\xi|^2] < +\infty$, and $a(\cdot) \in \mathcal{U}^\mu$. Then:*

(i) *There exists a unique variational solution of (3.322) $X(\cdot) \in L^2(\Omega; C([0,T];H))$, and the energy equality holds $\mathbb{P}$-a.e.*

$$\begin{aligned}|X(s)|^2 = |\xi|^2 + 2\int_0^s \langle A(r,a(r))X(r), X(r)\rangle_{\langle V',V\rangle} dr + 2\int_0^s \langle b(r,X(r),a(r)), X(r)\rangle dr \\ +2\int_0^s \langle \sigma(r,X(r),a(r))dW_Q(r), X(r)\rangle + \int_0^s \|\sigma(r,X(r),a(r))\|^2_{\mathcal{L}_2(\Xi_0,H)} dr \quad s \in [0,T].\end{aligned} \tag{3.325}$$

(ii) *If μ_1 is another generalized reference probability space, ξ_1 is a $\mathscr{F}_0^{\mu_1}$ measurable H-valued random variable such that $\mathbb{E}^{\mu_1}[|\xi_1|^2] < +\infty$, $a_1(\cdot) \in \mathcal{U}^\mu$, and*

$$\mathcal{L}_{\mathbb{P}_1}(\xi_1, a_1(\cdot), W_{Q,1}(\cdot)) = \mathcal{L}_{\mathbb{P}}(\xi, a(\cdot), W_Q(\cdot)),$$

then

$$\mathcal{L}_{\mathbb{P}_1}(a_1(\cdot), X_1(\cdot)) = \mathcal{L}_{\mathbb{P}}(a(\cdot), X(\cdot)), \tag{3.326}$$

where $X_1(\cdot)$ is the variational solution of (3.322) in μ_1 with control $a_1(\cdot)$ and initial condition ξ_1.

PROOF. We sketch the proof. The complete proof of the first part of the theorem can be found in [93], pages 168–183 or [220, 296, 298, 385, 406]. Let $\{v_1, v_2, ...\}$ be an orthonormal basis of H composed of elements of V. We set $H_n := \mathrm{span}\{v_1, ..., v_n\}$, and define

$$P_n w := \sum_{k=1}^n \langle w, v_k\rangle_{\langle V',V\rangle} v_k, \quad w \in V'.$$

If $w \in H$ we have

$$P_n w = \sum_{k=1}^n \langle w, v_k\rangle v_k$$

so P_n is an extension to V' of the orthogonal projection in H onto H_n. We set

$$\begin{cases} A^n(s,\alpha)v := P_n A(s,\alpha)v, \ b^n(s,v,\alpha) := P_n b(s,v,\alpha), \ \sigma^n(s,v,\alpha) := P_n\sigma(s,v,\alpha) \\ \xi^n := P_n\xi. \end{cases}$$

(One can also project the Wiener process on a finite dimensional subspace but it is not necessary.) Since the above functions are Lipschitz continuous in v on H_n, standard theory guarantees that there exists a unique strong solution (in the sense of Definition 1.112) $X^n(\cdot)$ of

$$\begin{cases} dX^n(s) = (A^n(s,a(s))X^n(s) + b^n(s,X^n(s),a(s)))ds + \sigma^n(s,X^n(s),a(s)))dW_Q(s) \\ X^n(0) = \xi^n \end{cases} \tag{3.327}$$

which satisfies $X^n(\cdot) \in L^2_\mu(\Omega; C([0,T];H)) \cap M^2_\mu(0,T;V)$. Moreover, Ito's formula gives $\mathbb{P}$-a.e.

$$\begin{aligned} |X^n(s)|^2 &= |\xi^n|^2 + 2\int_0^s \langle A^n(r,a(r))X^n(r), X^n(r)\rangle dr \\ &+ 2\int_0^s \langle b^n(r,X^n(r),a(r)), X^n(r)\rangle dr + 2\int_0^s \langle \sigma^n(r,X^n(r),a(r))dW_Q(r), X^n(r)\rangle \\ &+ \int_0^s \|\sigma^n(r,X^n(r),a(r))\|^2_{\mathcal{L}_2(\Xi_0,H)} dr, \end{aligned} \tag{3.328}$$

and using (3.328) and the assumptions one shows

$$\mathbb{E}\left[\sup_{0\le s\le T} |X^n(s)|^2 + \int_0^T |X^n(s)|^2_V ds\right] \le M \quad \text{for all } n.$$

Therefore, up to subsequences still denoted by X^n, b^n, σ^n, we obtain that there exist $X(\cdot)$, $\tilde{b}(\cdot) \in M^2_\mu(0,T;H)$, $\tilde{X}(\cdot) \in M^2_\mu(0,T;V)$, $\tilde{\sigma}(\cdot) \in \mathcal{N}^2_Q(0,T;H)$, and a $\mathscr{F}_T$ measurable random variable $\eta \in L^2(\Omega;H)$ such that

$$X^n(\cdot) \rightharpoonup \tilde{X}(\cdot) \text{ in } M^2_\mu(0,T;V), \quad X^n(T) \rightharpoonup \eta \text{ in } L^2(\Omega;H),$$

$$X^n(\cdot) \rightharpoonup X(\cdot), b^n(r,X^n(r),a(r)) \rightharpoonup \tilde{b}(\cdot) \text{ in } M^2_\mu(0,T;H),$$

$$\sigma^n(r,X^n(r),a(r)) \rightharpoonup \tilde{\sigma}(\cdot) \text{ in } \mathcal{N}^2_Q(0,T;H).$$

Obviously $\tilde{X}(\cdot)$ is an equivalent version of $X(\cdot)$. Passing to the limit as $n \to +\infty$ one obtains that $X(\cdot)$ is a variational solution of the linear equation

$$\begin{cases} dX(s) = (A(s,a(s))\tilde{X}(s) + \tilde{b}(s)ds + \tilde{\sigma}(s)dW_Q(s) \\ X(0) = \xi, \end{cases}$$

$X(T) = \eta$, and, by Theorem 3.101, $X(\cdot) \in L^2_\mu(\Omega; C([0,T];H))$ and it satisfies $\mathbb{P}$-a.e.

$$\begin{aligned} |X(s)|^2 = |\xi|^2 + 2\int_0^s \langle A(r,a(r))X(r), X(r)\rangle_{\langle V',V\rangle} dr + 2\int_0^s \langle \tilde{b}(r), X(r)\rangle dr \\ +2\int_0^s \langle \tilde{\sigma}(r)dW_Q(r), X(r)\rangle + \int_0^s \|\tilde{\sigma}(r)\|^2_{\mathcal{L}_2(U_0,H)} dr. \end{aligned}$$

One then uses monotonicity arguments to prove that

$$\tilde{b}(r) = b(r,X(r),a(r)), \ \tilde{\sigma}(r) = \sigma(r,X(r),a(r)) \quad dt \otimes \mathbb{P} \text{ a.e.}$$

and hence $X(\cdot)$ is a variational solution of (3.322) and (3.325) holds. Moreover it also follows from these arguments (see for instance [296] for details) that $\mathbb{E}[|X^n(T)|^2] \to \mathbb{E}[|X(T)|^2]$ and so $X^n(T) \to X(T)$ in $L^2(\Omega;H)$. Replacing interval

$[0, T]$ by another interval $[0, t], 0 < t < T$ the same arguments give $X^n(t) \to X(t)$ in $L^2(\Omega; H)$ for all $0 \le t \le T$.

The uniqueness of variational solution in the generalized reference probability space μ follows from Theorem 3.101, elementary estimates using the assumptions on the coefficients, and Gronwall's lemma.

Finally, if $X_1(\cdot)$ is the variational solution in the generalized reference probability space μ_1 and $X_1^n(\cdot)$ are the solutions of the approximating problems (3.327) in this space then we have

$$\mathcal{L}_{\mathbb{P}_1}(a_1(\cdot), X_1^n(\cdot)) = \mathcal{L}_{\mathbb{P}}(a(\cdot), X^n(\cdot)),$$

and thus (3.326) follows since $X_1^n(t) \to X_1(t)$ in $L^2(\Omega_1; H)$ and $X^n(t) \to X(t)$ in $L^2(\Omega; H)$ for every $t \in [0, T]$. $\square$

3.11.2. Weighted Sobolev spaces. We denote the norm and the inner product in $\mathbb{R}^d$ by $|\cdot|_{\mathbb{R}^d}$ and $\langle \cdot, \cdot \rangle_{\mathbb{R}^d}$ respectively. Given $k \in \mathbb{N}$ we denote by $H^k := H^k\left(\mathbb{R}^d\right)$ the standard Sobolev space on $\mathbb{R}^d$. We remind that $H^0 = L^2\left(\mathbb{R}^d\right)$, and the inner product in H^0 will be denoted by $\langle \cdot, \cdot \rangle_0$. Let $B := (-\Delta + I)^{-1}$, where $D(\Delta) = H^2$. We equip H^k with the inner product

$$\langle x, y \rangle_k := \langle B^{-k/2} x, B^{-k/2} y \rangle_0.$$

This inner product gives the norm $|x|_k = |B^{-k/2} x|$ which is equivalent to the standard norm in H^k given by

$$\left(\sum_{|\alpha| \le k} \int_{\mathbb{R}^d} |\partial^\alpha x(\xi)|^2 \xi \right)^{\frac{1}{2}}.$$

The topological dual space of H^k is denoted by H^{-k}. Except when explicitly stated we always identify H^0 with its dual. The space H^{-k} can be identified with the completion of H^0 under the norm

$$|x|_{-k} := |B^{k/2} x| = \langle B^k x, x \rangle_0^{1/2}$$

and then $B^{1/2}$ (after a natural extension) is an isometry between H^k and H^{k+1}, $k \in \mathbb{Z}$. The space H^{-k}, $k \in \mathbb{N}$ is a Hilbert space equipped with the inner product

$$\langle x, y \rangle_{-k} := \langle B^{k/2} x, B^{k/2} y \rangle_0.$$

The duality pairing between H^{-k} and $H^k, k \in \mathbb{N}$ is denoted by $\langle \cdot, \cdot \rangle_{\langle H^{-k}, H^k \rangle}$. We have $\langle a, b \rangle_{\langle H^{-k}, H^k \rangle} = \left\langle B^{k/2} a, B^{-k/2} b \right\rangle_0$. Observe also that, for $k \in \mathbb{Z}$, the adjoint of the operator $B^{1/2} : H^k \to H^{k+1}$ is $B^{1/2} : H^{-k-1} \to H^{-k}$.

Let $k = 0, 1, 2$. Given a strictly positive real-valued function $\rho \in C^2(\mathbb{R}^d)$, we define the *weighted Sobolev space* $H^k_\rho(\mathbb{R}^d)$ (or simply H^k_ρ) to be the completion of $C_c^\infty(\mathbb{R}^d)$ with respect to the *weighted norm*

$$|x|_{k,\rho} := |\rho x|_k.$$

The space H^k_ρ can also be defined as the space of all measurable functions $x : \mathbb{R}^d \to \mathbb{R}$ such that $\rho(\cdot) x(\cdot) \in H^k$. We recall that the norm $|x|_{k,\rho}$ is equivalent to the norm given by

$$\left(\sum_{|\alpha| \le k} \int_{\mathbb{R}^d} \left| \partial^\alpha \left[\rho(\xi) x(\xi) \right] \right|^2 d\xi \right)^{1/2}.$$

We denote by C_ρ the isometry $C_\rho : H^k_\rho \to H^k$ defined as $(C_\rho x)(\xi) = \rho(\xi) x(\xi)$, and by $C_{1/\rho} = C_\rho^{-1} : H^k \to H^k_\rho$ its inverse: $\left(C_{1/\rho} x\right)(\xi) = \rho^{-1}(\xi) x(\xi)$. We observe that H^k_ρ is a Hilbert space with the inner product $\langle x, y \rangle_{k,\rho} = \langle C_\rho x, C_\rho y \rangle_k$. We denote the topological dual space of H^k_ρ by H^{-k}_ρ and, identifying H^0_ρ with its dual, we have

$$H^k_\rho \subset H^0_\rho = [H^0_\rho]' \subset \left[H^k_\rho\right]' = H^{-k}_\rho, \quad k \geq 0. \tag{3.329}$$

We always use this identification, except when explicitly stated.

The adjoint C^*_ρ of C_ρ is an isometry $C^*_\rho : H^{-k} \longmapsto H^{-k}_\rho$. Observe that C^*_ρ can be identified with $C_{1/\rho}$.

To simplify notation we write $X_k := H^k_\rho$.

Let $B_\rho := C_{1/\rho}\left[(-\Delta + I)^{-1}\right] C_\rho = C_{1/\rho} B C_\rho$. Similarly to the case of non-weighted spaces X_{-k} can be identified with the completion of X_0 under the norm $|x|^2_{-k,\rho} := \langle B^k_\rho x, x\rangle_{0,\rho} = \langle B^k C_\rho x, C_\rho x\rangle_0$ and then $B^{1/2}_\rho$ is an isometry between X_{-2}, X_{-1}, X_0, X_1 and X_{-1}, X_0, X_1, X_2 respectively. We remark that $B^{-1}_\rho = C_{1/\rho} B^{-1} C_\rho$, $B^{1/2}_\rho = C_{1/\rho} B^{1/2} C_\rho$, $B^{-1/2}_\rho = C_{1/\rho} B^{-1/2} C_\rho$. Thus $|x|_{-k,\rho} = |B^{k/2}_\rho x|_{0,\rho}$ and $|x|_{k,\rho} = |B^{-k/2}_\rho x|_{0,\rho}$. The duality pairing between X_{-k} and X_k is denoted by $\langle \cdot,\cdot\rangle_{\langle X_{-k}, X_k\rangle}$. We have

$$\langle a, b\rangle_{\langle X^{-k}, X^k\rangle} = \langle B^{k/2}_\rho a, B^{-k/2}_\rho b\rangle_{0,\rho} = \langle C_\rho a, C_\rho b\rangle_{\langle H^{-k}, H^k\rangle}.$$

In what follows we consider weight functions ρ of the form

$$\rho_\beta(\xi) = 1 + |\xi|^\beta_{\mathbb{R}^d}, \qquad \beta > 2. \tag{3.330}$$

With such choice of ρ it can be showed that $X_k \subset H^k$, and if $\beta > d/2$ then $X_0 \subset L^1\left(\mathbb{R}^d\right)$ and $X_k \subset W^{k,1}\left(\mathbb{R}^d\right)$.

3.11.3. Optimal control of the Duncan-Mortensen-Zakai equation. We now study the optimal control problem for the DMZ equation derived in Section 2.6.6. We study it in the weighted space X_0 using the formalism of abstract control problems. Let $T > 0$. The control set Λ was originally a subset of $\mathbb{R}^n$ but we will consider Λ to be a more general Polish space. For every $0 \leq t \leq T$, the reference and generalized reference probability spaces are defined by Definitions 2.7 and 1.95 where W_Q is now just a standard Wiener process in $\mathbb{R}^m$ (i.e. $\Xi = \mathbb{R}^m, Q = I$) which is denoted by W. The classes of admissible controls with respect to all reference and generalized reference probability spaces are defined as always by $\mathcal{U}_t$ and $\overline{\mathcal{U}}_t$. We remark that for a reference probability space $\mu = \left(\Omega, \mathscr{F}, \{\mathscr{F}^t_s\}_{s\in[t,T]}, \mathbb{P}, W\right)$, $\mathbb{P}$ now corresponds to $\bar{\mathbb{P}}$ in Section 2.6.6, $\mathscr{F}^t_s$ to $\mathscr{F}^{y_1,t}_s$, and W to y_1. Without loss of generality we will always assume that the Q-Wiener processes in the reference probability spaces have everywhere continuous paths.

Recall from Section 2.6.6 that for every $a \in \Lambda$ we have the differential operators A_a and S^k_a $(k = 1, \ldots, m)$

$$(A_a x)(\xi) = \sum_{i,j=1}^d \partial_i\left[a_{i,j}(\xi, a)\partial_j x(\xi)\right] + \sum_{i=1}^d \partial_i\left[b_i(\xi, a)x(\xi)\right], \tag{3.331}$$

$$\left(S^k_a x\right)(\xi) = \sum_{i=1}^d d_{ik}(\xi, a)\partial_i x(\xi) + e_k(\xi, a)x(\xi); \quad k = 1, ..., m. \tag{3.332}$$

Typically we set $D(A_a) = X_2, D(S^k_a) = X_1,\ k = 1, ..., m$, however the operators will be considered with different domains in different Gelfand triples. Having in mind the original Hypothesis 2.46, we assume the following hypothesis.

Hypothesis 3.103

(i) *Λ is a compact metric space.*

(ii) *The coefficients*

$$\left(a_{ij}\right)_{i,j=1,\dots,d}, \left(b_i\right)_{i=1,\dots,d}, c, \left(d_{ik}\right)_{i=1,\dots,d;\ k=1,\dots,m}, \left(e_k\right)_{k=1,\dots,m} : \mathbb{R}^d \times \Lambda \to \mathbb{R}$$

are continuous in (ξ, a) *and, as functions of* ξ *are in* $C_b^2\left(\mathbb{R}^d\right)$ *for every* a, *with their norms in* $C_b^2\left(\mathbb{R}^d\right)$ *bounded uniformly in* $a \in \Lambda$. *Moreover, there exists a constant* $\lambda > 0$ *such that*

$$\sum_{i,j=1}^{d} \left(a_{i,j}(\xi, a) - \frac{1}{2} \sum_{k=1}^{m} d_{ik}(\xi, a) d_{jk}(\xi, a) \right) z_i z_j \geq \lambda |z|^2 \tag{3.333}$$

for every $a \in \Lambda$ *and* $\xi, z \in \mathbb{R}^d$.

(iii) *The weight* ρ *is of the form (3.330).*

For every $a(\cdot) \in \mathcal{U}_t$ the DMZ equation is considered in X_0:

$$\begin{cases} dY(s) = A_{a(s)} Y(s) ds + \sum_{k=1}^{m} S_{a(s)}^k Y(s) dW_k(s), & s > t \\ Y(t) = x \in X_0. \end{cases} \tag{3.334}$$

We consider the cost functional

$$J(t, x; a(\cdot)) = \mathbb{E} \left\{ \int_t^T l(Y(s), a(s)) ds + g(Y(T)) \right\}$$

and we make the following assumptions about the cost functions l and g.

HYPOTHESIS 3.104

(i) $l : X_0 \times \Lambda \to \mathbb{R}$ *and* $g : X_0 \to \mathbb{R}$ *are continuous and there exist* $C > 0$ *and* $\gamma < 2$ *such that*

$$|l(x, a)|, |g(x)| \leq C \left(1 + |x|_{0,\rho}^{\gamma} \right)$$

for every $(x, a) \in X_0 \times \Lambda$;

(ii) *for every* $R > 0$ *there exists a modulus* ω_R *such that*

$$|l(x, a) - l(y, a)| \leq \omega_R \left(|x - y|_{0,\rho} \right), \quad |g(x) - g(y)| \leq \omega_R \left(|x - y|_{-1,\rho} \right) \tag{3.335}$$

for every $x, y \in B_{X_0}(0, R)$, $a \in \Lambda$.

The optimal control problem in the weak formulation we study consists in minimizing the cost $J(t, x; a(\cdot))$ over all admissible controls $a(\cdot) \in \mathcal{U}_t$.

The associated HJB equation in X_0 has the form

$$\begin{cases} v_t + \inf_{a \in \Lambda} \left\{ \frac{1}{2} \sum_{k=1}^{m} \left\langle D^2 v S_a^k x, S_a^k x \right\rangle_{\rho,0} + \left\langle A_a x, Dv \right\rangle_{\rho,0} + l(x, a) \right\} = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0, T) \times X_0, \\ v(T, x) = g(x), \end{cases} \tag{3.336}$$

and thevalue function is

$$V(t, x) = \inf_{a(\cdot) \in \mathcal{U}_t} J(t, x; a(\cdot)). \tag{3.337}$$

Let us now describe which restrictions may be placed on the original separated problem of Section 2.6.6 so that the current assumptions be satisfied.

First the law of the initial datum η of equation (2.129) must have density x in $L_\rho^2(\mathbb{R}^d)$. To guarantee that the density is also in $L^1(\mathbb{R}^d)$ we should assume $\beta > d/2$. The density x is polynomially decreasing when $|\xi|_{\mathbb{R}^d} \to +\infty$. This is of course a further restriction with respect to assuming only $x \in L^1\left(\mathbb{R}^d\right)$ but it is satisfied in many practical cases, for instance when the starting distribution is normal. One can consult e.g. [34], pages 36, 204, for the use of x being Gaussian or [34], pages

82, 167, for other integrability assumptions on x (see also [468], [369] on this). Regarding the cost functional 2.134 we have (recall that we now use $\mathbb{E}$ in place of $\bar{\mathbb{E}}$)

$$J(t,x;a(\cdot)) = \mathbb{E}\left\{\int_t^T \langle l_1(\cdot,a(s)), Y(s)\rangle_0\, ds + \langle g_1(\cdot), Y(T)\rangle_0\right\}$$

$$= \mathbb{E}\left\{\int_t^T \left\langle (1/\rho^2) l_1(\cdot,a(s)), Y(s)\right\rangle_{0,\rho} ds + \left\langle (1/\rho^2) g_1(\cdot), Y(T)\right\rangle_{0,\rho}\right\}$$

$$= \mathbb{E}\left\{\int_t^T l(Y(s),a(s))ds + g(Y(T))\right\},$$

where we set

$$l\,(x,a) = \langle l_1\,(\cdot,a)\,,x\rangle_0 = \left\langle \frac{1}{\rho^2} l_1\,(\cdot,a)\,,x\right\rangle_{0,\rho},$$

$$g\,(x) = \langle g_1\,(\cdot)\,,x\rangle_0 = \left\langle \frac{1}{\rho^2} g_1\,(\cdot)\,,x\right\rangle_{0,\rho}. \tag{3.338}$$

It is easy to see that Hypothesis 3.104 is satisfied if the functions $l_1 : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}$ and $g_1 : \mathbb{R}^d \to \mathbb{R}$ are continuous and $\sup_{a\in\Lambda} |\frac{1}{\rho} l_1\,(\cdot,a)\,|_0 + |\frac{1}{\rho} g_1\,(\cdot)\,|_0 < +\infty$, since in this case the function g is weakly sequentially continuous in X_0. For instance, if

$$l_1(\xi,\alpha) = \langle M\xi,\xi\rangle_{\mathbb{R}^d} + \langle N\alpha,\alpha\rangle_{\mathbb{R}^n}, \quad g_1(\xi) = \langle G\xi,\xi\rangle_{\mathbb{R}^d},$$

where M, N and G are suitable non-negative definite matrices, $\Lambda = B_{\mathbb{R}^n}(0,R)$, Hypothesis 3.104 is satisfied if $\beta > 2 + d/2$. This is the main advantage of using weighted spaces. When the initial density is, say, polynomially decreasing at infinity, we can deal with polynomially growing cost functions. This is not possible if we took $\rho = 1$. Finally we mention that in the absence of density the separated problem has to be studied in the space of measures.

3.11.4. Estimates for the DMZ equation.

LEMMA 3.105 *Let Hypothesis 3.103 hold. Then:*

(i) *The DMZ equation satisfies the assumptions of Hypothesis 3.99 for the Gelfand triple (X_1, X_0, X_{-1}) and also for (X_2, X_1, X_0).*
(ii) *There exist constants $\bar{\lambda} > 0, K \geq 0$ such that for all $a \in \Lambda$*

$$\langle A_a x, x\rangle_{\langle X_{-1},X_1\rangle} + \frac{1}{2}\sum_{k=1}^m \left\langle S_a^k x, S_a^k x\right\rangle_{0,\rho} \leq -\bar{\lambda}\,|x|_{1,\rho}^2 + K\,|x|_{0,\rho}^2, \quad x\in X_1, \tag{3.339}$$

$$\left\langle A_a x, B_\rho^{-1} x\right\rangle_{0,\rho} + \frac{1}{2}\sum_{k=1}^m \left\langle B_\rho^{-1} S_a^k x, S_a^k x\right\rangle_{\langle X_{-1},X_1\rangle} \leq -\bar{\lambda}\,|x|_{2,\rho}^2 + K\,|x|_{1,\rho}^2, \quad x\in X_2, \tag{3.340}$$

$$\langle A_a x, B_\rho x\rangle_{\langle X_{-2},X_2\rangle} + \frac{1}{2}\sum_{k=1}^m \left\langle B_\rho S_a^k x, S_a^k x\right\rangle_{\langle X_1,X_{-1}\rangle} \leq -\bar{\lambda}\,|x|_{0,\rho}^2 + K\,|x|_{-1,\rho}^2, \quad x\in X_0. \tag{3.341}$$

PROOF. Part (i) follows from direct computations and estimates in (ii). Part (ii) is proved in [245], Lemma 3.3. □

We also record for future use that

$$\sup_{a\in\Lambda,k=1,\dots,m}\left(\|B_\rho A_a\|_{\mathcal{L}(X_0)}+\|B_\rho^{1/2}S_a\|_{\mathcal{L}(X_0)}\right)\le C \tag{3.342}$$

PROPOSITION 3.106 *Assume that Hypothesis 3.103 hold. Let $0\le t\le T$, let μ be a generalized reference probability space, $a(\cdot)\in\mathcal{U}_t^\mu$ and $x\in L^2(\Omega;X_0)$ be an $\mathscr{F}_t$ measurable random variable. Then there exists a unique variational solution $Y(s):=Y(\cdot;t,x,a(\cdot))\in L^2(\Omega;C([t,T];H))$ of the state equation (3.334). Moreover we have:*

- $\mathbb{P}$ *a.s.*

$$\begin{aligned}|Y(s)|_{0,\rho}^2&=|x|_{0,\rho}^2+2\int_t^s\left\langle A_{a(r)}Y(r),Y(r)\right\rangle_{\langle X_{-1},X_1\rangle}dr\\&+\sum_{k=1}^m\int_t^s\left\langle S_{a(r)}^kY(r),Y(r)\right\rangle_{0,\rho}dW_k(r)\\&+\sum_{k=1}^m\int_t^s\left\langle S_{a(r)}^kY(r),S_{a(r)}^kY(r)\right\rangle_{0,\rho}dr,\quad s\in[t,T].\end{aligned}\tag{3.343}$$

  In particular

$$\begin{aligned}\mathbb{E}\,|Y(s)|_{0,\rho}^2&=\mathbb{E}\,|x|_{0,\rho}^2+2\mathbb{E}\int_t^s\left\langle A_{a(r)}Y(r),Y(r)\right\rangle_{\langle X_{-1},X_1\rangle}dr\\&\quad+\sum_{k=1}^m\mathbb{E}\int_t^s\left\langle S_{a(r)}^kY(r),S_{a(r)}^kY(r)\right\rangle_{0,\rho}dr.\end{aligned}\tag{3.344}$$

- *There exists a constant $C>0$ independent of $\mu,a(\cdot)\in\mathcal{U}_t^\mu$ and x such that*

$$\mathbb{E}\,|Y(s)|_{0,\rho}^2\le\mathbb{E}\,|x|_{0,\rho}^2\,(1+C(s-t)),\tag{3.345}$$

$$\mathbb{E}\int_t^T|Y(s)|_{1,\rho}^2\,ds\le C\mathbb{E}\,|x|_{0,\rho}^2\,.\tag{3.346}$$

- *The conclusion of Theorem 3.102-(ii) (with 0 replaced by t) is satisfied.*

PROOF. The results follows from Theorem 3.102. □

The following proposition collects various estimates for solutions of (3.334).

PROPOSITION 3.107 *Assume that Hypothesis 3.103 hold and let $0\le t\le T$. Let $a(\cdot)\in\mathcal{U}_t$ and $x\in X_0$. Then:*

(i) *There exists a constant $C>0$ independent of $a(\cdot)\in\mathcal{U}_t$ and x such that*

$$\mathbb{E}\,|Y(s)|_{-1,\rho}^2\le|x|_{-1,\rho}^2\,(1+C(s-t)),\tag{3.347}$$

$$\mathbb{E}\int_t^T|Y(s)|_{0,\rho}^2\,ds\le C\,|x|_{-1,\rho}^2\,.\tag{3.348}$$

$$\mathbb{E}\,|Y(s)-x|_{-1,\rho}^2\le C\,(s-t)\,|x|_{0,\rho}^2\,,\tag{3.349}$$

$$\mathbb{E}\int_t^s|Y(r)-x|_{0,\rho}^2\,dr\le C\,(s-t)\,|x|_{0,\rho}^2\,.\tag{3.350}$$

There is a modulus σ_x, independent of $a(\cdot)\in\mathcal{U}_t$ such that

$$\mathbb{E}\,|Y(s)-x|_{0,\rho}^2\le\sigma_x(s-t).\tag{3.351}$$

(ii) *If in addition $x \in X_1$, then $Y(s)$ is a strong solution and there exists a constant $C > 0$ independent of $a(\cdot) \in \mathcal{U}_t$ and x such that*

$$\mathbb{E}\,|Y(s)|^2_{1,\rho} \le |x|^2_{1,\rho}\,(1 + C(s-t)), \tag{3.352}$$

$$\mathbb{E}\int_t^T |Y(s)|^2_{2,\rho}\,ds \le C\,|x|^2_{1,\rho}\,, \tag{3.353}$$

$$\mathbb{E}\,|Y(s) - x|^2_{0,\rho} \le C\,|x|^2_{1,\rho}\,(s-t), \tag{3.354}$$

$$\mathbb{E}\int_t^s |Y(r) - x|^2_{1,\rho}\,dr \le C(s-t)\,|x|^2_{1,\rho}\,. \tag{3.355}$$

There is a modulus σ_x, independent of $a(\cdot) \in \mathcal{U}_t$ such that

$$\mathbb{E}\,|Y(s) - x|^2_{1,\rho} \le \sigma_x(s-t). \tag{3.356}$$

Proof.
(i). By Itô's formula we have

$$\begin{aligned}
\mathbb{E}\,|Y(s)|^2_{-1,\rho} &= \mathbb{E}\left|B_\rho^{1/2}Y(s)\right|^2_{0,\rho} \\
&= \mathbb{E}\left|B_\rho^{1/2}x\right|^2_{0,\rho} + 2\mathbb{E}\int_t^s \left\langle B_\rho^{1/2}A_{a(r)}Y(r), B_\rho^{1/2}Y(r)\right\rangle_{0,\rho} dr \\
&\quad + \sum_{k=1}^m \mathbb{E}\int_t^s \left\langle B_\rho^{1/2}S^k_{a(r)}Y(r), B_\rho^{1/2}S^k_{a(r)}Y(r)\right\rangle_{0,\rho} dr.
\end{aligned} \tag{3.357}$$

Since

$$\left\langle B_\rho^{1/2}A_{a(r)}Y(r), B_\rho^{1/2}Y(r)\right\rangle_{0,\rho} = \left\langle A_{a(r)}Y(r), B_\rho Y(r)\right\rangle_{\langle X_{-2},X_2\rangle}$$

and

$$\left\langle B_\rho^{1/2}S^k_{a(r)}Y(r), B_\rho^{1/2}S^k_{a(r)}Y(r)\right\rangle_{0,\rho} = \left\langle B_\rho S^k_{a(r)}Y(r), S^k_{a(r)}Y(r)\right\rangle_{\langle X_1,X_{-1}\rangle}$$

we have, thanks to (3.341),

$$\mathbb{E}\,|Y(s)|^2_{-1,\rho} + 2\bar\lambda\int_t^s \mathbb{E}\,|Y(r)|^2_{0,\rho}\,dr \le |x|^2_{-1,\rho} + 2K\int_t^s \mathbb{E}\,|Y(r)|^2_{-1,\rho}\,dr.$$

Estimates (3.347)-(3.348) follow by applying Gronwall's inequality.

To show (3.349), we have

$$\mathbb{E}\,|Y(s) - x|^2_{-1,\rho} = \mathbb{E}\,|Y(s)|^2_{-1,\rho} + |x|^2_{-1,\rho} - 2\mathbb{E}\,\langle Y(s), B_\rho x\rangle_{0,\rho}$$

which gives, by (3.357) and by the definition of variational solution,

$$\begin{aligned}
&\mathbb{E}\,|Y(s) - x|^2_{-1,\rho} \\
= \;& 2\mathbb{E}\int_t^s \left[\left\langle A_{a(r)}Y(r), B_\rho Y(r)\right\rangle_{\langle X_{-1},X_1\rangle} + \frac{1}{2}\sum_{k=1}^m \left\langle B_\rho S^k_{a(r)}Y(r), S^k_{a(r)}Y(r)\right\rangle_{0,\rho}\right] dr \\
& -2\mathbb{E}\int_t^s \left\langle B_\rho A_{a(r)}Y(r), x\right\rangle_{0,\rho}.
\end{aligned}$$

Therefore, by (3.341),

$$\begin{aligned}
&\mathbb{E}\,|Y(s) - x|^2_{-1,\rho} + 2\bar\lambda\mathbb{E}\int_t^s |Y(r)|^2_{0,\rho}\,dr \\
\le\;& 2K\mathbb{E}\int_t^s |Y(r)|^2_{-1,\rho}\,dr + 2\mathbb{E}\int_t^s |x|_{0,\rho}\left|B_\rho A_{a(r)}Y(r)\right|_{0,\rho} dr
\end{aligned}$$

which, upon using (3.342), (3.345) and straightforward calculations, yields

$$\mathbb{E}\,|Y(s) - x|^2_{-1,\rho} + \bar\lambda\mathbb{E}\int_t^s |Y(r)|^2_{0,\rho}\,dr \le C\,(s-t)\left[|x|^2_{-1,\rho} + |x|^2_{0,\rho}\right]$$

for some constant $C > 0$. This proves (3.349). Estimate (3.350) follows upon noticing that

$$\mathbb{E}\int_t^s |Y(r)-x|_{0,\rho}^2\, dr \le 2\mathbb{E}\int_t^s \left(|Y(r)|_{0,\rho}^2 + |x|_{0,\rho}^2\right) dr \le C\,(s-t)\,|x|_{0,\rho}^2. \quad (3.358)$$

To prove (3.351), we assume by contradiction that it is not satisfied. In this case there are $a_n(\cdot) \in \mathcal{U}_t$ (which we can assume to be $\mathscr{F}_s^{t,0}$-predictable) and $t_n \to t$ such that $\mathbb{E}|Y_n(t_n) - x|_{0,\rho}^2 \not\to 0$ as $n \to +\infty$. Because of Corollary 2.21 and Proposition 3.106 we can assume that all $a_n(\cdot)$ are defined on the same reference probability space. Since $\mathbb{E}\,|Y(t_n)|_{0,\rho}^2$ is bounded, there exists a subsequence, still denoted by $t_n, t_n \to 0$ as $n \to +\infty$, and an element $\bar{Y}$ of $L^2(\Omega; X_0)$ such that, as $n \to +\infty$

$$Y(t_n) \rightharpoonup \bar{Y}, \quad \text{weakly in } L^2(\Omega; X_0)$$

and hence also weakly in $L^2(\Omega; X_{-1})$. Since by (3.349)

$$Y(t_n) \to x, \quad \text{strongly in } L^2(\Omega; X_{-1})$$

as $n \to +\infty$, we obtain $\bar{Y} = x$. This, plus the fact that $\mathbb{E}\,|Y(t_n)|_{0,\rho}^2 \to |x|_{0,\rho}^2$ as $n \to +\infty$ provided by (3.345), implies that $Y(t_n) \to x$, strongly in $L^2(\Omega; X_0)$ which gives a contradiction.

(ii). The existence of the strong solution is known (see [295, 296, 298]) and can be obtained similarly to Proposition 3.106 by applying it to the Gelfand triple (X_2, X_1, X_0). One now obtains

$$\begin{aligned}\mathbb{E}\,|Y(s)|_{1,\rho}^2 \quad = \quad & \mathbb{E}\,|x|_{1,\rho}^2 + 2\mathbb{E}\int_t^s \left\langle A_{a(r)}Y(r), Y(r)\right\rangle_{\langle X_0, X_2\rangle} dr \\ & + \sum_{k=1}^m \mathbb{E}\int_t^s \left\langle S_{a(r)}^k Y(r), S_{a(r)}^k Y(r)\right\rangle_{1,\rho} dr,\end{aligned}$$

which, upon using (3.340) and applying the same arguments as those in the proof of (i), yields (3.352) and (3.353). The proof of the final three estimates is analogous to the similar ones proved in (i). ☐

3.11.5. Viscosity solutions. The definition of solution for (3.336) is similar to the general definition given in Section 3.3. However here we have unbounded first and second order terms so the equation is different. We also make use of the coercivity of the operators A_a.

Definition 3.108 *A function ψ is a test function if $\psi = \varphi + \delta(t)|x|_{0,\rho}^2$, where:*

(i) *$\varphi \in C^{1,2}((0,T)\times X_0)$ is B_ρ-lower semicontinuous, and $\varphi_t \in UC((0,T)\times X_0)$, $D\varphi \in UC((0,T)\times X_0; X_2)$, $D^2\varphi \in BUC((0,T)\times X_0; \mathcal{L}(X_{-1}; X_1))$.*

(ii) *$\delta \in C[0,T] \cap C^1(0,T)$ is such that $\delta > 0$.*

Definition 3.109 *A locally bounded B_ρ-upper (respectively, lower) semicontinuous function $u : (0,T]\times X_0 \to \mathbb{R}$ is a viscosity subsolution (respectively, supersolution) of (3.336) if $u(T,x) \le g(x)$ (respectively, $u(T,x) \ge g(x)$) on X_0 and for every test function ψ, whenever $u - \psi$ (respectively $u + \psi$) has a global maximum (respectively, minimum) at $(t,x) \in (0,T)\times X_0$, then $x \in X_{1,\rho}$ and*

$$\psi_t(t,x) + \inf_{a\in\Lambda}\left\{\frac{1}{2}\sum_{k=1}^m \left\langle D^2\psi(t,x)S_a^k x, S_a^k x\right\rangle_{0,\rho} + \langle A_a x, D\psi(t,x)\rangle_{\langle X_{-1}, X_1\rangle} + f(x,a)\right\} \ge 0,$$

respectively

$$-\psi_t(t,x)+\inf_{a\in\Lambda}\Bigg\{\frac{1}{2}\sum_{k=1}^m\left\langle -D^2\psi(t,x)S_a^kx,S_a^kx\right\rangle_{0,\rho}$$
$$+\langle A_ax,-D\psi(t,x)\rangle_{\langle X_{-1},X_1\rangle}+f(x,a)\Bigg\}\le 0. \quad (3.359)$$

A function is a viscosity solution if it is both a viscosity subsolution and a viscosity supersolution.

The main difference between this and the definition in Section 3.3 is that we require that the point x where the maximum/minimum occurs belongs to a smaller subspace X_1 (of more regular functions. This is possible because of the coercivity of the operators A_a. In this way all terms appearing in the equation are well defined and there is no need to discard any of them. Such definitions originated for first order equations in [71, 106]. Compared to Definition 3.32 we put more conditions on φ and restricted the class of radial functions. The role of A^* in Definition 3.32 is now played by B_ρ^{-1}. The radial test functions are quadratic since we only consider solutions with smaller growth rate at infinity, which is a reasonable assumption for value functions coming from separated problems (see (3.338)). We remark that the definition of viscosity solution here is different from the definition given in [245]) where it was required that $\varphi\in UC^{1,2}\left((0,T)\times X_{-1}\right)$. Both allow to prove the same results. We decided to change the definition to make it more in line with the presentation of the material in the book.

If u has less than quadratic growth in x as $|x|_{0,\rho}\to+\infty$ then the maxima and minima in the definition of solution can be assumed to be strict: if $u-(\varphi+\delta(t)\,|x|_{0,\rho}^2)$ has a global maximum at $(\hat t,\hat x)$ and $\lambda\in C^2([0,+\infty))$ is such that $\lambda>0$, $\lambda(r)=r^4$ if $r\le 1$ and $\lambda(r)=1$ if $r\ge 2$, then it is easy to see that $u-(\varphi+\delta(t)\,|x|_{0,\rho}^2)-\lambda(|x-\hat x|_{-1,\rho})-(t-\hat t)^2$ has a strict global maximum at $(\hat t,\hat x)$.

It is easy to see that $\langle A_ax,D\varphi(t,x)\rangle_{\langle X_{-1},X_1\rangle}=\left\langle B_\rho A_ax,B_\rho^{-1}D\varphi(t,x)\right\rangle_{-,\rho}$ and $\left\langle D^2\varphi(t,x)S_a^kx,S_a^kx\right\rangle_{0,\rho}=\left\langle B_\rho^{-1/2}D^2\varphi(t,x)B_\rho^{-1/2}B_\rho^{1/2}S_a^kx,B_\rho^{1/2}S_a^kx\right\rangle_{0,\rho}$, $k=1,...,m$. Moreover

$$B_\rho^{-1}D\varphi\in UC\left((0,T)\times X_0;X_0\right),\quad B_\rho^{-1/2}D^2\varphi B_\rho^{-1/2}\in BUC\left((0,T)\times X_0;\mathcal{L}(X_0)\right) \quad (3.360)$$

We also remark that Itô's formula holds for the test functions. For the radial part of test function this follows from (3.343) and Itô's formula. As regards φ, the easiest way to see it is to use the fact that if $Y(s):=Y(\cdot;t,x,a(\cdot))$ is the solution of equation (3.334) on $[t,T],x\in X_0$ then $Y(\cdot)$ is in fact a strong solution on any interval $[s,T],s>t$. Thus if φ is a test function as above and $a(\cdot)\in\mathcal{U}_t$ then by the usual Itô's formula we have for $t<s<\eta$

$$\mathbb{E}\varphi(\eta,Y(\eta))=\mathbb{E}\varphi(s,Y(s))+\mathbb{E}\int_s^\eta\Bigg[\varphi_t\left(r,Y\left(r\right)\right)+\left\langle A_{a(r)}Y\left(r\right),D\varphi\left(r,Y\left(r\right)\right)\right\rangle_{0,\rho}$$
$$+\frac{1}{2}\sum_{k=1}^m\left\langle D^2\varphi\left(r,Y\left(r\right)\right)S_{a(r)}^kY\left(r\right),S_{a(s)}^kY\left(r\right)\right\rangle_{0,\rho}\Bigg]dr$$
$$\to\varphi(t,x)+\mathbb{E}\int_t^\eta\Bigg[\varphi_t\left(r,Y\left(r\right)\right)+\left\langle A_{a(r)}Y\left(r\right),D\varphi\left(r,Y\left(r\right)\right)\right\rangle_{\langle X_{-1},X_1\rangle}$$
$$+\frac{1}{2}\sum_{k=1}^m\left\langle D^2\varphi\left(r,Y\left(r\right)\right)S_{a(r)}^kY\left(r\right),S_{a(r)}^kY\left(r\right)\right\rangle_{0,\rho}\Bigg]dr$$

as $s \to t$ using for instance (3.345) and (3.351), since by (3.342)

$$|\varphi_t(r, Y(r))| + \left|\left\langle A_{a(r)} Y(r), D\varphi(r, Y(r))\right\rangle_{\langle X_{-1}, X_1\rangle}\right| \le C\left(1 + |Y(r)|_{0,\rho}^2\right),$$

$$\left|\left\langle D^2\varphi(r, Y(r)) S_{a(r)}^k Y(r), S_{a(r)}^k Y(r)\right\rangle_{0,\rho}\right| \le C|Y(r)|_{0,\rho}^2, \quad k = 1, ..., m.$$

Itô's formulas for test functions from [245] are proved in [364].

We prove comparison principle. We use the notation of Section 3.2: $\{e_n\}_{n=1}^\infty \subset X_0$ is now an orthonormal basis in X^{-1}, $X^N = \mathrm{span}\{e_1, ..., e_N\}$, P_N is the orthogonal projection from X_{-1} onto X^N, $Q_N = I - P_N$, and $Y^N = Q_N X_{-1}$. We have an orthogonal decomposition $X_{-1} = X^N \times Y^N$, and for $x \in X_{-1}$ we write $x = (P_N x, Q_N x)$.

THEOREM 3.110 *Let Hypotheses 3.103 and 3.104 hold. Let $u, v : X_0 \to \mathbb{R}$ be respectively a viscosity subsolution, and a viscosity supersolution of (3.336) (as defined in Definition 3.109). Let*

$$\limsup_{|x|_{0,\rho}\to\infty} \frac{u(t,x)}{|x|_{0,\rho}^2} = 0, \quad \limsup_{|x|_{0,\rho}\to\infty} \frac{-v(t,x)}{|x|_{0,\rho}^2} = 0. \tag{3.361}$$

uniformly for $t \in [0, T]$, and

$$\begin{cases} (i) & \lim_{t\uparrow T}(u(t,x) - g(x))^+ = 0 \\ (ii) & \lim_{t\uparrow T}(v(t,x) - g(x))^- = 0 \end{cases} \tag{3.362}$$

uniformly on bounded subsets of X_0. Then $u \le v$.

PROOF. Without loss of generality we can assume that u and $-v$ are bounded from above and such that

$$\lim_{|x|_{0,\rho}\to\infty} u(t,x) = -\infty, \qquad \lim_{|x|_{0,\rho}\to\infty} v(t,x) = +\infty. \tag{3.363}$$

To see this we notice that if K is the constant from (3.339) then for every $\eta > 0$

$$u_\eta(t,x) = u(t,x) - \eta e^{2K(T-t)}|x|_{0,\rho}^2, \quad v_\eta(t,x) = v(t,x) + \eta e^{2K(T-t)}|x|_{0,\rho}^2$$

are respectively viscosity sub- and supersolutions of (3.336) and satisfy (3.362). This follows from (3.339) since, denoting $h(t,x) = \eta e^{2K(T-t)}|x|_{0,\rho}^2$, we have

$$h_t + \sup_{a\in\Lambda}\left\{\frac{1}{2}\sum_{k=1}^m \left\langle D^2 h S_a^k x, S_a^k x\right\rangle_{0,\rho} + \langle A_a x, Dh\rangle_{\langle X_{-1}, X_1\rangle}\right\} \le -\eta K e^{2K(T-t)}|x|_{0,\rho}^2 \le 0$$

on $(0, T) \times X_0$. The functions $u_\eta, -v_\eta$ satisfy (3.363). Therefore, if we can prove that $u_\eta \le v_\eta$ for every $\eta > 0$ we recover $u \le v$ by letting $\eta \to 0$.

The proof basically follows the lines of the proof of Theorem 3.50. Suppose that $u \not\le v$. Let for $\mu, \epsilon, \delta, \beta > 0$,

$$\Psi(t,s,x,y) := u(t,x) - v(s,y) - \frac{\mu}{t} - \frac{\mu}{t} - \frac{|x-y|_{-1,\rho}^2}{2\epsilon} - \delta(|x|_{0,\rho}^2 + |y|_{0,\rho}^2) - \frac{(t-s)^2}{2\beta}.$$

For every $n \in \mathbb{N}$ there exist p_n, $q_n \in X_0, a_n, b_n \in \mathbb{R}$ such that $|p_n|_{0,\rho}, |q_n|_{0,\rho}, |a_n|, |b_n| \le 1/n$, and

$$\Psi(t,s,x,y) + \langle B_\rho p_n, x\rangle_{0,\rho} + \langle B_\rho q_n, y\rangle_{0,\rho} + a_n t + b_n s$$

has a strict global maximum at $(\bar{t}, \bar{s}, \bar{x}, \bar{y})$ over $(0, T] \times (0, T] \times X_0 \times X_0$. Arguing like in the proof of Theorem 3.50 we obtain

$$\limsup_{\beta\to 0}\limsup_{n\to\infty} \frac{(\bar{t} - \bar{s})^2}{2\beta} = 0 \quad \text{for fixed } \mu, \epsilon, \delta, \tag{3.364}$$

$$|\bar{x}|_{0,\rho} + |\bar{y}|_{0,\rho} \leq R \quad \text{for some } R, \text{ independently of } \mu, \epsilon, \delta, \beta, n, \tag{3.365}$$

and

$$\limsup_{\epsilon\to 0} \limsup_{\delta\to 0} \limsup_{\beta\to 0} \limsup_{n\to\infty} \frac{|\bar{x}-\bar{y}|^2_{-1,\rho}}{2\epsilon} = 0 \quad \text{for fixed } \mu. \tag{3.366}$$

Therefore, it follows from (3.362), (3.363), and (3.364) that $0 < \bar{t}$, and $\bar{s} < T$. We now fix $N \in \mathbb{N}$. Defining

$$u_1(t,x) = u(t,x) - \frac{\mu}{t} - \frac{\langle B_\rho Q_N(\bar{x}-\bar{y}), x\rangle_{0,\rho}}{\epsilon} - \frac{|Q_N(x-\bar{x})|^2_{-1,\rho}}{\epsilon} + \frac{|Q_N(\bar{x}-\bar{y})|^2_{-1,\rho}}{2\epsilon} - \delta |x|^2_{0,\rho} + a_n t + \langle Bp_n, x\rangle_{0,\rho}$$

and

$$v_1(s,y) = v(s,y) + \frac{\mu}{s} - \frac{\langle B_\rho Q_N(\bar{x}-\bar{y}), y\rangle_{0,\rho}}{\epsilon} + \frac{|Q_N(y-\bar{y})|^2_{-1,\rho}}{\epsilon} + \delta |y|^2_{0,\rho} - b_n s - \langle Bq_n, y\rangle_{0,\rho}.$$

we see that

$$u_1(t,x) - v_1(s,y) - \frac{1}{2\epsilon}|P_N(x-y)|^2_{-1,\rho} - \frac{1}{2\beta}|t-s|^2$$

has a strict global maximum at $(\bar{t},\bar{s},\bar{x},\bar{y})$. It now follows from Corollary 3.29 and the proof of Theorem 3.27 that for every $\nu > 1$ there exist test functions φ_i, and ψ_i, $i = 1, 2, ...$, such that

$$u(t,x) - \varphi_i(t,x)$$

has a global maximum at some point (t_i, x_i),

$$v(s,y) - \psi_i(s,y)$$

has a global minimum at some point (s_i, y_i), and

$$\big(t_i, x_i, u(t_i,x_i), D\varphi_i(t_i,x_i), D^2\varphi_i(t_i,x_i)\big) \xrightarrow{k\to\infty} \left(\bar{t}, \bar{x}, u(\bar{t},\bar{x}), \frac{B_\rho(\bar{x}_N - \bar{y}_N)}{\epsilon}, L_N\right) \quad \text{in } \mathbb{R}\times X_0 \times \mathbb{R} \times X_2 \times \mathcal{L}(X_{-1}, X_1), \tag{3.367}$$

$$\big(s_i, y_i, v(s_i,y_i), D\psi_i(s_i,y_i), D^2\psi_i(s_i,y_i)\big) \xrightarrow{k\to\infty} \left(\bar{s}, \bar{y}, v(\bar{s},\bar{y}), \frac{B_\rho(\bar{x}_N - \bar{y}_N)}{\epsilon}, M_N\right) \quad \text{in } \mathbb{R}\times X_0 \times \mathbb{R} \times X_2 \times \mathcal{L}(X_{-1}, X_1), \tag{3.368}$$

where $L_N = P_N^* L_N P_N$, $M_N = P_N^* M_N P_N$ and

$$\begin{pmatrix} L_N & 0 \\ 0 & -M_N \end{pmatrix} \leq \frac{\nu}{\epsilon} \begin{pmatrix} B_\rho P_N & -B_\rho P_N \\ -B_\rho P_N & B_\rho P_N \end{pmatrix}. \tag{3.369}$$

Using the definition of viscosity subsolution we thus obtain

$$\inf_{a\in\Lambda} \Bigg\{ \frac{1}{2}\sum_{k=1}^m \langle (D^2\varphi_i(t_i,x_i) + \frac{2}{\epsilon}B_\rho Q_N + 2\delta I) S_a^k x_i, S_a^k x_i\rangle_{0,\rho} + \langle A_a x_i, D\varphi_i(t_i,x_i) + \frac{B_\rho Q_N(\bar{x}-\bar{y})}{\epsilon} + \frac{2B_\rho Q_N(x_i - \bar{x})}{\epsilon} + 2\delta x_i - B_\rho p_n\rangle_{\langle X_{-1},X_1\rangle} + f(x_i,a)\Bigg\} - a_n + (\varphi_i)_t(t_i,x_i) \geq \frac{\mu}{T^2} \tag{3.370}$$

(We remark that the function $\frac{\mu}{t}$ can be modified around 0 so that it is part of a test function.)

We now pass to the limit in (3.370) as $i \to \infty$. We notice that by (3.339), for every $a \in \Lambda$

$$\sum_{k=1}^{m} \langle S_a^k x_i, S_a^k x_i \rangle_0 + 2\langle A_a x_i, x_i \rangle_{\langle X_{-1}, X_1 \rangle} \le 2\delta K |x_i|_{0,\rho}^2 \to 2\delta K |\bar{x}|_{0,\rho}^2$$

as $i \to \infty$. (In fact one can prove $x_i \rightharpoonup \bar{x}$ in X_1.) Moreover we observe that by (3.360), $B_\rho^{-1/2} D^2 \varphi_i(t_i, x_i) B_\rho^{-1/2} \to B_\rho^{-1/2} L_N B_\rho^{-1/2}$ in $\mathcal{L}(X_0)$. Using these, (3.367), and (3.342), i.e. that $\|B_\rho A_a\|_{\mathcal{L}(X_0)}, \|B_\rho^{\frac{1}{2}} S_a^k\|_{\mathcal{L}(X_0)} \le C$ independently of $a \in \Lambda, 1 \le k \le m$, we obtain upon passing to $\limsup_{i \to \infty}$ in (3.370) that

$$\begin{aligned} &- a_n + \frac{\bar{t} - \bar{s}}{\beta} + \inf_{a \in \Lambda} \Bigg\{ \frac{1}{2} \sum_{k=1}^{m} \langle L_N + \frac{2}{\epsilon} B_\rho Q_N) S_a^k \bar{x}, S_a^k \bar{x} \rangle_{0,\rho} \\ &\quad + \langle A_a \bar{x}, \frac{B_\rho(\bar{x} - \bar{y})}{\epsilon} - B_\rho p_n \rangle_{\langle X_{-1}, X_1 \rangle} + f(\bar{x}, a) \Bigg\} + 2\delta K |\bar{x}|_{0,\rho}^2 \ge \frac{\mu}{T^2}. \end{aligned} \tag{3.371}$$

We obtain similarly for the supersolution v

$$\begin{aligned} &b_n + \frac{\bar{t} - \bar{s}}{\beta} + \inf_{a \in \Lambda} \Bigg\{ \frac{1}{2} \sum_{k=1}^{m} \langle M_N - \frac{2}{\epsilon} B_\rho Q_N) S_a^k \bar{y}, S_a^k \bar{y} \rangle_{0,\rho} \\ &\quad + \langle A_a \bar{y}, \frac{B_\rho(\bar{x} - \bar{y})}{\epsilon} + B_\rho q_n \rangle_{\langle X_{-1}, X_1 \rangle} + f(\bar{y}, a) \Bigg\} - 2\delta K |\bar{y}|_{0,\rho}^2 \le -\frac{\mu}{T^2}. \end{aligned} \tag{3.372}$$

By Hypothesis 3.103 the closures of the sets $\{S_a^k \bar{x} : a \in \Lambda, 1 \le k \le m\}$ and $\{S_a^k \bar{y} : a \in \Lambda, 1 \le k \le m\}$ are compact in X_0, and hence in X_{-1}. This yields

$$\sup\{|B_\rho^{1/2} Q_N S_a^k \bar{x}|_{0,\rho} + |B_\rho^{1/2} Q_N S_a^k \bar{y}|_{0,\rho} : a \in \Lambda, 1 \le k \le m\} \to 0 \tag{3.373}$$

as $N \to \infty$. Moreover, (3.369) implies that

$$\langle L_N S_a^k \bar{x}, S_a^k \bar{x} \rangle_{0,\rho} - \langle M_N S_a^k \bar{y}, S_a^k \bar{y} \rangle_{0,\rho} \le \frac{\nu}{2\epsilon} \langle B_\rho S_a^k (\bar{x} - \bar{y}), S_a^k (\bar{x} - \bar{y}) \rangle_{0,\rho}. \tag{3.374}$$

Therefore, subtracting (3.371) from (3.372) and using (3.335), (3.373), and (3.374), we have

$$\begin{aligned} &\inf_{a \in \Lambda} \Bigg\{ -\frac{\nu}{2\epsilon} \sum_{k=1}^{m} \langle B_\rho S_a^k (\bar{x} - \bar{y}), S_a^k (\bar{x} - \bar{y}) \rangle_{0,\rho} - \frac{1}{\epsilon} \langle A_a (\bar{x} - \bar{y}), B_\rho (\bar{x} - \bar{y}) \rangle_{\langle X_{-1}, X_1 \rangle} \Bigg\} \\ &+ a_n + b_n - \omega_R(|\bar{x} - \bar{y}|_0) - 2\delta K (|\bar{x}|_{0,\rho}^2 + |\bar{y}|_{0,\rho}^2) - \sigma(1/N, n) \le -\frac{2\mu}{T^2} \end{aligned}$$

for some local modulus σ. Now, if ν is is close to 1, it follows from (3.341) that

$$\begin{aligned} &a_n + b_n + \frac{\bar{\lambda}}{2\epsilon} |\bar{x} - \bar{y}|_{0,\rho}^2 - \frac{K}{\epsilon} |\bar{x} - \bar{y}|_{-1,\rho}^2 \\ &- \omega_R(|\bar{x} - \bar{y}|_{0,\rho}) - 2\delta K (|\bar{x}|_{0,\rho}^2 + |\bar{y}|_{0,\rho}^2) - \sigma(1/N, n) \le -\frac{2\mu}{T^2}. \end{aligned} \tag{3.375}$$

Since ω_R is a modulus we have

$$\lim_{\epsilon \to 0} \inf_{r \ge 0} \left(\frac{\bar{\lambda}}{2\epsilon} r^2 - \omega_R(r) \right) = 0. \tag{3.376}$$

Therefore we obtain a contradiction in (3.375) after sending $N \to \infty$, $n \to \infty$, $\beta \to 0$, $\delta \to 0$, $\epsilon \to 0$ in the above order, and using (3.365), (3.366), and (3.376). $\square$

3.11.6. Value function and existence of solutions. In this subsection we show that the value function V defined by (3.337) is the unique viscosity solution of (3.336).

PROPOSITION 3.111 *Assume that Hypotheses 3.103 and 3.104 are satisfied. Then for every $R > 0$ there exists a modulus σ_R such that*

$$|V(t,x)-V(s,y)| \le \sigma_R(|t-s|+|x-y|_{-1,\rho}), \quad t,s \in [0,T], |x|_{0,\rho}, |y|_{0,\rho} \le R, \tag{3.377}$$

and there is $C > 0$ such that

$$|V(t,x)| \le C\left(1+|x|^\gamma_{0,\rho}\right), \quad t \in [0,T], x \in X_0. \tag{3.378}$$

Moreover the Dynamic Programming Principle (2.25) is satisfied.

PROOF. We only sketch the proof since it is very similar to the proof of Proposition 3.61. Using similar arguments it follows from (3.345), (3.347), (3.348), linearity of the DMZ equation (3.334), and Hypothesis 3.104 that (3.378) holds and there are moduli σ^1_R such that

$$|J(t,x;a(\cdot))-J(t,y;a(\cdot))| \le \sigma^1_R(|x-y|_{-1,\rho}), \quad t \in [0,T], |x|_{0,\rho}, |y|_{0,\rho} \le R, a(\cdot) \in \mathcal{U}_t, \tag{3.379}$$

and thus the same inequality is satisfied by V. We now claim that the Dynamic Programming Principle holds. To do this we need to check that the assumptions of Hypothesis 2.12 are satisfied. Parts $(A0)$ and $(A2)$ follow from the definition of variational solution, properties of stochastic integrals and standard manipulations. Part $(A1)$ was proved in Proposition 3.106 (i.e. Theorem 3.102-(ii)). Part $(A3)$ can be proved similarly to the proof of Proposition 2.16. It thus follows from Theorem 2.24 that for every $x \in X_0, 0 \le t \le \eta \le T$,

$$V(t,x) = \inf_{a(\cdot)\in\mathcal{U}_t} \mathbb{E}\left[\int_t^\eta l(Y(s),a(s))ds + V(\eta, Y(\eta))\right]. \tag{3.380}$$

Using (3.380) we again argue as in the proof of Proposition 3.61 using (3.345), (3.349), (3.378), and (3.379) to obtain that there exist moduli σ^2_R such that

$$|V(t,x)-V(s,x)| \le \sigma^2_R(|t-s|), \quad t,s \in [0,T], |x|_{0,\rho} \le R. \tag{3.381}$$

Obviously (3.379) and (3.381) produce (3.377). □

THEOREM 3.112 *Assume that Hypotheses 3.103 and 3.104 are true. Then the value function V is the unique viscosity solution of the HJB equation (3.336) among functions satisfying*

$$\limsup_{|x|_{0,\rho}\to\infty} \frac{|u(t,x)|}{|x|^2_{0,\rho}} = 0 \quad \textit{uniformly for } t \in [0,T],$$

and

$$\lim_{t\uparrow T} |u(t,x)-g(x)| = 0 \quad \textit{uniformly on bounded subsets of } X_0.$$

PROOF. The uniqueness is a consequence of Theorem 3.110 and Proposition 3.111. Therefore it remains to show that V is a viscosity solution of (3.336).

We only argue about the supersolution property as the subsolution part is easier. Suppose that $V + (\varphi + \delta(t)\,|x|^2_{0,\rho})$ has a global minimum at $(t_0,x_0) \in (0,T)\times X_0$.

We need to prove that $x_0 \in X_1$. For every $(t,x) \in (0,T)\times X_0$

$$V(t,x) - V(t_0,x_0) \ge -(\varphi(t,x)-\varphi(t_0,x_0)) - \left(\delta(t)\,|x|^2_{0,\rho} - \delta(t_0)\,|x_0|^2_{0,\rho}\right). \tag{3.382}$$

By the dynamic programming principle for every $\epsilon > 0$ there exists and $a_\epsilon(\cdot) \in \mathcal{U}_{t_0}$ such that, writing $Y_\epsilon(s)$ for $Y(s; t_0, x_0, a_\epsilon(\cdot))$, we have

$$V(t_0, x_0) + \epsilon^2 > \mathbb{E}\left[\int_{t_0}^{t_0+\epsilon} l(Y_\epsilon(s), a_\epsilon(s))\, ds + V(t_0+\epsilon, Y_\epsilon(t_0+\epsilon))\right].$$

In light of Corollary 2.21 and Proposition 3.106, without loss of generality we can assume that all $a_\epsilon(\cdot)$ are defined on the same reference probability space. We have, by (3.382),

$$\epsilon^2 - \mathbb{E}\int_{t_0}^{t_0+\epsilon} l(Y_\epsilon(s), a_\epsilon(s))\, ds \geq \mathbb{E}\left[V(t_0+\epsilon, Y_\epsilon(t_0+\epsilon)) - V(t_0, x_0)\right]$$

$$\geq -\mathbb{E}\left[\varphi(t_0+\epsilon, Y_\epsilon(t_0+\epsilon)) - \varphi(t_0, x_0)\right] - \mathbb{E}\left[\delta(t_0+\epsilon)\, |Y_\epsilon(t_0+\varepsilon)|_{0,\rho}^2 - \delta(t_0)\, |x_0|_{0,\rho}^2\right]$$

and, by Itô's formula

$$\begin{aligned}
&\epsilon - \mathbb{E}\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon} l(Y_\epsilon(s), \alpha_\epsilon(s))\, ds \\
&\geq -\mathbb{E}\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon}\Bigg[\varphi_t(s, Y_\epsilon(s)) + \left\langle A_{a_\epsilon(s)} Y_\epsilon(s), D\varphi(s, Y_\epsilon(s))\right\rangle_{\langle X_{-1}, X_1\rangle} \\
&\qquad + \frac{1}{2}\sum_{k=1}^m \left\langle D^2\varphi(s, Y_\epsilon(s))\, S^k_{a_\epsilon(s)} Y(s), S^k_{a_\epsilon(s)} Y_\epsilon(s)\right\rangle_{0,\rho}\Bigg] ds \qquad (3.383) \\
&\quad - \mathbb{E}\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon}\Bigg[\delta'(s)\, |Y_\epsilon(s)|_{0,\rho}^2 + 2\delta(s)\Bigg[\left\langle A_{a_\epsilon(s)} Y_\epsilon(s), Y_\epsilon(s)\right\rangle_{\langle X_{-1}, X_1\rangle} \\
&\qquad + \frac{1}{2}\sum_{k=1}^m \left\langle S^k_{a_\epsilon(s)} Y_\epsilon(s), S^k_{a_\epsilon(s)} Y_\epsilon(s)\right\rangle_{0,\rho}\Bigg]\Bigg] ds.
\end{aligned}$$

By (3.339) we have

$$\begin{aligned}
&-2\delta(s)\left[\left\langle A_{a_\epsilon(s)} Y_\epsilon(s), Y_\epsilon(s)\right\rangle_{\langle X_{-1}, X_1\rangle} + \frac{1}{2}\sum_{k=1}^m \left\langle S^k_{a_\epsilon(s)} Y_\epsilon(s), S^k_{a_\epsilon(s)} Y_\epsilon(s)\right\rangle_{0,\rho}\right] \\
&\qquad \geq 2\delta(s)\left[\bar{\lambda}\, |Y_\epsilon(s)|_{1,\rho}^2 - K\, |Y_\epsilon(s)|_{0,\rho}^2\right].
\end{aligned}$$

Moreover

$$|\varphi_t(s, Y_\epsilon(s))| + \left|\left\langle A_{a_\epsilon(s)} Y_\epsilon(s), D\varphi(s, Y_\epsilon(s))\right\rangle_{\langle X_{-1}, X_1\rangle}\right| \leq C_1(1 + |Y_\epsilon(s)|_{0,\rho}^2,$$

$$\left|\sum_{k=1}^m \left\langle D^2\varphi(s, Y_\epsilon(s))\, S^k_{a_\epsilon(s)} Y_\epsilon(s), S^k_{a_\epsilon(s)} Y_\epsilon(s)\right\rangle_{0,\rho}\right| \leq C_2\, |Y_\epsilon(s)|_0^2$$

Therefore, using Hypothesis 3.104, (3.345), and $\delta(s) \geq \gamma > 0$ for s close to t_0 for some $\gamma > 0$, we obtain

$$2\bar{\lambda}\gamma\mathbb{E}\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon} |Y_\epsilon(s)|_{1,\rho}^2\, ds \leq C_3\left[1 + \mathbb{E}\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon} |Y_\epsilon(s)|_{0,\rho}^2\, ds\right] \leq C_4.$$

Take now $\epsilon = 1/n$ and set $Y_n(s) := Y\left(s; t_0, x_0, a_{1/n}\right)$. The above inequality yields

$$n\int_{t_0}^{t_0+1/n} \mathbb{E}\, |Y_n(s)|_1^2\, ds \leq C_5$$

so that, along a sequence $t_n \in (t_0, t_0 + 1/n)$

$$\mathbb{E}\, |Y_n(t_n)|_{1,\rho}^2 \leq C_5$$

and thus, along a subsequence, still denoted by t_n, we have

$$Y_n(t_n) \rightharpoonup \bar{Y}$$

weakly in $L^2(\Omega; X_1)$ for some $\bar{Y} \in L^2(\Omega; X_1)$. This also clearly implies weak convergence in $L^2(\Omega; X_0)$. However, by (3.351), $Y_n(t_n) \to x_0$ strongly (and weakly) in $L^2(\Omega; X_0)$. Thus $\bar{Y} = x_0 \in X_1$.

Having established that $x_0 \in X_1$, we now go back to (3.383), use the properties of test functions and estimates of Proposition 3.107 to obtain

$$\begin{aligned}
\epsilon \geq & -\mathbb{E}\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon}\Bigg[\varphi_t(t_0,x_0) + \left\langle A_{a_\epsilon(s)}x_0, D\varphi(t_0,x_0)\right\rangle_{\langle X_{-1},X_1\rangle} \\
& +\frac{1}{2}\sum_{k=1}^{m}\left\langle D^2\varphi(t_0,x_0)\, S^k_{a_\epsilon(s)}x_0, S^k_{a_\epsilon(s)}x_0\right\rangle_{0,\rho}\Bigg]ds \\
& -\mathbb{E}\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon}\Bigg[\delta'(t_0)\,|x_0|^2_{0,\rho} + 2\delta(t_0)\Bigg[\left\langle A_{a_\epsilon(s)}x_0, x_0\right\rangle_{\langle X_{-1},X_1\rangle} \\
& +\frac{1}{2}\sum_{k=1}^{m}\left\langle S^k_{a_\epsilon(s)}x_0, S^k_{a_\epsilon(s)}x_0\right\rangle_{0,\rho}\Bigg] - l\left(x_0,\alpha_\epsilon(s)\right)\Bigg]ds - \gamma(\epsilon), \\
\geq & \ \mathbb{E}\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon}\Bigg[-\psi_t(t,x) + \inf_{a\in\Lambda}\Bigg\{\frac{1}{2}\sum_{k=1}^{m}\left\langle -D^2\psi(t,x)S^k_a x, S^k_a x\right\rangle_{0,\rho} \\
& +\langle A_a x, -D\psi(t,x)\rangle_{\langle X_{-1},X_1\rangle} + f(x,a)\Bigg\}\Bigg]ds - \gamma(\epsilon),
\end{aligned}$$

where $\lim_{\epsilon\to 0}\gamma(\epsilon) = 0$. It remains to let $\epsilon \to 0$. We refer the reader to the proof of Theorem 5.4 in [245] for more details. □

3.12. HJB equation for a boundary control problem

In this section we discuss how the theory of viscosity solutions can be applied to solve HJB equations coming from stochastic boundary control problems discussed in Section 2.6.2. We will only consider time independent problems. Suppose that H is a real, separable Hilbert space and A is an operator in H satisfying the following hypothesis.

HYPOTHESIS 3.113 *$A : D(A) \subset H \to H$ is a (densely defined) self-adjoint operator, there exists $a > 0$ such that $\langle Ax, x\rangle \leq -a|x|^2$ for all $x \in D(A)$ and A^{-1} is compact.*

Hypothesis 3.113 implies in particular that A is the infinitesimal generator of an analytic semigroup with compact resolvent satisfying $\|e^{tA}\| \leq e^{-at}$ for all $t \geq 0$ and that there is an orthonormal basis of H composed of eigenvectors of A such that the corresponding sequence of (positive) eigenvalues diverges to $+\infty$ as $n \to \infty$. Moreover the fractional powers $(-A)^\gamma, \gamma > 0$ are well defined, and if $\gamma \in (0,1]$ and $\alpha \in (0,\gamma)$, a well known interpolation inequality (see e.g. [375], p. 73-74) gives us that for every $\sigma > 0$ there exists $C_\sigma > 0$ such that

$$|(-A)^\alpha x| \leq \sigma|(-A)^\gamma x| + C_\sigma|x|, \qquad \text{for every } x \in D((-A)^\gamma). \tag{3.384}$$

The HJB equations introduced in Section 2.6.2 (see (2.105)) have the form

$$\lambda v - \langle Ax, Dv\rangle + F(x, Dv, D^2v) = 0, \quad x \in H, \tag{3.385}$$

where $F : Z \subset H \times H \times S(H) \to \mathbb{R}$, $\lambda > 0$. In particular $F(x,p,X)$ may only be defined if $p \in D((-A)^\beta)$ for some $\beta > 0$ and may be undefined if X is not of trace class. The unboundedness in the first order terms comes from the boundary control term rewritten as a distributed control terms, and the unboundedness in the second derivative terms comes from noise with non-nuclear covariance in the control problem. Such second order unboundedness has not been discussed so far in this book and indeed it is not easy to handle by the viscosity solution methods. Here we suggest one way to do it. The idea is the following. To deal with the unboundedness in the first and second derivatives we introduce a change of variables $x = (-A)^{\frac{\beta}{2}} y, \beta > 0$. Then the function $u(y) := v((-A)^{\frac{\beta}{2}} y)$ should formally solve

$$\lambda u - \langle Ay, Dv\rangle + F((-A)^{\frac{\beta}{2}} y, (-A)^{-\frac{\beta}{2}} Dv, (-A)^{-\frac{\beta}{2}} D^2 v(-A)^{-\frac{\beta}{2}}) = 0. \qquad (3.386)$$

This equation contains less unbounded terms and is easier to handle in spite of the additional difficulty created by the presence of the new unbounded term $A^{\frac{\beta}{2}} y$. We will define a viscosity solution of (3.385) to be a function v such that $u(\cdot) \stackrel{def}{=} v(A^{\frac{\beta}{2}} \cdot)$ is a viscosity solution of (3.386). We will make this idea rigorous in the next section. The definition is meaningful, indeed, when (3.385) comes from a stochastic boundary control problem, v and u can be respectively characterized as the value functions of their control problems.

3.12.1. Definition of viscosity solution. We first consider the following HJB equation

$$\lambda u - \langle Ay, Du\rangle + G(y, Du, D^2 u) = 0, \qquad y \in H, \qquad (3.387)$$

where $G : D((-A)^{\frac{\beta}{2}}) \times D((-A)^{\frac{\beta}{2}}) \times S(H) \to \mathbb{R}$.

DEFINITION 3.114 *We say that a function ψ is a test function if $\psi(x) = \varphi(x) + \delta|x|^2$, where $\delta > 0$ and*

(i) $\varphi \in C^2(H)$ *and is weakly sequentially lower semicontinuous on H.*
(ii) $D\varphi \in UC(H,H) \cap UC\left(D((-A)^{\frac{1}{2}-\epsilon}), D((-A)^{\frac{1}{2}})\right)$ *for some $\epsilon = \epsilon(\varphi) > 0$.*
(iii) $D^2\varphi \in BUC(H, S(H))$.

DEFINITION 3.115 *We say that a function $w : H \to \mathbb{R}$ is a viscosity subsolution of* (3.387) *if w is weakly sequentially upper semicontinuous on H, and whenever $w - \psi$ has a local maximum at x for a test function ψ, then*

$$x \in D((-A)^{\frac{1}{2}})$$

and

$$\lambda w(x) + \langle (-A)^{\frac{1}{2}} x, (-A)^{\frac{1}{2}} D\varphi(x)\rangle + 2\delta |(-A)^{\frac{1}{2}} x|^2 + G\left(x, D\psi(x), D^2\psi(x)\right) \le 0.$$

We say that w is a viscosity supersolution of (3.387) *if w is weakly sequentially lower semicontinuous on H, and whenever $w + \psi$ has a local minimum at x for a test function ψ, then*

$$x \in D((-A)^{\frac{1}{2}})$$

and

$$\lambda w(x) - \langle (-A)^{\frac{1}{2}} y, (-A)^{\frac{1}{2}} D\varphi(x)\rangle - 2\delta |(-A)^{\frac{1}{2}} x|^2 + G\left(x, -D\psi(x), -D^2\psi(x)\right) \ge 0.$$

We say that w is a viscosity solution of (3.387) *if it is both a viscosity subsolution and a supersolution.*

Suppose now that F from (3.385) is such that $F : H \times D((-A)^{\beta}) \times (-A)^{-\frac{\beta}{2}} S(H)(-A)^{-\frac{\beta}{2}} \to \mathbb{R}$. We define

$$G_F(z,p,S) \stackrel{def}{=} F\left((-A)^{\frac{\beta}{2}} z, (-A)^{-\frac{\beta}{2}} p, (-A)^{-\frac{\beta}{2}} S (-A)^{-\frac{\beta}{2}}\right), \qquad (3.388)$$

DEFINITION 3.116 *A bounded continuous function $v : H \to \mathbb{R}$ is said to be a viscosity solution of equation (3.385) if the function*

$$u(y) \stackrel{def}{=} v((-A)^{\frac{\beta}{2}} y)$$

is a viscosity solution of the equation

$$\lambda u - \langle Ay, Du\rangle + G_F(y, Du, D^2u) = 0, \qquad y \in H. \tag{3.389}$$

Similarly we define a viscosity subsolution and a supersolution of (3.385).

We remark that the function v is uniquely determined once u has been characterized on $D((-A)^{\frac{\beta}{2}})$.

3.12.2. Comparison and existence theorem. For $\gamma > 0$ we denote by $H_{-\gamma}$ the completion of H in the norm $|x|_{-\gamma} = |(-A)^{-\frac{\gamma}{2}}x|$, and for $0 < \gamma \leq 2$, and $D((-A)^{\frac{\gamma}{2}})$ is equipped with the norm $|x|_\gamma = |(-A)^{\frac{\gamma}{2}}x|$. For $N > 2$ let H_N be finite dimensional subspaces of H generated by eigenvectors of $(-A)^{-1}$ corresponding to the eigenvalues which are greater than or equal to $1/N$. Denote by P_N the orthogonal projection in H_{-1} onto H_N, $Q_N = I - P_N$, and $H_N^\perp = Q_N H$. P_N and Q_N are also orthogonal projections in H. We then have an orthogonal decomposition $H = H_N \times H_N^\perp$ and we will write $x = (x_N, x_N^\perp) = (P_N x, Q_N x)$.

We assume:

HYPOTHESIS 3.117

(i) *There exists $\beta \in (0,1)$ such that the function $G : D((-A)^{\frac{\beta}{2}}) \times D((-A)^{\frac{\beta}{2}}) \times S(H) \to \mathbb{R}$ is uniformly continuous (in the topology of $D((-A)^{\frac{\beta}{2}}) \times D((-A)^{\frac{\beta}{2}}) \times S(H)$) on bounded sets of $D((-A)^{\frac{\beta}{2}}) \times D((-A)^{\frac{\beta}{2}}) \times S(H)$.*

(ii) *$G(y,p,S_1) \leq G(y,p,S_2)$ if $S_1 \geq S_2$, for all $y, p \in D((-A)^{\frac{\beta}{2}})$.*

(iii) *There exists a modulus ρ such that*

$$|G(y,p,S_1) - G(y,q,S_2)| \\ \leq \rho\Big((1 + |(-A)^{\frac{\beta}{2}}y|)|(-A)^{\frac{\beta}{2}}(p-q)| + (1 + |(-A)^{\frac{\beta}{2}}y|^2)\|S_1 - S_2\|\Big)$$

for all $y, p, q \in D((-A)^{\frac{\beta}{2}})$ and $S_1, S_2 \in S(H)$.

(iv) *There exist $0 < \eta < 1 - \beta$ and a modulus ω such that, for all $N > 2$, $\epsilon > 0$,*

$$G\bigg(x, \frac{(-A)^{-\eta}(x-y)}{\epsilon}, Z\bigg) - G\bigg(y, \frac{(-A)^{-\eta}(x-y)}{\epsilon}, Y\bigg) \\ \geq -\omega\bigg(|(-A)^{\frac{\beta}{2}}(x-y)|\bigg(1 + \frac{|(-A)^{\frac{\beta}{2}}(x-y)|}{\epsilon}\bigg)\bigg)$$

for all $x, y \in D((-A)^{\frac{\beta}{2}})$ and $Z, Y \in S(H)$, $Z = P_N Z P_N$, $Y = P_N Y P_N$ such that

$$\begin{pmatrix} Z & 0 \\ 0 & -Y \end{pmatrix} \leq \frac{3}{\epsilon} \begin{pmatrix} (-A)^{-\eta}P_N & -(-A)^{-\eta}P_N \\ -(-A)^{-\eta}P_N & (-A)^{-\eta}P_N \end{pmatrix} \tag{3.390}$$

(v) *Let η be from (iv). For every $R < +\infty, |\lambda| \le R,\ p, x \in D((-A)^{\frac{\beta}{2}})$*

$$\sup\left\{|G(x,p,S+\lambda Q_N) - G(x,p,S)| : \ \|S\| \le R, S = P_N S P_N\right\} \to 0. \tag{3.391}$$

as $N \to \infty$.

Some of the conditions of Hypothesis 3.117 can be weakened. By the properties of moduli, Hypothesis 3.117-(*iii*) guarantees that there exists a constant C such that, for every y, p, S,

$$|G(y,p,S)| \le C\left(1 + (1 + |(-A)^{\frac{\beta}{2}}y|)|(-A)^{\frac{\beta}{2}}p| + (1 + |(-A)^{\frac{\beta}{2}}y|^2)\|S\|\right) + |G(y,0,0)|, \tag{3.392}$$

and conditions (*i*), (*iv*) of Hypothesis 3.117 imply that there is C_1 such that

$$|G(y,0,0)| \le C_1\left(1 + \left|(-A)^{\frac{\beta}{2}}y\right|\right). \tag{3.393}$$

Theorem 3.118 *Let Hypotheses 3.113 and 3.117 be satisfied. Then:*
Comparison: Let $u, -v \le M$ for some constant M. If u is a viscosity subsolution of (3.387) and v is a viscosity supersolution of (3.387) then $u \le v$ on H. Moreover, if u is a viscosity solution then

$$|u(x) - u(y)| \le m(|(-A)^{-\frac{\eta}{2}}(x-y)|) \tag{3.394}$$

for all $x, y \in H$ and some modulus m, where η is the constant in (iv).
Existence: If

$$\sup_{x \in D(A^{\frac{\beta}{2}})} |G(x,0,0)| = K < \infty, \tag{3.395}$$

then there exists a unique viscosity solution $u \in BUC(H_{-\eta})$ of (3.387).

Proof. *Comparison.* Let $\epsilon, \delta > 0$. We set

$$\Phi(x,y) = u(x) - v(y) - \frac{|(-A)^{-\frac{\eta}{2}}(x-y)|^2}{2\epsilon} - \frac{\delta}{2}|x|^2 - \frac{\delta}{2}|y|^2$$

Since $u - v$ is bounded from above and weakly sequentially upper-semicontinuous in $H \times H$, Φ must attain its maximum at some point $(\bar{x}, \bar{y}) \in D((-A)^{\frac{1}{2}}) \times D((-A)^{\frac{1}{2}})$ (which can be assumed to be strict by subtracting for instance $\mu(|(-A)^{-1}(x-\bar{x})|^2 + |(-A)^{-1}(y-\bar{y})|^2)$ and then letting $\mu \to 0$). Moreover, arguing similarly as in the proof of Theorem 3.56, we have

$$\lim_{\delta \to 0}\left(\delta|\bar{x}|^2 + \delta|\bar{y}|^2\right) = 0 \quad \text{for every fixed } \epsilon > 0, \tag{3.396}$$

$$\lim_{\epsilon \to 0}\limsup_{\delta \to 0}\left(\frac{|(-A)^{-\frac{\eta}{2}}(\bar{x}-\bar{y})|^2}{\epsilon}\right) = 0. \tag{3.397}$$

Then (see the proof of Theorem 3.50), defining

$$\begin{aligned} u_1(x) = u(x) - \frac{\langle x, Q_N(-A)^{-\eta}Q_N(\bar{x}-\bar{y})\rangle}{\epsilon} &+ \frac{\langle Q_N(-A)^{-\eta}Q_N(\bar{x}-\bar{y}), \bar{x}-\bar{y}\rangle}{2\varepsilon} \\ &- \frac{|(-A)^{-\frac{\eta}{2}}Q_N(x-\bar{x})|^2}{\epsilon} - \frac{\delta}{2}|x|^2, \end{aligned}$$

$$v_1(y) = v(y) - \frac{\langle y, Q_N(-A)^{-\eta}Q_N(\bar{x}-\bar{y})\rangle}{\varepsilon} + \frac{|(-A)^{-\frac{\eta}{2}}Q_N(y-\bar{y})|^2}{\epsilon} + \frac{\delta}{2}|y|^2,$$

it follows that the function

$$\widetilde{\Phi}(x,y) \stackrel{def}{=} u_1(x) - v_1(y) - \frac{|(-A)^{-\frac{\eta}{2}}P_N(x-y)|^2}{2\epsilon} \tag{3.398}$$

always satisfies $\widetilde{\Phi} \le \Phi$ and $\widetilde{\Phi}$ attains a strict global maximum at $(\bar{x}, \bar{y})$, where $\widetilde{\Phi}(\bar{x}, \bar{y}) = \Phi(\bar{x}, \bar{y})$. Using Corollary 3.28 (with $B = (-A)^{-\eta}$) we thus obtain functions $\varphi_n, -\psi_n$ satisfying conditions $(i) - (iii)$ of Definition 3.114 such that

$$u_1(x) - \varphi_n(x)$$

has a global maximum at some point x^n,

$$v_1(y) + \psi_n(y)$$

has a global minimum at some point y^n, and

$$\big(x^n, u_1(x^n), D\varphi_n(x^n), D^2\varphi_n(x^n)\big) \xrightarrow{n\to\infty} \left(\bar{x}, u_1(\bar{x}), \frac{(-A)^{-\eta}P_N(\bar{x}-\bar{y})}{\epsilon}, Z_N\right) \quad \text{in } H \times \mathbb{R} \times D(-A) \times \mathcal{L}(H;H), \tag{3.399}$$

$$\big(y^n, v_1(y^n), -D\psi_n(y^n), -D^2\psi_n(y^n)\big) \xrightarrow{n\to\infty} \left(\bar{y}, v(\bar{y}), \frac{(-A)^{-\eta}P_N(\bar{x}-\bar{y})}{\epsilon}, Y_N\right) \quad \text{in } H \times \mathbb{R} \times D(-A) \times \mathcal{L}(H;H), \tag{3.400}$$

for some $Z_N, Y_N \in S(H)$ such that $Z_N = P_N Z_N P_N$, $Y_N = P_N Y_N P_N$, they satisfy (3.390) and $\|Z_N\| + \|Y_N\| \le C_\epsilon$ for some constant C_ϵ.

Therefore, by the definition of viscosity subsolution, $x^n \in D((-A)^{\frac{1}{2}})$ and

$$\begin{aligned}
\lambda u(x^n) + \Bigg\langle (-A)^{\frac{1}{2}}x^n, (-A)^{\frac{1}{2}}D\varphi_n(x^n) + \frac{(-A)^{\frac{1}{2}-\eta}Q_N(\bar{x}-\bar{y})}{\epsilon} \\
+ \frac{2(-A)^{\frac{1}{2}-\eta}Q_N(x^n-\bar{x})}{\epsilon}\Bigg\rangle + \delta|(-A)^{\frac{1}{2}}x^n|^2 \\
+G\Bigg(x^n, D\varphi_n(x^n) + \frac{(-A)^{-\eta}Q_N(\bar{x}-\bar{y})}{\epsilon} + \frac{2(-A)^{-\eta}Q_N(x^n-\bar{x})}{\epsilon} + \delta x^n, \\
D^2\varphi_n(x^n) + \frac{2\|(-A)^{-\eta}\|Q_N}{\epsilon} + \delta I\Bigg) \le 0.
\end{aligned} \tag{3.401}$$

Thus, using (3.399), (3.392), (3.393), and (3.384), it follows from (3.401) that $|(-A)^{\frac{1}{2}}x^n|$ are bounded independently of n which implies, thanks to (3.399) that

$$(-A)^{\frac{1}{2}}x^n \rightharpoonup (-A)^{\frac{1}{2}}\bar{x} \quad \text{as } n \to +\infty. \tag{3.402}$$

Since $(-A)^{\frac{\beta-1}{2}}$ and $(-A)^{-\frac{\eta}{2}}$ are compact we conclude that, as $n \to +\infty$,

$$(-A)^{\frac{\beta}{2}}x^n = (-A)^{\frac{\beta-1}{2}}((-A)^{\frac{1}{2}}x^n) \to (-A)^{\frac{\beta}{2}}x \quad \text{and } (-A)^{\frac{1-\eta}{2}}x^n \to (-A)^{\frac{1-\eta}{2}}\bar{x}. \tag{3.403}$$

Using (3.399), (3.402), (3.403), and the weak sequential lower semicontinuity of the norm we thus obtain

$$\begin{aligned}
&\left\langle (-A)^{\frac{1-\eta}{2}}\bar{x}, \frac{(-A)^{\frac{1-\eta}{2}}(\bar{x}-\bar{y})}{\epsilon}\right\rangle + \delta|(-A)^{\frac{1}{2}}\bar{x}|^2 \\
&\le \liminf_{n\to\infty}\Bigg[\Bigg\langle (-A)^{\frac{1}{2}}x^n, (-A)^{\frac{1}{2}}D\varphi_n(x^n) + \frac{(-A)^{\frac{1}{2}-\eta}Q_N(\bar{x}-\bar{y})}{\epsilon} \\
&\qquad + \frac{2(-A)^{\frac{1}{2}-\eta}Q_N(x^n-\bar{x})}{\epsilon}\Bigg\rangle + \delta|(-A)^{\frac{1}{2}}x^n|^2\Bigg]
\end{aligned}$$

and then letting $n \to \infty$ in (3.401) yields

$$\lambda u(\bar{x}) + \left\langle (-A)^{\frac{1-\eta}{2}} \bar{x}, \frac{(-A)^{\frac{1-\eta}{2}}(\bar{x} - \bar{y})}{\epsilon} \right\rangle + \delta |(-A)^{\frac{1}{2}} \bar{x}|^2$$
$$+ G\left(\bar{x}, \frac{(-A)^{-\eta}(\bar{x} - \bar{y})}{\epsilon} + \delta \bar{x}, Z_N + \frac{2\|(-A)^{-\eta}\| Q_N}{\epsilon} + \delta I\right) \le 0. \tag{3.404}$$

Using Hypothesis 3.117-(iii) we have

$$G\left(\bar{x}, \frac{(-A)^{-\eta}(\bar{x} - \bar{y})}{\epsilon}, Z_N + \frac{2\|(-A)^{-\eta}\| Q_N}{\epsilon}\right) - \rho\left(c\delta(1 + |(-A)^{\frac{\beta}{2}} \bar{x}|^2)\right)$$
$$\le G\left(\bar{x}, \frac{(-A)^{-\eta}(\bar{x} - \bar{y})}{\epsilon} + \delta \bar{x}, Z_N + \frac{2\|(-A)^{-\eta}\| Q_N}{\epsilon} + \delta I\right) \tag{3.405}$$

for some constant $c > 0$. Now, given $\tau > 0$, let K_τ be such that $\rho(s) \le \tau + K_\tau s$. Applying (3.384) with $\alpha = \beta/2$ and $\gamma = 1/2$ we obtain

$$\rho\left(c\delta(1 + |(-A)^{\frac{\beta}{2}} \bar{x}|^2)\right) \le \delta |(-A)^{\frac{1}{2}} \bar{x}|^2 + \delta C_\tau |\bar{x}|^2 + \tau + K_\tau d\delta$$

for some constant $C_\tau > 0$ independent of δ and ϵ. It then follows from (3.396) that

$$\limsup_{\delta \to 0} \left(\rho\left(c\delta(1 + |(-A)^{\frac{\beta}{2}} \bar{x}|^2)\right) - \delta |(-A)^{\frac{1}{2}} \bar{x}|^2\right) \le 0. \tag{3.406}$$

Using (3.405), (3.406) and (3.391) in (3.404) we thus obtain

$$\lambda u(\bar{x}) + \left\langle (-A)^{\frac{1-\eta}{2}} \bar{x}, \frac{(-A)^{\frac{1-\eta}{2}}(\bar{x} - \bar{y})}{\epsilon} \right\rangle + G\left(\bar{x}, \frac{(-A)^{-\eta}(\bar{x} - \bar{y})}{\epsilon}, Z_N\right) \tag{3.407}$$
$$\le \omega_1(\epsilon, \delta; N) + \omega_2(\epsilon; \delta),$$

where $\lim_{N \to \infty} \omega_1(\epsilon, \delta; N) = 0, \lim_{\delta \to 0} \omega_2(\epsilon; \delta) = 0$. Similarly we obtain

$$\lambda v(\bar{y}) + \left\langle (-A)^{\frac{1-\eta}{2}} \bar{y}, \frac{(-A)^{\frac{1-\eta}{2}}(\bar{x} - \bar{y})}{\epsilon} \right\rangle + G\left(\bar{y}, \frac{(-A)^{-\eta}(\bar{x} - \bar{y})}{\epsilon}, Y_N\right) \tag{3.408}$$
$$\ge -\omega_1(\epsilon, \delta; N) - \omega_2(\epsilon; \delta).$$

We subtract (3.408) from (3.407), use Hypothesis 3.117-(iv), and let $N \to +\infty$ to conclude

$$\lambda(u(\bar{x}) - v(\bar{y})) \le \omega\left(|(-A)^{\frac{\beta}{2}}(\bar{x} - \bar{y})|\left(1 + \frac{|(-A)^{\frac{\beta}{2}}(\bar{x} - \bar{y})|}{\epsilon}\right)\right)$$
$$- \frac{|(-A)^{\frac{1-\eta}{2}}(\bar{x} - \bar{y})|^2}{\epsilon} + 2\omega_2(\epsilon; \delta). \tag{3.409}$$

By (3.384), for every $\sigma > 0$

$$|(-A)^{\frac{\beta}{2}}(\bar{x} - \bar{y})| \le \sigma |(-A)^{\frac{1-\eta}{2}}(\bar{x} - \bar{y})| + C_\sigma |(-A)^{-\frac{\eta}{2}}(\bar{x} - \bar{y})|. \tag{3.410}$$

Since for every $\alpha > 0$, $\omega(s) \le \alpha/2 + K_\alpha s$, if σ is sufficiently small, we obtain after elementary calculations

$$\lambda(u(\bar{x}) - v(\bar{y})) \le \alpha + \tilde{K}_\alpha \frac{|(-A)^{-\frac{\eta}{2}}(\bar{x} - \bar{y})|^2}{\epsilon} + 2\omega_2(\epsilon; \delta). \tag{3.411}$$

By (3.397) this implies

$$\limsup_{\epsilon \to 0} \limsup_{\delta \to 0} (u(\bar{x}) - v(\bar{y})) \le \frac{\alpha}{\lambda}$$

for all $\alpha > 0$, which gives $u \le v$ in H, since for all $x \in H$ we have

$$\Phi(x,x) \le \Phi(\bar{x},\bar{y}) \le u(\bar{x}) - v(\bar{y}).$$

If u is a solution, we can set $u = v$ in the proof to obtain that for all $x, y \in H$

$$u(x) - u(y) - \frac{|(-A)^{-\frac{\eta}{2}}(x-y)|^2}{2\epsilon} = \lim_{\delta\to 0} \Phi(x,y) \le \limsup_{\delta\to 0}(u(\bar{x}) - u(\bar{y})) \le \rho_1(\epsilon)$$

for some modulus ρ_1 in light of (3.411). This proves (3.394).

Existence. Existence of a viscosity solution will be proved by the method of finite dimensional approximations similar to this of Section 3.7. We consider for $N > 2$ the approximating equations

$$\lambda u_N - \langle Ax, Du_N\rangle + G(x, Du_N, D^2u_N) = 0 \quad \text{in } H_N. \tag{3.412}$$

We notice that for every $\gamma > 0$, $(-A)^\gamma x = P_N(-A)^\gamma x, (-A)^{-\gamma}x = P_N(-A)^{-\gamma}x$ for $x \in H_N$, and thus (3.412) satisfies Hypotheses 3.113 and 3.117 with constants and moduli independent of N. Since $\underline{u}(x) = -K/\lambda$ is a viscosity subsolution and $\overline{u}(x) = K/\lambda$ is a viscosity supersolution of (3.412), it follows from the finite dimensional Perron's method that (3.412) has a (unique) bounded viscosity solution u_N such that $\|u_N\|_0 \le K/\lambda$.

We will prove that there exists a modulus $\tilde{\sigma}_\eta$ independent of N such that

$$|u_N(x) - u_N(y)| \le \tilde{\sigma}_\eta(|x-y|_{-\eta})$$

for all $x, y \in H_N$. To do this we adapt the technique of Section 3.7.

For every $\epsilon > 0$ let K_ϵ be such that $\omega(r) \le \lambda\epsilon/2 + K_\epsilon r$. For $L > K/\lambda + 1$ we set

$$\psi_L(r) = 2Lr^{\frac{1}{2L}}.$$

The function $\psi_L \in C^2(0,\infty)$, is increasing, concave, $\psi_L'(r) \ge 1$ for $0 < r \le 1$, $\psi_L(0) = 0$, $\psi_L(1) > 2(K/\lambda + 1)$, and

$$\psi_L(r) > L\left(\psi_L'(r)r + r\right) \quad \text{for } 0 \le r \le 1. \tag{3.413}$$

We will show that, for every $\epsilon > 0$ there exists $L = L_\epsilon$ such that

$$u_N(x) - u_N(y) \le \psi_L(|(-A)^{-\frac{\eta}{2}}(x-y)|) + \epsilon \qquad \text{for every } x, y \in H_N. \tag{3.414}$$

Set $\Delta = \left\{(x,y) \in H \times H : \ |(-A)^{-\frac{\eta}{2}}(x-y)| < 1\right\}$. It is clear from the properties of ψ_L, that for $(x,y) \notin \Delta$ (3.414) is always satisfied independently of L. Assume now by contradiction that (3.414) is false. Then, for any $L > \frac{K}{\lambda} + 1$ we have, for small $\delta > 0$,

$$\sup_{(x,y)\in H_N\times H_N} \left(u_N(x) - u_N(y) - \psi_L(|(-A)^{-\frac{\eta}{2}}(x-y)|) - \epsilon - \frac{\delta}{2}|x|^2 - \frac{\delta}{2}|y|^2 \right) > 0 \tag{3.415}$$

and is attained at $(\bar{x},\bar{y}) \in \Delta$ such that $\bar{x} \ne \bar{y}$. Denote $s = |(-A)^{-\frac{\eta}{2}}(\bar{x}-\bar{y})|$.

Repeating the arguments from the proof of Proposition 3.81 that led to (3.218) and then the arguments from the just finished proof of comparison we obtain that there exist $Z, Y \in S(H_N)$ such that

$$\begin{pmatrix} Z & 0 \\ 0 & -Y \end{pmatrix} \le \frac{2\psi_L'(s)}{s} \begin{pmatrix} (-A)^{-\eta}P_N & -(-A)^{-\eta}P_N \\ -(-A)^{-\eta}P_N & (-A)^{-\eta}P_N \end{pmatrix}$$

and

$$\begin{aligned}\lambda(u_N(\bar{x})-u_N(\bar{y})) &\le -\frac{\psi_L'(s)}{s}|(-A)^{\frac{1-\eta}{2}}(\bar{x}-\bar{y})|^2+G(\bar{y},\frac{\psi_L'(s)}{s}(-A)^{-\eta}(\bar{x}-\bar{y}),Y)\\ &\quad -G(\bar{x},\frac{\psi_L'(s)}{s}(-A)^{-\eta}(\bar{x}-\bar{y}),Z)+\rho(L;\delta)\\ &\le -\frac{\psi_L'(s)}{s}|(-A)^{\frac{1-\eta}{2}}(\bar{x}-\bar{y})|^2+\frac{\lambda\epsilon}{4}\\ &\quad +K_\epsilon\left(|(-A)^{\frac{\beta}{2}}(\bar{x}-\bar{y})|\left(1+\frac{\psi_L'(s)}{s}|(-A)^{\frac{\beta}{2}}(\bar{x}-\bar{y})|\right)\right)+\rho(L;\delta),\end{aligned}$$

where $\lim_{\delta\to 0}\rho(L;\delta)=0$. Therefore using (3.410) with sufficiently small σ it follows that

$$\lambda(u_N(\bar{x})-u_N(\bar{y}))\le\frac{\lambda\epsilon}{2}+C_\varepsilon(\psi_L'(s)s+s)+\rho(L;\delta),$$

where C_ϵ only depends on K_ϵ and the interpolation constant but not on L. Choosing $L=C_\epsilon/\lambda$, using (3.413), and letting $\delta\to 0$ we arrive at

$$u_N(\bar{x})-u_N(\bar{y})\le\frac{\epsilon}{2}+\psi_L(s).$$

This is a contradiction since we obviously have by (3.415)

$$\psi_L(s)+\varepsilon\le u_N(\bar{x})-u_N(\bar{y}).$$

Hence we obtain the existence of the required modulus of continuity $\tilde{\sigma}_\eta$.

Now set $v_N(x)=u_N(P_Nx)$. Since $(-A)^{-\frac{\eta}{2}}$ is compact we are in a position to apply Arzela-Ascoli theorem to find a subsequence (still denoted by v_N) converging uniformly on bounded sets of H to a function u that obviously satisfies the same estimates as u_N's. It remains to show that u solves the limiting equation (3.387). To this end let $u-\psi$ have a maximum at $\hat{x}$ (which we may assume to be strict) for some test function $\psi(x)=\varphi(x)+\delta|x|^2$. It follows from the local uniform convergence of the v_N and the strictness of the maximum at $\hat{x}$ that there exists a sequence $\hat{x}_N=P_N\hat{x}_N\to\hat{x}$ as $N\to\infty$ such that, for every $x\in H_N$,

$$v_N(x)-\varphi(x)-\delta|x|^2\le v_N(\hat{x}_N)-\varphi(\hat{x}_N)-\delta|\hat{x}_N|^2.$$

Therefore, since $AP_N=P_NA$,

$$\begin{aligned}&\lambda u_N(\hat{x}_N)+\langle(-A)^{\frac12}\hat{x}_N,(-A)^{\frac12}D\varphi(\hat{x}_N)\rangle+2\delta|(-A)^{\frac12}\hat{x}_N|^2\\ &+G\left(\hat{x}_N,P_ND\varphi(\hat{x}_N)+2\delta\hat{x}_N,P_N(D^2\varphi(\hat{x}_N)+2\delta I)P_N\right)\le 0.\end{aligned}\tag{3.416}$$

Since φ is a test function we have

$$|(-A)^{\frac12}D\varphi(\hat{x}_N)|\le C_1+C_2|(-A)^{\frac12-\epsilon}\hat{x}_N|\tag{3.417}$$

for some independent constants C_1,C_2. Also, by (3.384), (3.392), (3.395) and (3.417),

$$\begin{aligned}&|G\left(\hat{x}_N,P_ND\varphi(\hat{x}_N)+2\delta\hat{x}_N,P_N(D^2\varphi(\hat{x}_N)+2\delta I)P_N\right)|\\ &\le C_3\left(1+|(-A)^{\frac{\beta}{2}}\hat{x}_N|^2+|(-A)^{\frac12-\epsilon}\hat{x}_N|^2\right)\le C_4+\frac{\delta}{2}|(-A)^{\frac12}\hat{x}_N|^2.\end{aligned}$$

Using this, (3.417) and (3.384), we therefore obtain from (3.416) that

$$|(-A)^{\frac12}\hat{x}_N|\le C_5$$

for some constant C_5 independent of N. Thus $(-A)^{\frac12}\hat{x}_N\rightharpoonup(-A)^{\frac12}\hat{x}$ (so $\hat{x}\in D((-A)^{\frac12})$) and hence

$$(-A)^{\frac{\beta}{2}}\hat{x}_N\to(-A)^{\frac{\beta}{2}}\hat{x},\quad\text{and}\quad(-A)^{\frac12}D\varphi(\hat{x}_N)\to(-A)^{\frac12}D\varphi(\hat{x}).$$

These convergences and Lemma 3.85 allow us to pass to the limit in (3.416) as $N \to \infty$ to conclude that

$$\lambda u(\hat{x})+\langle(-A)^{\frac{1}{2}}\hat{x},(-A)^{\frac{1}{2}}D\varphi(\hat{x})\rangle+2\delta|(-A)^{\frac{1}{2}}\hat{x}|^2+G\left(\hat{x},D\varphi(\hat{x})+\delta\hat{x},D^2\varphi(\hat{x})+\delta I\right)\le 0.$$

The proof of the supersolution property is analogous. $\square$

3.12.3. Stochastic control problem. We present an application of the results of the previous section to an abstract infinite horizon stochastic optimal control problem which includes the class of problems discussed in Section 2.6.2 and may come from a boundary control problem of Dirichlet type with distributed controls. We take the usual setup. Let H, Ξ be real separable Hilbert spaces, and $Q \in \mathcal{L}^+(\Xi)$. Let $\Lambda = \Lambda_1 \times \tilde{\Lambda}_2$, where Λ_1, Λ_2 are real separable Hilbert spaces, and $\tilde{\Lambda}_2$ is a closed bounded subset of Λ_2. We set $R := \sup_{a_2 \in \tilde{\Lambda}_2} |a_2|_{\Lambda_2}$. Given a reference probability space $\nu = (\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s\ge 0}, \mathbb{P}, W_Q)$ we have the set of admissible controls

$$\begin{aligned}\mathcal{U}^\nu := \big\{ & a(\cdot) = (a_1(\cdot), a_2(\cdot)) : [0,+\infty)\times\Omega \to \Lambda : \\ & a_1(\cdot), a_2(\cdot) \text{ are } \mathscr{F}_s - \text{progressively measurable}\big\},\end{aligned} \tag{3.418}$$

and we denote $\mathcal{U} := \bigcup_\nu \mathcal{U}^\nu$ to be the set of all admissible controls.

We control the state given by the SDE

$$\begin{cases} dX(t) = \left[AX(t) + b(X(t), a_1(t)) + (-A)^\beta C a_2(t)\right] dt \\ \qquad\qquad\qquad +\sigma(X(t), a_1(t)) dW_Q(t), \quad t > 0 \\ X(0) = x_0 \in H, \end{cases} \tag{3.419}$$

i.e.

$$\begin{aligned} X(t) = e^{tA}x_0 &+ \int_0^t e^{(t-s)A} b(X(s), a_1(s)) ds + (-A)^\beta \int_0^t e^{(t-s)A} C a_2(s) ds \\ &+ \int_0^t (-A)^{\frac{\beta}{2}} e^{(t-s)A} (-A)^{-\frac{\beta}{2}} \sigma(X(s), a_1(s)) dW_Q(s), \end{aligned} \tag{3.420}$$

and try to minimize the cost functional

$$J(x_0; a(\cdot)) = \mathbb{E} \int_0^{+\infty} e^{-\lambda t} l\left(X(t; x_0, a(\cdot)), a(t)\right) dt, \tag{3.421}$$

over all admissible controls $a(\cdot) \in \mathcal{U}$. We denote by v the value function for this problem. We assume that A satisfies Hypothesis 3.113 and $\lambda > 0$. We also make the following assumptions.

HYPOTHESIS 3.119

(i) *The function b is continuous from $H \times \Lambda_1$ to H and there exists a constant $c_0 > 0$ such that*

$|b(x, a_1)| \le c_0(1 + |x|)$ *for all* $x \in H, a_1 \in \Lambda_1$,

$|b(x_1, a_1) - b(x_2, a_1)| \le c_0|x_1 - x_2|$ *for all* $x_1, x_2 \in H, a_1 \in \Lambda_1$.

(ii) $C \in \mathcal{L}(\Lambda_2, H)$ *and* $\beta \in \left(\frac{3}{4}, 1\right)$.

(iii) $\sigma : H \times \Lambda_1 \to \mathcal{L}(\Xi_0, H)$, *the map* $(-A)^{-\frac{\beta}{2}}\sigma : H \times \Lambda_1 \to \mathcal{L}_2(\Xi_0, H)$ *is continuous and moreover there exists a constant $K_1 > 0$ such that*

$$\|(-A)^{-\frac{\beta}{2}}\sigma(x, a_1)\|_{\mathcal{L}_2(\Xi_0, H)} \le K_1(1 + |x|)$$

for all $x \in H$, $a_1 \in \Lambda_1$,

$$\|(-A)^{-\frac{\beta}{2}}[\sigma(x_1, a_1) - \sigma(x_2, a_1)]\|_{\mathcal{L}_2(\Xi_0, H)} \le K_1|x_1 - x_2|$$

for all $x_1, x_2 \in H$, $a_1 \in \Lambda_1$.

(iv) *For all* $x \in H$

$$\lim_{N\to+\infty} \sup_{a_1\in\Lambda_1} \|Q_N(-A)^{-\frac{\beta}{2}}\sigma(x,a_1)\|_{\mathcal{L}_2(\Xi_0,H)} = 0.$$

(v) $l \in C(H\times\Lambda)$ *and*

$|l(x,a)| \le C_l$, *for all* $(x,a)\in H\times\Lambda$,

$|l(x_1,a)-l(x_2,a)| \le \omega_l(|x_1-x_2|)$, *for all* $a\in\Lambda,\ x_1,x_2\in H$,

for some positive constant C_l *and modulus* ω_l.

Remark 3.120

(1) Hypotheses 3.119-*(iii),(iv)* are satisfied if we assume, for example, that there exists a constant $K_2>0$ such that

$$\|\sigma(x,a_1)\|_{\mathcal{L}(\Xi_0,H)} \le K_2(1+|x|)$$

for all $x\in H$, $a_1\in\Lambda_1$

$$\|\sigma(x_1,a_1)-\sigma(x_2,a_1)\|_{\mathcal{L}(\Xi_0,H)} \le K_2|x_1-x_2|$$

for all $x_1,x_2\in H$, $a_1\in\Lambda_1$, and if the operator $(-A)^{-\beta}$ is trace class.

(2) Hypothesis 3.119-*(iv)* is satisfied if, for instance, for every $x\in H$ there exists $\eta\in(0,\beta/2)$ such that $(-A)^{-\eta}\sigma(x,a_1)$ is bounded in $\mathcal{L}_2(\Xi_0,H)$ independently of $a_1\in\Lambda_1$.

■

It is a consequence of Theorem 1.135 that for every generalized reference probability space $\mu=(\Omega,\mathscr{F},\{\mathscr{F}_s\}_{s\ge0},\mathbb{P},W_Q)$, $T>0$, $a(\cdot)\in\mathcal{U}^\mu$, and $\xi\in L^2(\Omega,\mathscr{F}_0,\mathbb{P})$, equation (3.419) with $X(0)=\xi$ has a unique mild solution in $\mathcal{H}_2^\mu(0,T;H)$ with continuous trajectories. Following the strategy described in Remark 2.17 and using Proposition 1.136 one can then argue that all the assumptions needed to prove DPP for the problem are satisfied. However we will not look directly into this, since we need to study the transformed HJB equation 3.389 and the optimal control problem associated with it.

Let us first look how the state equation is transformed by the change of variables. If $X(\cdot;x_0,a(\cdot))$ satisfies (3.419) then $Y(\cdot)=Y(\cdot;y_0,a(\cdot)):=(-A)^{-\frac{\beta}{2}}X(\cdot;x_0,a(\cdot))$ satisfies the equation

$$\begin{cases} dY(t) = \Big[AY(t) + (-A)^{-\frac{\beta}{2}}b((-A)^{\frac{\beta}{2}}Y(t),a_1(t)) + (-A)^{\frac{\beta}{2}}Ca_2(t)\Big]\,dt \\ \qquad\qquad +(-A)^{-\frac{\beta}{2}}\sigma((-A)^{\frac{\beta}{2}}Y(t),a_1(t))dW_Q(t) \\ \\ Y(0)=y_0=(-A)^{-\frac{\beta}{2}}x_0\in H, \end{cases} \tag{3.422}$$

which is understood in its mild form

$$Y(t) = e^{tA}y_0 + \int_0^t e^{(t-s)A}(-A)^{-\frac{\beta}{2}}b((-A)^{\frac{\beta}{2}}Y(s),a_1(s))ds \\ + (-A)^{\frac{\beta}{2}}\textstyle\int_0^t e^{(t-s)A}Ca_2(s)ds + \displaystyle\int_0^t e^{(t-s)A}(-A)^{-\frac{\beta}{2}}\sigma((-A)^{\frac{\beta}{2}}Y(s),a_1(s))dW_Q(s). \tag{3.423}$$

We are now minimizing the cost functional

$$\tilde{J}(y_0; a(\cdot)) = \mathbb{E} \int_0^{+\infty} e^{-\lambda t} l\left((-A)^{\frac{\beta}{2}} Y(t; y_0, a(\cdot)), a(t) \right) dt, \tag{3.424}$$

over all admissible controls and we denote by u the value function for this problem. The HJB equation associated with this new control problem is of the form (3.389) with $G : D((-A)^{\frac{\beta}{2}}) \times D((-A)^{\frac{\beta}{2}}) \times S(H) \to \mathbb{R}$ given by

$$\begin{aligned} &G(y, q, S) \\ &= \sup_{a \in \Lambda} \Big\{ -\frac{1}{2} \mathrm{Tr}\ \left[((-A)^{-\frac{\beta}{2}} \sigma((-A)^{\frac{\beta}{2}} y, a_1) Q^{\frac{1}{2}})((-A)^{-\frac{\beta}{2}} \sigma((-A)^{\frac{\beta}{2}} y, a_1) Q^{\frac{1}{2}})^* S \right] \\ &\qquad - \left\langle b((-A)^{\frac{\beta}{2}} y, a_1), (-A)^{-\frac{\beta}{2}} q \right\rangle - \left\langle C a_2, (-A)^{\frac{\beta}{2}} p \right\rangle - l((-A)^{\frac{\beta}{2}} y, a) \Big\}. \end{aligned} \tag{3.425}$$

We will see that the value functions v and u are linked by the relation $v(x) = u((-A)^{-\frac{\beta}{2}} x)$ for $x \in H$. Thus u should correspond to an HJB equation (3.385) with a Hamiltonian $F : H \times D((-A)^{\beta}) \times (-A)^{-\frac{\beta}{2}} S(H)(-A)^{-\frac{\beta}{2}} \to \mathbb{R}$ such that $G = G_F$, where G_F is given by (3.388). An easy calculation shows that this is true if

$$\begin{aligned} F(x, p, S) = \sup_{a \in \Lambda} \Big\{ &-\frac{1}{2} \mathrm{Tr}\ \left[(\sigma(x, a_1) Q^{\frac{1}{2}})^* S (\sigma(x, a_1) Q^{\frac{1}{2}}) \right] \\ &- \langle b(x, a_1), q \rangle - \left\langle C a_2, (-A)^{\beta} p \right\rangle - l(x, a) \Big\}. \end{aligned} \tag{3.426}$$

This is just the formal Hamiltonian corresponding to the original control problem, however notice that the second order terms in F are written in a slightly different form since we do not know that $\sigma(x, a_1) Q^{\frac{1}{2}} \in \mathcal{L}_2(U, H)$. We remark that if either $(\sigma(x, a_1) Q^{\frac{1}{2}})^* S$ or $S(\sigma(x, a_1) Q^{\frac{1}{2}})$ is trace class then (see Appendix B.3)

$$\mathrm{Tr}\ \left[(\sigma(x, a_1) Q^{\frac{1}{2}})^* S (\sigma(x, a_1) Q^{\frac{1}{2}}) \right] = \mathrm{Tr}\ \left[(\sigma(x, a_1) Q^{\frac{1}{2}})(\sigma(x, a_1) Q^{\frac{1}{2}})^* S \right].$$

PROPOSITION 3.121 *Assume that Hypotheses 3.113 and 3.119 hold. Let* $\mu = (\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s \geq 0}, \mathbb{P}, W_Q)$ *be a generalized reference probability space,* $T > 0$, $a(\cdot) \in \mathcal{U}^{\mu}$, *and* $\xi \in L^2(\Omega, \mathscr{F}_0, \mathbb{P})$. *Then equation (3.422) with* $Y(0) = \xi$ *has a unique mild solution* $Y(\cdot) = Y(\cdot\,; \xi, a(\cdot))$ *in* $M^2_{\mu}(0, T; D((-A)^{\frac{\beta}{2}}))$ *with continuous trajectories in* H.

PROOF. The proof of existence and uniqueness will follow from the contraction mapping principle. Assume first that $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$, $2 \leq p < 2/\beta$. For $Z \in M^p_{\mu}(0, T; D((-A)^{\frac{\beta}{2}}))$ we define a map $\mathcal{K}$ on $M^p_{\mu}(0, T; D((-A)^{\frac{\beta}{2}})$ by

$$\begin{aligned} &\mathcal{K}[Z](t) = e^{tA} \xi + \int_0^t e^{(t-s)A} (-A)^{-\frac{\beta}{2}} b((-A)^{\frac{\beta}{2}} Z(s), a_1(s)) ds \\ &+ (-A)^{\frac{\beta}{2}} \int_0^t e^{(t-s)A} C a_2(s) ds + \int_0^t e^{(t-s)A} (-A)^{-\frac{\beta}{2}} \sigma((-A)^{\frac{\beta}{2}} Z(s), a_1(s)) dW_Q(s). \end{aligned}$$

Thus, thanks to Hypothesis 3.119(i) and (B.15), for suitable constants $C_1, C_2 > 0$,

$$\begin{aligned} &|(-A)^{\frac{\beta}{2}} \mathcal{K}[Z](t)| \leq |(-A)^{\frac{\beta}{2}} e^{tA} \xi| + C_1 \int_0^t e^{-a(t-s)} [1 + |(-A)^{\frac{\beta}{2}} Z(s)|] ds \\ &+ C_2 R \int_0^t \frac{e^{-a(t-s)}}{(t-s)^{\beta}} ds + \left| \int_0^t (-A)^{\frac{\beta}{2}} e^{(t-s)A} (-A)^{-\frac{\beta}{2}} \sigma((-A)^{\frac{\beta}{2}} Z(s), a_1(s)) dW_Q(s) \right|. \end{aligned}$$

Then, taking the expectation of the p-th power of the terms of this last inequality and using (1.106) and Hypothesis 3.119(iii) we get

$$\begin{aligned} \mathbb{E}\left|(-A)^{\frac{\beta}{2}}\mathcal{K}[Z](t)\right|^p &\le C_3\bigg[\tfrac{1}{t^{\frac{p\beta}{2}}}\mathbb{E}|\xi|^p + 1 \\ + \int_0^t \mathbb{E}|(-A)^{\frac{\beta}{2}}Z(s)|^p ds &+ \int_0^t \frac{1}{(t-s)^{\frac{p\beta}{2}}}[1+\mathbb{E}|(-A)^{\frac{\beta}{2}}Z(s)|^p]ds\bigg]. \end{aligned} \tag{3.427}$$

Therefore

$$\begin{aligned} |\mathcal{K}[Z]|^p_{M^p_\mu(0,T;D((-A)^{\frac{\beta}{2}}))} &= \int_0^T \mathbb{E}\left|(-A)^{\frac{\beta}{2}}\mathcal{K}[Z](t)\right|^p dt \\ &\le C_4(T)\left[\mathbb{E}|\xi|^p + 1 + \int_0^T\int_0^t\left[1+\frac{1}{(t-s)^{\frac{p\beta}{2}}}\right]\mathbb{E}|(-A)^{\frac{\beta}{2}}Z(s)|^p ds\right] \\ &\le C_4(T)\left[\mathbb{E}|\xi|^p + 1 + \int_0^T \mathbb{E}|(-A)^{\frac{\beta}{2}}Z(s)|^p\int_s^T\left[1+\frac{1}{(t-s)^{\frac{p\beta}{2}}}\right]dtds\right] \\ &\le C_5(T)\left[\mathbb{E}|\xi|^p + 1 + |Z|^p_{M^p_\mu(0,T;D((-A)^{\frac{\beta}{2}}))}\right]. \end{aligned} \tag{3.428}$$

We now prove that $\mathcal{K}$ is a contraction on $M^p_\mu(0,T;D((-A)^{\frac{\beta}{2}}))$ if T is sufficiently small. Let $Z_1(\cdot), Z_2(\cdot) \in M^p_\mu(0,T;D((-A)^{\frac{\beta}{2}}))$. Then, arguing as above, we have

$$\begin{aligned} &\mathbb{E}\left|(-A)^{\frac{\beta}{2}}\left(\mathcal{K}[Z_1](t) - \mathcal{K}[Z_2](t)\right)\right|^p \\ &\qquad \le C_6\int_0^t\left(1+\frac{1}{(t-s)^{\frac{p\beta}{2}}}\right)\mathbb{E}\left|(-A)^{\frac{\beta}{2}}(Z_1(s)-Z_2(s))\right|^p ds \end{aligned} \tag{3.429}$$

for some constant $C_6 > 0$, which implies

$$|\mathcal{K}[Z_1] - \mathcal{K}[Z_2]|_{M^p_\mu(0,T;D((-A)^{\frac{\beta}{2}}))} \le C_7\left(T + T^{1-\frac{p\beta}{2}}\right)^{\frac{1}{2}}|Z_1 - Z_2|_{M^p_\mu(0,T;D((-A)^{\frac{\beta}{2}}))}.$$

Thus $\mathcal{K}$ is a contraction on $M^p_\mu(0,T;D((-A)^{\frac{\beta}{2}}))$ for small $T > 0$. The existence of a unique solution in $M^p_\mu(0,T;D((-A)^{\frac{\beta}{2}}))$ for any $T > 0$ follows by repeating the argument a finite number of times. The second and third terms of the right hand side of (3.423) have continuous trajectories by Lemma 1.110 whereas the stochastic integral there has continuous trajectories if $p > 2$ by Proposition 1.111. To prove that the solution has continuous trajectories if $\xi \in L^2(\Omega, \mathscr{F}_0, \mathbb{P})$ we argue as in the proof of Theorem 1.135. We approximate ξ by random variables

$$\xi_n = \begin{cases} \xi & \text{if } |\xi| \le n \\ 0 & \text{if } |\xi| > n. \end{cases}$$

The solutions $Y(\cdot;\xi_n, a(\cdot))$ have continuous trajectories in H, $\mathbb{P}$-a.s. and since the solutions are obtained by fixed point, one can show that $Y(\cdot;\xi,a(\cdot)) = Y(\cdot;\xi_n,a(\cdot))$, $\mathbb{P}$-a.s. on $\{\omega : |\xi(\omega)| \le n\}$. □

PROPOSITION 3.122 *Assume that Hypotheses 3.113 and 3.119 hold. Let $T > 0$, $y_0 \in H$. Then there exists a constant $C(T, |y_0|) \ge 0$ such that, for all $t \in (0,T]$ and $a(\cdot) \in \mathcal{U}$,*

$$\mathbb{E}|(-A)^{\frac{\beta}{2}}Y(t;y_0,a(\cdot))|^2 \le C(T,|y_0|)\frac{1}{t^\beta}. \tag{3.430}$$

Moreover, for every $\gamma \in (0, 1-\beta)$, there exists a constant $C_\gamma(T, |y_0|) \geq 0$ such that, for all $t \in (0, T]$ and $a(\cdot) \in \mathcal{U}$,

$$\int_0^t \frac{1}{(t-s)^{\beta+\gamma}} \mathbb{E}|(-A)^{\frac{\beta}{2}} Y(s; y_0, a(\cdot))|^2 ds \leq C_\gamma(T, |y_0|) \frac{1}{t^\beta}. \tag{3.431}$$

PROOF. Estimate (3.427) for $p = 2$ applied to the solution $Y(\cdot) = Y(\cdot; y_0, a(\cdot))$ implies

$$\mathbb{E}\left|(-A)^{\frac{\beta}{2}} Y(t)\right|^2 \leq C(T)\left[\frac{1}{t^\beta}(|y_0|^2 + 1) + \int_0^t \frac{1}{(t-s)^\beta} \mathbb{E}|(-A)^{\frac{\beta}{2}} Y(s)|^2 ds\right].$$

Estimate (3.430) thus follows from Proposition D.25. Now

$$\int_0^t \frac{1}{(t-s)^{\beta+\gamma}} \mathbb{E}|(-A)^{\frac{\beta}{2}} Y(s)|^2 ds \leq C(T, |y_0|) \int_0^t \frac{1}{(t-s)^{\beta+\gamma}} \frac{1}{s^\beta} ds \leq C_\gamma(T, |y_0|) \frac{1}{t^\beta}$$

since

$$\int_0^t \frac{1}{(t-s)^{\beta+\gamma}} \frac{1}{s^\beta} ds = t^{1-(2\beta+\gamma)} \int_0^1 \frac{1}{(1-s)^{\beta+\gamma}} \frac{1}{s^\beta} ds$$

and this last integral is bounded and $1 - (2\beta + \gamma) > -\beta$. □

THEOREM 3.123 *Assume that Hypotheses 3.113, 3.119 hold. Then the value function v is the unique $BUC(H_{-\eta})$ viscosity solution (for every $\eta \in (0,1)$) of the HJB equation (3.385) with the Hamiltonian F given by (3.426). Moreover, the dynamic programming principle holds for u and v, i.e. for $x \in H$ and all $T > 0$,*

$$v(x) = \inf_{a(\cdot) \in \mathcal{U}} \mathbb{E}\left\{\int_0^T e^{-\lambda t} l(X(t; x, a(\cdot)), a(t)) dt + e^{-\lambda T} v(X(T; x, a(\cdot)))\right\}$$

and

$$u(y) = \inf_{a(\cdot) \in \mathcal{U}} \mathbb{E}\left\{\int_0^T e^{-\lambda t} l((-A)^{\frac{\beta}{2}} Y(t; y, a(\cdot)), a(t)) dt + e^{-\lambda T} u(Y(T; y, a(\cdot)))\right\}.$$

We will prove this theorem by the approximation argument used in the proof of existence of Theorem 3.118. We consider for $N \geq 1$ the following SDE approximating the state equation (3.422).

$$\begin{cases} dY_N(t) = \Big[P_N A Y_N(t) + (-A)^{-\frac{\beta}{2}} P_N b((-A)^{\frac{\beta}{2}} Y(t), a_1(t)) + (-A)^{\frac{\beta}{2}} P_N C a_2(t)\Big] dt \\ \qquad\qquad + (-A)^{-\frac{\beta}{2}} P_N \sigma((-A)^{\frac{\beta}{2}} Y(t), a_1(t)) dW_Q(t) \\ Y_N(0) = P_N y_0 \in H_N. \end{cases} \tag{3.432}$$

These are finite dimensional SDE (even though the noise is infinite dimensional) in the spaces H_N which have unique strong solutions $Y_N(\cdot) = Y_N(\cdot; y_0, a(\cdot))$ and good continuous dependence estimates like these in Sections 1.4.3 and 3.1.2 with respect to the norm in H_N. The solutions $Y_N(\cdot)$ can be also written in the mild form

$$\begin{aligned} Y_N(t) = e^{tA} P_N y_0 + \int_0^t e^{(t-s)A} (-A)^{-\frac{\beta}{2}} P_N b((-A)^{\frac{\beta}{2}} Y_N(s), a_1(s)) ds \\ + (-A)^{\frac{\beta}{2}} \int_0^t e^{(t-s)A} P_N C a_2(s) ds \\ + \int_0^t e^{(t-s)A} (-A)^{-\frac{\beta}{2}} P_N \sigma((-A)^{\frac{\beta}{2}} Y_N(s), a_1(s)) dW_Q(s). \end{aligned} \tag{3.433}$$

Lemma 3.124 *Let Hypotheses 3.113, 3.119 hold, $y_0 \in H$, and $T > 0$. Then*

$$\lim_{N\to+\infty} \sup_{a(\cdot)\in\mathcal{U}} |Y_N(\cdot; y_0, a(\cdot)) - Y(\cdot; y_0, a(\cdot))|_{M^2(0,T;D((-A)^{\frac{\beta}{2}}))} = 0.$$

Proof. We denote $Y(\cdot) = Y(\cdot; y_0, a(\cdot)), Y_N(\cdot) = Y_N(\cdot; y_0, a(\cdot))$ and fix $\gamma \in (0, 1-\beta)$. Recall that P_N, Q_N commute with $-A$, its fractional powers and e^{tA}. We have

$$\begin{aligned}
(-A)^{\frac{\beta}{2}}(Y_N(t) - Y(t)) &= -(-A)^{\frac{\beta}{2}} e^{tA} Q_N y_0 \\
&- Q_N(-A)^{-\frac{\gamma}{2}} \int_0^t (-A)^{\frac{\gamma}{2}} e^{(t-s)A} b((-A)^{\frac{\beta}{2}} Y(s), a_1(s)) ds \\
&- Q_N(-A)^{-\frac{\gamma}{2}} \int_0^t (-A)^{\beta+\frac{\gamma}{2}} e^{(t-s)A} C a_2(s) ds \\
&- Q_N(-A)^{-\frac{\gamma}{2}} \int_0^t (-A)^{\frac{\beta+\gamma}{2}} e^{(t-s)A} (-A)^{\frac{\beta}{2}} \sigma((-A)^{\frac{\beta}{2}} Y(s), a_1(s)) dW(s) \\
&+ \int_0^t e^{(t-s)A} P_N [b((-A)^{\frac{\beta}{2}} Y_N(s), a_1(s)) - b((-A)^{\frac{\beta}{2}} Y(s), a_1(s))] ds \\
&+ \int_0^t (-A)^{\frac{\beta}{2}} e^{(t-s)A} (-A)^{\frac{\beta}{2}} P_N [\sigma((-A)^{\frac{\beta}{2}} Y_N(s), a_1(s)) - \sigma((-A)^{\frac{\beta}{2}} Y(s), \alpha_1(s))] dW_Q(s),
\end{aligned}$$

which yields, for a suitable $C_\gamma(T) > 0$,

$$\begin{aligned}
&\mathbb{E}|(-A)^{\frac{\beta}{2}}(Y_N(t) - Y(t))|^2 \\
&\le C_\gamma(T) \left[\frac{1}{t^\beta} |Q_N y_0|^2 + \|Q_N A^{-\frac{\gamma}{2}}\|^2 \left(1 + \int_0^t \left(1 + \frac{1}{(t-s)^{\beta+\gamma}} \right) \mathbb{E}|(-A)^{\frac{\beta}{2}} Y(s)|^2 ds \right) \right. \\
&\quad \left. + \int_0^t \left(1 + \frac{1}{(t-s)^\beta} \right) \mathbb{E}|(-A)^{\frac{\beta}{2}}(Y_N(s) - Y(s))|^2 ds \right].
\end{aligned}$$

Since $A^{-\gamma/2}$ is compact, $\|Q_N A^{-\gamma/2}\| \to 0$ as $N \to +\infty$, and by using (3.431) we thus deduce

$$\begin{aligned}
\mathbb{E}|(-A)^{\frac{\beta}{2}}(Y_N(t) - Y(t))|^2 &\le C_{\gamma,T,y_0}(N) \left(1 + \frac{1}{t^\beta} \right) \\
&\quad + \tilde{C}_{\gamma,T} \int_0^t \frac{1}{(t-s)^\beta} \mathbb{E}|(-A)^{\frac{\beta}{2}}(Y_N(s) - Y(s))|^2 ds,
\end{aligned}$$

where $C_{\gamma,T,y_0}(N) \to 0$ as $N \to +\infty$. Using Proposition D.25 we thus obtain

$$\mathbb{E}|(-A)^{\frac{\beta}{2}}(Y_N(t) - Y(t))|^2 \le C_{\gamma,T,y_0}(N) M \frac{1}{t^\beta} \tag{3.434}$$

for some constant M independent of N. This implies the claim. □

Proof of Theorem 3.123. We notice that under our assumptions equation (3.389) with G given by (3.425) has a unique viscosity solution in $BUC(H_{-\eta})$ for every $\eta \in (0, 1-\beta)$. To verify that the value function u is the solution we consider the approximating problems

$$\lambda u_N - \langle Ax, Du_N \rangle + G(x, Du_N, D^2 u_N) = 0 \quad \text{in} \quad H_N. \tag{3.435}$$

Equation (3.435) is the one used in the proof of Theorem 3.118 and it is easy to see that it is the equation in H_N corresponding to the control problem with evolution given by (3.432). Therefore, by the results of Section 3.6.3, the function

$$u_N(y_0) = \inf_{a(\cdot)\in\mathcal{U}} \mathbb{E} \int_0^{+\infty} e^{-\lambda t} l((-A)^{\frac{\beta}{2}} Y_N(t; y_0, a(\cdot)), a(t)) dt \tag{3.436}$$

belongs to $BUC(H_N)$, it satisfies the dynamic programming principle, i.e. for every $y_0 \in H_N$, $T \geq 0$

$$u_N(y_0) = \inf_{a(\cdot)\in\mathcal{U}} \mathbb{E}\bigg\{ \int_0^T e^{-\lambda t} l((-A)^{\frac{\beta}{2}} Y_N(t; y_0, a(\cdot)), a(t))dt \\ + e^{-\lambda T} u_N(Y_N(T; y_0, a(\cdot)))\bigg\}, \quad (3.437)$$

and u_N is the unique viscosity solution of (3.435) in $BUC(H_N)$.

Since for every $y_0 \in H$, $Y_N(t; y_0, a(\cdot)) = Y_N(t; P_N y_0, a(\cdot))$, extending u_N to H by putting $u_N(y) = u_N(P_N y)$ we obtain (3.436) and (3.435) for every $y_0 \in H$. Moreover, from the proof of existence of Theorem 3.118, we know that for every $\eta \in (0, 1-\beta)$ and $N \geq 1$

$$\|u_N\|_0 \leq \frac{C_l}{\lambda}, \qquad |u_N(x) - u_N(y)| \leq \tilde{\sigma}_\eta(|x-y|_{-\eta}) \quad (3.438)$$

for some modulus $\tilde{\sigma}_\eta$ and $u_N \to \overline{u}$ uniformly on bounded sets, where $\overline{u}$ is the unique viscosity solution of (3.389) in $BUC(H_{-\eta}), \eta \in (0, 1-\beta)$.

We need to show that $u = \overline{u}$. We will prove that u_N converges pointwise to u as $N \to \infty$. Let $y_0 \in H$. For every $T > 0$,

$$|u_N(y_0) - u(y_0)| \\ \leq \sup_{a(\cdot)\in\mathcal{U}} \int_0^T e^{-\lambda t} \mathbb{E}\omega_l(|(-A)^{\frac{\beta}{2}}(Y_N(t; y_0, a(\cdot)) - Y(t; y_0, a(\cdot)))|)dt + 2C_l \frac{e^{-\lambda T}}{\lambda}.$$

Let $\epsilon > 0$ and $T_\epsilon > 0$ be such that such that $2C_l e^{-\lambda T_\epsilon}/\lambda \leq \epsilon$. If $\omega_l(s) \leq \epsilon + K_\epsilon s, s \geq 0$, we obtain by Cauchy-Schwarz inequality

$$\int_0^{T_\epsilon} e^{-\lambda t} \mathbb{E}\omega_l(|(-A)^{\frac{\beta}{2}}(Y_N(t; y_0, a(\cdot)) - Y(t; y_0, a(\cdot)))|)dt \\ \leq \frac{\epsilon}{\lambda} + \frac{K_\epsilon}{\lambda} |Y_N(\cdot; y_0, a(\cdot)) - Y(\cdot; y_0, a(\cdot))|_{M^2(0,T_\epsilon;D((-A)^{\frac{\beta}{2}}))}$$

for all $N \geq 1$ and all $a(\cdot) \in \mathcal{U}$. The conclusion thus follows by letting $N \to +\infty$ and using Lemma 3.124 since ϵ is arbitrary.

It remains to show the dynamic programming principle for u. By (3.437), we have

$$\left| u_N(y_0) - \inf_{a(\cdot)\in\mathcal{U}} \mathbb{E}\left\{ \int_0^T e^{-\lambda t} l((-A)^{\frac{\beta}{2}} Y(t; y_0, a(\cdot)), a(t))dt + e^{-\lambda T} u(Y(T; y_0, a(\cdot)))\right\}\right| \\ \leq \sup_{a(\cdot)\in\mathcal{U}} \mathbb{E} \int_0^T e^{-\lambda t} \omega_l(|(-A)^{\frac{\beta}{2}}(Y_N(t; y_0, a(\cdot)) - Y(t; y_0, a(\cdot)))|)dt \\ + e^{-\lambda T} \sup_{a(\cdot)\in\mathcal{U}} \mathbb{E}|u_N(Y_N(T; y_0, a(\cdot))) - u(y(T; y_0, a(\cdot)))|.$$

The first term of the right-hand side converges to 0 when N goes to infinity by the same argument as this used in the previous paragraph. For the second term, we proceed as follows:

$$\mathbb{E}|u_N(Y_N(T; y_0, a(\cdot))) - u(Y(T; y_0, a(\cdot)))| \\ \leq \mathbb{E}|u_N(Y_N(T; y_0, a(\cdot))) - u_N(Y(T; y_0, a(\cdot)))| \\ + \mathbb{E}|u_N(Y(T; y_0, a(\cdot))) - u(Y(T; y_0, a(\cdot)))|.$$

The first term of the right-hand side converges to 0 uniformly in $a(\cdot) \in \mathcal{U}$ when N goes to infinity by (3.434) and (3.438). It remains to prove that

$$\sup_{a(\cdot)\in\mathcal{U}} \mathbb{E}|u_N(Y(T;y_0,a(\cdot))) - u(Y(T;y_0,a(\cdot)))|$$

goes to 0 when N goes to infinity. By Proposition 3.122, estimate (3.430), $\mathbb{E}|Y(T;y_0,a(\cdot))|^2$ is bounded by a constant $\tilde{C}(T,|y_0|) > 0$ which does not depend on $a(\cdot) \in \mathcal{U}$. Hence, for all $R > 0$,

$$\mathbb{P}\left\{|Y(T;y_0,a(\cdot))| > R\right\} \le \frac{\tilde{C}(T,|y_0|)}{R^2}.$$

Let $\epsilon > 0$ and choose $R_\epsilon > 0$ sufficiently large so that this probability be smaller than ϵ. Then

$$\sup_{a(\cdot)\in\mathcal{U}} \mathbb{E}|u_N(Y(T;y_0,a(\cdot))) - u(Y(T;y_0,a(\cdot)))| \le \frac{2C_l}{\lambda}\epsilon + \sup_{|y|\le R_\epsilon} |u_N(y) - u(y)|.$$

We conclude by letting $N \to +\infty$ since ϵ was arbitrary.

Finally we observe that for $x_0 \in H$ and $y_0 = (-A)^{-\frac{\beta}{2}}x_0$, the mild solutions $X(\cdot)$ of (3.420) and $Y(\cdot)$ of (3.423) are related by $Y(\cdot) = (-A)^{-\frac{\beta}{2}}X(\cdot)$. Therefore we have $v(x_0) = u((-A)^{-\frac{\beta}{2}}x_0)$. Thus the dynamic programming principle also holds for v, and by definition v is the unique viscosity solution in $BUC(H_{-\eta}), \eta \in (0,1)$ of the HJB equation (3.385) with F given by (3.426). □

Remark 3.125 We would obtain the same results if instead of the change of variables $y = (-A)^{-\frac{\beta}{2}}x$ we applied the change of variables $y = (-A)^{-\frac{\gamma}{2}}x$ for $\beta \le \gamma < 1$. This may be beneficial for a boundary control problem with the Neumann boundary condition, where we have $\beta < 1/2$, as it may help make the second order terms satisfy Hypothesis 3.119-(iii)(iv), which would then have $(-A)^{-\frac{\beta}{2}}$ replaced by $(-A)^{-\frac{\gamma}{2}}$ there (see also the example below). ■

We now discuss a specific example of a stochastic boundary control problem with Dirichlet boundary conditions. It is more general than the one from Section 2.6.2 since it also contains distributed controls and allows multiplicative noise Good examples of deterministic boundary control problems can be found in [309]. The results of this section would apply to suitable stochastic perturbations of examples belonging to the "first abstract class" in [309].

Let $\mathcal{O} \subset \mathbb{R}^N$ be an open, connected and bounded set with smooth boundary. Consider, as in Section 2.6.2, the following stochastic controlled PDE

$$\begin{cases} \dfrac{\partial x}{\partial t}(t,\xi) = \Delta_\xi x(t,\xi) + f_1\left(x(t,\xi),\alpha_1(t,\xi)\right) \\ \qquad\qquad\qquad + f_2\left(x(t,\xi),\alpha_1(t,\xi)\right)\dot{W}_Q(t,\xi) & \text{in } (0,\infty)\times\mathcal{O} \\ x(0,\xi) = x_0(\xi) & \text{on } \mathcal{O} \\ x(t,\xi) = \alpha_2(t,\xi) & \text{on } (0,\infty)\times\partial\mathcal{O}, \end{cases} \tag{3.439}$$

where W_Q is a Q-Wiener process, $Q \in \mathcal{L}^+(L^2(\mathcal{O}))$, $x_0 \in L^2(\mathcal{O})$, and $f_1, f_2 : \mathbb{R}^2 \to \mathbb{R}$. We take $H = L^2(\mathcal{O})$, $\Lambda_1 = L^2(\mathcal{O})$ and Λ_2 to be the closed ball centered at 0 with radius R in $L^2(\partial\mathcal{O})$, and assume that the control $a(t) = (a_1(t), a_2(t)) := (\alpha_1(t,\cdot), \alpha_2(t,\cdot))$ belongs to $\mathcal{U}$ as defined in (3.418). As it was discussed in Section 2.6.2, (3.439) can be rewritten as an abstract stochastic evolution equation (3.419)-(3.420), where A is the Laplace operator with zero Dirichlet boundary conditions,

C is the the Dirichlet operator, and

$$b(x,a_1)(\xi)=f_1(x(\xi),a_1(\xi)),\ [\sigma(x,a_1)y](\xi)=f_2(x(\xi),a_1(\xi))y(\xi).$$

Suppose that f_1, f_2 satisfy for $i=1,2$

$$|f_i(r,s)|\le c_1(1+|r|) \text{ for all } r,s\in\mathbb{R},$$
$$|f_i(r_1,s)-f_i(r_2,s)|\le c_1|r_1-r_2| \text{ for all } r_1,r_2,s\in\mathbb{R}.$$

It is then easy to see that b satisfies Hypothesis 3.119-(i). As regards B, suppose that $Q=I$ and $N=1$. Let $\{e_k\}$ be the orthonormal basis of eigenvectors of A. In this case $-Ae_k=ck^2e_k$ where $c>0$. Moreover the e_k are bounded in $L^\infty(\mathcal{O})$, uniformly in k. Therefore we obtain

$$\begin{aligned}
&\|(-A)^{-\frac{\beta}{2}}\sigma(x,a_1)\|^2_{\mathcal{L}_2(H)}=\sum_{k=1}^{+\infty}|(-A)^{-\frac{\beta}{2}}\sigma(x,a_1)e_k|^2\\
&=c^{-\beta}\sum_{k=1}^{+\infty}\sum_{h=1}^{+\infty}h^{-2\beta}\langle\sigma(x,a_1)e_k,e_h\rangle^2=c^{-\beta}\sum_{h=1}^{+\infty}h^{-2\beta}|\sigma(x,a_1)e_h|^2\\
&=c^{-\beta}\sum_{h=1}^{+\infty}h^{-2\beta}\int_{\mathcal{O}}|f_2(x(\xi),a_1(\xi))e_h(\xi)|^2d\xi\le C_1(1+|x|^2),
\end{aligned}$$

where we used that $\langle\sigma(x,a_1)e_k,e_h\rangle=\langle e_k,\sigma(x,a_1)e_h\rangle$ to justify the third equality above, and the fact that $\beta>1/2$. However the above computation does not work for for $N\ge 2$, where stronger assumptions either on f_2 or Q need to be imposed. It also does not work in the Neumann case when $\beta<1/2$, even if $f_2\equiv 1$, and this is why it may be beneficial to use the change of variables $y=(-A)^{-\frac{\gamma}{2}}x$ for $\beta\le\gamma<1$ as discussed in Remark 3.125. The other conditions of Hypothesis 3.119-(iii),(iv) are checked similarly. If f_2 is constant (and thus so is B), Hypothesis 3.119-(iii),(iv) holds if we assume that there is an orthonormal basis $\{e_k\}$ of H such that

$$Ae_k=-\lambda_k e_k,\quad Qe_k=\beta_k e_k,\quad k\in\mathbb{N},$$

where $\{\lambda_k\}$ is a sequence of positive numbers increasing to $+\infty$ while $\{\beta_k\}$ is a bounded sequence of nonnegative real numbers and

$$\sum_{k=1}^{\infty}\frac{\beta_k}{\lambda_k^{\beta}}<+\infty.$$

Since for the Laplace operator A we have $\lambda_k\approx k^{\frac{2}{N}}$ as $k\to+\infty$, this condition is fulfilled if for some $\epsilon>0$, $\beta_k\le Ck^{\frac{2\beta}{N}-1-\epsilon}$. When Q is invertible this is possible only for $N=1$. However Q can have finite rank.

If the original cost functional was given by

$$\mathbb{E}\int_0^{+\infty}e^{-\lambda t}\int_{\mathcal{O}}f_3(x(t,\xi),\alpha_1(t,\xi))d\xi dt,$$

where $\lambda>0$ and $f_3:\mathbb{R}^2\to\mathbb{R}$, then the cost functional for the abstract evolution system (3.419)-(3.420) is given by (3.421), where

$$l(x,a_1):=\int_{\mathcal{O}}f_3(x(\xi),a_1(\xi))d\xi.$$

If $f_3\in BUC(\mathbb{R}^2)$ then l satisfies Hypothesis 3.119-(v). The original cost functional can be more general and depend explicitly on the boundary control α_2.

3.13. HJB equations for control of stochastic Navier-Stokes equations

In this section we present another special class of equations which can be studied by viscosity solution methods, which however require modification of the general definition of viscosity solution from Section 3.3 and the techniques of Sections 3.5 and 3.6. We will study the second order HJB equations that arise in problems of optimal control of stochastic Navier-Stokes equations. No much is known about equations of this type. Kolmogorov equations for stochastic Navier-Stokes equations have been studied by Komech and Vishik (see [443] and the references therein) and more recently in [25, 26, 188, 401] for two-dimensional stochastic Navier-Stokes equations and by Da Prato and Debussche [117] for the three-dimensional case. Only existence of strict and mild solutions has been proved in [117]. A semilinear equation associated to a special optimal control problem has been investigated by Da Prato and Debussche in [115] from the point of view of mild solutions. Some of these results have been generalized to the three dimensional case in [328]. The mild solution approach of [115] is discussed in Section 4.9. The viscosity solution approach is more general in a sense that it can handle more complicated cost functionals and applies to stochastic optimal control problems with the associated HJB equations that are fully nonlinear in the gradient variable and the noise. On the other hand the covariance operator of the Wiener process here must be of trace class and thus the viscosity solution approach cannot cover non-degenerate cases studied in [115], where regular solutions were obtained and a formula for optimal feedback was derived.

We will consider an optimal control problem for the 2-dimensional stochastic Navier-Stokes equations with periodic boundary conditions in the setting of an abstract stochastic evolution equation for the velocity vector field discussed in Section 2.6.5. Let $\mathcal{O} = [0, L] \times [0, L]$, and let $\nu > 0$. We define the spaces

$$V = \left\{ x \in H_p^1\left(\mathcal{O}; \mathbb{R}^2\right), \operatorname{div} x = 0, \int_{\mathcal{O}} x = 0 \right\},$$

$$H = \text{ the closure of } V \text{ in } L^2\left(\mathcal{O}; \mathbb{R}^2\right),$$

where for an integer $k \geq 1$, $H_p^k\left(\mathcal{O}; \mathbb{R}^2\right)$ is the space of $\mathbb{R}^2$ valued functions x that are in $H_{\text{loc}}^k\left(\mathbb{R}^2; \mathbb{R}^2\right)$ and such that $x(y + Le_i) = x(y)$ for every $y \in \mathbb{R}^2$ and $i = 1, 2$. We will denote by $\langle \cdot, \cdot \rangle$, and $|\cdot|$ respectively the inner product and the norm in $L^2\left(\mathcal{O}; \mathbb{R}^2\right)$. The space H inherits the same inner product and norm, and V has the norm inherited from $H_p^1\left(\mathcal{O}; \mathbb{R}^2\right)$. Let P_H be the orthogonal projection in $L^2\left(\mathcal{O}; \mathbb{R}^2\right)$ onto H. Define $Ax = P_H \Delta x$ with the domain $D(A) = H_p^2\left(\mathcal{O}; \mathbb{R}^2\right) \cap V$, and $B(x, y) = P_H[(x \cdot \nabla)y]$ for $x, y \in V$. The operator A is maximal dissipative, self-adjoint, and $(-A)^{-1}$ is compact. For $\gamma = 1, 2$ we denote $V_\gamma := D((-A)^{\frac{\gamma}{2}})$, equipped with the norm

$$|x|_\gamma := |(-A)^{\frac{\gamma}{2}} x|. \tag{3.440}$$

The space V_1 coincides with V. Recall that

$$\int_{\mathcal{O}} |\text{curl } x(\xi)|^2 d\xi = \int_{\mathcal{O}} |\nabla x(\xi)|^2 d\xi, \quad \text{for } x \in V.$$

Hence $|x|_1$-norm is equivalent to

$$\left(\int_{\mathcal{O}} |\text{curl } x(\xi)|^2 d\xi \right)^{1/2}.$$

The dual space V' of V can be identified with the space V_{-1}, which is the completion of H with respect to the norm

$$|x|_{-1} := |(-A)^{-\frac{1}{2}} x|.$$

The duality is then given by

$$\langle x, y\rangle_{\langle V', V\rangle} = \langle (-A)^{-\frac{1}{2}}x, (-A)^{\frac{1}{2}}y\rangle.$$

Let $T > 0$, Λ be a complete separable metric space. For every $0 \le t < T$, reference probability space $\nu = (\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}, W_Q)$, where W_Q ia an H-valued Q-Wiener process with $Q \in \mathcal{L}_1^+(H)$, and $a(\cdot) \in \mathcal{U}_t^\mu$, the abstract controlled stochastic Navier-Stokes (SNS) equations describe the evolution of the velocity vector field $X : [t, T] \times \mathcal{O} \times \Omega \to \mathbb{R}^2$ that satisfies the stochastic evolution equation

$$\begin{cases} dX(s) = (AX(s) - B\left(X(s), X(s)\right) + f\left(s, a(s)\right))\, ds + dW_Q(s) \quad \text{in } (t, T] \times H, \\ \\ X(t) = x \in H, \end{cases} \tag{3.441}$$

where $f : [0, T] \times \Lambda \to V$. (We remark that without loss of generality we set the viscosity coefficient in front of A to be 1.) The optimal control problem consists in the minimization, over all controls $a(\cdot) \in \mathcal{U}_t$, of a cost functional

$$J(t, x; a(\cdot)) = \mathbb{E}\left\{\int_t^T l(s, X(s), a(s))ds + g(X(T))\right\}.$$

The value function

$$v(t, x) = \inf_{a(\cdot)\in\mathcal{U}_t} J(t, x; a(\cdot)), \tag{3.442}$$

and the associated Hamilton-Jacobi-Bellman equation is

$$\begin{cases} u_t + \dfrac{1}{2}\mathrm{Tr}\left(QD^2u\right) + \langle Ax - B(x, x), Du\rangle + F(t, x, Du) = 0, \\ \\ u(T, x) = g(x) \qquad\qquad \text{for } (t, x) \in (0, T) \times H, \end{cases} \tag{3.443}$$

where the Hamiltonian function F is defined by

$$F(t, x, p) := \inf_{a\in\Lambda} \left\{\langle f(t, a), p\rangle + l(t, x, a)\right\}. \tag{3.444}$$

It is convenient to introduce the trilinear form $b(\cdot, \cdot, \cdot) : V \times V \times V \to \mathbb{R}$. It is defined as

$$b\left(x, y, z\right) = \int_{\mathcal{O}} z(\xi) \cdot (x(\xi) \cdot \nabla_\xi) y(\xi)\, d\xi = \langle B(x, y), z\rangle.$$

It is a continuous operator on $V \times V \times V$ but it can also be extended to a continuous map in different topologies, for instance it is also continuous on $V \times V_2 \times H$ (see [436] and (3.449) below.) The incompressibility condition gives the standard orthogonality relations

$$b\left(x, y, z\right) = -b\left(x, z, y\right), \quad b\left(x, y, y\right) = 0. \tag{3.445}$$

Also, because of the periodic boundary conditions (see for instance [436]),

$$b\left(x, x, Ax\right) = 0 \quad \text{for } x \in V_2. \tag{3.446}$$

We will be using the following inequalities. If $x, y, z \in V$ then

$$\left|b\left(x, y, z\right)\right| \le C\left|x\right|^{1/2}\left|x\right|_1^{1/2}\left|y\right|_1\left|z\right|^{1/2}\left|z\right|_1^{1/2}, \tag{3.447}$$

which gives when $z = x$

$$\left|b\left(x, y, x\right)\right| \le C\left|x\right|\left|x\right|_1\left|y\right|_1. \tag{3.448}$$

Also, if $x \in V$, $y \in V_2$, $z \in H$, then

$$\left|b\left(x, y, z\right)\right| \le C\left|x\right|_1\left|y\right|_2\left|z\right|. \tag{3.449}$$

We will assume the following hypothesis throughout the rest of this section

Hypothesis 3.126

(i) $(-A)^{\frac{1}{2}}Q^{\frac{1}{2}} \in \mathcal{L}_2(H)$.
(ii) *The function* $f : [0,T] \times \Lambda \to V$ *is continuous and there is* $R \geq 0$ *such that*

$$|f(t,a)|_1 \leq R \quad \text{for all } t \in [0,T], a \in \Lambda. \tag{3.450}$$

We remark that Hypothesis 3.126-(i) is equivalent to the requirement that $\mathrm{Tr}(Q_1) < +\infty$, where $Q_1 := (-A)^{\frac{1}{2}}Q(-A)^{\frac{1}{2}}$. By this we mean that $(-A)^{\frac{1}{2}}Q(-A)^{\frac{1}{2}}$ is densely defined and it extends to a bounded operator, still denoted by Q_1, belonging to $\mathcal{L}_1^+(H)$.

3.13.1. Estimates for controlled SNS equations. We will be using the notions of variational and strong solutions of the SNS equations (3.441). The definition of a variational solution is the same as this in Section 3.11.1, however since the generic equation (3.322) there is slightly different from (3.441), we repeat the definition below.

DEFINITION 3.127 *Let* $0 \leq t < T$. *Let* $\mu = \left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s\in[t,T]}, \mathbb{P}, W_Q\right)$ *be a generalized reference probability space, let Hypothesis 3.126 be satisfied. Let* ξ *be an* $\mathscr{F}_t$ *measurable* H*-valued random variable such that* $\mathbb{E}^\mu[|\xi|^2] < +\infty$, *and let* $a(\cdot) \in \mathcal{U}_t^\mu$.

- *A process* $X(\cdot) \in M_\mu^2(t,T;H)$ *is called a* variational solution *of (3.441) with initial condition* $X(t) = \xi$ *if*

$$\mathbb{E}\left[\int_t^T |X(r)|_V^2 dr\right] < +\infty$$

and for every $\phi \in V$ *we have*

$$\langle X(s), \phi\rangle = \langle \xi, \phi\rangle + \int_t^s \langle AX(r) - B(X(r), X(r)) + f(r, a(r)), \phi\rangle_{\langle V', V\rangle} dr$$
$$+ \int_t^s \langle dW_Q(r), \phi\rangle \quad \text{for each } s \in [t,T], \mathbb{P}\text{-a.e..}$$

- *A process* $X(\cdot) \in M_\mu^2(t,T;H)$ *is called a* strong solution *of (3.441) with initial condition* $X(t) = \xi$ *if*

$$\mathbb{E}\left[\int_t^T |X(r)|_{V_2}^2 dr\right] < +\infty$$

and we have

$$X(s) = \xi + \int_t^s (AX(r) - B(X(r), X(r)) + f(r, a(r)))dr + \int_t^s dW_Q(r)$$

for each $s \in [t,T], \mathbb{P}$*-a.e..*

PROPOSITION 3.128 *Let* $0 \leq t < T$ *and* $p \geq 2$. *Let* $\mu = \left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s\in[t,T]}, \mathbb{P}, W_Q\right)$ *be a generalized reference probability space, and let Hypothesis 3.126 be satisfied. Let* ξ *be an* $\mathscr{F}_t^t$ *measurable* H*-valued random variable such that* $\mathbb{E}^\mu[|\xi|^p] < +\infty$, *and let* $a(\cdot) \in \mathcal{U}_t^\mu$. *Then:*

(i) *There exists a unique variational solution* $X(\cdot) = X(\cdot; t, \xi, a(\cdot))$ *of (3.441) with initial condition* $X(t) = \xi$. *The solution has continuous trajectories and satisfies for* $t \leq s \leq T$

$$\mathbb{E}|X(s)|^p + \mathbb{E}\int_t^s |X(\tau)|_1^2 |X(\tau)|^{p-2} d\tau \leq \mathbb{E}|\xi|^p + C(p,R,Q)(s-t). \tag{3.451}$$

and

$$\mathbb{E}\left[\sup_{t\le s\le T}|X(s)|^p\right] \le C(p,T,R,Q)\left(1+\mathbb{E}|\xi|^p\right). \tag{3.452}$$

(ii) *If $\mathbb{E}|\xi|_1^p < +\infty$, then the variational solution $X(\cdot) = X(\cdot;t,\xi,a(\cdot))$ is a strong solution with trajectories continuous in V. Moreover we have for $t \le s \le T$*

$$\mathbb{E}|X(s)|_1^p + \mathbb{E}\int_t^s |X(\tau)|_2^2|X(\tau)|_1^{p-2}d\tau \le \mathbb{E}|\xi|_1^p + C(p,R,Q_1)(s-t) \tag{3.453}$$

and

$$\mathbb{E}\left[\sup_{t\le s\le T}|X(s)|_1^p\right] \le C(p,T,R,Q_1)\left(1+\mathbb{E}|\xi|_1^p\right). \tag{3.454}$$

(iii) *If μ_1 is another generalized reference probability space, ξ_1 is a $\mathscr{F}_t^{t,\mu_1}$ measurable H-valued random variable such that $\mathbb{E}^{\mu_1}[|\xi_1|^p] < +\infty$, $a_1(\cdot) \in \mathcal{U}_t^{\mu_1}$, and*

$$\mathcal{L}_{\mathbb{P}_1}(\xi_1, a_1(\cdot), W_{Q,1}(\cdot)) = \mathcal{L}_{\mathbb{P}}(\xi, a(\cdot), W_Q(\cdot)),$$

then

$$\mathcal{L}_{\mathbb{P}_1}(a_1(\cdot), X_1(\cdot)) = \mathcal{L}_{\mathbb{P}}(a(\cdot), X(\cdot)), \tag{3.455}$$

where $X_1(\cdot) = X_1(\cdot;t,\xi_1,a_1(\cdot))$ is the variational solution of (3.441) in μ_1 with control $a_1(\cdot)$ and initial condition ξ_1.

Proof. (i) The general strategy of the poof of part (i) is similar to the proof of Theorem 3.102. More precisely, part (i) is proved in [342], Proposition 3.3 (see also [93] for a similar proof and [127, 443] for related results and estimates). We sketch the main points of the proof since we will need them to explain parts (ii) and (iii).

Let $\{e_1, e_2, ...\}$ be the orthonormal basis of H composed of eigenvectors of A, $H_n := \mathrm{span}\{e_1, ..., e_n\}$, and P_n be the orthogonal projection in H onto H_n. In this case P_n extends to the orthogonal projection in V' onto H_n. Also we have $P_nA = AP_n$. Let $X^n(\cdot)$ be the unique strong solution of

$$\begin{cases} dX^n(s) = (P_nAX^n(s) - P_nB(X^n(s), X^n(s)) + P_nf(s,a(s)))ds + P_ndW_Q(s) \\ X^n(t) = P_n\xi \end{cases} \tag{3.456}$$

We first assume that $p \ge 8$. If follows using Itô's formula (see also [342]) that we have

$$\mathbb{E}\left[\sup_{t\le s\le T}|X^n(s)|^p + \int_t^T |X^n(s)|_V^2(1+|X^n(s)|^{p-2})ds\right] \le M \quad \text{for all } n.$$

It can also be deduced from the estimates for $X^n(\cdot)$ obtained from Itô's formula, and a martingale inequality, that the norms of $X^n(\cdot)$ in $L^8(\Omega; L^4((t,T)\times\mathcal{O}))$ are bounded uniformly in n. Therefore, there exists a process $X(\cdot) \in M_\mu^2(t,T;V) \cap L^p(\Omega; L^\infty(t,T;H))$, and a $\mathscr{F}_T$ measurable random variable $\eta \in L^2(\Omega; H)$ such that (up to a subsequence and identifying $X(\cdot)$ with its versions)

$$X^n(\cdot) \rightharpoonup X(\cdot) \text{ in } M_\mu^2(t,T;V), \quad X^n(T) \rightharpoonup \eta \text{ in } L^2(\Omega;H),$$
$$X^n(\cdot) \to X(\cdot) \text{ weak star in } L^p(\Omega; L^\infty(t,T;H)).$$

We can also assume that $X(\cdot) \in L^8(\Omega; L^4((t,T)\times\mathcal{O}))$. Passing to the limit as $n \to +\infty$ we obtain that there is a process $F_0(\cdot) \in M_\mu^2(t,T;V')$ such that, up to a subsequence,

$$P_nAX^n(\cdot) - P_nB(X^n(\cdot), X^n(\cdot)) \rightharpoonup F_0(\cdot) \text{ in } M_\mu^2(t,T;V'),$$

$X(\cdot)$ is a variational solution of

$$\begin{cases} dX(s) = (F_0(s) + f(s, a(s)))ds + dW_Q(s) \\ X(t) = \xi, \end{cases}$$

$X(T) = \eta$, and, by Theorem 3.101, $X(\cdot) \in L^p(\Omega; C([t,T]; H))$ and $\mathbb{P}$-a.e.

$$|X(s)|^2 = |\xi|^2 + 2\int_t^s \langle F_0(r) + f(r, a(r)), X(r)\rangle_{\langle V', V\rangle} dr \\ +2\int_t^s \langle dW_Q(r), X(r)\rangle + \mathrm{Tr}(Q)(s-t).$$

One then uses an argument based on the monotonicity of the operator $-Ax+B(x,x)$ on balls in $L^4(\mathcal{O})$ (see [342]) to show that $F_0(\cdot) = AX(\cdot) - B(X(\cdot), X(\cdot))$, i.e. $X(\cdot)$ is a variational solution of (3.441), and thus using (3.445), we have

$$|X(s)|^2 = |\xi|^2 - 2\int_t^s |X(r)|_1^2 dr + 2\int_t^s \langle f(r, a(r)), X(r)\rangle dr \\ +2\int_t^s \langle dW_Q(r), X(r)\rangle + \mathrm{Tr}(Q)(s-t). \tag{3.457}$$

(The monotonicity argument uses the fact that $X(\cdot) \in L^8(\Omega; L^4((t,T)\times\mathcal{O}))$.) We also have a similar identity as (3.457) for $|X^n(s)|^2$. Taking expectation in both of them for $s = T$, passing to the limit as $n \to +\infty$, and recalling that $X^n(T) \rightharpoonup X(T)$, we deduce $\mathbb{E}|X^n(T)|^2 \to \mathbb{E}|X(T)|^2$, which gives $\mathbb{E}|X^n(T) - X(T)|^2 \to 0$ as $n \to +\infty$. Replacing T by $s \in (t,T)$, the same arguments give $\mathbb{E}|X^n(s) - X(s)|^2 \to 0$.

Estimates (3.451) and (3.452) can now be proved by applying Itô's formula to the function $\varphi(r) = r^{p/2}$ and using identity (3.457).

Uniqueness of variational solutions for any $p \geq 2$ follows from Proposition 3.129-(i).

Let now $2 \leq p < 8$. For $n \geq 1$ we denote $\Omega^n := \{\omega \in \Omega : |\xi(\omega)| \leq n\}$ and $\xi^n := \xi \mathbf{1}_{\Omega^n}$. Then $X(\cdot; t, \xi^n, a(\cdot)) = X(\cdot; t, \xi^m, a(\cdot))$ on Ω^n if $n \leq m$, and estimates (3.451) and (3.452) are true for the processes $X(\cdot; t, \xi^n, a(\cdot)), n \geq 1$. Therefore, the process $X(s) := \lim_{n\to+\infty} X(\cdot; t, \xi^n, a(\cdot))$ is well defined, (3.451) and (3.452) for $X(\cdot; t, \xi^n, a(\cdot))$ follow from Fatou's lemma, and it is easy to see that $X(\cdot)$ is a variational solution of (3.441).

(ii) Let $2p \geq 8$ and $X^n(\cdot)$ be the processes from part (i). It follows from Itô's formula, Hypothesis 3.126 and (3.446), that there is $M \geq 0$ such that the processes $X^n(\cdot)$ satisfy in this case

$$\mathbb{E}\left[\sup_{t\leq s\leq T} |X^n(s)|_1^p + \int_t^T |X^n(s)|_2^2 (1 + |X^n(s)|_1^{p-2}) ds\right] \leq M \quad \text{for all } n.$$

Therefore by passing to a weak limit we obtain that the variational solution from part (i) satisfies $X(\cdot) \in L^p(\Omega; L^\infty(t,T;V)) \cap M^2_\mu(t,T;V_2)$, and thus it is a strong solution. Moreover $(-A)^{\frac{1}{2}}(AX(r) - B(X(r), X(r)) + f(r, a(r))) \in M^2_\mu(t,T;V_{-1})$, the process

$$(-A)^{\frac{1}{2}} X(s) = (-A)^{\frac{1}{2}}\xi \\ + \int_t^s (-A)^{\frac{1}{2}}(AX(r) - B(X(r), X(r)) + f(r, a(r)))\, dr + \int_t^s (-A)^{\frac{1}{2}} dW_Q(r)$$

is a continuous process with values in V_{-1}, and $(-A)^{\frac{1}{2}}X(\cdot) \in M^2_\mu(t,T;V)$. Thus, by Theorem 3.101, $X(\cdot) \in L^p(\Omega; C([t,T];V))$ and $\mathbb{P}$-a.e.

$$
\begin{aligned}
|X(s)|_1^2 = |\xi|_1^2 - 2\int_t^s \left(|AX(r)|^2 + \langle f(r,a(r)), AX(r)\rangle\right) dr \\
-2\int_t^s \langle dW_Q(r), AX(r)\rangle + \mathrm{Tr}(Q_1)(s-t).
\end{aligned}
\tag{3.458}
$$

Estimates (3.453) and (3.454) now follow by standard arguments applying Itô's formula to the function $\varphi(r) = r^{p/2}$ and using identity (3.458). For $2 \le p < 8$ we proceed as in the proof of part (i).

(iii) Similarly to the proof of part (ii) of Theorem 3.102, if $p \ge 8$ and $X_1(\cdot)$ is the variational solution in the generalized reference probability space μ_1 and $X_1^n(\cdot)$ are the solutions of the approximating problems (3.456) in this space, then

$$
\mathcal{L}_{\mathbb{P}_1}(a_1(\cdot), X_1^n(\cdot)) = \mathcal{L}_{\mathbb{P}}(a(\cdot), X^n(\cdot)),
$$

and thus (3.455) follows since $X_1^n(s) \to X_1(s)$ in $L^2(\Omega_1; H)$ and $X^n(s) \to X(s)$ in $L^2(\Omega; H)$ for every $s \in [t,T]$. For $2 \le p < 8$ we have

$$
\mathcal{L}_{\mathbb{P}_1}(a_1(\cdot), X_1(\cdot; t, \xi_1^n, a_1(\cdot))) = \mathcal{L}_{\mathbb{P}}(a(\cdot), X(\cdot; t, \xi^n, a(\cdot))),
$$

which gives the claim in the limit as $n \to +\infty$. □

Without loss of generality we will always assume from now on that the Q-Wiener processes in the reference probability spaces have everywhere continuous paths.

PROPOSITION 3.129 *Let $0 \le t < T$ and $p \ge 2$. Let $\nu = \left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s\in[t,T]}, \mathbb{P}, W_Q\right)$ be a reference probability space, and let Hypothesis 3.126 be satisfied. Let ξ, η be $\mathscr{F}_t^t$ measurable H-valued random variables such that $\mathbb{E}^\nu[|\xi|^p + |\eta|^p] < +\infty$, and let $a(\cdot) \in \mathcal{U}_t^\nu$. Then:*

(i) *There exists a constant C independent of $t, \xi, \eta, a(\cdot)$ and μ, such that a.s. on Ω*

$$
|X(s) - Y(s)|^2 + \int_t^s |X(\tau) - Y(\tau)|_1^2 d\tau \le |\xi - \eta|^2 \exp\left\{\int_t^s C|X(\tau)|_1^2 d\tau\right\} \tag{3.459}
$$

for all $s \in [t,T]$, where $X(\cdot) = X(\cdot; t, \xi, a(\cdot)), Y(\cdot) = Y(\cdot; t, \eta, a(\cdot))$ are solutions of (3.441) with initial conditions $X(t) = \xi$ and $Y(t) = \eta$.

(ii) *If $|x|_1 \le R_1$ then there exists a constant $C = C(p, T, R, R_1, Q)$ such that*

$$
\mathbb{E}|X(s) - x|^p \le C(p, T, R, R_1, Q)(s-t), \tag{3.460}
$$

where $X(\cdot) = X(\cdot; t, x, a(\cdot))$.

(iii) *For every initial condition $x \in V$ there exists a modulus ω, independent of the reference probability spaces ν and controls $a(\cdot) \in \mathcal{U}_t^\nu$, such that*

$$
\mathbb{E}|X(s) - x|_1^2 \le \omega_x(s-t), \tag{3.461}
$$

where $X(\cdot) = X(\cdot; t, x, a(\cdot))$.

PROOF. (i) Denote $Z(s) = X(s) - Y(s)$. Then $Z(\cdot)$ satisfies

$$
Z(s) = \xi - \eta + \int_t^s AZ(\tau)d\tau + \int_t^s [B(Y(\tau), Y(\tau)) - B(X(\tau), X(\tau))]d\tau.
$$

Hence using (3.445) and (3.448) we obtain

$$
\begin{aligned}
|Z(s)|^2 &= |\xi-\eta|^2 - 2\int_t^s |Z(\tau)|_1^2 d\tau - \int_t^s b(Z(\tau),X(\tau),Z(\tau))d\tau \\
&\le |\xi-\eta|^2 - 2\nu\int_t^s |Z(\tau)|_1^2 d\tau + \int_t^s C|Z(\tau)|_1|X(\tau)|_1|Z(\tau)|d\tau \\
&\le |\xi-\eta|^2 - \nu\int_t^s |Z(\tau)|_1^2 d\tau + C\int_t^s |X(\tau)|_1^2|Z(\tau)|^2 d\tau. \qquad (3.462)
\end{aligned}
$$

Here we have used Young's inequality. Then it follows from Gronwall's lemma that

$$
|Z(s)|^2 \le |\xi-\eta|^2 \exp\{\int_t^s C|X(\tau)|_1^2 d\tau\} \quad \mathbb{P} \text{ a.s.}
$$

Plugging this back into (3.462) yields (3.459) with another constant C.

(ii) Denote $Y(s) = X(s) - x$. Then

$$
Y(s) = \int_t^s \left(AX(\tau) - B\left(X(\tau),X(\tau)\right) + f\left(\tau,a(\tau)\right)\right)d\tau + \int_t^s dW(\tau).
$$

Therefore applying Itô's formula, taking expectation, and using (3.445) and the Cauchy-Schwartz inequality, we obtain

$$
\begin{aligned}
\mathbb{E}|Y(s)|^p &\le \mathbb{E}\int_t^s p\langle AX(\tau) - B\left(X(\tau),X(\tau)\right) + f\left(\tau,a(\tau)\right)), Y(\tau)\rangle |Y(\tau)|^{p-2}d\tau \\
&+ \mathbb{E}\int_t^s \frac{p(p-1)}{2}\mathrm{tr}(Q)|Y(\tau)|^{p-2}d\tau \\
&\le -\frac{p}{2}\mathbb{E}\int_t^s |X(\tau)|_1^2|Y(\tau)|^{p-2}d\tau + C_p\mathbb{E}\int_t^s |x|_1^2|Y(\tau)|^{p-2}d\tau \\
&+ C(p,R,R_1,Q)\mathbb{E}\int_t^s (|Y(\tau)|^{p-1} + |Y(\tau)|^{p-2})d\tau \\
&+ p\mathbb{E}\int_t^s |b(X(\tau),X(\tau),x)||Y(\tau)|^{p-2}d\tau.
\end{aligned}
$$

Since

$$
|b(X(\tau),X(\tau),x)| \le C|X(\tau)|_1|X(\tau)||x|_1 \le \frac{1}{2}|X(\tau)|_1^2 + \frac{C^2}{2}|X(\tau)|^2|x|_1^2,
$$

plugging this into the previous inequality and using (3.452) finally yields

$$
\begin{aligned}
\mathbb{E}|Y(s)|^p &\le C(p,R,R_1,Q)\mathbb{E}\int_t^s (|Y(\tau)|^{p-1} + |Y(\tau)|^{p-2} + |X(\tau)|^2|Y(\tau)|^{p-2})d\tau \\
&\le C(p,T,R,R_1,Q)(s-t).
\end{aligned}
$$

(iii) If (3.461) is not satisfied then there are $\epsilon > 0, a_n(\cdot) \in \mathcal{U}_t$ (which we can assume to be $\mathscr{F}_s^{t,0}$-predictable) and $s_n \to t$ such that $\mathbb{E}|X_n(s_n) - x|_1^2 \ge \epsilon$ for all $n \ge 1$, where $X_n(\cdot) = X(\cdot;t,x,a_n(\cdot)$. By Corollary 2.21 and Proposition 3.128-(iii), we can assume that all $a_n(\cdot)$ are defined on the same reference probability space.

However it follows from (3.460) and (3.453) that, up to a subsequence, we have

$$
X_n(s_n) \to x \text{ strongly in } L^2(\Omega;H) \text{ and weakly in } L^2(\Omega;V).
$$

Since the weak sequential convergence in $L^2(\Omega;V)$ implies

$$
|x|_1^2 \le \liminf_{n\to\infty} \mathbb{E}|X(s_n)|_1^2,
$$

this, together with (3.453), implies $|x|_1^2 = \lim_{n\to\infty}\mathbb{E}|X(s_n)|_1^2$. Therefore $X(s_n) \to x$ strongly in $L^2(\Omega;V)$, contrary to our assumption. □

3.13.2. Value function. In this section we show continuity properties of the value function of the stochastic optimal control problem and the Dynamic Programming Principle.

To minimize non-essential technical difficulties we will assume that the running cost function l is independent of t. The case of l depending on t is a straightforward extension of the methods presented here. The continuity of the value function is not entirely trivial since continuous dependence estimates in the mean for solutions of the stochastic Navier-Stokes equation depend on exponential moments of solutions (3.459) and these seem to be bounded only for a short time (see [443, Corollary XI.3.1], also [425]). We make the following assumptions about the cost functions l and g.

HYPOTHESIS 3.130 *The functions $l : V \times \Lambda \to \mathbb{R}$, and $g : H \to \mathbb{R}$ are continuous and there exist $k \geq 0$ and for every $r > 0$ a modulus σ_r such that*

$$|l(x, a)|, |g(x)| \leq C(1 + |x|_1^k) \quad \textit{for all } x \in V, a \in \Lambda, \tag{3.463}$$

$$|l(x, a) - l(y, a)| \leq \sigma_r (|x - y|_1) \quad \textit{if } |x|_1, |y|_1 \leq r, a \in \Lambda \tag{3.464}$$

$$|g(x) - g(y)| \leq \sigma_r (|x - y|) \quad \textit{if } |x|_1, |y|_1 \leq r. \tag{3.465}$$

PROPOSITION 3.131 *Let Hypotheses 3.126 and 3.130 be satisfied. Then:*

(i) *For every $r > 0$ there exists a modulus ω_r such that for every $t \in [0, T], a(\cdot) \in \mathcal{U}_t$*

$$|J(t, x; a(\cdot)) - J(t, y; a(\cdot))| \leq \omega_r (|x - y|) \quad \textit{if } |x|_1, |y|_1 \leq r. \tag{3.466}$$

(ii) *The value function v satisfies the Dynamic Programming Principle, i.e. for every $0 \leq t \leq \eta \leq T$ and $x \in V$,*

$$v(t, x) = \inf_{a(\cdot) \in \mathcal{U}_t} \mathbb{E} \left\{ \int_t^\eta l(X(s; t, x, a(\cdot)), a(s)) ds + v(\eta, X(\eta; t, x, a(\cdot))) \right\}. \tag{3.467}$$

(iii) *For every $r > 0$ there exists a modulus ω_r such that*

$$|v(t_1, x) - v(t_2, y)| \leq \omega_r (|t_1 - t_2| + |x - y|) \tag{3.468}$$

for all $t_1, t_2 \in [0, T]$ and $|x|_1, |y|_1 \leq r$, and there exists $C \geq 0$ such that

$$|v(t, x)| \leq C(1 + |x|_1^k) \tag{3.469}$$

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

PROOF. (i) Let $x, y \in V, t \in [0, T]$ and $a(\cdot) \in \mathcal{U}_t$. For every $m > 0$ let D_m be a constant such that $\sigma_m(s) \leq \frac{1}{m} + D_m s$. Denote $X(s) = X(s; t, x, a(\cdot)), Y(s) = Y(s; t, y, a(\cdot))$, and $A_m = \{\omega \in \Omega : \max_{t \leq s \leq T} |X(s)|_1 \leq m\}, B_m = \{\omega \in \Omega : \max_{t \leq s \leq T} |Y(s)|_1 \leq m\}$. Then, using (3.454), (3.459), (3.463) and (3.464), we obtain

$$\mathbb{E} \int_t^T |l(X(s), a(s)) - l(Y(s), a(s))| ds \leq \frac{T}{m} + \mathbb{E} \int_t^T D_m |X(s) - Y(s)|_1 \, \mathbf{1}_{A_m \cap B_m} ds$$
$$+ \mathbb{E} \int_t^T C(2 + |X(s)|_1^k + |Y(s)|_1^k) \, \mathbf{1}_{\Omega \setminus (A_m \cap B_m)} ds$$
$$\leq \frac{T}{m} + D_m |x - y| \mathbb{E} \int_t^T \exp \left\{ C \int_t^s |X(\tau)|_1^2 d\tau \right\} \mathbf{1}_{A_m} \, ds$$
$$+ \int_t^T C \left(2 + (\mathbb{E}|X(s)|_1^{2k})^{\frac{1}{2}} + (\mathbb{E}|Y(s)|_1^{2k})^{\frac{1}{2}} \right) \left((\mathbb{P}(\Omega \setminus A_m))^{\frac{1}{2}} + (\mathbb{P}(\Omega \setminus B_m))^{\frac{1}{2}} \right) ds$$
$$\leq \frac{T}{m} + D_m T |x - y| e^{CTm^2} + C_1(p, T, R, Q_1)(1 + |x|_1^k + |y|_1^k) \frac{1 + |x|_1 + |y|_1}{m}.$$

Applying the same process to estimate $|g(X(T)) - g(Y(T))|$ we therefore obtain that for every $r, m > 0$ there exist constants c_m, d_r such that for every $t \in [0,T]$ and $a(\cdot) \in \mathcal{U}_t$

$$|J(t,x;a(\cdot)) - J(t,y;a(\cdot))| \le \frac{d_r}{m} + c_m|x-y| \quad \text{if } |x|_1, |y|_1 \le r.$$

Estimate (3.466) now follows by taking the infimum over all $m > 0$.

(ii) We need to show that the problem satisfies the assumptions of Hypothesis 2.12. However here the statement of the DPP is restricted to points in V but the filtrations are still generated by Q-Wiener processes with values in H. Thus to proceed with the proof of the DPP described in Section 2.3 it is enough to assume in Hypothesis 2.12 that the random variable ξ there satisfies $\mathbb{E}^\mu|\xi|_1^2 < +\infty$. We recall that in this case we have strong solutions so conditions $(A0)$ and $(A2)$ follow from the definition of solution and standard arguments. In particular $(A0)$ follows from (3.459) and $(A2)$ from an obvious generalization of (3.459). Condition $(A1)$ follows from Proposition 3.128-(iii). We point out that if $\mathcal{L}_{\mathbb{P}_1}(X_1(\cdot;t_1,x,a_1(\cdot)),a_1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(X_2(\cdot;t_1,x,a_2(\cdot)),a_2(\cdot))$ as processes with values in $H \times \Lambda$ then, by Lemma 1.17-(i), they have the same laws as processes with values in $V \times \Lambda$. The proof of condition $(A3)$ starts as the proof of it in Proposition 2.16 however the arguments are now obvious since here the stochastic integral is just $W_{t_1}(s)$.

(iii) We notice that (3.466) implies

$$|v(t,x) - v(t,y)| \le \omega_r(|x-y|) \tag{3.470}$$

for all $t \in [0,T]$ and $|x|_1, |y|_1 \le r$. Moreover (3.469) is a direct consequence of (3.454) and (3.463).

Let now $0 \le t_1 < t_2 \le T, x \in V, |x|_1 \le r$. We will denote $X(s) = X(s;t_1,x,a(\cdot))$. Using (3.454), (3.460), (3.467), (3.469), (3.470), we obtain for $m > r$,

$$\begin{aligned}
|v(t_1,x) - v(t_2,x)| &\le \sup_{a(\cdot)\in\mathcal{U}_{t_1}} \mathbb{E}\int_{t_1}^{t_2} (1+|X(s)|_1^k)ds \\
&\qquad + \sup_{a(\cdot)\in\mathcal{U}_{t_1}} \mathbb{E}|v(t_2,X(t_2)) - v(t_2,x)| \\
&\le C(R,T,Q_1,r)(t_2-t_1) + \sup_{a(\cdot)\in\mathcal{U}_{t_1}} \left\{\mathbb{E}\left(C(1+|X(t_2)|_1^k + |x|_1^k)\mathbf{1}_{\{|X(t_2)|_1>m\}}\right)\right\} \\
&\qquad + \sup_{a(\cdot)\in\mathcal{U}_{t_1}} \mathbb{E}\sigma_m(|X(t_2)-x|) \\
&\le C(R,T,Q_1,r)(t_2-t_1) + C(R,T,Q_1)(1+|x|_1^k)\frac{1+|x|_1}{m} \\
&\qquad + \sigma_m\left(C(R,Q,r)(t_2-t_1)^{\frac{1}{2}}\right).
\end{aligned}$$

(We also used that σ_m above can be assumed to be concave.) The result now follows by taking the infimum over $m > r$. □

3.13.3. Viscosity solutions and comparison theorem. Since we only have continuity of the value function on $[0,T] \times V$, the definition of viscosity solution has to be restricted to this space. From the point of view of the HJB equation it might be better to set it up in this space, however because of the associated control problem, we want to keep H as our reference space. We achieve it by a proper choice of test functions. By using a special radial function of $|\cdot|_1$ as test function we first restrict the points where maxima or minima occur in the definition of viscosity sub/solution to be in $(0,T) \times V$. Then we require that the points where the maxima/minima occur belong to $(0,T) \times V_2$. Having this property we can interpret all terms appearing in the HJB equation. In this way we gain some coercive terms which had to be discarded in the generic definition given in Section 3.3, which are very useful in the proof of comparison principle. The definition is meaningful as we are able to show, using properties of the Navier-Stokes equation and the coercivity of the operator $-A$, that the value function is a viscosity solution. The definition of viscosity solution here is thus similar to the one used in Sections 3.11 and 3.12, however we use a radial test function of a different type. If different continuity requirements were imposed in Hypothesis 3.130, we would have different continuity properties of the value function, and then we could work with a definition of viscosity solution which resembles more the definition from Section 3.11, as it was done for first order equations in [243].

DEFINITION 3.132 *A function ψ is a test function for equation (3.443) if $\psi = \varphi + \delta(t)(1+|x|_1^2)^m$, where*

(i) $\varphi \in C^{1,2}((0,T)\times H)$, *and is such that $\varphi, \varphi_t, D\varphi, D^2\varphi$ are uniformly continuous on $[\epsilon, T-\epsilon]\times H$ for every $\epsilon > 0$.*

(ii) $\delta \in C^1((0,T))$ *is such that $\delta > 0$ on $(0,T)$, and $m \geq 1$.*

The function $h(t,x) = \delta(t)(1+|x|_1^2)^m$ is not Fréchet differentiable in H. Therefore the terms involving Dh and D^2h, in particular $\langle Ax - B(x,x), Dh(t,x)\rangle$ and $\mathrm{Tr}(QD^2h(t,x))$, have to be understood properly. We define

$$Dh(t,x) := -\delta(t)\left(2m(1+|x|_1^2)^{m-1}Ax\right),$$

and we will write

$$D\psi := D\varphi + Dh$$

even though this is a slight abuse of notation. Then, if $(t,x) \in (0,T)\times V_2$, $D\psi(t,x)$ makes sense, and so does the term $\langle Ax - B(x,x), D\psi(t,x)\rangle$. As regards the term $\mathrm{Tr}(QD^2\psi(t,x))$, without defining $D^2h(t,x)$, we interpret is by defining

$$\begin{aligned}\mathrm{Tr}(QD^2\psi(t,x)) := \mathrm{Tr}(QD^2\varphi(t,x)) + \delta(t)\big(&2m(1+|x|_1^2)^{m-1}\mathrm{Tr}(Q_1)\\ &+4m(m-1)(1+|x|_1^2)^{m-2}|Q^{\frac12}Ax|^2\big).\end{aligned}$$

It will be seen in the next section that the above interpretations appear as direct consequences of Itô's formula applied to h.

We give a definition of viscosity solution for a general equation (3.443) where the Hamiltonian function F is not necessarily given by (3.444). Thus we assume in this section that $F : [0,T]\times V\times H \to \mathbb{R}$ is any function.

DEFINITION 3.133 *A weakly sequentially upper-semicontinuous (respectively, lower-semicontinuous) function $u : (0,T]\times V \to \mathbb{R}$ is called a viscosity subsolution (respectively, supersolution) of (3.443) if $u(T,y) \leq h(y)$ (respectively, $u(T,y) \geq h(y)$) for all $y \in V$ and if, for every test function ψ, whenever $u-\psi$ has a global maximum (respectively $u+\psi$ has a global minimum) over $(0,T)\times V$ at (t,x), then $x \in V_2$ and*

$$\psi_t(t,x) + \frac12\mathrm{Tr}(QD^2\psi(t,x)) + \langle Ax - B(x,x), D\psi(t,x)\rangle + F(t,x,D\psi(t,x)) \geq 0$$

(respectively,

$$-\psi_t(t,x) - \frac12\mathrm{Tr}(QD^2\psi(t,x)) - \langle Ax - B(x,x), D\psi(t,x)\rangle + F(t,x,-D\psi(t,x)) \leq 0.)$$

A function u is a viscosity solution of (3.443) if it is both a viscosity subsolution and a viscosity supersolution of (3.443).

HYPOTHESIS 3.134 $F : [0,T]\times V\times H \to \mathbb{R}$ *and there exist a modulus of continuity ω, and moduli ω_r such that for every $r > 0$ we have*

$$|F(t,x,p) - F(t,y,p)| \leq \omega_r(|x-y|_1) + \omega(|x-y|_1|p|), \quad \text{if } |x|_1, |y|_1 \leq r, \quad (3.471)$$

$$|F(t,x,p) - F(t,x,q)| \le \omega\left((1+|x|_1)|p-q|\right), \tag{3.472}$$

$$|F(t,x,p) - F(s,x,p)| \le \omega_r(|t-s|), \ \ \textit{if } |x|_1, |p|_1 \le r, \tag{3.473}$$

$$|g(x) - g(y)| \le \omega_r(|x-y|), \ \ \textit{if } |x|_1, |y|_1 \le r. \tag{3.474}$$

THEOREM 3.135 *Let Hypothesis 3.134 hold. Let $u, v : (0,T] \times V \to \mathbb{R}$ be, respectively, a viscosity subsolution, and a viscosity supersolution of (3.443). Let*

$$u(t,x), \ -v(t,x), \ |g(x)| \le C(1+|x|_1^k) \tag{3.475}$$

for some $k \ge 0$. Then $u \le v$ on $(0,T] \times V$.

PROOF. We observe that weak sequential upper-semicontinuity of u and weak sequential lower-semicontinuity of v imply that

$$\begin{cases} \lim_{t\uparrow T}(u(t,x) - g(x))^+ = 0 \\ \lim_{t\uparrow T}(v(t,x) - g(x))^- = 0 \end{cases} \tag{3.476}$$

uniformly on bounded subsets of V. We define for $\mu > 0$,

$$u_\mu(t,x) = u(t,x) - \frac{\mu}{t}, \quad v_\mu(t,x) = v(t,x) + \frac{\mu}{t}.$$

Then u_μ and v_μ are, respectively, a viscosity subsolution, and a viscosity supersolution of

$$(u_\mu)_t + \frac{1}{2}\mathrm{Tr}(QD^2u_\mu) + \langle Ax - B(x,x), Du_\mu\rangle + F(t,x,Du_\mu) = \frac{\mu}{T^2}$$

and

$$(v_\mu)_t + \frac{1}{2}\mathrm{Tr}(QD^2u_\mu) + \langle Ax - B(x,x), Dv_\mu\rangle + F(t,x,Dv_\mu) = -\frac{\mu}{T^2}.$$

Let m be a number such that $m \ge 1$ and $2m \ge k+1$. For $0 < \epsilon, \delta, \beta \le 1$, we consider the function

$$\begin{aligned}\Phi(t,s,x,y) &= u_\mu(t,x) - v_\mu(s,y) - \frac{|x-y|^2}{2\epsilon} \\ &\quad -\delta e^{K_\mu(T-t)}(1+|x|_1^2)^m - \delta e^{K_\mu(T-s)}(1+|y|_1^2)^m - \frac{(t-s)^2}{2\beta}\end{aligned}$$

and set

$$\Phi(t,s,x,y) = -\infty \quad \text{if } x, y \notin V,$$

The constant K_μ will be chosen later. Obviously $\Phi \to -\infty$ as $\max(|x|_1, |y|_1) \to +\infty$. We claim that Φ is weakly sequentially upper-semicontinuous on $(0,T] \times (0,T] \times H \times H$.

It is well known that functions $x \to (1+|x|_1^2)^m, y \to (1+|y|_1^2)^m$ and $|x-y|^2$ are weakly sequentially lower-semicontinuous, respectively in H and $H \times H$. To show that, say,

$$u_\mu(t,x) - \delta e^{K_\mu(T-t)}(1+|x|_1^2)^m$$

is weakly sequentially upper-semicontinuous on $(0,T] \times H$, we suppose that this is not the case, i.e. there exist sequences $t_n \to t \in (0,T]$, $x_n \rightharpoonup x \in H$ such that

$$\limsup_{n\to\infty}\left(u_\mu(t_n,x_n) - \delta e^{K_\mu(T-t_n)}(1+|x_n|_1^2)^m\right) > u_\mu(t,x) - \delta e^{K_\mu(T-t)}(1+|x|_1^2)^m.$$

If $\liminf_{n\to\infty}|x_n|_1 = +\infty$, this is impossible by (3.475). So there must exist a subsequence (still denoted by (t_n, x_n)) such that $\limsup_{n\to\infty}|x_n|_1 < +\infty$. But then we have $x_n \rightharpoonup x$ in V, which contradicts the weak sequential upper-semicontinuity of u_μ.

Therefore Φ has a global maximum over $(0,T] \times (0,T] \times H \times H$ at some point $(\bar{t}, \bar{s}, \bar{x}, \bar{y}) \in (0,T] \times (0,T] \times V \times V$, where $|\bar{x}|_1, |\bar{y}|_1$ are bounded independently of

ϵ, β for a fixed δ. We can assume that the maximum is strict. By the definition of viscosity solution, $\bar{x}, \bar{y} \in V_2$. Moreover it is standard to notice that

$$\lim_{\beta \to 0} \frac{(\bar{t} - \bar{s})^2}{2\beta} = 0 \quad \text{for fixed } \delta, \epsilon, \tag{3.477}$$

and

$$\lim_{\epsilon \to 0} \limsup_{\beta \to 0} \frac{|\bar{x} - \bar{y}|^2}{2\epsilon} = 0 \quad \text{for fixed } \delta. \tag{3.478}$$

If $u \not\leq v$ it then follows from (3.478), (3.477), (3.474) and (3.476) that for small μ and δ, we have $\bar{t}, \bar{s} < T$ if β and ϵ are sufficiently small.

We use the projections from the proof of Proposition 3.128. Let $\{e_1, e_2, ...\}$ be the orthonormal basis of H composed of eigenvectors of A, $H_N := \operatorname{span}\{e_1, ..., e_N\}$, P_N be the orthogonal projection in H onto H_N, and $Q_N = I - P_N$ for $N \geq 2$. We define

$$\hat{u}(t,x) = u_\mu(t,x) - \frac{\langle x, Q_N(\bar{x} - \bar{y})\rangle}{\epsilon} + \frac{|Q_N(\bar{x} - \bar{y})|^2}{2\epsilon} - \frac{|Q_N(x - \bar{x})|^2}{\epsilon} - \delta e^{K_\mu (T-t)}(1 + |x|_1^2)^m,$$

$$v_1(s,y) = v_\mu(s,y) - \frac{\langle y, Q_N(\bar{x} - \bar{y})\rangle}{\epsilon} + \frac{|Q_N(y - \bar{y})|^2}{\epsilon} + \delta e^{K_\mu (T-s)}(1 + |y|_1^2)^m,$$

and we set $\hat{u}(t,x) = -\infty, \hat{v}(s,y) = +\infty$ if $x, y \notin V$. Then $\hat{u}, \hat{v}$ are respectively weakly sequentially upper- and lower-semicontinuous on $(0,T] \times H$ and it is easy to see (as in the proof of Theorem 3.50) that

$$\hat{u}(t,x) - \hat{v}(s,y) - \frac{|P_N(x-y)|^2}{2\epsilon} - \frac{(t-s)^2}{2\beta}$$

attains a strict global maximum over $(0,T] \times (0,T] \times H \times H$ at $(\bar{t}, \bar{s}, \bar{x}, \bar{y})$. Moreover the functions $\hat{u}, -\hat{v}$ satisfy (3.55). Therefore the assumptions of Corollary 3.29 are satisfied for $B = I$ there. Therefore there exist functions $\varphi_k, \psi_k \in C^2((0,T) \times H)$ for $k = 1, 2, ...$ such that $\varphi_k, (\varphi_k)_t, D\varphi_k, D^2\varphi_k, \psi_k, (\psi_k)_t, D\psi_k, D^2\psi_k$ are bounded and uniformly continuous, and such that

$$\hat{u}(t,x) - \varphi_k(t,x)$$

has a global maximum at some point $(t_k, x_k) \in (0,T) \times V$,

$$\hat{v}(s,y) - \psi_k(s,y)$$

has a global minimum at some point $(s_k, y_k) \in (0,T) \times V$, and

$$\big(t_k, x_k, \hat{u}(t_k,x_k), (\varphi_k)_t(t_k,x_k), D\varphi_k(t_k,x_k), D^2\varphi_k(t_k,x_k)\big) \xrightarrow[\mathbb{R}\times H\times\mathbb{R}\times H\times\mathcal{L}(H)]{k\to\infty} \left(\bar{t}, \bar{x}, \hat{u}(\bar{t},\bar{x}), \frac{\bar{t}-\bar{s}}{\beta}, \frac{P_N(\bar{x}-\bar{y})}{\epsilon}, X_N\right) \tag{3.479}$$

$$\big(s_k, y_k, \hat{v}(s_k,y_k), (\psi_k)_t(s_k,y_k), D\psi_k(s_k,y_k), D^2\psi_k(s_k,y_k)\big) \xrightarrow[\mathbb{R}\times H\times\mathbb{R}\times H\times\mathcal{L}(H)]{k\to\infty} \left(\bar{s}, \bar{y}, \hat{v}(\bar{s},\bar{y}), \frac{\bar{t}-\bar{s}}{\beta}, \frac{P_N(\bar{x}-\bar{y})}{\epsilon}, Y_N\right), \tag{3.480}$$

where $X_N = P_N X_N P_N, Y_N = P_N Y_N P_N$ and $X_N \leq Y_N$. Moreover, since the functions $\hat{u}, -\hat{v}$ are weakly sequentially upper-semicontinuous, it follows from the proof of Corollary 3.29 (see the proof of Theorem 3.27) that $\varphi_k(t,x) = \varphi_k(t, P_N x), \psi_k(s,y) = \psi_k(s, P_N y)$, and thus in particular we have

$$D\varphi_k(t_k,x_k) \to \frac{P_N(\bar{x}-\bar{y})}{\epsilon}, \quad D\psi_k(s_k,y_k) \to \frac{P_N(\bar{x}-\bar{y})}{\epsilon} \quad \text{in } H_1. \tag{3.481}$$

In addition it is easy to see that we must have $|x_k|_1, |y_k|_1 \le C$ for some C. Therefore $x_k \rightharpoonup \bar{x}, y_k \rightharpoonup \bar{y}$ in V. This, together with (3.479), (3.479), and the fact that $(\bar{t}, \bar{s}, \bar{x}, \bar{y})$ is the maximum point of Φ, implies

$$|x_k|_1 \to |\bar{x}|_1,\ |y_k|_1 \to |\bar{y}|_1,$$

which in turns gives

$$x_k \to \bar{x},\ y_k \to \bar{y} \quad \text{in } V. \tag{3.482}$$

By the definition of viscosity solution, we have $x_k, y_k \in V_2$, and

$$\begin{aligned}
- \quad & \delta K_\mu e^{K_\mu(T-t_k)}(1+|x_k|_1^2)^m + (\varphi_k)_t(t_k, x_k) \\
+ \quad & \frac{\delta}{2} e^{K_\mu(T-t_k)} \Big(2m\mathrm{Tr}(Q_1)(1+|x_k|_1^2)^{m-1} + 4m(m-1)|Q^{\frac{1}{2}} A x_k|^2 (1+|x_k|_1^2)^{m-2}\Big) \\
+ \quad & \frac{1}{2}\mathrm{Tr}\left(QD^2\varphi_k(t_k,x_k) + 2QQ_N\right) \\
+ \quad & \Big\langle Ax_k, D\varphi_k(t_k,x_k) + \frac{Q_N(\bar{x}-\bar{y})}{\epsilon} + \frac{2Q_N(x_k-\bar{x})}{\epsilon} \\
& \qquad -2m\delta e^{K_\mu(T-t_k)}(1+|x_k|_1^2)^{m-1}Ax_k \Big\rangle \\
- \quad & b\left(x_k, x_k, D\varphi_k(t_k,x_k) + \frac{Q_N(\bar{x}-\bar{y})}{\epsilon} + \frac{2Q_N(x_k-\bar{x})}{\epsilon}\right) \\
+ \quad & F\Bigg(t_k, x_k, D\varphi_k(t_k,x_k) + \frac{Q_N(\bar{x}-\bar{y})}{\epsilon} + \frac{2Q_N(x_k-\bar{x})}{\epsilon} \\
& \qquad -2m\delta e^{K_\mu(T-t_k)}(1+|x_k|_1^2)^{m-1}Ax_k\Bigg) \ge \frac{\mu}{T^2}.
\end{aligned} \tag{3.483}$$

Above we have used (3.446) to get $b\left(x_k, x_k, Ax_k\right) = 0$. We now want to pass to the limit as $k \to \infty$. Let C_μ be a constant such that

$$\omega(s) \le \frac{\mu}{2T^2} + C_\mu s.$$

It then follows from (3.472) that

$$\begin{aligned}
&\Bigg|F\Bigg(t_k, x_k, D\varphi_k(t_k,x_k) + \frac{Q_N(\bar{x}-\bar{y})}{\epsilon} + \frac{2Q_N(x_k-\bar{x})}{\epsilon} \\
&\qquad -2m\delta e^{K_\mu(T-t_k)}(1+|x_k|_1^2)^{m-1}Ax_k\Bigg) \\
&\quad -F\left(t_k, x_k, D\varphi_k(t_k,x_k) + \frac{Q_N(\bar{x}-\bar{y})}{\epsilon} + \frac{2Q_N(x_k-\bar{x})}{\epsilon}\right)\Bigg| \\
&\quad \le \frac{\mu}{2T^2} + C_\mu(1+|x_k|_1)2m\delta e^{K_\mu(T-t_k)}(1+|x_k|_1^2)^{m-1}|Ax_k|.
\end{aligned}$$

Moreover

$$\begin{aligned}
& C_\mu(1+|x_k|_1)2m\delta e^{K_\mu(T-t_k)}(1+|x_k|_1^2)^{m-1}|Ax_k| \\
+ \quad & \frac{\delta}{2} e^{K_\mu(T-t_k)} \Big(2m\mathrm{Tr}(Q_1)(1+|x_k|_1^2)^{m-1} + 4m(m-1)|Q^{\frac{1}{2}} A x_k|^2 (1+|x_k|_1^2)^{m-2}\Big) \\
\le \quad & 2m\delta C_\mu^2 e^{K_\mu(T-t_k)}(1+|x_k|_1^2)^m + m\delta e^{K_\mu(T-t_k)}|Ax_k|^2(1+|x_k|_1^2)^{m-1} \\
+ \quad & \delta e^{K_\mu(T-t_k)} m(2m-1)\mathrm{Tr}(Q_1)(1+|x_k|_1^2)^m \\
\le \quad & m\delta e^{K_\mu(T-t_k)}|Ax_k|^2(1+|x_k|_1^2)^{m-1} \\
+ \quad & \delta e^{K_\mu(T-t_k)}\left(2mC_\mu^2 + m(2m-1)\mathrm{Tr}(Q_1)\right)(1+|x_k|_1^2)^m.
\end{aligned} \tag{3.484}$$

Therefore, choosing $K_\mu = 1+2(2mC_\mu^2+m(2m-1)\mathrm{Tr}(Q_1))$ we obtain from (3.483) and (3.484) that

$$\begin{aligned}
- \quad & \frac{\delta}{2}K_\mu e^{K_\mu(T-t_k)}(1+|x_k|_1^2)^m + (\varphi_k)_t(t_k,x_k)\\
+ \quad & \frac{1}{2}\mathrm{Tr}\left(QD^2\varphi_k(t_k,x_k)+2QQ_N\right)\\
+ \quad & \Big\langle Ax_k, D\varphi_k(t_k,x_k)+\frac{Q_N(\bar{x}-\bar{y})}{\epsilon}+\frac{2Q_N(x_k-\bar{x})}{\epsilon}\\
& \qquad -m\delta e^{K_\mu(T-t_k)}Ax_k(1+|x_k|_1^2)^{m-1}\Big\rangle\\
- \quad & b\left(x_k,x_k,D\varphi_k(t_k,x_k)+\frac{Q_N(\bar{x}-\bar{y})}{\epsilon}+\frac{2Q_N(x_k-\bar{x})}{\epsilon}\right)\\
+ \quad & F\left(t_k,x_k,D\varphi_k(t_k,x_k)+\frac{Q_N(\bar{x}-\bar{y})}{\epsilon}+\frac{2Q_N(x_k-\bar{x})}{\epsilon}\right) \geq \frac{\mu}{2T^2}. \qquad (3.485)
\end{aligned}$$

Using (3.479), (3.481), (3.482), (3.471)-(3.473), and the continuity of b on $V\times V\times V$, we obtain from (3.485) that the norms $|Ax_k|$ are bounded and therefore $x_k \rightharpoonup \bar{x}$ in V_2. Therefore, using the above again, we can pass to the $\limsup$ as $n\to\infty$ in (3.485) to get

$$\begin{aligned}
- \quad & \frac{\delta}{2}K_\mu e^{K_\mu(T-\bar{t})}(1+|\bar{x}|_1^2)^m + \frac{\bar{t}-\bar{s}}{\beta} + \frac{1}{2}\mathrm{Tr}\left(QX_N+2QQ_N\right)\\
+ \quad & \langle A\bar{x}, \frac{\bar{x}-\bar{y}}{\epsilon}\rangle - b\left(\bar{x},\bar{x},\frac{\bar{x}-\bar{y}}{\epsilon}\right) + F\left(\bar{t},\bar{x},\frac{\bar{x}-\bar{y}}{\epsilon}\right) \geq \frac{\mu}{2T^2}. \qquad (3.486)
\end{aligned}$$

Similarly, we obtain

$$\begin{aligned}
& \frac{\delta}{2}K_\mu e^{K_\mu(T-\bar{s})}(1+|\bar{y}|_1^2)^m + \frac{\bar{t}-\bar{s}}{\beta} + \frac{1}{2}\mathrm{Tr}\left(QY_N-2QQ_N\right)\\
& +\langle A\bar{y}, \frac{\bar{x}-\bar{y}}{\epsilon}\rangle - b\left(\bar{y},\bar{y},\frac{\bar{x}-\bar{y}}{\epsilon}\right) + F\left(\bar{s},\bar{y},\frac{\bar{x}-\bar{y}}{\epsilon}\right) \leq -\frac{\mu}{2T^2}. \qquad (3.487)
\end{aligned}$$

Combining (3.486) and (3.487), using $X_N\leq Y_N$, and then sending $N\to\infty$ yields

$$\begin{aligned}
& \frac{\delta}{2}\left((1+|\bar{x}|_1^2)^m+(1+|\bar{y}|_1^2)^m\right) + \frac{|\bar{x}-\bar{y}|_1^2}{\epsilon}\\
& +b\left(\bar{x},\bar{x},\frac{\bar{x}-\bar{y}}{\epsilon}\right) - b\left(\bar{y},\bar{y},\frac{\bar{x}-\bar{y}}{\epsilon}\right)\\
& +F\left(\bar{t},\bar{x},\frac{\bar{x}-\bar{y}}{\epsilon}\right) - F\left(\bar{s},\bar{y},\frac{\bar{x}-\bar{y}}{\epsilon}\right) \leq -\frac{\mu}{T^2}. \qquad (3.488)
\end{aligned}$$

To estimate the trinilear form terms we use (3.445), (3.448), and then (3.478) to produce

$$\begin{aligned}
& \left|b\left(\bar{x},\bar{x},\frac{\bar{x}-\bar{y}}{\epsilon}\right) - b\left(\bar{y},\bar{y},\frac{\bar{x}-\bar{y}}{\epsilon}\right)\right|\\
& = \frac{1}{\epsilon}\left|b\left(\bar{x}-\bar{y},\bar{x},\bar{x}-\bar{y}\right)\right| \leq \frac{C}{\epsilon}|\bar{x}|_1|\bar{x}-\bar{y}||\bar{x}-\bar{y}|_1 \qquad (3.489)\\
& \leq \frac{\delta}{2}|\bar{x}|_1^2 + C_\delta\frac{|\bar{x}-\bar{y}|^2}{\epsilon}\frac{|\bar{x}-\bar{y}|_1^2}{\epsilon} \leq \frac{\delta}{2}(1+|\bar{x}|_1^2)^m + \sigma_2(\beta,\epsilon;\delta,\mu)\frac{|\bar{x}-\bar{y}|_1^2}{\epsilon},
\end{aligned}$$

where, for fixed μ,δ, $\lim_{\epsilon\to0}\limsup_{\beta\to0}\sigma_2(\beta,\epsilon;\delta,\mu)=0$.

Finally we need to estimate the terms containing F. We know that for μ and δ fixed $|\bar{x}|_1,|\bar{y}|_1\leq R_\delta$ for some $R_\delta>0$. Let $D_{\mu,\delta}$ be a constant such that

$$\omega_{R_\delta}(s) \leq \frac{\mu}{4T^2} + D_{\mu,\delta}s,$$

and denote $R_{\delta,\epsilon} := 2R_\delta/\epsilon$. Then (3.471), (3.473), (3.477) and (3.478) imply

$$\begin{aligned}
&\left| F\left(\bar{t}, \bar{x}, \frac{\bar{x}-\bar{y}}{\epsilon}\right) - F\left(\bar{s}, \bar{y}, \frac{\bar{x}-\bar{y}}{\epsilon}\right) \right| \\
&\leq \omega_{R_{\delta,\epsilon}}(|\bar{t}-\bar{s}|) + \omega_{R_\delta}(|\bar{x}-\bar{y}|_1) + \omega\left(|\bar{x}-\bar{y}|_1 \frac{|\bar{x}-\bar{y}|}{\epsilon}\right) \\
&\leq \omega_{R_{\delta,\epsilon}}(|\bar{t}-\bar{s}|) + \frac{3\mu}{4T^2} + D_{\mu,\delta}|\bar{x}-\bar{y}|_1 + C_\mu |\bar{x}-\bar{y}|_1 \frac{|\bar{x}-\bar{y}|}{\epsilon} \\
&\leq \omega_{R_{\delta,\epsilon}}(|\bar{t}-\bar{s}|) + \frac{3\mu}{4T^2} + D_{\mu,\delta}|\bar{x}-\bar{y}|_1 + \frac{|\bar{x}-\bar{y}|_1^2}{2\epsilon} + 2C_\mu^2 \frac{|\bar{x}-\bar{y}|^2}{\epsilon} \\
&\leq \frac{3\mu}{4T^2} + \sigma_3(\beta,\epsilon;\delta,\mu) + D_{\mu,\delta}|\bar{x}-\bar{y}|_1 + \frac{|\bar{x}-\bar{y}|_1^2}{2\epsilon},
\end{aligned} \tag{3.490}$$

where, for fixed μ, δ, $\lim_{\epsilon\to 0} \limsup_{\beta\to 0} \sigma_3(\beta,\epsilon;\delta,\mu) = 0$.

Therefore, using (3.489) and (3.490) in (3.488), we obtain

$$\left(\frac{1}{2} - \sigma_2(\beta,\epsilon;\delta,\mu)\right) \frac{|\bar{x}-\bar{y}|_1^2}{\epsilon} - D_{\mu,\delta}|\bar{x}-\bar{y}|_1 \leq -\frac{\mu}{4T^2} + \sigma_3(\beta,\epsilon,\delta;\mu). \tag{3.491}$$

We now notice that if ϵ and β are small, then $\frac{1}{2} - \sigma_2(\beta,\epsilon;\delta,\mu) > \frac{1}{4}$ and that

$$\lim_{\epsilon\to 0} \inf_{r>0} \left(\frac{r^2}{4\epsilon} - D_{\mu,\delta} r\right) = 0.$$

Therefore, it remains to take $\liminf_{\epsilon\to 0} \liminf_{\beta\to 0}$ in (3.491) to obtain a contradiction, which proves that we must have $u \leq v$. $\square$

3.13.4. Existence of viscosity solutions. We go back to the HJB equation (3.443) with the Hamiltonian function F defined by (3.444) and show that the value function of the associated stochastic optimal control problem is its viscosity solution.

THEOREM 3.136 *Let Hypotheses 3.126 and 3.130 be satisfied, and let in addition $f : [0,T] \times \Lambda \to V$ be such that $f(\cdot, a)$ is uniformly continuous, uniformly for $a \in \Lambda$. Then the value function v defined by (3.442) is the unique viscosity solution of the HJB equation (3.443)-(3.444) within the class of viscosity solutions u satisfying*

$$|u(t,x)| \leq C(1 + |x|_1^k), \quad (t,x) \in (0,T] \times V,$$

for some $k \geq 0$.

PROOF. First of all we notice that under our assumptions, the Hamiltonian F in (3.444) satisfies Hypothesis 3.134. Moreover, by Proposition 3.131, the value function v satisfies (3.468), (3.469), and the Dynamic Programming Principle (3.467). In particular v is weakly sequentially continuous on $(0,T] \times V$. Therefore, if v is a viscosity solution of (3.443)-(3.444), the uniqueness part is a direct consequence of Theorem 3.135. We will only show that the value function is a viscosity supersolution. The proof that v is a viscosity subsolution is easier and uses the same techniques. To this end, let $\psi(t,x) = \varphi(t,x) + \delta(t)(1 + |x|_1^2)^m$ be a test function and let $v + \psi$ have a global minimum at $(t_0, x_0) \in (0,T) \times V$.

Step 1. We need to show that $x_0 \in V$. By (3.467), for every $\epsilon > 0$ there exists $a_\epsilon(\cdot) \in \mathcal{U}_{t_0}$ such that, writing $X_\epsilon(s)$ for $X(s; t_0, x_0, a_\epsilon(\cdot))$, we have

$$v(t_0, x_0) + \epsilon^2 > \mathbb{E}\left\{\int_{t_0}^{t_0+\epsilon} l(X_\epsilon(s), a_\epsilon(s))\, ds + v(t_0+\epsilon, X_\epsilon(t_0+\epsilon))\right\}.$$

We can assume that a_ϵ is $\mathscr{F}_s^{t_0,0}$-predictable and thus, by Corollary 2.21 and Proposition 3.128-(iii), we can assume that all $a_\epsilon(\cdot)$ are defined on the same reference probability space ν, i.e. $a_\epsilon(\cdot) \in \mathcal{U}_{t_0}^{\nu}$. Since for every $(t,x) \in (0,T) \times V$

$$v(t,x) - v(t_0,x_0) \geq -\varphi(t,x) + \varphi(t_0,x_0) - \delta(t)(1+|x|_1^2)^m + \delta(t_0)(1+|x_0|_1^2)^m,$$

we have

$$\begin{aligned}
&\epsilon^2 - \mathbb{E}\int_{t_0}^{t_0+\epsilon} l(X_\epsilon(s), a_\epsilon(s))\,ds \geq \mathbb{E}\left[v(t_0+\epsilon, X_\epsilon(t_0+\epsilon)) - v(t_0,x_0)\right] \\
&\geq \mathbb{E}\Big[-\varphi(t_0+\epsilon, X_\epsilon(t_0+\epsilon)) + \varphi(t_0,x_0) \\
&\qquad -\delta(t_0+\epsilon)\left(1+|X_\epsilon(t_0+\epsilon)|_1^2\right)^m + \delta(t_0)\left(1+|x_0|_1^2\right)^m\Big].
\end{aligned}$$

Set $\lambda = \inf_{t\in[t_0,t_0+\epsilon_0]} \delta(t)$ for some fixed $\epsilon_0 > 0$, and take $\epsilon < \epsilon_0$. Applying Itô's formula to $\varphi(s, X_\epsilon(s))$ and $\delta(s)\left(1+|X_\epsilon(s)|_1^2\right)^m$, together with identity (3.458), in the inequality above, and then dividing both sides by ϵ, we obtain

$$\begin{aligned}
&\epsilon - \frac{1}{\epsilon}\mathbb{E}\int_{t_0}^{t_0+\epsilon} l(X_\epsilon(s), a_\epsilon(s))\,ds \\
&\geq -\frac{1}{\epsilon}\mathbb{E}\Big[\int_{t_0}^{t_0+\epsilon}\Big(\varphi_t(s, X_\epsilon(s)) + \langle AX_\epsilon(s) - B(X_\epsilon(s), X_\epsilon(s)), D\varphi(s, X_\epsilon(s))\rangle \\
&\qquad + \langle f(s, a_\epsilon(s)), D\varphi(s, X_\epsilon(s))\rangle + \frac{1}{2}\mathrm{Tr}\left(QD^2\varphi(s, X_\epsilon(s))\right)\Big)ds\Big] \\
&-\frac{1}{\epsilon}\mathbb{E}\Big[\int_{t_0}^{t_0+\epsilon}\Big(\delta'(s)\left(1+|X_\epsilon(s)|_1^2\right)^m + m\mathrm{Tr}(Q_1)\left(1+|X_\epsilon(s)|_1^2\right)^{m-1} \\
&\qquad -2m\delta(s)\left(|AX_\epsilon(s)|^2 + \langle f(s, a_\epsilon(s)), AX_\epsilon(s)\rangle\right)\left(1+|X_\epsilon(s)|_1^2\right)^{m-1} \\
&\qquad +2m(m-1)|Q^{\frac{1}{2}}AX_\epsilon(s)|^2\left(1+|X_\epsilon(s)|_1^2\right)^{m-2}\Big)ds\Big].
\end{aligned} \tag{3.492}$$

By the definition of λ it then follows that

$$\begin{aligned}
&\frac{2m\lambda}{\epsilon}\mathbb{E}\int_{t_0}^{t_0+\epsilon} |X_\epsilon(s)|_2^2\left(1+|X_\epsilon(s)|_1^2\right)^{m-1}ds \\
&\leq \epsilon + \frac{1}{\epsilon}\mathbb{E}\Big[\int_{t_0}^{t_0+\epsilon}\Big(-l(X_\epsilon(s), a_\epsilon(s)) + \varphi_t(s, X_\epsilon(s)) \\
&\qquad + \langle AX_\epsilon(s) - B(X_\epsilon(s), X_\epsilon(s)), D\varphi(s, X_\epsilon(s))\rangle \\
&\qquad + \langle f(s, a_\epsilon(s)), D\varphi(s, X_\epsilon(s))\rangle + \frac{1}{2}\mathrm{Tr}\left(QD^2\varphi(s, X_\epsilon(s))\right)\Big)ds\Big] \\
&+\frac{1}{\epsilon}\mathbb{E}\Big[\int_{t_0}^{t_0+\epsilon}\Big(\delta'(s)\left(1+|X_\epsilon(s)|_1^2\right)^m + m\mathrm{Tr}(Q_1)\left(1+|X_\epsilon(s)|_1^2\right)^{m-1} \\
&\qquad -2m\delta(s)\langle f(s, a_\epsilon(s)), AX_\epsilon(s)\rangle\left(1+|X_\epsilon(s)|_1^2\right)^{m-1} \\
&\qquad +2m(m-1)|Q^{\frac{1}{2}}AX_\epsilon(s)|^2\left(1+|X_\epsilon(s)|_1^2\right)^{m-2}\Big)ds\Big].
\end{aligned} \tag{3.493}$$

We now have

$$\begin{gathered}
|l(X_\epsilon(s), a_\epsilon(s))| \leq C\left(1+|X_\epsilon(s)|_1^k\right), \\
|\varphi_t(s, X_\epsilon(s))| \leq C(1+|X_\epsilon(s)|), \\
|\langle AX_\epsilon(s), D\varphi(s, X_\epsilon(s))\rangle| \leq \frac{\lambda}{2}|X_\epsilon(s)|_2^2 + C\left(1+|X_\epsilon(s)|^2\right), \\
|\langle B(X_\epsilon(s), X_\epsilon(s)), D\varphi(s, X_\epsilon(s))\rangle| = |b(X_\epsilon(s), X_\epsilon(s), D\varphi(s, X_\epsilon(s)))| \\
\leq C|X_\epsilon(s)|_1|X_\epsilon(s)|_2(1+|X_\epsilon(s)|) \leq \frac{\lambda}{2}|X_\epsilon(s)|_2^2 + C\left(1+|X_\epsilon(s)|_1^4\right),
\end{gathered}$$

$$|\text{Tr}\left(QD^2\varphi\left(s,X_\epsilon\left(s\right)\right)\right)|,|\left\langle f\left(s,a_\epsilon\left(s\right)\right),D\varphi\left(s,X_\epsilon\left(s\right)\right)\right\rangle|\le C(1+|X_\epsilon\left(s\right)|),$$
$$|\left\langle f\left(s,a_\epsilon\left(s\right)\right),AX_\epsilon\left(s\right)\right\rangle|\left(1+|X_\epsilon\left(s\right)|_1^2\right)^{m-1}\le C\left(1+|X_\epsilon\left(s\right)|_1^2\right)^m,$$

and

$$|Q^{\frac{1}{2}}AX_\epsilon\left(s\right)|^2\left(1+|X_\epsilon\left(s\right)|_1^2\right)^{m-2}\le C\left(1+|X_\epsilon\left(s\right)|_1^2\right)^{m-1}.$$

Employing the above estimates in (3.493) and then using (3.454) yields

$$\frac{\lambda}{\epsilon}\int_{t_0}^{t_0+\epsilon}\mathbb{E}|X_\epsilon\left(s\right)|_2^2\left(1+|X_\epsilon\left(s\right)|_1^2\right)^{m-1}ds\le C \tag{3.494}$$

for some constant C independent of ϵ. Therefore there exist sequences $\epsilon_n\to 0$ and $t_n\in(t_0,t_0+\epsilon_n)$ such that

$$\mathbb{E}|X_{\epsilon_n}\left(t_n\right)|_2^2\le C,$$

and thus there exist subsequences, still denoted by ϵ_n, t_n, such that

$$X_{\epsilon_n}\left(t_n\right)\rightharpoonup\bar{x}\quad\text{weakly in } L^2(\Omega^\nu;V_2)$$

for some $\bar{x}\in L^2(\Omega^\nu;V_2)$ (and thus also weakly in $L^2(\Omega^\nu;H)$). However, by (3.460), $X_{\epsilon_n}\left(t_n\right)\to x_0$ strongly in $L^2(\Omega^\nu;H)$. Therefore, by the uniqueness of the weak limit in $L^2(\Omega^\nu;H)$, it follows that $x_0=\bar{x}\in V_2$.

Step 2. We now prove the supersolution inequality. We need to "pass to the limit" as $\epsilon\to 0$ in (3.492), at least along a subsequence. This operation is rather standard for most of the terms, more precisely for those that only use convergence in the norms of H and V. To explain how we deal with the easy terms, let us consider the cost term.

Let $r\ge|x_0|_1$. Then, using (3.454), (3.461), (3.463) and (3.464), we have

$$\begin{aligned}
&\left|\frac{1}{\epsilon}\mathbb{E}\int_{t_0}^{t_0+\epsilon}\left[l\left(X_\epsilon\left(s\right),a_\epsilon\left(s\right)\right)ds-l\left(x_0,a_\epsilon\left(s\right)\right)\right]ds\right|\\
&\quad\le\frac{1}{\epsilon}\mathbb{E}\int_{t_0}^{t_0+\epsilon}\sigma_r\left(|X_\epsilon\left(s\right)-x_0|_1\right)ds\\
&\qquad+\frac{1}{\epsilon}\mathbb{E}\int_{t_0}^{t_0+\epsilon}C\left(1+|X_\epsilon\left(s\right)|_1^k+|x_0|_1^k\right)\mathbf{1}_{\{|X_\epsilon(s)|_1>r\}}ds\\
&\quad\le\sigma_r\left(\frac{1}{\epsilon}\mathbb{E}\int_{t_0}^{t_0+\epsilon}|X_\epsilon\left(s\right)-x_0|_1ds\right)+C\left(1+|x_0|_1^k\right)\frac{1+|x_0|_1}{r}\\
&\quad\le\sigma_r\left(\frac{1}{\epsilon}\mathbb{E}\int_{t_0}^{t_0+\epsilon}\sqrt{\omega_{x_0}(\epsilon)}ds\right)+C\left(1+|x_0|_1^k\right)\frac{1+|x_0|_1}{r}.
\end{aligned}\tag{3.495}$$

The above implies that

$$\left|\frac{1}{\epsilon}\mathbb{E}\int_{t_0}^{t_0+\epsilon}\left[l\left(X_\epsilon\left(s\right),a_\epsilon\left(s\right)\right)ds-l\left(x_0,a_\epsilon\left(s\right)\right)\right]ds\right|\le\gamma(\epsilon),\tag{3.496}$$

where $\lim_{\epsilon\to 0}\gamma(\epsilon)=0$. Arguing like in (3.495), and using that $(-A)^{\frac{1}{2}}f$ is bounded in H, $(-A)^{\frac{1}{2}}f(\cdot,a)$ is uniformly continuous with values in H, uniformly for $a\in\Lambda$, and $Q^{\frac{1}{2}}(-A)^{\frac{1}{2}}$ extends to a bounded operator in H, we can deal with all the terms in (3.492), except the terms

$$-\frac{1}{\epsilon}\mathbb{E}\int_{t_0}^{t_0+\epsilon}\left\langle AX_\epsilon\left(s\right),D\varphi\left(s,X_\epsilon\left(s\right)\right)\right\rangle,\tag{3.497}$$

$$\frac{1}{\epsilon}\mathbb{E}\int_{t_0}^{t_0+\epsilon}\left\langle B\left(X_\epsilon\left(s\right),X_\epsilon\left(s\right)\right),D\varphi\left(s,X_\epsilon\left(s\right)\right)\right\rangle,\tag{3.498}$$

$$\frac{1}{\epsilon}\mathbb{E}\int_{t_0}^{t_0+\epsilon}2m\delta\left(s\right)|AX_\epsilon(s)|^2\left(1+|X_\epsilon\left(s\right)|_1^2\right)^{m-1}\tag{3.499}$$

which require special consideration.

We first notice that

$$\mathbb{E}\left|\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon}\sqrt{\delta(s)}AX_\epsilon\left(s\right)\left(1+|X_\epsilon\left(s\right)|_1^2\right)^{\frac{m-1}{2}}ds\right|^2$$
$$\leq \mathbb{E}\frac{1}{\epsilon}\int_{t_0}^{t_0+\epsilon}\delta(s)|X_\epsilon\left(s\right)|_2^2\left(1+|X_\epsilon\left(s\right)|_1^2\right)^{m-1}ds \leq C \qquad (3.500)$$

by (3.494). Therefore, there exists a sequence $\epsilon_n \to 0$ and $Y \in L^2(\Omega^\nu, H)$ such that

$$Y_n := \frac{1}{\epsilon_n}\int_{t_0}^{t_0+\epsilon_n}\sqrt{\delta(s)}AX_{\epsilon_n}\left(s\right)\left(1+|X_{\epsilon_n}\left(s\right)|_1^2\right)^{\frac{m-1}{2}}ds \rightharpoonup Y \quad \text{in } L^2(\Omega^\nu, H)$$

as $n \to \infty$. However, using arguments similar to these in (3.495), it is easy to see that

$$A^{-1}Y_n = \frac{1}{\epsilon_n}\int_{t_0}^{t_0+\epsilon_n}\sqrt{\delta(s)}X_{\epsilon_n}(s)\left(1+|X_{\epsilon_n}(s)|_1^2\right)^{\frac{m-1}{2}}ds \to \sqrt{\delta(t_0)}x_0\left(1+|x_0|_1^2\right)^{\frac{m-1}{2}}$$

strongly in $L^2(\Omega^\nu, H)$. Therefore it follows that

$$Y = \sqrt{\delta(t_0)}Ax_0\left(1+|x_0|_1^2\right)^{\frac{m-1}{2}}.$$

Then, using the first inequality of (3.500), we get

$$\liminf_{n\to\infty}\mathbb{E}\frac{1}{\epsilon_n}\int_{t_0}^{t_0+\epsilon_n}\delta(s)|AX_{\epsilon_n}\left(s\right)|^2\left(1+|X_{\epsilon_n}\left(s\right)|_1^2\right)^{m-1}ds$$
$$\geq \delta(t_0)|Ax_0|^2\left(1+|x_0|_1^2\right)^{m-1}. \qquad (3.501)$$

This takes care of the term (3.499). The same argument also shows that we can assume that

$$\frac{1}{\epsilon_n}\int_{t_0}^{t_0+\epsilon_n}AX_{\epsilon_n}\left(s\right)ds \rightharpoonup Ax_0 \quad \text{in } L^2(\Omega^\nu, H) \text{ as } n \to \infty. \qquad (3.502)$$

As regards (3.497), denoting by ω_φ a modulus of continuity of $D\varphi$, we have by (3.494), (3.460), and (3.502)

$$\left|\frac{1}{\epsilon_n}\mathbb{E}_{t_0}\int_{t_0}^{t_0+\epsilon_n}\left\langle AX_{\epsilon_n}\left(s\right), D\varphi\left(s, X_{\epsilon_n}\left(s\right)\right)\right\rangle ds - \left\langle Ax_0, D\varphi\left(t_0, x_0\right)\right\rangle\right|$$
$$\leq \frac{1}{\epsilon_n}\int_{t_0}^{t_0+\epsilon_n}\left(\mathbb{E}|AX_{\epsilon_n}\left(s\right)|^2\right)^{\frac{1}{2}}\left(\mathbb{E}\left(\omega_\varphi\left(\epsilon_n+|X_{\epsilon_n}\left(s\right)-x_0|\right)\right)^2\right)^{\frac{1}{2}}ds \qquad (3.503)$$
$$+\left|\mathbb{E}\left\langle\frac{1}{\epsilon_n}\int_{t_0}^{t_0+\epsilon_n}AX_{\epsilon_n}\left(s\right)ds - Ax_0, D\varphi\left(t_0, x_0\right)\right\rangle\right| \to 0 \quad \text{as } n \to \infty.$$

Finally for (3.498), using (3.449), (3.494), (3.454), (3.460), (3.461), and (3.502),

$$\begin{aligned}
&\left|\frac{1}{\epsilon_n}\mathbb{E}\int_{t_0}^{t_0+\epsilon_n} b\left(X_{\epsilon_n}(s), X_{\epsilon_n}(s), D\varphi\left(s, X_{\epsilon_n}(s)\right)\right) ds - b\left(x_0, x_0, D\varphi\left(t_0, x_0\right)\right)\right| \\
&\leq \frac{1}{\epsilon_n}\mathbb{E}\int_{t_0}^{t_0+\epsilon_n} |X_{\epsilon_n}(s)|_1 |X_{\epsilon_n}(s)|_2 \omega_\varphi\left(\epsilon_n + |X_{\epsilon_n}(s) - x_0|\right) ds \\
&+\frac{1}{\epsilon_n}\mathbb{E}\int_{t_0}^{t_0+\epsilon_n} |X_{\epsilon_n}(s) - x_0|_1 |X_{\epsilon_n}(s)|_2 |D\varphi\left(t_0, x_0\right)| ds \\
&+\left|\frac{1}{\epsilon_n}\mathbb{E}\int_{t_0}^{t_0+\epsilon_n} b\left(x_0, X_{\epsilon_n}(s) - x_0, D\varphi\left(t_0, x_0\right)\right) ds\right| \\
&\leq \frac{1}{\epsilon_n}\int_{t_0}^{t_0+\epsilon_n} \left(\mathbb{E}|X_{\epsilon_n}(s)|_2^2\right)^{\frac{1}{2}} \left(\mathbb{E}|X_{\epsilon_n}(s)|_1^4\right)^{\frac{1}{4}} \left(\mathbb{E}\left(\omega_\varphi\left(\epsilon_n + |X_{\epsilon_n}(s) - x_0|\right)\right)^4\right)^{\frac{1}{4}} ds \\
&+C\frac{1}{\epsilon_n}\int_{t_0}^{t_0+\epsilon_n} \left(\mathbb{E}|X_{\epsilon_n}(s)|_2^2\right)^{\frac{1}{2}} \left(\mathbb{E}|X_{\epsilon_n}(s) - x_0|_1^2\right)^{\frac{1}{2}} ds \qquad (3.504) \\
&+\left|\mathbb{E} b\left(x_0, \frac{1}{\epsilon_n}\int_{t_0}^{t_0+\epsilon_n} X_{\epsilon_n}(s) ds - x_0, D\varphi\left(t_0, X_{\epsilon_n}(s)\right)\right)\right| \to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

In particular the last term goes to zero since, by (3.502),

$$\frac{1}{\epsilon_n}\int_{t_0}^{t_0+\epsilon_n} X_{\epsilon_n}(s) ds \rightharpoonup x_0 \quad \text{in } L^2(\Omega^\nu, V_2) \text{ as } n \to \infty$$

and

$$Z \to b\left(x_0, Z, D\varphi\left(t_0, x_0\right)\right)$$

is a bounded linear functional on $L^2(\Omega^\nu, V_2)$.

Therefore, using (3.496) (and similar estimates for other standard terms), (3.501), (3.503), and (3.504) in (3.492), we obtain for small ϵ_n that

$$-\psi_t(t_0, x_0) - \frac{1}{2}\mathrm{Tr}\left(QD^2\psi(t_0, x_0)\right) - \left\langle Ax_0 - B\left(x_0, x_0\right), D\psi\left(t, x_0\right)\right\rangle$$

$$+\frac{1}{\epsilon_n}\mathbb{E}\int_{t_0}^{t_0+\epsilon_n} \left[\left\langle f\left(t_0, a_\epsilon(s)\right), -D\psi\left(t_0, x_0\right)\right\rangle + l\left(x_0, a_\epsilon(s)\right)\right] ds \leq \omega_1\left(\epsilon_n\right)$$

for some modulus ω_1. It now remains to take the infimum over $a \in \Lambda$ inside the integral and then send $n \to \infty$. □

Example 3.137 The following example satisfies the assumptions of Theorem 3.136. Let

$$l(x, a) = |\mathrm{curl}\ x|^2 + \frac{1}{2}|a|^2,$$

$$g(x) = |x|^2,$$

$$f(t, a) = Ka,$$

where $K \in \mathcal{L}(H; V)$ and $\Lambda = B_H(0, R) \subset H$. Such a control and the singular kernel of K can be approximately realized by a suitable Lorentz force distribution in electrically conducting fluids such as liquid metals and salt water. The Hamiltonian function is then

$$F(x, p) = |\mathrm{curl}\ x|^2 + h(K^*p),$$

where $h(\cdot): H \to R$ is given by

$$h(z) := \inf_{a\in\Lambda}\left\{\langle a, z\rangle_H + \frac{1}{2}|a|^2\right\}$$

and K^* is the adjoint of K considered as an operator from H to H. We can in fact explicitly obtain h as

$$h(z) = \begin{cases} -\frac{1}{2}|z|^2 & \text{for } |z| \le R \\ \\ -R|z| + \frac{1}{2}R^2 & \text{for } |z| > R. \end{cases}$$

We also remark that the optimal feedback control here is given formally as

$$\tilde{a}(t) = \Upsilon(K^* Du(t, x(t))),$$

where

$$\Upsilon(z) := Dh(z) = \begin{cases} -z & \text{for } |z| \le R, \\ \\ -z\frac{R}{|z|} & \text{for } |z| > R. \end{cases}$$

Under additional conditions on Q, optimal feedback control for this example are discussed in Section 4.9.1.2 using mild solutions. ■

3.14. Bibliographical notes

The material of Section 3.1.1 about B-continuity is based on [103, 104] and [396]. The formulation and the proof of the exponential moment estimates of Proposition 3.18 is taken from [425] while the rest of Section 3.1.2 mostly follows [104, 286]. More general formulations of Theorem 3.25 are in [420, 312]. Corollary 3.26 was first introduced in [104]. Other smooth or partially smooth perturbed optimization principles can be found in [48, 133, 293].

A first version of a maximum principle for semicontinuous functions in Hilbert spaces appeared in [321]. It was an infinite dimensional version of a maximum principle in domains of $\mathbb{R}^n$ (see e.g. [101]) and was applicable to a class of bounded second order equations (3.68). By a reduction to a finite dimensional case and the use of the finite dimensional maximum principle it provided test functions whose second order derivatives satisfied proper inequalities on finite dimensional subspaces, and with the remaining parts of second order derivatives becoming negligible for the class of equations considered as the dimension of the finite dimensional subspaces increased to $+\infty$. A corrected and simplified proof of this result, which also included its time dependent version, based on the use of so called partial sup-convolutions appeared in [102]. These maximum principles have been adapted to unbounded equations first in [272, 421, 422] and later in various settings in [242, 244, 245, 286, 288]. The version stated in Theorem 3.27 is general and new as we tried to formulate it in a way that would be more directly applicable to various classes of equations. Its proof draws on the collective body of work from the above cited papers. A scaling reduction to obtain Corollary 3.29 was introduced in [272]. A different type of time dependent maximum principle, similar in the spirit to its finite dimensional version in [101], is in [102], see Remark 3.30 for more on this.

The definition of viscosity solution similar to the one presented in Section 3.3 was introduced in [421, 422]. It was based on the notion of B-continuous viscosity solution developed by Crandall and P. L. Lions in [103, 104]. An earlier paper [320] dealing with a specific second order HJB equation for an optimal control of a Zakai equation also used some ideas of the B-continuous viscosity solution. The material of Section 3.3 is mostly based on [421, 422, 286] however the definitions of viscosity solutions are more general. Lemma 3.37 is taken from [288]. A different definition of viscosity solution for second order equations in Hilbert spaces was proposed by Ishii in [272]. It was related to the definition for first order equations in [271]. Ishii's viscosity sub/super-solutions are allowed to be discontinuous and the definition uses a special (convex) function to deal with the unboundedness in the equation. The function is related to the equation and can be thought of as some energy function for the controlled deterministic/stochastic PDE related to the HJB equation. The advantage of this definition is that viscosity solutions can be relatively easily obtained by Perron's method and the unbounded operator A (together with other terms) can be non-linear. However the definition seems to be difficult to apply to control problems and no attempts have been made in this direction. The idea of using special functions as part of test functions in the definition of viscosity solution to exploit the coercivity of the operator A also appeared in [71, 106] (see also [183, 243]) and later for second order equations in [242, 244, 245, 288], see Sections 3.9, 3.11, 3.12, 3.13.

We only briefly mentioned bounded equations in Section 3.3.1. The definition of viscosity solution in Section 3.3.1 is taken from [321] where the theory of such equations was developed for equations satisfying Hypothesis 3.47. In this paper also an equivalent definition using second-order jets is discussed. For equations that do not satisfy Hypothesis 3.47 a stronger definition of viscosity solution was introduced in [319]. It allowed for more general test functions which are not necessarily twice Fréchet differentiable. Both papers contain comparison and existence results. Uniqueness of solutions is obtained in [319] by a combination of stochastic and analytic techniques. Perron's method is discussed in [321] while connections with stochastic optimal control are discussed in [319]. In particular [319] contains proofs of sub- and super-optimality inequalities of dynamic programming. Regularity results for bounded equations and their obstacle problems have been obtained in [319, 426]. Existence and uniqueness results for bounded equations can also be found in [290], in particular one can find there proofs of comparison principles using parabolic maximum principle from [102] which allows to relax the way viscosity subsolutions and supersolutions attain the initial/terminal values. A Dirichlet boundary value problem for a linear equation was investigated in [286] and a risk-sensitive control problem in [424]. An obstacle problem related to optimal stopping and pricing of American options was studied in [219]. Classical results for bounded linear equations can be found in [129].

First comparison theorems for B-continuous viscosity sub/super-solutions of equations discussed in Section 3.5 were proved in [421, 422]. These works dealt with the case of compact operator B and the proofs of comparison in most part relied on the combination of techniques developed for first order equations in [103, 104] and the maximum principle arguments of [102, 321]. Slightly different techniques but also based on the maximum principle of [321] were used to prove comparison principle with Ishii's definition of solution in [272]. Some of Ishii's methods were later used in other comparison proofs. Comparison results for equations of Section 3.5 without the compactness assumption on B are in [286] and the proofs of comparison in various special cases are contained in [242, 244, 245, 288, 292]. Comparison theorem for equations with quadratic gradient terms is in [425]. The papers [272, 288] show comparison for discontinuous viscosity sub- and super-solutions. The material of Section 3.5 is to some extent new and incorporates formulations and techniques of [272, 286, 421, 422]. The statements of Theorems 3.50, 3.54, 3.56, 3.58 are new and include general growth conditions for viscosity sub- and supersolutions. The proofs of the above comparison theorems are also to some extent new. Comparison theorems with general growth conditions for B-continuous viscosity sub/super-solutions of first order equations can be found in [312].

Direct proofs that value functions of stochastic optimal control problems in Hilbert spaces are viscosity solutions of their HJB equations were done in various cases in [245, 319, 320]. An early attempt in this direction was also made in [358]. The general finite time horizon optimal control problem (3.125)-(3.126) and its connection to B-continuous viscosity solutions of equation (3.128)-(3.129) was studied in [286]. Our presentation expands and generalizes [286]. Some results which were part of the folklore of the theory are stated in Section 3.6 for the first time. We presented continuity properties of the value functions in both finite and infinite horizon cases and under both weak and strong B-conditions. Only the weak B-condition case was discussed in [286]. The use of dynamic programming principle is also fully explained and the proofs that value functions are viscosity solutions of the associated HJB equations are done in all cases. We tried to include all the details. The proof of a stronger version of the dynamic programming principle in the stopping time formulation in Section 3.6.2 uses some arguments from [349].

The material of Section 3.7 about finite dimensional approximations is based on the results of [421, 422], however it contains some improvements of the results and their proofs. The method of finite dimensional approximations provides a way to construct B-continuous viscosity solutions for equations which may not be HJB equations related to optimal control problems, for instance for Isaacs equations. It requires however that the operator B be compact. The method, together with its basic techniques, was introduced in [103] for first order equations, and was later generalized to second order equations in [421, 422]. A version of this method was also used in [242]. Lemma 3.83-(ii) was proved in [103] by viscosity solution arguments. Our proof uses direct functional analytic arguments. Other proofs of existence employ Perron's method (see comments in this section in the paragraph about Perron's method). For Isaacs equations, probabilistic representation formulas can be obtained [192, 361, 363, 423].

Section 3.8 on singular perturbations is based on [425]. Singular perturbation problems in finite dimensional spaces have been studied extensively by viscosity solution methods and the reader can consult [30, 195] for results and references. The problems have not been widely investigated yet in Hilbert or other infinite dimensional spaces. Nisio studied such problem in [363] in connection with a risk sensitive control problem. Also a singular limit problem related to a risk-sensitive control problem with bounded evolution in a Hilbert space was studied in [424]. In [425] convergence of viscosity solutions of singularly perturbed HJB equations was used to investigate large deviation problems for stochastic PDE perturbed by small noise. The case of integro-PDE was studied in [428]. Both papers [425, 428] use a general PDE approach to large deviations developed in [184] (see also [180, 181, 182]).

Perron's method for viscosity solutions of PDE in finite dimensional spaces was introduced by Ishii in [269] (see also [101]). It was extended to bounded equations in Hilbert spaces in [321]. Perron's method provides another way to obtain existence of viscosity solutions for equations that may not necessarily be of the HJB type, for instance for Isaacs equations. For unbounded first and second order equations it was shown to work with Ishii's definitions of viscosity solution [271, 272] and with the Tataru-Crandall-Lions definition of viscosity solution [105]. For B-continuous viscosity solutions Perron's method was introduced in [288] under an assumption that the unbounded operator A has some coercivity properties. Perron's method requires the notion of a discontinuous viscosity solution so in [288] a more general definition of a discontinuous viscosity solution using B-semicontinuous envelopes was introduced. This definition borrowed an idea from the definitions in [271, 272] of combining the upper and lower-semicontinuous envelopes with the radial test functions. The method of half-relaxed limits of Barles-Perthame requires compactness. The fact that it may not work in infinite dimensional spaces was noticed in [9, 424] (see Example 3.43 in this book). A version of half-relaxed limits presented in Section 3.9 was developed in [288] where more results on Perron's method and half-relaxed limits can be found.

The material of Section 3.10 is based on [287]. Infinite dimensional Black-Scholes equation was analyzed in [227], where existence of smooth solutions was proved for smooth data, and an obstacle problem for the Black-Scholes equation was studied in [454] from the point of view of Bellman's inclusions. A similar obstacle problem for a related model was studied in [219]. A non-local Black-Scholes-Barenblatt equation associated with the HJMM model driven by a Lévy type noise was investigated in [427].

Section 3.11 follows [245]. First result about viscosity solutions of the HJB equation for control of the DMZ equation appeared in [320], where the equation was studied in a standard L^2 space and the operators S_a^k were bounded multiplication operators. The paper [320] used a combination of probabilistic and analytic techniques to deal with the uniqueness of viscosity solution of the HJB equation. In [259] it was shown that the value function is a viscosity solution in a very weak sense when the HJB equation was considered in the space of measures. (A regularity result for a related equation in the space of measures was obtained in [261].) Other paper on the subject is [15]. The approach of [245] used the theory of B-continuous viscosity solutions of [421, 422] together with an idea that originated in [71, 106] (also [271, 272] had related ideas) to use a special radial function and the coercivity of operators in the equation to "improve" the points where the maxima/minima occur in the definition of viscosity solution. This idea of using special energy function related to the underlying controlled state equation as a part of test functions was also used in may cases for first and second order equations [242, 243, 244, 288, 183]. The viscosity solution approach of [245] is also presented in [364], where a different proof of the dynamic programming principle is given. The book [364] discusses in details, in Chapters 5 and 6, a partially observed optimal control problem, the separated problem, the optimal control of the DMZ equation, and other related material. It complements the material in Sections 2.6.6 and 3.11 and gives a slightly different perspective. Our short introduction to variational solutions in Section 3.11.1 is based on [93, 220, 296, 298, 385]. Semi-linear stochastic parabolic equations and the DMZ equation are also discussed in [364], where Itô's formulas are proved for the original test functions used in [245], which are similar but different from the test functions in Section 3.11.5. Our presentation of the various energy and continuous dependence estimates for the DMZ equation follows, with small changes, [245]. The reader can also find similar results in [364]. The material on viscosity solutions and the value function of Sections 3.11.5 and 3.11.6 has some differences from [245] as we merged it into the presentation of the book and made some improvements and corrections.

Section 3.12 is based on [242] and fills some missing details there. The HJB equation in this section has second-order coefficients which are not trace class so the equation is also unbounded in the second-order terms. Since the equation is fully nonlinear it cannot be dealt with by the techniques of mild solutions. A change of variables is done to convert the equation to a one with bounded second-order terms. This is a rather ad hoc technique. Viscosity solutions of HJB equations with unbounded second-order terms coming from control problems with state equations driven by cylindrical Wiener processes, have not yet been studied systematically. The definition of viscosity solution for the converted equation is similar to this in [71] and uses a special radial function to guarantee that the points where the maxima/minima occur belong to a better space (see the previous paragraph for the discussion about the origins of such definition). A Cauchy problem for equations similar to the "converted" equation was also studied in [439] using the techniques of [242]. Apart from [242], boundary control problems and their associated HJB equations have been studied by viscosity solution techniques in [451] for the stochastic case (with noise at the boundary) and in [68, 70, 71, 163, 164] for the deterministic case. Second-order HJB equations and stochastic boundary control/noise problems have been investigated by mild solutions and Backward SDE in [131, 136, 165, 235, 338, 448, 464, 465], some of them also in connection with stochastic delay equations.

The material of Section 3.13 follows [244]. The definition of viscosity solution is similar to these of [242, 245] (see previous comments about origins of these definitions) however it uses different energy function (a radial function of the $|\cdot|_1$ norm) which reduces the equation to a subspace of the Hilbert space H. In this respect the definition is similar to the definitions in [271, 272]. A stationary equation similar to (3.443) was also investigated in [438]. Viscosity solution approaches to first-order HJB equations associated to optimal control of deterministic Navier-Stokes equations are in [243, 412] (see also [419] for earlier attempts). A PDE-viscosity solution approach to large deviations of stochastic two-dimensional Navier-Stokes equations with small noise intensities is considered in [425] where convergence of viscosity solutions of singularly perturbed HJB equations is studied. For results on Kolmogorov and HJB equations by other approaches [26, 25, 115, 117, 188, 328, 401, 443] we refer to Section 4.9.1, and the short discussion at the beginning of Section 3.13.

We did not discuss explicitly in the book Isaacs equations in Hilbert spaces which are associated to zero-sum two-player stochastic differential games. For Isaacs equations, one can prove existence of viscosity solutions by showing directly that the associated upper/lower value function of the game is a viscosity solution of the upper/lower Isaacs equation. Such results can be found in [192, 361, 363, 423]. This however is not easy since the proof of the dynamic programming principle is very complicated and thus only limited results are available. Related results on risk-sensitive stochastic control and differential games can be found in [359, 360, 362, 363, 424].

Other types of equations can be studied using the theory of viscosity solutions presented in this chapter, for instance obstacle problems for HJB equations related to optimal stopping problems, HJB equations for ergodic control problems. Comparison proofs easily extend to the case of obstacle problems. Explicit literature however is limited. Obstacle problems for bounded equations have been studied in [219, 319, 426]. HJB equations for ergodic control have not been investigated by viscosity solutions. In [226] they have been studied by the perturbation approach to mild solutions and in [206] by BSDE. Likewise HJB equations for singular control problems and for state constraints problems have not been studied in the infinite dimensional stochastic case in the viscosity solution framework. A singular stochastic control problem with delay is studied by other methods in [5] and [177]. Concerning infinite dimensional state constraints control problems and viscosity solutions of the associated HJB equations, the readers may check [172, 173] for the stochastic case and [67, 291, 175, 176, 178] for the deterministic case. A stochastic viability problem for a subset of a Hilbert space was studied by viscosity solutions in [60].

Also viscosity solution approach to HJB equations for optimal control of stochastic delay equations has not been fully explored. Some results on the subject are in [172, 173, 462, 463], see also [460, 461, 175, 176, 178] for the deterministic case and first order HJB equations. For other methods applicable to HJB equations for control of stochastic delay equations we refer the reader to Chapter 5 (in particular Sections 5.5 and 5.6) and Chapter 6 (in particular Section 6.5).

Another unexplored area is viscosity solutions of HJB equations with unbounded second-order terms which come from control problems with state equations driven by Q-Wiener processes with $\mathrm{Tr}(Q) = +\infty$, i.e. such that we may have $\mathrm{Tr}[(\sigma(t,x,a)Q^{\frac{1}{2}})(\sigma(t,x,a)Q^{\frac{1}{2}})^*] = +\infty$. So far [242] has been the only paper on the subject in a specific case. Up to now viscosity solution theory handles well fully nonlinear but "degenerate" equations while the theory of mild solutions handles well semilinear but "nondegenerate" equations. One would expect that the theory of viscosity solutions can be extended to the fully nonlinear "nondegenerate" HJB equations (see also the comments about [98] in the paragraph below discussing path dependent PDE).

There are also very few explicit results on viscosity solutions of boundary value problems in Hilbert spaces. Only some Dirichlet boundary value problems were studied. There exist comparison theorems (see Section 3.5), however quations studied by viscosity solutions are "degenerate" and hence construction of barriers at the boundary is not easy. Thus value functions may not be continuous up to the boundary unless some conditions are imposed on the drift. A Dirichlet boundary value problem for a bounded linear equation was investigated in [286] and a boundary value problem for a bounded HJB equation related to a risk-sensitive control problem was studied in [424]. Some results about value functions in bounded sets are sketched in [319]. A related paper for first order HJB equations is [67]. Results using approaches of mild and L^2 solutions are limited to linear equations. We refer the reader to [28, 29, 120, 121, 122, 123, 129, 390, 391, 430] and the references there for more. In [28, 29, 123] Neumann boundary value problems are considered.

An interesting direction in the evolution of the notion of viscosity solution in infinite dimensional spaces may come from the concept of path dependent PDE. Path dependent PDE come from the study of problems driven by path dependent SDE. In the finite dimensional spaces the notion of path dependent viscosity solution was introduced in [148] and this notion was extended to infinite dimensional spaces in [98]. In the Markovian case this approach gives an alternative way to treat HJB equations studied in this book. Its advantage is that it avoids the use of the maximum principle and thus it can be applied to "nondegenerate" equations, the continuity assumptions in the $|\cdot|_{-1}$ norm can be dropped, and the operator A does not need to be maximal dissipative. It is not clear if this method can be applied to fully nonlinear equations.

An emerging area of development for second-order equations in infinite dimensional spaces seems to be related to PDE in spaces of probability measures, in particular in the Wasserstein space. A second-order HJB equation in the space of probability measures was studied in [261] in connection with partially observed control and regularity of solutions was proved. Similar results were obtained for first order equations in [260]. Following the program described in [184], first-order HJB equations can be used to study large deviations for empirical measures of stochastic particle systems. Results in this direction are in [183, 184, 185, 186] (see also the references therein). It would be natural to investigate second-order equations related to the Laplace limit for such problems. Equations in the space probability measures also appear in the form of the so-called Master Equations of Mean Field Games [37, 36, 59, 72, 73, 74, 86, 314, 216]. These are non-local equations, which in the case of second-order Mean Field Games are of second-order. So far only limited results about existence and in some cases uniqueness of classical and strong solutions of first and second-order Master Equations was shown in [59, 74, 86, 216]. Also an approach proposed by P. L. Lions [72, 314] allows to convert an equation in the Wasserstein space to an equation in the Hilbert space L^2, where measures with finite second moments become random variables in L^2 with given laws.

Another emerging direction is HJB integro-PDE in Hilbert spaces which are associated to optimal control problems with state equations driven by Lévy processes or random measures. Viscosity solutions have been introduced for such non-local equations in [428, 429, 427]. Comparison theorems have been proved in [429] and existence of viscosity solutions and optimal control problems have been studied in [427]. Some linear non-local PDE and properties of transition semigroups for processes with jumps have been studied by other methods in [8, 311, 378, 392, 393, 394].

APPENDIX A

Notations and Function Spaces

In this appendix, we list the main notation and the definitions of the basic functions spaces used throughout the book. Other notation and definitions of functions spaces that are used only in specific chapters are not recalled here and are introduced directly in their chapters.

A.1. Basic Notation

If X is a Banach space we denote its norm by $|\cdot|_X$. If this space is also Hilbert, we denote its inner product by $\langle\cdot,\cdot\rangle_X$. Given $R > 0$, $B_X(\bar{x}, R)$ denotes the closed ball in X centered at $\bar{x}$ of radius R. We will omit the subscript X if the context is clear. The dual space of X, i.e. the space of all continuous linear functionals, will be denoted by X^*. The (operator) norm in X^* will be denoted by $|\cdot|_{X^*}$, and the duality will be denoted by $\langle\cdot,\cdot\rangle_{\langle X^*,X\rangle}$.

If a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ converges to $x \in X$ in the norm topology we write $x_n \to x$. If it converges weakly we write $x_n \rightharpoonup x$.

If X is a Hilbert space and $\{e_k\}_{k\in\mathbb{N}}$ is an orthonormal basis of X we use, for $x \in X$, the notation $x_k := \langle x, e_k\rangle$. Unless stated explicitly, we will always identify its dual X^* with X through the standard Riesz identification.

Given a second Banach space Y with norm $|\cdot|_Y$ (and inner product $\langle\cdot,\cdot\rangle_Y$ if it is also Hilbert) we denote by $\mathcal{L}(X;Y)$ the set of all bounded (continuous) linear operators $T : X \to Y$ with norm $\|T\|_{\mathcal{L}(X;Y)} := \sup_{x\in X, x\neq 0} \frac{|Tx|_Y}{|x|_X}$ (or simply $\|T\|$), using for simplicity the notation $\mathcal{L}(X)$ when $X = Y$. $\mathcal{L}(X)$ is a Banach algebra with identity element I_X (simply I if unambiguous).

Given a linear (possibly unbounded) operator $T : D(T) \subset X \to Y$ such that $D(T)$ is dense in X we will denote its adjoint operator by $T^* : D(T^*) \subset Y^* \to X^*$. If X is a Hilbert space we will denote by $\mathcal{S}(X) \subset \mathcal{L}(X)$ the space of all bounded self-adjoint operators on X.

For $k = 1, 2, ...$ we denote by X^k the product space $X \times X \times \cdots \times X$ (k times) endowed with the norm $|(x_1, ..., x_k)|_{X^k} := \left(|x_1|^2 + ... + |x_k|^2\right)^{1/2}$ and by $\mathcal{L}^k(X;Y)$ the set of all bounded multilinear operators $T : X^k \to Y$ with norm $\|T\|_{\mathcal{L}^k(X;Y)} := \sup_{x\in X^k, x\neq 0} \frac{|T(x_1,...,x_n)|_Y}{|(x_1,...,x_n)|_{X^k}}$ using for simplicity the notation $\mathcal{L}^k(X)$ when $X = Y$. It is known (see e.g. [198] p. 318, Theorem A.2.6) that $\mathcal{L}^k(X;Y)$ is isometrically isomorphic to the space $\mathcal{L}(X, \mathcal{L}(X, \ldots, \mathcal{L}(X;Y)))$.

Given a complex number $\lambda \in \mathbb{C}$ we denote by $\mathrm{Re}\lambda$ and $\mathrm{Im}\lambda$ respectively its real and imaginary parts.

A.2. Function spaces

Let Y and Z be two Banach spaces and let $X \subset Z$ be endowed with the induced topology. We denote by $B(X;Y)$, $B_b(X;Y)$, $C(X;Y)$, $UC(X;Y)$, $C_b(X;Y)$ and $UC_b(X;Y)$ the sets of all functions $\varphi : X \to Y$ which are, respectively, Borel measurable, Borel measurable and bounded, continuous, uniformly continuous, continuous and bounded, uniformly continuous and bounded on X. The spaces $B_b(X;Y)$,

$C_b(X;Y)$ and $UC_b(X;Y)$ are Banach spaces with the usual norm

$$\|\varphi\|_0 = \sup_{x\in X} |\varphi(x)|_Y.$$

We denote by $USC(X;Y)$ (respectively, $LSC(X;Y)$) the space of all upper semicontinuous (respectively, lower semicontinuous) functions $f : X \to Y$.

If $Y = \mathbb{R}$, we will simply write $B_b(X)$, $C(X)$, $UC(X)$, $C_b(X)$, $UC_b(X)$, $USC(X)$ and $LSC(X)$ for $B_b(X;\mathbb{R})$, $C(X;\mathbb{R})$, $UC(X;\mathbb{R})$, $C_b(X;\mathbb{R})$, $UC_b(X;\mathbb{R})$, $USC(X;\mathbb{R})$ and $LSC(X;\mathbb{R})$.

For a given $m > 0$ we define $B_m(X,Y)$ (respectively, $C_m(X,Y)$ and $UC_m(X,Y)$) to be the set of all functions $\phi \in B(X,Y)$ such that the function

$$\psi(x) := \frac{\phi(x)}{(1+|x|^2)^{m/2}} \tag{A.1}$$

belongs to $B_b(X,Y)$ (respectively $C_b(X,Y)$ and $UC_b(X,Y)$). These spaces of functions that have at most polynomial growth of order m are Banach spaces when they are endowed with the norm

$$N(\phi) := \sup_{x\in X} \frac{|\phi(x)|}{(1+|x|^2)^{m/2}}.$$

We will write $\|\phi\|_{B_m(X,Y)}$, $\|\phi\|_{C_m(X,Y)}$, $\|\phi\|_{UC_m(X,Y)}$ to denote these norms, or simply $\|\phi\|_{B_m}$, $\|\phi\|_{C_m}$, $\|\phi\|_{UC_m}$ when the spaces are clear from the context. The above definition is meaningful also when $m = 0$ and in such case the spaces $B_m(X,Y)$, $C_m(X,Y)$ and $UC_m(X,Y)$ reduce to $B_b(X,Y)$, $C_b(X,Y)$ and $UC_b(X,Y)$. We do not use the notation $B_0(X,Y)$, $C_0(X,Y)$ and $UC_0(X,Y)$ to avoid confusion with respect to the standard notation in the literature. However we often consider the spaces $B_m(X,Y)$, $C_m(X,Y)$ and $UC_m(X,Y)$ for $m \geq 0$ meaning that also the case of the spaces $B_b(X,Y)$, $C_b(X,Y)$ and $UC_b(X,Y)$ is included.

Let Y be a Banach space and X be an open subset of a Banach space Z. For $k \in \mathbb{N}$, we denote by $B^k(X;Y)$ (respectively $B^k_b(X;Y)$) the set of all functions $\varphi : X \to Y$ which are Borel measurable (respectively Borel measurable and bounded) on X, together with all their Fréchet derivatives (see Section D.2) up to the order k. If $\varphi \in B^k_b(X;Y)$ the l-th Fréchet derivative of φ is denoted by $D^l\varphi$ (or simply $D\varphi$ when $l = 1$). We set

$$\|\varphi\|_k = \|\varphi\|_0 + \sum_{l=1}^{k} \sup_{x\in X} \left\|D^l\varphi(x)\right\|_{\mathcal{L}^l(Z;Y)}.$$

Similarly we define the space $C^k(X;Y)$ (respectively, $C^k_b(X;Y)$, $UC^k(X;Y)$, $UC^k_b(X;Y)$) to be the set of all functions $\varphi : X \to Y$ which are continuous (respectively continuous and bounded, uniformly continuous, uniformly continuous and bounded) on X together with all their Fréchet derivatives up to the order k. z

For $k \in \mathbb{N}$ and $m > 0$ we denote by $B^k_m(X;Y)$ (respectively, $C^k_m(X,Y)$ and $UC^k_m(X,Y)$) the set of all functions $\phi \in B(X,Y)$ such that the corresponding function ψ defined in (A.1) belongs to $B^k_b(X,Y)$ (respectively $C^k_b(X,Y)$ and $UC^k_b(X,Y)$). We set

$$\|\phi\|_{B^k_m(X;Y)} = \|\psi\|_{B^k_b(X;Y)}$$

where ψ is the function defined in (A.1) corresponding to ϕ. With this norm $B^k_b(X,Y)$, $C^k_b(X,Y)$, $UC^k_b(X,Y)$ are all Banach spaces. For the last two we will denote the norms by $\|\phi\|_{C^k_m(X,Y)}$, $\|\phi\|_{UC^k_m(X,Y)}$.

If $Y = \mathbb{R}$, we write $B^k_b(X)$, $B^k_m(X)$, $C^k(X)$, $C^k_b(X)$, $C^k_m(X)$, $UC^k(X)$, $UC^k_b(X)$, $UC^k_m(X)$ instead of $B^k_b(X;\mathbb{R})$, $B^k_m(X;\mathbb{R})$, $C^k(X;\mathbb{R})$, $C^k_b(X;\mathbb{R})$, $C^k_m(X;\mathbb{R})$, $UC^k(X;\mathbb{R})$, $UC^k_b(X;\mathbb{R})$, $UC^k_m(X;\mathbb{R})$, respectively.

For $\alpha \in (0,1]$ we denote by $C^{0,\alpha}(X;Y)$ the space of all Hölder continuous functions from X to Y endowed with the semi-norm

$$[\varphi]_{0,\alpha} = \sup\left\{\frac{|\varphi(x)-\varphi(y)|_Y}{|x-y|_Z^\alpha};\ x,y\in X;\ x\neq y\right\}.$$

The space $C_b^{0,\alpha}(X;Y) := C_b(X;Y)\cap C^{0,\alpha}(X;Y)$ is a Banach space with the norm

$$\|\varphi\|_{0,\alpha} = \|\varphi\|_0 + [\varphi]_{0,\alpha}$$

and is contained in $UC_b(X;Y)$. If $\alpha = 1$ the space $C^{0,1}(X;Y)$ is the space of all Lipschitz continuous functions from X to Y. If X is open, convex and $\varphi \in C^{0,1}(X;Y)$ is Fréchet differentiable in X then the derivative $D\varphi$ is bounded and

$$[\varphi]_{0,1} = \sup_{x\in X}\|D\varphi(x)\|_{\mathcal{L}(Z;Y)}.$$

We set, for $\alpha \in (0,1]$,

$$C^{1,\alpha}(X;Y) := \left\{\varphi \in C^1(X;Y)\ :\ [D\varphi]_{0,\alpha} < \infty\right\}. \tag{A.2}$$

The space $C_b^{1,\alpha}(X;Y) := C_b^1(X;Y)\cap C^{1,\alpha}(X;Y)$ is a Banach space with the norm

$$\|\varphi\|_{1,\alpha} = \|\varphi\|_1 + [D\varphi]_{0,\alpha}.$$

Similarly,

$$C^{2,\alpha}(X;Y) := \left\{\varphi \in C^2(X;Y)\ :\ [D^2\varphi]_{0,\alpha} < \infty\right\}. \tag{A.3}$$

The space $C_b^{2,\alpha}(X;Y) := C_b^2(X;Y)\cap C^{2,\alpha}(X;Y)$ is also a Banach space with the norm

$$\|\varphi\|_{2,\alpha} = \|\varphi\|_2 + [D^2\varphi]_{0,\alpha}.$$

If $Y = \mathbb{R}$ the above spaces are denoted by $C^{0,\alpha}(X)$, $C_b^{0,\alpha}(X)$, $C^{1,\alpha}(X)$, $C_b^{1,\alpha}(X)$, $C^{2,\alpha}(X)$, $C_b^{2,\alpha}(X)$.

If $\alpha \in (0,1]$, we say that $\varphi : X \to Y$ is Hölder (Lipschitz when $\alpha = 1$) continuous on bounded subsets of X, if $\varphi \in C^{0,\alpha}(B(0,R)\cap X;Y)$ for every $R > 0$ i.e. if the semi-norm

$$[\varphi]_{0,\alpha,R} := \sup\left\{\frac{|\varphi(x)-\varphi(y)|_Y}{|x-y|_Z^\alpha};\ x,y\in B(0,R)\cap X;\ x\neq y\right\}$$

is finite for every $R > 0$. The space of such functions is denoted by $C_{\mathrm{loc}}^{0,\alpha}(X;Y)$. Similarly we define the spaces $C_{\mathrm{loc}}^{1,\alpha}(X;Y)$ and $C_{\mathrm{loc}}^{2,\alpha}(X;Y)$.

If $Y = \mathbb{R}$ we write $C^{k,\alpha}(X)$ instead of $C^{k,\alpha}(X;\mathbb{R})$ and $C_{\mathrm{loc}}^{k,\alpha}(X)$ instead of $C_{\mathrm{loc}}^{k,\alpha}(X;\mathbb{R})$, for $k = 0,1,2$, $\alpha \in (0,1]$.

Let now X be a real separable Hilbert space and Y be a real Banach space. As usual we set, for an open subset $\mathcal{O}$ of $\mathbb{R}^n$, and $k \in \mathbb{N}\cup\{\infty\}$,

$$C_0^k(\mathcal{O};Y) := \left\{f \in C^k(\mathcal{O};Y) : f \quad \text{has compact support in } \mathcal{O}\ \right\}.$$

Following [91] Section 0 and [225] Section 2 we define various spaces of cylindrical functions. We set, for $k \in \mathbb{N}\cup\{\infty\}$,

$$\begin{aligned}\mathcal{F}C_0^k(X;Y) := \big\{\varphi : X \to Y : \exists n\in\mathbb{N},\ x_1,...,x_n \in X,\ f\in C_0^k(\mathbb{R}^n;Y)\\ \text{such that } \varphi(x) = f(\langle x,x_1\rangle,...,\langle x,x_n\rangle),\ \forall x\in X\big\}.\end{aligned} \tag{A.4}$$

We denote $\mathcal{F}C_0^k(X;\mathbb{R})$ simply by $\mathcal{F}C_0^k(X)$.

Given $k \in \mathbb{N}\cup\{\infty\}$ and a linear closed operator with dense domain $B : D(B) \subseteq X \to X$, we define

$$\begin{aligned}\mathcal{F}C_0^{k,B}(X) = \big\{\varphi : X \to \mathbb{R} : \exists n\in\mathbb{N},\ x_1,...,x_n \in D(B),\ f\in C_0^k(\mathbb{R}^n;\mathbb{R})\\ \text{such that } \varphi(x) = f(\langle x,x_1\rangle,...,\langle x,x_n\rangle),\ \forall x\in X\big\}\end{aligned} \tag{A.5}$$

and

$$\mathcal{F}C_b^{k,B}(X) := \big\{\varphi : X \to \mathbb{R} : \exists n \in \mathbb{N},\ x_1, ..., x_n \in D(B),\ f \in C_b^k(\mathbb{R}^n;\mathbb{R}) \\ \text{such that } \varphi(x) = f(\langle x, x_1\rangle, ..., \langle x, x_n\rangle),\ \forall x \in H\big\}. \quad \text{(A.6)}$$

Moreover if $\mathcal{E} = \{e_k\}_{k\in\mathbb{N}}$ is an orthonormal basis of X, we introduce, for $k \in \mathbb{N} \cup \{\infty\}$,

$$\mathcal{E}C_0^k(X;Y) := \big\{\varphi : X \to Y : \exists n \in \mathbb{N},\ f \in C_0^k(\mathbb{R}^n;Y) \text{ such that} \\ \varphi(x) = f(\langle x, e_1\rangle, ..., \langle x, e_n\rangle),\ \forall x \in X\big\} \quad \text{(A.7)}$$

and, if $B : D(B) \subseteq X \to X$ is a linear closed operator with dense domain and $\mathcal{E} \subseteq D(B)$,

$$\mathcal{E}C_0^{k,B}(X) := \big\{\varphi : X \to \mathbb{R} : \exists n \in \mathbb{N},\ f \in C_0^k(\mathbb{R}^n) \text{ such that} \\ \varphi(x) = f(\langle x, e_1\rangle, ..., \langle x, e_n\rangle),\ \forall x \in X\big\}. \quad \text{(A.8)}$$

Let $\mathcal{O}$ be an open subset of $\mathbb{R}^n$ and $p \in [1, +\infty)$. We denote by $L^p(\mathcal{O})$, the set of all real valued measurable functions[1] $f : \mathcal{O} \to \mathbb{R}$ with $\int_{\mathcal{O}} |f(\xi)|^p d\xi < +\infty$ (classical Lebesgue integral); $L^p(\mathcal{O})$ is a Banach space with the usual norm $|f|_{L^p(\mathcal{O})} := \left[\int_{\mathcal{O}} |f(\xi)|^p d\xi\right]^{1/p}$. We denote by $L^p_{\mathrm{loc}}(\mathcal{O})$ the set of all measurable functions $f : \mathcal{O} \to \mathbb{R}$ such that $\int_K |f(\xi)|^p d\xi < +\infty$ for every compact subset K of $\mathcal{O}$. The space $L^\infty(\mathcal{O})$ is the quotient space of $B_b(\mathcal{O})$ with respect to the relation of being equal a.e. and is a Banach space with the usual ess sup norm. (Obviously the above spaces can be defined for more general sets $\mathcal{O}$.)

We denote by $W^{k,p}(\mathcal{O})$ ($k \in \mathbb{N}$, $p \in [1, +\infty]$) the usual Sobolev space of real valued functions whose distributional derivatives, up to the order k, are p-th power integrable (or essentially bounded if $p = +\infty$). Moreover $W_0^{k,p}(\mathcal{O})$ is the closure of $C_0^\infty(\mathcal{O})$ in $W^{k,p}(\mathcal{O})$. Following the standard convention, sometimes we will write $H^k(\mathcal{O})$ for $W^{k,2}(\mathcal{O})$ and $H_0^k(\mathcal{O})$ for $W_0^{k,2}(\mathcal{O})$. Similarly, for $\alpha \geq 0$ and $p \in [1, +\infty]$, we denote the fractional Sobolev spaces by $W^{\alpha,p}(\mathcal{O})$, $W_0^{\alpha,p}(\mathcal{O})$ (or $H^\alpha(\mathcal{O})$, $H_0^\alpha(\mathcal{O})$ when $p = 2$) defined in the usual way (see e.g. [1] Chapter VII). By duality then one defines, for every $\alpha > 0$, the negative order spaces $H^{-\alpha}(\mathcal{O})$ setting $(H_0^\alpha(\mathcal{O}))^* = H^{-\alpha}(\mathcal{O})$ (see e.g. [313] Section 1.12).

If the boundary $\partial\mathcal{O}$ is a C^∞ manifold of dimension $n-1$ in $\mathbb{R}^n$ then also the spaces $L^p(\partial\mathcal{O})$, $H^\alpha(\partial\mathcal{O})$, $W^{\alpha,p}(\partial\mathcal{O})$ can be defined. We refer to [313], Section 1.7 or [1], Chapter VII, p. 215.

If Y is a real, separable Banach space and $a < b$, we define the space $W^{1,p}(a, b; Y)$ ($p \in [1, \infty]$) to be the set of all functions $f \in L^p(a, b; Y)$ whose weak derivative f' (see [435] Chapter III, §1) exists and belongs to $L^p(a, b; Y)$. It is a Banach space equipped with the norm $|f|_{W^{1,p}(a,b;Y)} := |f|_{L^p(a,b;Y)} + |f'|_{L^p(a,b;Y)}$.

Let X be a subset of a Banach space Z, Y be a Banach space and I be an interval in $\mathbb{R}$. We define the space

$$UC_b^x(I \times X; Y) := \Big\{\varphi \in C_b(I \times X; Y) : \\ \varphi(t, x) \text{ is uniformly continuous in } x, \text{ uniformly with respect to } t \in I\Big\}. \quad \text{(A.9)}$$

If $Y = \mathbb{R}$, we write $UC_b^x(I \times X)$ instead of $UC_b^x(I \times X; \mathbb{R})$.

[1]That is, the set of all equivalence classes of such functions with respect to the relation of a.e. equality.

REMARK A.1 We recall some useful properties of $UC_b^x(I \times X; Y)$ (see [82] for more).

(1) If $I \subseteq \mathbb{R}$ is compact and $u \in UC_b^x(I \times X; Y)$, then for every compact set $K \subseteq X$, the restriction $u|_{I \times K}$ of u to $I \times K$ belongs to $UC_b(I \times K; Y)$. Thus, for every compact set $K \subseteq X$,
$$u|_{I\times K} \in C_b(I; C_b(K; Y)). \tag{A.10}$$
In particular $u(\cdot, x)$ is uniformly continuous on I, uniformly with respect to $x \in K$, namely, for every $t, s \in I$, we have
$$\sup_{x\in K} |u(t,x) - u(s,x)|_Y \le \rho_K(|t-s|), \tag{A.11}$$
where ρ_K is a modulus of continuity (see Appendix D.1) depending on the compact set K.

(2) $UC_b(I \times X; Y) \subseteq UC_b^x(I \times X; Y) \subseteq C_b(I \times X; Y)$. In particular, as the uniform continuity is stable with respect to the convergence in the norm $\|\cdot\|_0$, the space $UC_b^x(I \times X; Y)$ is a closed subspace of $C_b(I \times X; Y)$. On the other hand, if X is the whole space, $\varphi \in UC_b(X)$ and $\{e^{tA},\ t \ge 0\}$ is a strongly continuous (not uniformly continuous) semigroup on X, then
$$w(t,x) = \varphi(e^{tA}x)$$
is a natural example of a function belonging to $UC_b^x(I \times X; Y)$ but not to $UC_b(I \times X; Y)$.

■

For a given $m > 0$ we define $B_m(I \times X, Y)$ (respectively, $C_m(I \times X, Y)$ and $UC_m(I \times X, Y)$) to be the set of all functions $\phi \in B(I \times X, Y)$ such that the function
$$\psi(t,x) := \frac{\phi(t,x)}{(1+|x|^2)^{m/2}} \tag{A.12}$$
belongs to $B_b(I \times X, Y)$ (respectively $C_b(I \times X, Y)$ and $UC_b(I \times X, Y)$). These spaces of functions that have at most polynomial growth of order m in the variable x are Banach spaces when they are endowed with the norm
$$N(\phi) := \sup_{(t,x)\in I\times X} \frac{|\phi(t,x)|}{(1+|x|^2)^{m/2}}.$$
We will write $\|\phi\|_{B_m(I\times X,Y)}$, $\|\phi\|_{C_m(I\times X,Y)}$, $\|\phi\|_{UC_m(I\times X,Y)}$ to denote these norms.

Following [76], we introduce the space
$$UC_m^x(I \times X; Y) := \Big\{\varphi \in C(I \times X; Y) \ : \\ \psi(t,x) := \frac{\phi(t,x)}{(1+|x|^2)^{m/2}} \in UC_b^x(I \times X; Y)\Big\}. \tag{A.13}$$
It has properties similar to these described in Remark A.1 for $u \in UC_b^x(I \times X; Y)$.

Let X be an open subset of a Banach space Z, and Y be a Banach space. Let $I \subset \mathbb{R}$ be open. Given a function $u \in C(I \times X; Y)$ which is l times Fréchet differentiable in t and k times Fréchet differentiable in x, we denote its l-th partial derivative in t by $D_t^l u$, and its k-th partial derivative in x by $D_x^k u$. For the low order derivatives we use the symbols u_t, Du, D^2u, and so on.

We denote by $C^{l,k}(I \times X; Y)$ (respectively, by $C_b^{l,k}(I \times X; Y)$, $UC^{l,k}(I \times X; Y)$, $UC_b^{l,k}(I \times X; Y)$) the space of all functions $\varphi : I \times X \to Y$ that are Fréchet differentiable l times in t and k times in x and which are continuous (respectively continuous and bounded, uniformly continuous, uniformly continuous and bounded), together with their Fréchet derivatives up to these orders. If I is an interval which is not open then $C^{l,k}(I \times X; Y)$ is the space of functions from $I \times X \to Y$ whose restrictions to $I^o \times X$ (where I^o is the interior of I) belong to $C^{l,k}(I^o \times X; Y)$ and such that the functions and all their derivatives extend continuously to $I \times X$. The spaces $C_b^{l,k}(I \times X; Y)$, $UC^{l,k}(I \times X; Y)$, $UC^{l,k}(I \times X; Y)$ and $UC_b^{l,k}(I \times X; Y)$ are defined similarly.

$C_b^{l,k}(I \times X; Y)$ and $UC_b^{l,k}(I \times X; Y)$ are Banach spaces endowed with the norm

$$|||\varphi|||_{l,k} = \sup_{x \in X} \|\varphi(\cdot, x)\|_l + \sup_{t \in I} \|\varphi(t, \cdot)\|_k$$

We define $C_m^{l,k}(I \times X; Y)$ (respectively $UC_m^{l,k}(I \times X; Y)$) to be the set of all functions $\phi : I \times X \to Y$ such that the function ψ in (A.12) belongs to $C_b^{l,k}(I \times X; Y)$ (respectively to $UC_b^{l,k}(I \times X; Y)$). Such spaces are Banach spaces endowed with the norm

$$\|\phi\|_{C_m^{l,k}(I \times X; Y)} := \|\psi\|_{C_b^{l,k}(I \times X; Y)}$$

where ψ is the function given in (A.12) corresponding to ϕ.

If $Y = \mathbb{R}$ we will use the notation $C^{l,k}(I \times X), C_b^{l,k}(I \times X), UC^{l,k}(I \times X)$ and $UC_b^{l,k}(I \times X)$ for the above spaces.

Let I be an interval in $\mathbb{R}$, X be a real separable Hilbert space and Y be a real Banach space. Similarly to the time independent case we set, for $k \in \mathbb{N} \cup \{\infty\}$,

$$C_0^k(I \times \mathbb{R}^n; Y) := \left\{ f \in C^k(I \times \mathbb{R}^n; Y) : f \quad \text{has compact support in } I \times \mathbb{R}^n . \right\}$$

If $k \in \mathbb{N} \cup \{\infty\}$, we denote by $\mathcal{F}C_0^k(I \times X; Y)$ the space

$$\mathcal{F}C_0^k(I \times X; Y) := \big\{ \varphi : I \times X \to Y : \exists n \in \mathbb{N},\ x_1, ..x_n \in X,\ f \in C_0^k(I \times \mathbb{R}^n; Y), \\ \text{such that } \varphi(x) = f(t, \langle x, x_1 \rangle, ..., \langle x, x_n \rangle),\ \forall (t,x) \in I \times X \big\} \tag{A.14}$$

Similarly, for $k \in \mathbb{N} \cup \{\infty\}$ and a linear, densely defined closed operator $B : D(B) \subseteq X \to X$, we define

$$\mathcal{F}C_0^{k,B}(I \times X) = \big\{ \varphi : I \times X \to \mathbb{R} : \exists n \in \mathbb{N},\ x_1, ..x_n \in D(B),\ f \in C_0^k(I \times \mathbb{R}^n; \mathbb{R}) \\ \text{such that } \varphi(x) = f(\langle x, x_1 \rangle, ..., \langle x, x_n \rangle),\ \forall x \in X \big\} \tag{A.15}$$

Finally, given an orthonormal basis $\mathcal{E} = \{e_k\}_{n \in \mathbb{N}}$ of X, we define, for $k \in \mathbb{N} \cup \{\infty\}$,

$$\mathcal{E}C_0^k(I \times X; Y) := \big\{ \varphi : I \times X \to Y : \exists n \in \mathbb{N},\ f \in C_0^k(I \times \mathbb{R}^n; Y), \text{ such that} \\ \varphi(x) = f(t, \langle x, e_1 \rangle, ..., \langle x, e_n \rangle),\ \forall (t,x) \in I \times X \big\}. \tag{A.16}$$

APPENDIX B

Linear operators and C_0-semigroups

All spaces considered in the book are real. However the spectral theory has to be done in complex spaces and thus some results presented here require the use of complex spaces. To accommodate real Hilbert and Banach spaces we thus use complexification of spaces and operators, which for Hilbert spaces can be done in a natural way. If H is a real Hilbert space, its complexification H_c is defined by

$$H_c := \{\tilde{x} = x + iy : x, y \in H\}$$

with standard operations

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

and with the inner product

$$\langle (x+iy), (z+iw) \rangle_c := \langle x, z \rangle_H + \langle y, w \rangle_H + i(\langle y, z \rangle_H - \langle x, w \rangle_H).$$

Thus $|x + iy|_{H_c} = |(x, y)|_{H \times H}$. A real Banach space E is complexified in the same way, however, except for special cases, the product norm is no longer a norm because the homogeneity condition fails. To define a norm in E_c we first compute

$$e^{it}(x + iy) = (x \cos t - y \sin t) + i(x \sin t + y \cos t)$$

and then define

$$|x + iy|_{E_c} := \sup_{0 \le t \le 2\pi} |(x \cos t - y \sin t, x \sin t + y \cos t)|_{E \times E}.$$

We refer the reader to [351, 413] for more on complexification.

A linear operator $T : D(T) \subset E \to E$ is complexified by setting

$$D(T_c) := \{x + iy : x, y \in D(T)\}, \quad T_c(x + iy) := Tx + iTy.$$

It is easy to see that $T \in \mathcal{L}(E)$ if and only if $T_c \in \mathcal{L}(E_c)$, and moreover $\|T\| = \|T_c\|$. Also T is invertible if and only if T_c is invertible. It is a standard convention which will not be repeated, that the spectrum and the resolvent set of T are understood to be the spectrum and the resolvent set of T_c. This is how the statements here should be understood in the context of real Hilbert and Banach spaces.

Throughout Appendix B, E will be a Banach space endowed with the norm $|\cdot|_E$ and H will be a Hilbert space endowed with the inner product $\langle \cdot, \cdot \rangle_H$ and the norm $|\cdot|_H$.

B.1. Linear operators

For an operator T we denote by $D(T)$ its domain, by $R(T)$ its range and by $\ker T$ its kernel (or null space).

Definition B.1 (Pseudoinverse) *If E is a uniformly convex Banach space, Z is a Banach space, and $T \in \mathcal{L}(E; Z)$, the* pseudoinverse T^{-1} *of T is the linear operator defined on $T(E) \subseteq Z$ that associates to every element z in $T(E)$ the element in $T^{-1}(z)$ with minimum norm (for existence of such an element see* [144] *II.4.29 page 74). Notice that If E is a Hilbert space then we have*

$$R(T^{-1}) = (\ker T)^{\perp}.$$

The following result is taken from [130], Proposition B.1, p. 429, where the reader can find its proof.

PROPOSITION B.2 *Let E, E_1, E_2 be three Hilbert spaces, let $A_1 : E_1 \to E$, $A_2 : E_2 \to E$ be linear bounded operators, let $A_1^* : E \to E_1$ and $A_2^* : E \to E_2$ be their adjoints and finally let $A_1^{-1} : R(A_1) \subseteq E \to E_1$, $A_2^{-1} : R(A_2) \subseteq E \to E_2$ be the respective pseudoinverses. Then we have:*

(i) *$R(A_1) \subseteq R(A_2)$ if and only if there exists a constant $k > 0$ such that*

$$|A_1^\star x|_{E_1} \le k|A_2^\star x|_{E_2} \qquad \forall x \in E.$$

(ii) *If*

$$|A_1^\star x|_{E_1} = |A_2^\star x|_{E_2} \qquad \forall x \in E,$$

then $R(A_1)$=$R(A_2)$ and

$$|A_1^{-1}x|_{E_1} = |A_2^{-2}x|_{E_2} \qquad \forall x \in R(A_1).$$

DEFINITION B.3 (Closed operator, Graph norm) *Let E and Z be two Banach spaces. A linear operator $A : D(A) \subset E \to Z$, is said to be* closed *if its graph*

$$\{(x, y) \in D(A) \times Z \ : \ y = Ax\}$$

is closed in $E \times Z$. Given a closed operator $A : D(A) \subset E \to Z$ the graph norm *on $D(A)$ is defined as follows:*

$$|x|_{D(A)} := \left(|x|_E^2 + |Ax|_Z^2\right)^{\frac{1}{2}} \qquad \textit{for all } x \in D(A).$$

Sometimes an equivalent norm $|x|_{D(A)} := |x|_E + |Ax|_Z$ is also used.

PROPOSITION B.4 *Let E and Z be two Banach spaces. If $A : D(A) \subset E \to Z$ is a linear, closed operator, then $D(A)$ with the graph norm is a Banach space. Moreover, if $D(A)$ is endowed with the graph norm, $A : D(A) \to Z$ is continuous.*

DEFINITION B.5 (Resolvent) *Consider a linear, closed operator $A : D(A) \subset E \to E$ and define, for $\lambda \in \mathbb{C}$, the operator $(\lambda I - A) : D(A) \to E$. The* resolvent set *of A is defined as follows:*

$$\varrho(A) := \left\{\lambda \in \mathbb{C} \ : \ (\lambda I - A) \text{ is invertible and } (\lambda I - A)^{-1} \in \mathcal{L}(E)\right\}.$$

$$\sigma(A) := \mathbb{C} \setminus \varrho(A)$$

is called the spectrum *of A.*

LEMMA B.6 *Let $A : D(A) \subset E \to E$ be a linear, closed operator. Then the resolvent set $\varrho(A)$ is open.*

PROOF. See [450], Theorem 1 page 201. □

DEFINITION B.7 (Closable operator, closure of an operator) *Let E and Z be two Banach spaces. A linear operator $A : D(A) \subset E \to Z$ is said to be* closable *if the closure of its graph in $E \times Z$ is the graph of some (closed) operator. Such an operator is called the* closure *of A and it is denoted by $\overline{A}$.*

REMARK B.8 It is easy to see that $A : D(A) \subset E \to Z$ is closable if and only if, for any sequence x_n in $D(A)$ and any $y \in Z$ such that

$$x_n \xrightarrow[E]{n\to\infty} 0 \qquad \text{and} \qquad T(x_n) \xrightarrow[Z]{n\to\infty} y,$$

we have $y = 0$. ■

Definition B.9 (Core of a closed operator) *Let E and Z be two Banach spaces. Consider a closed linear operator $A : D(A) \subset E \to Z$. A linear subspace Y of $D(A)$ is said to be a* core *for A if it is dense in $D(A)$ (endowed with its graph norm). In other words Y is a core for A if and only if $Y \subseteq D(A)$ and, for any $x \in D(A)$ there exists a sequence x_n of elements of Y such that $x_n \to x$ and $Ax_n \to Ax$.*

B.2. Dissipative operators

Definition B.10 (Duality mapping) *Let E be a Banach space and E^* be its dual. The function $\mathcal{J} \colon E \to 2^{E^*}$, defined by*

$$\mathcal{J}(x) := \left\{ x^* \in E \ : \ \langle x^*, x \rangle_{\langle E^*, E \rangle} = |x|_E^2 = |x^*|_{E^*}^2 \right\}, \qquad \textit{for } x \in E,$$

is called the duality mapping*.*

The duality mapping is in general multivalued and, for all $x \in E$, $J(x) \neq \emptyset$. If the dual E^* is strictly convex, and in particular if E is a Hilbert space, $\mathcal{J}$ is single-valued (see Section 1.1 of [20] and in particular Theorem 1.2).

Lemma B.11 (Kato's Lemma) *Let E be a Banach space and $x, y \in E$. There exists $w \in \mathcal{J}(x)$ such that $\langle w, y \rangle_{\langle E^*, E \rangle} \geq 0$ if and only if*

$$|x|_E \leq |x + \lambda y|_E$$

for any $\lambda > 0$.

Proof. See [20], Lemma 3.1, page 98. □

In the remainder of this section A will denote a possibly non-linear operator.

Definition B.12 (Dissipative operators) *Let E be a Banach space. An operator $A : D(A) \subset E \to E$ is called* dissipative *if*

$$\langle w, A(x) - A(y) \rangle_{E^*, E} \leq 0, \quad \textit{for all } x, y \in D(A) \textit{ and some } w \in \mathcal{J}(x - y). \tag{B.1}$$

*A dissipative operator A is said to be m-*dissipative *if $R(I - A) = E$. A dissipative operator A is said to be* maximal dissipative *if there does not exist any proper dissipative extension of A.*

In the specific case of a linear operator A in a Hilbert space H the expression (B.1) can be rephrased as $\langle Ax, x \rangle_H \leq 0$ for all $x \in D(A)$. If the Hilbert space is complex one considers its real part: $\mathrm{Re}\, \langle Ax, x \rangle_H \leq 0$.

Definition B.13 (Accretive operators) *Let E be a Banach space. An operator $A : D(A) \subset E \to E$ is called* accretive *(respectively m-accretive,* maximal accretive*) if $-A$ is dissipative (respectively m-dissipative, maximal dissipative).*

Remark B.14 It easily follows from the definition that any m-dissipative operator is maximal dissipative. In case of (possibly non-linear) operators in Hilbert spaces the two properties are equivalent, see Remark 3.1 page 101 of [20]. ■

Proposition B.15 *An operator $A : D(A) \subset E \to E$ is dissipative if and only if, for any $\lambda > 0$, for any $x, y \in D(A)$,*

$$|x - y|_E \leq |x - y - \lambda(A(x) - A(y))|_E \,,$$

or equivalently, if there exists $\lambda > 0$ with such a property.

Proof. See Proposition 3.1, page 98, of [20]. □

REMARK B.16 Proposition B.15 can be rewritten in an obvious way in the following form: $A : D(A) \subset E \to E$ is dissipative if and only if for any $\lambda > 0$, for any $x, y \in D(A)$,

$$|x - y|_E \le \frac{1}{\lambda} \left|(\lambda x - A(x)) - (\lambda y - A(y))\right|_E ,$$

or equivalently, if there exists $\lambda > 0$ with such a property. ■

PROPOSITION B.17 *A dissipative operator $A : D(A) \subset E \to E$ is m-dissipative if and only if $R(I - \lambda A) = E$ for all (equivalently, for some) $\lambda > 0$.*

PROOF. See Proposition 3.3 page 99 of [20]. □

PROPOSITION B.18 *Any linear maximal dissipative operator A in a Hilbert space H is closed if and only if it has a dense domain.*

PROOF. See [383], Theorem 1.1.1, p. 200-201 and Lemma 1.1.3, p. 201. □

PROPOSITION B.19 *Let E be a Banach space. Let $\lambda_0 > 0$ and*

$$F\colon [\lambda_0, +\infty) \to C^{0,1}(E; E)$$

be a function such that, for any $\lambda \ge \lambda_0$, $F(\lambda)$ is injective and, for any $\lambda, \mu \ge \lambda_0$,

$$F(\lambda) = F(\mu) \circ (I_E + (\mu - \lambda)F(\lambda)) .$$

Then there exists a unique operator $A\colon D(A) \subseteq E \to E$ such that, for any $\lambda \ge 0$,

$$F(\lambda) = (\lambda I - A)^{-1}.$$

If moreover

$$|F(\lambda)x - F(\lambda)y| \le \frac{1}{\lambda}|x - y|, \qquad \text{for all } \lambda \ge \lambda_0,\ x, y \in E,$$

then A is m-dissipative.

PROOF. Proposition I.3.3, page 13 of [108] ensures the existence of the operator A. The second part about the m-dissipativity of A follows from Proposition II.9.6 of [108]. □

We refer to Chapter 3 of [20], Appendix D of [130], Chapter 5 of [127], Chapter 3 of [132], [160], [375], [458] and [459] for more about dissipative and accretive operators.

B.3. Trace class and Hilbert-Schmidt operators

Throughout this section H, U, V will denote real, separable Hilbert spaces, $\langle \cdot, \cdot \rangle_H, \langle \cdot, \cdot \rangle_U, \langle \cdot, \cdot \rangle_V$, $|\cdot|_H, |\cdot|_U, |\cdot|_V$ will be respectively the inner products in H, U and V and the related norms.

DEFINITION B.20 *A linear operator $T \in \mathcal{L}(U, H)$ is called* nuclear *or* trace class *if T can be represented in the form*

$$T(z) = \sum_{k=1}^{+\infty} b_k \langle z, a_k \rangle_U \qquad \text{for any } z \in U,$$

where a_k and b_k are two sequences of elements respectively in U and H such that $\sum_{k=1}^{+\infty} |a_k|_U |b_k|_H < +\infty$. We denote the set of the nuclear operators from U to H by $\mathcal{L}_1(U, H)$. We write $\mathcal{L}_1(H)$ instead of $\mathcal{L}_1(H, H)$.

PROPOSITION B.21 *$\mathcal{L}_1(U,H)$ is a separable Banach space with respect to the norm*

$$\|T\|_{\mathcal{L}_1(U,H)} := \inf\Bigg\{ \sum_{k=1}^{+\infty} |a_k|_U |b_k|_H \;:\; \{a_k\}\subseteq U,\, \{b_k\}\subseteq H, \text{ and } T(z) = \sum_{k=1}^{+\infty} b_k \langle z, a_k\rangle_U, \forall z\in U \Bigg\}. \tag{B.2}$$

PROOF. See [409] Proposition 2.8, page 21 and the subsequent observations. □

PROPOSITION B.22 *Given $T\in\mathcal{L}_1(H)$ and an orthonormal basis $\{e_k\}$ of H, the series*

$$\sum_{k\in\mathbb{N}} \langle Te_k, e_k\rangle_H$$

converges absolutely and its sum does not depend on the choice of the basis $\{e_k\}$.

PROOF. See [380], pages 357-358 after the proof of the Proposition A.4. □

DEFINITION B.23 *Given $T\in\mathcal{L}_1(H)$ and any orthonormal basis $\{e_k\}$ of H,*

$$\mathrm{Tr}(T) := \sum_{k\in\mathbb{N}} \langle Te_k, e_k\rangle_H$$

is called the trace *of T.*

We have $|\mathrm{Tr}(T)| \le \|T\|_{\mathcal{L}_1(H)}$ (see e.g. [380], page 357).

DEFINITION B.24 *Let $\{e_k\}_{k\in\mathbb{N}}$ an orthonormal basis of U. The space of* Hilbert-Schmidt operators *$\mathcal{L}_2(U,H)$ from U to H is defined by*

$$\mathcal{L}_2(U,H) := \left\{ T\in\mathcal{L}(U,H) \;:\; \sum_{k\in\mathbb{N}} |Te_k|_H^2 < +\infty \right\}. \tag{B.3}$$

We write $\mathcal{L}_2(H)$ for $\mathcal{L}_2(H,H)$.

PROPOSITION B.25 *The space of Hilbert-Schmidt operators $\mathcal{L}_2(U,H)$ does not depend on the choice of orthonormal basis $\{e_k\}_{k\in\mathbb{N}}$. It is a separable Hilbert space if endowed with the inner product*

$$\langle S,T\rangle_2 := \sum_{k\in\mathbb{N}}^{\infty} \langle Se_k, Te_k\rangle_H, \qquad S,T\in\mathcal{L}_2(U,H).$$

The inner product is independent of the choice of basis.

PROOF. See Appendix C of [130] after Proposition C.3. □

The following proposition follows easily from the definition of the space of Hilbert-Schmidt operators and elementary calculations.

PROPOSITION B.26

(i) *$T\in\mathcal{L}_2(U,H)$ if and only if $T^*\in\mathcal{L}_2(H,U)$. Moreover $\|T\|_{\mathcal{L}_2(U,H)} = \|T^*\|_{\mathcal{L}_2(H,U)}$.*

(ii) *If $T\in\mathcal{L}_2(U,H)$ and $S\in\mathcal{L}(H,V)$ then $ST\in\mathcal{L}_2(U,V)$ and*

$$\|ST\|_{\mathcal{L}_2(U,V)} \le \|S\|_{\mathcal{L}(H,V)} \|T\|_{\mathcal{L}_2(U,H)}.$$

If $T\in\mathcal{L}(U,H)$ and $S\in\mathcal{L}_2(H,V)$ then $ST\in\mathcal{L}_2(U,V)$ and

$$\|ST\|_{\mathcal{L}_2(U,V)} \le \|S\|_{\mathcal{L}_2(H,V)} \|T\|_{\mathcal{L}(U,H)}.$$

DEFINITION B.27 *Let $U \subseteq H$. The embedding $U \subseteq H$ is said to be* Hilbert-Schmidt *if for some orthonormal basis $\{e_k\}_{k\in\mathbb{N}}$ of U, we have*

$$\sum_{k\in\mathbb{N}} |e_k|_H^2 < +\infty.$$

Thanks to Proposition B.25 this definition does not depend on the choice of basis $\{e_k\}$.

PROPOSITION B.28 *The following properties hold:*

(i) *If $T \in \mathcal{L}_1(U,H)$ and $S \in \mathcal{L}(H,V)$ then ST is in $\mathcal{L}_1(U,V)$ and*

$$\|ST\|_{\mathcal{L}_1(U,V)} \le \|S\|_{\mathcal{L}(H,V)} \|T\|_{\mathcal{L}_1(U,H)}.$$

If $T \in \mathcal{L}(U,H)$ and $S \in \mathcal{L}_1(H,V)$ then ST is in $\mathcal{L}_1(U,V)$ and

$$\|ST\|_{\mathcal{L}_1(U,V)} \le \|S\|_{\mathcal{L}_1(H,V)} \|T\|_{\mathcal{L}(U,H)}.$$

(ii) *If $T \in \mathcal{L}_1(U,H)$ and $S \in \mathcal{L}(H,U)$ (respectively, $T \in \mathcal{L}(U,H)$ and $S \in \mathcal{L}_1(H,U)$) then ST is in $\mathcal{L}_1(U)$, TS is in $\mathcal{L}_1(H)$ and*

$$\mathrm{Tr}(ST) = \mathrm{Tr}(TS).$$

(iii) *If $T \in \mathcal{L}_2(U,H), S \in \mathcal{L}_2(H,V)$ then $ST \in \mathcal{L}_1(U,V)$ and*

$$\|ST\|_{\mathcal{L}_1(U,V)} \le \|S\|_{\mathcal{L}_2(H,V)} \|T\|_{\mathcal{L}_2(U,H)}.$$

Moreover, if $U = V$, then $\mathrm{Tr}(ST) = \mathrm{Tr}(TS)$.

(iv) $\mathcal{L}_1(U,H) \subseteq \mathcal{L}_2(U,H)$.

(v) *If $T \in \mathcal{L}_2(U,H)$ then T is compact.*

PROOF. Most claims are proved in Appendix A.2 of [380]. More precisely: (i) is proved in Propositions 4, page 356, (ii) is proved in Proposition A.5-(i), page 358, while (iv) and (v) are proved in Proposition A.6, page 359. The proof of (iii) also repeats the proof of Proposition A.5-(ii), page 358, however we include it here for completeness. Consider an orthonormal basis $\{f_k\}_{k\in\mathbb{N}}$ of H. For every $z \in U$ we have

$$Tz = \sum_{k\in\mathbb{N}} \langle Tz, f_k\rangle_H f_k = \sum_{k\in\mathbb{N}} \langle z, T^* f_k\rangle_U f_k$$

so

$$STz = \sum_{k\in\mathbb{N}} \langle z, T^* f_k\rangle_U Sf_k \tag{B.4}$$

and then, using the definition of $\mathcal{L}_1$-norm given in (B.2) and the Cauchy-Schwarz inequality, we get

$$\begin{aligned}\|ST\|_{\mathcal{L}_1(U,V)} &\le \sum_{k\in\mathbb{N}} |T^* f_k|_U |Sf_k|_V \le \left(\sum_{k\in\mathbb{N}} |T^* f_k|_U^2\right)^{1/2} \left(\sum_{k\in\mathbb{N}} |Sf_k|_V^2\right)^{1/2} \\ &= \|T^*\|_{\mathcal{L}_2(H,U)} \|S\|_{\mathcal{L}_2(H,V)} = \|T\|_{\mathcal{L}_2(U,H)} \|S\|_{\mathcal{L}_2(H,V)}, \end{aligned} \tag{B.5}$$

where in the last step we used Proposition B.26-(i). It also follows from (B.4) and the Parseval identity that

$$\mathrm{Tr}(ST) = \sum_{k\in\mathbb{N}} \langle Sf_k, T^* f_k\rangle_U = \sum_{k\in\mathbb{N}} \langle TSf_k, f_k\rangle_H = \mathrm{Tr}(TS).$$

□

Observe that in general, if $T, S \in \mathcal{L}(H)$ and $TS \in \mathcal{L}_1(H)$, this does not necessarily imply $ST \in \mathcal{L}_1(H)$. Indeed consider two operators defined on $H \times H$ by

$$T = \begin{pmatrix} I & I \\ -I & -I \end{pmatrix}, \quad S = \begin{pmatrix} -I & I \\ I & -I \end{pmatrix},$$

where I stands for the identity operator. Then

$$TS = 0, \quad ST = 2\begin{pmatrix} -I & -I \\ I & I \end{pmatrix},$$

however ST is not trace class if H is infinite dimensional. Another example is given in [129], page 6.

NOTATION B.29 We set

$$\mathcal{L}^+(H) := \{T \in \mathcal{S}(H) \ : \ \langle Tx, x\rangle_H \geq 0 \ \ \forall x \in H\}$$

and

$$\mathcal{L}_1^+(H) := \mathcal{L}_1(H) \cap \mathcal{L}^+(H).$$

■

The operators in $\mathcal{L}^+(H)$ are called *positive operators* on H. A positive operator T on H is called *strictly positive* if it satisfies $\langle Tx, x\rangle_H > 0$ for all $x \in H$.

PROPOSITION B.30 *An operator $T \in \mathcal{L}^+(H)$ is nuclear if and only if*

$$\sum_{k\in\mathbb{N}} \langle Te_k, e_k\rangle_H < +\infty.$$

for an orthonormal basis $\{e_n\}$ on H. Moreover in this case $\mathrm{Tr}(T) = \|T\|_{\mathcal{L}_1(H)}$.

PROOF. See [130], Proposition C.3. □

B.4. C_0-semigroups and related results

B.4.1. Basic definitions.

DEFINITION B.31 (C_0-semigroup) *A map $S : [0, +\infty) \to \mathcal{L}(E)$ is called a* C_0-semigroup *(or a strongly continuous semigroup) on E if the following three conditions are satisfied:*

(i) $S(0) = I$.
(ii) *For all $s, t \in [0, +\infty)$, $S(t)S(s) = S(t + s)$.*
(iii) *For all $x \in E$, the map $t \mapsto S(t)x$ is continuous from $[0, +\infty)$ to E*[1].

For C_0-semigroups we will use the notation $\{S(t), t \geq 0\}$ or simply $S(t)$.

DEFINITION B.32 (Generator of a C_0-semigroup) *Let $S(t)$ be a C_0-semigroup on E. The linear operator $A : D(A) \subset E \to E$ defined as*

$$\begin{cases} D(A) := \left\{x \in E \ : \ \frac{S(t)x - x}{t} \ \text{ has a limit in } E \text{ when } t \to 0^+\right\} \\ Ax := \lim_{t\to 0^+} \frac{S(t)x - x}{t} \end{cases}$$

is called the infinitesimal generator *of $S(t)$.*

PROPOSITION B.33 *Let $S(t)$ be a C_0-semigroup on E. Then there exist $M \geq 1$ and $\omega \in \mathbb{R}$ such that*

$$\|S(t)\| \leq Me^{\omega t}, \qquad \text{for } t \geq 0. \tag{B.6}$$

PROOF. See for instance [375], Theorem 2.2, Chapter 1, page 4. □

The infimum of all ω such that (B.6) is satisfied for some M_ω is called the *type* of the C_0-semigroup $S(t)$ and is denoted by ω_0, see [35], Part II, Section 2.2. We have $\omega_0 \in [-\infty, +\infty)$. If $\omega_0 < 0$ we say that the C_0-semigroup $S(t)$ is of *negative type* and if $\omega_0 > 0$ we say that the C_0-semigroup $S(t)$ is of *positive type*.

[1] Equivalently one can ask here that $\lim_{t\searrow 0} S(t)x = x$ for all $x \in E$, see e.g. [375] Corollary 2.3, p.4.

Definition B.34 (Contraction semigroup) *A C_0-semigroup $S(t)$ on E is called a* C_0-semigroup of contractions*, if (B.6) holds with $M = 1, \omega = 0$.*

Definition B.35 (Pseudo-contraction semigroup) *A C_0-semigroup $S(t)$ on E is called a* C_0-semigroup of pseudo-contractions *if (B.6) holds with $M = 1$ for some $\omega \in \mathbb{R}$.*

Definition B.36 (Uniformly bounded semigroup) *A C_0-semigroup $S(t)$ on E is called* uniformly bounded *if (B.6) holds with $\omega = 0$ for some $M \geq 1$.*

We remark that the complexification $S_c(t)$ of a C_0-semigroup on E is a C_0-semigroup on E_c whose generator is the complexification A_c of the generator A of $S(t)$.

B.4.2. Hille-Yosida Theorem and Yosida approximation.

Theorem B.37 (Hille-Yosida) *A linear operator $A : D(A) \subset E \to E$ is the infinitesimal generator of a C_0-semigroup $S(t)$ on E satisfying (B.6) if and only if*

(1) *A is a closed and $D(A)$ is dense in E,*
(2) *$(\lambda I - A)$ is invertible and $(\lambda I - A)^{-1} \in \mathcal{L}(E)$ for every $\lambda > \omega$, and*

$$\|((\lambda I - A)^{-1})^k\| \leq M\,(\lambda - \omega)^{-k} \quad \text{for all } k \in \mathbb{N} \text{ and } \lambda > \omega.$$

Proof. See [222], Theorem 2.13, p. 20 or [375], Theorem 5.3. □

In fact, see [375], Remark 5.4, we have the following.

Remark B.38 For complex spaces condition (2) in Theorem B.37 can be replaced by $\{\lambda \in \mathbb{C} \,:\, \mathrm{Re}\lambda > \omega\} \subset \varrho(A)$ and, for all $k \in \mathbb{N}$ and $\lambda \in \mathbb{C}, \mathrm{Re}\lambda > \omega$,

$$\|((\lambda I - A)^{-1})^k\| \leq M\,(\mathrm{Re}\lambda - \omega)^{-k}.$$

■

Remark B.39 Considering, if needed, $\tilde{S}(t) := e^{-(\omega+\epsilon)t} S(t), \epsilon > 0$, we can always restrict to uniformly bounded C_0-semigroups having invertible generators. In particular, condition (2) of Theorem B.37 can be assumed to hold with $\omega = 0$. ■

Definition B.40 (Yosida approximations) *Let $A : D(A) \subset E \to E$ be the infinitesimal generator of a C_0-semigroup $S(t)$ on E satisfying (B.6). For $n \in \mathbb{N}$ greater than ω, define*

$$J_n = n(nI - A)^{-1}. \tag{B.7}$$

The Yosida approximation *of A is defined as follows:*

$$A_n := n^2(nI - A)^{-1} - nI = AJ_n \in \mathcal{L}(E). \tag{B.8}$$

Lemma B.41 *Consider A and J_n be as in Definition B.40. Then*

$$\|J_n\| \leq \frac{Mn}{n - \omega} \qquad \textit{for all } n > \omega. \tag{B.9}$$

Proof. It follows directly from condition (2) ($k = 1$) of Theorem B.37. □

Proposition B.42 *Let $A : D(A) \subset E \to E$ be the infinitesimal generator of a C_0-semigroup on E. Let J_n and A_n be as in Definition B.40. Then*

$$J_n x \xrightarrow[E]{n\to\infty} x \qquad \textit{for all } x \in E \tag{B.10}$$

and

$$A_n x \xrightarrow[E]{n\to\infty} Ax \qquad \textit{for all } x \in D(A).$$

Proof. See [130], Proposition A.4, page 409. □

PROPOSITION B.43 *Let $A : D(A) \subset E \to E$ be the infinitesimal generator of a C_0-semigroup $S(t)$ on E satisfying (B.6). Let A_n be the Yosida approximation of A. Define, for $x \in E$ and $t \geq 0$,*

$$e^{tA_n}x := \sum_{j\in\mathbb{N}} \frac{t^j A_n^j x}{j!}.$$

Then,

$$\|e^{tA_n}\| \leq Me^{t\frac{n\omega}{n-\omega}}. \tag{B.11}$$

and

$$S(t)x = \lim_{n\to\infty} e^{tA_n}x \qquad \textit{for every } x \in E \tag{B.12}$$

uniformly on bounded subsets of $[0, +\infty)$.

PROOF. See [375], Theorem 5.5, p. 21 and [35], Step 2 of the proof of Theorem 2.5, p. 102-103. □

Expression (B.12) shows how to construct explicitly the semigroup generated by a linear operator A.

NOTATION B.44 The semigroup generated by A will be denoted by e^{tA}. ■

THEOREM B.45 (Lumer-Phillips) *Let H be a separable Hilbert space. Given a linear operator $A : D(A) \subset H \to H$, the following facts are equivalent:*

(1) *A is the generator of a C_0-semigroup of contractions on H.*
(2) *$\overline{D(A)} = H$ and A is maximal dissipative.*
(3) *$\overline{D(A)} = H$ and A^* is maximal dissipative.*
(4) *$\overline{D(A)} = H, A$ is dissipative and $R(\lambda_0 I - A) = H$ for some $\lambda_0 > 0$.*

PROOF. See [383], Theorem 1.1.3, page 203, Theorem 1.4.2, page 214, and [375], Theorem 4.3, page 14. □

The following results is a corollary of the Trotter-Kato Theorem.

PROPOSITION B.46 (Trotter-Kato) *Let $S(t), S_n(t)$, $n \in \mathbb{N}$, be strongly continuous semigroups on E with generators A and A_n, respectively. Assume that $D(A) \subset D(A_n)$ for every $n \in \mathbb{N}$ and that*

$$\|S(t)\|, \|S_n(t)\| \leq Me^{\omega t}, \qquad \forall t \geq 0,\ n \in \mathbb{N}$$

and some constants $M \geq 1$ and $\omega \in \mathbb{R}$. If $A_n x \to Ax$ for every $x \in D(A)$ then

$$S_n(t)x \to S(t)x, \qquad \forall t \geq 0,\ x \in X,$$

and the limit is uniform in t for t in bounded intervals.

PROOF. See [375], Theorem 4.5, p. 88, or [160], Theorem 4.8, p. 209. □

PROPOSITION B.47 *Let $S(t)$ be a strongly a continuous semigroup on E with the generator A. Let $Y \subseteq D(A)$ be a subspace of $D(A)$. Assume that*

(i) *Y is dense on E,*
(ii) *$S(t)(Y) \subseteq Y$ for all $t \geq 0$.*

Then Y is a core for A.

PROOF. See [110], Proposition A.19, page 204. □

B.4.3. Analytic semigroups and fractional powers of generators. Throughout this section H is a separable Hilbert space. The material about analytic semigroups requires that H be complex. The statements for real H should be understood with the convention that $H, A, S(t)$ are their complexifications $H_c, A_c, S_c(t)$.

DEFINITION B.48 (Differentiable semigroup) *A C_0-semigroup $S(t)$ on H is called* differentiable, *if for every $x \in H$, $t \to S(t)x$ is differentiable for $t > 0$.*

DEFINITION B.49 (Analytic semigroup) *A C_0-semigroup $S(t)$ on H is called* analytic *if it has an extension $G(z)$ to a sector of the form $\Delta := \{z \in \mathbb{C} \, : \, a < \arg(z) < b\}$ for some $a < 0 < b$ with the following properties:*

(i) *$z \mapsto G(z)$ is analytic on Δ.*
(ii) *$G(0) = I$ and $\lim_{\substack{z \in \Delta \\ z \to 0}} G(z)x = x$ for every $x \in H$.*
(iii) *$G(z_1 + z_2) = G(z_1)G(z_2)$ for $z_1, z_2 \in \Delta$.*

THEOREM B.50 *Let $S(t)$ be a uniformly bounded C_0-semigroup on H and let A be the generator of $S(t)$. Assume that $0 \in \varrho(A)$. The following are equivalent:*

(1) *$S(t)$ can be extended to an analytic semigroup in a sector $\Delta_\delta = \{z \in \mathbb{C} \, : \, |\arg(z)| < \delta\}$, and $\|S(z)\|$ is uniformly bounded in every subsector $\Delta_{\delta'} = \{z \in \mathbb{C} \, : \, |\arg(z)| \le \delta'\}, \delta' < \delta$.*
(2) *There exists $\delta \in \left(0, \frac{\pi}{2}\right)$ and $B > 0$ such that*

$$\Sigma := \left\{\lambda \, : \, |\arg(\lambda)| < \frac{\pi}{2} + \delta\right\} \cup \{0\} \subset \varrho(A)$$

and, for every $\lambda \in \Sigma \setminus \{0\}$,

$$\|(\lambda I - A)^{-1}\| \le \frac{B}{|\lambda|}.$$

(3) *$S(t)$ is differentiable and there exists a constant $C > 0$ such that*

$$\|AS(t)\| \le \frac{C}{t} \quad \textit{for } t > 0.$$

PROOF. See [375] Theorem 5.2, page 61. □

THEOREM B.51 *Consider a linear operator $A : D(A) \subset H \to H$ that generates a uniformly bounded analytic C_0-semigroup $S(t)$, and $0 \in \varrho(A)$. Define, for $\alpha < 0$,*

$$(-A)^\alpha := \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{-\alpha-1} S(t) dt, \tag{B.13}$$

where $\Gamma(\cdot)$ is the Gamma function, and set $(-A)^0 := I$. Then:

(i) *The integral in (B.13) converges in norm and $(-A)^\alpha$ is a well defined operator in $\mathcal{L}(H)$.*
(ii) *$(-A)^\alpha(-A)^\beta = (-A)^{\alpha+\beta}$ for $\alpha, \beta \le 0$.*
(iii) *$(-A)^\alpha$ is injective.*

PROOF. See [375] pages 70-72. □

The operator $(-A)^\alpha$ can also be defined by

$$(-A)^\alpha = -\frac{1}{2\pi i} \int_C \lambda^\alpha (\lambda I + A)^{-1} d\lambda, \tag{B.14}$$

where the path C is in $\varrho(-A)$ and goes from $\infty e^{-i\theta}$ to $\infty e^{i\theta}$ for $\theta \in \left(\frac{\pi}{2}, \pi\right)$, avoiding the non-positive real axis. Using (B.14) one can define the fractional powers for more general operators, as it is done in classical references like [431] or [375]. We are only interested in the analytic case in this book.

Definition B.52 *Let A be as in Theorem B.51. The* fractional powers of $-A$ *are defined as follows:*

(i) *If $\alpha < 0$, $D((-A)^\alpha) = H$ and $(-A)^\alpha$ is defined by (B.13).*
(ii) *If $\alpha = 0$, $D((-A)^0) = H$ and $(-A)^0 := I$.*
(iii) *If $\alpha > 0$, $D((-A)^\alpha) = R((-A)^{-\alpha})$ and $(-A)^\alpha := ((-A)^{-\alpha})^{-1}$.*

Theorem B.53 *Let A be as in Theorem B.51. The fractional powers of $-A$ satisfy the following properties:*

(i) *For all positive α, $D((-A)^\alpha)$ is dense in H and $(-A)^\alpha$ is a closed operator.*
(ii) *If $\alpha \le \beta$ then $D((-A)^\beta) \subset D((-A)^\alpha)$.*
(iii) *For every real numbers α, β, $(-A)^\alpha(-A)^\beta = (-A)^{\alpha+\beta}$ on $D((-A)^{\beta\vee(\alpha+\beta)})$.*

Proof. See [375], Theorem 6.8, page 72. □

Theorem B.54 *Let A be as in Theorem B.51. The following hold:*

(i) *$e^{tA}(H) \subset D((-A)^\alpha)$ for every $t > 0$ and $\alpha \ge 0$.*
(ii) *$e^{tA}(-A)^\alpha x = (-A)^\alpha e^{tA} x$ or every $x \in D((-A)^\alpha)$.*
(iii) *There exists an $M_\alpha > 0$ such that for every $t > 0$, the operator $(-A)^\alpha e^{tA} : H \to H$ is bounded and*

$$\|(-A)^\alpha e^{tA}\| \le M_\alpha t^{-\alpha} e^{-at}. \tag{B.15}$$

(iv) *If $\alpha \in (0, 1]$ then for every $x \in D((-A)^\alpha)$*

$$|e^{tA}x - x|_H \le C_\alpha t^\alpha |(-A)^\alpha x|_H$$

for some constant C_α independent of x.

Proof. See [375], Theorem 6.13, page 74. □

B.5. π-convergence, $\mathcal{K}$-convergence, π- and $\mathcal{K}$-continuous semigroups

In this section, where H is always a real separable Hilbert space, we introduce the notions of π-convergence, $\mathcal{K}$-convergence, π-continuous and $\mathcal{K}$-continuous semigroups and we recall their basic properties as well as other related notions. For further results and details we refer the reader to [386, 387, 389] and [129], Section 6.3, for π-convergence and π-continuous semigroups; to [75, 76, 82] and the appendix of [79], for $\mathcal{K}$-convergence and $\mathcal{K}$-continuous (also called weakly continuous, see e.g. [75]) semigroups. We also recall the paper [124] that deals with semigroups which are not strongly continuous.

Most of the literature on the present subject deals with the spaces $C_b(H)$ and $UC_b(H)$, except for [76, 225] which deal with $UC_m(H)$ and [225] which deals with $C_m(H)$ in the case of $\mathcal{K}$-continuous semigroups, and the final part of [386] (p. 293-294, see also [387], Section 6.5) which shows how to extend the results to the space $B_b(H)$, in the case of π-continuous semigroups. Here we present the results for π-continuous and $\mathcal{K}$-continuous semigroups mainly in the spaces $C_m(H)$ and $UC_m(H)$ because these are the spaces most commonly used in this book. In some cases we will also deal with $B_m(H)$.

B.5.1. π-convergence and $\mathcal{K}$-convergence. The definition of π-convergence can be found e.g. in [161], page 111, where it is called *bp*-convergence (bounded-pointwise), and in [386]; the former in spaces of continuous and bounded functions, the latter in spaces of uniformly continuous and bounded functions. For

$\mathcal{K}$-convergence in $UC_b(H)$ space the reader is referred to [75], [82] and, for the $C_m(H)$ (respectively, $UC_m(H)$) framework, to [225] (respectively, [76][2]).

We state all the definitions in this section considering $B_m(H)$ ($m \geq 0$) as the environment space. The same definitions hold if the basic space $B_m(H)$ is replaced by $C_m(H)$ or $UC_m(H)$ ($m \geq 0$).

Definition B.55 *[π-convergence] Let $m \geq 0$. A sequence $(f_n) \subseteq B_m(H)$ is said to be π-convergent to $f \in B_m(H)$ and we will write*

$$f_n \xrightarrow{\pi} f \qquad \text{or} \qquad f = \pi\text{-}\lim_{n\to\infty} f_n$$

if the following conditions hold:

(i) $\sup_{n\in\mathbb{N}} \|f_n\|_{B_m} < +\infty$.
(ii) $\lim_{n\to\infty} f_n(x) = f(x)$ *for any* $x \in H$.

Moreover, given $I \subseteq \mathbb{R}$, $t_0 \in I$, a family $(f_t)_{t\in I\setminus\{t_0\}} \subset B_m(H)$ and $f \in B_m(H)$ we write

$$f_t \xrightarrow[t\to t_0]{\pi} f \qquad \text{or} \qquad f = \pi\text{-}\lim_{t\to t_0} f_n$$

if, for any sequence t_n of elements of $I \setminus \{t_0\}$ converging to t_0, we have $f_{t_n} \xrightarrow{\pi} f$.

Similarly, given $I \subseteq \mathbb{R}$, a sequence $(f_n) \subseteq B_m(I \times H)$ is said to be π-convergent to $f \in B_m(I \times H)$ and we will write

$$f_n \xrightarrow[n\to\infty]{\pi} f \qquad \text{or} \qquad f = \pi\text{-}\lim_{n\to\infty} f_n$$

if the following conditions hold:

(i) $\sup_{n\in\mathbb{N}} \|f_n\|_{B_m} < +\infty$.
(ii) $\lim_{n\to\infty} f_n(t,x) = f(t,x)$ *for any* $(t,x) \in I \times H$.

Definition B.56 *[$\mathcal{K}$-convergence] A sequence $(f_n) \subset B_m(H)$ is said to be $\mathcal{K}$-convergent to $f \in B_m(H)$ if*

$$\begin{cases} \sup\limits_{n\in\mathbb{N}} \|f_n\|_{B_m} < +\infty, \\ \lim\limits_{n\to+\infty} \sup\limits_{x\in K} |f_n(x) - f(x)| = 0 \end{cases} \tag{B.16}$$

for every compact set $K \subset H$. In this case we will write

$$f_n \xrightarrow[n\to\infty]{\mathcal{K}} f \qquad \text{or} \qquad f = \mathcal{K}\text{-}\lim_{n\to+\infty} f_n.$$

Moreover, given $I \subseteq \mathbb{R}$, $t_0 \in I$ and a family $(f_t)_{t\in I\setminus\{t_0\}} \subset B_m(H)$, where $I \subseteq \mathbb{R}$ and $f \in B_m(H)$, we write

$$f_t \xrightarrow[t\to t_0]{\mathcal{K}} f \qquad \text{or} \qquad f = \mathcal{K}\text{-}\lim_{t\to t_0} f_t$$

if, for any sequence t_n of elements of $I \setminus \{t_0\}$ converging to t_0, we have $f_{t_n} \xrightarrow{\mathcal{K}} f$.

In a similar way, given $I \subseteq \mathbb{R}$, a sequence $(f_n) \subset B_m(I \times H)$ is said to be $\mathcal{K}$-convergent to $f \in B_m(I \times H)$ if

$$\begin{cases} \sup\limits_{n\in\mathbb{N}} \|f_n\|_{B_m} < +\infty, \\ \\ \lim\limits_{n\to+\infty} \sup\limits_{(t,x)\in I_0\times K} |f_n(t,x) - f(t,x)| = 0 \end{cases} \tag{B.17}$$

[2] In this paper the $\mathcal{K}$-convergence in $UC_m(H)$ is called $\mathcal{K}_m$-convergence.

for all compact sets $I_0 \subseteq I$ and $K \subset H$[3]. In this case we will write as before

$$f_n \xrightarrow[n\to\infty]{\mathcal{K}} f \qquad or \qquad f = \mathcal{K}\text{-}\lim_{n\to+\infty} f_n.$$

The two convergencies that we have just introduced can be immediately generalized to the case of sequences in $B_m(H,Y)$ or in $B_m(I \times H, Y)$ where Y is a given Banach space. Moreover they induce a series of related concepts like those of closedness and density. In the two following definitions we recall some of them taken from [386, 387] and [82].

DEFINITION B.57 *Let $m \geq 0$. A subset Y of $B_m(H)$ (respectively $C_m(H)$, $UC_m(H)$) is said to be π-closed if, for any sequence $(f_n) \subset Y$ and $f \in B_m(H)$ (respectively $C_m(H)$, $UC_m(H)$) such that $f_n \xrightarrow{\pi} f$, we have $f \in Y$.*

A subset Y of $B_m(H)$ (respectively $C_m(H)$, $UC_m(H)$) is said to be π-dense in $B_m(H)$ (respectively $C_m(H)$, $UC_m(H)$) if, for any $f \in B_m(H)$ (respectively $C_m(H)$, $UC_m(H)$), there exists $(f_n) \subseteq Y$ such that $f_n \xrightarrow{\pi} f$.

A linear operator $\mathcal{A} : D(\mathcal{A}) \subset B_m(H) \to B_m(H)$ is said to be π-closed if, given a sequence $(f_n) \subseteq D(\mathcal{A})$, the following condition holds:

$$\Big(f_n \xrightarrow{\pi} f \qquad and \qquad \mathcal{A}f_n \xrightarrow{\pi} g\Big) \Rightarrow \Big(f \in D(\mathcal{A}) \qquad and \qquad \mathcal{A}f = g\Big).$$

The same if $\mathcal{A} : D(\mathcal{A}) \subset C_m(H) \to C_m(H)$ or $\mathcal{A} : D(\mathcal{A}) \subset UC_m(H) \to UC_m(H)$.

DEFINITION B.58 *Let $m \geq 0$. A subset Y of $B_m(H)$ (respectively $C_m(H)$, $UC_m(H)$) is said to be $\mathcal{K}$-closed if for any sequence $(f_n) \subset Y$ and $f \in B_m(H)$ (respectively $C_m(H)$, $UC_m(H)$) such that*

$$\mathcal{K}\text{-}\lim_{n\to+\infty} f_n = f,$$

we have $f \in Y$.

A subset Y of $B_m(H)$ (respectively $C_m(H)$, $UC_m(H)$) is said to be $\mathcal{K}$-dense if for any $f \in B_m(H)$ (respectively $C_m(H)$, $UC_m(H)$) there exists a sequence $(f_n) \subseteq Y$ such that

$$f = \mathcal{K}\text{-}\lim_{n\to+\infty} f_n.$$

A linear operator $\mathcal{A} : D(\mathcal{A}) \subset B_m(H) \to B_m(H)$ is said to be $\mathcal{K}$-closed if, given a sequence $(f_n) \subset D(\mathcal{A})$ such that

$$\mathcal{K}\text{-}\lim_{n\to+\infty} f_n = f \quad and \quad \mathcal{K}\text{-}\lim_{n\to+\infty} \mathcal{A}f_n = g,$$

we have

$$f \in D(\mathcal{A}) \quad and \quad \mathcal{A}f = g.$$

The same if $\mathcal{A} : D(\mathcal{A}) \subset C_m(H) \to C_m(H)$ or $\mathcal{A} : D(\mathcal{A}) \subset UC_m(H) \to UC_m(H)$.

Let $\mathcal{A} : D(\mathcal{A}) \subset B_m(H) \to B_m(H)$ and $\mathcal{B} : D(\mathcal{B}) \subset B_m(H) \to B_m(H)$ be two linear operators and assume that $\mathcal{A} \subset \mathcal{B}$ and that $\mathcal{B}$ is $\mathcal{K}$-closed. We say that $\mathcal{B}$ is the $\mathcal{K}$-closure of $\mathcal{A}$, and we write $\mathcal{B} = \overline{\mathcal{A}}^{\mathcal{K}}$, if for every $f \in D(\mathcal{B})$ there exists a sequence $(f_n) \subset D(\mathcal{A})$ such that

$$\begin{cases} \mathcal{K}\text{-}\lim\limits_{n\to+\infty} f_n = f \\ \\ \mathcal{K}\text{-}\lim\limits_{n\to+\infty} \mathcal{A}f_n = \mathcal{B}f. \end{cases} \tag{B.18}$$

[3] In the literature (see e.g. [82, 233, 234]) this definition is given taking the supremum over $I \times K$ even when I is not compact. We prefer to use the above definition as it is more coherent with the concept of $\mathcal{K}$-convergence and it does not change anything in the results and in the proofs.

The same if $\mathcal{A} : D(\mathcal{A}) \subset C_m(H) \to C_m(H)$, $\mathcal{B} : D(\mathcal{B}) \subset C_m(H) \to C_m(H)$ *or if* $\mathcal{A} : D(\mathcal{A}) \subset UC_m(H) \to UC_m(H)$, $\mathcal{B} : D(\mathcal{B}) \subset UC_m(H) \to UC_m(H)$.

Motivated by Theorem 4.5 of [225] we introduce the notions of $\mathcal{K}$-core and π-core, .

DEFINITION B.59 ($\mathcal{K}$-core and π-core) *Let* $m \geq 0$. *Let the operator* $\mathcal{A}\colon D(\mathcal{A}) \subseteq B_m(H) \to B_m(H)$ *be* $\mathcal{K}$*-closed (respectively* π*-closed). A linear subspace* Y *of* $D(\mathcal{A})$ *is a* $\mathcal{K}$*-core (respectively, a* π*-core) for* $\mathcal{A}$ *if for any* $f \in D(\mathcal{A})$ *there exists a sequence* f_n *of elements of* Y *such that*

$$\mathcal{K}\text{-}\lim_{n\to+\infty} f_n = f \qquad \textit{and} \qquad \mathcal{K}\text{-}\lim_{n\to+\infty} \mathcal{A}f_n = \mathcal{A}f$$

(respectively,

$$\pi\text{-}\lim_{n\to+\infty} f_n = f \qquad \textit{and} \qquad \pi\text{-}\lim_{n\to+\infty} \mathcal{A}f_n = \mathcal{A}f).$$

The same holds if $\mathcal{A} : D(\mathcal{A}) \subset C_m(H) \to C_m(H)$ *or* $\mathcal{A} : D(\mathcal{A}) \subset UC_m(H) \to UC_m(H)$.

REMARK B.60 Let $(f_n) \subseteq B_m(H)$ be a sequence which π-converges to a function $f : H \to \mathbb{R}$. Since measurability preserves over pointwise limits (see e.g. Lemma 1.8) then it must be $f \in B_m(H)$.

Similarly, since converging sequences in H are compact, one can prove that, if $(f_n) \subseteq C_m(H)$ is a sequence which $\mathcal{K}$-converges to a function $f : H \to \mathbb{R}$, then $f \in C_m(H)$ (i.e. $C_m(H)$ is $\mathcal{K}$-closed in $B_m(H)$).

On the other hand, if $f_n \xrightarrow[n\to\infty]{\mathcal{K}} f$ and $(f_n) \subseteq UC_m(H)$ it is not true, in general, that $f \in UC_m(H)$. Indeed, using Lemma B.77 below, one easily sees that $UC_m(H)$ is $\mathcal{K}$-dense in $C_m(H)$. Similarly, if $f_n \xrightarrow[n\to\infty]{\pi} f$ and $(f_n) \subseteq C_m(H)$ it is not true, in general, that $f \in C_m(H)$. ■

REMARK B.61 Concerning π-convergence, in [386], Theorem 2.2 (see also [387], Theorem 6.2.3), the author introduces a "natural" Hausdorff locally convex topology τ_0, not metrizable and not sequentially complete in $UC_b(H)$, whose convergent sequences are exactly π-convergent sequences. As stated in [386], end of Section 5, the same holds in $C_b(H)$. The lack of completeness of τ_0 relies on the fact that continuity (and, a fortiori, uniform continuity) is not preserved under π-convergence, as observed in the previous remark.

Similarly, concerning $\mathcal{K}$-convergence, in [225] it is shown (see in particular Proposition 2.3) that in the so-called *mixed topology* $\tau_{\mathcal{M}}$, which is a locally convex and complete one, introduced in [447], convergent sequences are precisely $\mathcal{K}$-convergent sequences. The result is obtained in $C_m(H)$ for $m \geq 0$. Completeness of $\tau_{\mathcal{M}}$ relies on the fact that continuity is preserved under $\mathcal{K}$-convergence, as observed in the previous remark. However, completeness of such topology is not guaranteed if we consider it on $UC_m(H)$, as observed in the introduction of [225].

In [161] p. 495-496 a topology in the space $B_b(H)$ is given whose convergent sequences coincide with π-convergent sequences.

It is clear that alls the concepts introduced in Definitions B.57, B.58 and B.59 can also be seen as topological concepts in the topologies just described. ■

REMARK B.62 In [327] (Definition 2.1 an Remark 2.6, see also [392][Section 5.2]) the author also considers π-convergence for multisequences and defines the concepts of π-closedness, π-density (Definition 2.5 there) and π-core (Definition 2.10) with multisequences. It seems that using multisequences the theory works as well since the dominated convergence is still preserved under π-convergent multisequences.

Since in this book we do not need such more general setting we keep the definitions with single sequences which, apriori, are not equivalent to the analogous definitions in [327]. ■

B.5.2. π- and $\mathcal{K}$-continuous semigroups and their generators. We use the definitions of the previous section to introduce π-continuous and $\mathcal{K}$-continuous semigroups. The theory of such semigroups has been developed in the literature mainly using $UC_b(H)$ as the environment space but the definitions and the results can be easily adapted to the $C_b(H)$ (or also to the $C_m(H)$ and $UC_m(H)$) framework. We will present the $UC_b(H)$ setting, making few remarks on how to generalize the results to $C_b(H)$ (or also to $C_m(H)$ and $UC_m(H)$).

B.5.2.1. *The definitions.*

DEFINITION B.63 (π-continuous semigroup) *Let $S(t)$ be a semigroup of bounded linear operators on $UC_b(H)$, namely, for any $f \in UC_b(H)$ and $s, t \in \mathbb{R}^+$, $S(t+s)f = S(t)S(s)f$ and $S(0)f = f$. We say that $S(t)$ is a π-continuous semigroup on $UC_b(H)$ of class $\mathcal{G}_\pi(M, \omega)$ if the following conditions hold:*

(i) *There exist $M \geq 1$ and $\omega \in \mathbb{R}$ s.t.*
$$\|S(t)\|_{\mathcal{L}(UC_b(H))} \leq M e^{\omega t}, \qquad t \geq 0.$$

(ii) *For any $(f_n) \subseteq UC_b(H)$, $f \in UC_b(H)$ s.t. $f_n \xrightarrow{\pi} f$ we have $S(t)f_n \xrightarrow[n\to\infty]{\pi} S(t)f$ for all $t \geq 0$.*

(iii) *For any $x \in H$ and $f \in UC_b(H)$, the map $[0, +\infty) \to \mathbb{R}$, $t \mapsto (S(t)f)(x)$ is continuous.*

DEFINITION B.64 ($\mathcal{K}$-continuous semigroup) *Let $S(t)$ be a semigroup of bounded linear operators on $UC_b(H)$, namely, for any $f \in UC_b(H)$ and $s, t \in \mathbb{R}^+$, $S(t+s)f = S(t)S(s)f$ and $S(0)f = f$. We say that $S(t)$ is a $\mathcal{K}$-continuous semigroup (or a weakly continuous semigroup) of class $\mathcal{G}_\mathcal{K}(M, \omega)$ on $UC_b(H)$ if the following conditions hold:*

(i) *There exist $M \geq 1$ and $\omega \in \mathbb{R}$ s.t.*
$$\|S(t)\|_{\mathcal{L}(UC_b(H))} \leq M e^{\omega t}, \qquad t \geq 0.$$

(ii) *For any $(f_n) \subseteq UC_b(H)$, $f \in UC_b(H)$ s.t. $f_n \xrightarrow{\mathcal{K}} f$ we have $S(t)f_n \xrightarrow[n\to\infty]{\mathcal{K}} S(t)f$ for all $t \geq 0$. The limit is uniform in $t \in [0, T]$ for any $T > 0$.*

(iii) *For every $f \in UC_b(H)$ and $t_0 \geq 0$ we have*
$$\mathcal{K}\text{-}\lim_{t \to t_0} S(t)f = S(t_0)f.$$

(iv) *For any $T > 0$ and $f \in UC_b(H)$, the family of functions*
$$\{\, S(t)f \,:\, t \in [0, T] \,\}$$
is equi-uniformly continuous in $UC_b(H)$.

We now give some observations concerning:

- The relationship between the above two definitions and the definition of a strongly continuous semigroup.
- The extension of the above two definitions to more general spaces.
- The relationship of the notions of $\mathcal{K}$ and π-continuous semigroups with strongly continuous semigroups in coarser topologies.
- The main reason why they are introduced: to deal with transition semigroups.

REMARK B.65 The above definitions were introduced first in [386, 387] (for π-semigroups) and in [75] (for $\mathcal{K}$-semigroups, there called *weakly continuous*). We observe the following.

(1) Differently from the case of strongly continuous semigroups (see Proposition B.33), condition (i) has to be included in both definitions above because it is not known if it follows from other assumptions.

(2) Conditions (ii) in both definitions gives a kind of continuity with respect to the π or $\mathcal{K}$-convergence. In condition (ii) for a $\mathcal{K}$-continuous semigroup (following [75]) the limit is required to be uniform in $t \in [0, T]$. Such a requirement is avoided in the definition given in [76]: we keep it here since it is verified in all cases we consider. Also this is not the case for π-continuous semigroups. There are examples of π-continuous semigroups where the limit is not uniform (see Remark 6.2.2 in [387]).

(3) Conditions (iii) in both definitions are analogues of condition (iii) in Definition B.31. However here the equivalence mentioned there in the footnote is not obvious. Indeed condition (iii) for a π-continuous semigroup clearly implies that

$$\pi\text{-}\lim_{t \to 0^+} S(t)f = f, \tag{B.19}$$

while the opposite is not obvious: the above implies right continuity of the map $t \mapsto (S(t)f)(x)$ but it is not known if also left continuity holds (see Remark 6.2.5 (a) in [387]). In [386] Remark 2.5 (see also Remark 6.2.5 (b) in [387]), the author proves that the equivalence holds for π-semigroups defined on $C_b(M)$ for a compact metric space M. Similarly for a $\mathcal{K}$-continuous semigroup is not obvious if condition (iii) for $t_0 = 0$ implies that the same holds true for all $t_0 \geq 0$.

The reason why in both cases the stronger conditions in the definitions are used is that the continuity of the map $t \mapsto S(t)f(x)$ is useful to simplify various steps in the proofs. Moreover the stronger conditions are verified in the cases we consider in this book. We finally observe that in the original definition of $\mathcal{K}$-continuous semigroup (see [75], Definition 2.1) the weaker condition (B.19) is used but it seems that the author still uses continuity of the map $t \mapsto S(t)f(x)$ (proof of Proposition 3.1, [75]).

(4) Condition (iv) for a $\mathcal{K}$-continuous semigroup extends the regularity in time of the function $(t, x) \to (S(t)f)(x)$ requiring uniformity in time of the modulus of continuity in x. We point out that in the original definition of $\mathcal{K}$-continuous semigroups in [75] the uniformity in conditions (ii) and (iv) are required on $[0, +\infty)$ since in this paper the author deals with semigroups of negative type which we do not want to assume here.

■

REMARK B.66 The definition of a π-continuous semigroup can also be made in $C_b(H)$ or in $B_b(H)$ by simply substituting $UC_b(H)$ with $C_b(H)$ or $B_b(H)$ (see [386] pages 293-294). Similarly we can define them in $UC_m(H)$, $C_m(H)$, $B_m(H)$.

Concerning $\mathcal{K}$-continuous semigroups, they can be easily defined in $UC_m(H)$ by simply substituting $UC_b(H)$ with $UC_m(H)$ as it is done in Section 2.2 of [76]. On the other hand, to define $\mathcal{K}$-continuous semigroups in $C_b(H)$ and in $C_m(H)$ one should also substantially modify (or erase) condition (iv) of Definition B.64. This is done in Theorem 4.1 if [225] where (iv) is implicitly substituted with the local equicontinuity of the family of operators $S(t)$, $t \geq 0$. The extension of the definition of $\mathcal{K}$-continuous semigroups to spaces $B_b(H)$ or $B_m(H)$ is not considered in the literature. ■

Remark B.67 Recalling Remark B.61, consider a π-continuous (respectively $\mathcal{K}$-continuous) semigroup $S(t)$ acting on the space $UC_b(H)$ endowed with the topology τ_0 generated by the π-convergence (respectively $\tau_{\mathcal{M}}$ generated by the $\mathcal{K}$-convergence). By Definition B.63-(ii) (respectively B.64-(ii)), for every $t \geq 0$, $S(t)$ is sequentially continuous but it is not known if it is also continuous. In [225] the authors show that the transition semigroups (introduced in the following section) are strongly continuous with respect to the mixed topology $\tau_{\mathcal{M}}$ so, in particular they are continuous in such topology. ■

Remark B.68 The reason why $\mathcal{K}$-continuous semigroups, and later, π-continuous semigroups, were introduced (in [75] and in [387], respectively) is the need for studying[4] Markov transition semigroups associated with finite and infinite dimensional SDE since such semigroups are naturally not C_0-semigroups: as shown for instance in Example 6.1 of [75], already in spatial dimension 1, the Ornstein-Uhlenbeck semigroup is not strongly continuous. Nevertheless by a simple application of the dominated convergence theorem, one can see that all Markov transition semigroups defined in (1.95) are π-continuous semigroups (see also Definition 3.5 of [386] and the subsequent comments). Moreover, with a slightly more complicated proof it can also be proved that such semigroups are $\mathcal{K}$-continuous.

On the other hand, it is not true, as one may expect, that all strongly continuous semigroups are also π-continuous or $\mathcal{K}$-continuous. Indeed it is shown in [387], Proposition 6.2.4 and the subsequent observation, that the class of uniformly continuous semigroups on $UC_b(\mathbb{R})$ is not contained in the class of π-continuous semigroups nor in that of $\mathcal{K}$-continuous semigroups. Moreover, even though the class of π-continuous semigroups has been introduced to study a wider set of problems (see Remark 6.2.2 of [387]), it is not clear if all $\mathcal{K}$-continuous semigroup are also π-continuous semigroups. ■

B.5.2.2. *The generators.* Similarly to what we have for C_0-semigroups, we can define the generators of π-continuous and $\mathcal{K}$-continuous semigroups. Given a semigroup of bounded operators $S(t)$, we will write

$$\Delta_h := \frac{S(h) - I}{h}.$$

Definition B.69 *Let $S(t)$ be a π-continuous semigroup on $UC_b(H)$. We define the infinitesimal generator $\mathcal{A}$ of $S(t)$ as follows:*

$$\begin{cases} D(\mathcal{A}) := \{f \in UC_b(H) \ : \ \textit{there exists } g \in UC_b(H) \textit{ s.t. } \pi\text{- }\lim_{t\to 0^+} \Delta_h f = g\} \\ (\mathcal{A}f)(x) := \lim_{t\to 0} \Delta_t f(x), \ \textit{for } x \in H. \end{cases}$$

Definition B.70 *Let $S(t)$ be a $\mathcal{K}$-continuous semigroup on $UC_b(H)$. We define the infinitesimal generator $\mathcal{A}$ of $S(t)$ as follows:*

$$\begin{cases} D(\mathcal{A}) := \{f \in UC_b(H) \ : \ \textit{there exists } g \in UC_b(H) \textit{ s.t. } \mathcal{K}\text{- }\lim_{t\to 0^+} \Delta_h f = g\} \\ (\mathcal{A}f)(x) := \lim_{t\to 0} \Delta_t f(x), \ \textit{for } x \in H. \end{cases}$$

Remark B.71 In [75, 79, 80, 82] the generator $\mathcal{A}$ of a $\mathcal{K}$-continuous semigroup $S(t)$ is defined in a different way, by using the resolvent. Indeed $\mathcal{A}$ is the unique

[4] Other approaches are possible to deal with such semigroups (see e.g. the theory of semigroups on general locally convex spaces [277, 278]) but, as remarked in the introduction of [386], the use of such a theory for the above goal would be much more complicated.

closed linear operator such that, for all $\lambda > \omega$, $f \in UC_b(H)$, $x \in H$, we have

$$(\lambda I - \mathcal{A})^{-1} f(x) = \int_0^{+\infty} e^{-\lambda t} S(t) f(x) dt.$$

In fact the two definitions are equivalent. The equivalence is implicitly proved for π-continuous semigroups in [387][Proposition 6.2.11] (see also [386][Proposition 2.5]) with a proof that easily adapts to $\mathcal{K}$-continuous semigroups too. Moreover the equivalence of the two definitions is also proved in [225][Remark 4.3], for $\mathcal{K}$-continuous transition semigroups (which is the case of interest in this book).

In [387], Theorem 6.2.13, the author shows that, if a π-continuous semigroup is also $\mathcal{K}$-continuous, the two generators coincide. ■

We finally observe that the generators can be introduced, exactly in the same way, if we consider, as said in Remark B.66, π- continuous semigroups on $UC_m(H)$, $C_m(H)$, $B_m(H)$ and $\mathcal{K}$-continuous semigroups in $UC_m(H)$, $C_m(H)$, for some $m \geq 0$.

B.5.2.3. *Cauchy problems for π-continuous and $\mathcal{K}$-continuous semigroups.* The following propositions establish the relationship between π-continuous/$\mathcal{K}$-continuous semigroups, their generators and homogeneous Cauchy problems.

PROPOSITION B.72 *Let $\mathcal{A}$ be the generator of a π-continuous semigroup $S(t)$ on $UC_b(H)$. Then:*

- (i) *$D(\mathcal{A})$ is π-dense in $UC_b(H)$.*
- (ii) *$\mathcal{A}$ is a closed and π-closed operator on $UC_b(H)$.*
- (iii) *For any $f \in D(\mathcal{A})$:*
  - (a) *$S(t)f \in D(\mathcal{A})$ and $\mathcal{A}S(t)f = S(t)\mathcal{A}f$, for any $t \geq 0$.*
  - (b) *For any $x \in H$ the mapping*

$$(0, +\infty) \to \mathbb{R}, \quad t \mapsto (S(t)f)(x)$$

is continuously differentiable and

$$\frac{d}{dt} S(t) f(x) = \mathcal{A}(S(t)f)(x)$$

for all $t > 0$.

PROOF. See [387] Proposition 6.2.9 and 6.2.7. Closedness of $\mathcal{A}$ follows from Remark B.71. □

PROPOSITION B.73 *Let $\mathcal{A}$ be the generator of a $\mathcal{K}$-continuous semigroup $S(t)$ on $UC_b(H)$. Then:*

- (i) *$D(\mathcal{A})$ is $\mathcal{K}$-dense in $UC_b(H)$.*
- (ii) *$\mathcal{A}$ is a closed and $\mathcal{K}$-closed operator on $UC_b(H)$.*
- (iii) *For every $f \in D(\mathcal{A})$:*
  - (a) *$S(t)f \in D(\mathcal{A})$ and $\mathcal{A}S(t)f = S(t)\mathcal{A}f$, for any $t \geq 0$.*
  - (b) *For any $x \in H$ the mapping:*

$$(0, +\infty) \to \mathbb{R}, \quad t \mapsto (S(t)f)(x)$$

is continuously differentiable and

$$\frac{d}{dt} S(t) f(x) = \mathcal{A}(S(t)f)(x)$$

for all $t > 0$.

PROOF. See [82] Proposition 2.9 and Remark 2.10. Closedness of $\mathcal{A}$ follows from Remark B.71. □

REMARK B.74 In the case of a $\mathcal{K}$-continuous semigroup with generator $\mathcal{A}$, if we consider the function $u\colon [0,T] \times H \to \mathbb{R}$ defined as

$$(t, x) \mapsto (S(t)\varphi)(x), \tag{B.20}$$

which should be the natural solution of the homogeneous Cauchy problem corresponding to the operator $\mathcal{A}$, then u belongs to $UC_b^x([0,T]\times H)$[5] and not, in general, to $C([0,T]; UC_b(H)) = UC_b([0,T] \times H)$. Indeed, if for every datum $\varphi \in UC_b(H)$, $u \in C([0,T]; UC_b(H))$, then $\mathcal{A}$ generates a C_0-semigroup. Similarly, in the case of π-continuous semigroups the function in (B.20) does not belong to $UC_b([0,T] \times H)$ in general, but it belongs in a natural way to a space called $C_\pi([0,T]; UC_b(H))$ (requiring global boundedness and separate continuity), introduced in Definition 7.2.1 on page 161 of [387], which is strictly bigger than $UC_b^x([0,T] \times H)$.

Moreover, in general the $UC_b(H)$-valued function $t \mapsto e^{t\mathcal{A}}\varphi$ is not even measurable. In fact measurability of this map implies strong continuity of the semigroup (see [450], pages 233-234). ■

The above propositions can be extended, as said in Remark B.66, to the case of π- continuous semigroups in $UC_m(H)$, $C_m(H)$, $B_m(H)$, and $\mathcal{K}$-continuous semigroups in $UC_m(H)$, $C_m(H)$, for some $m \geq 0$.

We now consider non-homogeneous Cauchy problems. Similarly to the case of C_0-semigroups (see e.g. [35] Chapter II.1.3), given a π-continuous or a $\mathcal{K}$-continuous semigroup $S(t)$ and the related generator $\mathcal{A}$, one can consider the following non homogeneous Cauchy problem on $UC_b(H)$[6]:

$$\begin{cases} \frac{d}{dt}v(t) = \mathcal{A}v(t) + f(t), \quad t \in (0,T], \\ v(0) = \varphi \in UC_b(H), \end{cases} \tag{B.21}$$

where $f\colon (0,T] \to UC_b(H)$.

In analogy with what is usually done for C_0-semigroups, the mild solution of problem (B.21) is, by definition, the function $u : [0,T] \times H \to \mathbb{R}$ given by

$$u(t,x) = [S(t)\varphi](x) + \int_0^t [S(t-s)f(s)](x)ds. \tag{B.22}$$

REMARK B.75 This definition assumes that, for every $t > 0$ and $x \in H$, the map

$$[0,t] \to \mathbb{R}, \qquad s \mapsto [S(t-s)f(s)](x)$$

is measurable. This is true under mild assumptions on f but one has to be careful since, even when f above is constant, as recalled in Remark B.74, for a fixed $t \in [0,T]$, the function

$$[0,t] \to UC_b(H), \qquad s \mapsto S(t-s)f$$

may not be measurable if the semigroup $S(t)$ is not strongly continuous. ■

Results on uniqueness and regularity of mild solutions and their relationship with other concepts of solutions (in particular approximations by classical solutions) are contained in [82, 76] for the case of $\mathcal{K}$-continuous semigroups and in Chapter 7 of [387, 389] for π-continuous semigroups. They are applied mainly to obtain suitable approximations of mild solutions of Kolmogorov equations which are used in Chapter 4. Such results are presented in Section B.7 for the Kolmogorov equations we are interested in.

[5]Indeed as remarked in [82], Remark 2.6, this is equivalent to requiring that points (iii) and (iv) of Definition B.64 be satisfied.

[6]Clearly, extending the concepts of π-semigroups and $\mathcal{K}$-semigroups it is possible to consider the same problem in the spaces $C_b(H)$, $C_m(H)$, $UC_m(H)$, $B_m(H)$.

Several other results on π-continuous and $\mathcal{K}$-continuous semigroups are obtained in [386, 387, 389] and [75, 76, 82]. We recall in particular the possibility of proving an analogue of Hille-Yosida theorem in both cases.

B.6. Approximation of continuous functions through $\mathcal{K}$-convergence

Recall that H is a real separable Hilbert space. When $\dim H = +\infty$, the space $UC_b^2(H)$ is not dense in $UC_b(H)$ (see [354]). We refer to Appendix D.3 for more on approximations in Hilbert spaces. However, in many cases (in particular to prove verification theorems for optimal control problems), we need to be able to approximate functions in $C_m(H)$, $m \geq 0$, (or $UC_m(H)$) by smooth (at least C^2) functions with special properties. We can substitute uniform convergence by π-convergence or $\mathcal{K}$-convergence which are good enough to apply the dominated convergence theorem. We start with the following lemma which is a small variation of Lemma 5.2 of [82].

LEMMA B.76 *Let A be the generator of the strongly continuous semigroup $S(t) = e^{tA}$ in H. Let J_n as in (B.7) and let $\{P_n\}_{n\in\mathbb{N}}$ be a sequence of finite dimensional orthogonal projections on H strongly convergent to the identity operator. Then, for all compact subsets $I_0 \subset \mathbb{R}$ and $K \subset H$, the sets*

$$\mathcal{R}_1(K) =: \{\, P_n x \,:\, x \in K,\ n \in \mathbb{N} \,\} \subset H,$$

$$\mathcal{R}_1(I_0 \times K) =: \{\, (t, P_n x) \,:\, t \in I_0,\ x \in K,\ n \in \mathbb{N} \,\} \subset \mathbb{R} \times H,$$

$$\mathcal{R}_2(K) =: \{\, P_n J_n x \,:\, x \in K,\ n \in \mathbb{N} \,\} \subset H$$

are relatively compact.

PROOF. We show the claim for $\mathcal{R}_2(K)$. The argument for $\mathcal{R}_1(K)$ and $\mathcal{R}_1(I_0 \times K)$ is the same and easier. Let $\left\{P_{n_j} J_{n_j} x_j\right\}_{j\in\mathbb{N}}$ be a sequence in $\mathcal{R}_2(K)$. From compactness of K it follows that there exists an increasing subsequence j_k and $\bar{x} \in K$ such that

$$x_{j_k} \to \bar{x}, \quad \text{as } k \to +\infty.$$

If there exists $C > 0$ such that $n_{j_k} \leq C$, for all j_k, then we can suppose $n_{j_k} = \bar{n}$, for all j_k and then

$$\lim_{k\to+\infty} P_{n_{j_k}} J_{n_{j_k}} x_{j_k} = P_{\bar{n}} J_{\bar{n}} \bar{x}.$$

Otherwise, let us suppose $\lim_{k\to+\infty} n_{j_k} = +\infty$. Then, using (B.9) and (B.10),

$$\begin{aligned} &|P_{n_{j_k}} J_{n_{j_k}} x_{j_k} - \bar{x}| \leq |P_{n_{j_k}} J_{n_{j_k}} (x_{j_k} - \bar{x})| + |P_{n_{j_k}} (J_{n_{j_k}} \bar{x} - \bar{x})| + |P_{n_{j_k}} \bar{x} - \bar{x}| \\ &\leq \frac{M n_{j_k}}{n_{j_k} + \omega} |x_{j_k} - \bar{x}| + |J_{n_{j_k}} \bar{x} - \bar{x}| + |P_{n_{j_k}} \bar{x} - \bar{x}| \to 0 \quad \text{as } k \to +\infty. \end{aligned}$$

□

The following lemma generalizes Lemma 2.6, page 25 in [225].

LEMMA B.77 *Let $m \geq 0$ and let $I \subseteq \mathbb{R}$ be an interval. Then $\mathcal{F}C_0^\infty(H)$ is $\mathcal{K}$-dense in $C_m(H)$ and $\mathcal{F}C_0^\infty(I \times H)$ is $\mathcal{K}$-dense in $C_m(I \times H)$. Moreover, for every closed operator B with dense domain we also have that $\mathcal{F}C_0^{\infty,B}(H)$ is $\mathcal{K}$-dense in $C_m(H)$ and $\mathcal{F}C_0^{\infty,B}(I \times H)$ is $\mathcal{K}$-dense in $C_m(I \times H)$.*

PROOF. Take an orthonormal basis $\mathcal{E} = (e_n)_{n\in\mathbb{N}}$ of H and, for $x \in H$, let $x = \sum_{i=1}^{\infty} x_i e_i$. For every $n \in \mathbb{N}$ let P_n be the orthogonal projection onto the n dimensional subspace of H spanned by $\{e_1, \ldots e_n\}$. Define

$$\Pi_n : H \to \mathbb{R}^n, \qquad \Pi x = (x_1, \ldots, x_n),$$

$$Q_n : \mathbb{R}^n \to H, \qquad Q_n(x_1, \ldots, x_n) = x_1 e_1 + \cdots + x_n e_n,$$

and recall that $P_n = Q_n \circ \Pi_n$. Let $\varphi \in C_m(H)$. Given a family of C^∞ mollifiers $\eta_k : \mathbb{R}^n \to \mathbb{R}$ with support in $B(0, 1/k)$, we define the regularizing convolutions (as e.g. [379])

$$\psi_k^n(x) = \int_{\mathbb{R}^n} \varphi(Q_n y)\eta_k(\Pi_n x - y)dy = \int_{\mathbb{R}^n} \varphi(P_n x - Q_n y)\eta_k(y)dy.$$

Observe that, by the definition, we have for $x \in H$,

$$|\psi_n^k(x)| \le \sup_{|x_1| \le |P_n x| + 1/k} |\varphi(x_1)| \le \sup_{|x_1| \le |x| + 1/k} |\varphi(x_1)|,$$

so

$$\begin{aligned} \frac{|\psi_k^n(x)|}{(1+|x|^2)^{m/2}} &\le \sup_{|x_1| \le |x| + 1/k} \left\{ \frac{|\varphi(x_1)|}{(1+|x_1|^2)^{m/2}} \frac{(1+|x_1|^2)^{m/2}}{(1+|x|^2)^{m/2}} \right\} \\ &\le \|\varphi\|_{C_m(H)} \sup_{|x_1| \le |x| + 1/k} \frac{(1+|x_1|^2)^{m/2}}{(1+|x|^2)^{m/2}}. \end{aligned}$$

Now it is easy to see that

$$\sup_{|x_1| \le |x| + 1/k} \frac{1+|x_1|^2}{1+|x|^2} = 1 + \rho(1/k)$$

for some modulus ρ. Hence we get

$$\|\psi_k^n\|_{C_m(H)} \le \|\varphi\|_{C_m(H)} (1 + \rho(1/k))^{m/2}. \tag{B.23}$$

Now, from the properties of finite dimensional convolutions, we easily observe that, setting $\xi_n(x) := \varphi(P_n x)$, the sequence $\{\psi_k^n\}_{k \in \mathbb{N}}$ converges to ξ_n uniformly on bounded sets of $\mathbb{R}^n$. For every $n \in \mathbb{N}$, let $k(n) \in \mathbb{N}$ be such that

$$\sup_{|x| \le n} |\psi_{k(n)}^n(x) - \xi_n(x)| \le \frac{1}{n}. \tag{B.24}$$

Therefore, if we set $\psi_n = \psi_{k(n)}^n$, we have, for any compact set $K \subset H$ and $n \ge \sup_{x \in K} |x|$,

$$\begin{aligned} \sup_{x \in K} |\psi_n(x) - \varphi(x)| &\le \sup_{x \in K} |\psi_n(x) - \xi_n(x)| + \sup_{x \in K} |\xi_n(x) - \varphi(x)| \\ &\le \frac{1}{n} + \sup_{x \in K} |\xi_n(x) - \varphi(x)|. \end{aligned} \tag{B.25}$$

By Lemma B.76 and the continuity of φ it follows that, for all $x \in K$ and $n \in \mathbb{N}$,

$$|\xi_n(x) - \varphi(x)| \le \rho_{\mathcal{R}_1(K)} |P_n x - x|, \tag{B.26}$$

where $\rho_{\mathcal{R}_1(K)}$ is the modulus of continuity of φ over the compact set $\mathcal{R}_1(K)$. Since $\lim_{n \to +\infty} \sup_{x \in K} |P_n x - x| = 0$, it then follows from (B.23) and (B.25), that

$$\mathcal{K}\text{-}\lim_{n \to +\infty} \psi_n = \varphi. \tag{B.27}$$

To end the proof of the first statement in the time independent case it is enough to define $\varphi_n(x) := \psi_n(x)\theta(|P_n x|^2/n^2)$, where $\theta \in C^\infty(\mathbb{R})$ is a cut-off function such that $0 \le \theta \le 1$, $\theta(r) = 1$ for $|r| \le 1$ and $\theta(r) = 0$ for $|r| \ge 2$.

Consider now the time dependent case. Let $f \in C_m(I \times H)$. First let $I = \mathbb{R}$, and define f_n following the same procedure used for φ_n. Given a family of C^∞ mollifiers $\eta_k : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ with support in $B(0, 1/k) \subset \mathbb{R} \times \mathbb{R}^n$, we define the regularizing convolutions

$$\begin{aligned} g_k^n(t, x) &= \int_{\mathbb{R}^{n+1}} f(s, Q_n y)\eta_k(t - s, \Pi_n x - y)dsdy \\ &= \int_{\mathbb{R}^{n+1}} f(t - s, P_n x - Q_n y)\eta_k(s, y)dsdy. \end{aligned}$$

Similarly to the time independent case we prove that

$$\|g_k^n\|_{C_m(I\times H)} \le \|f\|_{C_m(I\times H)}(1+\rho(1/k))^{m/2}. \tag{B.28}$$

Furthermore we define $h_n(t,x) := f(t,P_n x)$ and $g_n(t,x) = g^n_{k(n)}(t,x)$ as we did in the time independent case for ξ_n and ψ_n, respectively. Hence we obtain

$$\mathcal{K}\text{-}\lim_{n\to+\infty} g_n = f. \tag{B.29}$$

Then, defining $f_n(t,x) := g_n(t,x)\theta(|P_n x|^2/n^2)$ with θ as in the time independent case, we get the claim.

If $I=[a,b]$ for $a<b\in\mathbb{R}$ we first define $\bar f(t,x) = f(a,x)$ for $t<a$ and $\bar f(t,x) = f(b,x)$ for $t>b$. We then take the approximations $\bar f_n$ defined above and finally define f_n as the restrictions of $\bar f_n$ to $I\times H$. In other cases, we first take a sequence of suitable approximations. For example, if $I=(a,b)$ is open we first define $\tilde f_h$ on $\mathbb{R}\times H$ for $h>4/(b-a)$ by $\tilde f_h(t,x) = \chi_h(t) f(t,x)$, where χ is continuous, $0\le\chi_h\le 1$, $\chi_h(t)=1$ for $t\in[a+2/h, b-2/h]$, $\chi_h(t)=0$ for $t\notin[a+1/h,b-1/h]$. Then for every h we approximate the functions $\tilde f_h$ by a sequence $\{f_{h,n}\}_n$ chosen as above. The diagonal sequence $\{f_n\}_n := \{f_{n,n}\}_n$ $\mathcal{K}$-converges to f. Indeed, given I_0 and K compact subsets of (a,b) and H, respectively, we have

$$\sup_{I_0\times K}|f_{n,n}(t,x)-f(t,x)| \le \sup_{I_0\times K}|f_{n,n}(x)-\tilde f_n(x)| + \sup_{I_0\times K}|\tilde f_n(x)-f(x)|.$$

The second term of the right hand side converges to 0 since, by construction, f_n $\mathcal{K}$-converges to f; for the first term we observe that, replicating (B.25) in this case, it remains to estimate $\sup_{I_0\times K}|\tilde f_n(t,P_n x)-\tilde f_n(t,x)|$. By construction, if n is sufficiently large, then $\tilde f_n = f$ on $I_0\times K$. Hence, arguing as in (B.26) and using the compactness of $\mathcal{R}_1(I_0\times K)$ from Lemma B.76, we get the claim.

Regarding the second statement, it is enough to take an orthonormal basis $\mathcal{E}$ which is contained in $D(B)$, it always exists since $D(B)$ is dense. □

B.7. Approximation of solutions of Kolmogorov equations through $\mathcal{K}$-convergence

Let H be a real separable Hilbert space and $T>0$. We consider the following initial value problem for a linear Kolmogorov equation[7]

$$\begin{cases} u_t = \dfrac{1}{2}\operatorname{Tr}[QD^2u] + \langle x, A^*Du\rangle + Cu + f(t,x) = 0, & (t,x)\in(0,T]\times H,\\[2ex] u(0,x)=\varphi(x),\ x\in H, \end{cases} \tag{B.30}$$

where $C\in\mathbb{R}$, $\varphi\in C(H)$, $f\in C((0,T]\times H)$ and A, Q satisfy the following.

Hypothesis B.78

(i) *A is the infinitesimal generator of a strongly continuous semigroup $\{e^{tA}, t\ge 0\}$ on H with $\|e^{tA}\|\le Me^{\omega t}$ for all $t\ge 0$ for given $M\ge 1$, $\omega\in\mathbb{R}$.*

(ii) *$Q\in\mathcal{L}^+(H)$ is such that, setting, for all $x\in H$, $Q_t x = \int_0^t e^{sA}Qe^{sA^*}x\,ds$, the operator Q_t is nuclear for all $t\ge 0$.*

[7]Often (see e.g. [129] Chapter 6) the term $\langle x, A^*Dv\rangle$ is replaced by the term $\langle Ax, Dv\rangle$ that is well defined if $x\in D(A)$. Here we use the former expression because we will look for more regular solutions, having the derivative in $D(A^*)$ (see the notion of *classical solution* used in Section 6.2 of [129]).

If we call, formally, $\mathcal{A}$ to be the operator associated to the second and third terms of (B.30) and, still formally, R_t to be the corresponding semigroup, then we can rewrite (B.30) in the following mild form

$$u(t,x) = e^{Ct} R_t[\varphi](x) + \int_0^t e^{C(t-s)} R_{t-s}[f(s,\cdot)](x)ds \tag{B.31}$$

and call the function on the right hand side the *mild solution* of (B.30). In what follows we provide some useful approximation results for solutions of such equations. Similar results are available in the literature (see e.g. [82, 233, 387]) but only when the data φ and f are bounded and in a slightly different setting. Here we allow the data to have polynomial growth and take a setting which is more suitable for the purpose of this book (see Remark B.95 for more on this).

In the next three subsections we first present (Subsection B.7.1) suitable definitions of classical and strong solutions of (B.30); then (Subsection B.7.2) we define precisely R_t, the mild solutions, and connect the generator $\mathcal{A}$ of R_t with (B.30) through the operators $\mathcal{A}_0$ and $\hat{\mathcal{A}}_0$; finally, (Subsection B.7.3) we show a useful approximation result.

B.7.1. Classical and strong solutions of (B.30). Let us introduce the operator $\mathcal{A}_0$ on $C(H)$ as follows (compare it with the operators $\mathcal{A}_1$ in Section 4.5 and $\mathcal{A}_0$ in Section 4.6):

$$\begin{cases} D(\mathcal{A}_0) = & \Big\{\phi \in UC_b^2(H) \,:\, A^* D\phi \in UC_b(H,H),\ D^2\phi \in UC_b(H,\mathcal{L}_1(H))\Big\} \\ \\ \mathcal{A}_0\phi = & \frac{1}{2}\mathrm{Tr}\,[QD^2\phi] + \langle x, A^*D\phi\rangle\,. \end{cases} \tag{B.32}$$

We also consider the restriction $\hat{\mathcal{A}}_0 \subseteq \mathcal{A}_0$ defined as in [82][Section 5] (see also [389][Section 5] or [387][Section 7.4.3]),

$$\begin{cases} D(\hat{\mathcal{A}}_0) = & \Big\{\phi \in UC_b^2(H) \,:\, A^* D\phi \in UC_b(H,H), \\ & \qquad \langle x, A^*D\phi\rangle \in UC_b(H) \text{ and } D^2\phi \in UC_b(H,\mathcal{L}_1(H))\Big\} \\ \\ \hat{\mathcal{A}}_0\phi = & \frac{1}{2}\mathrm{Tr}\,[QD^2\phi] + \langle x, A^*D\phi\rangle\,. \end{cases} \tag{B.33}$$

REMARK B.79 The operator $\hat{\mathcal{A}}_0$ can be seen as an unbounded operator on $C_m(H)$ for all $m \geq 0$ since, by the definition of $D(\hat{\mathcal{A}}_0)$, one easily sees that $\hat{\mathcal{A}}_0\phi \in C_b(H)$ (and so it also belongs to $C_m(H)$) for all $\phi \in D(\hat{\mathcal{A}}_0)$. On the other hand this is not the case for $\mathcal{A}_0$. Indeed, for a generic $\phi \in D(\mathcal{A}_0)$ we can only say that $\mathcal{A}_0\phi \in C_m(H)$ for $m \geq 1$ since the term $\langle x, A^*D\phi\rangle$ may be unbounded. The same considerations holds if we consider the operators in $UC_m(H)$.

Both operators are used to define suitable approximations of the solution of (B.30). In most of the literature only $\hat{\mathcal{A}}_0$ is used. Here we see that, for approximation purposes, also $\mathcal{A}_0$ can be used with some advantages. ■

We endow $D(\mathcal{A}_0)$ with the norm

$$\|\phi\|_{D(\mathcal{A}_0)} := \|\phi\|_0 + \|D\phi\|_0 + \|A^*D\phi\|_0 + \sup_{x\in H}\|D^2\phi(x)\|_{\mathcal{L}_1(H)}, \tag{B.34}$$

while in $D(\hat{\mathcal{A}}_0)$ we take the norm

$$\|\phi\|_{D(\hat{\mathcal{A}}_0)} := \|\phi\|_0 + \|D\phi\|_0 + \|A^*D\phi\|_0 + \|\,\langle x, A^*D\phi\rangle\,\|_0 + \sup_{x\in H}\|D^2\phi(x)\|_{\mathcal{L}_1(H)}. \tag{B.35}$$

Arguing as in Theorem 2.7 of [126], it can be proved that both $D(\mathcal{A}_0)$ and $D(\hat{\mathcal{A}}_0)$ with the above norms are Banach spaces[8]. However the Banach structure is not essential for our purposes even if it simplifies the notation.

We now give two notions of classical solutions of (B.30). The first, a more restrictive one, uses the operator $\hat{\mathcal{A}}_0$, and is in line with what is done e.g. in Definition 4.6 of [233] or Definition 4.1 of [234].

DEFINITION B.80 *$u \in C_b([0,T] \times H)$ is a classical solution of (B.30) in $D(\hat{\mathcal{A}}_0)$ if*

$$\begin{cases} u(\cdot,x) \in C^1([0,T]), \quad \forall x \in H \\ u(t,\cdot) \in D(\hat{\mathcal{A}}_0) \text{ for any } t \in [0,T] \text{ and } \sup_{t\in[0,T]} \|u(t,\cdot)\|_{D(\hat{\mathcal{A}}_0)} < +\infty \\ Du,\ D^2u,\ A^*Du,\ \hat{\mathcal{A}}_0 u \in C_b([0,T] \times H) \end{cases} \tag{B.36}$$

and u satisfies, for every $x \in H$, equation (B.30).

The second definition is similar to Definition 4.87 and uses the operator $\mathcal{A}_0$.

DEFINITION B.81 *$u \in C_b([0,T] \times H)$ is a classical solution of (B.30) in $D(\mathcal{A}_0)$ if*

$$\begin{cases} u(\cdot,x) \in C^1([0,T]), \quad \forall x \in H \\ u(t,\cdot) \in D(\mathcal{A}_0) \text{ for any } t \in [0,T] \text{ and } \sup_{t\in[0,T]} \|u(t,\cdot)\|_{D(\mathcal{A}_0)} < +\infty \\ Du,\ D^2u,\ A^*Du,\ \in C_b([0,T] \times H) \end{cases} \tag{B.37}$$

and u satisfies, for every $x \in H$, equation (B.30).

REMARK B.82 Concerning the above definitions we observe the following.

(1) If we compare these definitions with the one given in Section 6.2 of [129], we see that the classical solutions in the sense of both Definitions B.80 and B.81 are more regular. Indeed the goal here is not (as in Section 6.2 of [129]) to prove that mild solutions are classical, but to approximate mild solutions with classical solutions to which we can apply Itô's/Dynkin formula (see Section 1.7); hence we want the approximating solutions to be as regular as possible.

(2) The definitions above are the same no matter if we work in the spaces $C_m(H)$ (as we mainly do) or $UC_m(H)$, $(m \geq 0)$.

(3) Note that, if u is a classical solution from Definition B.80, then $\hat{\mathcal{A}}_0 u$, and consequently also u_t, belong to $C_b([0,T] \times H)$. On the other hand, if u is a classical solution from Definition B.81, then $\hat{\mathcal{A}}_0 u$, and then also u_t, only belong to $C_1([0,T] \times H)$.

(4) If we compare Definition B.80 with Definition 4.6 of [233] and Definition 4.1 of [234] we see that, in the last two, the functions $u, u_t, Du, D^2u, A^*Du, \hat{\mathcal{A}}_0 u$ are required to belong to $UC_b^x([0,T] \times H)$. The reason is that elements of this space possess, roughly speaking, the maximal joint regularity one can expect even when $f = 0, C = 0$ and $\varphi \in UC_b^\infty(H)$. Indeed in such case the mild solution is $u(t,x) = R_{T-t}\varphi(x)$ (where R_t is the Ornstein-Uhlenbeck semigroup corresponding to A and Q) which is in $UC_b^x([0,T] \times H)$ but not in $UC_b([0,T] \times H)$, see Subsection B.7.2. Here we decided to ask for less joint regularity, i.e. $u, Du, D^2u, A^*Du \in C_b([0,T] \times H)$ since this last space is more commonly used in the literature on PDEs in infinite dimensions and

[8] The definition of the domain of the operator studied in this paper is slightly different but the arguments there can be adapted easily.

since all the results we need, in particular Itô's/Dynkin formula, still hold.

(5) The definitions implicitly require regularity of data. Indeed the second requirement of (B.36) implies that $\varphi \in D(\hat{\mathcal{A}}_0)$ (and similarly for (B.37)) while, from the last requirement we see that necessarily $f \in C_b([0,T] \times H)$.

■

We now pass to the definition of a strong solution i.e. approximation of classical solutions. It is substantially a variation of Definition 4.7 of [233] (see also Definition 4.3 of [234]) in the sense that we use, as underlying space, the space C_m instead of the space UC_b. The definition below is a special case of Definition 4.89 and uses some definitions given in Chapter 4 which, for the reader's convenience, we repeat in the forms needed here.

We recall first the definition of the class of weights $\mathcal{I}_1$ given in (4.17):

$$\mathcal{I}_1 := \Big\{\eta(\cdot) : (0,+\infty) \to (0,+\infty) \text{ integrable on } (0,T) \text{ for all } T > 0\Big\}. \tag{B.38}$$

Then we recall (see (4.21)) that, given $\eta \in \mathcal{I}_1$, a function $f : (0,T] \times H \to Z$ (where Z is a given real separable Hilbert space) belongs to $C_{m,\eta}((0,T] \times H; Z)$ if

$$f \in C_m([\tau, T] \times H; Z) \ \ \forall \tau \in (0,T) \text{ and } \eta^{-1} f \in C_m((0,T] \times H; Z). \tag{B.39}$$

When $Z = \mathbb{R}$, we omit it as usual, simply writing $C_{m,\eta}((0,T] \times H)$. Furthermore we give a variant of the definition of $\mathcal{K}$-convergence (see Definition 4.88).

Definition B.83 *Let $m \geq 0$, $\eta \in \mathcal{I}_1$ and let Z be a real Hilbert space. We say that a sequence $\{f_n\}_{n\in\mathbb{N}} \subseteq C_{m,\eta}((0,T] \times H; Z)$ $\mathcal{K}$-converges to $f \in C_{m,\eta}((0,T] \times H; Z)$ if*

$$\begin{cases} \sup_{n\in\mathbb{N}} \|f_n\|_{C_{m,\eta}((0,T]\times H;Z)} < +\infty, \\ \lim_{n\to+\infty} \sup_{(t,x)\in(0,T]\times K} \eta(t)^{-1}|f_n(t,x) - f(t,x)| = 0, \end{cases} \tag{B.40}$$

for every compact set $K \subset H$. In such case we write $\mathcal{K}$-$\lim_{n\to+\infty} f_n = f$ in $C_{m,\eta}((0,T] \times H; Z)$.

Below are the definitions of $\mathcal{K}$-strong solutions.

Definition B.84 *Let $m \geq 0$ and $\eta \in \mathcal{I}_1$. Let $\varphi \in C_m(H)$ and $f \in C_{m,\eta}((0,T] \times H)$. We say that a function $u \in C_m([0,T] \times H)$ is a $\mathcal{K}$-strong solution in $D(\hat{\mathcal{A}}_0)$ of (B.30) if $u(t,\cdot)$ is Fréchet differentiable for any $t \in (0,T]$ and there exist three sequences $\{u_n\} \subset C_b([0,T] \times H)$, $\{\varphi_n\} \subset D(\hat{\mathcal{A}}_0)$, $\{f_n\} \subset C_{m,\eta}((0,T] \times H)$ such that:*

(i) *For every $n \in \mathbb{N}$, u_n is a classical solution in $D(\hat{\mathcal{A}}_0)$ (from Definition B.80) of*

$$\begin{cases} w_t + \hat{\mathcal{A}}_0 w + Cw + f_n = 0 \\ w(T) = \varphi_n. \end{cases} \tag{B.41}$$

(ii) *The following limits hold*

$$\begin{cases} \mathcal{K}\text{-}\lim_{n\to+\infty} \varphi_n = \varphi & \text{in } \ C_m(H) \\ \mathcal{K}\text{-}\lim_{n\to+\infty} u_n = u & \text{in } \ C_m([0,T] \times H) \end{cases}$$

and, for some $\eta_1 \in \mathcal{I}_1$ (possibly different from η),

$$\begin{cases} \mathcal{K}\text{-}\lim_{n\to+\infty} f_n = f & \text{in } \ C_{m,\eta}((0,T] \times H) \\ \mathcal{K}\text{-}\lim_{n\to+\infty} Du_n = Du & \text{in } \ C_{m,\eta_1}((0,T] \times H; H). \end{cases}$$

Finally we say that a function $u \in C_m([0,T] \times H)$ is a $\mathcal{K}$-strong solution in $D(\mathcal{A}_0)$ of (B.30) if all the above holds substituting $D(\hat{\mathcal{A}}_0)$ with $D(\mathcal{A}_0)$.

REMARK B.85 The spirit of this definition is substantially the one of the so-called strong solutions in Friedrichs sense (terminology dating back to [200]) for abstract Cauchy problems, see e.g. Definition 4.1.1 of [323]. It is useful to connect the concept of mild solution with the one of classical solution proving that mild solutions are indeed strong solutions. We remark two key points here: first, the convergences in (ii) are asked to hold in the $\mathcal{K}$-sense, and second, the convergence of the derivatives is also required. The reason for the former is that uniform convergence in the whole H would be a requirement which is too strong as $UC_b^2(H)$ (and so $D(\mathcal{A}_0)$) is not dense in $UC_b(H)$ when H is infinite dimensional, see [354] and also [387], Remark 2.2.11. On the other hand the use of π-convergence (as in [389][Section 4] or [387][Chapter 7]) is possible but in the cases we treat here it would give weaker results, as $\mathcal{K}$-convergence can always be proved when data are continuous. The latter is motivated by the use of this definition for applications to HJB equations (as in Chapter 4) which are written in the mild form of (B.31) with f depending on u and Du (see e.g. (4.1) and (4.5)). Hence convergence of the derivatives allows to pass to the limit in such a mild form. ■

B.7.2. The Ornstein-Uhlenbeck semigroup associated to (B.30), the mild solutions and their properties. Assume that Hypothesis B.78 holds and take a generalized reference probability space $\mu = (\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s\geq 0}, \mathbb{P}, W)$, where W is a Wiener process on H with covariance operator I (the identity).

Consider the following linear stochastic equation on H

$$\begin{cases} dX(t) = AX(t)\,dt + \sqrt{Q}\,dW(t) \\ X(0) = x. \end{cases} \tag{B.42}$$

Under Hypothesis B.78 the problem (B.42) has an H-valued mild solution from Definition 1.113 (see Theorem 1.144). Such solution is mean square continuous (see Proposition 1.138 or also [130][Theorem 5.2-(i)]) and is denoted by $X(t,x)$. We denote by $\{R_t, t \geq 0\}$ or simply by R_t, the corresponding transition semigroup (see Section 1.6) on $B_m(H)$:

$$R_t\varphi(x) = \mathbb{E}\left[\varphi(X(t,x))\right] = \int_H \varphi(y)\,\mathcal{N}(e^{tA}x, Q_t)(dy) = \int_H \varphi(y+e^{tA}x)\,\mathcal{N}(0,Q_t)(dy), \tag{B.43}$$

where $\mathcal{N}(z,Q)$ is the Gaussian measure introduced in Definition 1.57. See Subsections 4.3.1.1 and 4.3.1.2 for more on this.

We immediately see that, when $\varphi \in B_m(H)$ and $f \in B_m([0,T] \times H)$, the function u in (B.31) is well defined. Thus we can give the precise definition of the mild solution of equation (B.30).

DEFINITION B.86 *Given $\varphi \in B_m(H)$ and $f \in B_m([0,T] \times H)$, the function u given by (B.31) is well defined on $[0,T]\times H$ and is called the mild solution of (B.30).*

We now explain the connection between equation (B.30) and the notion of the mild solution. We start with the following result.

PROPOSITION B.87 *Under Hypothesis B.78, $\{R_t, t \geq 0\}$, is a $\mathcal{K}$-continuous semigroup and a π-continuous semigroup in $C_m(H)$ and $UC_m(H)$ ($m \geq 0$). Moreover it is not strongly continuous in these spaces unless $A = 0$. Finally, as $t \to 0^+$, we have $R_t\phi \to \phi$ in $UC_b(H)$ if and only if $\phi(e^{tA}\cdot) \to \phi(\cdot)$ in $UC_b(H)$.*

PROOF. For the $\mathcal{K}$-continuity, see [75], Proposition 6.2, and [387][Section 6.3.3], when $m = 0$. Moreover, see [76] for the $UC_m(H)$ case and Theorem 4.1 of

[225] for the case of $C_m(H)$ (recalling the change in the definition mentioned in Remark B.66).

For the π-continuity, see [387][Section 6.3.3], and [386][p. 293-294]: there the proof is done in $UC_b(H)$ but it can be easily extended to the cases of $C_m(H)$ and $UC_m(H)$.

Regarding the non-strong continuity, see [75][Example 6.1] and [387][Section 6.3.3]. Finally, for the last statement, one can see [129], Proposition 6.3.1. □

From now on $\mathcal{A}^m$ will denote the generator of the transition semigroup R_t as a $\mathcal{K}$-continuous semigroup in $C_m(H)$ (or $UC_m(H)$ when specified), as defined in the previous section. We first give a useful lemma.

LEMMA B.88 *Let Hypothesis B.78 be satisfied.*

(i) *For all $t \geq 0$, the operator R_t maps $D(\mathcal{A}_0)$ into itself, and for all $T > 0$ there exists a constant $L_T > 0$ such that*

$$\|R_t[\phi]\|_{D(\mathcal{A}_0)} \leq L_T \|\phi\|_{D(\mathcal{A}_0)}, \quad \forall\, t \in [0,T], \ \phi \in D(\mathcal{A}_0). \tag{B.44}$$

The same holds for $D(\hat{\mathcal{A}}_0)$.

(ii) *For all $m \geq 1$ we have $\mathcal{A}_0 \subseteq \mathcal{A}^m$ and, for any $\phi \in D(\mathcal{A}_0)$ we have*

$$\frac{d}{dt}(R_t\phi) = \mathcal{A}_0 R_t \phi = R_t \mathcal{A}_0 \phi, \quad t \geq 0. \tag{B.45}$$

Similarly, for all $m \geq 0$ $\hat{\mathcal{A}}_0 \subseteq \mathcal{A}^m$ and, for any $\phi \in D(\hat{\mathcal{A}}_0)$ we have

$$\frac{d}{dt}(R_t\phi) = \hat{\mathcal{A}}_0 R_t \phi = R_t \hat{\mathcal{A}}_0 \phi, \quad t \geq 0. \tag{B.46}$$

(iii) *Let $T > 0$. For every $f \in C_b([0,T] \times H)$ such that, for all $t \in [0,T]$, $f(t,\cdot) \in D(\mathcal{A}_0)$ and*

$$\|f(t,\cdot)\|_{D(\mathcal{A}_0)} \leq g_0(t)$$

for a suitable $g_0 \in L^1(0,T;\mathbb{R}^+)$, let us set

$$g(t,x) = \int_0^t R_{t-s}[f(s,\cdot)](x)\, ds, \quad \forall\, x \in H. \tag{B.47}$$

Then $g \in C_b([0,T] \times H)$ and for every $t \in [0,T]$, $g(t,\cdot) \in D(\mathcal{A}_0)$. Moreover

$$\|g(t,\cdot)\|_{D(\mathcal{A}_0)} \leq L_T \|g_0\|_{L^1}, \quad \forall\, t \in [0,T]. \tag{B.48}$$

The same holds if we replace $\mathcal{A}_0$ by $\hat{\mathcal{A}}_0$.

(iv) *Let $m \geq 0$ and let $\lambda > m(\omega \vee 0)$, where ω is given in Hypothesis B.78. Then $(\lambda I - \mathcal{A}^m)^{-1}$ exists and it is a bounded operator. Moreover it maps $D(\mathcal{A}_0)$ into itself. The same is true for $D(\hat{\mathcal{A}}_0)$.*

PROOF. *Proof of (i).*

If $\phi \in D(\mathcal{A}_0)$ we have, using the last equality of (B.43) and straightforward computations, for $x, h, k \in H$,

$$\langle DR_t[\phi](x), h\rangle = \int_H \left\langle D\phi(y + e^{tA}x), e^{tA}h\right\rangle \mathcal{N}(0,Q_t)(dy) = R_t[\left\langle D\phi(\cdot), e^{tA}h\right\rangle], \tag{B.49}$$

$$\langle A^*DR_t[\phi](x), h\rangle = \int_H \left\langle A^*D\phi(y + e^{tA}x), e^{tA}h\right\rangle \mathcal{N}(0,Q_t)(dy) = R_t[\left\langle A^*D\phi(\cdot), e^{tA}h\right\rangle], \tag{B.50}$$

$$\left\langle D^2R_t[\phi](x)h, k\right\rangle = \int_H \left\langle D^2\phi(y + e^{tA}x)e^{tA}k, e^{tA}h\right\rangle \mathcal{N}(0,Q_t)(dy) = R_t[\left\langle D^2\phi(\cdot)e^{tA}k, e^{tA}h\right\rangle]. \tag{B.51}$$

and, for a given orthonormal basis $\{e_n\}$,

$$\operatorname{Tr}\left(D^2 R_t[\phi](x)\right) = \sum_{n=0}^{+\infty} \left\langle D^2 R_t[\phi](x)e_n, e_n \right\rangle = \int_H \sum_{n=0}^{+\infty} \left\langle D^2\phi(y+e^{tA}x)e^{tA}e_n, e^{tA}e_n \right\rangle \mathcal{N}(0,Q_t)(dy). \tag{B.52}$$

Hence by simple computations, we get that $R_t[\phi] \in D(\mathcal{A}_0)$ and

$$\|R_t[\phi]\|_{D(\mathcal{A}_0)} \le \|\phi\|_0 + Me^{\omega t}[\|D\phi\|_0 + \|A^* D\phi\|_0] + M^2 e^{2\omega t}\|D\phi\|_{\mathcal{L}_1(H)},$$

which gives the claim. In the case of $D(\hat{\mathcal{A}}_0)$ we also need to estimate the term $\langle A^* DR_t[\phi](x), x\rangle$. We have, thanks to (B.50),

$$\begin{aligned} \langle A^* DR_t[\phi](x), x\rangle &= \int_H \left\langle A^* D\phi(y+e^{tA}x), y+e^{tA}x \right\rangle \mathcal{N}(0,Q_t)(dy) \\ &- \int_H \left\langle A^* D\phi(y+e^{tA}x), y \right\rangle \mathcal{N}(0,Q_t)(dy). \end{aligned}$$

Since, by the Hölder inequality and by Proposition 1.58,

$$\left| \int_H \left\langle A^* D\phi(y+e^{tA}x), y \right\rangle \mathcal{N}(0,Q_t)(dy) \right| \le \|A^* D\phi\|_0 (\operatorname{Tr}(Q_t))^{1/2},$$

we obtain

$$\sup_{x\in H} | < A^* DR_t[\phi](x), x > | \le \sup_{x\in H} | < A^* D\phi(x), x > | + \|A^* D\phi\|_0 (\operatorname{Tr}(Q_t))^{1/2}.$$

which gives the claim.

Proof of (ii).

From Dynkin's formula of Proposition 1.159 we easily get that (here $X(s,x)$ is the solution of (B.42))

$$\frac{R_h[\phi](x) - \phi(x)}{t} = \frac{1}{h}\mathbb{E}\left(\phi(X(h,x)) - \phi(x)\right) =$$

$$= \frac{1}{h}\mathbb{E}\int_0^h \left(\frac{1}{2}\operatorname{Tr}\left[QD^2\phi(X(s,x))\right] + \langle A^* D\phi(X(s,x)), X(s,x)\rangle\right) ds.$$

It thus follows from the dominated convergence theorem that

$$\lim_{h\to 0} \frac{R_h[\phi](x) - \phi(x)}{t} = \mathcal{A}_0\phi(x)$$

which, by the definition of generator (Definition B.70), implies $\mathcal{A}_0 \subset \mathcal{A}^m$ and $\hat{\mathcal{A}}_0 \subset \mathcal{A}^m$, too. Now equations (B.45) and (B.46) follow immediately from Proposition B.73-(iii) and from point (i) of this proposition.

Proof of (iii).[9] Since $f \in C_b([0,T]\times H)$, using the first equality in (B.43) and the mean square continuity of $X(t,x)$, one can prove exactly as in Proposition 4.38 -(ii), that also $g \in C_b([0,T]\times H)$. Moreover, since $f(t,\cdot) \in D(\mathcal{A}_0)$, for all $t\in[0,T]$, by part (i) it follows that, for $0\le s\le t\le T$, $R_{t-s}[f(s,\cdot)] \in D(\mathcal{A}_0)$ and

$$\|R_{t-s}[f(s,\cdot)]\|_{D(\mathcal{A}_0)} \le L_T\|f(s,\cdot)\|_{D(\mathcal{A}_0)} \le L_T g_0(s).$$

Now, using that the derivative operator is closed, we get that $g(t,\cdot)$ is Fréchet differentiable and, for $t\in[0,T]$, $x,h\in H$,

$$\langle Dg(t,x), h\rangle = \int_0^t \langle DR_{t-s}[f(s,\cdot)](x), h\rangle\, ds = \int_0^t R_{t-s}\left[\left\langle Df(s,\cdot), e^{(t-s)A}h\right\rangle\right](x) ds, \tag{B.53}$$

where we exploited (B.49) in the last equality. The integrand is Borel measurable in s since it is the limit of the difference quotients; the continuity of $Dg(t,x)$ is proved

[9] This proof is partially similar to the proof of Lemma 4.5 of [389].

again as in Proposition 4.38-(ii). Similarly, using (B.50) we get for $t \in [0,T]$, $x, h \in H$,

$$\langle A^* Dg(t,x), h\rangle = \int_0^t \langle A^* DR_{t-s}[f(s,\cdot)](x), h\rangle \, ds = \int_0^t R_{t-s}\left[\left\langle A^* Df(s,\cdot), e^{(t-s)A}h\right\rangle\right](x) ds. \tag{B.54}$$

Moreover, iterating the argument and exploiting (B.49) in the last equality, we prove that $g(t,\cdot)$ is twice Fréchet differentiable and, for $t \in [0,T]$, $x, h, k \in H$,

$$\left\langle D^2 g(t,x)h, k\right\rangle = \int_0^t \left\langle D^2 R_{t-s}[f(s,\cdot)](x)h, k\right\rangle ds = \int_0^t R_{t-s}\left[\left\langle D^2 f(s,\cdot) e^{(t-s)A}h, e^{(t-s)A}k\right\rangle\right](x) ds. \tag{B.55}$$

The integrand is Borel measurable in s since it is the limit of the difference quotients and the continuity of $D^2 g(t,x)h$ for each h is proved again as in Proposition 4.38-(ii). This does not even imply the measurability of $Dg(t,x)$ in general, but it is enough to prove the continuity of the trace. Indeed

$$\mathrm{Tr}(D^2 g(t,x)) = \sum_{n=0}^{+\infty} \left\langle D^2 g(t,x) e_n, e_n\right\rangle = \tag{B.56}$$

$$\int_0^t \int_H \sum_{n=0}^{+\infty} \left\langle D^2 f(s, y + e^{tA}x) e^{(t-s)A} e_n, e^{(t-s)A} e_n\right\rangle \mathcal{N}(0, Q_{t-s})(dy) ds \tag{B.57}$$

As before the integrand is Borel measurable in (s,y), hence continuity can be proved as in Proposition 4.38-(ii).

In the case of $D(\hat{\mathcal{A}}_0)$ we also need to estimate the term $\langle A^* Dg(t,x), x\rangle$. We have, thanks to (B.54),

$$\begin{aligned}
\langle A^* Dg(t,x), x\rangle &= \int_0^t \langle A^* DR_{t-s}[f(s,\cdot)](x), x\rangle \, ds = \int_0^t R_{t-s}\left[\left\langle A^* Df(s,\cdot), e^{(t-s)A}x\right\rangle\right](x) ds \\
&= \int_0^t \int_H \left\langle A^* Df(s, y + e^{(t-s)A}x), y + e^{(t-s)A}x\right\rangle \mathcal{N}(0, Q_{t-s})(dy) ds \\
&- \int_0^t \int_H \left\langle A^* D\phi(y + e^{(t-s)A}x), y\right\rangle \mathcal{N}(0, Q_{t-s})(dy) ds.
\end{aligned}$$

The conclusion now follows as in part (i) using Hölder's inequality and Proposition 1.58.

Proof of (iv). Recall first that, for any $\phi \in C_m(H)$, $x \in H$, the function $t \to e^{-\lambda t} R_t[\phi](x)$ is integrable when $\lambda > m(\omega \vee 0)$,due to the first estimate of Theorem 4.32. Thus by Remark B.71, we have that λ is in the resolvent set of $\mathcal{A}^m$ and

$$(\lambda I - \mathcal{A}^m)^{-1}\phi = \int_0^{+\infty} e^{-\lambda t} R_t[\phi](x) dt.$$

Taking $\phi \in D(\mathcal{A}_0)$ the required conclusion follows using the same arguments as these in the proof of part (iii), which here are even easier since we do not have the time dependency of the integral. □

We now pass to the following result.

Proposition B.89 *Let Hypothesis B.78 be satisfied and let $T > 0$.*

(i) *Let $\varphi \in D(\mathcal{A}_0)$ and $f \in C_b([0,T] \times H)$ such that, for all $t \in [0,T]$, $f(t,\cdot) \in D(\mathcal{A}_0)$ and*

$$\|f(t,\cdot)\|_{D(\mathcal{A}_0)} \le g_0(t)$$

for a suitable $g_0 \in L^1(0,T;\mathbb{R}^+)$. Then the function u defined in (B.31) is a classical solution in $D(\mathcal{A}_0)$ of (B.30).

(ii) *If the assumption of point (i) hold with $D(\hat{\mathcal{A}}_0)$ in place of $D(\mathcal{A}_0)$ then the function u defined in (B.31) is a classical solution in $D(\hat{\mathcal{A}}_0)$ of (B.30).*

PROOF. We take $C = 0$ as the case $C \neq 0$ can be obtained by a straightforward change of variable. We start proving (i). As a first step we need to prove that u satisfies the regularity required in (B.37). Indeed, by Lemma B.88, in particular (B.44) and (B.48), it immediately follows that the second line of (B.37) is true. The third line of (B.37) is obtained, for the first term in (B.31), from (B.49), (B.50), (B.51), (B.52), and for the convolution term, from (B.53), (B.54), (B.55), (B.56). Hence continuity is immediately deduced. The first line of (B.37) follows from the next computation which also shows that u satisfies (B.30).

We apply Dynkin's formula of Proposition 1.159 to the process $\varphi(X(s,x))$ getting, on the interval $[0,t]$,

$$\mathbb{E}\varphi(X(t,x)) = \varphi(x) + \mathbb{E}\int_t^T \langle A^* D\varphi(X(s,x)), X(s,x)\rangle\, ds \\ + \frac{1}{2}\mathrm{Tr}\left[QD^2\varphi(X(s,x))\right]\Big] ds. \quad \text{(B.58)}$$

Now we compute the left derivative of u at $t = 0$: first we observe that, by the definition of u,

$$\frac{u(h,x) - u(0,x)}{h} = \frac{R_h[\varphi](x) - \varphi(x)}{h} + \frac{1}{h}\int_0^h R_{h-s}[f(s,\cdot)](x) ds.$$

By the continuity of the integrands in (B.58) we obtain that, when $h \searrow 0$, the first term of the above right hand side converges to

$$\langle A^* D\varphi(x), x\rangle + \frac{1}{2}\mathrm{Tr}[QD^2\varphi(x)]$$

Similarly by the continuity of the integrand in the second term of the right hand side we get that this term, when $h \searrow 0$, converges to $f(0,x)$. We then have, denoting by D_t^+ the right time derivative,

$$D_t^+ u(0,x) = \lim_{h\searrow 0} \frac{u(h,x) - u(0,x)}{h} = \\ = \langle A^* D\varphi(x), x\rangle + \frac{1}{2}\mathrm{Tr}[QD^2\varphi(x)] + f(0,x)$$

so the equation is satisfied for $t = 0$. For $t > 0$ we observe that, by the semigroup property (see Theorem 1.152),

$$u(t+h,x) = R_h[u(t,\cdot)](x) + \int_t^{t+h} R_{t+h-s}[f(s,\cdot)](x) ds,$$

hence we have

$$\frac{u(t+h,x) - u(t,x)}{h} = \frac{R_h[u(t,\cdot)](x) - u(t,x)}{h} + \frac{1}{h}\int_t^{t+h} R_{t+h-s}[f(s,\cdot)](x) ds$$

and, arguing as for the case $t = 0$ but replacing φ by $u(t,\cdot)$, we get

$$D_t^+ u(t,x) = \langle x, A^* Du(t,x)\rangle + \frac{1}{2}\mathrm{Tr}\left[QD^2 u(t,x)\right] + f(t,x).$$

Now the right hand side of the above identity is a continuous function on $[0,T]\times H$ and consequently, by Lemma 3.2.4 of [129], $u_n(\cdot,x)$ is continuously differentiable and satisfies equation (B.30).

The proof of part (ii) is almost exactly the same. The only difference is that, thanks to the regularity of the data, also $\hat{\mathcal{A}}_0 u$ and u_t are bounded. $\square$

B.7.3. The approximation results. We start with the following result about operators $\mathcal{A}_0$, $\hat{\mathcal{A}}_0$ and $\mathcal{A}^m$ (recall that ω below is the one given by Hypothesis B.78).

PROPOSITION B.90 *Let Hypothesis B.78 hold. We have the following:*

(i) *For any $m \geq 1$, we have $\mathcal{A}_0 \subseteq \mathcal{A}^m$, and $D(\mathcal{A}_0)$ is a $\mathcal{K}$-core for $\mathcal{A}^m$ in $C_m(H)$. In particular $\mathcal{A}_0$ is $\mathcal{K}$-closable in $C_m(H)$ and its $\mathcal{K}$-closure $\overline{\mathcal{A}_0}^{\mathcal{K}}$ coincides with $\mathcal{A}^m$.*
Moreover, for all $m \geq 0$ and $\lambda > m(\omega \vee 0)$, the set

$$(\lambda I - \mathcal{A}^m)^{-1} \left(\mathcal{F}C_0^{\infty,A^*}(H)\right) := \left\{\phi \in C_b(H) \ : \ (\lambda I - \mathcal{A}^m)\phi \in \mathcal{F}C_0^{\infty,A^*}(H)\right\}$$

is contained in $D(\mathcal{A}_0)$ and it is always a core for $\mathcal{A}^m$.
Finally, the set $\mathcal{F}C_0^{\infty,A^*}(H) \subseteq D(\mathcal{A}_0)$ is always a core for $\mathcal{A}^m$ when $m \geq 1$, but not, in general, when $m \in [0, 1)$.

(ii) *For any $m \geq 0$, we have $\hat{\mathcal{A}}_0 \subseteq \mathcal{A}^m$, $D(\hat{\mathcal{A}}_0)$ is a $\mathcal{K}$-core for $\mathcal{A}^m$ in $C_m(H)$ In particular $\hat{\mathcal{A}}_0$ is $\mathcal{K}$-closable in $C_m(H)$ and its $\mathcal{K}$-closure $\overline{\hat{\mathcal{A}}_0}^{\mathcal{K}}$ coincides with $\mathcal{A}^m$.*
Moreover, for all $m \geq 0$ and $\lambda > m(\omega \vee 0)$, the set $(\lambda I - \mathcal{A}^m)^{-1} \left(\mathcal{F}C_0^{\infty,A^*}(H)\right)$ is contained in $D(\hat{\mathcal{A}}_0)$ and it is always a core for $\mathcal{A}^m$ when $m \geq 0$.
Finally the set $\mathcal{F}C_0^{\infty,A^*}(H)$ is in general not contained in $D(\hat{\mathcal{A}}_0)$.

PROOF. *Proof of part (i).*
For any $m \geq 1$, we have $\mathcal{A}_0 \subseteq \mathcal{A}^m$ by Lemma B.88-(ii). To prove that $D(\mathcal{A}_0)$ is a $\mathcal{K}$-core for $\mathcal{A}^m$ in $C_m(H)$ we take any $\phi \in D(\mathcal{A}^m)$ and set, for some $\lambda > m(\omega \vee 0)$ $g := (\lambda I - \mathcal{A}^m)\phi \in C_m(H)$. Then we consider the approximating sequence $\{g_n\}$ in $\mathcal{F}C_0^{\infty,A^*}(H) \subseteq D(\mathcal{A}_0)$ given by Lemma B.77. By construction we have $g_n \xrightarrow{\mathcal{K}} g$. Define $\phi_n := (\lambda I - \mathcal{A}^m)^{-1} g_n$. By Lemma B.88-(iv) we have $\phi_n \in D(\mathcal{A}_0)$ and, by the dominated convergence theorem, we get $\phi_n \xrightarrow{\mathcal{K}} \phi$. Moreover

$$\mathcal{A}_0\phi_n = \mathcal{A}^m\phi_n = \lambda\phi_n - (\lambda I - \mathcal{A}^m)\phi_n = \lambda\phi_n - g_n,$$

hence also $\mathcal{A}_0\phi_n \xrightarrow{\mathcal{K}} \mathcal{A}^m\phi$.

Clearly $\mathcal{F}C_0^{\infty,A^*}(H) \subseteq D(\mathcal{A}_0)$ and, since $(\lambda I - \mathcal{A}^m)(D(\mathcal{A}_0)) \subseteq D(\mathcal{A}_0)$ (Lemma B.88- (iv)), then it must be

$$(\lambda I - \mathcal{A}^m)^{-1} \left(\mathcal{F}C_0^{\infty,A^*}(H)\right) \subseteq D(\mathcal{A}_0).$$

The choice of the sequence above implies that such set is a core for $\mathcal{A}^m$.

The fact that the set $\mathcal{F}C_0^{\infty,A^*}(H)$ is a core for $\mathcal{A}^m$ when $m \geq 1$ follows using Lemma 4.4 in [225]. This lemma generalizes a well known result about cores (see e.g. [160][Proposition 1.7, p.53] implying that a $\mathcal{K}$-dense subspace $\mathcal{D}$ of $C_m(H)$ which is invariant for R_t is always a core. Indeed let $\phi \in \mathcal{F}C_0^{\infty,A^*}(H)$ and let $f : \mathbb{R}^n \to \mathbb{R}$ and $x_1, \dots x_n \in D(A^*)$ be such that

$$\phi(x) = f(\langle x, x_1\rangle, \dots, \langle x, x_n\rangle).$$

Then

$$\begin{aligned}
R_t[\phi](x) &= \int_H \phi(y + e^{tA}x)\mathcal{N}(0,Q_t)(dy) \\
&= \int_H f\left(\left\langle y + e^{tA}x, x_1\right\rangle, \dots, \left\langle y + e^{tA}x, x_n\right\rangle\right)\mathcal{N}(0,Q_t)(dy) \\
&= \int_H f\left(\langle y, x_1\rangle + \left\langle x, e^{tA^*}x_1\right\rangle, \dots, \langle y, x_n\rangle + \left\langle x, e^{tA^*}x_n\right\rangle\right)\mathcal{N}(0,Q_t)(dy) \\
&= g\left(\left\langle x, e^{tA^*}x_1\right\rangle, \dots, \left\langle x, e^{tA^*}x_n\right\rangle\right)
\end{aligned}$$

for a suitable function $g \in \mathcal{F}C_0^{\infty,A^*}(H)$. Since $x_1, \dots x_n \in D(A^*)$ we have $e^{tA^*}x_1, \dots e^{tA^*}x_n \in D(A^*)$ and the claim is proved.

However when $m \in [0,1)$ it is not true in general that, for $\phi \in \mathcal{F}C_0^{\infty,A^*}(H)$, one has $\mathcal{A}_0\phi \in C_m(H)$ so the core property in general fails contrary to what is stated in[225][Theorem 4.5] (see on this [392][Remark 5.11]).

The proof of part (ii) is exactly the same except for the fact that, since in general $\mathcal{F}C_0^{\infty,A^*}(H) \not\subseteq D(\hat{\mathcal{A}}_0)$, we need to prove directly that

$$(\lambda I - \mathcal{A}^m)^{-1}\left(\mathcal{F}C_0^{\infty,A^*}(H)\right) \subseteq D(\hat{\mathcal{A}}_0).$$

In fact we can prove more, as in Proposition 4.6 of [389], that

$$(\lambda I - \mathcal{A}^m)^{-1}\left(D(\mathcal{A}_0)\right) \subseteq D(\hat{\mathcal{A}}_0). \tag{B.59}$$

To do this it is enough to show that, given $\phi \in (\lambda I - \mathcal{A}^m)^{-1}\left(D(\mathcal{A}_0)\right)$, the map $x \mapsto \langle A^*D\phi(x), x\rangle$ is bounded, which is equivalent to prove that $\mathcal{A}_0\phi$ is bounded, since the second order term is always bounded when $\phi \in D(\mathcal{A}_0)$. Let $f \in D(\mathcal{A}_0)$ and let $\phi := (\lambda I - \mathcal{A}^m)^{-1}f$. Then $f = (\lambda I - \mathcal{A}^m)\phi$. Since $\phi \in D(\mathcal{A}_0)$ we also have $f = (\lambda I - \mathcal{A}_0)\phi$ which gives $\mathcal{A}_0\phi = \lambda\phi - f$. Since the right hand side is bounded, then also $\mathcal{A}_0\phi$ is bounded. □

We are going to use, in addition to Hypothesis B.78, the following assumption (compare it with Hypotheses 4.22 and 4.25).

HYPOTHESIS B.91 *The operators A and Q satisfy the following:*

(i) *For all $t > 0$ we have $e^{tA}(H) \subseteq Q_t^{1/2}(H)$.*

(ii) *Defining $\Gamma(t) := Q_t^{-1/2}e^{tA}$ we have that the map $t \mapsto \|\Gamma(t)\|$ (which is always decreasing) belongs to $\mathcal{I}_1$.*

REMARK B.92 If Hypothesis B.91 holds, then the semigroup R_t enjoys the smoothing property stated in Theorem 4.32 with the estimate, for all $f \in C_m(H)$,

$$\|DR_tf\|_{C_m(H)} \le C(m)e^{m(\omega\vee 0)t}\|\Gamma(t)\|\,\|f\|_{C_m(H)}. \tag{B.60}$$

for some constant $C(m) \ge 1$. In particular Hypothesis B.91 implies that, for $m \ge 0$, $D(\mathcal{A}^m) \subseteq C_m^1(H)$. Indeed let $\phi \in D(\mathcal{A}^m)$. Then, taking $\lambda > m(\omega \vee 0)$ we must have, for some $f \in C_m(H)$,

$$\phi(x) = (\lambda I - \mathcal{A}^m)^{-1}f(x) = \int_0^{+\infty} e^{-\lambda t}R_tf(x)dt, \qquad x \in H.$$

Estimate (B.60) and the closedness of the derivative operator imply then

$$D\phi(x) = \int_0^{+\infty} e^{-\lambda t}DR_tf(x)dt, \qquad x \in H.$$

This fact also implies that the approximating sequence $\{\phi_n\}$ of elements in $D(\mathcal{A}_0)$ in the proof of the previous proposition can be chosen such that

$$\phi_n \xrightarrow{\mathcal{K}} \phi, \quad \mathcal{A}_0\phi_n \xrightarrow{\mathcal{K}} \mathcal{A}\phi, \quad D\phi_n \xrightarrow{\mathcal{K}} D\phi.$$

■

Here is the final result.

THEOREM B.93 *Let Hypotheses B.78 and B.91 hold. Let* $m \geq 0$ *and* $\eta(t) = \|\Gamma(t)\|$ *for* $t > 0$.

(i) *Let* $\varphi \in C_m(H)$ *and* $f \in C_{m,\eta}((0,T] \times H)$. *Then the mild solution* u *(given by (B.31)) of the Cauchy problem (B.30) is also a* $\mathcal{K}$*-strong solution of (B.30) both in* $D(\mathcal{A}_0)$ *and in* $D(\hat{\mathcal{A}}_0)$.

(ii) *If in addition* $\varphi \in C^1_m(H)$, $f \in C_m([0,T] \times H)$, f *is differentiable in the* x *variable and* $Df \in C_{m,\eta}((0,T] \times H; H)$, *then the approximating sequences* u_n, φ_n, f_n *defining the* $\mathcal{K}$*-strong solution* u *in part (i) (both in* $D(\mathcal{A}_0)$ *and in* $D(\hat{\mathcal{A}}_0)$*) can be chosen such that, for some* $\eta_1 \in \mathcal{I}_1$,

$$\begin{cases} \mathcal{K}\text{-}\lim_{n\to+\infty} D\varphi_n = D\varphi, & \text{in } C_m(H;H), \\ \mathcal{K}\text{-}\lim_{n\to+\infty} Df_n = Df, & \text{in } C_{m,\eta}((0,T] \times H; H), \\ \mathcal{K}\text{-}\lim_{n\to+\infty} Du_n = Du, & \text{in } C_m((0,T] \times H; H), \\ \mathcal{K}\text{-}\lim_{n\to+\infty} D^2u_nh = D^2uh & \text{in } C_{m,\eta_1}((0,T] \times H; H), \quad \forall h \in H. \end{cases}$$

PROOF. A similar result is proved in [233][Proposition 4.10] when $m = 0$. Here we give a complete proof.

Proof of (i). We first approximate φ by a sequence $\{\varphi_n\} \subset \mathcal{F}C_0^{\infty,A^*}(H)$ given by Lemma B.77. Moreover setting $\tilde{f} = \eta^{-1}f$ we approximate $\tilde{f} \in C_m((0,T] \times H)$ by a sequence $\{\tilde{f}_n\} \subset \mathcal{F}C_0^{\infty,A^*}((0,T] \times H)$ given by Lemma B.77. We then define $f_n := \eta\tilde{f}_n$ and set

$$u_n(t,x) := R_t[\varphi_n](x) + \int_0^t R_{t-s}[f_n(s,\cdot)](x)ds \tag{B.61}$$

By construction we know that φ_n and f_n satisfy the assumptions of part (i) of Proposition B.89 and so u_n is a classical solution of (B.30) in $D(\mathcal{A}_0)$. Moreover, still by construction, we have

$$\begin{cases} \mathcal{K}\text{-}\lim_{n\to+\infty} \varphi_n = \varphi, & \text{in } C_m(H), \\ \mathcal{K}\text{-}\lim_{n\to+\infty} f_n = f, & \text{in } C_{m,\eta}((0,T] \times H). \end{cases}$$

By Proposition 4.38 (i) and (ii) we get $u \in C_m([0,T] \times H)$ while using dominated convergence we obtain $\mathcal{K}\text{-}\lim_{n\to+\infty} u_n = u$ in $C_m([0,T] \times H)$. Up to now we did not use Hypothesis B.91. We use it now to show that, as required, $u(t;\cdot)$ is Fréchet differentiable for any $t \in (0,T]$ and that Du_n $\mathcal{K}$-converges to Du. Indeed, using Proposition 4.38, Remark 4.39 and the closedness of the derivative operator we get the required differentiability and that

$$Du(t,x) = DR_t[\varphi](x) + \int_0^t DR_{t-s}[f(s,\cdot)](x)ds.$$

Note that $Du \in C_{m,\eta_1}([0,T] \times H; H)$, where

$$\eta_1(t) = \eta(t) \vee \left(\int_0^t \eta(s)\eta(t-s)ds \right).$$

Finally the required convergence of Du_n to Du in $C_{m,\eta_1}([0,T] \times H; H)$ follows using the representation formula (4.62) for the derivatives of $R_t[\varphi]$ and $R_{t-s}[f(s,\cdot)]$, and then applying the dominated convergence theorem.

The above approximating sequences u_n, φ_n, f_n are not suitable, in general, for u to be a $\mathcal{K}$-strong solution in $D(\hat{\mathcal{A}}_0)$ since the sequences φ_n and $f_n(t,\cdot)$ may not belong to $D(\hat{\mathcal{A}}_0)$, as we discussed in the proof of Proposition B.90. To get an approximating sequence $\bar{u}_n$ for $\mathcal{K}$-strong solution in $D(\hat{\mathcal{A}}_0)$ it is enough to define, as in [389][Section 4],

$$\bar{u}_n(t,x) := J_n u_n(t,x) = R_t[J_n\varphi_n](x) + \int_0^t R_{t-s}[J_n f_n(s,\cdot)](x)ds,$$

where, as in (B.7), we set $J_n := n(nI - \mathcal{A}^m)^{-1}$. By Lemma B.88-(iv) we easily see that $\bar{u}_n(t,\cdot) \in D(\hat{\mathcal{A}}_0)$ for all $t \in [0,T]$. All the other required properties can be proved exactly as we did for u_n.

Proof of (ii). Assume now that, in addition, $\varphi \in C^1_m(H)$, $f \in C_m([0,T] \times H)$, f is differentiable in the x variable and $Df \in C_{m,\eta}((0,T] \times H; H)$. Take sequences $\{\varphi_n\} \subset \mathcal{F}C_0^{\infty,A^*}(H)$ and $\{f_n\} \subset \mathcal{F}C_0^{\infty,A^*}((0,T] \times H)$ given by Lemma B.77 and take u_n given by (B.61). It is not difficult to see, looking at the proof of Lemma B.77, that for such sequences we also have, under the assumptions on φ and f,

$$\begin{cases} \mathcal{K}\text{-}\lim_{n\to+\infty} D\varphi_n = D\varphi, & in\ C_m(H;H), \\ \\ \mathcal{K}\text{-}\lim_{n\to+\infty} Df_n = Df, & in\ C_{m,\eta}((0,T] \times H; H) \end{cases}$$

To prove the regularity of u and the remaining convergences we observe first that, using (B.49) and (B.53), we have, for $t \in (0,T]$, $x, h \in H$,

$$\langle DR_t[\varphi](x), h\rangle = R_t[\langle D\varphi(\cdot), e^{tA}h\rangle], \tag{B.62}$$

$$\left\langle D\int_0^t R_{t-s}[f(s,\cdot)](x)ds, h \right\rangle = \int_0^t R_{t-s}\left[\left\langle Df(s,\cdot), e^{(t-s)A}h\right\rangle\right](x)ds, \tag{B.63}$$

so the required regularity of u follows again using the smoothing property and the estimates of Theorem 4.32. On the other hand the formulae (B.62) and (B.63) also hold for φ_n and f_n, hence the required convergences simply hold by applying the dominated convergence theorem as in the first part of the proof.

The proof of (ii) for the strong solution in $D(\hat{\mathcal{A}}_0)$ follows the same arguments and replacing φ_n by $J_n\varphi_n$ and $f_n(s,\cdot)$ by $J_n f_n(s,\cdot)$, $s \in (0,T]$. □

REMARK B.94 In Definition B.84 (of strong solution), as noted in Remark B.85, we also required convergence of the derivatives. If we did not ask such convergence, then part (i) of Theorem B.93 would be true without using Hypothesis B.91, as noted in the body of the proof. In such a case the first part of Theorem B.93 may be seen as a particular case of a more general result for a class of abstract $\mathcal{K}$-continuous semigroups (see [82] Theorem 4.10 for the case $m = 0$). ■

REMARK B.95 We point out some general observations that can be extracted from the results of this section.

First of all, when one needs to approximate continuous functions in infinite dimensions by C^2 functions through $\mathcal{K}$-convergence, a straightforward way (not the only one, see e.g. [389][Section 4] for an alternative approach) is to use cylindrical functions (as in Lemma B.77). The use of cylindrical functions allows to find approximating sequences required in the definition of strong solutions, as in Theorem B.93, however it is not adequate to find approximations belonging to $D(\hat{\mathcal{A}}_0)$ since the map

$$x \mapsto \langle A^* D\phi(x), x\rangle$$

may be unbounded for $\phi \in \mathcal{F}C_0^{\infty,A^*}(H)$. This also implies, as observed in [392][Remark 5.11], that $\mathcal{F}C_0^{\infty,A^*}(H)$ is in general not a $\mathcal{K}$-core for $\mathcal{A}^m$, $m \in [0, 1)$.

Thus we decided to use separately the two operators $\mathcal{A}_0$ and $\hat{\mathcal{A}}_0$. They have their advantages and drawbacks. Indeed, using $\mathcal{A}_0$ allows to use directly the cylindrical functions to find approximations to the mild solutions of (B.30) but, since the $\mathcal{K}$-closure of $\mathcal{A}_0$ is $\mathcal{A}^m$ only for $m \geq 1$, it is not suitable to look at the generator $\mathcal{A}^m$ for $m \in [0, 1)$ (Proposition B.90). On the other hand, using $\hat{\mathcal{A}}_0$ calls for more complicated approximations but allows to study the generator $\mathcal{A}^m$ for all $m \geq 0$. For these reasons $\mathcal{A}_0$ is used when we want to find strong solutions to HJB equations in Sections 4.5 and 4.7, while $\hat{\mathcal{A}}_0$ is used when we use the generator $\mathcal{A}^m$ in the spaces $C_b(H)$ to study mild solutions to the infinite horizon problem in Subsection 4.6.2. ■

APPENDIX C

Parabolic equations with non-homogeneous boundary conditions

In this section we show how to rewrite some classes of parabolic equations with control and noise on the boundary using the infinite dimensional formalism. We focus on two particular cases: Dirichlet and Neumann boundary conditions. Throughout the section $\mathcal{O}$ will denote a bounded domain (open and connected) in $\mathbb{R}^d$ with regular boundary $\partial\mathcal{O}$.

We begin introducing and recalling some properties of the Dirichlet and Neumann maps.

C.1. Dirichlet and Neumann maps

Consider the Laplace equation with Dirichlet boundary condition

$$\begin{cases} \Delta_\xi y(\xi) = 0, & \xi \in \mathcal{O} \\ y(\xi) = \gamma(\xi), & \xi \in \partial\mathcal{O}. \end{cases} \tag{C.1}$$

THEOREM C.1 *Given $s \geq 0$ and a boundary condition $\gamma \in H^s(\partial\mathcal{O})$, there exists a unique solution $D\gamma \in H^{s+1/2}(\mathcal{O})$ of the problem (C.1). Moreover, for all $s \geq 0$, the operator*

$$\begin{cases} D : H^s(\partial\mathcal{O}) \to H^{s+1/2}(\mathcal{O}) \\ \gamma \mapsto D\gamma \end{cases} \tag{C.2}$$

is continuous.

PROOF. See [313] Theorem 5.4, page 165, Theorem 6.6, page 177 and Theorem 7.3, page 187. □

DEFINITION C.2 *The operator D introduced in (C.2) is called the Dirichlet map.*

PROPOSITION C.3 *Consider the heat equation with zero Dirichlet boundary condition*

$$\begin{cases} \frac{\partial}{\partial s} y(s,\xi) = \Delta_\xi y(s,\xi), & (s,\xi) \in (0,T) \times \mathcal{O} \\ y(s,\xi) = 0, & (s,\xi) \in (0,T) \times \partial\mathcal{O} \\ y(0,\xi) = x(\xi), & \xi \in \mathcal{O}, \end{cases} \tag{C.3}$$

where the initial datum x belongs to $L^2(\mathcal{O})$. For $s \geq 0$ denote by $S_D(s)x$ the solution of (C.3) at time s. Then $S_D(t)$ is a C_0-semigroup on $L^2(\mathcal{O})$. Its generator A_D is given by:

$$\begin{cases} D(A_D) := H_0^1(\mathcal{O}) \cap H^2(\mathcal{O}) \\ A_D\phi := \Delta\phi. \end{cases} \tag{C.4}$$

A_D is maximal dissipative, self-adjoint, invertible, $A_D^{-1} \in \mathcal{L}(L^2(\mathcal{O}))$, and $S_D(t)$ is analytic.

PROOF. This is a standard result. We refer for instance to Theorem 12.40 of [397] for the proof that the C_0-semigroup $S_D(t)$ is analytic and that $A_D^{-1} \in \mathcal{L}(L^2(\mathcal{O}))$ (it is included in the definition of an analytic semigroup given there). $\square$

PROPOSITION C.4 *The Dirichlet map D introduced in (C.2) is continuous as a linear map between the spaces $L^2(\partial\mathcal{O})$ and $D((-A_D)^{1/4-\epsilon})$ for all $\epsilon > 0$:*

$$D : L^2(\partial\mathcal{O}) \to D((-A_D)^{1/4-\epsilon}). \tag{C.5}$$

PROOF. See [308], Section 6.1. $\square$

Similarly we consider the following problem with Neumann boundary condition:

$$\begin{cases} \Delta_\xi y(\xi) = \lambda y(\xi), & \xi \in \mathcal{O} \\ \frac{\partial}{\partial n} y(\xi) = \gamma(\xi), & \xi \in \partial\mathcal{O}, \end{cases} \tag{C.6}$$

where n is the outward unit normal vector and $\frac{\partial}{\partial n}$ is the normal derivative.

THEOREM C.5 *Let $\lambda > 0$. Given any $s \geq 0$ and a boundary condition $\gamma \in H^s(\partial\mathcal{O})$, there exists a unique solution $N_\lambda\gamma \in H^{s+3/2}(\mathcal{O})$ of the problem (C.6). Moreover, for all $s \geq 0$, the operator*

$$\begin{cases} N_\lambda : H^s(\partial\mathcal{O}) \to H^{s+3/2}(\mathcal{O}) \\ \gamma \mapsto N_\lambda\gamma \end{cases} \tag{C.7}$$

is continuous.

PROOF. See [313] Theorem 5.4, page 165, Theorem 6.6, page 177 and Theorem 7.3, page 187. $\square$

DEFINITION C.6 *The operator N_λ introduced in (C.7) is called the Neumann map.*

PROPOSITION C.7 *Consider the following heat equation with zero Neumann boundary condition*

$$\begin{cases} \frac{\partial}{\partial s} y(s,\xi) = \Delta_\xi y(s,\xi), & (s,\xi) \in (0,T) \times \mathcal{O} \\ \frac{\partial}{\partial n} y(s,\xi) = 0, & (s,\xi) \in (0,T) \times \partial\mathcal{O} \\ y(0,\xi) = x(\xi), & \xi \in \mathcal{O} \end{cases} \tag{C.8}$$

where the initial datum x belongs to $L^2(\mathcal{O})$. For $s \geq 0$ denote by $S_N(s)x$ the solution of (C.8) at time s. Then $S_N(t)$ is a C_0-semigroup on $L^2(\mathcal{O})$. Its generator A_N is given by

$$\begin{cases} D(A_N) := \left\{ u \in H^2(\mathcal{O}) \ : \ \frac{\partial u}{\partial n} = 0 \quad \text{on } \partial\mathcal{O} \right\} \\ A_N\phi := \Delta\phi. \end{cases} \tag{C.9}$$

If $\lambda \geq 0$ then $(A_N - \lambda I)$ is maximal dissipative, self-adjoint, and $e^{t(A_N-\lambda I)}$ is analytic. If $\lambda > 0$ then $(A_N - \lambda I)$ is invertible and $(A_N - \lambda I)^{-1} \in \mathcal{L}(L^2(\mathcal{O}))$.

PROOF. For the proof of analyticity, see for instance [2]. $\square$

PROPOSITION C.8 *For any $\epsilon > 0$, the Neumann map N_λ introduced in (C.7) is continuous as a linear map between the spaces $L^2(\partial\mathcal{O})$ and $D((-A_N + \lambda I)^{3/4-\epsilon})$:*

$$N_\lambda : L^2(\partial\mathcal{O}) \to D((-A_N + \lambda I)^{3/4-\epsilon}). \tag{C.10}$$

PROOF. See [308], Section 6.1. $\square$

We remark that, see [308], Section 6.1, in fact we have

$$\begin{aligned} &D((-A_D)^\alpha) = H^{2\alpha}(\mathcal{O}), \quad 0 < \alpha < \frac{1}{4}, \\ &D((-A_D)^\alpha) = H_0^{2\alpha}(\mathcal{O}), \quad 0 < \alpha < \frac{3}{4}, \alpha \neq \frac{1}{4}, \\ &D((-A_N + \lambda I)^\alpha) = H^{2\alpha}(\mathcal{O}), \quad 0 < \alpha < \frac{3}{4}, \lambda > 0. \end{aligned} \tag{C.11}$$

Moreover, for $\alpha \geq 0$,

$$|x|_{H^{2\alpha}(\mathcal{O})} \leq C_\alpha |(-A_D)^\alpha x|, \ x \in D((-A_D)^\alpha), \tag{C.12}$$

$$|x|_{H^{2\alpha}(\mathcal{O})} \leq C_{\alpha,\lambda} |(-A_N + \lambda I)^\alpha x|, \ x \in D((-A_N + \lambda I)^\alpha), \ \lambda > 0. \tag{C.13}$$

We also recall the Sobolev embeddings (see e.g. [1], Theorem 7.57).

THEOREM C.9 *Let $\mathcal{O}$ be a domain in $\mathbb{R}^d$ with C^1 boundary, $s > 0, 1 < p < +\infty$. The following embeddings are continuous:*

- *If $d > sp$ then $W^{s,p}(\mathcal{O}) \hookrightarrow L^q(\mathcal{O})$ for $p \leq q \leq dp/(d - sp)$.*
- *If $d = sp$ then $W^{s,p}(\mathcal{O}) \hookrightarrow L^q(\mathcal{O})$ for $p \leq q < +\infty$.*
- *If $d < (s - j)p$ for some nonnegative integer j then $W^{s,p}(\mathcal{O}) \hookrightarrow C_b^j(\mathcal{O})$.*

For more on Sobolev embeddings we refer to [1], Chapters V and VII.

C.2. Non-zero boundary conditions, the Dirichlet case

In this and in the following subsections, we show how to rewrite some classes of parabolic equations with control and noise on the boundary using the infinite dimensional formalism.

We will always work on the time interval $[t, T]$ where t and T are such that $0 \leq t < T$. We consider the initial datum at time t instead of time 0 to be consistent with the notation we use in Chapter 2.

We consider the following problem:

$$\begin{cases} \frac{\partial}{\partial s} y(s,\xi) = \Delta_\xi y(s,\xi) + f(s,\xi), & (s,\xi) \in (t,T) \times \mathcal{O} \\ y(s,\xi) = \gamma(s,\xi), & (s,\xi) \in (t,T) \times \partial\mathcal{O} \\ y(t,\xi) = x(\xi), & \xi \in \mathcal{O}. \end{cases} \tag{C.14}$$

Until the end of the section, H and Λ denote respectively the Hilbert spaces $L^2(\mathcal{O})$ and $L^2(\partial\mathcal{O})$, and A_D is the operator defined in (C.4). Recall from Theorem C.1 that the Dirichlet map $D : \Lambda \to H$ is continuous.

LEMMA C.10 *Assume that A is the generator of a C_0-semigroup on H. Suppose that $\eta \in W^{1,1}(t,T;H)$, $x \in D(A)$ and*

$$z(s) = e^{(s-t)A}x + \int_t^s e^{(s-r)A}\eta(r)dr, \qquad \text{for } s \in [t,T].$$

Then $z \in C^1([t,T],H) \cap C([t,T],D(A))$ and

$$Az(s) = Ae^{(s-t)A}x + \int_t^s e^{(s-r)A}\frac{d}{dr}\eta(r)dr + e^{(s-t)A}\eta(t) - \eta(s), \qquad \text{for } s \in [t,T].$$

PROOF. See [127] Lemma 13.2.2, Chapter 13, page 242. □

PROPOSITION C.11 *Assume that $y \in C^\infty([t,T] \times \overline{\mathcal{O}})$ is a classical solution of (C.14). Then, denoting $X(s) = y(s,\cdot)$, the solution can be written as*

$$X(s) = e^{(s-t)A_D}x \ - A_D \int_t^s e^{(s-r)A_D} D\gamma(r)dr + \int_t^s e^{(s-r)A_D} f(r)dr. \tag{C.15}$$

PROOF. We follow a well known procedure, see for instance [127] Chapter 13 or [312] Chapter 9, Section 1.1.

Since y is smooth, by classical theory we have that $D\gamma(s)$ is smooth and moreover $\frac{d}{ds}D\gamma(s) = D\frac{d}{ds}\gamma(s)$ for $t < s < T$. In particular the function

$$z(s) := X(s) - D\gamma(s) \qquad \text{for } s \in [t, T] \tag{C.16}$$

is in $C^2((t,T);H) \cap C([t,T];D(A_D))$ and

$$\begin{aligned}\frac{d}{ds}z(s) = \frac{d}{ds}X(s) - \frac{d}{ds}D\gamma(s) &= \Delta_\xi X(s) - \Delta_\xi D\gamma(s) + \Delta_\xi D\gamma(s)\\ &- \frac{d}{ds}D\gamma(s) + f(s) = A_D z(s) - D\frac{d}{ds}\gamma(s) + f(s),\end{aligned}$$

where in the last equality we used that $\Delta_\xi D\gamma(s) = 0$. Observe that the expression $A_D z(s)$ is well defined because $z(s)$ belongs to $D(A_D)$ while neither $X(s)$ nor $D\gamma(s)$ are contained in $D(A_D)$. In particular, $z(\cdot)$ is a strict solution (see [35], Definition 3.1, page 129) of the following evolution equation in H:

$$\begin{cases} \frac{d}{ds}z(s) = A_D z(s) - D\frac{d}{ds}\gamma(s) + f(s) \\ z(t) = x - D\gamma(t). \end{cases} \tag{C.17}$$

The strict solution of (C.17) can be written in the mild form (see e.g. [35] Chapter II.1, Lemma 3.2, page 135)

$$z(s) = e^{(s-t)A_D}\left[x - D\gamma(t)\right] + \int_t^s e^{(s-r)A_D}\left[-D\frac{d}{dr}\gamma(r) + f(r)\right] dr.$$

Therefore

$$\begin{aligned} X(s) &= z(s) + D\gamma(s) \\ &= e^{(s-t)A_D}\left[x - D\gamma(t)\right] + \int_t^s e^{(s-r)A_D}\left[-D\frac{d}{dr}\gamma(r) + f(r)\right] dr + D\gamma(s). \end{aligned} \tag{C.18}$$

which, upon using Lemma C.10, yields

$$X(s) = e^{(s-t)A_D}x - A_D\int_t^s e^{(s-r)A_D}D\gamma(r)dr + \int_t^s e^{(s-r)A_D}f(r)dr.$$

□

Observe that equation (C.15) can be seen as a mild form of equation

$$\begin{cases} \frac{d}{ds}X(s) = A_D X(s) - A_D D\gamma(s) + f(s) \\ X(t) = x \in H. \end{cases} \tag{C.19}$$

Define for $\epsilon > 0$

$$G_D := (-A_D)^{1/4-\epsilon}D. \tag{C.20}$$

By (C.5), $G_D \in \mathcal{L}(\Lambda, H)$. Thus (if $\gamma(\cdot) \in L^1(t,T;\Lambda)$)

$$A_D\int_t^s e^{(s-r)A_D}D\gamma(r)dr = \int_t^s (-A_D)^{3/4+\epsilon}e^{(r-t)A_D}G_D\gamma(r)dr. \tag{C.21}$$

Hence we can rewrite

$$X(s) = e^{(s-t)A_D}x - \int_t^s (-A_D)^{3/4+\epsilon}e^{(s-r)A_D}D\gamma(r)dr + \int_t^s e^{(s-r)A_D}f(r)dr. \tag{C.22}$$

NOTATION C.12 Expression (C.22) is called the *mild form* of equation (C.19) (and of (C.14)). ■

NOTATION C.13 We used the letter X for the unknown in the equation to be consistent with the notation we use in Chapter 2 where the variable y is only used for the equation in the PDE form. ■

Let $Q \in \mathcal{L}^+(H)$[1] and let $\left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s\in[t,T]}, \mathbb{P}, W_Q\right)$ be a generalized reference probability space.

We consider now the following stochastic parabolic equation:

$$\begin{cases} dy(s,\xi) = [\Delta_\xi y(s,\xi) + f(s,y(s,\xi))]\,ds + g(s,y(s,\xi))dW_Q(s)(\xi), \text{ on } (t,T)\times\mathcal{O} \\ y(s,\xi) = \gamma(s,\xi), \text{ on } (t,T)\times\partial\mathcal{O} \\ y(t,\xi) = x(\xi), \text{ on } \mathcal{O} \end{cases} \tag{C.23}$$

where $f, g : [t,T]\times\mathbb{R}\times\Omega \to \mathbb{R}$ and $\gamma : [t,T]\times\partial\mathcal{O}\times\Omega \to \mathbb{R}$ are appropriately measurable functions.

Denote $b(s,y)(\cdot) := f(s,y(s,\cdot))$ and $[\sigma(s,y)z](\cdot) := g(s,y(s,\cdot))z(\cdot)$ and consider the following integral equation

$$X(s) = e^{(s-t)A_D}x + \int_t^s e^{(s-r)A_D}b(r,X(r))dr + \int_t^s (-A_D)^{3/4+\epsilon}e^{(s-r)A_D}G_D\gamma(r)dr + \int_t^s e^{(s-r)A_D}\sigma(r,X(r))dW_Q(r) \ \ \mathbb{P}\text{-a.e.}. \tag{C.24}$$

Similarly to the deterministic case it can be viewed as the mild form of the equation

$$\begin{cases} dX(s) = [A_D X(s) - A_D D\gamma(s) + b(s,X(s))]\,ds + \sigma(s,X(s))dW_Q(s) \\ X(t) = x \in H. \end{cases} \tag{C.25}$$

NOTATION C.14 Equation (C.24) is called the *mild form* of equation (C.23) (and of equation (C.25)). Its solution is given by Definition 1.113, see Remark 1.114. Thanks to (C.20) we can rewrite the term $-A_D D$ in (C.25) as $(-A_D)^{3/4+\epsilon}G_D$ in (C.24). ■

Conditions under which some equations of the form (C.23) have unique mild solutions are given in Theorem 1.135 part (a). If $\sigma \notin \mathcal{L}_2(\Xi_0, H)$ then the stochastic term in (C.24) must be given proper interpretation and the same is also true about (C.33).

C.3. Non-zero boundary conditions, the Neumann case

The Neumann case is similar to the Dirichlet case which was explained in Subsection C.2.

We consider the following problem:

$$\begin{cases} \frac{\partial}{\partial s}y(s,\xi) = \Delta_\xi y(s,\xi) + f(s,\xi), & (s,\xi)\in(t,T)\times\mathcal{O} \\ \frac{\partial}{\partial n}y(s,\xi) = \gamma(s,\xi), & (s,\xi)\in(t,T)\times\partial\mathcal{O} \\ y(t,\xi) = x(\xi). \end{cases} \tag{C.26}$$

As before we denote respectively by H and Λ the Hilbert spaces $L^2(\mathcal{O})$ and $L^2(\partial\mathcal{O})$. A_N is the generator of the C_0-semigroup associated to heat equations with zero Neumann boundary conditions defined in (C.9), and $\lambda > 0$. Using the same arguments as these in the proof of Theorem C.15 one can show the following proposition.

[1]With the notation of Chapter 1, this means that we assume in this case $\Xi = H$.

PROPOSITION C.15 *If $y \in C^\infty([t,T] \times \overline{\mathcal{O}})$ is a classical solution of (C.26) then $X(s) := y(s,\cdot)$ can be written as*

$$X(s) = e^{(s-t)A_N}x - (A_N - \lambda I)\int_t^s e^{(s-r)A_N} N_\lambda \gamma(r)dr + \int_t^s e^{(s-r)A_N} f(r)dr. \quad \text{(C.27)}$$

The previous expression can be seen as the mild form of the equation

$$\begin{cases} \frac{d}{ds}X(s) = A_N X(s) + (\lambda I - A_N)N_\lambda \gamma(s) + f(s) \\ X(t) = x \in H. \end{cases} \quad \text{(C.28)}$$

If $\gamma \in L^1(t,T;\Lambda)$, thanks to (C.10) we have

$$-(A_N - \lambda I)\int_t^s e^{(s-r)A_N} N_\lambda \gamma(r)dr = \int_t^s (\lambda I - A_N)^{1/4+\epsilon} e^{(s-r)A_N} G_N \gamma(r)dr, \quad \text{(C.29)}$$

where

$$G_N := (\lambda I - A_N)^{3/4-\epsilon} N_\lambda \in \mathcal{L}(\Lambda, H). \quad \text{(C.30)}$$

Therefore we can rewrite (C.27) as

$$X(s) = e^{(s-t)A_N}x + \int_t^s (\lambda I - A_N)^{1/4+\epsilon} e^{(s-r)A_N} G_N \gamma(r)dr + \int_t^s e^{(s-r)A_N} f(r)dr. \quad \text{(C.31)}$$

NOTATION C.16 Equation (C.31) is called the *mild solution* of (C.26) and (C.28). ■

To define mild form of the stochastic parabolic equation

$$\begin{cases} dy(s,\xi) = \left[\Delta_\xi y(s,\xi) + f(s,y(s,\xi))\right] ds + g(s,y(s,\xi))dW_Q(s)(\xi), \text{ on } (t,T)\times\mathcal{O} \\ \frac{\partial}{\partial n} y(s,\xi) = \gamma(s,\xi), \text{ on } (t,T)\times\partial\mathcal{O} \\ y(t,\xi) = x(\xi), \text{ on } \mathcal{O} \end{cases} \quad \text{(C.32)}$$

we consider the integral equation

$$X(s) = e^{(s-t)A_N}x + \int_t^s e^{(s-r)A_N} b(r,X(r))dr + \int_t^s (\lambda I - A_N)^{1/4+\epsilon} e^{(s-r)A_N} G_N \gamma(r)dr + \int_t^s e^{(s-r)A_N}\sigma(r,X(r))dW_Q(r) \ \mathbb{P}\text{-a.e.}, \quad \text{(C.33)}$$

where $b(s,y)(\cdot) := f(s,y(\cdot))$ and $[\sigma(s,y)z](\cdot) := g(s,y(\cdot))z(\cdot)$. Conditions under which some equations of the form (C.32) have unique mild solutions are given in Theorem 1.135 part (b).

Equation (C.33) is in fact the *mild form* of the problem

$$\begin{cases} dX(s) = \left[A_N X(s) + (\lambda I - A_N)N_\lambda \gamma(s) + b(s,X(s))\right] ds + \sigma(s,X(s))dW_Q(s) \\ X(t) = x. \end{cases} \quad \text{(C.34)}$$

NOTATION C.17 Equation (C.33) is called the *mild form* of equation (C.32) (and of equation (C.34)). Thanks to (C.30) we can rewrite the term $(\lambda I - A_N)N_\lambda$ in (C.34) as $(\lambda I - A_N)^{1/4+\epsilon} G_N$ in (C.33). ■

Conditions under which some equations of the form (C.32) have unique mild solutions are given in Theorem 1.135 part (b).

C.4. Boundary noise, Neumann case

Let H, Λ, Q, and $\left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s\in[t,T]}, \mathbb{P}, W_Q\right)$ be defined as in Subsection C.2. We consider the following problem:

$$\begin{cases} \frac{\partial}{\partial s} y(s,\xi) = \Delta_\xi y(s,\xi) + f(s, y(s,\xi)), & (s,\xi) \in (t,T)\times\mathcal{O} \\ \frac{\partial}{\partial n} y(s,\xi) = h(s,y(s,\xi))\frac{dW_Q(s)}{ds} + g(s,\xi), & (s,\xi)\in(t,T)\times\partial\mathcal{O} \\ y(t,\xi) = x(\xi) \end{cases} \tag{C.35}$$

where $f, h : [t,T]\times\mathbb{R}\times\Omega \to \mathbb{R}$ and $g : [t,T]\times\partial\mathcal{O}\times\Omega\to\mathbb{R}$ are appropriately measurable functions.

As in Subsection C.3, $G_N := (\lambda I - A_N)^{3/4-\epsilon} N_\lambda \in \mathcal{L}(\Lambda, H)$, $\lambda > 0$, A_N is defined by (C.9) and N_λ by (C.7).

To rewrite the equation in an infinite-dimensional setting in H, we follow the approach used in Subsections C.2 and C.3. The idea is to consider formally the boundary term as a (particularly irregular) boundary condition corresponding to γ appearing in (C.26). So, denoting as before $b(s,y)(\cdot) := f(s, y(\cdot))$ and $[\sigma(s,y)z](\cdot) := h(s,y(\cdot))z(\cdot)$, we define the mild form of (C.35) as

$$\begin{aligned} X(s) = e^{(s-t)A_N}x + \int_t^s e^{(s-r)A_N} b(r, X(r))dr \\ + \int_t^s (\lambda I - A_N)^{1/4+\epsilon} e^{(s-r)A_N} G_N g(r) dr \\ + \int_t^s (\lambda I - A_N)^{1/4+\epsilon} e^{(s-r)A_N} G_N \sigma(r, X(r)) dW_Q(r) \quad \mathbb{P}\text{-a.e.} \end{aligned} \tag{C.36}$$

The above expression can also be referred to as the *mild form* of the infinite dimensional problem

$$\begin{cases} dX(s) = [A_N X(s) + (\lambda I - A_N)N_\lambda g(s) + b(s, X(s))]\, ds + (\lambda I - A_N) N_\lambda \sigma(s, X(s)) dW_Q(s) \\ X(t) = x \in H, \end{cases} \tag{C.37}$$

where, thanks to (C.30), the term $(\lambda I - A_N)N_\lambda g(s)$ is rewritten as $(\lambda I - A_N)^{1/4+\epsilon} G_N g(s)$ and the stochastic term is rewritten similarly.

Theorem 1.135 part (b) provides conditions under which some equations of the form (C.35) have unique continuous mild solutions.

NOTATION C.18 The solution of (C.36) is called the *mild solution* of (C.35) (and of (C.37)). ■

C.5. Boundary noise, Dirichlet case

To apply the approach of Subsection C.4 to the Dirichlet-boundary-noise version of the problem (C.35) we would need to give a meaning to the term

$$(-A_D)^{3/4+\epsilon} \int_t^s e^{(s-r)A_D} G_D \sigma(r, X(r)) dW_Q(r),$$

where A_D and $G_D \in \mathcal{L}(\Lambda, H)$ are introduced respectively in (C.4) and (C.20). However estimate (B.15) does not allow to prove the convergence of the integral

$$\int_t^T \left\| (-A_D)^{3/4+\epsilon} e^{(T-r)A_D} G_D \right\|^2_{\mathcal{L}_2(\Lambda, H)} dr.$$

One way to resolve this problem is to look for a solution of the stochastic PDE in the completion of H in the weaker norm $|x|_* := |(-A_D)^{-\alpha}x|$ for some $\alpha > 0$. It is, roughly speaking, a space of distributions. This approach was used e.g. in [125], see in particular Proposition 3.1, page 176.

Another way, used e.g. by [6, 46, 165], is to look for a solution in a weighted L^2 space. Consider for example a simple case of the following problem on the positive half-line

$$\begin{cases} \frac{\partial}{\partial s}y(s,\xi) = \frac{\partial^2}{\partial \xi^2}y(s,\xi) + f(s,y(s,\xi)), & (s,\xi) \in (t,T)\times\mathbb{R}^+ \\ y(s,0) = h(s,y(t,0))\frac{dW(s)}{ds} + g(s), & s \in (t,T) \\ y(t,\xi) = x(\xi), \end{cases} \tag{C.38}$$

where W a one-dimensional Brownian motion.

Define the weight $\eta(\xi) := \min\{1, \xi^{1+\theta}\}$ for some $\theta > 0$ and the weighted L^2 space

$$L^2_\eta := \left\{ p : [0,+\infty) \to \mathbb{R} \text{ measurable } : \int_0^{+\infty} p^2(\xi)\eta(\xi)d\xi < \infty \right\}.$$

It is a real separable Hilbert space. The inner product in L^2_η is given by

$$\langle p, q\rangle_\eta := \int_0^{+\infty} p(\xi)q(\xi)\eta(\xi)d\xi$$

and it induces the usual norm

$$|p|_\eta = \left(\int_0^{+\infty} p(\xi)^2\eta(\xi)d\xi\right)^{1/2}.$$

Proposition C.19 *The heat semigroup with zero Dirichlet-boundary conditions extends to a C_0-semigroup on L^2_η. The semigroup is analytic. Denote its generator by A_η. For $\lambda > 0$, $(\lambda I - A_\eta)$ is invertible and the Dirichlet map*

$$D_\lambda a = \phi \iff \begin{cases} (\lambda I - \partial_x^2)\phi[\xi] = 0 & \text{for all } \xi > 0 \\ \phi(0) = a \end{cases}$$

is linear and continuous from $\mathbb{R}$ to $D((\lambda I - A_\eta)^\alpha)$ for all $\alpha \in [0, 1/2 + \theta/4)$.

Proof. See Proposition 2.1 and Lemma 2.2 in [165]. □

Using the above proposition we can proceed in a fashion similar to that of Subsection C.4. We fix $\lambda > 0$ and we choose

$$\alpha_\theta := \frac{1}{2} + \frac{\theta}{8}.$$

We define $G_\eta := (\lambda I - A_\eta)^{\alpha_\theta}D_\lambda \in \mathcal{L}(\mathbb{R}; L^2_\eta)$ and rewrite (C.38) as an integral equation

$$\begin{aligned} X(s) = e^{(s-t)A_\eta}x + \int_t^s e^{(s-r)A_\eta}b(r,X(r))dr \\ + \int_t^s (\lambda I - A_\eta)^{(1-\alpha_\theta)}e^{(s-r)A_\eta}G_\eta g(r)dr \\ + \int_t^s (\lambda I - A_\eta)^{(1-\alpha_\theta)}e^{(s-r)A_\eta}G_\eta\sigma(r,X(r))dW_Q(r) \ \mathbb{P}\text{-a.e.} \end{aligned} \tag{C.39}$$

This integral equation is in fact the mild form of the evolution equation in $L^2(t,T;\mathbb{R})$

$$\begin{cases} dX(s) = \left[A_\eta X(s) + (\lambda I - A_\eta)^{(1-\alpha_\theta)} G_\eta g(s) + b(s, X(s))\right] ds \\ \qquad\qquad\qquad\qquad\qquad\qquad + (\lambda I - A_\eta)^{(1-\alpha_\theta)} G_\eta \sigma(s, X(s)) dW_Q(s) \\ X(t) = x \in L^2(t,T;\mathbb{R}), \end{cases} \tag{C.40}$$

and we remark that $(\lambda I - A_\eta)^{(1-\alpha_\theta)} G_\eta = (\lambda I - A_\eta) D_\lambda$.

One can use Theorem 1.135 part (b) to obtain existence and uniqueness of mild solution $X(\cdot)$ of (C.40).

APPENDIX D

Functions, derivatives and approximations

In this appendix we collect some standard definitions related to functions which are used in the text and are recalled here for the reader's convenience. In the second part of this appendix we recall the definition and some properties of the sup-inf convolutions.

D.1. Continuity properties, modulus of continuity

DEFINITION D.1 (Modulus) *We say that a function $\rho : [0, +\infty) \to [0, +\infty)$ is a* modulus (of continuity) *if ρ is continuous, nondecreasing, subadditive, and $\rho(0) = 0$.*

In the literature the subadditivity property in the definition of a modulus is not always required and continuity is sometimes required only at 0. The following theorem shows that one can use Definition D.1 without loss of generality.

THEOREM D.2 *Consider $\eta : \mathbb{R}^+ \to \mathbb{R}^+$, non-decreasing, such that $\lim_{s\to 0} \eta(s) = \eta(0) = 0$, and*

$$\lim_{s\to\infty} \frac{\eta(s)}{s} < +\infty.$$

Then there exists a modulus ρ (in the sense of Definition D.1) such that $\rho(s) \geq \eta(s)$ for all $s \in \mathbb{R}^+$.

PROOF. See Theorem 1 page 406 in [11]. □

The modulus ρ in Theorem D.2 can be also assumed to be concave. Thus we will always assume that a modulus is concave.

REMARK D.3 Thanks to the subadditivity, we have the following property for any modulus $\rho(\cdot)$: given any $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that

$$\rho(r) \leq \varepsilon + C_\varepsilon r \qquad \text{for every } r \geq 0. \tag{D.1}$$

■

DEFINITION D.4 (Local Modulus) *A function $\rho : [0, +\infty) \times [0, +\infty) \to [0, +\infty)$ is called a* local modulus *if the following three conditions are satisfied:*

(i) *ρ is continuous and nondecreasing in both variables.*
(ii) *ρ is subadditive in the first variable.*
(iii) *$\rho(0, r) = 0$ for every $r \geq 0$.*

DEFINITION D.5 *Consider two Banach spaces E_0 and E_1. Given $\varphi \in UC(E_0; E_1)$, we define its* modulus of continuity *$\rho[\varphi](\cdot)$ as follows:*

$$\rho[\varphi](\epsilon) := \sup_{x,y\in E_0} \left\{ |\varphi(x) - \varphi(y)|_{E_1} \ : \ |x - y|_{E_0} \leq \varepsilon \right\}, \qquad \text{for } \varepsilon \geq 0. \tag{D.2}$$

Of course $|\varphi(x) - \varphi(y)|_{E_1} \le \rho[\varphi](|x-y|_{E_0})$ for every $x, y \in E_0$. Any modulus with such property is also called a *modulus of continuity of* φ. In particular, for every $\varphi \in UC(E_0; E_1)$ we can always find two positive constants C_0, C_1 such that

$$|\varphi(x)|_{E_1} \le C_0 + C_1|x|_{E_0}, \qquad \text{for every } x \in E_0.$$

DEFINITION D.6 (Local boundedness from above (below)) *Let $G \subset E_0$. A function $u : G \to \mathbb{R}$ is said to be* locally bounded from above (respectively, below) *if, for every $R > 0$, u is bounded from above (below) on $G \cap (B_{E_0}(0, R))$. u is said to be locally bounded if it is both locally bounded from above and from below.*

DEFINITION D.7 (Local uniform continuity) *Let $G \subset E_0$. A function $u : G \to \mathbb{R}$ is said to be* locally uniformly continuous *if, for every $R > 0$, its restriction to $G \cap (B_{E_0}(0, R))$ is uniformly continuous.*

DEFINITION D.8 (Local uniform convergence) *Given $G \subset E_0$, and $u_n, u \in C(G)$, we say that u_n converge locally uniformly to u if, for every $R > 0$, the restrictions of u_n to $G \cap (B_{E_0}(0, R))$ converge uniformly to the restriction of u to $G \cap (B_{E_0}(0, R))$.*

DEFINITION D.9 (Upper/lower semicontinuous envelope) *Let $G \subset E_0$. Consider a function $u : G \to \overline{\mathbb{R}}$. Its* upper (respectively, lower) semicontinuous envelope *is defined as follows:*

$$\begin{cases} u^* : G \to \overline{\mathbb{R}} \\ u^*(x) := \limsup\limits_{\substack{y \in G \\ y \to x}} u(y) \qquad \textit{(respectively, } u_*(x) := \liminf\limits_{\substack{y \in G \\ y \to x}} u(y)\textit{).} \end{cases}$$

DEFINITION D.10 (Strict maximum/minimum) *Let $G \subset E_0$. We say that a function $u : G \to \mathbb{R} \cup \{-\infty\}$ (respectively, $u : G \to \mathbb{R} \cup \{+\infty\}$) has a* strict maximum (respectively, strict minimum) *at $x \in G$ if u has a maximum (respectively, minimum) at x and whenever $x_n \in G$ are such that $u(x_n) \to u(x)$ then $x_n \to x$. Similarly we define a strict local maximum and a strict local minimum.*

D.2. Fréchet and Gâteaux derivatives

Throughout this section Y, Z are Banach spaces, X is an open subset of Z, and H is a real separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$.

DEFINITION D.11 (Fréchet derivative) *A function $u : X \to Y$ is said to be* Fréchet differentiable *at a point $\bar{x} \in X$ if there exists a linear functional $Du(\bar{x}) \in \mathcal{L}(Z; Y)$ such that*

$$\lim_{|x-\bar{x}|_Z \to 0} \frac{|u(x) - u(\bar{x}) - Du(\bar{x})(x - \bar{x})|_Y}{|x - \bar{x}|_Z} = 0.$$

NOTATION D.12 We consider a continuous function $u : X \to Y$ having Fréchet derivative at all $x \in X$. If the function

$$\begin{cases} X \to \mathcal{L}(Z, Y) \\ \bar{x} \mapsto Du(\bar{x}) \end{cases}$$

is continuous ($\mathcal{L}(Z, Y)$ is endowed with the operator norm) then u is said to be *continuously (Fréchet) differentiable* on X. ■

The k-th order Fréchet derivative $D^k u$ of u is defined inductively

$$D^1 u = Du, \qquad D^k u = D(D^{k-1} f),\ k = 1, 2, 3, \dots$$

(see e.g. [198] p.186). If $D^k u$ is defined at a point $\bar{x} \in X$ then we say that u is k times Fréchet differentiable at $\bar{x}$ and we call $D^k u(\bar{x})$ the k-th Fréchet derivative of u at $\bar{x}$. Clearly $D^k u(\bar{x}) \in \mathcal{L}(Z, \mathcal{L}(Z, \ldots, \mathcal{L}(Z, Y)) \ldots)$. Since this space is isometrically isomorphic to $\mathcal{L}^k(Z;Y)$, we will always consider $D^k u(\bar{x})$ as an element of this last space endowed with its natural norm $\|T\|_{\mathcal{L}^k(Z;Y)} := \sup_{z \in Z^k, z \neq 0} \frac{|T(z_1,\ldots,z_n)|_Y}{|(z_1,\ldots,z_n)|_{Z^k}}$.

If X is an open subset of H and $Y = \mathbb{R}$ then thanks to the Riesz Representation Theorem, we can identify the linear functional $Du(\bar{x})$ with the element $y \in H$ such that $\langle y, x \rangle = Du(\bar{x})(x)$ for all $x \in H$. Abusing slightly notation, we will denote y by $Du(\bar{x})$. The second order derivative $D^2 u(\bar{x})$ can be identified with a symmetric bilinear form in $\mathcal{L}^2(H;\mathbb{R})$ (see e.g. [198] 3.5.4 and 3.5.5, p.190). So we can identify $D^2 u(\bar{x})$ with the unique $T \in \mathcal{S}(H)$ having the property that

$$\langle Tx, y \rangle = D^2 u(\bar{x})(x)(y) = D^2 u(\bar{x})(y)(x) \qquad \text{for all } x, y \in H.$$

With an abuse of notation we will again denote T by $D^2 u(\bar{x})$.

The following is a special case of the Generalized Taylor's Theorem.

THEOREM D.13 *Let the function $f : U(x) \subset H \to \mathbb{R}$ be defined on an open neighborhood $U(x)$, of x, and let $f \in C^2(U(x))$. Then, for all h in a neighborhood of the origin in H, we have*

$$f(x+h) = f(x) + \langle Df(x), h \rangle + \langle D^2 f(x)h, h \rangle + o(|h|^2)$$

PROOF. See [457], page 148. □

We now introduce now the notion of Gâteaux derivative and we give a special case of the Generalized Taylor's Theorem for Gâteaux differentiable functions.

DEFINITION D.14 (Gâteaux derivative) *A function $u : X \to Y$ is said to be* Gâteaux differentiable *at a point $\bar{x} \in X$ if there exists a linear functional $\nabla u(\bar{x}) \in \mathcal{L}(Z;Y)$ such that, for any $y \in Z$ with $|y|_Z = 1$,*

$$\lim_{t \to 0^+} \frac{|u(\bar{x}+ty) - u(\bar{x}) - t\nabla u(\bar{x})(y)|_Y}{t} = 0. \tag{D.3}$$

THEOREM D.15 *Let the function $f : U(x) \subset H \to \mathbb{R}$ be defined on an open, convex neighborhood $U(x)$ of x. Consider $h \in Z$ such that $x + h \in U(x)$. Suppose that f is Gâteaux differentiable at any point $x \in U(x)$ and that the mapping*

$$y \mapsto \nabla f(y)(h)$$

is continuous on $U(x)$. Then

$$f(x+h) = f(x) + \int_0^1 \nabla f(x+th)(h)dt.$$

PROOF. See Theorem 4.A, page 148 of [457]. □

Other type of derivatives are also used in the book. The basic material on directional derivatives can be found in Section 6.1.2. The concepts of G-directional derivative, G-Gâteaux derivative and G-Fréchet derivative, used in Chapters 4 and 5, are introduced in Section 4.2.1.

D.3. Inf-Sup convolutions

We recall here some results from [310] on approximations of bounded uniformly continuous functions in Hilbert spaces. Consider a real and separable Hilbert space H with inner product $\langle \cdot, \cdot \rangle$, and $\dim(H) = +\infty$. Since closed balls are not compact in H, functions in $C_b(H)$ cannot be approximated uniformly on bounded sets by functions in $UC_b(H)$ and consequently by functions in $C_b^1(H)$. It was proved in [302] that $UC_b(H) \cap C^\infty(H)$ is dense in $UC_b(H)$. However, it was observed in [354] that, differently from what we have in the finite dimensional case, $UC_b^2(H)$ is not a dense subset of $UC_b(H)$. Nevertheless the inclusion $UC_b^{1,1}(H) \subset UC_b(H)$ remains dense. Lasry and P. L. Lions introduced in [310] an explicit way to approximate functions in $UC_b(H)$ by elements of $UC_b^{1,1}(H)$, the so-called *inf-sup-convolutions* and *sup-inf-convolutions*. These explicit approximations have many other interesting properties, for instance they preserve order and commute with translations. More information about them can be found in [310] and [129]. We also remark that P. L. Lions proved in [319] that functions in $UC_b^{1,1}(H)$ can be uniformly approximated by functions in $UC_b^{1,1}(H)$ with uniformly continuous second order partial derivatives, for which Itô's formula can be applied (see the proof of Lemma IV.1 and Lemma III.2 in [319]). The technique of [319] is based on limits of mollifications over increasing finite dimensional subspaces of H. Also in [387, 388] it is proved that the space $UC_s^2(H)$, the subspace of $UC_b^{1,1}(H)$ admitting weakly uniformly continuous second Hadamard derivative, is dense in $UC_b(H)$.

DEFINITION D.16 (Semiconvex and semiconcave functions) *A function $u : H \to \mathbb{R}$ is said to be* semiconcave *(respectively,* semiconvex*) if there exists a constant $M \geq 0$ such that*

$$x \mapsto u(x) - M|x|^2 \qquad \textit{(respectively, } x \mapsto u(x) + M|x|^2\textit{)}$$

is concave (respectively, convex).

DEFINITION D.17 (Sup-convolution and Inf-convolution) *Given $u \in C_b(H)$ and $\epsilon > 0$, we define the* inf-convolution *of u as*

$$u_\epsilon(x) := \inf_{y \in H} \left(u(y) + \frac{|x-y|^2}{2\epsilon} \right), \qquad x \in H$$

and its sup-convolution *as*

$$u^\epsilon(x) := \sup_{y \in H} \left(u(y) - \frac{|x-y|^2}{2\epsilon} \right), \qquad x \in H.$$

We set $u_0(x) = u^0(x) := u(x)$.

PROPOSITION D.18 *For all $\epsilon, \delta \geq 0$ and $u \in UC_b(H)$, $u_\epsilon, u^\epsilon \in UC_b(H)$, and $(u_\epsilon)_\delta = u_{\epsilon+\delta}, (u^\epsilon)^\delta = u^{\epsilon+\delta}$. Moreover, for all $u, v \in UC_b(H)$ and $\epsilon > 0$, we have the following properties:*

(1) $\|u_\epsilon - v_\epsilon\|_0 \leq \|u - v\|_0$ and $\|u^\epsilon - v^\epsilon\|_0 \leq \|u - v\|_0$.
(2) $\rho[u_\epsilon] \leq \rho[u]$ and $\rho[u^\epsilon] \leq \rho[u]$.
(3) $\lim_{\epsilon \to 0} u_\epsilon = u$ and $\lim_{\epsilon \to 0} u^\epsilon = u$ *in $UC_b(H)$, more precisely*

$$\|u_\epsilon - u\|_0 \leq \rho[u](2\sqrt{\epsilon \|u\|_0}) \text{ and } \|u^\epsilon - u\|_0 \leq \rho[u](2\sqrt{\epsilon \|u\|_0}).$$

(4) $x \mapsto u_\epsilon(x) - \frac{|x|^2}{2\epsilon}$ *is concave.*
(5) $x \mapsto u^\epsilon(x) + \frac{|x|^2}{2\epsilon}$ *is convex.*
(6) u_ϵ *and u^ϵ are Lipschitz continuous, and, for all $x, y \in H$,*

$$\frac{|u_\epsilon(x) - u_\epsilon(y)|}{|x-y|}, \ \frac{|u^\epsilon(x) - u^\epsilon(y)|}{|x-y|} \leq \frac{2\sqrt{\|u\|_0}}{\sqrt{\epsilon}}$$

PROOF. Most of the claims are proved in [129], Propositions C.3.2, C.3.3, and C.3.4. We sketch the proof of (6) for u^ϵ.

Let $x, y \in H$. We define the function $v : \mathbb{R} \to \mathbb{R}$ by

$$v(t) := u^\epsilon(x + t(y - x)).$$

Since u^ϵ is semiconvex, v is semiconvex and hence locally Lipschitz and differentiable a.e. Let $0 < \bar{t} < 1$ be a point of differentiability of v and let $\varphi \in C^1(\mathbb{R})$ be such that $v - \varphi$ has a strict maximum at $\bar{t}$. By Theorem 3.25, for every n there are $a_n \in \mathbb{R}, p_n \in H, |a_n| + |p_n| < 1/n$ such that

$$u(z) - \frac{|x + t(y - x) - z|^2}{2\epsilon} - \varphi(t) + \langle p_n, z\rangle + a_n t \tag{D.4}$$

has a global maximum over $[0, 1] \times H$ at some point (t_n, z_n). Since the maximum of $v - \varphi$ at $\bar{t}$ was strict, it is easy to see that we must have $t_n \to \bar{t}$ and $|z_n| \leq C, n \in \mathbb{N}$ for some C. Moreover, setting $z = z_n$ and differentiating the function in (D.4) with respect to t we obtain

$$D\varphi(t_n) = \frac{\langle x + t_n(y - x) - z_n, x - y\rangle}{\epsilon} + a_n. \tag{D.5}$$

Taking $z = x + t_n(y - x)$ we get

$$u(x + t_n(y - x)) + \langle p_n, x + t_n(y - x)\rangle \leq u(z_n) - \frac{|x + t_n(y - x) - z_n|^2}{2\epsilon} + \langle p_n, z_n\rangle,$$

which implies

$$\frac{|x + t_n(y - x) - z_n|^2}{2\epsilon} \leq u(z_n) - u(x + t_n(y - x)) + \langle p_n, z_n - x - t_n(y - x)\rangle \leq 2\|u\|_0 + \frac{C_1}{n}.$$

Using this and letting $n \to +\infty$ in (D.5) we thus obtain

$$|D\varphi(\bar{t})| \leq \frac{2\sqrt{\|u\|_0}}{\sqrt{\epsilon}}|y - x|. \tag{D.6}$$

We can thus conclude

$$|u^\epsilon(y) - u^\epsilon(x)| = |v(1) - v(0)| \leq \int_0^1 |v'(t)|dt \leq \frac{2\sqrt{\|u\|_0}}{\sqrt{\epsilon}}|y - x|.$$

□

We remark that the above proof shows that one can in fact obtain

$$|u(y) - u(x)| \leq \frac{t_\epsilon}{\sqrt{\epsilon}}|y - x|,$$

where t_ϵ is defined in Proposition D.21.

We also remark that it follows from semiconvexity of u_ϵ that if $v'(t)$ exists, we must have

$$v'(t) = \langle p, y - x\rangle \quad \text{for every } p \in D^- u^\epsilon(x + t(y - x)),$$

where $D^- u^\epsilon(x+t(y-x))$ is the subdifferential of u^ϵ at $(x+t(y-x))$ (see Definition E.1). Moreover, $D^- u^\epsilon(z)$ is non-empty for every z. In fact for a semiconvex function $w : H \to \mathbb{R}$,

$$D^- w(z) = \overline{\mathrm{conv}}\{p : Dw(z_n) \rightharpoonup p, z_n \to z\}, \tag{D.7}$$

where z_n above are points of Fréchet differentiability of w (see [48], page 522). Lipschitz functions on H are Fréchet differentiable on a dense subset of H by Preiss's theorem.

Finally we remark that if u^ϵ is differentiable at some point $\bar{x}$ then $Du^\epsilon(\bar{x}) = (\bar{y} - \bar{x})/\epsilon$, where $\bar{y}$ is a unique point such that

$$u^\epsilon(\bar{x}) = u(\bar{y}) - \frac{|\bar{x} - \bar{y}|^2}{2\epsilon}. \tag{D.8}$$

To show this let $\varphi \in C^1(H)$ be such that $u^\epsilon - \varphi$ has a global strict maximum at $\bar{x}$ and $\varphi(x) = |x|^2$ if $|x|$ is large enough. By Theorem 3.25 for every n there are $p_n, q_n \in H, |p_n| + |q_n| < 1/n$ such that

$$u(y) - \frac{|x - y|^2}{2\epsilon} - \varphi(x) + \langle p_n, x\rangle + \langle q_n, y\rangle \tag{D.9}$$

has a global maximum over $H \times H$ at some point (x_n, y_n). Since the maximum of $u^\epsilon - \varphi$ at $\bar{x}$ was strict, it is easy to see that we must have $x_n \to \bar{x}$ and

$$u(y_n) - \frac{|x_n - y_n|^2}{2\epsilon} \to u^\epsilon(\bar{x})$$

as $n \to +\infty$. Moreover, setting $y = y_n$ and differentiating the function in (D.9) with respect to x we obtain

$$D\varphi(x_n) = -\frac{x_n - y_n}{\epsilon} + p_n,$$

which implies that $\lim_{n\to+\infty} y_n = \bar{y}$ for some point $\bar{y}$,

$$Du^\epsilon(\bar{x}) = D\varphi(\bar{x}) = \frac{\bar{y} - \bar{x}}{\epsilon}, \tag{D.10}$$

and (D.8) is satisfied. Since (D.10) holds for every point $\bar{y}$ satisfying (D.8), it must be unique.

DEFINITION D.19 (Inf-sup and sup-inf-convolutions) *Given $\varepsilon > 0$ and a function $u \in C_b(H)$, we define the* inf-sup convolution *(respectively,* sup-inf-convolution*) of u as*

$$\underline{u}_\epsilon := (u_\epsilon)^{\frac{\epsilon}{2}} \qquad \text{(respectively, } \overline{u}_\epsilon := (u^\epsilon)_{\frac{\epsilon}{2}}\text{)}.$$

More explicitly we have, for $x \in H$,

$$\underline{u}_\epsilon(x) = \sup_{z\in H} \inf_{y\in H} \left(u(y) + \frac{1}{2\epsilon}|z-y|^2 - \frac{1}{\epsilon}|z-x|^2 \right)$$

(respectively,

$$\overline{u}_\epsilon(x) = \inf_{z\in H} \sup_{y\in H} \left(u(y) - \frac{1}{2\epsilon}|z-y|^2 + \frac{1}{\epsilon}|z-x|^2 \right).)$$

PROPOSITION D.20 *The inf-sup-convolution and the sup-inf-convolution preserve the order. In other words, given $u, v \in UC_b(H)$ such that*

$$u(x) \geq v(x), \qquad \text{for all } x \in H,$$

we have $\underline{u}_\varepsilon(x) \geq \underline{v}_\varepsilon(x)$ and $\overline{u}_\varepsilon(x) \geq \overline{v}_\varepsilon(x)$ for all $x \in H$. Moreover the inf-sup-convolution and the sup-inf-convolution commute with translations, i.e. for every $y \in H$ and translation $\tau_y : x \mapsto x - y$, we have $\overline{(\tau_y u)}_\varepsilon(x) = (\tau_y \overline{u}_\varepsilon)(x)$ and $\overline{(\tau_y u)}_\varepsilon(x) = (\tau_y \overline{u}_\varepsilon)(x)$ for all $x \in H$.

PROOF. See [310]. □

PROPOSITION D.21 *Let $u \in UC_b(H)$. Let t_ϵ be the maximum positive root of the equation*

$$t_\epsilon^2 = 2\epsilon\rho[u](t_\epsilon),$$

which implies that $t_\epsilon \epsilon^{-1/2} \to 0$ as $\epsilon \to 0$. Then $\underline{u}_\epsilon$ and $\overline{u}_\epsilon$ belong to $UC_b^{1,1}(H)$ and, for all $x, y \in H$, the following properties hold:

(1) $\inf_{y\in H} u(y) \leq \underline{u}_\epsilon(x) \leq u(x) \leq \overline{u}_\epsilon(x) \leq \sup_{y\in H} u(y)$.
(2) $|\underline{u}_\epsilon(x) - \underline{u}_\epsilon(y)| \leq \rho[u](|x-y|)$ and $|\overline{u}_\epsilon(x) - \overline{u}_\epsilon(y)| \leq \rho[u](|x-y|)$.
(3) $\|\underline{u}_\epsilon - u\|_0, \|\overline{u}_\epsilon - u\|_0 \leq \rho[u](t_\epsilon)$.
(4) $\|D\underline{u}_\epsilon\|_0, \|D\overline{u}_\epsilon\|_0 \leq \dfrac{t_\epsilon}{\epsilon}$.
(5) $|D\underline{u}_\epsilon(x) - D\underline{u}_\epsilon(y)|, |D\overline{u}_\epsilon(x) - D\overline{u}_\epsilon(y)| \leq \frac{2}{\epsilon}|x-y|$.

PROOF. See [310], pages 260-261. □

REMARK D.22 In fact $x \mapsto \underline{u}_\epsilon(x) - \frac{|x|^2}{2\epsilon}$ is concave and $x \mapsto \underline{u}_\epsilon(x) + \frac{|x|^2}{\epsilon}$ is convex. Similarly $x \mapsto \overline{u}_\epsilon(x) - \frac{|x|^2}{\epsilon}$ is concave and $x \mapsto \overline{u}_\epsilon(x) + \frac{|x|^2}{2\epsilon}$ is convex. ■

The inf-sup and sup-inf-convolutions can also be used to approximate more general functions.

PROPOSITION D.23 *If $u \in C(H)$ and there exists $C > 0$ such that $|u(x)| \leq C(1 + |x|^2)$ then, for all ϵ small enough, the inf-sup convolution $\underline{u}_\epsilon$ and the sup-inf convolution $\overline{u}_\epsilon$ are well defined, they belong to $C^{1,1}(H)$ and they converge pointwise to u when $\epsilon \to 0$. Moreover, if u is locally uniformly continuous then $\underline{u}_\epsilon$ and $\overline{u}_\epsilon$ converge to u locally uniformly.*

PROOF. See [310] page 261 (iii). □

We remark that to obtain pointwise convergence in the above proposition one can replace $u \in C(H)$ by $u \in USC(H)$ for $\overline{u}_\epsilon$, and by $u \in LSC(H)$ for $\underline{u}_\epsilon$.

D.4. Two versions of Gronwall's Lemma

We recall two versions of Gronwall's Lemma. The first is a well known result while the second is more specialized. (See e.g. [161] or [198] pages 95-97 for similar versions of Gronwall's Lemma). .

PROPOSITION D.24 (Gronwall's Lemma 1) *Let $T \in [0, +\infty) \cup \{+\infty\}$ and $I = [t_0, T)$. Let $a(\cdot)$ be a nonnegative, measurable, nondecreasing function on I. Let $b(\cdot)$ be a non-negative, locally integrable function on I. Suppose that $u(\cdot)$ is a non-negative function such that $b(\cdot)u(\cdot)$ is locally integrable on I. Assume also that*

$$u(s) \leq a(s) + \int_{t_0}^{s} b(r)u(r)dr, \qquad \text{for a.e. } s \in I. \tag{D.11}$$

Then

$$u(s) \leq a(s) \exp\left(\int_{t_0}^{s} b(r)dr\right), \qquad \text{for a.e. } s \in I,$$

PROOF. Define

$$v(s) = \exp\left(-\int_{t_0}^{s} b(r)dr\right) \int_{t_0}^{s} b(r)u(r)dr, \qquad s \in I.$$

Then, for a.e. $s \in I$, $v'(s)$ exists and

$$v'(s) = \left(u(s) - \int_{t_0}^{s} b(r)u(r)dr\right) b(s) \exp\left(-\int_{t_0}^{s} b(r)dr\right).$$

So, using (D.11) and integrating we have

$$v(s) \leq \int_{t_0}^{s} a(r)b(r) \exp\left(-\int_{t_0}^{r} b(\tau)d\tau\right) dr.$$

Now, since

$$\int_{t_0}^{s} b(r)u(r)dr = v(s) \exp\left(\int_{t_0}^{s} b(r)dr\right), \qquad s \in I,$$

by (D.11) we get

$$\begin{aligned} u(s) &\leq a(s) + \int_{t_0}^{s} b(r)u(r)dr = a(s) + \exp\left(\int_{t_0}^{s} b(r)dr\right) v(s) \\ &\leq a(s) + \exp\left(\int_{t_0}^{s} b(r)dr\right) \int_{t_0}^{s} a(r)b(r) \exp\left(-\int_{t_0}^{r} b(\tau)d\tau\right) dr \\ &= a(s) + \int_{t_0}^{s} a(r)b(r) \exp\left(\int_{r}^{s} b(\tau)d\tau\right) dr. \end{aligned}$$

Since the function $a(\cdot)$ is nondecreasing the above implies

$$u(s) \le a(s) + \int_{t_0}^{s} a(s) b(r) \exp\left(\int_r^s b(\tau) d\tau \right) dr$$
$$= a(s) + \left(-a(s) \exp\left(\int_r^s b(\tau) d\tau \right) \right) \Big|_{r=t_0}^{r=s} = a(s) \exp\left(\int_{t_0}^{s} b(r) dr \right)$$

which is the claim. ☐

PROPOSITION D.25 (Gronwall's Lemma 2) *Let $T \in [0, +\infty) \cup \{+\infty\}$, $b \ge 0$, $\beta > 0$. Let $a(\cdot)$ be a non-negative, locally integrable function on $[0, T)$. Suppose that $u(\cdot)$ is a non-negative, locally integrable function on $[0, T)$ such that*

$$u(s) \le a(s) + b \int_0^s (s-r)^{\beta - 1} u(r) dr, \qquad \text{for a.e. } s \in [0, T].$$

Then

$$u(s) \le a(s) + \theta \int_0^s E'_\beta(\theta(s-r)) a(r) dr, \qquad \text{for a.e. } s \in [0, T],$$

where, for $s > 0$, $\Gamma(s) := \int_0^{+\infty} r^{s-1} e^{-r} dr$, and θ and $E_\beta(s)$ are defined as

$$\theta := (b \Gamma(\beta))^{1/\beta}$$

and

$$E_\beta(s) := \sum_{n=0}^{+\infty} \frac{s^{n\beta}}{\Gamma(n\beta + 1)}.$$

The function $E'_\beta(s) = \frac{d}{ds} E_\beta(s)$ has the following properties: (i) $E'_\beta(s) = \frac{s^{\beta-1}}{\Gamma(\beta)} + o(s^{\beta-1})$ as $s \to 0^+$, (ii) $E'_\beta(s) = \frac{e^s}{\beta} + o(e^s)$ as $s \to +\infty$.

As a particular case, if $a, b, T \in \mathbb{R}^+$ and $\alpha, \beta \in [0, 1)$, there exists $M \in \mathbb{R}$ (depending on b, T, α and β) such that any integrable function $u : [0, T] \to \mathbb{R}$ such that

$$0 \le u(s) \le a s^{-\alpha} + b \int_0^s (s-r)^{-\beta} u(r) dr, \qquad \text{for a.e. } s \in [0, T],$$

satisfies

$$0 \le u(s) \le a M s^{-\alpha}, \qquad \text{for a.e. } s \in [0, T].$$

PROOF. See Lemma 7.1.1, Chapter 7, page 188 of [256]. The second claim is again in [256], Subsection 1.2.1, page 6. ☐

APPENDIX E

Viscosity solutions in $\mathbb{R}^N$

We collect some basic definitions and results about viscosity solutions in finite dimensional spaces. We refer the reader to [101] and the books [30, 31, 195] for more information on the subject.

E.1. Second order jets

Let $\mathcal{O}$ be an open subset of $\mathbb{R}^N$.

DEFINITION E.1 (Sub- and superdifferentials) *Let $u : \mathcal{O} \to \mathbb{R}$ be an upper semicontinuous function. The* superdifferential *of u at a point $\bar{x} \in \mathcal{O}$ is defined as*

$$D^+u(\bar{x}) := \left\{ p \in \mathbb{R}^N \ : \ \limsup_{y\to\bar{x}, y\in\mathcal{O}} \frac{u(y) - u(\bar{x}) - \langle p, y - \bar{x}\rangle}{|y - \bar{x}|} \le 0 \right\}.$$

Similarly, given a lower semicontinuous function $u : \mathcal{O} \to \mathbb{R}$, the subdifferential *of u at a point $\bar{x} \in \mathcal{O}$ is defined as*

$$D^-u(\bar{x}) := \left\{ p \in \mathbb{R}^N \ : \ \liminf_{y\to\bar{x}, y\in\mathcal{O}} \frac{u(y) - u(\bar{x}) - \langle p, y - \bar{x}\rangle}{|y - \bar{x}|} \ge 0 \right\}.$$

LEMMA E.2 *Let $u : \mathcal{O} \to \mathbb{R}$ be an upper semicontinuous function, and $\bar{x} \in \mathcal{O}$. Then $p \in D^+u(\bar{x})$ if and only if there exists a function $\phi \in C^1(\mathbb{R}^N)$ such that $u - \phi$ attains a strict global maximum at $\bar{x}$ and*

$$(\phi(\bar{x}), D\phi(\bar{x})) = (u(\bar{x}), p).$$

Similarly, if $u : \mathcal{O} \to \mathbb{R}$ is a lower semicontinuous function, and $\bar{x} \in \mathcal{O}$, then $p \in D^-u(\bar{x})$ if and only if there exists a function $\phi \in C^1(\mathbb{R}^N)$ such that $u - \phi$ attains a strict global minimum at $\bar{x}$ and

$$(\phi(\bar{x}), D\phi(\bar{x})) = (u(\bar{x}), p).$$

PROOF. See [449], Lemma 2.7, page 173 or [162], page 544. □

We remark that Definition E.1 is exactly the same and Lemma E.2 is true if $\mathbb{R}^N$ is replaced by a real Hilbert space.

DEFINITION E.3 (Second order sub- and superjets) *Let $u : \mathcal{O} \to \mathbb{R}$ be an upper semicontinuous function, and $\bar{x} \in \mathbb{R}^N$. The set*

$$J^{2,+}u(\bar{x}) := \Bigg\{ (p, X) \in \mathbb{R}^N \times S(\mathbb{R}^N) \ : \\ \limsup_{y\to\bar{x}, y\in\mathcal{O}} \frac{u(y) - u(\bar{x}) - \langle p, y - \bar{x}\rangle - \frac{1}{2}\langle X(y - \bar{x}), (y - \bar{x})\rangle}{|y - \bar{x}|^2} \le 0 \Bigg\}$$

is called the second order superjet of u at $\bar{x}$. *Similarly, given a lower semicontinuous function* $u : \mathcal{O} \to \mathbb{R}$, *and* $\bar{x} \in \mathcal{O}$, *the set*

$$J^{2,-}u(\bar{x}) := \Bigg\{ (p, X) \in \mathbb{R}^N \times S(\mathbb{R}^N) \ : \ \liminf_{y \to \bar{x}, y \in \mathcal{O}} \frac{u(y) - u(\bar{x}) - \langle p, y - \bar{x} \rangle - \frac{1}{2} \langle X(y - \bar{x}), (y - \bar{x}) \rangle}{|y - \bar{x}|^2} \geq 0 \Bigg\}$$

is called the second order subjet of u at $\bar{x}$.

Lemma E.4 *Let* $u : \mathcal{O} \to \mathbb{R}$ *be an upper semicontinuous function, and* $\bar{x} \in \mathcal{O}$. *Then* (p, X) *belongs to* $J^{2,+}u(\bar{x})$ *if and only if there exists a function* $\phi \in C^2(\mathbb{R}^N)$ *such that* $u - \phi$ *attains a strict global maximum at* $\bar{x}$ *and*

$$\big(\phi(\bar{x}), D\phi(\bar{x}), D^2\phi(\bar{x})\big) = (u(\bar{x}), p, X).$$

Similarly, if $u : \mathcal{O} \to \mathbb{R}$ *is a lower semicontinuous function, and* $\bar{x} \in \mathcal{O}$, *then* $(p, X) \in J^{2,-}u(\bar{x})$ *if and only if there exists a function* $\phi \in C^2(\mathbb{R}^N)$ *such that* $u - \phi$ *attains a strict global minimum at* $\bar{x}$ *and*

$$\big(\phi(\bar{x}), D\phi(\bar{x}), D^2\phi(\bar{x})\big) = (u(\bar{x}), p, X).$$

Proof. See [449] Lemma 5.4 page 193 or [195], Lemma 4.1, page 211. □

Definition E.5 (Closure of second order sub- and superjets) *Let* $u : \mathcal{O} \to \mathbb{R}$ *be an upper semicontinuous function, and* $\bar{x} \in \mathcal{O}$. *We define*

$$\overline{J}^{2,+}u(\bar{x}) := \Bigg\{ (p, X) \in \mathbb{R}^N \times S(\mathbb{R}^N) \ : \ \textit{there exist } x_n \in \mathcal{O} \ \textit{and } (p_n, X_n) \in J^{2,+}u(x_n) \ \textit{s.t.} \ (x_n, u(x_n), p_n, X_n) \xrightarrow{n \to \infty} (\bar{x}, u(\bar{x}), p, X) \Bigg\}.$$

Similarly, given a lower semicontinuous function $u : \mathcal{O} \to \mathbb{R}$ *and* $\bar{x} \in \mathcal{O}$, *we define*

$$\overline{J}^{2,-}u(\bar{x}) := \Bigg\{ (p, X) \in \mathbb{R}^N \times S(\mathbb{R}^N) \ : \ \textit{there exist } x_n \in \mathcal{O} \ \textit{and } (p_n, X_n) \in J^{2,-}u(x_n) \ \textit{s.t.} \ (x_n, u(x_n), p_n, X_n) \xrightarrow{n \to \infty} (\bar{x}, u(\bar{x}), p, X) \Bigg\}.$$

Remark E.6 Note that the definition is a little different from what one would expect as closures of set-valued mappings. Indeed we also ask $u(x_n) \to u(\bar{x})$. This form of the closures of the semijets was first introduced in [273]. ■

Definition E.7 (Parabolic second order sub- and superjets) *Let* $T > 0$. *Let* $u : (0, T) \times \mathcal{O} \to \mathbb{R}$ *be an upper semicontinuous function, and* $(\bar{t}, \bar{x}) \in (0, T) \times \mathcal{O}$. *The set*

$$\mathcal{P}^{2,+}u(\bar{t}, \bar{x}) := \Bigg\{ (a, p, X) \in \mathbb{R} \times \mathbb{R}^N \times S(\mathbb{R}^N) \ : \ \limsup_{(s,y) \to (\bar{t}, \bar{x})} \frac{u(s, y) - u(\bar{t}, \bar{x}) - a(s - \bar{t}) - \langle p, y - \bar{x} \rangle - \frac{1}{2} \langle X(y - \bar{x}), (y - \bar{x}) \rangle}{|s - \bar{t}| + |y - \bar{x}|^2} \leq 0 \Bigg\}$$

is called the parabolic second order superjet of u at $(\bar{t},\bar{x})$. *Similarly, given a lower semicontinuous function* $u:(0,T)\times\mathcal{O}\to\mathbb{R}$, *and* $(\bar{t},\bar{x})\in(0,T)\times\mathcal{O}$, *the set*

$$\mathcal{P}^{2,-}u(\bar{t},\bar{x}):=\Big\{(a,p,X)\in\mathbb{R}\times\mathbb{R}^N\times S(\mathbb{R}^N)\ :$$
$$\liminf_{(s,y)\to(\bar{t},\bar{x})}\frac{u(s,y)-u(\bar{t},\bar{x})-a(s-\bar{t})-\langle p,y-\bar{x}\rangle-\frac{1}{2}\langle X(y-\bar{x}),(y-\bar{x})\rangle}{|s-\bar{t}|+|y-\bar{x}|^2}\geq 0\Big\}$$

is called the parabolic second order subjet of u at $(\bar{t},\bar{x})$.

The closures of the parabolic second order sub- and superjets $\overline{\mathcal{P}}^{2,-}u(\bar{t},\bar{x}),\overline{\mathcal{P}}^{2,+}u(\bar{t},\bar{x})$ are defined in the same way as $\overline{J}^{2,-}u(\bar{x}),\overline{J}^{2,+}u(\bar{x})$. A parabolic analogue of Lemma E.4 is also true, see [195], Lemma 4.1, page 211.

E.2. Definition of viscosity solution

$\mathcal{O}$ be an open subset of $\mathbb{R}^N$. Consider an equation

$$F(x,u,Du,D^2u)=0\quad\text{in }\mathcal{O},\tag{E.1}$$

where $F:\mathcal{O}\times\mathbb{R}\times\mathbb{R}^N\times S(\mathbb{R}^N)\to\mathbb{R}$ is continuous, non-decreasing in the second variable, and degenerate elliptic, i.e. for every $(x,r,p)\in\mathcal{O}\times\mathbb{R}\times\mathbb{R}^N$, and $X,Y\in S(\mathbb{R}^N)$,

$$F(x,r,p,X)\leq F(x,r,p,Y)\quad\text{if }X\geq Y.$$

Definition E.8 *An upper semi-continuous function $u:\mathcal{O}\to\mathbb{R}$ is a viscosity subsolution of (E.1) if*

$$F(x,u(x),p,X)\leq 0\quad\textit{if } x\in\mathcal{O}\textit{ and }(p,X)\in J^{2,+}u(x).$$

A lower semi-continuous function $u:\mathcal{O}\to\mathbb{R}$ is a viscosity supersolution of (E.1) if

$$F(x,u(x),p,X)\geq 0\quad\textit{if } x\in\mathcal{O}\textit{ and }(p,X)\in J^{2,-}u(x).$$

A viscosity solution of (E.1) is a function which is both a viscosity subsolution and a viscosity supersolution of (E.1).

We remark that since F is continuous, we obtain an equivalent definition if $J^{2,+}u(x),J^{2,-}u(x)$ in Definition E.8 are replaced respectively by $\overline{J}^{2,+}u(x),\overline{J}^{2,-}u(x)$. Moreover, in light of Lemma E.4, Definition E.8 is equivalent to the following definition using test functions.

Definition E.9 *An upper semi-continuous function $u:\mathcal{O}\to\mathbb{R}$ is a viscosity subsolution of (E.1) if whenever $u-\varphi$ has a local maximum at a point $x\in\mathcal{O}$ for a test function $\varphi\in C^2(\mathcal{O})$ then*

$$F(x,u(x),D\varphi(x),D^2\varphi(x))\leq 0.$$

A lower semi-continuous function $u:\mathcal{O}\to\mathbb{R}$ is a viscosity supersolution of (E.1) if whenever $u-\varphi$ has a local minimum at a point $x\in\mathcal{O}$ for a test function $\varphi\in C^2(\mathcal{O})$ then

$$F(x,u(x),D\varphi(x),D^2\varphi(x))\geq 0.$$

A viscosity solution of (E.1) is a function which is both a viscosity subsolution and a viscosity supersolution of (E.1).

Viscosity sub/supersolutions and solutions of parabolic initial value problems

$$\begin{cases} u_t+F(t,x,u,Du,D^2u)=0 & \text{in }(0,T)\times\mathcal{O},\\ u(0,x)=g(x) & \text{on }\mathcal{O}\end{cases}\tag{E.2}$$

are defined in the same way if we replace $J^{2,-}u(\bar{x}), J^{2,+}u(\bar{x})$ in Definition E.8 by $\mathcal{P}^{2,-}u(\bar{t},\bar{x}), \mathcal{P}^{2,+}u(\bar{t},\bar{x})$ and use test functions φ which are once continuously differentiable in t and twice continuously differentiable in x on $(0,T)\times\mathcal{O}$ in Definition E.9.

It is often useful to use the notion of a *discontinuous viscosity solution.* A function u is a discontinuous viscosity subsolution if u^* is a viscosity subsolution, and u is a discontinuous viscosity supersolution if u_* is a viscosity supersolution.

E.3. Finite dimensional maximum principles

The following form of the finite-dimensional maximum principle was introduced in [100] and is sometimes referred to as Crandall-Ishii lemma.

THEOREM E.10 (Maximum principle) *Let $N\in\mathbb{N}$ and $\mathcal{O}$ be an open subset of $\mathbb{R}^N$. Let $u_i:\mathcal{O}\to\mathbb{R}$, $i=1,2$, be two upper semicontinuous functions, and $\phi\in C^2(\mathcal{O}\times\mathcal{O})$. Set, for $x=(x_1,x_2)\in\mathcal{O}\times\mathcal{O}$,*

$$w(x):=u_1(x_1)+u_2(x_2).$$

Suppose that $w-\phi$ has a local maximum at $\bar{x}=(\bar{x}_1,\bar{x}_2)\in\mathcal{O}$. Then, for each $\varepsilon>0$, there exist $X_i\in S(\mathbb{R}^N)$ such that

$$(D_{x_i}\phi(\bar{x}),X_i)\in\overline{J}^{2,+}u_i(\bar{x}_i)\qquad \textit{for } i=1,2$$

and

$$-\left(\frac{1}{\varepsilon}+\left\|D^2\phi(\bar{x})\right\|\right)I\le\begin{pmatrix}X_1&0\\0&X_2\end{pmatrix}\le D^2\phi(\bar{x})+\varepsilon\left(D^2\phi(\bar{x})\right)^2.$$

PROOF. Theorem E.10 is a particular case of Theorem 3.2 of [101]. Its proof is given in the appendix of [101]. □

The following is a parabolic version of Theorem E.10 and is taken from [100], see also [101], Theorem 8.2 or [195], Theorem 6.1, page 216.

THEOREM E.11 (Parabolic maximum principle) *Let $T>0, N\in\mathbb{N}$, and $\mathcal{O}$ be an open subset of $\mathbb{R}^N$. Let $u_i:(0,T)\times\mathcal{O}\to\mathbb{R}$, $i=1,2$ be two upper semicontinuous functions, and $\phi:(0,T)\times\mathcal{O}\times\mathcal{O}\to\mathbb{R}$ be once continuously differentiable in t and twice continuously differentiable in $x=(x_1,x_2)\in\mathbb{R}^{2N}$. Set, for $(t,x)=(t,x_1,x_2)\in(0,T)\times\mathcal{O}\times\mathcal{O}$,*

$$w(t,x):=u_1(t,x_1)+u_2(t,x_2).$$

Suppose that $w-\phi$ has a local maximum at $(\bar{t},\bar{x})=(\bar{t},\bar{x}_1,\bar{x}_2)\in(0,T)\times\mathbb{R}^{2N}$. Assume moreover that there is $r>0$ such that for every $M>0$ there is $C>0$ such that for $i=1,2$

$$\begin{cases} b_i\le C \textit{ whenever } (b_i,p_i,X_i)\in\mathcal{P}^{2,+}u_i(t,x_i),\\ |x_i-\bar{x}_i|+|t-\bar{t}|\le r \textit{ and } |u_i(t,x_i)|+|p_i|+\|X_i\|\le M.\end{cases}\tag{E.3}$$

Then, for each $\varepsilon>0$, there exist $b_i\in\mathbb{R}, X_i\in S(\mathbb{R}^N)$ such that

$$(b_i,D_{x_i}\phi(\bar{t},\bar{x}),X_i)\in\overline{\mathcal{P}}^{2,+}u_i(\bar{t},\bar{x}_i)\quad \textit{for } i=1,2,\quad b_1+b_2=\varphi_t(\bar{t},\bar{x}),$$

and

$$-\left(\frac{1}{\varepsilon}+\left\|D^2\phi(\bar{t},\bar{x})\right\|\right)I\le\begin{pmatrix}X_1&0\\0&X_2\end{pmatrix}\le D^2\phi(\bar{t},\bar{x})+\varepsilon\left(D^2\phi(\bar{t},\bar{x})\right)^2.$$

We remark that the somewhat strange looking condition E.3 is satisfied if u_1 is a viscosity subsolution and $-u_2$ is a viscosity supersolution of a parabolic equation.

E.4. Perron's method

Perron's method is an easy and very general procedure to obtain existence of viscosity solutions. Consider a parabolic initial value problem (E.2), where $\mathcal{O} = \mathbb{R}^N$ and $F : [0, T] \times \mathbb{R}^N \times \mathbb{R} \times \mathbb{R}^N \times S(\mathbb{R}^N) \to \mathbb{R}$ is continuous, non-decreasing in the third variable, and degenerate elliptic. This is the only case that will be used in this book.

Suppose that we have a viscosity supersolution $\overline{u}$ of (E.2) and a viscosity subsolution $\underline{u}$ of (E.2) such that $\underline{u} \leq \overline{u}$ and $\underline{u}(0, x) = \overline{u}(0, x) = g(x)$. Suppose moreover that the equation satisfies the following comparison property: If u is a viscosity subsolution of (E.2) and v is a viscosity supersolution of (E.2) such that $\underline{u}_* \leq u, v \leq \overline{u}^*$, then $u \leq v$. We then have the following theorem. Its proof follows standard arguments, see for instance [101], Section 4, pages 22-24.

THEOREM E.12 (Perron's method) *If the assumptions of this subsection are satisfied then the function*

$$w(t, x) = \sup\{u(t, x) : \underline{u} \leq u \leq \overline{u},\ u \textit{ is a viscosity subsolution of } (E.2)\}$$

is a viscosity solution of (E.2).

We remark that when applying Perron's method it is often more convenient to use the notion of discontinuous viscosity solution. The comparison property is then not needed and one always has that the function w, defined as the supremum of discontinuous viscosity subsolutions u such that $\underline{u} \leq u \leq \overline{u}$ is a discontinuous viscosity solution.

Bibliography

1. R. A. Adams, *Sobolev spaces*, Pure and Applied Mathematics, no. 62, Academic Press, New York, 1975.
2. S. Agmon, *On the eigenfunctions and on the eigenvalues of general elliptic boundary value problems*, Comm. Pure Appl. Math. **15** (1962), no. 2, 119–147.
3. N. U. Ahmed, *Optimal control of ∞-dimensional stochastic systems via generalized solutions of HJB equations*, Discuss. Math. Differ. Incl. Control Optim. **21** (2001), no. 1, 97–126.
4. ———, *Generalized solutions of HJB equations applied to stochastic control on Hilbert space*, Nonlinear Anal. **54** (2003), no. 3, 495–523.
5. R. Aid, S. Federico, H. Pham, and B. Villeneuve, *Explicit investment rules with time-to-build and uncertainty*, J. Econom. Dynam. Control **51** (2015), 240–256.
6. E. Alòs and S. Bonaccorsi, *Stochastic partial differential equations with Dirichlet white-noise boundary conditions*, Ann. Inst. H. Poincaré Probab. Statist. **38** (2002), no. 2, 125–154.
7. A. Ambrosetti and G. Prodi, *A primer of nonlinear analysis*, Cambridge Studies in Advanced Mathematics, vol. 34, Cambridge University Press, Cambridge, 1995, Corrected reprint of the 1993 original.
8. D. Applebaum, *On the infinitesimal generators of Ornstein-Uhlenbeck processes with jumps in Hilbert space*, Potential Anal. **26** (2007), no. 1, 79–100.
9. M. Arisawa, H. Ishii, and P. L. Lions, *A characterization of the existence of solutions for Hamilton-Jacobi equations in ergodic control problems with applications*, Appl. Math. Optim. **42** (2000), no. 1, 35–50.
10. L. Arnold, *Mathematical models of chemical reactions*, Stochastic systems: the mathematics of filtering and identification and applications, Reidel, Dordrecht, 1981, pp. 111–134.
11. N. Aronszajn and P. Panitchpakdi, *Extension of uniformly continuous transformations and hyperconvex metric spaces*, Pacific J. Math. **6** (1956), 405–439.
12. R. B. Ash, *Probability and measure theory*, second ed., Harcourt/Academic Press, Burlington, 2000.
13. J.-P. Aubin and H. Frankowska, *Set-valued analysis*, Modern Birkhäuser Classics, Birkhäuser Boston, Inc., Boston, MA, 2009, Reprint of the 1990 edition.
14. M. Avellaneda, A. Levy, and A. Paras, *Pricing and hedging derivative securities in markets with uncertain volatilities*, Appl. Math. Finance **2** (1995), no. 2, 73–88.
15. F. Baghéry, I. Turpin, and Y. Ouknine, *Some remark on optimal stochastic control with partial information*, Stoch. Anal. Appl. **23** (2005), no. 6, 1305–1320.
16. S. Bahlali, *Necessary and sufficient optimality conditions for relaxed and strict control problems*, SIAM J. Control Optim. **47** (2008), no. 4, 2078–2095.
17. A. V. Balakrishnan, *Applied functional analysis*, second ed., Applications of Mathematics, vol. 3, Springer, New York, 1981.
18. M. Bambi, *Endogenous growth and time-to-build: the AK case*, J. Econom. Dynam. Control **32** (2008), no. 4, 1015–1040.
19. M. Bambi, G. Fabbri, and F. Gozzi, *Optimal policy and consumption smoothing effects in the time-to-build AK model*, Econ. Theor. **50** (2012), no. 3, 635–669.
20. V. Barbu, *Nonlinear differential equations of monotone types in Banach spaces*, Springer Monographs in Mathematics, Springer, Berlin, 2010.

21. V. Barbu and G. Da Prato, *Global existence for the Hamilton-Jacobi equations in Hilbert space*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **8** (1981), no. 2, 257–284.
22. ———, *A direct method for studying the dynamic programming equation for controlled diffusion processes in Hilbert spaces*, Numer. Funct. Anal. Optim. **4** (1981/82), no. 1, 23–43.
23. ———, *Hamilton-Jacobi equations in Hilbert spaces*, Research Notes in Mathematics, vol. 86, Pitman, Boston, 1983.
24. ———, *Solution of the Bellman equation associated with an infinite-dimensional stochastic control problem and synthesis of optimal control*, SIAM J. Control Optim. **21** (1983), no. 4, 531–550.
25. ———, *The Kolmogorov equation for a 2D-Navier-Stokes stochastic flow in a channel*, Nonlinear Anal. **69** (2008), no. 3, 940–949.
26. V. Barbu, G. Da Prato, and A. Debussche, *The Kolmogorov equation associated to the stochastic Navier-Stokes equations in 2D*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **7** (2004), no. 2, 163–182.
27. V. Barbu, G. Da Prato, and C. Popa, *Existence and uniqueness of the dynamic programming equation in Hilbert space*, Nonlinear Anal. **7** (1983), no. 3, 283–299.
28. V. Barbu, G. Da Prato, and L. Tubaro, *Kolmogorov equation associated to the stochastic reflection problem on a smooth convex set of a Hilbert space*, Ann. Probab. **37** (2009), no. 4, 1427–1458.
29. ———, *Kolmogorov equation associated to the stochastic reflection problem on a smooth convex set of a Hilbert space II*, Ann. Inst. H. Poincaré Probab. Statist. **47** (2011), no. 3, 699–724.
30. M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Systems and Control: Foundations and Applications, Birkhäuser Boston, Boston, 1997.
31. G. Barles, *Solutions de viscosité des équations de Hamilton-Jacobi*, Mathématiques and Applications, vol. 17, Springer, Paris, 1994.
32. A. Bátkai and S. Piazzera, *Semigroups for delay equations*, Research Notes in Mathematics, vol. 10, Peters, Wellesley, 2005.
33. A. Bensoussan, *Stochastic maximum principle for distributed parameter systems*, J. Franklin Inst. **315** (1983), no. 5, 387–406.
34. ———, *Stochastic control of partially observable systems*, Cambridge University Press, Cambridge, 1992.
35. A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter, *Representation and control of infinite dimensional systems*, second ed., Systems and Control: Foundations and Applications, Birkhäuser, Boston, 2007.
36. A. Bensoussan, J. Frehse, and S. C. P. Yam, *On the interpretation of the Master equation*, Preprint arXiv:1503.07754, 2015.
37. ———, *The Master equation in mean field theory*, J. Math. Pures Appl. **103** (2015), no. 6, 1441–1474.
38. A. Bensoussan and J.-L. Lions, *Applications of variational inequalities in stochastic control*, Studies in Mathematics and its Applications, vol. 12, North-Holland, Amsterdam, 1982.
39. A. Bensoussan and R. Temam, *Équations stochastiques du type Navier-Stokes*, J. Funct. Anal. **13** (1973), no. 2, 195–222.
40. D. P. Bertsekas, *Dynamic programming and optimal control. Vol. 1*, Athena Scientific, Belmont, 1995.
41. ———, *Dynamic programming and optimal control. Vol. 2*, Athena Scientific, Belmont, 1995.
42. P. Billingsley, *Probability and measure*, third ed., Wiley Series in Probability and Mathematical Statistics, Wiley, New York, 1995.
43. J.-M. Bismut, *Martingales, the Malliavin calculus and hypoellipticity under general Hörmander's conditions*, Probab. Theory Related Fields **56** (1981), no. 4, 469–505.
44. V. I. Bogachev, *Measure theory. Vol. I and II*, Springer, Berlin, 2007.

45. S. Bonaccorsi and M. Fuhrman, *Regularity results for infinite dimensional diffusions. a Malliavin calculus approach*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **10** (1999), no. 1, 35–45.
46. S. Bonaccorsi and G. Guatteri, *Stochastic partial differential equations in bounded domains with Dirichlet boundary conditions*, Stoch. Stoch. Rep. **74** (2002), no. 1-2, 349–370.
47. V. S. Borkar, *Optimal control of diffusion processes*, Wiley, New York, 1989.
48. J. M. Borwein and D. Preiss, *A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions*, Trans. Amer. Math. Soc. **303** (1987), no. 2, 517–527.
49. B. Bouchard, N.-M. Dang, and C.-A. Lehalle, *Optimal control of trading algorithms: a general impulse control approach*, SIAM J. Financial Math **2** (2011), no. 1, 404–438.
50. B. Bouchard and M. Nutz, *Weak dynamic programming for generalized state constraints*, SIAM J. Control Optim. **50** (2012), no. 6, 3344–3373.
51. B. Bouchard and N. Touzi, *Weak dynamic programming principle for viscosity solutions*, SIAM J. Control Optim. **49** (2011), no. 3, 948–962.
52. N. Bourbaki, *Éléments de mathématique. Intégration. Chapitres 1-4*, Springer, Berlin, 2007.
53. A. J. V. Brandão, E. Fernández-Cara, P. M. D. Magalhães, and M. A. Rojas-Medar, *Theoretical analysis and control results for the FitzHugh-Nagumo equation*, Electron. J. Differential Equations (2008), Paper No. 164, 20 pp., electronic only.
54. P. Briand and F. Confortola, *BSDEs with stochastic Lipschitz condition and quadratic PDEs in Hilbert spaces*, Stochastic Process. Appl. **118** (2008), no. 5, 818–838.
55. ———, *Quadratic BSDEs with random terminal time and elliptic PDEs in infinite dimension*, Electron. J. Probab. **13** (2008), Paper no. 54, 1529–1561.
56. P. Briand, B. Delyon, Y. Hu, É. Pardoux, and L. Stoica, *L^p solutions of backward stochastic differential equations*, Stochastic Process. Appl. **108** (2003), no. 1, 109–129.
57. P. Briand and Y. Hu, *Stability of BSDEs with random terminal time and homogenization of semilinear elliptic PDEs*, J. Funct. Anal. **155** (1998), no. 2, 455–494.
58. ———, *BSDE with quadratic growth and unbounded terminal value*, Probab. Theory Related Fields **136** (2006), no. 4, 604–618.
59. R. Buckdahn, J. Li, S. Peng, and C. Rainer, *Mean-field stochastic differential equations and associated PDEs*, Preprint arXiv:1407.1215, 2014.
60. R. Buckdahn, M. Quincampoix, and G. Tessitore, *Controlled stochastic differential equations under constraints in infinite dimensional spaces*, SIAM J. Control Optim. **47** (2008), no. 1, 218–250.
61. J. M. Burgers, *A mathematical model illustrating the theory of turbulence*, Advances in Applied Mechanics (R. von Mises and T. von Kármán, eds.), Academic Press, New York, 1948, pp. 171–199.
62. ———, *The nonlinear diffusion equation: asymptotic solutions and statistical problems*, Springer, Berlin, 1974.
63. P. Cannarsa and G. Da Prato, *Second-order Hamilton-Jacobi equations in infinite dimensions*, SIAM J. Control Optim. **29** (1991), no. 2, 474–492.
64. ———, *Direct solution of a second order Hamilton-Jacobi equation in Hilbert spaces*, Stochastic partial differential equations and applications, Pitman Research Notes in Mathematics Series, vol. 268, Longman, Harlow, 1992, pp. 72–85.
65. P. Cannarsa and G. Di Blasio, *A direct approach to infinite-dimensional Hamilton-Jacobi equations and applications to convex control with state constraints*, Differential Integral Equations **8** (1995), no. 2, 225–246.
66. P. Cannarsa and H. Frankowska, *Value function and optimality conditions for semilinear control problems*, Appl. Math. Optim. **26** (1992), no. 2, 139–169.
67. P. Cannarsa, F. Gozzi, and H. M. Soner, *A boundary value problem for Hamilton-Jacobi equations in Hilbert spaces*, Appl. Math. Optim. **24** (1991), no. 2, 197–220.
68. ———, *A dynamic programming approach to nonlinear boundary control problems of parabolic type*, J. Funct. Anal. **117** (1993), no. 1, 25–61.

69. P. Cannarsa and C. Sinestrari, *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, Birkhäuser, Boston, MA, 2004.
70. P. Cannarsa and M. E. Tessitore, *Cauchy problem for the dynamic programming equation of boundary control*, Boundary control and variation, Lecture Notes in Pure and Applied Mathematics, vol. 163, Dekker, New York, 1994, pp. 13–26.
71. ______, *Infinite-dimensional Hamilton-Jacobi equations and Dirichlet boundary control problems of parabolic type*, SIAM J. Control Optim. **34** (1996), no. 6, 1831–1847.
72. P. Cardaliaguet, *Notes on mean field games (from P-L. Lions' lectures at Collège de France)*, Available at https://www.ceremade.dauphine.fr/∼cardalia/MFG20130420.pdf, 2013.
73. P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions, *The master equation and the convergence problem in mean field games*, (2015), Preprint arXiv:1509.02505.
74. R. Carmona and F. Delarue, *The master equation for large population equilibriums*, Stochastic Analysis and Applications 2014 (D. Crisan, B. Hambly, and T. Zariphopoulou, eds.), 2014, pp. 77–128.
75. S. Cerrai, *A Hille-Yosida theorem for weakly continuous semigroups*, Semigroup Forum **49** (1994), no. 3, 349–367.
76. ______, *Weakly continuous semigroups in the space of functions with polynomial growth*, Dynam. Syst. Appl **4** (1995), 351–372.
77. ______, *Differentiability of Markov semigroups for stochastic reaction-diffusion equations and applications to control*, Stoch. Proc. appl. **83** (1999), no. 1, 15–37.
78. ______, *Smoothing properties of transition semigroups relative to SDEs with values in Banach spaces*, Probab. Theory Related Fields **113** (1999), no. 1, 85–114.
79. ______, *Optimal control problems for stochastic reaction-diffusion systems with non-Lipschitz coefficients*, SIAM J. Control Optim. **39** (2001), no. 6, 1779–1816.
80. ______, *Second order PDE's in finite and infinite dimension: a probabilistic approach*, Lecture Notes in Mathematics, vol. 1762, Springer, Berlin, 2001.
81. ______, *Stationary Hamilton-Jacobi equations in Hilbert spaces and applications to a stochastic optimal control problem*, SIAM J. Control Optim. **40** (2001), no. 3, 824–852.
82. S. Cerrai and F. Gozzi, *Strong solutions of Cauchy problems associated to weakly continuous semigroups*, Differential Integral Equations **8** (1995), no. 3, 465–486.
83. D. H. Chambers, R. J. Adrian, P. Moin, D. S. Stewart, and H. J. Sung, *Karhunen–Loéve expansion of Burgers' model of turbulence*, Phys. Fluids **31** (1988), 25–73.
84. C. D. Charalambous and J. L. Hibey, *First passage risk-sensitive criterion for stochastic evolutions*, Proceedings of the IEEE American Control Conference, IEEE, 1995, pp. 2449–2450.
85. C. D. Charalambous, D. S. Naidu, and K. L. Moore, *Risk-sensitive control, differential games, and limiting problems in infinite dimensions*, Proceedings of the 33rd IEEE Conference on Decision and Control, IEEE, 1994, pp. 2184–2186.
86. J.-F. Chassagneux, D. Crisan, and F. Delarue, *A probabilistic approach to classical solutions of the Master equation for large population equilibria*, Preprint arXiv:1411.3009v2, 2015.
87. P. Cheridito, H. M. Soner, and N. Touzi, *The multi-dimensional super-replication problem under gamma constraints*, Ann. Inst. H. Poincaré Anal. Non Linéaire **22** (2005), no. 5, 633–666.
88. H. Choi, R. Temam, P. Moin, and J. Kim, *Feedback control for unsteady flow and its application to the stochastic Burgers equation*, J. Fluid Mech. **253** (1993), 509–543.
89. A. Chojnowska-Michalik, *Representation theorem for general stochastic delay equations*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **26** (1978), no. 7, 635–642.
90. ______, *Transition semigroups for stochastic semilinear equations on Hilbert spaces*, Dissertationes Math. (Rozprawy Mat.) **396** (2001), 1–59.
91. A. Chojnowska-Michalik and B. Gołdys, *Existence, uniqueness and invariant measures for stochastic semilinear equations on Hilbert spaces*, Probab. Theory Related Fields **102** (1995), no. 3, 331–356.

92. P.-L. Chow, *Infinite-dimensional Kolmogorov equations in Gauss-Sobolev spaces*, Stochastic Anal. Appl. **14** (1996), no. 3, 257–282.
93. ______, *Stochastic partial differential equations*, Chapman & Hall Applied Mathematics and Nonlinear Science Series, Chapman & Hall, Raton, 2007.
94. P.-L. Chow and J.-L. Menaldi, *Infinite-dimensional Hamilton-Jacobi-Bellman equations in Gauss-Sobolev spaces*, Nonlinear Anal. **29** (1997), no. 4, 415–426.
95. F. H. Clarke, *Functional analysis, calculus of variations and optimal control*, Springer, London, 2013.
96. F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth analysis and control theory.*, Springer, New York, 1998.
97. F. Confortola, *Dissipative backward stochastic differential equations with locally Lipschitz nonlinearity*, Stochastic Process. Appl. **117** (2007), no. 5, 613–628.
98. A. Cosso, S. Federico, F. Gozzi, M. Rosestolato, and N. Touzi, *Path-dependent equations and viscosity solutions in infinite dimension*, Preprint arXiv:1502.05648, 2015.
99. M. G. Crandall, *Semidifferentials, quadratic forms and fully nonlinear elliptic equations of second order*, Ann. Inst. H. Poincaré Anal. Non Linéaire **6** (1989), no. 6, 419–435.
100. M. G. Crandall and H. Ishii, *The maximum principle for semicontinuous functions*, Differ. Integral Equ. **3** (1990), no. 6, 1001–1014.
101. M. G. Crandall, H. Ishii, and P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.) **27** (1992), no. 1, 1–67.
102. M. G. Crandall, M. Kocan, and A. Święch , *On partial sup-convolutions, a lemma of P.-L. Lions and viscosity solutions in Hilbert spaces*, Adv. Math. Sci. Appl. **3** (1993/94), no. Special Issue, 1–15.
103. M. G. Crandall and P.-L. Lions, *Viscosity solutions of Hamilton-Jacobi equations in infinite dimensions. IV. Hamiltonians with unbounded linear terms*, J. Funct. Anal. **90** (1990), no. 2, 237–283.
104. ______, *Viscosity solutions of Hamilton-Jacobi equations in infinite dimensions. V. Unbounded linear terms and B-continuous solutions*, J. Funct. Anal. **97** (1991), no. 2, 417–465.
105. ______, *Hamilton-Jacobi equations in infinite dimensions. VI. Nonlinear A and Tataru's method refined*, Evolution equations, control theory, and biomathematics, Lecture Notes in Pure and Applied Mathematics, vol. 155, Dekker, New York, 1994, pp. 51–89.
106. ______, *Viscosity solutions of Hamilton-Jacobi equations in infinite dimensions. VII. The HJB equation is not always satisfied*, J. Funct. Anal. **125** (1994), no. 1, 111–148.
107. A. Cretarola, F. Gozzi, H. Pham, and P. Tankov, *Optimal consumption policies in illiquid markets*, Finance Stoch. **15** (2011), no. 1, 85–115.
108. G. Da Prato, *Applications croissantes et équations d'évolution dans les espaces de Banach*, Academic Press, New York, 1976.
109. ______, *Some results on Bellman equation in Hilbert spaces*, SIAM J. Control Optim. **23** (1985), no. 1, 61–71.
110. ______, *An introduction to infinite-dimensional analysis*, Universitext, Springer, Berlin, 2006.
111. ______, *Introduction to stochastic analysis and Malliavin calculus*, Edizioni della Normale, Pisa, 2014.
112. G. Da Prato and A. Debussche, *Differentiability of the transition semigroup of the stochastic Burgers equation, and application to the corresponding Hamilton-Jacobi equation*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **9** (1998), no. 4, 267–277.
113. ______, *Control of the stochastic Burgers model of turbulence*, SIAM J. Control Optim. **37** (1999), no. 4, 1123–1149.
114. ______, *Dynamic programming for the stochastic Burgers equation*, Ann. Mat. Pura Appl. **178** (2000), no. 1, 143–174.
115. ______, *Dynamic programming for the stochastic Navier-Stokes equations*, M2AN Math. Model. Numer. Anal. **34** (2000), no. 2, 459–475.

116. ———, *Two-Dimensional Navier-Stokes Equations Driven by a Space-Time White Noise*, J. Funct. Anal. **196** (2002), no. 1, 180–210.
117. ———, *Ergodicity for the 3D stochastic Navier-Stokes equations*, J. Math Pure. Appl. **82** (2003), no. 8, 877–947.
118. ———, *m-dissipativity of Kolmogorov operators corresponding to Burgers equations with space-time white noise*, Potential Anal. **26** (2007), no. 1, 31–55.
119. G. Da Prato, A. Debussche, and R. Temam, *Stochastic Burgers' equation*, NoDEA Nonlinear Differential Equations Appl. **1** (1994), no. 4, 389–402.
120. G. Da Prato, B. Goldys, and J. Zabczyk, *Ornstein-Uhlenbeck semigroups in open sets of Hilbert spaces*, C. R. Acad. Sci. Paris Sér. I Math. **325** (1997), no. 4, 433–438.
121. G. Da Prato and A. Lunardi, *On the Dirichlet semigroup for Ornstein-Uhlenbeck operators in subsets of Hilbert spaces*, J. Funct. Anal. **259** (2010), no. 10, 2642–2672.
122. ———, *Maximal L^2 regularity for Dirichlet problems in Hilbert spaces*, J. Math. Pures Appl. **99** (2013), no. 6, 741–765.
123. ———, *Maximal Sobolev regularity in Neumann problems for gradient systems in infinite dimensional domains*, Preprint arXiv:1309.6519, 2013.
124. G. Da Prato and E. Sinestrari, *Differential operators with nondense domain*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **14** (1987), no. 2, 285–344.
125. G. Da Prato and J. Zabczyk, *Evolution equations with white-noise boundary conditions*, Stochastics Stochastics Rep. **42** (1993), no. 3-4, 167–182.
126. ———, *Regular densities of invariant measures in Hilbert spaces*, J. Funct. Anal. **130** (1995), no. 2, 427–449.
127. ———, *Ergodicity for infinite-dimensional systems*, London Mathematical Society Lecture Note Series, vol. 229, Cambridge University Press, Cambridge, 1996.
128. ———, *Differentiability of the Feynman-Kac semigroup and a control application*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **8** (1997), no. 3, 183–188.
129. ———, *Second order partial differential equations in Hilbert spaces*, London Mathematical Society Lecture Note Series, vol. 293, Cambridge University Press, Cambridge, 2002.
130. ———, *Stochastic equations in infinite dimensions*, vol. 152, Cambridge university press, 2014.
131. A. Debussche, M. Fuhrman, and G. Tessitore, *Optimal control of a stochastic heat equation with boundary-noise and boundary-control*, ESAIM Control Optim. Calc. Var. **13** (2007), no. 1, 178–205.
132. K. Deimling, *Nonlinear functional analysis*, Springer, Berlin, 1985.
133. R. Deville, G. Godefroy, and V. Zizler, *A smooth variational principle with applications to Hamilton-Jacobi equations in infinite dimensions*, J. Funct. Anal. **111** (1993), no. 1, 197–212.
134. G. Di Blasio, *Global solutions for a class of Hamilton-Jacobi equations in Hilbert spaces*, Numer. Funct. Anal. Optim. **8** (1986), no. 3-4, 261–300.
135. ———, *Optimal control with infinite horizon for distributed parameter systems with constrained controls*, SIAM J. Control Optim. **29** (1991), no. 4, 909–925.
136. C. Di Girolami and F. Gozzi, *Solutions of second order HJB equations in Hilbert spaces via smoothing property*, working paper.
137. J. Diestel and J. J. Uhl, *Vector measures*, Mathematical Surveys and Monographs, vol. 15, American Mathematical Society, Providence, 1977.
138. N. Dinculeanu, *Integration on locally compact spaces*, Springer, 1974.
139. ———, *Vector integration and stochastic integration in Banach spaces*, Pure and Applied Mathematics, Wiley, New York, 2000.
140. J. L. Doob, *Measure theory*, Graduate Texts in Mathematics, vol. 143, Springer, New York, 1994.
141. K. Du and Q. Meng, *A general maximum principle for optimal control of stochastic evolution equations*, Preprint arXiv:1206.3649, 2012.
142. ———, *A maximum principle for optimal control of stochastic evolution equations*, SIAM J. Control Optim. **51** (2013), no. 6, 4343–4362.

143. E. Duncan, *Probability densities for diffusion processes*, Tech. Report Stanford Electronics Labs. 7001-4, University of California, Stanford, May 1967.
144. N. Dunford and J. T. Schwartz, *Linear operators. Part I*, Interscience, New York, 1958.
145. P. Dupuis and W. M. McEneaney, *Risk-sensitive and robust escape criteria*, SIAM J. Control Optim. **35** (1997), no. 6, 2021–2049.
146. W. E, K. Khanin, A. Mazel, and Y. Sinai, *Invariant measures for Burgers equation with stochastic forcing*, Ann. of Math. **151** (2000), no. 3, 877–960.
147. I. Ekeland and G. Lebourg, *Generic Frechet-differentiability and perturbed optimization problems in Banach spaces*, Trans. Amer. Math. Soc. **224** (1976), no. 2, 193–216.
148. I. Ekren, C. Keller, N. Touzi, and J. Zhang, *On viscosity solutions of path dependent PDEs*, Ann. Probab. **42** (2014), no. 1, 204–236.
149. N. El Karoui, *Les aspects probabilistes du contrôle stochastique*, École d'été de probabilités de Saint-Flour IX-1979, Lecture Notes in Mathematics, vol. 876, Springer, Berlin, 1981, pp. 74–238.
150. ———, *Backward stochastic differential equations: a general introduction*, Backward stochastic differential equations (Paris, 1995–1996), Pitman Res. Notes Math. Ser., vol. 364, Longman, Harlow, 1997, pp. 7–26.
151. N. El Karoui, M. Jeanblanc-Picqué, and S. E. Shreve, *Robustness of the Black and Scholes formula*, Math. Finance **8** (1998), no. 2, 93–126.
152. N. El Karoui and L. Mazliak (eds.), *Backward stochastic differential equations*, Pitman Research Notes in Mathematics Series, vol. 364, Longman, Harlow, 1997.
153. N. El Karoui, D. Nguyen, and M. Jeanblanc-Picqué, *Compactification methods in the control of degenerate diffusions: existence of an optimal control*, Stochastics **20** (1987), no. 3, 169–219.
154. N. El Karoui, S. Peng, and M. C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance **7** (1997), no. 1, 1–71.
155. N. El Karoui and X. Tan, *Capacities, measurable selection and dynamic programming part ii: Application in stochastic control problems*, Preprint arXiv:1310.3364, 2013.
156. R. J. Elliott, *Stochastic calculus and applications*, Wiley Classics Library, Springer, Berlin, 1982.
157. ———, *Filtering and control for point process observations*, Recent Advances in Stochastic Calculus (J. Baras and V. Mirelli, eds.), Progress in Automation and Information Systems, vol. 91, Springer, Berlin, 1990, pp. 1–27.
158. I. Elsanosi, B. Øksendal, and A. Sulem, *Some solvable stochastic control problems with delay*, Stochastics Stochastics Rep. **71** (2000), no. 1-2, 69–89.
159. K. D. Elworthy and X.-M. Li, *Formulae for the derivatives of heat semigroups*, J. Funct. Anal. **125** (1994), no. 1, 252–286.
160. K. J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*, Graduate Texts in Mathematics, vol. 194, Springer, Berlin, 1999.
161. S. N. Ethier and T. G. Kurtz, *Markov processes. Characterization and convergence*, Wiley Series in Probability and Statistics, Wiley, New York, 1986.
162. L. C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, 1998.
163. G. Fabbri, *First order HJB equations in Hilbert spaces and applications*, Ph.D. thesis, Università di Roma - La Sapienza, 2006.
164. ———, *A viscosity solution approach to the infinite-dimensional HJB equation related to a boundary control problem in a transport equation*, SIAM J. Control Optim. **47** (2008), no. 2, 1022–1052.
165. G. Fabbri and B. Goldys, *An LQ problem for the heat equation on the halfline with dirichlet boundary control and noise*, SIAM J. Control Optim. **48** (2009), no. 6, 1473–1488.
166. G. Fabbri, F. Gozzi, and A. Święch , *Verification theorem and construction of ε-optimal controls for control of abstract evolution equations*, J. Convex Anal. **17** (2010), no. 2, 611–642.

167. S. Faggian, *Boundary-control problems with convex cost and dynamic programming in infinite dimension. I. The maximum principle*, Differential Integral Equations **17** (2004), no. 9-10, 1149–1174.
168. ______, *Boundary control problems with convex cost and dynamic programming in infinite dimension. II. Existence for HJB*, Discrete Contin. Dyn. Syst. **12** (2005), no. 2, 323–346.
169. ______, *Regular solutions of first-order Hamilton-Jacobi equations for boundary control problems and applications to economics*, Appl. Math. Optim. **51** (2005), no. 2, 123–162.
170. ______, *Hamilton-Jacobi equations arising from boundary control problems with state constraints*, SIAM J. Control Optim. **47** (2008), no. 4, 2157–2178.
171. S. Faggian and F. Gozzi, *Optimal investment models with vintage capital: Dynamic programming approach*, J. Math. Econ. **46** (2010), no. 4, 416–437.
172. S. Federico, *Stochastic optimal control problems for pension funds management*, Ph.D. thesis, Scuola Normale Superiore, Pisa, 2009.
173. ______, *A stochastic control problem with delay arising in a pension fund model*, Finance Stoch. **15** (2011), no. 3, 421–459.
174. S. Federico, P. Gassiat, and F. Gozzi, *Impact of time illiquidity in a mixed market without full observation*, Preprint arXiv:1211.1285. To appear on *Math. Finance*, 2012.
175. S. Federico, B. Goldys, and F. Gozzi, *HJB equations for the optimal control of differential equations with delays and state constraints, I: regularity of viscosity solutions*, SIAM J. Control Optim. **48** (2010), no. 8, 4910–4937.
176. ______, *HJB equations for the optimal control of differential equations with delays and state constraints, II: verification and optimal feedbacks*, SIAM J. Control Optim. **49** (2011), no. 6, 2378–2414.
177. S. Federico and H. Pham, *Characterization of the optimal boundaries in reversible investment problems*, SIAM J. Control Optim. **52** (2014), no. 4, 2180–2223.
178. S. Federico and E. Tacconi, *Dynamic programming for optimal control problems with delays in the control variable*, SIAM J. Control Optim. **52** (2014), no. 2, 1203–1236.
179. S. Federico and P. Tankov, *Finite-dimensional representations for controlled diffusions with delay*, Appl. Math. Optim. **71** (2015), no. 1, 165–194.
180. J. Feng, *Martingale problems for large deviations of Markov processes*, Stochastic Process. Appl. **81** (1999), no. 2, 165–216.
181. ______, *Large deviation for a stochastic Cahn-Hilliard equation*, Methods Funct. Anal. Topology **9** (2003), no. 4, 333–356.
182. ______, *Large deviation for diffusions and Hamilton-Jacobi equation in Hilbert spaces*, Ann. Probab. **34** (2006), no. 1, 321–385.
183. J. Feng and M. Katsoulakis, *A comparison principle for Hamilton-Jacobi equations related to controlled gradient flows in infinite dimensions*, Arch. Rational Mech. Anal. **192** (2009), no. 2, 275–310.
184. J. Feng and T. G. Kurtz, *Large deviations for stochastic processes*, Mathematical Surveys and Monographs, vol. 131, American Mathematical Society, Providence, 2006.
185. J. Feng and T. Nguyen, *Hamilton-Jacobi equations in space of measures associated with a system of conservation laws*, J. Math. Pures Appl. **97** (2012), no. 4, 318–390.
186. J. Feng and A. Święch , *Optimal control for a mixed flow of Hamiltonian and gradient type in space of probability measures (with Appendix B by Atanas Stefanov)*, Trans. Am. Math. Soc. **365** (2013), no. 8, 3987–4039.
187. F. Flandoli, *An introduction to 3D stochastic fluid dynamics*, SPDE in hydrodynamic: recent progress and prospects, Lecture Notes in Mathematics, vol. 1942, Springer, Berlin, 2008, pp. 51–150.
188. F. Flandoli and F. Gozzi, *Kolmogorov equation associated to a stochastic Navier-Stokes equation*, J. Funct. Anal. **160** (1998), no. 1, 312–336.
189. F. Flandoli and G. Zanco, *An infinite-dimensional approach to path-dependent kolmogorov's equations*, Preprint arXiv:1312.6165, 2013.

190. W. H. Fleming, *Nonlinear semigroup for controlled partially observed diffusions*, SIAM J. Control Optim. **20** (1982), no. 2, 286–301.
191. W. H. Fleming and M. Nisio, *On the existence of optimal stochastic controls*, Indiana Univ. Math. J. **15** (1966), 777–794.
192. ______, *Differential games for stochastic partial differential equations*, Nagoya Math. J. **131** (1993), 75–107.
193. W. H. Fleming and É. Pardoux, *Optimal control of partially observed diffusions*, SIAM J. Control Optim. **20** (1982), no. 2, 261–285.
194. W. H. Fleming and R. W. Rishel, *Deterministic and stochastic optimal control*, Applications of Mathematics, vol. 1, Springer, Berlin, 1975.
195. W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, second ed., Stochastic Modelling and Applied Probability, vol. 25, Springer, New York, 2006.
196. W. H Fleming and P. E. Souganidis, *On the existence of value-functions of 2-player, zero-sum stochastic differential-games*, Indiana Univ. Math. J. **38** (1989), no. 2, 293–314.
197. W. H. Fleming and D. Vermes, *Convex duality approach to the optimal control of diffusions*, SIAM J. Control Optim. **27** (1989), no. 5, 1136–1155.
198. T. M. Flett, *Differential analysis: differentiation, differential equations, and differential inequalities*, Cambridge University Press, Cambridge, 1980.
199. G. B. Folland, *Real analysis*, second ed., Pure and Applied Mathematics, Wiley, New York, 1999.
200. K. O. Friedrichs, *The identity of weak and strong extensions of differential operators*, Trans. Am. Math. Soc. **55** (1944), no. 1, 132–151.
201. K. Frieler and C. Knoche, *Solutions of stochastic differential equations in infinite dimensional Hilbert spaces and their dependence on initial data*, Diploma Thesis, Bielefeld University. BiBoS-Preprint E02-04-083., 2001.
202. M. Fuhrman, *Smoothing properties of nonlinear stochastic equations in Hilbert spaces*, NoDEA Nonlinear Differential Equations Appl. **3** (1996), no. 4, 445–464.
203. ______, *On a class of stochastic equations in hilbert spaces: solvability and smoothing properties*, Stoch. Anal. Appl. **17** (1999), no. 1, 43–69.
204. M. Fuhrman and Y. Hu, *Infinite horizon BSDEs in infinite dimensions with continuous driver and applications*, J. Evol. Equ. **6** (2006), no. 3, 459–484.
205. M. Fuhrman, Y. Hu, and G. Tessitore, *Stochastic control and BSDEs with quadratic growth*, Control theory and related topics, World Scientific Publishing, Hackensack, 2007, pp. 80–86.
206. ______, *Ergodic BSDES and optimal ergodic control in Banach spaces*, SIAM J. Control Optim. **48** (2009), no. 3, 1542–1566.
207. ______, *Stochastic maximum principle for optimal control of SPDEs*, C. R. Acad. Sci. Paris Sér. I Math. **350** (2012), no. 13, 683–688.
208. M. Fuhrman, Y. Hu, and G Tessitore, *Stochastic maximum principle for optimal control of SPDEs*, Appl. Math. Optim. **68** (2013), no. 2, 181–217.
209. M. Fuhrman, F. Masiero, and G. Tessitore, *Stochastic equations with delay: optimal control via BSDEs and regular solutions of Hamilton-Jacobi-Bellman equations*, SIAM J. Control Optim. **48** (2010), no. 7, 4624–4651.
210. M. Fuhrman and G. Tessitore, *The Bismut-Elworthy formula for backward SDEs and applications to nonlinear Kolmogorov equations and control in infinite dimensional spaces*, Stoch. Stoch. Rep. **74** (2002), no. 1-2, 429–464.
211. ______, *Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control*, Ann. Probab. **30** (2002), no. 3, 1397–1465.
212. ______, *Infinite horizon backward stochastic differential equations and elliptic equations in Hilbert spaces*, Ann. Probab. **32** (2004), no. 1B, 607–660.
213. ______, *Generalized directional gradients, backward stochastic differential equations and mild solutions of semilinear parabolic equations*, Appl. Math. Optim. **51** (2005), no. 3, 279–332.

214. T. Funaki, *Random motion of strings and related stochastic evolution equations*, Nagoya Math. J. **89** (1983), 129–193.
215. A. V. Fursikov, *Optimal control of distributed systems. Theory and applications*, Translations of Mathematical Monographs, vol. 187, American Mathematical Society, Providence, 2000, Translated from the Russian.
216. W. Gangbo and A. Święch , *Existence of a solution to an equation arising from the theory of Mean Field Games*, J. Differential Equations **259** (2015), no. 11, 6573–6643.
217. P. Gassiat, F. Gozzi, and H. Pham, *Dynamic programming for an investment/consumption problem in illiquid markets with regime switching*, Stochastic Analysis and Control (A. C. Michalik, S. Peszat, and L. Stettner, eds.), 2014.
218. ———, *Investment/consumption problem in illiquid markets with regimes switching*, SIAM J. Control Optim. **52** (2014), no. 3, 1761–1786.
219. D. Gątarek and A. Święch , *Optimal stopping in Hilbert spaces and pricing of American options*, Math. Methods Oper. Res. **50** (1999), no. 1, 135–147.
220. L. Gawarecki and V. Mandrekar, *Stochastic differential equations in infinite dimensions with applications to stochastic partial differential equations*, Probability and its Applications, Springer, Heidelberg, 2011.
221. H. Geman, N. El Karoui, and J.-C. Rochet, *Changes of numéraire, changes of probability measure and option pricing*, J. Appl. Probab. **32** (1995), no. 2, 443–458.
222. J. A. Goldstein, *Semigroups of linear operators and applications*, Oxford Mathematical Monographs, Oxford University Press, New York, 1985.
223. B. Goldys and F. Gozzi, *Second order parabolic Hamilton-Jacobi-Bellman equations in Hilbert spaces and stochastic control: L^2_μ approach*, Stochastic Process. Appl. **116** (2006), no. 12, 1932–1963.
224. B. Goldys, F. Gozzi, and J. M. A. M. van Neerven, *On closability of directional gradients*, Potential Anal. **18** (2003), no. 4, 289–310.
225. B. Goldys and M. Kocan, *Diffusion semigroups in spaces of continuous functions with mixed topology*, J. Differential Equations **173** (2001), no. 1, 17–39.
226. B. Goldys and B. Maslowski, *Ergodic control of semilinear stochastic equations and the Hamilton-Jacobi equation*, J. Math. Anal. Appl. **234** (1999), no. 2, 592–631.
227. B. Goldys and M. Musiela, *On partial differential equations related to term structure models*, Preprint, Univ. New South Wales, 1996.
228. B. Goldys, M. Musiela, and D. Sondermann, *Lognormality of rates and term structure models*, Stochastic Anal. Appl. **18** (2000), no. 3, 375–396.
229. F. Gozzi, *Some results for an infinite horizon control problem governed by a semilinear state equation*, Control and estimation of distributed parameter systems, International Series of Numerical Mathematics, vol. 91, Birkhäuser, Basel, 1989, pp. 145–163.
230. ———, *Some results for an optimal control problem with semilinear state equation*, SIAM J. Control Optim. **29** (1991), no. 4, 751–768.
231. ———, *Regularity of solutions of a second order Hamilton-Jacobi equation and application to a control problem*, Comm. Partial Differential Equations **20** (1995), no. 5-6, 775–826.
232. ———, *Global regular solutions of second order Hamilton-Jacobi equations in Hilbert spaces with locally Lipschitz nonlinearities*, J. Math. Anal. Appl. **198** (1996), no. 2, 399–443.
233. ———, *Strong solutions for Kolmogorov equation in Hilbert spaces*, Partial differential equation methods in control and shape analysis, Lecture Notes in Pure and Applied Mathematics, vol. 188, Dekker, New York, 1997, pp. 163–187.
234. ———, *Second order Hamilton-Jacobi equations in Hilbert spaces and stochastic optimal control*, Ph.D. thesis, Scuola Normale Superiore, Pisa, 1998.
235. ———, *Second order Hamilton-Jacobi equations in Hilbert spaces and stochastic optimal control*, Stochastic partial differential equations and applications, Lecture Notes in Pure and Applied Mathematics, vol. 227, Dekker, New York, 2002, pp. 255–285.
236. ———, *Smoothing properties of nonlinear transition semigroups: case of Lipschitz nonlinearities*, J. Evol. Equ. **6** (2006), no. 4, 711–743.

237. F. Gozzi and P. Loreti, *Regularity of the minimum time function and minimum energy problems: the linear case*, SIAM J. Control Optim. **37** (1999), no. 4, 1195–1221.
238. F. Gozzi and C. Marinelli, *Stochastic optimal control of delay equations arising in advertising models*, Stochastic partial differential equations and applications VII, Lecture Notes in Pure and Applied Mathematics, vol. 245, Chapman & Hall, Raton, 2006, pp. 133–148.
239. F. Gozzi, C. Marinelli, and S. Savin, *On controlled linear diffusions with delay in a model of optimal advertising under uncertainty with memory effects*, J. Optim. Theory Appl. **142** (2009), no. 2, 291–321.
240. F. Gozzi and F. Masiero, *Stochastic optimal control with delay in the control: solution through partial smoothing*, Preprint arXiv:1506.06013, 2015.
241. F. Gozzi and E. Rouy, *Regular solutions of second-order stationary Hamilton-Jacobi equations*, J. Differential Equations **130** (1996), no. 1, 201–234.
242. F. Gozzi, E. Rouy, and A. Święch , *Second order Hamilton-Jacobi equations in Hilbert spaces and stochastic boundary control*, SIAM J. Control Optim. **38** (2000), no. 2, 400–430.
243. F. Gozzi, S. S. Sritharan, and A. Święch , *Viscosity solutions of dynamic-programming equations for the optimal control of the two-dimensional Navier-Stokes equations*, Arch. Rational Mech. Anal. **163** (2002), no. 4, 295–327.
244. ______ , *Bellman equations associated to the optimal feedback control of stochastic Navier-Stokes equations*, Comm. Pure Appl. Math. **58** (2005), no. 5, 671–700.
245. F. Gozzi and A. Święch , *Hamilton-Jacobi-Bellman equations for the optimal control of the Duncan-Mortensen-Zakai equation*, J. Funct. Anal. **172** (2000), no. 2, 466–510.
246. F. Gozzi, A. Święch , and X. Y. Zhou, *A corrected proof of the stochastic verification theorem within the framework of viscosity solutions*, SIAM J. Control Optim. **43** (2005), no. 6, 2009–2019.
247. ______ , *Erratum: "A corrected proof of the stochastic verification theorem within the framework of viscosity solutions"*, SIAM J. Control Optim. **48** (2010), no. 6, 4177–4179.
248. F. Gozzi and T. Vargiolu, *On the superreplication approach for european interest rates derivatives*, Stochastic Analysis, Random Fields and Applications III (M. Dozzi C. Dalang and F. Russo, eds.), Progress in Probability, vol. 52, Birkhauser, Boston, 2002, pp. 173–188.
249. ______ , *Superreplication of european multiasset derivatives with bounded stochastic volatility*, Math. Methods Oper. Res. **55** (2002), no. 1, 69–91.
250. A. Grorud and É. Pardoux, *Intégrales hilbertiennes anticipantes par rapport à un processus de Wiener cylindrique et calcul stochastique associé*, Appl. Math. Optim. **25** (1992), no. 1, 31–49.
251. L. Gross, *Potential theory on Hilbert space*, J. Funct. Anal. **1** (1967), no. 2, 123–181.
252. G. Guatteri, *On a class of forward-backward stochastic differential systems in infinite dimensions*, J. Appl. Math. Stoch. Anal. (2007), Paper no. 42640, 33 pp.
253. ______ , *Stochastic maximum principle for SPDEs with noise and control on the boundary*, Systems Control Lett. **60** (2011), no. 3, 198–204.
254. U. G. Haussmann and J.-P. Lepeltier, *On the existence of optimal controls*, SIAM J. Control Optim. **28** (1990), no. 4, 851–902.
255. T. Havârneanu, *Existence for the dynamic programming equation of control diffusion processes in Hilbert space*, Nonlinear Anal. **9** (1985), no. 6, 619–629.
256. D. Henry, *Geometric theory of semilinear parabolic equations*, Lecture Notes in Mathematics, vol. 840, Springer, Berlin, 1981.
257. O. Hijab, *Partially observed control of Markov processes I*, Stochastics **28** (1989), no. 2, 123–144.
258. ______ , *Partially observed control of Markov processes II*, Stochastics **28** (1989), no. 3, 247–262.
259. ______ , *Partially observed control of Markov processes III*, Ann. Probab. **18** (1990), no. 3, 1099–1125.
260. ______ , *Infinite-dimensional Hamilton-Jacobi equations with large zeroth-order coefficient*, J. Funct. Anal. **97** (1991), no. 2, 311–326.

261. ______, *Partially observed control of Markov processes, IV*, J. Funct. Anal. **109** (1992), no. 2, 215–256.
262. M. Hinze and S. Volkwein, *Analysis of instantaneous control for the Burgers equation*, Nonlinear Anal. **50** (2002), no. 1, 1–26.
263. Y. Hu and S. Peng, *Maximum principle for semilinear stochastic evolution control systems*, Stoch. Stoch. Rep. **33** (1990), no. 3-4, 159–180.
264. ______, *Adapted solution of a backward semilinear stochastic evolution equation*, Stochastic Anal. Appl. **9** (1991), no. 4, 445–459.
265. Y. Hu and G. Tessitore, *BSDE on an infinite horizon and elliptic PDEs in infinite dimension*, NoDEA Nonlinear Differential Equations Appl. **14** (2007), no. 5-6, 825–846.
266. J. Huang and J. Shi, *Maximum principle for optimal control of fully coupled forward-backward stochastic differential delayed equations*, ESAIM Control Optim. Calc. Var. **18** (2012), no. 4, 1073–1096.
267. N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, second ed., North-Holland Mathematical Library, vol. 24, North-Holland, Amsterdam, 1989.
268. I. Iscoe, M . B. Marcus, D. McDonald, M. Talagrand, and J. Zinn, *Continuity of L^2-valued Ornstein-Uhlenbeck processes*, Ann. Probab. **18** (1990), no. 1, 68–84.
269. H. Ishii, *Perron's method for Hamilton-Jacobi equations*, Duke Math. J. **55** (1987), no. 2, 369–384.
270. ______, *On uniqueness and existence of viscosity solutions of fully nonlinear second-order elliptic PDEs*, Comm. Pure Appl. Math. **42** (1989), no. 1, 15–45.
271. ______, *Viscosity solutions for a class of Hamilton-Jacobi equations in Hilbert spaces*, J. Funct. Anal. **105** (1992), no. 2, 301–341.
272. ______, *Viscosity solutions of nonlinear second-order partial differential equations in Hilbert spaces*, Comm. Partial Differential Equations **18** (1993), no. 3-4, 601–650.
273. H. Ishii and P.-L. Lions, *Viscosity solutions of fully nonlinear second-order elliptic partial differential equations*, J. Differential Equations **83** (1990), no. 1, 26–78.
274. K. Itô and M. Nisio, *On stationary solutions of a stochastic differential equation*, J. Math. Kyoto Univ. **4** (1964), no. 1, 1–75.
275. A. F. Ivanov, Y. I. Kazmerchuk, and A. V. Swishchuk, *Theory, stochastic stability and applications of stochastic delay differential equations: a survey of results*, Differential Equations Dynam. Systems **11** (2003), no. 1-2, 55–115.
276. M. R. James, J. S. Baras, and R. J. Elliott, *Output feedback risk-sensitive control and differential games for continuous-time nonlinear systems*, Proceedings of the 32nd IEEE Conference on Decision and Control, IEEE, 1993, pp. 3357–3360.
277. B. Jefferies, *Weakly integrable semigroups on locally convex spaces*, J. Funct. Anal. **66** (1986), no. 3, 347–364.
278. ______, *The generation of weakly integrable semigroups*, J. Funct. Anal. **73** (1987), no. 1, 195–215.
279. D. T. Jeng, *Forced model equation for turbulence*, Phys. Fluids **12** (1969), no. 10, 2006.
280. R. Jensen, *The maximum principle for viscosity solutions of fully nonlinear second order partial differential equations*, Arch. Rational Mech. Anal. **101** (1988), no. 1, 1–27.
281. O. Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications, Springer, New York, 2002.
282. G. Kallianpur and J. Xiong, *Stochastic differential equations in infinite-dimensional spaces*, Institute of Mathematical Statistics Lecture Notes, vol. 26, Institute of Mathematical Statistics, Hayward, 1995.
283. I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, Graduate Texts in Mathematics, vol. 113, Springer, New York, 1988.
284. M. Kardar, G. Parisi, and Y. C. Zhang, *Dynamic scaling of growing interfaces*, Phys. Rev. Lett. **56** (1986), no. 9, 889–892.
285. T. Kato, *Perturbation theory for linear operators*, second ed., Classics in Mathematics, Springer, Berlin, 1995.

286. D. Kelome, *Viscosity solution of second order equations in a separable Hilbert space and applications to stochastic optimal control*, Ph.D. thesis, Georgia Institute of Technology, 2002.
287. D. Kelome and A. Święch , *Viscosity solutions of an infinite-dimensional Black-Scholes-Barenblatt equation*, Appl. Math. Optim. **47** (2003), no. 3, 253–278.
288. ______ , *Perron's method and the method of relaxed limits for "unbounded" PDE in Hilbert spaces*, Studia Math. **176** (2006), no. 3, 249–277.
289. M. Kobylanski, *Backward stochastic differential equations and partial differential equations with quadratic growth*, Ann. Probab. **28** (2000), no. 2, 558–602.
290. M. Kocan, *Some aspects of the theory of viscosity solutions of fully nonlinear partial differential equations in infinite dimensions*, Ph.D. thesis, University of California, Santa Barbara, 1994.
291. M. Kocan and P. Soravia, *A viscosity approach to infinite-dimensional Hamilton-Jacobi equations arising in optimal control with state constraints*, SIAM J. Control Optim. **36** (1998), no. 4, 1348–1375.
292. M. Kocan and A. Święch , *Second order unbounded parabolic equations in separated form*, Studia Math. **115** (1995), no. 3, 291–310.
293. ______ , *Perturbed optimization on product spaces*, Nonlinear Anal. **26** (1996), no. 1, 81–90.
294. N. V. Krylov, *Controlled diffusion processes*, Applications of Mathematics, vol. 14, Springer, New York, 1980, Translated from the Russian.
295. N. V. Krylov and B. L. Rozovskiĭ, *On the Cauchy problem for linear stochastic partial differential equations*, Math. USSR, Izv. **11** (1977), no. 6, 1267–1284.
296. ______ , *Stochastic evolution equations*, J. Sov. Math. **16** (1981), no. 4, 1233–1277, Original paper in Russian: Sovremennye Problemy Matematiki. Noveishie Dostizheniya **14** (1979): 71–146.
297. ______ , *Stochastic partial differential equations and diffusion processes*, Russ. Math. Surv. **37** (1982), no. 6, 81–105, Original paper in Russian: Uspekhi Mat. Nauk **37** (228) no. 6, 75–95.
298. ______ , *Stochastic evolution equations*, Stochastic Differential Equations: Theory and Applications (P. H. Baxendale and S. V. Lototsky, eds.), Interdisciplinary Mathematical Sciences, vol. 2, World Scientific Publishing, Hackensack, 2007, pp. 1–69.
299. S. Kuksin and A. Shirikyan, *Mathematics of 2D statistical hydrodynamics*, preliminary version.
300. H. H. Kuo, *Gaussian measures in Hilbert spaces*, Lecture Notes in Mathematics, vol. 463, Springer, Berlin, 1975.
301. T. G. Kurtz, *Martingale problems for controlled processes*, Stochastic modelling and filtering, Lecture Notes in Control and Information Sciences, vol. 91, Springer, Berlin, 1987, pp. 75–90.
302. J. Kurzweil, *On approximation in real Banach spaces*, Studia Math. **14** (1954), no. 2, 214–231.
303. H. J. Kushner, *On the dynamical equations of conditional probability density functions, with applications to optimal stochastic control theory*, J. Math. Anal. Appl. **8** (1964), no. 2, 332–344.
304. ______ , *Stochastic stability and control*, Mathematics in Science and Engineering, vol. 33, North-Holland, Amsterdam, 1967.
305. O. A. Ladyzhenskaya, *On integral inequalities, the convergence of approximate methods, and the solution of linear elliptic operators*, Vest. Leningrad State Univ. **7** (1958), 60–69.
306. S. Lang, *Real and functional analysis*, Graduate Texts in Mathematics, vol. 142, Springer, New York, 1993.
307. B. Larssen and N. H. Risebro, *When are HJB-equations in stochastic control of delay systems finite dimensional?*, Stochastic Anal. Appl. **21** (2003), no. 3, 643–671.
308. I. Lasiecka, *Unified theory for abstract parabolic boundary problems - a semigroup approach*, Appl. Math. Optim. **6** (1980), no. 1, 287–333.
309. I. Lasiecka and R. Triggiani, *Differential and algebraic Riccati equations with application to boundary/point control problems: continuous theory and approximation*

theory, Lecture Notes in Control and Information Sciences, vol. 164, Springer, Berlin, 1991.
310. J.-M. Lasry and P.-L. Lions, *A remark on regularization in Hilbert spaces*, Israel J. Math. **55** (1986), no. 3, 257–266.
311. P. Lescot and M. Röckner, *Perturbations of generalized Mehler semigroups and applications to stochastic heat equations with Levy noise and singular drift*, Potential Anal. **20** (2004), no. 4, 317–344.
312. X. J. Li and J. M. Yong, *Optimal control theory for infinite-dimensional systems*, Systems and Control: Foundations and Applications, Birkhäuser, Boston, 1995.
313. J.-L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications*, vol. 1, Springer, Berlin, 1972.
314. P.-L. Lions, *Mean field games*, Course at the Collège de France. Available in video: http://www.college-de-france.fr/site/pierre-louis-lions/_audiovideos.htm.
315. ———, *Une inégalité pour les opérateurs elliptiques du second ordre*, Ann. Mat. Pura Appl. **127** (1981), no. 1, 1–11.
316. ———, *Generalized solutions of Hamilton-Jacobi equations*, Research Notes in Mathematics, vol. 69, Pitman, Boston, 1982.
317. ———, *Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. I. The dynamic programming principle and applications*, Comm. Partial Differential Equations **8** (1983), no. 10, 1101–1174.
318. ———, *Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. II. Viscosity solutions and uniqueness*, Comm. Partial Differential Equations **8** (1983), no. 11, 1229–1276.
319. ———, *Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. I. The case of bounded stochastic evolutions*, Acta Math. **161** (1988), no. 3-4, 243–278.
320. ———, *Viscosity solutions of fully nonlinear second order equations and optimal stochastic control in infinite dimensions. II. Optimal control of Zakai's equation*, Stochastic partial differential equations and applications II, Lecture Notes in Mathematics, vol. 1390, Springer, Berlin, 1989, pp. 147–170.
321. ———, *Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. III. Uniqueness of viscosity solutions for general second-order equations*, J. Funct. Anal. **86** (1989), no. 1, 1–18.
322. Q. Lü and X. Zhang, *General Pontryagin-type stochastic maximum principle and backward stochastic evolution equations in infinite dimensions*, SpringerBriefs in Mathematics, Springer, Berlin, 2014.
323. A. Lunardi, *Analytic semigroups and optimal regularity in parabolic problems*, Progress in Nonlinear Differential Equations and their Applications, vol. 16, Birkhäuser, Basel, 1995.
324. T. J. Lyons, *Uncertain volatility and the risk-free synthesis of derivatives*, Appl. Math. Finance **2** (1995), no. 2, 117–133.
325. J. Ma and J. Yong, *Forward-backward stochastic differential equations and their applications*, Lecture Notes in Mathematics, vol. 1702, Springer, Berlin, 1999.
326. Z. M. Ma and M. Röckner, *Introduction to the theory of (nonsymmetric) Dirichlet forms*, Universitext, Springer, Berlin, 1992.
327. L. Manca, *Kolmogorov equations for measures*, J. Evol. Equ. **8** (2008), no. 2, 231–262.
328. ———, *On the dynamic programming approach for the 3D Navier-Stokes equations*, Appl. Math. Optim. **57** (2008), no. 3, 329–348.
329. ———, *The Kolmogorov operator associated to a Burgers SPDE in spaces of continuous functions*, Potential Anal. **32** (2010), no. 1, 67–99.
330. C. Marinelli, *On stochastic modelling and optimal control in advertising*, Ph.D. thesis, Graduate School of Business, Columbia University, 2004.
331. C. Marinelli and S. Savin, *Optimal distributed dynamic advertising*, J. Optim. Theory Appl. **137** (2008), no. 3, 569–591.
332. F. Masiero, *Semilinear Kolmogorov equations and applications to stochastic optimal control*, Ph.D. thesis, Dipartimento di Matematica, Universitá di Milano, 2003.

333. ______, *Semilinear Kolmogorov equations and applications to stochastic optimal control*, Appl. Math. Optim. **51** (2005), no. 2, 201–250.
334. ______, *Infinite horizon stochastic optimal control problems with degenerate noise and elliptic equations in Hilbert spaces*, Appl. Math. Optim. **55** (2007), no. 3, 285–326.
335. ______, *Regularizing properties for transition semigroups and semilinear parabolic equations in Banach spaces*, Electron. J. Probab. **12** (2007), Paper no. 13, 387–419.
336. ______, *Stochastic optimal control for the stochastic heat equation with exponentially growing coefficients and with control and noise on a subdomain*, Stoch. Anal. Appl. **26** (2008), no. 4, 877–902.
337. ______, *Stochastic optimal control problems and parabolic equations in Banach spaces*, SIAM J. Control Optim. **47** (2008), no. 1, 251–300.
338. ______, *A stochastic optimal control problem for the heat equation on the halfline with Dirichlet boundary-noise and boundary-control*, Appl. Math. Optim. **62** (2010), no. 2, 253–294.
339. ______, *Hamilton Jacobi Bellman equations in infinite dimensions with quadratic and superquadratic Hamiltonian*, Discret. Contin. Dyn. S. **32** (2012), no. 1, 223–263.
340. ______, *A Bismut Elworthy formula for quadratic BSDEs*, Preprint arXiv:1404.2098, 2014.
341. F. Masiero and A. Richou, *HJB equations in infinite dimension with locally Lipschitz Hamiltonian and unbounded terminal condition*, J. Differential Equations **257** (2014), no. 6, 1989–2034.
342. J.-L. Menaldi and S. S. Sritharan, *Stochastic 2D Navier-Stokes equation*, Appl. Math. Optim. **46** (2002), no. 1, 31–53.
343. Q. Meng and P. Shi, *Stochastic optimal control for backward stochastic partial differential systems*, J. Math. Anal. Appl. **402** (2013), no. 2, 758–771.
344. M. Métivier, *Semimartingales: a course on stochastic processes*, De Gruyter Studies in Mathematics, vol. 2, Walter de Gruyter, Berlin, 1982.
345. M. Métivier and J. Pellaumail, *Stochastic integration*, Probability and Mathematical Statistics, Academic Press, New York, 1980.
346. P. A. Meyer, *Probability and potentials*, Blaisdell, New York, 1966.
347. S. E. A. Mohammed, *Stochastic functional differential equations*, Research Notes in Mathematics, no. 99, Pitman, Boston, 1984.
348. ______, *Stochastic differential systems with memory: Theory, examples and applications.*, Stochastic analysis and related topics VI. Proceedings of the 6th Oslo-Silivri workshop, Birkhäuser, Boston, 1998, pp. 1–77.
349. H. Morimoto, *Stochastic control and mathematical modelling, applications in economics*, Encyclopedia of Mathematics and its Applications, vol. 131, Cambridge University Press, Cambridge, 2010.
350. R. E. Mortensen, *Optimal control of continuous-time stochastic systems*, Tech. Report Elektronics Research Laboratory Report 66-1, Univ. of Calif., Berkeley, August 1966.
351. G. A. Muñoz, Y. Sarantopoulos, and A. Tonge, *Complexifications of real Banach spaces, polynomials and multilinear maps*, Studia Math. **134** (1999), no. 1, 1–33.
352. M. Musiela, *Stochastic PDEs and term structure models*, Journees Internationales de Finance, IGR-AFFI, La Baule, 1993.
353. M. Musiela and M. Rutkowski, *Martingale methods in financial modelling*, second ed., Stochastic Modelling and Applied Probability, vol. 36, Springer, Berlin, 2005.
354. A. S. Nemirovskiĭ and S. M. Semenov, *The polynomial approximation of functions on Hilbert space*, Mat. Sb. (N.S.) **92(134)** (1973), no. 2(10), 257–281.
355. J. Neveu, *Discrete-parameter martingales*, Holland Mathematical Library, vol. 10, North-Holland, Amsterdam, 1975.
356. M. Nisio, *Some remarks on stochastic optimal controls*, Proceedings of the Third Japan-USSR symposium on probability theory, Lecture Notes in Mathematics, vol. 550, Springer, Berlin, 1976, pp. 446–460.
357. ______, *Lectures on stochastic control theory*, ISI Lecture Notes, vol. 9, Macmillan, Delhi, 1981.

358. ———, *Optimal control for stochastic partial differential equations and viscosity solutions of Bellman equations*, Nagoya Math. J. **123** (1991), 13–37.
359. ———, *On sensitive control for stochastic partial differential equations*, Stochastic analysis on infinite-dimensional spaces, Pitman Research Notes in Mathematics Series, vol. 310, Longman, Harlow, 1994, pp. 231–241.
360. ———, *On sensitive control and differential games in infinite-dimensional spaces*, Itô's stochastic calculus and probability theory, Springer, Tokyo, 1996, pp. 281–292.
361. ———, *On infinite-dimensional stochastic differential games*, Osaka J. Math. **35** (1998), no. 1, 15–33.
362. ———, *Game approach to risk sensitive control for stochastic evolution systems*, Stochastic analysis, control, optimization and applications, Systems and Control: Foundations and Applications, Birkhäuser, Boston, 1999, pp. 115–134.
363. ———, *On value function of stochastic differential games in infinite dimensions and its application to sensitive control*, Osaka J. Math. **36** (1999), no. 2, 465–483.
364. ———, *Stochastic control theory: Dynamic programming principle*, Probability Theory and Stochastic Modelling, vol. 72, Springer, Berlin, 2015.
365. D. Nualart, *The Malliavin calculus and related topics*, Probability and its Applications (New York), Springer-Verlag, New York, 1995.
366. D. Nualart and É. Pardoux, *Stochastic calculus with anticipating integrands*, Probab. Theory Related Fields **78** (1988), no. 4, 535–581.
367. B. Øksendal, A. Sulem, and T. Zhang, *Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations*, Adv. Appl. Probab. **43** (2011), no. 2, 572–596.
368. M. Ondreját, *Uniqueness for stochastic evolution equations in Banach spaces*, Dissertationes Math. (Rozprawy Mat.) **426** (2004), 1–63.
369. É. Pardoux, *Stochastic partial differential equations and filtering of diffusion processes*, Stochastics **3** (1979), no. 2, 127–167.
370. ———, *Equations of non-linear filtering and application to stochastic control with partial observation*, Nonlinear filtering and stochastic control, Lecture Notes in Mathematics, vol. 972, Springer, Berlin, 1982, pp. 208–248.
371. ———, *Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic PDEs of second order*, Stochastic analysis and related topics VI, Progress in Probability, vol. 42, Birkhäuser, Boston, 1998, pp. 79–127.
372. É. Pardoux and S. G. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett. **14** (1990), no. 1, 55–61.
373. É. Pardoux and A. Răşcanu, *Backward stochastic differential equations with subdifferential operator and related variational inequalities*, Stochastic Process. Appl. **76** (1998), no. 2, 191–215.
374. K. R. Parthasarathy, *Probability measures on metric spaces*, Probability and Mathematical Statistics, vol. 3, Academic Press, New York, 1967.
375. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Applied Mathematical Sciences, vol. 44, Springer, New York, 1983.
376. S. Peng, *A general stochastic maximum principle for optimal control problems*, SIAM J. Control Optim. **28** (1990), no. 4, 966–979.
377. ———, *Backward stochastic differential equations and applications to optimal control*, Appl. Math. Optim. **27** (1993), no. 2, 125–144.
378. S. Peszat, *Lévy-Ornstein-Uhlenbeck transition semigroup as second quantized operator*, J. Funct. Anal. **260** (2011), no. 12, 3457–3473.
379. S. Peszat and J. Zabczyk, *Strong Feller property and irreducibility for diffusions on Hilbert spaces*, Ann. Probab. **23** (1995), no. 1, 157–172.
380. ———, *Stochastic partial differential equations with Lévy noise*, Encyclopedia of Mathematics and its Applications, vol. 113, Cambridge University Press, Cambridge, 2007.
381. B. J. Pettis, *On integration in vector spaces*, Trans. Amer. Math. Soc. **44** (1938), no. 2, 277–304.

382. H. Pham, *Continuous-time stochastic control and optimization with financial applications*, Springer, Berlin, 2009.
383. R. S. Phillips, *Dissipative operators and hyperbolic systems of partial differential equations*, Trans. Amer. Math. Soc. **90** (1959), no. 1, 193–254.
384. G. Da Prato and J. Zabczyk, *Smoothing properties of transition semigroups in hilbert spaces*, Stochastics **35** (1991), no. 2, 63–77.
385. C. Prévôt and M. Röckner, *A concise course on stochastic partial differential equations*, Lecture Notes in Mathematics, vol. 1905, Springer, Berlin, 2007.
386. E. Priola, *On a class of Markov type semigroups in spaces of uniformly continuous and bounded functions*, Studia Math. **136** (1999), no. 3, 271–295.
387. ______, *Partial differential equations with infinitely many variables*, Ph.D. thesis, Tesi di Dottorato, Universita degli Studi di Milano, 1999.
388. ______, *Uniform approximation of uniformly continuous and bounded functions on Banach spaces*, Dynam. Systems Appl. **9** (2000), no. 2, 181–197.
389. ______, *The Cauchy problem for a class of Markov-type semigroups*, Commun. Appl. Anal. **5** (2001), no. 1, 49–76.
390. ______, *Schauder estimates for a homogeneous Dirichlet problem in a half-space of a Hilbert space*, Nonlinear Anal. **44** (2001), no. 5, 679–702.
391. ______, *Dirichlet problems in a half-space of a Hilbert space*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **5** (2002), no. 2, 257–291.
392. E. Priola and S. Tracà, *On the Cauchy problem for non-local Ornstein-Uhlenbeck operators*, Preprint arXiv:1505.01876, 2015.
393. E. Priola and J. Zabczyk, *Liouville theorems for non-local operators*, J. Funct. Anal. **216** (2004), no. 2, 455–490.
394. ______, *Structural properties of semilinear SPDEs driven by cylindrical stable processes*, Probab. Theory Related Fields **149** (2011), no. 1-2, 97–137.
395. P. E. Protter, *Stochastic integration and differential equations*, Stochastic Modeling and Applied Probability, vol. 21, Springer, Berlin, 2004.
396. M. Renardy, *Polar decomposition of positive operators and a problem of Crandall and Lions*, Appl. Anal. **57** (1995), no. 3-4, 383–385.
397. M. Renardy and R. C. Rogers, *An introduction to partial differential equations*, second ed., Texts in Applied Mathematics, vol. 13, Springer, New York, 2004.
398. D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der Mathematischen Wissenschaften, vol. 293, Springer, Berlin, 1999.
399. M. Röckner and Z. Sobol, *A new approach to Kolmogorov equations in infinite dimensions and applications to stochastic generalized Burgers equations*, C. R., Math., Acad. Sci. Paris **338** (2004), no. 12, 945–949.
400. ______, *Kolmogorov equations in infinite dimensions: well-posedness and regularity of solutions, with applications to stochastic generalized Burgers equations*, Ann. Probab. **34** (2006), no. 2, 663–727.
401. ______, *A new approach to Kolmogorov equations in infinite dimensions and applications to the stochastic 2D Navier-Stokes equation*, C. R., Math., Acad. Sci. Paris **345** (2007), no. 5, 289–292.
402. L. C. G. Rogers and D. Williams, *Diffusions, Markov processes and martingales. Volume one: foundations*, Wiley Series in Probability and Mathematical Statistics, Wiley, Chichester, 1994.
403. S. Romagnoli and T. Vargiolu, *Robustness of the Black-Scholes approach in the case of options on several assets*, Finance Stoch. **4** (2000), no. 3, 325–341.
404. M. Rosestolato, *On differentiability of solutions of infinite dimensional SDEs with respect to the initial condition*, In preparation.
405. M. Royer, *BSDEs with a random terminal time driven by a monotone generator and their links with PDEs*, Stoch. Stoch. Rep. **76** (2004), no. 4, 281–307.
406. B. L. Rozovskiĭ, *Stochastic evolution systems. Linear theory and applications to nonlinear filtering*, Kluwer Academic Publishers, Dordrecht, 1990, Transl. from the Russian.
407. W. Rudin, *Real and complex analysis*, McGraw-Hill, New York, 1970.

408. ______, *Functional analysis*, McGraw-Hill Series in Higher Mathematics, McGraw-Hill, New York, 1973.
409. R. A. Ryan, *Introduction to tensor products of Banach spaces*, Springer, Berlin, 2002.
410. M. Scheutzow, *Qualitative behaviour of stochastic delay equations with a bounded memory*, Stochastics **12** (1984), no. 1, 41–80.
411. T. I. Seidman and J. Yong, *How violent are fast controls? II*, Math. Control Signals Systems **9** (1996), no. 4, 327–340.
412. K. Shimano, *A class of Hamilton-Jacobi equations with unbounded coefficients in Hilbert spaces*, Appl. Math. Optim. **45** (2002), no. 1, 75–98.
413. I. Singer, *Bases in Banach spaces I*, Grundlehren der Mathematischen Wissenschaften, vol. 154, Springer, Berlin, 1970.
414. P. E. Sobolevskiĭ, *On equations with operators forming an acute angle*, Dokl. Akad. Nauk SSSR (N.S.) **116** (1957), 754–757.
415. H. M. Soner and N. Touzi, *Superreplication under gamma constraints*, SIAM J. Control Optim. **39** (2000), no. 1, 73–96.
416. ______, *Dynamic programming for stochastic target problems and geometric flows*, J. Eur. Math. Soc. **4** (2002), no. 3, 201–236.
417. ______, *Stochastic target problems, dynamic programming, and viscosity solutions*, SIAM J. Control Optim. **41** (2002), no. 2, 404–424.
418. ______, *The problem of super-replication under constraints*, Paris-Princeton Lectures on Mathematical Finance, 2002, Lecture Notes in Mathematics, vol. 1814, Springer, Berlin, 2003, pp. 133–172.
419. S. S. Sritharan, *An introduction to deterministic and stochastic control of viscous flow*, Optimal control of viscous flow, SIAM, Philadelphia, 1998, pp. 1–42.
420. C. Stegall, *Optimization of functions on certain subsets of Banach spaces*, Math. Ann. **236** (1978), 171–176.
421. A. Święch , *Viscosity solutions of fully nonlinear partial differential equations with "unbounded" terms in infinite dimensions*, Ph.D. thesis, University of California (Santa Barbara), July 1993.
422. ______, *"Unbounded" second order partial differential equations in infinite-dimensional Hilbert spaces*, Comm. Partial Differential Equations **19** (1994), no. 11-12, 1999–2036.
423. ______, *The existence of value functions of stochastic differential games for unbounded stochastic evolution*, Proceedings of the 34th IEEE Conference on Decision and Control, vol. 3, 1995, pp. 2289–2294.
424. ______, *Risk-sensitive control and differential games in infinite dimensions*, Nonlinear Anal. **50** (2002), no. 4, 509–522.
425. ______, *A PDE approach to large deviations in Hilbert spaces*, Stochastic Process. Appl. **119** (2009), no. 4, 1081–1123.
426. A. Święch and E. V. Teixeira, *Regularity for obstacle problems in infinite dimensional Hilbert spaces*, Adv. Math. **220** (2009), no. 3, 964–983.
427. A. Święch and J. Zabczyk, *Integro-PDE in Hilbert spaces: existence of viscosity solutions*, In preparation.
428. ______, *Large deviations for stochastic PDE with Lévy noise*, J. Funct. Anal. **260** (2011), no. 3, 674–723.
429. ______, *Uniqueness for integro-PDE in Hilbert spaces*, Potential Anal. **38** (2013), no. 1, 233–259.
430. A. Talarczyk, *Dirichlet problem for parabolic equations on Hilbert spaces*, Studia Math. **141** (2000), no. 2, 109–142.
431. H. Tanabe, *Equations of evolution*, Monographs and Studies in Mathematics, vol. 6, Pitman, Boston, 1979, Translated from the Japanese.
432. ______, *Functional analytic methods for partial differential equations*, Monographs and Textbooks in Pure and Applied Mathematics, vol. 204, Dekker, New York, 1997.
433. S. Tang and X. Li, *Maximum principle for optimal control of distributed parameter stochastic systems with random jumps*, Control of partial differential equations, Lecture Notes in Pure and Applied Mathematics, Dekker, New York, 1993, pp. 867–867.

434. ______, *Necessary conditions for optimal control of stochastic systems with random jumps*, SIAM J. Control Optim. **32** (1994), no. 5, 1447–1475.
435. R. Temam, *The Navier-Stokes equations: Theory and numerical analysis*, American Mathematical Society, Providence, 1977.
436. ______, *Navier-Stokes equations and nonlinear functional analysis*, Society for Industrial Mathematics, Providence, 1995.
437. ______, *Navier-Stokes equations: theory and numerical analysis*, AMS Chelsea Publishing, American Mathematical Society, Providence, 2001.
438. V. B. Tran, *The uniqueness of viscosity solutions of second order nonlinear partial differential equations in a Hilbert space of two-dimensional functions*, Acta Math. Vietnam **31** (2006), no. 2, 149–165.
439. V. B. Tran and D. V. Tran, *Viscosity solutions of the Cauchy problem for second-order nonlinear partial differential equations in Hilbert spaces*, Electron. J. Differential Equations (2006), Paper No. 47, 15pp., electronic only.
440. T. Vargiolu, *Finite dimensional approximations for the Musiela model in the Gaussian case*, Memoire de DEA Univ. Pierre et Marie Curie - Paris VI, 1997.
441. ______, *Invariant measures for the Musiela equation with deterministic diffusion term*, Finance Stoch. **3** (1999), no. 4, 483–492.
442. R. B. Vinter and R. H. Kwong, *The infinite time quadratic control problem for linear systems with state and control delays: an evolution equation approach*, SIAM J. Control Optim. **19** (1981), no. 1, 139–153.
443. M. J. Višik and A. V. Fursikov, *Mathematical problems of statistical hydromechanics*, Mathematics and its Applications (Soviet Series), vol. 9, Springer Netherlands, 1988.
444. J. B. Walsh, *An introduction to stochastic partial differential equations*, École d'été de probabilités de Saint-Flour XIV-1984, Lecture Notes in Mathematics, vol. 1180, Springer, Berlin, 1986, pp. 265–439.
445. J. Weidmann, *Linear operators in Hilbert spaces*, Graduate Texts in Mathematics, vol. 68, Springer, 2012.
446. D. Williams, *Probability with martingales*, Cambridge University Press, Cambridge, 1991.
447. A. Wiweger, *Linear spaces with mixed topology*, Studia Math. **20** (1961), 47–68.
448. D. Yang, *Optimal control problems for Lipschitz dissipative systems with boundary-noise and boundary-control*, J. Optim. Theory Appl. **165** (2015), no. 1, 14–29.
449. J. Yong and X. Y. Zhou, *Stochastic controls, Hamiltonian systems and HJB equations*, Applications of Mathematics, vol. 43, Springer, New York, 1999.
450. K. Yosida, *Functional analysis*, sixth ed., Grundlehren der Mathematischen Wissenschaften, vol. 123, Springer, Berlin, 1980.
451. H. Yu and B. Liu, *Properties of value function and existence of viscosity solution of HJB equation for stochastic boundary control problems*, J. Franklin Inst. **348** (2011), no. 8, 2108–2127.
452. J. Zabczyk, *Linear stochastic systems in Hilbert spaces: spectral properties and limit behavior*, Report of Institute of Mathematics, Polish Academy of Sciences 236, 1981, Also published in Banach Center Publications, **41** (1985), 591–609.
453. ______, *Parabolic equations on Hilbert spaces*, Stochastic PDE's and Kolmogorov equations in infinite dimensions, Lecture Notes in Mathematics, vol. 1715, Springer, Berlin, 1999, pp. 117–213.
454. ______, *Bellman's inclusions and excessive measures*, Probab. Math. Statist. **21** (2001), no. 1, 101–122.
455. ______, *Mathematical control theory: an introduction*, Modern Birkhäuser Classics, Birkhäuser, Boston, 2008.
456. M. Zakai, *On the optimal filtering of diffusion processes*, Z. Wahrscheinlichkeitstheor. Verw. Geb. **11** (1969), no. 3, 230–243.
457. E. Zeidler, *Nonlinear functional analysis and its applications. Volume I: fixed-point theorems*, Springer, New York, 1986.
458. ______, *Nonlinear functional analysis and its applications. Volume II/A: linear monotone operators*, Springer, New York, 1989.

459. ______, *Nonlinear functional analysis and its applications. Volume II/B: nonlinear monotone operators*, Springer, New York, 1989.
460. J. Zhou, *A class of delay optimal control problems and viscosity solutions to associated Hamilton-Jacobi-Bellman equations*, Preprint.
461. ______, *A class of infinite-horizon delay optimal control problems and a viscosity solution to the associated HJB equation*, Preprint.
462. ______, *Optimal control problems for stochastic differential equations with delays and a viscosity solution to the associated HJB equation*, Preprint.
463. ______, *Stochastic delay optimal control problems and viscosity solution to the associated Hamilton-Jacobi-Bellman equations*, Preprint.
464. ______, *Optimal control of a stochastic delay heat equation with boundary-noise and boundary-control*, Int. J. Control **87** (2014), no. 9, 1808–1821.
465. ______, *Optimal control of a stochastic delay partial differential equation with boundary-noise and boundary-control*, J. Dyn. Control Syst. **20** (2014), no. 4, 503–522.
466. J. Zhou and B. Liu, *The existence and uniqueness of the solution for nonlinear Kolmogorov equations*, J. Differential Equations **253** (2012), no. 11, 2873–2915.
467. J. Zhou and Z. Zhang, *Optimal control problems for stochastic delay evolution equations in Banach spaces*, Int. J. Control **84** (2011), no. 8, 1295–1309.
468. X. Y. Zhou, *On the existence of optimal relaxed controls of stochastic partial differential equations*, SIAM J. Control Optim. **30** (1992), no. 2, 247–261.
469. ______, *On the necessary conditions of optimal controls for stochastic partial differential equations*, SIAM J. Control Optim. **31** (1993), no. 6, 1462–1478.
470. ______, *Sufficient conditions of optimality for stochastic systems with controllable diffusions*, IEEE Trans. Auto. Control **41** (1996), no. 8, 1176–1179.