Chapter 5
Mild Solutions in L^2 Spaces

This chapter is devoted to the presentation of the L^2 theory for the existence and uniqueness of mild solutions for a class of second-order infinite-dimensional HJB equations in Hilbert spaces through a perturbation approach. As in the previous chapter, the concept of mild solution concerns the HJB equation in an integral form that uses the transition semigroup associated to the linear part of the equation.

In the previous chapter the perturbation approach was used in Banach spaces of regular (at least differentiable in the x variable, in a suitable sense) real-valued functions defined on a Hilbert space H. The space where we seek the solutions here is a space of functions which are square-integrable (with their x derivative defined in a suitable sense) with respect to a suitable reference measure m on H.

One of the main reasons for the development of the L^2 theory is the need to study HJB equations without the smoothing Hypothesis 4.76 about the behavior of the transition semigroup, which was used in the previous chapter (see Sect. 4.1 for a discussion). Indeed, once the existence of the reference measure is postulated, the estimates that allow us to ensure, in the L^2 framework, the applicability of a fixed point argument, can be proved under weaker assumptions (see Sect. 5.1 for details).

As for the mild solutions in spaces of continuous functions, the L^2 theory can be applied to obtain optimal synthesis. The class of applicable infinite-dimensional stochastic optimal control problems includes cases which cannot be treated in the context presented in Chap. 4, like the stochastic delay differential equations and first-order SPDEs. On the other hand, specific hypotheses ensuring the existence of the reference measure m and the compatibility of the Hamiltonian with it, need to be satisfied. Moreover, the synthesis provided by the L^2 theory is less regular.

The approach we describe was mostly developed in [3, 4, 125, 298]. We will mainly follow [298].

The chapter is organized as follows:

- In Sect. 5.1 we describe the main ideas of the L^2 method.

G. Fabbri et al., *Stochastic Optimal Control in Infinite Dimension*,
Probability Theory and Stochastic Modelling 82,
DOI 10.1007/978-3-319-53067-3_5

- In Sect. 5.2 we recall some classical results about invariant measures and other preliminary facts.
- Sections 5.3 and 5.4 are devoted to parabolic HJB equations. Section 5.3 contains existence and uniqueness results, while in Sect. 5.4 a result on approximation of mild solutions by classical solutions is provided.
- In Sect. 5.5 we apply the results of Sects. 5.3 and 5.4 to perform the optimal synthesis for stochastic optimal control problems, while in Sect. 5.6 we provide specific examples related to those of Chap. 2.
- In Sect. 5.7 we describe complementary results, mainly from [3, 4], which cover an additional class of problems. This section also contains existence and uniqueness results for a family of elliptic HJB equations without applications to control problems.
- Section 5.8 contains bibliographical notes.

5.1 Introduction to the Methods

We briefly sketch the main ideas of the method developed in the next sections. We consider a class of second-order infinite-dimensional HJB equations of the form

$$\begin{cases} v_t + \frac{1}{2} \operatorname{Tr} \left[Q D^2 v \right] + \langle Ax + b(x), Dv \rangle + F(t, x, Dv) + l(t, x) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad t \in [0, T), \ x \in D(A) \\ v(T, x) = g(x), \qquad x \in H, \end{cases} \tag{5.1}$$

and

$$\lambda v - \frac{1}{2} \operatorname{Tr} \left[Q D^2 v \right] - \langle Ax, Dv \rangle - F(x, Dv) = g, \qquad x \in H, \tag{5.2}$$

where $T > 0$ is fixed, A is the generator of a C_0-semigroup on a real separable Hilbert space H, $Q \in \mathcal{L}^+(H)$, and $b : H \to \mathbb{R}$, $l : [0, T] \times H \to \mathbb{R}$, $g : H \to \mathbb{R}$, $F : [0, T] \times H \times H \to \mathbb{R}$ (or $F : H \times H \to \mathbb{R}$) are measurable functions. Further hypotheses on b, l, g and F will be introduced later.

Since the results available in the literature up to now are mainly oriented towards the evolutionary HJB equation (5.1), we devote most of the chapter to the theory in this case, limiting the treatment of the stationary equation (5.2) to Sect. 5.7.3.

Given a reference measure on H, the basic idea is to introduce mild and strong solutions of (5.1) and (5.2) in the space of real square-integrable functions on $[0, T] \times H$ (or on H). If H were a finite-dimensional space, the Lebesgue measure would be the natural choice for the reference measure but in infinite dimension the situation is more delicate. We consider the following stochastic evolution equation

$$\begin{cases} dX(s) = (AX(s) + b(X(s)))\, ds + dW_Q(s), \quad s \geq 0, \\ X(0) = x \in H, \end{cases} \tag{5.3}$$

we suppose it admits a mild solution and an invariant measure m and we work in the space $L^2(H, \overline{\mathcal{B}}, m)$ where $\overline{\mathcal{B}}$ is the completion of the Borel σ-field $\mathcal{B}(H)$ with respect to m.

Under suitable assumptions on the operators A and Q and on the function b (see, e.g., [180] Chap. 9), the solution w of the following Kolmogorov equation

$$\begin{cases} w_t = \frac{1}{2}\ \mathrm{Tr}\,[QD^2 w] + \langle Ax + b(x), Dw \rangle\,, \\ w(0, x) = \phi(x) \end{cases} \tag{5.4}$$

can be associated to the transition semigroup P_t of the solution $X(\cdot; x)$ of (5.3) as follows:

$$w(t, x) = P_t[\phi](x) = \mathbb{E}\phi(X(t, x)) \tag{5.5}$$

for any bounded continuous ϕ.

The semigroup P_t extends to a strongly continuous semigroup of contractions on $L^2(H, \overline{\mathcal{B}}, m)$ with generator $\mathcal{A}$, whose explicit expression on regular functions is

$$\mathcal{A}\phi(x) = \frac{1}{2}\ \mathrm{Tr}\,[QD^2\phi] + \langle Ax + b(x), D\phi \rangle\,; \tag{5.6}$$

this fact is recalled in Lemma 5.37.

The original HJB equation (5.1) can be seen as a perturbation of (5.4) and, by formally applying the variation of parameters formula, it can be written in the following integral (mild) form

$$u(t, x) = P_{T-t}[g](\cdot) + \int_t^T P_{s-t}\,[l(s, \cdot) + F\,(s, \cdot, Du(s, \cdot))]\,(x)ds. \tag{5.7}$$

To prove the existence and uniqueness of mild solutions in spaces of continuous functions we needed, as a key assumption, a smoothing property for the transition semigroup P_t of the following form[1]: there exist $C > 0$ and $\theta \in (0, 1)$ such that for every $\varphi \in B_b(X)$, $s > t$, $x \in H$,

$$|DP_{t-s}[\varphi](x)| \leq C(1 \vee (s-t)^{-\theta})\|\varphi\|_0$$

(or a similar hypothesis which uses an operator G and an integrable function γ, see Sect. 4.1.1 for details). This assumption was needed to prove the existence and

[1] See Hypothesis 4.76.

uniqueness of the solution using a fixed point theorem in a Banach space of continuous and differentiable functions (see e.g. Theorem 4.80).

In the L^2 setting, an important role is played by the space $W_Q^{1,2}(H, m)$ which is, formally, the Sobolev space of functions which admit a weak derivative in $L^2\left(H, \overline{\mathcal{B}}, m\right)$, endowed with the norm

$$|\phi|^2_{W_Q^{1,2}} = \int_H |\phi|^2 dm + \int_H \left|Q^{1/2} D\phi\right|^2 dm.$$

In fact, the definition of such a space is more complicated (see Definition 5.11) due to the fact that the operator $Q^{1/2}D$ is not assumed to be closable in $L^2\left(H, \overline{\mathcal{B}}, m\right)$. We work in this framework because $Q^{1/2}D$ is not closable in some relevant cases, such as, for example, in the case of delay equations (see Sect. 5.6). The existence and uniqueness result is found by applying a fixed point argument in the space $L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ (see Theorem 5.35). In this new context a milder smoothing property is required (see estimate (5.36) in Proposition 5.20) and, thanks to the properties of the invariant measure m, it can be verified without strong requirements on the data A, b and Q. This is the main reason why the L^2 theory developed in the present chapter allows us to deal with equations and control problems which cannot be treated by the techniques of Chap. 4.

More precisely:

(i) We do not need any smoothing properties of the Ornstein–Uhlenbeck semigroup associated with (A, Q) (see Remark 5.21). Therefore we do not impose any restrictions on Q: it is possible, for example, to take Q a one-dimensional projection.
(ii) $g, l \in L^2(H, \overline{\mathcal{B}}, m)$: they are not necessarily continuous, bounded or with polynomial growth.

This generality comes at a price. Similarly to Chap. 6 and differently from Chap. 4, we can only deal with a class of Hamiltonians of the form $F(t, x, p) = F_0\left(t, x, Q^{1/2}p\right)$. If we look at this restriction in terms of the optimal control problems we can study, it means that we are only able to deal with problems where the control appears in the state equation via a term of the form $Q^{1/2}R(t, x, a(t))$ (see (5.78)). This assumption may seem restrictive, but in fact it is quite natural in many control problems when the operator Q is degenerate. It implies that the system should be controlled by feedback taking values in the same space in which the noise disturbing the system is concentrated. Let us note that if $Q^{1/2} = 0$ then both the control and the noise disappear. A natural interpretation of this fact is that the uncontrolled system is in fact deterministic and the noise is brought into the system only by the control.

Another drawback is the fact that mild solutions found in the setting of this chapter possess weaker regularity properties due to the choice of the spaces. In particular, if Q is very degenerate (e.g. a finite-dimensional projection) the measure substantially ignores most of the space H. However, despite this weak regularity, when (5.7) is the HJB equation related to a stochastic optimal control problem, one can characterize its solution as the value function of the problem and use it to perform the optimal synthesis.

5.2 Preliminaries and the Linear Problem

5.2.1 Notation

As usual we denote by H a real separable Hilbert space with the norm $|\cdot|$ and the inner product $\langle \cdot, \cdot \rangle$ and by Q an element of $\mathcal{L}^+(H)$. $\mathcal{B}(H)$ is the Borel σ-field of H. The function spaces $C(H)$, $UC(H)$, $C_b(H)$, $UC_b(H)$, $C_b(H, H)$, $C_b^k(H)$, $C_0^k(\mathbb{R}^n)$, ... are defined in Appendix A.

5.2.2 The Reference Measure m and the Main Assumptions on the Linear Part

We will work under the following set of assumptions.

Hypothesis 5.1 (A) A is the generator of a strongly continuous semigroup $\{e^{tA},\ t \geq 0\}$ on a real separable Hilbert space H. $M \geq 1$ and $\omega \in \mathbb{R}$ are two real constants such that

$$\left\| e^{tA} \right\| \leq M e^{\omega t}, \qquad \forall t \geq 0.$$

(B) $Q \in \mathcal{L}^+(H)$, and $\mu_0 = \left(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t \geq 0}, \mathbb{P}, W_Q\right)$ is every generalized reference probability space (see Definition 1.100).

(C) $e^{sA} Q e^{sA^*} \in \mathcal{L}_1(H)$ for all $s > 0$. Moreover, for every $t \geq 0$,

$$\int_0^t \operatorname{Tr}\left[e^{sA} Q e^{sA^*}\right] ds < +\infty,$$

so the symmetric positive operator

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

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

(D) The function $b : H \to H$ is continuous and Gâteaux differentiable, its Gâteaux differential ∇b is strongly continuous and

$$\|\nabla b\|_0 = \sup_{x \in H} \|\nabla b(x)\| \leq K < +\infty.$$

Proposition 5.2 *Let Hypothesis 5.1 be satisfied. Then:*

(i) The equation

$$\begin{cases} dX(s) = (AX(s) + b(X(s)))\, ds + dW_Q(s), & s \in [0, T], \\ X(0) = x \in H \end{cases} \tag{5.8}$$

has a unique mild solution $X(\cdot; x) \in \mathcal{H}_p^{\mu_0}(0, T; H)$ *(see Definition 1.126) for all* $p \geq 1$. *We also have*

$$\lim_{s \to 0} \mathbb{E}\, |X(s, x) - x|^2 = 0. \tag{5.9}$$

(ii) There exists a $\mathcal{B}([0, T]) \otimes \mathcal{B}(H) \otimes \mathscr{F} / \mathcal{B}(H)$*-measurable function*

$$\begin{cases} [0, T] \times H \times \Omega \to H \\ (s, x, \omega) \to \tilde{X}(s; x)(\omega) \end{cases}$$

such that, for every $x \in H$, $\tilde{X}(\cdot; x)$ *is a version of the solution* $X(\cdot; x)$. *Thus in the future we will not make a distinction between* $X(\cdot; x)$ *and* $\tilde{X}(\cdot; x)$.

Proof Part (i), except (5.9), is proved in Theorem 1.147 (observe that b is globally Lipschitz continuous thanks to Hypothesis 5.1-(D) and Theorem D.18). To prove (5.9) we can observe that, using Hypotheses 5.1-(A) and (D),

$$\mathbb{E}\, |X(s, x) - x|^2 \leq 3 \left| e^{sA} x - x \right|^2 + 3C \int_0^s \mathbb{E} \left(1 + |X(r)|^2 \right) dr + 3\mathbb{E} \left| W^A(s) \right|^2, \quad s \in [0, T],$$

where C is a constant depending only on b. The first term converges to zero when $s \to 0$, the second goes to zero because $X(\cdot; x) \in \mathcal{H}_2^{\mu_0}(0, T; H)$ while the term concerning the stochastic convolution converges to zero thanks to its mean square continuity ensured by Proposition 1.144.

Part (ii) is proved in Proposition 5.44 for a more general controlled version of the equation (even though Proposition 5.44 is in a later section, its proof is independent). □

The *transition semigroup* P_s, $s \geq 0$, associated to (5.8) is defined for every $\phi \in C_b(H)$ as[2]

$$\begin{cases} P_s[\phi] \colon H \to \mathbb{R} \\ P_s[\phi] \colon x \to \mathbb{E}\phi(X(s; x)), \end{cases} \tag{5.10}$$

[2]In Sect. 1.6 we define the semigroup directly on all the functions of $B_b(H)$. The arguments of the present chapter are more transparent if we start by defining the semigroup only on $C_b(H)$. Since it will be extended (Proposition 5.9) to $L^p(H, \overline{\mathcal{B}}, m)$, and (Lemma 5.10), for any $\phi \in L^p(H, \overline{\mathcal{B}}, m)$, $P_t[\phi](x) = \mathbb{E}\phi(X(t; x))$, the two approaches are equivalent.

where $X(s; x)$ is the solution of (5.8) at time s. It follows from Proposition 1.147 that $P_s(C_b(H)) \subset C_b(H)$ (see Theorem 1.162) and P_s has the semigroup property in $C_b(H)$ as was remarked in Corollary 1.158. Moreover, P_s does not depend on μ_0 so the theory developed in this chapter is independent of the choice of μ_0.

In the setting described by Hypothesis 5.1, we can introduce the notion of an invariant measure.

Definition 5.3 (*Invariant measure*) Let P_t be the transition semigroup introduced in (5.10). A probability measure m on $(H, \mathcal{B}(H))$ is said to be an *invariant measure* for (5.8) if, for any $\phi \in C_b(H)$ and $t \geq 0$,

$$\int_H P_t[\phi](x) dm(x) = \int_H \phi(x) dm(x). \tag{5.11}$$

If Hypothesis 5.1 holds, we formulate the following assumption.

Hypothesis 5.4 There exists an invariant measure m for Eq. (5.8). Moreover,

$$\int_H |x|^2 \, dm(x) < \infty. \tag{5.12}$$

We denote by $\overline{\mathcal{B}}$ the completion (see Sect. 1.1.1) of the Borel σ-field $\mathcal{B}(H)$ with respect to the measure m.

Notation 5.5 L^p spaces have been introduced in Sect. 1.1.3. In order to distinguish the norms in $L^p(H, \overline{\mathcal{B}}, m)$ and $L^p(H, \overline{\mathcal{B}}, m; H)$ (i.e., the L^p norms computed using the measure m) from other L^p-norms that appear in this chapter, we will denote them by $|\cdot|_{L^p_m}$ and by $|\cdot|_{L^p_{m,H}}$.

We first recall some density results that we will use frequently.

Lemma 5.6 *Suppose that A satisfies Hypothesis 5.1 (A). Denote by $\mathcal{E}_A(H)$ the linear subspace of $UC_b(H)$ given by the linear span of the set of all real parts of the functions $e^{i\langle x,h\rangle}$ for some $h \in D(A^*)$. Then, for any $f \in UC_b(H)$ there exists a multi-sequence $\left(f_{n_1,n_2,n_3}\right)_{n_1,n_2,n_3 \in \mathbb{N}}$ in $\mathcal{E}_A(H)$ such that*

$$\|f_{n_1,n_2,n_3}\|_0 \leq \|f\|_0, \quad \textit{for any}\, n_1, n_2, n_3 \in \mathbb{N}$$

and

$$\lim_{n_1\to+\infty} \lim_{n_2\to+\infty} \lim_{n_3\to+\infty} f_{n_1,n_2,n_3}(x) = f(x), \quad \textit{for any}\, x \in H.$$

Proof See Lemma 6.2.3, p. 112 in [179]. □

Lemma 5.7 *Given any bounded measure $\bar{m}$ defined on the Borel σ-field $\mathcal{B}(H)$ of H, denoting by $\overline{\mathcal{B}}_{\bar{m}}$ the completion of $\mathcal{B}(H)$ with respect to $\bar{m}$, we have the following density results:*

(i) *$UC_b(H)$ and $UC_b^k(H)$, for any integer $k > 0$, are dense in $L^2(H, \overline{\mathcal{B}}_{\bar{m}}, \bar{m})$.*

(ii) *Let A be the generator of a C_0-semigroup on H and let A^* be its adjoint. Then $\mathcal{F}C_0^{k,A^*}(H)$, defined in (A.4), is dense in $L^2(H, \overline{\mathcal{B}}_{\bar{m}}, \bar{m})$ for any integer $k \geq 0$.*

(iii) *For every $\psi \in L^2(0, T; L^2(H, \overline{\mathcal{B}}_{\bar{m}}, \bar{m}))$ there exists a sequence $\psi_n \colon [0, T] \to \mathcal{F}C_0^{2,A^*}(H)$ such that*

$$\begin{cases} \psi_n \in C\left([0,T], UC_b(H)\right), \\ D\psi_n, A^* D\psi_n \in C\left([0,T], UC_b(H,H)\right), \\ D^2\psi_n \in C\left([0,T], UC_b(H, \mathcal{L}_1(H))\right), \end{cases}$$

and

$$\psi_n \xrightarrow{n\to+\infty} \psi \quad \text{in } L^2\left(0, T; L^2\left(H, \overline{\mathcal{B}}_{\bar{m}}, \bar{m}\right)\right).$$

Proof Part (i): $UC_b(H)$ is dense in $L^2(H, \overline{\mathcal{B}}_{\bar{m}}, \bar{m})$ thanks to Theorem 1.34. The density of $UC_b^k(H)$ in $L^2(H, \overline{\mathcal{B}}_{\bar{m}}, \bar{m})$ for $k > 0$ will be proved below.

Part (ii): Given $f \in L^2(H, \overline{\mathcal{B}}_{\bar{m}}, \bar{m})$ and any $n \in \mathbb{N}$ we need to find $\tilde{f}_n \in \mathcal{F}C_0^{k,A^*}(H)$ with $|f - \tilde{f}_n|_{L^2(H,\overline{\mathcal{B}}_{\bar{m}},\bar{m})} \leq \frac{1}{n}$. Thanks to the already recalled density of $UC_b(H)$ in $L^2(H, \overline{\mathcal{B}}_{\bar{m}}, \bar{m})$ we can suppose that $f \in UC_b(H)$ and we can then consider an approximating multi-sequence $f_{n_1,n_2,n_3} \in \mathcal{E}_A(H)$ from Lemma 5.6. We define, for any $x \in H$, for $n_1 \in \mathbb{N}$, $f_{n_1}(x) := \lim_{n_2\to+\infty} \lim_{n_3\to+\infty} f_{n_1,n_2,n_3}(x)$ and, for $n_1, n_2 \in \mathbb{N}$, $f_{n_1,n_2}(x) := \lim_{n_3\to+\infty} f_{n_1,n_2,n_3}(x)$ so that, pointwise, $f = \lim_{n_1\to+\infty} f_{n_1}$. Using Egoroff's Theorem (Lemma 1.50-(iv)) we can find n_1 such that $|f - f_{n_1}|_{L^2(H,\overline{\mathcal{B}}_{\bar{m}},\bar{m})} \leq \frac{1}{6n}$, then n_2 such that $|f_{n_1} - f_{n_1,n_2}|_{L^2(H,\overline{\mathcal{B}}_{\bar{m}},\bar{m})} \leq \frac{1}{6n}$ and n_3 such that $|f_{n_1,n_2} - f_{n_1,n_2,n_3}|_{L^2(H,\overline{\mathcal{B}}_{\bar{m}},\bar{m})} \leq \frac{1}{6n}$. We denote such an f_{n_1,n_2,n_3} by f_n and we have $|f - f_n|_{L^2(H,\overline{\mathcal{B}}_{\bar{m}},\bar{m})} \leq \frac{1}{2n}$. The function f_n is a linear combination of real parts of functions $e^{i\langle x,h_i\rangle}$ for some $h_i \in D(A^*)$, $i = 1, \ldots, k_n$, so it does not belong to $\mathcal{F}C_0^{k,A^*}(H)$ and we need to modify it.

Let $\lambda \colon \mathbb{R} \to [0, 1]$ be a C^∞ function compactly supported in $(-2, 2)$ and identically equal to 1 in the interval $[-1, 1]$. We choose $\delta > 0$ and we replace the real part of each term $e^{i\langle x,h_i\rangle}$ in the linear combination by the real part of $e^{i\langle x,h_i\rangle}\lambda\left(\delta\langle x, h_i\rangle\right)$. We call the new function $\tilde{f}_n$. It belongs to $\mathcal{F}C_0^{k,A^*}(H)$ and if we choose δ small enough we have $|f_n - \tilde{f}_n|_{L^2(H,\overline{\mathcal{B}}_{\bar{m}},\bar{m})} \leq \frac{1}{2n}$. It then follows that $|f - \tilde{f}_n|_{L^2(H,\overline{\mathcal{B}}_{\bar{m}},\bar{m})} \leq \frac{1}{n}$.

The density of $UC_b^k(H)$ claimed in Part (i) now follows from Part (ii).

The proof of Part (iii) follows by applying the results of Part (ii) to the Hilbert space $\tilde{H} := \mathbb{R} \times H$, with the operator $\tilde{A} := \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}$ (having domain $\mathbb{R} \times D(A)$) and the measure $\tilde{m} := \mathbf{1}_{[0,T]}dt \otimes \bar{m}$, where dt is the Lebesgue measure on $\mathbb{R}$. □

Lemma 5.8 *The following results hold:*

(i) *If b satisfies Hypothesis 5.1-(D), there exists a sequence $(b_n) \subset C^2(H, H)$ such that*

$$\sup_n \|Db_n\|_0 \leq K < +\infty, \tag{5.13}$$

and for all $h, x \in H$ *and for any sequence* x_n *of elements of* H *converging to* x*,*

$$\lim_{n\to\infty} b_n(x_n) = b(x), \quad \lim_{n\to\infty} Db_n(x_n)(h) = \nabla b(x)(h).$$

(ii) If b *satisfies Hypothesis 5.1-(D) and* $\|b\|_0 < +\infty$*, then the sequence in Part (i) can be chosen such that*

$$\sup_n \|b_n\|_0 \le l < +\infty. \tag{5.14}$$

(iii) Given $\phi \in C_b^1(H)$*, there exists a sequence* $(\phi_n) \subset UC_b^2(H)$ *such that*

$$\sup_n \|\phi_n\|_0 \le l < +\infty, \quad \sup_n \|D\phi_n\|_0 \le l < +\infty, \tag{5.15}$$

and, for all $x \in H$*,*

$$\lim_{n\to\infty} \phi_n(x) = \phi(x), \quad \lim_{n\to\infty} D\phi_n(x) = D\phi(x).$$

Proof We only prove (i) since the proofs of (ii) and (iii) use the same arguments. The proof is based on a standard procedure of mollification over finite-dimensional subspaces (see e.g. the proof of Lemma 1.2, p. 164 of [486]). Take an orthonormal basis $\{e_n\}$ of H and, for $z \in H$, let $z = \sum_{i=1}^{\infty} z_i e_i$. 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

$$\Pi_n : H \to \mathbb{R}^n, \qquad \Pi_n z = (z_1, \ldots, z_n),$$

$$Q_n : \mathbb{R}^n \to H, \qquad Q_n(z_1, \ldots, z_n) = z_1 e_1 + \cdots + z_n e_n,$$

and recall that $P_n = Q_n \circ \Pi_n$. Given a family of C^∞ mollifiers $\eta_n : \mathbb{R}^n \to \mathbb{R}$ with support in $B(0, 1/n)$, we define

$$b_n(z) = \int_{\mathbb{R}^n} b(Q_n y)\eta_n(\Pi_n z - y)dy = \int_{\mathbb{R}^n} b(P_n z - Q_n y)\eta_n(y)dy.$$

From the first equality above, we easily conclude that $b_n \in C^\infty(H, H)$. We have, in particular,

$$b_n(x_n) = \int_{\mathbb{R}^n} b(P_n x_n - Q_n y)\eta_n(y)dy.$$

From this equation, the fact that $P_n x_n \to x$ and the continuity of b we can conclude that

$$\lim_{n\to\infty} b_n(x_n) = b(x).$$

Fix $z \in H$. For any $h \in H$ with $|h| = 1$ and $\tau > 0$ we have

$$\begin{aligned}\frac{b_n(z+\tau h)-b_n(z)}{\tau} &= \frac{1}{\tau}\int_{\mathbb{R}^n}\left[b(P_n(z+\tau h)-Q_n y)-b(P_n z-Q_n y)\right]\eta_n(y)dy\\ &= \int_{\mathbb{R}^n}\left[\int_0^1 \nabla b(P_n(z+r\tau h)-Q_n y)(h)dr\right]\eta_n(y)dy,\end{aligned} \tag{5.16}$$

where in last equality we used Theorem D.18. In particular,

$$\frac{b_n(x_n+\tau h)-b_n(x_n)}{\tau} = \int_{\mathbb{R}^n}\left[\int_0^1 \nabla b(P_n(x_n+r\tau h)-Q_n y)(h)dr\right]\eta_n(y)dy. \tag{5.17}$$

Since $b_n \in C^\infty(H,H)$ the left-hand side of the previous equality converges, when $\tau \to 0$, to $Db_n(x_n)(h)$ while, thanks to the strong continuity of ∇b, the right-hand side converges to $\int_{\mathbb{R}^n} \nabla b(P_n x_n - Q_n y)(h)\eta_n(y)dy$. Taking the limits of the two expressions when $n \to \infty$ we get (again thanks to the strong continuity of ∇b)

$$\lim_{n\to\infty} Db_n(x_n)(h) = \nabla b(x)(h).$$

Thanks to the last equality in (5.16), for any $z \in H$, we also have $\left|\frac{b_n(z+\tau h)-b_n(z)}{\tau}\right| \leq \|\nabla b\|_0$ and then, letting $\tau \to 0$, we obtain

$$\sup_n \|Db_n\|_0 \leq \|\nabla b\|_0.$$

□

Proposition 5.9 *Let $p \in [1,+\infty)$. Assume that Hypotheses 5.1 and 5.4 hold. Then P_t, defined on $C_b(H)$ by (5.10), extends to a strongly continuous semigroup of contractions on $L^p(H,\overline{\mathcal{B}},m)$. Moreover, for any $\phi \in L^p(H,\overline{\mathcal{B}},m)$ and $t \geq 0$, the relation (5.11) holds.*

Proof We follow the proof of Theorem 10.1.5, p. 209 of [179], where the statement is proved for the Ornstein–Uhlenbeck case. Given $\phi \in C_b(H)$, for any $x \in H$, thanks to Jensen's inequality we have $|P_t[\phi](x)|^p \leq |P_t\left[|\phi|^p\right](x)|$. Thus, since m is invariant,

$$\int_H |P_t[\phi](x)|^p dm(x) \leq \int_H |P_t\left[|\phi|^p\right](x)|dm(x) = \int_H |\phi|^p(x)dm(x),$$

where the last expression is finite since ϕ is bounded and m is a finite measure. Thanks to the density of $C_b(H)$ in $L^p(H,\overline{\mathcal{B}},m)$ (Theorem 1.34), P_t extends to a contraction on $L^p(H,\overline{\mathcal{B}},m)$ for any $t \geq 0$.

To prove the strong continuity we observe first that it follows easily from the Lebesgue dominated convergence theorem and (5.9) that for every $\phi \in C_b(H)$ and $x \in H$, we have $\lim_{t\to 0^+} P_t[\phi](x) = \phi(x)$. Moreover, since $\|P_t[\phi]\|_0 \leq \|\phi\|_0$, we then obtain, again using the Lebesgue dominated convergence theorem,

$$\lim_{t\to 0^+} P_t[\phi] = \phi \qquad \text{in } L^p(H,\overline{\mathcal{B}},m).$$

Since P_t is a semigroup of contractions on $L^p(H, \overline{\mathcal{B}}, m)$ this implies strong continuity for every $\phi \in L^p(H, \overline{\mathcal{B}}, m)$.

To show the last claim, let $\phi \in L^p(H, \overline{\mathcal{B}}, m)$ and let $\phi_n \in C_b(H)$ be a sequence such that $\phi_n \to \phi$ in $L^p(H, \overline{\mathcal{B}}, m)$. We have, in particular, $\int_H \phi_n(x)\, dm(x) \to \int_H \phi(x)\, dm(x)$. Moreover, since for any $t \geq 0$, $P_t \in \mathcal{L}(L^p(H, \overline{\mathcal{B}}, m))$, $P_t[\phi_n] \to P_t[\phi]$ in $L^p(H, \overline{\mathcal{B}}, m)$ and, in particular, $\int_H P_t[\phi_n](x)\, dm(x) \to \int_H P_t[\phi](x)\, dm(x)$, so (5.11) follows letting $n \to \infty$ because it holds for the elements of $C_b(H)$. □

In the previous proposition we extended, for any $t \geq 0$, the operator P_t to the whole space $L^p(H, \overline{\mathcal{B}}, m)$ by continuity. In other words, given $\phi \in L^p(H, \overline{\mathcal{B}}, m)$, $P_t[\phi]$ is defined as the limit in $L^p(H, \overline{\mathcal{B}}, m)$ of $P_t[\phi_n]$, where ϕ_n is a (any) sequence of elements of $C_b(H)$ converging to ϕ in $L^p(H, \overline{\mathcal{B}}, m)$. In the following lemma we show that this limit is indeed equal to $\mathbb{E}\phi(X(t; x))$ (which will be proved to be a well-defined expression) even for non-bounded and non-Borel measurable elements of $L^p(H, \overline{\mathcal{B}}, m)$.

Lemma 5.10 *Let $p \in [1, +\infty)$. Assume that Hypotheses 5.1 and 5.4 hold. Consider $\phi \in L^p(H, \overline{\mathcal{B}}, m)$ and $t \in [0, T]$. Then the function*

$$\begin{cases} H \times \Omega \to \mathbb{R} \\ (x, \omega) \to \phi(X(t; x)(\omega)) \end{cases}$$

is $\overline{\mathcal{B}(H) \otimes \mathscr{F}}/\mathcal{B}(\mathbb{R})$-measurable, where $\overline{\mathcal{B}(H) \otimes \mathscr{F}}$ is the completion of the σ-field $\mathcal{B}(H) \otimes \mathscr{F}$ w.r.t. the measure $m \otimes \mathbb{P}$. Moreover, $x \to \mathbb{E}\phi(X(t; x))$ is a $\overline{\mathcal{B}}/\mathcal{B}(\mathbb{R})$-measurable function and

$$P_t[\phi](x) = \mathbb{E}\phi(X(t; x)) \qquad \textit{for } m\textit{-a.e. } x \in H. \tag{5.18}$$

Proof Suppose first that ϕ is Borel-measurable and $\phi \geq 0$. By Proposition 5.44 we can assume that $(t, x, \omega) \to \phi(X(t; x)(\omega))$ is a $\mathcal{B}[0, T] \otimes \mathcal{B}(H) \otimes \mathscr{F}/\mathcal{B}(H)$-measurable function and then (see Lemma 1.8(iv)), for any $t \in [0, T]$, $(x, \omega) \to \phi(X(t; x)(\omega))$ is $\mathcal{B}(H) \otimes \mathscr{F}/\mathcal{B}(H)$-measurable so that the function $(x, \omega) \to \phi(X(t; x)(\omega))$, being the composition of a $\mathcal{B}(H) \otimes \mathscr{F}/\mathcal{B}(H)$-measurable function and a $\mathcal{B}(H)/\mathcal{B}(\mathbb{R})$-measurable function, is $\mathcal{B}(H) \otimes \mathscr{F}/\mathcal{B}(\mathbb{R})$-measurable. The (Borel) measurability of $x \to \mathbb{E}\phi(X(t; x))$ then follows (see e.g. Lemma 1.26, p. 14 of [370]). Moreover, if we consider $\phi_n := \phi \wedge n$, thanks to the monotone convergence theorem, we have

$$\mathbb{E}\phi(X(t; x)) = \lim_{n\to\infty} \mathbb{E}\phi_n(X(t; x)) = \lim_{n\to\infty} P_t[\phi_n](x), \qquad x \in H$$

(the limit can also be $+\infty$ for certain x). Since (again by monotone convergence) we have $\lim_{n\to\infty} \phi_n := \phi$ in $L^p(H, \overline{\mathcal{B}}, m)$, we also have

$$\lim_{n\to\infty} P_t[\phi_n] = P_t[\phi]$$

in $L^p(H, \overline{\mathcal{B}}, m)$ and then, extracting if necessary a subsequence, m-a.e. Thus we obtain (5.18).

As a second step we consider a positive $\phi \in L^p(H, \overline{\mathcal{B}}, m)$. By Lemma 1.16 we can find $\tilde{\phi} \in L^p(H, \mathcal{B}(H), m)$ and $V \in \mathcal{B}(H)$, $m(V) = 0$ such that $\phi(x) = \tilde{\phi}(x)$ for any $x \in H \setminus V$. Denoting by $\mathbf{1}_V$ the characteristic function of V we have

$$\int_H \mathbb{P}\{X(t;x)(\omega) \in V\}\, dm(x) = \int_H \mathbb{E}\left[\mathbf{1}_V(X(t;x))\right] dm(x)$$
$$= \int_H P_t[\mathbf{1}_V](x) dm(x) = \int_H \mathbf{1}_V(x) dm(x) = 0. \tag{5.19}$$

So the functions $(x, \omega) \to \phi(X(t;x)(\omega)$ and $(x, \omega) \to \tilde{\phi}(X(t;x)(\omega)$ disagree only on a subset of $H \times \Omega$ which has $m \otimes \mathbb{P}$-measure 0 and thus, since we have already observed that $(x, \omega) \to \tilde{\phi}(X(t;x)(\omega)$ is $\mathscr{F} \otimes \mathcal{B}/\mathcal{B}(\mathbb{R})$-measurable, $(x, \omega) \to \phi(X(t;x)(\omega)$ is $\overline{\mathscr{F} \otimes \mathcal{B}}/\mathcal{B}(\mathbb{R})$-measurable.

Therefore (see e.g. Theorem 2.39, p. 68 of [267]) $\mathbb{E}\left[\phi(X(t;x))\right] = \mathbb{E}\left[\tilde{\phi}(X(t;x))\right]$ is well defined for m-a.e. $x \in H$ and the function $x \to \mathbb{E}\left[\tilde{\phi}(X(t;x))\right]$ is $\overline{\mathcal{B}}/\mathcal{B}(\mathbb{R})$-measurable. However, for m-a.e. $x \in H$, $P_t[\phi](x) = P_t[\tilde{\phi}](x) = \mathbb{E}\left[\tilde{\phi}(X(t;x))\right] = \mathbb{E}\left[\phi(X(t;x))\right]$, which establishes (5.18).

The proof for a non-positive function follows by the previous arguments after decomposing the function into the sum of its positive and negative parts. □

5.2.3 The Operator $\mathcal{A}$

From now on we fix the constant p of Proposition 5.9 and Lemma 5.10 equal to 2 and work in the space $L^2(H, \overline{\mathcal{B}}, m)$.

Let Hypotheses 5.1 and 5.4 be satisfied and let P_t be defined as in (5.10). We denote by $\mathcal{A}$ the generator of P_t as a strongly continuous semigroup on $L^2(H, \overline{\mathcal{B}}, m)$ (see Proposition 5.9). Its domain is denoted by $D(\mathcal{A}) \subset L^2(H, \overline{\mathcal{B}}, m)$.

We will often use the elements of the space $\mathcal{F}C_0^{2,A^*}(H)$ to approximate less regular functions and it will be useful to know how to calculate explicitly the operator $\mathcal{A}$ on them. Indeed, as proved in Lemma 5.37, $\mathcal{F}C_0^{2,A^*}(H) \subset D(\mathcal{A})$ and for any $\phi \in \mathcal{F}C_0^{2,A^*}(H)$ we have

$$\mathcal{A}\phi(x) = \frac{1}{2} \operatorname{Tr}\left[QD^2\phi(x)\right] + \left\langle x, A^* D\phi(x)\right\rangle + \langle b(x), D\phi(x)\rangle. \tag{5.20}$$

5.2.4 The Gradient Operator D_Q and the Space $W_Q^{1,2}(H, m)$

Let Q be an operator satisfying Hypothesis 5.1-(B). We then introduce the following operator D_Q.

Definition 5.11 (*The operator D_Q and the space $W_Q^{1,2}(H, m)$*) We define the operator

$$D_Q\phi := Q^{1/2}D\phi, \quad \phi \in C_b^1(H), \tag{5.21}$$

where $D\phi$ denotes the Fréchet derivative of ϕ.

For $\phi \in C_b^1(H)$ we define the norm

$$|\phi|^2_{W_Q^{1,2}} = |\phi|^2_{L^2_m} + \left|D_Q\phi\right|^2_{L^2_{m,H}}.$$

The completion of $C_b^1(H)$ with respect to the norm $|\cdot|_{W_Q^{1,2}}$ will be denoted by $W_Q^{1,2}(H, m)$.

The space $W_Q^{1,2}(H, m)$ may be identified with the subspace of $L^2(H, \overline{\mathcal{B}}, m) \times L^2(H, \overline{\mathcal{B}}, m; H)$ which consists of all pairs

$$(\psi, \Psi) \in L^2(H, \overline{\mathcal{B}}, m) \times L^2(H, \overline{\mathcal{B}}, m; H)$$

such that there exists a sequence $(\phi_n) \subset C_b^1(H)$ with the property

$$\phi_n \to \psi, \quad \text{in} \quad L^2(H, \overline{\mathcal{B}}, m)$$

and

$$D_Q\phi_n \to \Psi, \quad \text{in} \quad L^2(H, \overline{\mathcal{B}}, m; H).$$

In the cases where the operator D_Q is closable (as an unbounded operator from its domain $C_b^1(H) \subset L^2(H, \overline{\mathcal{B}}, m)$ to $L^2(H, \overline{\mathcal{B}}, m; H)$), for any two pairs $(\psi_1, \Psi_1), (\psi_2, \Psi_2) \in W_Q^{1,2}(H, m)$ such that $\psi_1 = \psi_2$ in $L^2(H, \overline{\mathcal{B}}, m)$ we also have $\Psi_1 = \Psi_2$, so that $W_Q^{1,2}(H, m)$ is naturally embedded in $L^2(H, \overline{\mathcal{B}}, m)$.

If D_Q is not closable then we can find a sequence $(\phi_n) \subset C_b^1(H)$ such that

$$\phi_n \to 0 \quad \text{in} \quad L^2(H, \overline{\mathcal{B}}, m) \quad \text{and} \quad D_Q\phi_n \to \Phi \neq 0, \quad \text{in} \quad L^2(H, \overline{\mathcal{B}}, m; H).$$

Therefore, elements of $W_Q^{1,2}(H, m)$ cannot be identified, in general, with functions of $L^2(H, \overline{\mathcal{B}}, m)$ (e.g., the above element $(0, \Phi)$).[3] This means that the structure of

[3]For this reason, since we are interested in a definition that also works when the operator D_Q is non-closable, we do not work in the space $W^{1,2}(H, m)$ defined (see e.g. Chap. 9, p. 196 of [179]) as the linear space of all functions $\phi \in L^2(H, \overline{\mathcal{B}}, m)$ such that $D\phi \in L^2(H, \overline{\mathcal{B}}, m; H)$.

the Sobolev space changes significantly when we want to take into account the case of non-closable D_Q.

Observe that, in any case, even when D_Q is not closable, it can be extended to a well-defined continuous operator from $W_Q^{1,2}(H, m)$ (endowed with the norm described in Definition 5.11) to $L^2(H, \overline{\mathcal{B}}, m; H)$. Indeed, if $|\phi_n|^2_{W_Q^{1,2}} \to 0$ then $\left|D_Q\phi_n\right|^2_{L^2_{m,H}} \to 0$. We denote the continuous extension of D_Q from $W_Q^{1,2}(H, m)$ to $L^2(H, \overline{\mathcal{B}}, m; H)$ again by D_Q. When D_Q is not closable, considering the characterization of $W_Q^{1,2}(H, m)$ as a subspace of $L^2(H, \overline{\mathcal{B}}, m) \times L^2(H, \overline{\mathcal{B}}, m; H)$ and the notation described above, we have $D_Q(\psi, \Psi) = \Psi$.

The notation we use here is a little different from the one used in Chap. 4. Indeed, to be consistent with the notation of Chap. 4, we should write $D^{Q^{1/2}}$ instead of D_Q. We choose to use this notation for two reasons: it is simpler and, even if not very intuitive, it is fairly standard in the literature.

Sometimes in the literature the notation D_Q is used for different operators. We want to underline in particular the difference with respect to Chap. 9 of [179] where D_Q is used for the Malliavin derivative, which is again an operator of the form $Q^{\frac{1}{2}}D$ for some $Q \in \mathcal{L}_1^+(H)$. The difference is that, in our case Q is the covariance operator of the Wiener process, while in [179] it is the covariance operator of the (Gaussian) reference measure. When $b = 0$ and $\omega < 0$, the operator used in [179] is $Q_\infty = \int_0^{+\infty} e^{sA} Q e^{sA^*} ds$.

Remark 5.12 When (5.8) is linear (if $b = 0$) and $\omega < 0$, the problem of closability of D_Q can be approached using some characterizations that can be found in the literature. A negative result ensuring the non-closability of the operator is, for example, Theorem 3.5 of [299], which allows us to prove that D_Q is not closable, for example, in the two cases recalled in Sect. 5.6.

When the operator Q is injective, a characterization of closability is given by Theorem 6.1 of [299], which shows that the closability of the operator D_Q is equivalent to the closability of the operator $Z\colon D(Z) \subset H \to H$ given by

$$\begin{cases} D(Z) = Q_\infty^{1/2}(H) \\ Z\left(Q_\infty^{1/2}x\right) = Q^{\frac{1}{2}}x. \end{cases}$$

In the particular case considered, for example, in [3, 4, 125] (see also Example 4.46 and Sect. 4.8.3.1) the generator of the semigroup is

$$Ax = \sum_{n=1}^{+\infty} -\alpha_n \langle e_n, x \rangle e_n, \quad x \in D(A),$$

for some orthonormal basis $\{e_n\}$ and $0 < \alpha_1 \le \alpha_2 \le \alpha_3 \ldots$. Moreover Q is given by

$$Qx = \sum_{i=n}^{+\infty} q_n \langle e_n, x \rangle e_n, \quad x \in H,$$

for a sequence of positive eigenvalues q_n. The expression for Z is given, for any $y = Q_\infty^{1/2} x$, by

$$Zy = Q^{\frac{1}{2}} Q_\infty^{-1/2} y = \sum_{n=1}^{+\infty} \sqrt{q_n} \sqrt{\frac{2\alpha_n}{q_n}} \langle e_n, y \rangle e_n$$
$$= \sum_{n=1}^{+\infty} -\sqrt{2\alpha_n} \langle e_n, y \rangle e_n = \sqrt{2}(-A)^{1/2} y.$$

Thus, since $Q_\infty^{1/2}(H) \subset D\left((-A)^{1/2}\right)$ and since $(-A)^{1/2}$ is closed (see Theorem B.53-(i)), Z admits a closed extension and so (see Theorem 5.4(a), p. 91 of [569]) it is closable. Therefore, thanks to Theorem 6.1 of [299], the operator D_Q is closable. ∎

5.2.5 The Operator $\mathcal{R}$

Let Q be an operator satisfying Hypothesis 5.1-(B) and let D_Q be defined as in Definition 5.11. We introduce and study here the properties of the operator $\mathcal{R}$ defined below (Definition 5.19).

We begin by studying the regularity of the solution $X(\cdot; x)$ of (5.8) with respect to the initial datum. We use Proposition 6.7. The following lemma specifies it in the particular case we are interested in.

Lemma 5.13 *Let $\mathcal{H}_2^{\mu_0}(0, T; H)$ be the space defined in Definition 1.126. Let $\mathcal{K}\colon H \times \mathcal{H}_2^{\mu_0}(0, T; H) \to \mathcal{H}_2^{\mu_0}(0, T; H)$ be a continuous mapping satisfying, for some $\alpha \in [0, 1)$,*

$$|\mathcal{K}(x, X) - \mathcal{K}(x, Y)|_{\mathcal{H}_2^{\mu_0}(0,T;H)} \le \alpha \, |\mathcal{K}(x, X) - \mathcal{K}(x, Y)|_{\mathcal{H}_2^{\mu_0}(0,T;H)} \tag{5.22}$$

for all $x \in H$ and $X, Y \in \mathcal{H}_2^{\mu_0}(0, T; H)$. Then:

(i) There exists a unique mapping $\varphi\colon H \to \mathcal{H}_2^{\mu_0}(0, T; H)$ such that

$$\varphi(x) = \mathcal{K}(x, \varphi(x)), \qquad \text{for every } x \in H,$$

and it is continuous.

(ii) Suppose that, for any $(x, X) \in H \times \mathcal{H}_2^{\mu_0}(0, T; H)$ and for any $h \in H$ there exists the directional derivative of $\mathcal{K}$ with respect to x in the direction h and that, for any fixed h, the mapping

$$\begin{cases} H \times \mathcal{H}_2^{\mu_0}(0,T;H) \to \mathcal{H}_2^{\mu_0}(0,T;H) \\ (x,X) \to \nabla_x \mathcal{K}(x,X;h) \end{cases}$$

is continuous. Assume that, for any (x, X), $h \to \nabla_x \mathcal{K}(x, X; h)$ is continuous from H to $\mathcal{H}_2^{\mu_0}(0,T;H)$. Suppose also that for any $(x, X) \in H \times \mathcal{H}_2^{\mu_0}(0,T;H)$ and for any $Y \in \mathcal{H}_2^{\mu_0}(0,T;H)$ there exists the directional derivative of $\mathcal{K}$ with respect to X in the direction Y and that, for any fixed Y, the mapping

$$\begin{cases} H \times \mathcal{H}_2^{\mu_0}(0,T;H) \to \mathcal{H}_2^{\mu_0}(0,T;H) \\ (x,X) \to \nabla_X \mathcal{K}(x,X;Y) \end{cases}$$

is continuous. Assume that, for any (x, X), $Y \to \nabla_X \mathcal{K}(x, X; Y)$ is continuous from $\mathcal{H}_2^{\mu_0}(0,T;H)$ to $\mathcal{H}_2^{\mu_0}(0,T;H)$. Then, for any $x \in H$, there exists the Gâteaux derivative $\nabla\varphi(x)$. Moreover, $(x,h) \to \nabla\varphi(x)(h)$ is continuous as a mapping from $H \times H$ to $\mathcal{H}_2^{\mu_0}(0,T;H)$ and it satisfies the equation

$$\nabla\varphi(x)(h) = \nabla_x \mathcal{K}(x, \varphi(x); h) + \nabla_X \mathcal{K}(x, \varphi(x); \nabla\varphi(x)(h)), \qquad x, h \in H.$$

Proof This is a particular case of Proposition 6.7. In the claim of part (ii) we also made use of Lemma 6.4 (in a two-variable version) to verify the hypothesis "$F \in \mathcal{G}^{1,1}(X \times Y; X)$" of Proposition 6.7 for our spaces and of Lemma 6.3 to derive the continuity properties of $\nabla\varphi$. □

Lemma 5.14 *Let Hypothesis 5.1 be satisfied and let $x, h \in H$. Denote by $X(\cdot; x)$ the solution of (5.8). Then:*

(i) $X(\cdot; x)$ is Gâteaux differentiable as a mapping from H to $\mathcal{H}_2^{\mu_0}(0,T;H)$ and $x \to \nabla X(\cdot; x)$ is strongly continuous. For any $h \in H$ the (directional derivative) process $\zeta^{x,h}(\cdot) := \nabla X(\cdot; x)h$ is the unique mild solution in $\mathcal{H}_2^{\mu_0}(0,T;H)$ of the following equation

$$\begin{cases} \frac{d\zeta^{x,h}(s)}{ds} = (A + \nabla b(X(s;x))\, \zeta^{x,h}(s) \\ \zeta^{x,h}(0) = h \end{cases} \tag{5.23}$$

on $[0, T]$. The process $\zeta^{x,h}(\cdot)$ has $\mathbb{P}$-a.s. continuous trajectories.

(ii) There exist universal constants $\alpha, a > 0$, α also depends on K, such that

$$\left|\zeta^{x,h}(s)\right| \le a e^{\alpha s} |h|$$

for any $s \ge 0$. Therefore the solution to (5.23) defines, for any $x \in H$, $\omega \in \Omega$ and $s \ge 0$, a bounded operator $\zeta^x(s) : H \to H$, $\zeta^x(s)h = \zeta^{x,h}(s)$.

(iii) For any $h \in H$ there exists a $\mathcal{B}([0,T]) \otimes \mathcal{B}(H) \otimes \mathscr{F}/\mathcal{B}(H)$-measurable function

$$\begin{cases} [0,T] \times H \times \Omega \to H \\ (s,x,\omega) \to \tilde{\zeta}^{x,h}(s)(\omega) \end{cases} \tag{5.24}$$

such that, for every $x \in H$, $\tilde{\zeta}^{x,h}(\cdot)$ is a version of $\zeta^{x,h}(\cdot)$. Thus in the future we will not make a distinction between $\tilde{\zeta}^{x,h}(\cdot)$, $\zeta^{x,h}(\cdot)$, and $\nabla X(\cdot; x)h$.

Proof Since other similar results appearing in the book are proved for slightly different sets of hypotheses,[4] we provide the proofs.

To prove part (i), except for the $\mathbb{P}$-a.s. continuity of the trajectories of $\zeta^{x,h}(\cdot)$, we use Proposition 6.7 in the particular case stated in Lemma 5.13. The mapping $\mathcal{K}$ is defined as

$$\mathcal{K}(x, X)(s) = e^{sA}x + \int_0^s e^{(s-r)A} b(X(r))dr + W^A(s), \quad s \in [0, T],$$

where W^A is defined in (1.64). It is shown in the proof of Proposition 1.147 that if T is small enough then (5.22) is satisfied. The joint continuity of $\mathcal{K}$ is straightforward.

To verify the hypotheses of part (ii) of Lemma 5.13, we follow the arguments used in Sect. 9.1.1 of [180] (we repeat them because our hypotheses are a little different). The directional derivatives with respect to x are not a problem since one can easily see that $\nabla_x \mathcal{K}(x, X; h) = e^{\cdot A}h$ which is jointly continuous in all three variables.

As regards the directional derivative $\nabla_X \mathcal{K}(x, X; Y)$, we begin by showing that for any $X, Y \in \mathcal{H}_2^{\mu_0}(0, T; H)$ and any $x \in H$,

$$\nabla_X \mathcal{K}(x, X; Y)(s) = \int_0^s e^{(s-r)A} \nabla b(X(r))Y(r)dr, \quad s \in [0, T].$$

Indeed, we have

$$\sup_{s\in[0,T]} \mathbb{E} \left| \frac{1}{\varepsilon} \left(\mathcal{K}(x, X + \varepsilon Y) - \mathcal{K}(x, X)\right)(s) - \int_0^s e^{(s-r)A} \nabla b(X(r))Y(r)dr \right|^2$$
$$= \sup_{s\in[0,T]} \mathbb{E} \left| \int_0^s e^{(s-r)A} \left[\frac{1}{\varepsilon} \left(b(X(r) + \varepsilon Y(r)) - b(X(r)) - \nabla b(X(r))Y(r)\right) \right] dr \right|^2 .$$

Using Theorem D.18 the last expression above becomes

$$\sup_{s\in[0,T]} \mathbb{E} \left| \int_0^s e^{(s-r)A} \left[\int_0^1 \nabla b(X(r) + \theta\varepsilon Y(r))Y(r) - \nabla b(X(r))Y(r)d\theta \right] dr \right|^2$$
$$\le T \left(M \max\{e^{\omega T}, 1\}\right)^2 \mathbb{E} \int_0^T \left[\int_0^1 |\nabla b(X(r) + \theta\varepsilon Y(r))Y(r) - \nabla b(X(r))Y(r)|^2 \, d\theta \right] dr$$

which, thanks to the boundedness of ∇b and its strong continuity, converges to 0 when $\varepsilon \to 0$ by the Lebesgue dominated convergence theorem. We now prove the continuity properties of $\nabla_X \mathcal{K}(x, X; Y)$. We first fix (x, X) and we consider $Y_n \to Y$

[4] In particular, in Propositions 4.61 and 6.10 we work in $L^p_{\mathcal{P}}(\Omega; C([0, T], H))$, while here we use $\mathcal{H}_2^{\mu_0}(0, T; H)$. Indeed, in the mentioned propositions it is assumed that $\mathrm{Tr}\left[e^{sA}Qe^{sA^*}\right] \le C_\beta s^{-2\beta}$ for some $\beta \in [0, 1/2)$ and $C_\beta > 0$.

in $\mathcal{H}_2^{\mu_0}(0,T;H)$. We have, using Hypothesis 5.1 and Hölder's inequality,

$$
\begin{aligned}
&|\nabla_X \mathcal{K}(x, X; Y_n) - \nabla_X \mathcal{K}(x, X; Y)|^2_{\mathcal{H}_2^{\mu_0}(0,T;H)} \\
&\qquad = \sup_{s\in[0,T]} \mathbb{E}\left|\int_0^s e^{(s-r)A}\nabla b(X(r))(Y_n(r) - Y(r))dr\right|^2 \\
&\le K^2\left(M\max\{e^{\omega T}, 1\}\right)^2 \sup_{s\in[0,T]} s\mathbb{E}\int_0^s |Y_n(r) - Y(r)|^2 dr \le C|Y_n - Y|^2_{\mathcal{H}_2^{\mu_0}(0,T;H)} \to 0
\end{aligned}
$$

as $n \to +\infty$. To prove the strong continuity property we fix Y and suppose, to the contrary, that there are $\delta > 0$ and a sequence (x_n, X_n) such that $x_n \to x$ in H, $X_n \to X$ in $\mathcal{H}_2^{\mu_0}(0,T;H)$ but $|\nabla_X \mathcal{K}(x_n, X_n; Y) - \nabla_X \mathcal{K}(x, X; Y)|^2_{\mathcal{H}_2^{\mu_0}(0,T;H)} \ge \delta$ for any $n \in \mathbb{N}$. We have

$$
\begin{aligned}
&|\nabla_X \mathcal{K}(x_n, X_n; Y) - \nabla_X \mathcal{K}(x, X; Y)|^2_{\mathcal{H}_2^{\mu_0}(0,T;H)} \\
&\qquad = \sup_{s\in[0,T]} \mathbb{E}\left|\int_0^s e^{(s-r)A}\left[\nabla b(X_n(r)) - \nabla b(X(r))\right] Y(r)dr\right|^2 \\
&\qquad\qquad \le C\mathbb{E}\int_0^T \left|\left[\nabla b(X_n(r)) - \nabla b(X(r))\right] Y(r)\right|^2 dr,
\end{aligned}
$$

where C is a constant depending only on M, ω, T and K. For every $n \in \mathbb{N}$ the integrand in the last line above is dominated by $4K^2|Y(r)|^2$, moreover, since $X_n \to X$ in $\mathcal{H}_2^{\mu_0}(0,T;H)$ we can extract a subsequence X_{n_k} which converges to X, $dr \otimes \mathbb{P}$-a.e., and we can conclude using the Lebesgue dominated convergence theorem that $\left|\nabla_X \mathcal{K}(x_{n_k}, X_{n_k}; Y) - \nabla_X \mathcal{K}(x, X; Y)\right|^2_{\mathcal{H}_2^{\mu_0}(0,T;H)} \to 0$ as $k \to +\infty$, which contradicts our hypothesis.

Thus part (i) follows from Lemma 5.13. The continuity of the trajectories of $\zeta^{x,h}(\cdot)$ is a consequence of Lemma 1.115.

To prove part (ii) we observe that, thanks to Hypothesis 5.1 (A) and (D), we have, for all $s \in [0,T]$,

$$
\left|\zeta^{x,h}(s)\right| \le \left|M\max\{e^{\omega T}, 1\}\right| |h| + M\max\{e^{\omega T}, 1\}\int_0^s K|\zeta^{x,h}(r)|dr
$$

and hence the conclusion follows from Gronwall's lemma (Proposition D.29).

To prove the claim of part (iii), we use the result of Proposition 5.44 (even though Proposition 5.44 is in a later section, its proof is independent). Let $(s, x, \omega) \to \tilde{X}(s; x)(\omega)$ be the $\mathcal{B}([t,T]) \otimes \mathcal{B}(H) \otimes \mathscr{F}/\mathcal{B}(H)$-measurable function found in Proposition 5.44 (we consider here the case when $t = 0$ and $R = 0$). Observe that, by construction, $\tilde{X}$ satisfies (1.70) for any $s \in [0,T]$, any $x, y \in H$ and any $\omega \in \Omega$ and in particular it is continuous in the variable x for any choice of $(s,\omega) \in [0,T] \times \Omega$.

We denote by $\tilde{\zeta}^{x,h}(\cdot)$ the unique solution of

$$\tilde{\zeta}^{x,h}(s) = e^{As}h + \int_0^s e^{(s-r)A}\nabla b(\tilde{X}(r;x))\tilde{\zeta}^{x,h}(r)\,dr, \qquad s \in [0,T].$$

We remark that $\tilde{\zeta}^{x,h}(s)$ is defined for every $(s,\omega) \in [0,T] \times \Omega$. Since $\tilde{X}(\cdot;x)$ is a version of $X(\cdot;x)$, $\tilde{\zeta}^{x,h}(\cdot)$ is a version of $\zeta^{x,h}(\cdot)$. Moreover, we claim that, for any choice of $(s,\omega) \in [0,T] \times \Omega$, $\tilde{\zeta}^{x,h}(s)$ is continuous in the variable x. To prove this we fix $\omega \in \Omega$ and consider $x \in H$ and any sequence x_n in H converging to x. We have

$$\begin{aligned}\left|\tilde{\zeta}^{x,h}(s)(\omega) - \tilde{\zeta}^{x_n,h}(s)(\omega)\right| &\le I_1^n(s) + I_2^n(s)\\ &:= \left|\int_0^s e^{(s-r)A}\left(\nabla b(\tilde{X}(r;x))(\omega) - \nabla b(\tilde{X}(r;x_n))(\omega)\right)\tilde{\zeta}^{x,h}(r)(\omega)\,dr\right|\\ &+ \left|\int_0^s e^{(s-r)A}\nabla b(\tilde{X}(r;x_n))(\omega)\left(\tilde{\zeta}^{x,h}(r)(\omega) - \tilde{\zeta}^{x_n,h}(r)(\omega)\right)dr\right|, \quad s \in [0,T].\end{aligned}$$

$I_1^n(s)$ converges to zero, uniformly for $s \in [0,T]$, thanks to the Lebesgue dominated convergence theorem as Hypothesis 5.1 (A) and (D) and part (ii) give the uniform bound and the continuity of $x \to \tilde{X}(r;x)(\omega)$ gives the pointwise convergence. Thus the convergence (which is indeed uniform in s and thus even stronger than what we need) of $\left|\tilde{\zeta}^{x,h}(s)(\omega) - \tilde{\zeta}^{x_n,h}(s)(\omega)\right| \to 0$ follows from Gronwall's Lemma (using again Hypothesis 5.1 (A) and (D) which gives $|e^{(s-r)A}\nabla b(\tilde{X}(r;x_n))(\omega)| \le K\left(M \vee Me^{\omega T}\right)$ independently of s, r, n, ω).

Since $\tilde{\zeta}^{x,h}(\cdot)$ has continuous trajectories and is a version of $\zeta^{x,h}(\cdot) \in \mathcal{H}_2^{\mu_0}(0,T;H)$, it itself belongs to $\mathcal{H}_2^{\mu_0}(0,T;H)$. In particular, for every $x \in H$, $\tilde{\zeta}^{x,h}(\cdot)$ is $\mathcal{B}([0,T]) \otimes \mathscr{F}/\mathcal{B}(H)$-measurable as function of the variables s and ω. Moreover, we proved that, for any fixed $(s,\omega) \in [0,T] \times \Omega$, $\tilde{\zeta}^{x,h}(s)(\omega)$ is a continuous function of the variable x. It then follows from Lemma 1.18, that $\tilde{\zeta}^{x,h}(s)(\omega)$ is $\mathcal{B}([t,T]) \otimes \mathcal{B}(H) \otimes \mathscr{F}/\mathcal{B}(H)$-measurable. □

Lemma 5.15 *Assume that Hypotheses 5.1 and 5.4 hold. Fix $t \in [0,T]$. Given $\phi \in C_b^1(H)$, $P_t[\phi] \in C_b(H)$, $P_t[\phi]$ is Gâteaux differentiable at any $x \in H$ and*

$$\langle \nabla P_t[\phi](x), h\rangle = \mathbb{E}\left(\left\langle\left(\zeta^x(t)\right)^* D\phi(X(t;x)), h\right\rangle\right), \quad h \in H. \tag{5.25}$$

Moreover, $\nabla P_t[\phi]$ is strongly continuous and

$$\sup_{x \in H} |\nabla P_t[\phi](x)| < +\infty. \tag{5.26}$$

Proof The continuity of $P_t[\phi]$ follows from Theorem 1.162. Differentiating $P_t[\phi]$ and using its definition we obtain

$$\langle \nabla P_t[\phi](x), h\rangle = \mathbb{E}\,\langle D\phi(X(t;x)), \nabla(X(t;x))h\rangle$$

so (5.25) follows from Lemma 5.14. The strong continuity of the differential can be proved as follows. Given $h \in H$ and $t > 0$, consider a sequence x_n of elements of H converging to $x \in H$. We have

$$\begin{aligned}
&\langle \nabla P_t[\phi](x), h\rangle - \langle \nabla P_t[\phi](x_n), h\rangle \\
&\quad = \mathbb{E}\left(\left\langle D\phi(X(t;x)), \zeta^{x,h}(t)\right\rangle\right) - \mathbb{E}\left(\left\langle D\phi(X(t;x_n)), \zeta^{x_n,h}(t)\right\rangle\right) \\
&\quad \le \mathbb{E}\left|\left\langle D\phi(X(t;x)) - D\phi(X(t;x_n)), \zeta^{x,h}(t)\right\rangle\right| \\
&\qquad + \mathbb{E}\left|\left\langle D\phi(X(t;x_n)), \zeta^{x_n,h}(t) - \zeta^{x,h}(t)\right\rangle\right| \\
&\quad \le I_1(n) + I_2(n) := ae^{\alpha t}|h|\,\mathbb{E}\,|D\phi(X(t;x)) - D\phi(X(t;x_n))| \\
&\qquad + \|D\phi\|_0\,\mathbb{E}\left|\zeta^{x_n,h}(t) - \zeta^{x,h}(t)\right|,
\end{aligned}$$

where $ae^{\alpha t}|h|$ is introduced in Lemma 5.14.

$I_1(n)$ converges to 0 when $n \to +\infty$ thanks to the dominated convergence theorem, the boundedness and the continuity of $D\phi$ and (1.70). Observe that, since $\{x_n\}_{n\in\mathbb{N}}$ is countable, we can find a subset of Ω of measure 1 where (1.70) holds for any n (with x_n and x as ξ_1 and ξ_2, respectively, moreover $f(r)$ appearing in (1.70) is, in our case, just a positive constant independent of r).

For $I_2(n)$ observe that

$$\begin{aligned}
&\mathbb{E}\left|\zeta^{x_n,h}(t) - \zeta^{x,h}(t)\right| \\
&\quad = \mathbb{E}\left|\int_0^t e^{(t-s)A}\left(\nabla b(X(s,x_n))\zeta^{x_n,h}(s) - \nabla b(X(s,x))\zeta^{x,h}(s)\right) ds\right| \\
&\quad \le \mathbb{E}\Bigg[C\int_0^t \left|(\nabla b(X(s,x) - \nabla b(X(s,x_n)))\,\zeta^{x,h}(s)\right| ds \\
&\qquad + C\int_0^t \left|\nabla b(X(s,x_n))\left(\zeta^{x,h}(s) - \zeta^{x_n,h}(s)\right)\right| ds\Bigg] \\
&\quad \le \mathbb{E}\left[C\int_0^t \left|(\nabla b(X(s,x) - \nabla b(X(s,x_n)))\,\zeta^{x,h}(s)\right| ds\right] \\
&\qquad + CK\int_0^t \mathbb{E}\left|\left(\zeta^{x_n,h}(s)\right) - \left(\zeta^{x,h}(s)\right)\right| ds
\end{aligned} \tag{5.27}$$

for some positive constant C coming from Hypothesis 5.1-(A) and with K from Hypothesis 5.1-(D). Thanks to the strong continuity of ∇b, the boundedness of $\|\nabla b\|_0$ and of $|\zeta^{x,h}(s)|$ (Hypothesis 5.1-(D) and Lemma 5.14), (1.70) (recall again that we can find a subset of Ω of measure 1 where (1.70) holds for any n) and the dominated convergence theorem, the term

$$\mathbb{E}\left[C\int_0^t \left|(\nabla b(X(s,x) - \nabla b(X(s,x_n)))\,\zeta^{x,h}(s)\right| ds\right]$$

converges to 0 when $n \to \infty$. Thus we can apply Gronwall's Lemma to (5.27) and conclude that $I_2(n)$ converges to 0 when $n \to +\infty$. This concludes the proof of the strong continuity of $DP_t[\phi]$.

The bound (5.26) follows from the bound for the Gâteaux differential of X proved in Lemma 5.14 and the hypotheses on ϕ. □

Corollary 5.16 *Assume that Hypotheses 5.1 and 5.4 hold. For any $\phi \in C_b^1(H)$, $P_t[\phi] \in W_Q^{1,2}(H, m)$. In particular, $D_Q P_t[\phi]$ is well defined and it equals $Q^{1/2}\nabla P_t[\phi]$.*

Proof Thanks to Lemma 5.15, $P_t[\phi]$ satisfies Hypothesis 5.1-(D) and it is bounded so we can apply to it Lemma 5.8-(i)(ii). The conclusion follows by the characterization of $W_Q^{1,2}(H, m)$ given after Definition 5.11. □

Lemma 5.17 *Let Hypothesis 5.1 be satisfied, let b_n be as in Part (i) of Lemma 5.8, let $x \in H$ and $X(\cdot) = X(\cdot; x)$ be the solution of (5.8). The following hold:*

(i) If, for some sequence x_n converging to x in H, we denote by $X_n(\cdot) = X_n(\cdot; x_n)$ the unique solution of the equation

$$\begin{cases} dX_n(s) = (AX_n(s) + b_n(X_n(s)))\,dt + dW_Q(s), \\ X(0) = x_n, \end{cases} \tag{5.28}$$

then, for any $p > 1$,

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \mathbb{E}\,|X_n(t; x_n) - X(t; x)|^p = 0. \tag{5.29}$$

(ii) Let $X_n(\cdot)$, x_n be as in Part (i) above. Denote by $\zeta_n^{x_n,h}(\cdot)$ the solution of (5.23), where $X(\cdot)$ is replaced by $X_n(\cdot)$, b by b_n and x by x_n. Then, for any $p > 1$,

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \mathbb{E}\left(\sup_{|h|\le 1} \left|\zeta^{x,h}(t) - \zeta_n^{x_n,h}(t)\right|\right)^p = 0. \tag{5.30}$$

Proof For Part (i) we observe that for any $t \in [0, T]$

$$\begin{aligned} X_n(t; x_n) - X(t; x) &= e^{tA}(x_n - x) + \int_0^t e^{(t-s)A}(b_n(X_n(s; x_n)) - b(X(s; x)))ds \\ &= e^{tA}(x_n - x) + \int_0^t e^{(t-s)A}\big([b_n(X_n(s; x_n)) - b_n(X(s; x))] \\ &\qquad\qquad + [b_n(X(s; x)) - b(X(s; x))]\big)ds \end{aligned}$$

and thus

$$\begin{aligned}
&\mathbb{E}|X_n(t;x_n) - X(t;x)|^p \\
&\qquad \le C_T|x_n - x|^p + C_T\int_0^t l^p\mathbb{E}|X_n(s;x_n) - X(s;x)|^p ds \\
&\qquad\qquad + C_T\int_0^T \mathbb{E}|b_n(X(s;x)) - b(X(s;x))|^p ds
\end{aligned}$$

for a constant C_T depending on T. For any $s \in [0,T]$ the expression $\mathbb{E}|b_n(X(s;x)) - b(X(s;x))|^p$ converges to 0 thanks to Lemma 1.51 if we use Lemma 5.8 and the uniform moment estimates of (1.69). The claim thus follows by applying Gronwall's Lemma.

The argument for Part (ii) is similar. Indeed, for any $t \in [0,T]$,

$$\begin{aligned}
(\zeta_n^{x_n,h}(t) - \zeta^{x,h}(t)) &= \int_0^t e^{(t-s)A}[Db_n(X_n(s;x_n))\zeta_n^{x_n,h}(s) - \nabla b(X(s;x))\zeta^{x,h}(s)]ds \\
&= \int_0^t e^{(t-s)A}\big[Db_n(X_n(s;x_n)(\zeta_n^{x_n,h}(s) - \zeta^{x,h}(s)) \\
&\qquad + (Db_n(X_n(s;x_n)) - \nabla b(X(s;x)))\zeta^{x,h}(s)\big]ds.
\end{aligned}$$

So,

$$\begin{aligned}
\sup_{|h|\le 1}\left|\zeta_n^{x_n,h}(t) - \zeta^{x,h}(t)\right| \le C_T\int_0^t \Big[&\|Db_n(X_n(s;x_n))\| \sup_{|h|\le 1}\left|\zeta_n^{x_n,h}(s) - \zeta^{x,h}(s)\right| \\
&+ |(Db_n(X_n(s;x_n)) - \nabla b(X(s;x)))\zeta^{x,h}(s)|ds.
\end{aligned}$$

By taking the p-th powers and the expectations of the two sides and then using (5.13) we obtain, for a different constant C_T,

$$\mathbb{E}\left(\sup_{|h|\le 1}\left|\zeta_n^{x_n,h}(t) - \zeta^{x,h}(t)\right|\right)^p \le C_T\int_0^t K^p\mathbb{E}\left(\sup_{|h|\le 1}\left|\zeta_n^{x_n,h}(s) - \zeta^{x,h}(s)\right|\right)^p ds + I_n,$$

where

$$I_n := C_T\int_0^T \mathbb{E}\Big[|(Db_n(X_n(s;x_n)) - \nabla b(X(s;x)))\zeta^{x,h}(s)|^p\Big]ds.$$

All we need to do now is to prove that I_n converges to 0. Then the claim will be a direct consequence of Gronwall's Lemma. To show this it is enough to show that for any subsequence I_{n_k} there exists a sub-subsequence converging to 0.

Let us then consider a subsequence of X_n (denoted again by X_n). Thanks to (5.29),

$$\int_0^T \mathbb{E}\left[|X_n(s,x_n) - X(s,x)|^p\right]ds \xrightarrow{n\to\infty} 0$$

and then we can extract a subsequence (denoted again by X_n) such that $X_n(\cdot, x_n)(\cdot)$ converges $(ds \otimes \mathbb{P})$-a.e. to $X(\cdot, x)(\cdot)$ (ds denotes the Lebesgue measure on $\mathbb{R}$). So, using Lemma 5.8-(i), $|(Db_n(X_n(\cdot; x_n)) - \nabla b(X(\cdot; x)))\zeta^{x,h}(\cdot)|^p$ converges to 0, $(ds \otimes \mathbb{P})$-a.e. Since, by (5.13) and the bound on $|\zeta^{x,h}|$ given by Lemma 5.14, these functions are bounded uniformly in n, we can thus conclude using the dominated convergence theorem that $I_n \to 0$. □

Lemma 5.18 *Assume that Hypotheses 5.1 and 5.4 hold and $\phi \in C_b^1(H)$. Then, for any $t \in [0, T]$,*

$$\phi(X(t; x)) = P_t[\phi](x) + \int_0^t \left\langle \nabla P_{t-s}[\phi](X(s; x)), dW_Q(s) \right\rangle \qquad \mathbb{P}\ a.e. \tag{5.31}$$

Proof Step 1. The claim is proved for $b \in UC_b^2(H, H)$ and $\phi \in UC_b^2(H)$ in [582], Lemma 6.11, p. 181.

To extend the result to the general case, in the next step we will consider $\phi \in UC_b^2(H)$ and b which satisfies Hypothesis 5.1-(D), and in the third step we will prove the result in full generality.

Step 2. Consider $\phi \in UC_b^2(H)$ and b satisfying Hypothesis 5.1-(D). Let b_n be the sequence found in Part (i) of Lemma 5.8, $X_n(t; x)$ be the solution of (5.28) with $x_n = x$, and $P_t^n[\phi](x) = \mathbb{E}\phi(X_n(t; x))$, $t \geq 0$, be the corresponding transition semigroup.

Thanks to (5.29) (with $x_n = x$), up to extracting a subsequence,

$$\lim_{n \to +\infty} X_n(s; x) = X(s; x) \qquad \text{for } ds \otimes \mathbb{P}\text{-almost any } (s, \omega) \in [0, T] \times \Omega. \tag{5.32}$$

Observe now that, for any $x \in H$,

$$\lim_{n \to \infty} P_t^n[\phi](x) = P_t[\phi](x). \tag{5.33}$$

Indeed, we have

$$\left| P_t^n[\phi](x) - P_t[\phi](x) \right| = \mathbb{E}\left|\phi(X_n(t; x)) - \phi(X(t; x))\right| \leq C\mathbb{E}\left|X_n(t; x) - X(t; x)\right|^2$$

so the claim follows from (5.29). Observe also that by Lemma 5.15 we have

$$\nabla P_t[\phi](x) = \mathbb{E}\left(\left(\zeta^x(t)\right)^* D\phi(X(t; x))\right), \quad \nabla P_t^n[\phi](x) = \mathbb{E}\left(\left(\zeta_n^x(t)\right)^* D\phi(X_n(t; x))\right).$$

Thus, using (5.29), (5.30), and a universal bound on $\|\zeta_n^x(t)\|$ given by Lemma 5.14, we easily obtain that

$$\sup_{n,x,t}(|\nabla P_t^n[\phi](x)| + |\nabla P_t[\phi](x)|) \leq C \tag{5.34}$$

for some constant C. We can then conclude using the dominated convergence theorem if we can show that for almost every $s \in [0,t]$, $\lim_{n\to\infty} \nabla P^n_{t-s}[\phi](X_n(s;x)(\omega)) = \nabla P_{t-s}[\phi](X(s;x)(\omega))$ for $\mathbb{P}$-a.e. ω. In fact, we prove this convergence for any (s,ω) where (5.32) holds.

Given $(s,\omega) \in [0,t] \times \Omega$ where the convergence (5.32) holds, we rewrite it as $y_n := X_n(s,x)(\omega)$, $y_n \xrightarrow{n\to\infty} y := X(s,x)(\omega)$ in H. By Lemma 5.15,

$$\begin{aligned}
&\left|\nabla P^n_{t-s}[\phi](y_n) - \nabla P_{t-s}[\phi](y)\right| \\
&\qquad = \sup_{h\in H,\ |h|\le 1} \left|\left\langle \nabla P^n_{t-s}[\phi](y_n) - \nabla P_{t-s}[\phi](y), h\right\rangle\right| \\
&\qquad \le I_1^n + I_2^n := \sup_{h\in H,\ |h|\le 1} \left|\mathbb{E}\left\langle D\phi(y_n), \zeta_n^{y_n,h}(t-s) - \zeta^{y,h}(t-s)\right\rangle\right| \\
&\qquad\qquad + \sup_{h\in H,\ |h|\le 1} \left|\mathbb{E}\left\langle D\phi(y_n) - D\phi(y), \zeta^{y,h}(t-s)\right\rangle\right|.
\end{aligned}$$

We have

$$I_1^n \le \|D\phi\|_0 \left(\mathbb{E} \sup_{h\in H,\ |h|\le 1} \left|\zeta_n^{y_n,h}(t-s) - \zeta^{y,h}(t-s)\right|\right),$$

which converges to 0 by (5.30). Moreover, $I_2^n \to 0$ thanks to the boundedness of $\zeta^y(t-s)$ given by Lemma 5.14 and the continuity of $D\phi$.

The result is thus true for any $\phi \in UC_b^2(H)$ and b satisfying Hypothesis 5.1-(D).

Step 3. Assume now that b satisfies Hypothesis 5.1-(D) and $\phi \in C_b^1(H)$. Let ϕ_n be the approximating sequence described in Part (iii) of Lemma 5.8. We have, for any $x \in H$,

$$\lim_{n\to\infty} P_t[\phi_n](x) = P_t[\phi](x). \tag{5.35}$$

Indeed,

$$|P_t[\phi_n](x) - P_t[\phi](x)| = \mathbb{E}\,|\phi_n(X(t;x)) - \phi(X(t;x))|,$$

which converges to 0 thanks to Lemma 5.8-(iii) and the dominated convergence theorem. Moreover, for any $x \in H$, by Lemma 5.15,

$$\begin{aligned}
&|\nabla P_{t-s}[\phi_n](x) - \nabla P_{t-s}[\phi](x)| \\
&\qquad = \sup_{h\in H,\ |h|\le 1} |\langle \nabla P_{t-s}[\phi_n](x) - \nabla P_{t-s}[\phi](x), h\rangle| \\
&\qquad = \sup_{h\in H,\ |h|\le 1} \left|\mathbb{E}\left\langle D\phi_n(x) - D\phi(x), \zeta^{x,h}(t-s)\right\rangle\right| \\
&\qquad\qquad \le |D\phi_n(x) - D\phi(x)| \sup_{h\in H,\ |h|\le 1} \mathbb{E}\left|\zeta^{y,h}(t-s)\right|,
\end{aligned}$$

which converges to 0 thanks to Lemmas 5.8-(iii) and 5.14. □

We define the operator $\mathcal{R}$ as follows.

Definition 5.19 (*The operator $\mathcal{R}$*) Given $\phi \in C_b^1(H)$, we define for any $t \in [0, T]$,

$$(\mathcal{R}\phi)(t) := D_Q P_t[\phi].$$

The operator $\mathcal{R}$ is well defined thanks to Corollary 5.16.

The next proposition provides an identity which allows us to extend the operator $\mathcal{R}$ to the whole space $L^2(H, \overline{\mathcal{B}}, m)$.

Proposition 5.20 *Assume that Hypotheses 5.1 and 5.4 hold. For every $\phi \in C_b^1(H)$*

$$\int_0^T \left|D_Q P_t[\phi]\right|^2_{L^2_m} dt = |\phi|^2_{L^2_m} - |P_T[\phi]|^2_{L^2_m} . \tag{5.36}$$

Moreover, the operator $\mathcal{R}$ has a unique extension to a bounded operator

$$\mathcal{R} : L^2(H, \overline{\mathcal{B}}, m) \to L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m)\right),$$

with

$$|(\mathcal{R}\phi)|^2_{L^2\left(0,T;L^2(H,\overline{\mathcal{B}},m)\right)} = \int_0^T |(\mathcal{R}\phi)(t)|^2_{L^2_m} dt = |\phi|^2_{L^2_m} - |P_T[\phi]|^2_{L^2_m} \tag{5.37}$$

for any $\phi \in L^2(H, \overline{\mathcal{B}}, m)$.

Proof Let $\phi \in C_b^1(H)$. Then (5.31) yields

$$\mathbb{E}[\phi^2(X(T,x))] = (P_T[\phi](x))^2 + \int_0^T \mathbb{E}\left|Q^{1/2}\nabla P_{T-t}[\phi](X(t,x))\right|^2 dt.$$

Recall that, by Corollary 5.16, since $\phi \in C_b^1(H)$, we have $D_Q P_t[\phi] = Q^{1/2}\nabla P_t[\phi]$. Thus, integrating the previous identity with respect to m and rearranging the terms we get

$$\int_0^T \int_H \mathbb{E}\left|D_Q P_t[\phi](X(t,x))\right|^2 dm(x)dt$$
$$= \int_0^T \int_H \mathbb{E}[\phi^2(X(T,x))dm(x)] - \int_0^T \int_H (P_T[\phi](x))^2 dm(x),$$

so, by using the invariant measure property (5.11), we obtain (5.36) for all $\phi \in C_b^1(H)$. The result follows thanks to the density of $C_b^1(H)$ in $L^2(H, \overline{\mathcal{B}}, m)$ (Lemma 5.7-(i)). □

Remark 5.21 In the particular case where $b = 0$, the operator $\mathcal{A}$ reduces to the Ornstein–Uhlenbeck operator and the semigroup P_t is called the Ornstein–Uhlenbeck semigroup. In particular, if $\|e^{At}\| \le Me^{-\omega t}$ with $M \in \mathbb{R}$ and $\omega > 0$ (the condition assumed in the whole remark), the invariant measure for P_t is the Gaussian measure $N(0, Q_\infty)$, where

$$Q_\infty := \int_0^{+\infty} e^{sA} Q e^{sA^*} ds.$$

In this case there are links between the closability of the operator D_Q, the smoothing properties of the semigroup P_t and the characteristics of certain controllability problems:

(1) If we consider the following linear controlled system,

$$\frac{dX(t)}{dt} = AX(t) + Q^{1/2}a(t), \quad X(0) = 0, \tag{5.38}$$

the set of points of H that can be reached by the system in an infinite time using a control in the set $L^2(0, +\infty; H)$ is equal to $Q_\infty^{1/2}(H)$ (see [584], Theorem 2.3, page 210) and it can be proved (see [299], Theorem 6.1) that the closability of the operator D_Q is equivalent to the density of the set

$$\left\{x \in H : Q^{1/2}x \in Q_\infty^{1/2}(H)\right\}$$

in H.

(2) Fix $t > 0$. The null-controllability in time t of the system

$$\frac{dX(t)}{dt} = AX(t) + Q^{1/2}a(t), \qquad X(0) = x,$$

is defined as the capability, by choosing a suitable control in $L^2(0, t; H)$, of reaching at time t the point 0, given any initial condition $x \in H$. The null-controllability of the described system (see [584], Theorem 2.3, p. 210) is equivalent to the condition

$$e^{tA}(H) \subset Q_t^{1/2}(H).$$

This condition is equivalent (see Theorem 2.23, p. 53 of [180]) to the fact that all the transition probabilities are mutually absolutely continuous and (see Theorem 9.26, p. 260 and Remark 9.29, p. 265 of [180]) to the fact that the semigroup P_t is strong Feller (see Definition 1.159).

By the results of Sect. 4.3.1, given $\phi \in L^2(H, \overline{\mathcal{B}}, m)$, it can be seen that $\nabla P_t[\phi]$ is well defined for $t > 0$ if and only if (5.39) is satisfied (see Hypothesis 4.29, Remark 4.30 and Theorem 4.37). In this case (see Proposition 10.3.1, page 218 of [179]) the singularity of $|\nabla P_t[\phi]|_{L^2_{m,H}}$ at $t = 0^+$, similarly to the one of $|\nabla P_t[\phi]|_0$, is estimated from above by $\|\Gamma(t)\|$, where as in (4.59), $\Gamma(t) := Q_t^{-1/2}e^{tA}$. Similarly, $D_Q P_t[\phi]$ is well defined for $\phi \in L^2(H, \overline{\mathcal{B}}, m)$ and $t > 0$ if and only if

$$e^{tA}Q^{1/2}(H) \subset Q_t^{1/2}(H), \tag{5.39}$$

i.e. if and only if every point of $Q^{1/2}(H)$ is null controllable in time t (see again Hypothesis 4.29, Remark 4.30 and Theorem 4.41 when $G = Q^{1/2}$). In this case the singularity of $\left|D_Q P_t[\phi]\right|_{L^2_{m,H}}$ at 0^+ has the same behavior as the norm of the operator

$$\Gamma_{Q^{1/2}}(t) := Q_t^{-1/2} e^{tA} Q^{1/2}.$$

More on this subject can be found in [120], Sect. 10.3 of [179], Sect. 5.3 of [431, 432].

The observations of part (2) are useful to provide examples where the approach of the previous chapter cannot be applied while the theory of this chapter works. This is the case when the hypotheses of this chapter hold but (5.39) does not hold or when it holds but $\|\Gamma_{Q^{1/2}}(t)\|$ is not integrable at 0^+. Such examples are, for instance, delay equations (see Sect. 5.6.1), where the semigroup can never be strong Feller for t smaller than the delay appearing in the equation (r in Sect. 5.6.1) or certain classes of second-order SPDEs in the whole space, see Sect. 5.6.3. ■

Remark 5.22 If D_Q is closable in $L^2(H, \overline{\mathcal{B}}, m)$ then $\mathcal{R}(\phi)(t) = \overline{D_Q} P_t[\phi](t)$ for all $t > 0$ and $\phi \in L^2(H, \overline{\mathcal{B}}, m)$. In this case (5.36) is easier to obtain and the whole study of the HJB equation (5.1) is simpler. This is true, in particular, when Q is boundedly invertible. ■

5.2.6 Two Key Lemmas

Here we use Proposition 5.20 to provide two estimates that will be essential in the following. We begin with an estimate regarding the convolution of P_t.

Lemma 5.23 *Assume that Hypotheses 5.1 and 5.4 hold and let P_t be defined as in (5.10). Given $f \in L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m)\right)$ we define*

$$G_1 f(t) := \int_t^T P_{s-t}[f(s)]\, ds, \quad t \in [0, T],$$

and

$$G_2 f(t) := \int_t^T \mathcal{R}\left(f(s)\right)(s-t)\, ds, \quad t \in [0, T].$$

Then

$$\int_0^T |G_1 f(t)|^2_{L^2_m}\, dt \le T^2 \int_0^T |f(t)|^2_{L^2_m}\, dt, \tag{5.40}$$

$G_2 f(t) \in L^2(H, \overline{\mathcal{B}}, m; H)$ for almost every $t \in [0, T]$ and

$$\int_0^T |G_2 f(t)|^2_{L^2_{m,H}}\, dt \le T \int_0^T |f(t)|^2_{L^2_m}\, dt. \tag{5.41}$$

Proof For the first estimate, observe that

$$\int_0^T |G_1 f(t)|^2_{L^2_m}\, dt = \int_0^T \left| \int_t^T P_{s-t}[f(s)]ds \right|^2_{L^2_m} dt$$
$$\leq \int_0^T \left(\int_t^T |P_{s-t}[f(s)]|_{L^2_m}\, ds \right)^2 dt \leq \int_0^T \left(\int_0^T |f(s)|_{L^2_m}\, ds \right)^2 dt$$
$$\leq \int_0^T T \int_0^T |f(s)|^2_{L^2_m}\, dsdt = T^2 \int_0^T |f(s)|^2_{L^2_m}\, ds.$$

We prove the second inequality. Assume first that $f \in C_b^1([0,T] \times H)$ and $f(t) \in \mathcal{F}C_0^1(H)$ (defined in Sect. A.2) for all $t \geq 0$. Then $D_Q P_{s-t}[f(s)]$ is well defined for $s \geq t$ and so is $D_Q G_1(t)$ for $t > 0$. Moreover,

$$\int_0^T |G_2 f(t)|^2_{L^2_{m,H}}\, dt \leq \int_0^T \left(\int_t^T \left| D_Q P_{s-t}[f(s)] \right|_{L^2_{m,H}} ds \right)^2 dt$$

$$\leq \int_0^T T \int_t^T \left| D_Q P_{s-t}[f(s)] \right|^2_{L^2_{m,H}} dsdt = T \int_0^T \int_0^s \left| D_Q P_r[f(s)] \right|^2_{L^2_{m,H}} drds$$
$$\leq \int_0^T T \int_0^T \left| D_Q P_r[f(s)] \right|^2_{L^2_{m,H}} drdt.$$

Hence by (5.36),

$$\int_0^T |G_2(t)|^2_{L^2_{m,H}}\, dt \leq T \int_0^T |f(t)|^2_{L^2_m} dt.$$

If $f \in L^2\left(0,T; L^2(H, \overline{\mathcal{B}}, m)\right)$ is arbitrary, then, thanks to Lemma 5.7 applied to the space $[0,T] \times H$, there exists a sequence $f_n \in C_b^1([0,T] \times H)$, with $f_n(t) \in \mathcal{F}C_0^1(H)$ for any $t \in [0,T]$, which converges to f in $L^2\left(0,T; L^2(H, \overline{\mathcal{B}}, m)\right)$. Repeating the above arguments for

$$G_1^n(t) = \int_t^T P_{s-t}[f_n(s)]ds$$

we find that

$$\int_0^T \left| D_Q \left(G_1^n(t) - G_1^m(t) \right) \right|^2_{L^2_{m,H}} dt \leq T \int_0^T |f_n(t) - f_m(t)|^2_{L^2_m} dt.$$

Hence the sequence $D_Q G_1^n$ is convergent in $L^2\left(0,T; L^2(H, \overline{\mathcal{B}}, m; H)\right)$. Moreover, by the Fubini Theorem,

$$\int_0^T \left| D_Q G_1^n(t) - G_2(t) \right|^2_{L^2_{m,H}} dt$$

$$= \int_0^T \left| \int_t^T \left[D_Q P_{s-t}[f_n(s)]ds - \mathcal{R}\left(f(s)\right)(s-t) \right] ds \right|^2_{L^2_{m,H}} dt$$

$$\leq T \int_0^T ds \int_0^T \left| D_Q P_t[f_n(s)] - \mathcal{R}\left(f(s)\right)(t) \right|^2_{L^2_{m,H}} dt$$

$$= T \int_0^T ds \int_0^T \left| \mathcal{R}\left(f_n(s) - f(s)\right)(t) \right|^2_{L^2_{m,H}} dt,$$

which gives, by Proposition 5.20,

$$\int_0^T \left| D_Q G_1^n(t) - G_2(t) \right|^2_{L^2_{m,H}} dt$$

$$\leq T \int_0^T \left[|f_n(s) - f(s)|^2_{L^2_m} - |P_T\left[f_n(s) - f(s)\right]|^2_{L^2_m} \right] ds$$

$$\leq T \int_0^T |f_n(s) - f(s)|^2_{L^2_m} ds, \tag{5.42}$$

so that $D_Q G_1^n$ is convergent in $L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m; H)\right)$ to G_2 and (5.41) holds. □

The following corollary can be deduced from the proof of Lemma 5.23.

Corollary 5.24 *Assume that Hypotheses 5.1 and 5.4 hold. Let $f_n \to f$ be in $L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m)\right)$. Then, by (5.42), there exists a subsequence f_{n_k} such that for a.e. $(s, t) \in [0, T] \times [0, T]$ and $s \leq t$,*

$$D_Q P_{t-s}[f_{n_k}(s)] \to \mathcal{R}\left(f(s)\right)(t-s) \quad \text{in} \quad L^2(H, \overline{\mathcal{B}}, m; H).$$

This fact will be useful in Sect. 5.5.

We now extend the operator D_Q to all functions u that are mild solutions to suitable Cauchy problems.

Consider $g \in L^2\left(H, \overline{\mathcal{B}}, m\right)$ and $f \in L^2\left(0, T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right)$. Consider the Cauchy problem:

$$\begin{cases} u_t(t) + \mathcal{A}u(t) + f(t) = 0 \quad t \in [0, T), \\ u(T, x) = g(x) \end{cases} \tag{5.43}$$

and define the mild solution of (5.43) as

$$u(t) = P_{T-t}[g] + \int_t^T P_{s-t}[f(s)]ds, \quad t \in [0, T]. \tag{5.44}$$

We denote by $\Upsilon_{\mathcal{A}}(0, T)$ the set of all the functions in $L^2\left(0, T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right)$ that can be written in the form (5.44) for some f, g as above. The functions in $\Upsilon_{\mathcal{A}}(0, T)$ belong to $C\left([0, T], L^2\left(H, \overline{\mathcal{B}}, m\right)\right)$.

For the functions in $\Upsilon_{\mathcal{A}}(0, T)$ we define the operator $\tilde{D}_Q$ by

$$(\tilde{D}_Q u)(t) := \mathcal{R}(g)(T - t) + \int_t^T \mathcal{R}(f(s))(s - t)\, ds, \quad t \in [0, T]. \tag{5.45}$$

Observe that $\tilde{D}_Q$ is well defined on $\Upsilon_{\mathcal{A}}(0, T)$. Indeed, if we have $P_{T-t}[g_1] + \int_t^T P_{s-t}[f_1(s)]ds = P_{T-t}[g_2] + \int_t^T P_{s-t}[f_2(s)]ds$ then, taking $t = T$ we obtain $g_1 = g_2$ and then, $\int_t^T P_{s-t}[f_1(s)]ds = \int_t^T P_{s-t}[f_2(s)]ds$ so that $\int_t^T \mathcal{R}(f_1(s))(s - t)\, ds = \int_t^T \mathcal{R}(f_2(s))(s - t)\, ds$.

The following proposition gives a continuity result for $\tilde{D}_Q$.

Proposition 5.25 *Suppose that Hypotheses 5.1 and 5.4 hold. Consider two sequences* $g_n \subset L^2(H, m)$ *and* $f_n \subset L^2\left(0, T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right)$ *such that*

$$g_n \longrightarrow g \qquad in\ \ L^2\left(H, \overline{\mathcal{B}}, m\right),$$

$$f_n \longrightarrow f \qquad in\ \ L^2\left(0, T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right).$$

Then, setting

$$u_n(t) = P_{T-t}[g_n] + \int_t^T P_{s-t}[f_n(s)]ds, \quad t \in [0, T], \tag{5.46}$$

and

$$\tilde{D}_Q u_n(t) = \mathcal{R}(g_n)(T - t) + \int_t^T \mathcal{R}(f_n(s))(s - t)\, ds, \quad t \in [0, T],$$

we have

$$u_n \longrightarrow u \qquad in\ \ C\left([0, T], L^2\left(H, \overline{\mathcal{B}}, m\right)\right), \tag{5.47}$$

$$\tilde{D}_Q u_n \longrightarrow \tilde{D}_Q u \qquad in\ \ L^2\left(0, T; L^2\left(H, \overline{\mathcal{B}}, m; H\right)\right). \tag{5.48}$$

Proof We start with the first claim. Subtracting (5.44) from (5.46) we get

$$u_n(t) - u(t) = P_{T-t}[g_n - g] + \int_t^T P_{s-t}[f_n(s) - f(s)]\, ds$$

so that, by the strong continuity of P_t,

$$|u_n(t) - u(t)|^2_{L^2_m} \le C_T \left[|g_n - g|^2_{L^2_m} + \int_t^T |f_n(s) - f(s)|^2_{L^2_m} ds \right],$$

which gives (5.47) by taking the supremum over $[0, T]$. To prove (5.48) we observe that we have

$$\tilde{D}_Q(u_n(t) - u(t)) = \mathcal{R}(g_n - g)(T - t) + \int_t^T \mathcal{R}(f_n(s) - f(s))(s - t)\, ds$$

so that, by (5.37) and (5.41),

$$\int_0^T \left| \tilde{D}_Q u_n(t) - \tilde{D}_Q u(t) \right|^2_{L^2_{m,H}} \le |g_n - g|^2_{L^2_m} + T \int_0^T |f_n(s) - f(s)|^2_{L^2_m} ds,$$

which shows (5.48). □

Remark 5.26 If g and f are differentiable functions, the operator D_Q is well defined on the functions u of the form (5.44). In (5.45) we define the operator $\tilde{D}_Q$ on all the functions of the form (5.44), where $g \in L^2\left(H, \overline{\mathcal{B}}, m\right)$ and $f \in L^2\left(0, T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right)$. Thus Proposition 5.25 asserts that the operator $\tilde{D}_Q$ extends D_Q on $\Upsilon_{\mathcal{A}}(0, T)$ "without closability problems" if the functions in the approximating sequence have the form (5.46). ■

5.3 The HJB Equation

In this section we study the existence and uniqueness of solutions to the HJB equation[5]

$$\begin{cases} u_t + \mathcal{A}u + F_0\left(t, x, D_Q u\right) + l(t, x) = 0, \\ u(T, x) = g(x) \end{cases} \tag{5.49}$$

with $g \in L^2(H, \overline{\mathcal{B}}, m)$. Observe that this corresponds to F in (5.1) having the form $F(t, x, p) = F_0\left(t, x, Q^{1/2} p\right)$. We assume that the following conditions are satisfied.

Hypothesis 5.27 (A) $F_0 : [0, T] \times H \times H \to \mathbb{R}$ is $Leb \otimes \overline{\mathcal{B}} \otimes \overline{\mathcal{B}}/\mathcal{B}(\mathbb{R})$-measurable (where Leb is the σ-field of Lebesgue measurable sets in $\mathbb{R}$) and there exists an $L \in \mathbb{R}$ such that

[5]Following the notation we use for HJB equations throughout the book, in the first line of (5.49) we only explicitly mention the dependence on t and x of the functions F_0 and l while we do not do so for u_t, $D_Q u$ and $\mathcal{A}u$.

$$|F_0(t,x,p) - F_0(t,x,q)| \le L|p-q| \quad \text{and} \quad |F_0(t,x,p)| \le L(1+|p|) \tag{5.50}$$

for all $t \in [0,T]$ and $x, p, q \in H$.

(B) $l \in L^2\left(0,T; L^2(H,\overline{\mathcal{B}},m)\right)$ and $g \in L^2(H,\overline{\mathcal{B}},m)$.

Using the semigroup P_t defined in (5.10) and the variation of constants formula, as was done in Chap. 4, we can formally rewrite Eq. (5.49) in the following mild form:

$$u(t) = P_{T-t}[g] + \int_t^T P_{s-t}\left[F_0\left(s,\cdot,D_Q u(s)\right)\right]ds + \int_t^T P_{s-t}[l(s)]\,ds, \quad 0 \le t \le T, \tag{5.51}$$

where for simplicity we have written $D_Q u(s), l(s)$ for $D_Q u(s,\cdot), l(s,\cdot)$ and a similar convention is used later for other functions. We use this integral form to define a solution.

We will prove the existence of the solution of the HJB equation using a fixed point argument in the space $L^2\left(0,T; W_Q^{1,2}(H,m)\right)$. We can identify any element of $L^2\left(0,T; W_Q^{1,2}(H,m)\right)$ with an element (v,V) in $L^2\left(0,T; L^2(H,\overline{\mathcal{B}},m)\right) \times L^2\left(0,T; L^2(H,\overline{\mathcal{B}},m;H)\right)$. If $v(t) \in C_b^1(H)$ for almost every t, then $V(t) = D_Q v(t)$ for almost every t and the norm of $(v,V) = (v, D_Q v)$ in $L^2\left(0,T; W_Q^{1,2}(H,m)\right)$ can be written explicitly as follows

$$|(v, D_Q v)|^2_{L^2\left(0,T;W_Q^{1,2}\right)} = \int_0^T \left(|v(t)|^2_{L^2_m} + \left|D_Q v(t)\right|^2_{L^2_{m,H}}\right)dt.$$

To avoid any confusion in the notation we will always denote the elements of $L^2\left(0,T; W_Q^{1,2}(H,m)\right)$ as pairs.

Definition 5.28 By a *solution* of Eq. (5.51) (or *mild solution* of Eq. (5.49)), we mean a pair of functions

$$(u,U) \in L^2\left(0,T; W_Q^{1,2}(H,m)\right) \subset L^2\left(0,T; L^2(H,\overline{\mathcal{B}},m)\right) \times L^2\left(0,T; L^2(H,\overline{\mathcal{B}},m;H)\right)$$

such that, for a.e. $t \in [0,T]$ and m-a.e.

$$u(t) = P_{T-t}[g] + \int_t^T P_{s-t}\left[F_0\left(s,\cdot,U(s)\right)\right]ds + \int_t^T P_{s-t}[l(s)]\,ds, \tag{5.52}$$

and

$$U(t) = \mathcal{R}(g)(T-t) + \int_t^T \mathcal{R}\left(F_0(s,\cdot,U(s))\right)(s-t)ds + \int_t^T \mathcal{R}\left(l(s)\right)(s-t)ds. \tag{5.53}$$

Remark 5.29 By the definition of $\tilde{D}_Q u$ in (5.45) and Definition 5.28 we immediately get $U = \tilde{D}_Q u$. ■

Remark 5.30 If D_Q were closable, then it would be natural to define the solution of Eq. (5.51) as an element of $L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ such that (5.51) is satisfied for a.e. $t \in [0, T]$ and m-a.e. But D_Q may not be closable, so elements of $W_Q^{1,2}(H, m)$ are not functions in general, but pairs of functions belonging to the product space $L^2(H, \overline{\mathcal{B}}, m) \times L^2(H, \overline{\mathcal{B}}, m; H)$.

Note that the second equation (5.53) is an obvious consequence of (5.52) if the operator D_Q is closable and, in this case, $U = D_Q u$. ■

We will introduce a suitable nonlinear operator $\overline{\mathcal{M}}$ which will allow us to use the fixed point argument. It will be defined in terms of a certain operator $\mathcal{M}_1$ and its derivative. Both of these operators will be initially defined on a subspace of $L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m)\right)$ and then extended to $L^2\left(0, T; W_Q^{1,2}(H, m)\right)$. To make the distinction we will denote the extensions using the "overline": $\overline{\mathcal{M}_1}$ and $\overline{D_Q \mathcal{M}_1}$. As emphasized before, since the elements of $L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ can be identified with a subspace of $L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m)\right) \times L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m; H)\right)$, we will use a one-argument notation for the non-extended operators (e.g. $\mathcal{M}_1(u)$) and a two-argument notation for the extended ones (e.g. $\overline{\mathcal{M}_1}(u, U)$).

Given g, l and F_0 satisfying Hypothesis 5.27, we define the operator $\mathcal{M}_1$ as follows:

$$
\begin{cases}
D(\mathcal{M}_1) = \Big\{ v \in L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m)\right) \\
\qquad : \ v(t) \in C_b^1(H) \quad \text{for a.e. } t \text{ and } \quad |(v, D_Q v)|_{L^2\left(0,T; W_Q^{1,2}\right)} < \infty \Big\}, \\
\mathcal{M}_1 v(t) = P_{T-t}[g] + \int_t^T P_{s-t}\left[F_0\left(s, \cdot, D_Q v(s)\right)\right] ds + \int_t^T P_{s-t}[l(s)]\, ds, \quad t \le T.
\end{cases}
$$

Remark 5.31 If g, l and F_0 are regular enough, then we can directly define $D_Q \mathcal{M}_1$.

If $g \in L^2(H, \overline{\mathcal{B}}, m)$, $l \in L^2(0, T; L^2(H, \overline{\mathcal{B}}, m))$ and $F_0\left(s, x, D_Q v(s)\right) \in L^2(0, T; L^2(H, \overline{\mathcal{B}}, m))$ we can use Lemma 5.23 to define $\tilde{D}_Q \mathcal{M}_1 v \in L^2\left(0, T; L^2\left(H, \overline{\mathcal{B}}, m; H\right)\right)$ and it can be written as follows:

$$
\tilde{D}_Q \mathcal{M}_1 v(t) = \mathcal{R}(g)(T-t) + \int_t^T \mathcal{R}\left(F_0(s, \cdot, D_Q v(s))\right)(s-t) ds + \int_t^T \mathcal{R}\left(l(s)\right)(s-t) ds
$$

on $[0, T]$. In the following lemma we extend by continuity the operator $D_Q \mathcal{M}_1$ to $L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ obtaining $\overline{D_Q \mathcal{M}_1}$. Since the definitions of $\tilde{D}_Q \mathcal{M}_1 v$ and $\overline{D_Q \mathcal{M}_1}$ coincide on $D(\mathcal{M}_1)$, they coincide once $\overline{D_Q \mathcal{M}_1}$ is extended to $L^2\left(0, T; W_Q^{1,2}(H, m)\right)$. ■

Lemma 5.32 *Assume that Hypotheses 5.1, 5.4 and 5.27 hold. Then* $\mathcal{M}_1$ *extends to a Lipschitz mapping*

$$\overline{\mathcal{M}}_1 : L^2\left(0, T; W_Q^{1,2}(H, m)\right) \to L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m)\right)$$

with Lipschitz constant LT. *The mapping* $D_Q\mathcal{M}_1 : D(\mathcal{M}_1) \to L^2\left(0, T; L^2\left(H, \overline{\mathcal{B}}, m; H\right)\right)$ *extends to a Lipschitz mapping*

$$\overline{D_Q\mathcal{M}_1} : L^2\left(0, T; W_Q^{1,2}(H, m)\right) \to L^2\left(0, T; L^2\left(H, \overline{\mathcal{B}}, m; H\right)\right)$$

with Lipschitz constant $LT^{1/2}$.

Proof Since $|F_0(t, x, p)| \leq L(1 + |p|)$ for all $t \in [0, T]$ and $x, p \in H$, it follows from Lemma 5.23 that $\mathcal{M}_1 v \in L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m)\right)$ and $D_Q\mathcal{M}_1 v \in L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m)\right)$ for every $v \in D(\mathcal{M}_1)$.

Given v_1 and v_2 in $D(\mathcal{M}_1)$, we have

$$\mathcal{M}_1(v_1 - v_2)(t) = \int_t^T P_{s-t}\left[F_0\left(s, \cdot, D_Q v_1(s)\right) - F_0\left(s, \cdot, D_Q v_2(s)\right)\right] ds, \quad t \in [0, T],$$

and therefore, since $\|P_t\| \leq 1$ and by Hypothesis 5.27-(A),

$$|\mathcal{M}_1(v_1 - v_2)(t)|_{L^2_m} \leq L \int_t^T \left|D_Q v_1(s) - D_Q v_2(s)\right|_{L^2_{m,H}} ds, \quad t \in [0, T].$$

Hence,

$$\int_0^T |\mathcal{M}_1(v_1 - v_2)(t)|^2_{L^2_m} dt \leq L^2 T^2 \int_0^T \left|D_Q v_1(t) - D_Q v_2(t)\right|^2_{L^2_{m,H}} dt.$$

It follows that $\mathcal{M}_1$ may be extended to the whole space $L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ by continuity and the resulting mapping is Lipschitz continuous with constant LT. Similarly, for v_1 and v_2 in $D(\mathcal{M}_1)$ and $t \in [0, T]$,

$$D_Q\mathcal{M}_1(v_1 - v_2)(t) = \int_t^T D_Q P_{s-t}\left[F_0\left(s, \cdot, D_Q v_1(s)\right) - F_0\left(s, \cdot, D_Q v_2(s)\right)\right] ds.$$

Using the notation introduced in Lemma 5.23 we obtain

$$\begin{aligned}\int_0^T &\left|D_Q\mathcal{M}_1(v_1 - v_2)(t)\right|^2_{L^2_{m,H}} dt \\ &= \int_0^T \left|G_2\left(F_0\left(t, \cdot, D_Q v_1(t)\right) - F_0\left(t, \cdot, D_Q v_2(t)\right)\right)\right|^2_{L^2_{m,H}} dt\end{aligned}$$

$$\le T \int_0^T \left| F_0\left(t, \cdot, D_Q v_1(t)\right) - F_0\left(t, \cdot, D_Q v_2(t)\right)\right|^2_{L^2_m} dt$$
$$\le L^2 T \int_0^T \left| D_Q\left(v_1(t) - v_2(t)\right)\right|^2_{L^2_{m,H}} dt,$$

and therefore $D_Q\mathcal{M}_1$ extends to a Lipschitz continuous mapping on $L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ with constant $LT^{1/2}$. □

Remark 5.33 The operators $\overline{\mathcal{M}}_1$ and $\overline{D_Q\mathcal{M}}_1$ depend only on the second component of the elements of $L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ but it is convenient for us to define them on $L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ to apply the fixed point argument below. ■

Taking into account the extensions of the operators $\mathcal{M}_1$ and $D_Q\mathcal{M}_1$ provided by Lemma 5.32 we can define the operator

$$\begin{cases} \overline{\mathcal{M}} : L^2\left(0, T; W_Q^{1,2}(H, m)\right) \to L^2\left(0, T; W_Q^{1,2}(H, m)\right) \\ \overline{\mathcal{M}}(u, U) = (\overline{\mathcal{M}}_1(u, U), \overline{D_Q\mathcal{M}}_1(u, U)). \end{cases}$$

Remark 5.34 Using Proposition 5.20 and Lemma 5.23 we find that for a.e. $t \in [0, T]$,

$$\overline{\mathcal{M}}_1(u, U)(t) = P_{T-t}[g] + \int_t^T P_{s-t}\left[F_0(s, \cdot, U(s))\right] ds + \int_t^T P_{s-t}[l(s)] ds \tag{5.54}$$

and

$$\begin{aligned} &\overline{D_Q\mathcal{M}}_1(u, U)(t) \\ &= \mathcal{R}(g)(T-t) + \int_t^T \mathcal{R}\left(F_0(s, \cdot, U(s))\right)(s-t) ds + \int_t^T \mathcal{R}\left(l(s)\right)(s-t) ds. \end{aligned} \tag{5.55}$$

■

Theorem 5.35 *Assume that Hypotheses 5.1, 5.4 and 5.27 hold. Then for every $g \in L^2(H, \overline{\mathcal{B}}, m)$ there exists a unique mild solution (u, U) to Eq. (5.49) in the sense of Definition 5.28. Moreover, $u \in C\left([0, T], L^2(H, \overline{\mathcal{B}}, m)\right)$ and $U = \tilde{D}_Q u$.*

Proof We apply the Banach Fixed Point Theorem to the mapping $\overline{\mathcal{M}}$ in the space $L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ endowed with the norm $|\cdot|_{L^2\left(0,T;W_Q^{1,2}\right)}$ when T is sufficiently small. By Lemma 5.32, for any $(v_1, V_1), (v_2, V_2) \in L^2\left(0, T; W_Q^{1,2}(H, m)\right)$,

$$\int_0^T \left|\overline{\mathcal{M}}_1(v_1(t), V_1(t)) - \overline{\mathcal{M}}_1(v_2(t), V_2(t))\right|^2_{L^2_m} dt$$
$$\leq L^2T^2|(v_1, V_1) - (v_2, V_2)|^2_{L^2\left(0,T;W_Q^{1,2}\right)} \quad (5.56)$$

and

$$\int_0^T \left|\overline{D_Q\mathcal{M}}_1(v_1(t), V_1(t)) - \overline{D_Q\mathcal{M}}_1(v_2(t), V_2(t))\right|^2_{L^2_{m,H}} dt$$
$$\leq L^2T|(v_1, V_1) - (v_2, V_2)|^2_{L^2\left(0,T;W_Q^{1,2}\right)}. \quad (5.57)$$

From (5.56) and (5.57) we have

$$|\overline{\mathcal{M}}(v_1, V_1) - \overline{\mathcal{M}}(v_2, V_2)|_{L^2\left(0,T;W_Q^{1,2}\right)}$$
$$\leq L\sqrt{T(T+1)}|(v_1, V_1) - (v_2, V_2)|_{L^2\left(0,T;W_Q^{1,2}\right)}, \quad (5.58)$$

thus $\overline{\mathcal{M}}$ is a strict contraction for T sufficiently small. Thus we obtain a unique solution on a small time interval. The rest follows by standard iteration. Finally, denoting the solution by (u, U), since $F_0(s, \cdot, U(s)) \in L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m)\right)$ and P_t is a C_0-semigroup, we find that $u \in C\left([0, T], L^2(H, \overline{\mathcal{B}}, m)\right)$ thanks to (5.54).

The last statement is an immediate consequence of the definitions (see Remark 5.29). □

Remark 5.36 Observe that the uniqueness of the solution stated in Theorem 5.35 has to be understood with respect to the reference measure m whose support can also be very thin. This is one of the drawbacks of the method. For results about existence of non-degenerate invariant measures, see Sect. 5.6 and the comments in the bibliographical notes. ■

5.4 Approximation of Mild Solutions

We now show, following the approach of Chap. 4, that the mild solution of the HJB equation can be obtained as a limit of classical solutions. Thus we need to introduce the concept of a classical solution.

We introduce the operator $\mathcal{A}_1$ which is defined similarly to the operator $\mathcal{A}_1$ in (4.141):

$$\begin{cases} D(\mathcal{A}_1) = \left\{\phi \in UC_b^2(H)\ :\ A^*D\phi \in UC_b(H,H) \text{ and } D^2\phi \in UC_b(H,\mathcal{L}_1(H))\right\} \\ \mathcal{A}_1\phi = \frac{1}{2}\operatorname{Tr}[QD^2\phi] + \langle x, A^*D\phi\rangle + \langle b(x), D\phi\rangle\,. \end{cases} \tag{5.59}$$

It is easy to see that $D(\mathcal{A}_1)$ endowed with the norm

$$\|\phi\|_{D(\mathcal{A}_1)} := \|\phi\|_0 + \|D\phi\|_0 + \|A^*D\phi\|_0 + \sup_{x\in H}\|D^2\phi(x)\|_{\mathcal{L}_1(H)} \tag{5.60}$$

is a Banach space.

In Sect. 5.2.3 we introduced the operator $\mathcal{A}$ as the generator of the C_0-semigroup P_t on $L^2(H, \overline{\mathcal{B}}, m)$ (see Proposition 5.9). In the following lemma we study its relations with the operator $\mathcal{A}_1$.

Lemma 5.37 *Let Hypotheses 5.1 and 5.4 hold. Then:*

(i) $\mathcal{F}C_0^{2,A^*}(H) \subset D(\mathcal{A}_1)$.

(ii) $D(\mathcal{A}_1)$ *is embedded in* $D(\mathcal{A})$. *Moreover, for any* $\phi \in D(\mathcal{A}_1)$,

$$\mathcal{A}\phi(x) = \frac{1}{2}\operatorname{Tr}\left[QD^2\phi(x)\right] + \left\langle x, A^*D\phi(x)\right\rangle + \langle b(x), D\phi(x)\rangle\,. \tag{5.61}$$

(iii) If we consider the Banach space structure on $D(\mathcal{A}_1)$ *described above and the graph norm on* $D(\mathcal{A})$, *the embedding* $D(\mathcal{A}_1) \subset D(\mathcal{A})$ *is continuous.*

Proof *Part (i)* follows straightforwardly from the definitions of $\mathcal{F}C_0^{2,A^*}(H)$ and $D(\mathcal{A}_1)$.

Part (ii): We choose $\phi \in D(\mathcal{A}_1)$ and we start by showing that, for any $x \in H$,

$$\lim_{t\to 0}\frac{P_t[\phi](x) - \phi(x)}{t} = \frac{1}{2}\operatorname{Tr}\left[QD^2\phi(x)\right] + \left\langle x, A^*D\phi(x)\right\rangle + \langle b(x), D\phi(x)\rangle\,. \tag{5.62}$$

Indeed, applying Dynkin's formula (Proposition 1.169), we have

$$\begin{aligned} \frac{P_t[\phi](x) - \phi(x)}{t} &= \frac{\mathbb{E}\phi(X(t;x)) - \phi(x)}{t} \\ &= \frac{1}{t}\mathbb{E}\int_0^t \left[\frac{1}{2}\operatorname{Tr}\left[QD^2\phi(X(s;x))\right] + \left\langle X(s;x), A^*D\phi(X(s;x))\right\rangle\right. \\ &\qquad\qquad \left. + \langle b(X(s;x)), D\phi(X(s;x))\rangle\right] ds. \end{aligned} \tag{5.63}$$

We need to show that every term in the right-hand side of (5.63) converges to the corresponding one in (5.62). Let us look at the middle term. We define

$$I_t^1(x) := \frac{1}{t}\int_0^t \left[\left\langle X(s;x), A^*D\phi(X(s;x))\right\rangle - \left\langle x, A^*D\phi(x)\right\rangle\right] ds.$$

Let σ be a modulus of continuity of $A^* D\phi$ which we can assume to be concave. We have

$$\begin{aligned}\left|I_t^1(x)\right| \leq \frac{1}{t}\int_0^t \mathbb{E}\Big[\left|\left\langle X(s;x)-x, A^*D\phi(X(s;x))\right\rangle\right| \\ + \left|\left\langle x, A^*D\phi(X(s;x)) - A^*D\phi(x)\right\rangle\right| \Big] ds \\ \leq \frac{1}{t}\int_0^t \left(\|A^*D\phi\|_0 \mathbb{E}\,|X(s;x)-x| + |x|\mathbb{E}\sigma\left(|X(s;x)-x|\right)\right) ds \\ \leq \frac{1}{t}\int_0^t \left(\|A^*D\phi\|_0 \mathbb{E}\,|X(s;x)-x| + |x|\sigma\left(\mathbb{E}\,|X(s;x)-x|\right)\right) ds,\end{aligned}$$

where we used Jensen's inequality to obtain the last inequality. The last line above converges to 0 as $t \to 0$ by (5.9). The convergence of other terms in (5.62) is proved similarly.

We now need to show that the convergence takes place in $L^2(H, \overline{\mathcal{B}}, m)$. We see that, thanks to (1.69) and since $\|A^*D\phi\|_0$ is finite, we have $\sup_{t\in(0,1]}\left(\mathbb{E}[I_t^1(x)]\right)^2 \leq g(x) = C_1 + C_2|x|^2$ for some positive constants C_1, C_2. Since $g \in L^1(H, \overline{\mathcal{B}}, m)$ by (5.12), we can thus use the dominated convergence theorem to conclude that $\lim_{t\to 0}\mathbb{E}[I_t^1(\cdot)] = 0$ in $L^2(H, \overline{\mathcal{B}}, m)$. We argue similarly to get the convergence of the other terms. Therefore $\phi \in D(\mathcal{A})$ and (5.61) holds.

Part (iii): Given $\phi \in D(\mathcal{A}_1)$ we have

$$\begin{aligned}|\phi|^2_{D(\mathcal{A})} = |\phi|^2_{L^2_m} + |\mathcal{A}\phi|^2_{L^2_m} \\ \leq |\phi|^2_{L^2_m} + 3\left|\frac{1}{2}\operatorname{Tr}\left[QD^2\phi\right]\right|^2_{L^2_m} + 3\left|\left\langle\cdot, A^*D\phi(\cdot)\right\rangle\right|^2_{L^2_m} + 3\left|\langle b(\cdot), D\phi(\cdot)\rangle\right|^2_{L^2_m} \\ \leq \|\phi\|_0^2 + \frac{3}{4}\|Q\|^2_{\mathcal{L}(H)}\sup_{x\in H}\|D^2\phi(x)\|^2_{\mathcal{L}_1(H)} \\ +3\|A^*D\phi\|_0^2\int_H |x|^2\, dm(x) + 3\|D\phi\|_0^2\int_H \left(|b(0)| + K|x|\right)^2\, dm(x).\end{aligned}$$

Thanks to (5.12) there exists a constant C, depending only on m, b and Q such that the last expression is smaller than $C\|\phi\|^2_{D(\mathcal{A}_1)}$. This concludes the proof. □

The concept of a classical solution of (5.49) is also similar to the one introduced in Definition 4.129, however here we limit our interest to functions belonging to $\Upsilon_{\mathcal{A}}(0, T)$ to be able to define $\tilde{D}_Q$.

Definition 5.38 A function $u \in \Upsilon_{\mathcal{A}}(0, T)$ is a classical solution of (5.49) if u has the following regularity properties

$$\begin{cases} u(\cdot, x) \in C^1([0, T]), \quad \forall x \in H \text{ and } u_t \in C_b([0, T] \times H), \\ u(t, \cdot) \in D(\mathcal{A}_1), \ \forall t \in [0, T] \quad \text{and} \quad \sup_{t \in [0,T]} \|u(t, \cdot)\|_{D(\mathcal{A}_1)} < +\infty, \\ u, \mathcal{A}_1 u \in C_b([0, T] \times H), \\ Du, A^* Du, \tilde{D}_Q u \in C_b([0, T] \times H, H), \\ D^2 u \in C_b([0, T] \times H, \mathcal{L}_1(H)), \end{cases}$$

(where $\tilde{D}_Q$ is defined in (5.45)) and satisfies

$$\begin{cases} u_t + \mathcal{A}_1 u + F_0\left(t, x, \tilde{D}_Q u\right) + l(t, x) = 0, \quad t \in [0, T), \ for\, m - a.e.\, x \in H, \\ u(T, x) = g(x), \quad for\, m - a.e.\, x \in H. \end{cases} \tag{5.64}$$

Definition 5.39 A function $u \in \Upsilon_{\mathcal{A}}(0, T)$ is a strong solution of Eq. (5.49) if $(u, \tilde{D}_Q u) \in L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ and there exist sequences $(u_n), (l_n) \subset L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ and $g_n \subset \mathcal{F}C_0^{2,A^*}(H)$ such that for every $n \in \mathbb{N}$, u_n is the classical solution of the Cauchy problem

$$\begin{cases} w_t + \mathcal{A} w + F_0(t, x, D_Q w) + l_n(t, x) = 0, \\ \\ w(T, x) = g_n(x), \end{cases} \tag{5.65}$$

and the following limits hold as $n \to +\infty$:

$$\begin{array}{ll} g_n \longrightarrow g & \text{in } L^2\left(H, \overline{\mathcal{B}}, m\right) \\ l_n \longrightarrow l & \text{in } L^2\left(0, T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right) \\ u_n \longrightarrow u & \text{in } C\left([0, T], L^2\left(H, \overline{\mathcal{B}}, m\right)\right) \\ \tilde{D}_Q u_n \longrightarrow \tilde{D}_Q u & \text{in } L^2\left(0, T; L^2\left(H, \overline{\mathcal{B}}, m; H\right)\right). \end{array}$$

In principle we can have several strong solutions of Eq. (5.49), depending on the choice of the approximating sequences. Nevertheless we will see that in our case, if a strong solution exists, it is unique. See the discussion that follows the proof of Theorem 5.41 for more on this.

Theorem 5.40 *Assume that Hypotheses 5.1, 5.4 and 5.27 hold. If u is a strong solution of (5.49) then the pair $(u, U) := (u, \tilde{D}_Q u)$ is a mild solution of Eq. (5.49).*

Proof Let u_n, l_n, g_n be its approximating sequences as in Definition 5.39. Recalling that P_t is a strongly continuous semigroup on $L^2(H, \overline{\mathcal{B}}, m)$ (see Proposition 5.9), using Lemma 5.37 and the properties of classical solutions demanded in Definition 5.38, we can compute, for a fixed $t \in [0, T]$, the derivative in the variable s of $P_{s-t}[u_n(s)]$ (as a mapping from $[t, T]$ to $L^2(H, \overline{\mathcal{B}}, m)$). We get

$$\frac{d}{ds}P_{s-t}[u_n(s)] = P_{s-t}[\mathcal{A}u_n(s)] + P_{s-t}\left[\frac{d}{ds}u_n(s)\right]$$
$$= P_{s-t}[\mathcal{A}u_n(s)] + P_{s-t}\left[\left(-\mathcal{A}_1u_n(s) - F_0\left(s,\cdot,\tilde{D}_Qu_n(s)\right) - l_n(s)\right)\right]$$
$$= P_{s-t}\left[\left(-F_0\left(s,\cdot,\tilde{D}_Qu_n(s)\right) - l_n(s)\right)\right], \quad s \in [t, T].$$

Integrating both sides of this expression over $[t, T]$, using that $u_n(T) = g_n$ and reordering the terms we obtain for every n

$$u_n(t) = P_{T-t}[g_n] + \int_t^T P_{s-t}\left[F_0(s,\cdot,\tilde{D}_Qu_n(s)) + l_n(s)\right]ds.$$

Setting $\psi_n(s) = F_0(s,\cdot,\tilde{D}_Qu_n(s)) + l_n(s)$, the last expression becomes

$$u_n(t) = P_{T-t}[g_n] + \int_t^T P_{s-t}[\psi_n(s)]\,ds,$$

where $g_n \in \mathcal{F}C_0^{2,A^*}(H)$, $\psi_n \in L^2\left(0,T;L^2\left(H,\overline{\mathcal{B}},m\right)\right)$,

$$g_n \xrightarrow{n\to+\infty} g \quad \text{in } L^2(H,\overline{\mathcal{B}},m),$$

and, thanks to Hypothesis 5.27-(A),

$$\psi_n \xrightarrow{n\to+\infty} F_0(\cdot,\cdot,\tilde{D}_Qu) + l \quad \text{in } L^2\left(0,T;L^2\left(H,\overline{\mathcal{B}},m\right)\right).$$

We can now apply Proposition 5.25 and pass to the limit as $n \to +\infty$ to get the claim. □

Theorem 5.41 *Assume that Hypotheses 5.1, 5.4 and 5.27 hold and suppose $b = 0$. If the pair $(u, U) \in L^2\left(0,T;W_Q^{1,2}(H,m)\right)$ is a mild solution of Eq. (5.49) then $U = \tilde{D}_Qu$ and u is a strong solution of (5.49).*

Proof In the particular case $b = 0$ the semigroup P_t simplifies to the Ornstein–Uhlenbeck semigroup studied in Sect. B.7.2. The notation used in other parts of the book in this case is R_t but here, to be consistent with the general notation used in the chapter, we continue to denote the semigroup by P_t. Hypotheses 5.1-(A)-(B)-(C) imply Hypothesis B.79, needed in all the results of Sect. B.7 used in this proof. Observe that, if $b = 0$, the operator $\mathcal{A}_1$ defined in (5.59) reduces to the operator $\mathcal{A}_0$ defined in (B.36).

As argued in Remark 5.29 we immediately get $U = \tilde{D}_Qu$. Let g_n, ψ_n be two sequences such that

$$g_n \in \mathcal{F}C_0^{2,A^*}(H), \tag{5.66}$$

$$\psi_n \colon [0,T] \to \mathcal{F}C_0^{2,A^*}(H), \tag{5.67}$$

$$\psi_n \text{ and } \mathcal{A}_1\psi_n \text{ belong to } C([0,T], UC_b(H)), \tag{5.68}$$

$$g_n \overset{n\to+\infty}{\longrightarrow} g \quad \text{in } L^2(H, \overline{\mathcal{B}}, m) \tag{5.69}$$

and

$$\psi_n \overset{n\to+\infty}{\longrightarrow} F_0(\cdot,\cdot,\tilde{D}_Q u) + l \quad \text{in } L^2\left(0,T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right). \tag{5.70}$$

These sequences exist thanks to Lemma 5.7.

Since $(u, U) = (u, \tilde{D}_Q u)$ is a mild solution of (5.49) we have

$$u(t) = P_{T-t}[g] + \int_t^T P_{s-t}\left[F_0(s,\cdot,\tilde{D}_Q u(s)) + l(s)\right] ds.$$

If we set

$$u_n(t,x) = P_{T-t}[g_n] + \int_t^T P_{s-t}[\psi_n(s)]\, ds, \tag{5.71}$$

by Proposition 5.25 we obtain that

$$u_n \overset{n\to+\infty}{\longrightarrow} u \quad \text{in } C\left([0,T], L^2\left(H, \overline{\mathcal{B}}, m\right)\right), \tag{5.72}$$

$$\tilde{D}_Q u_n \overset{n\to+\infty}{\longrightarrow} \tilde{D}_Q u \quad \text{in } L^2\left(0,T; L^2\left(H, \overline{\mathcal{B}}, m; H\right)\right). \tag{5.73}$$

The latter, thanks to Hypotheses 5.27-(A), implies in particular that

$$F_0(\cdot,\cdot,\tilde{D}_Q u_n) \overset{n\to+\infty}{\longrightarrow} F_0(\cdot,\cdot,\tilde{D}_Q u) \quad \text{in } L^2\left(0,T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right).$$

So, thanks to (5.70), if we set

$$l_n = \psi_n - [F_0(\cdot,\cdot,\tilde{D}_Q u_n)], \tag{5.74}$$

we get

$$l_n \overset{n\to+\infty}{\longrightarrow} l \quad \text{in } L^2\left(0,T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right). \tag{5.75}$$

We can now apply Proposition B.91-(ii). Observe that the existence of the function g_0 demanded in the hypotheses of this proposition can be easily found thanks to (5.68) and the constant C in (B.33) and (B.35) is here equal to zero. The time is reversed (t in Proposition B.91 corresponds to our $T-t$ for any $t \in [0,T]$). It thus follows that u_n satisfies in the classical sense the approximating HJB equation

$$\begin{cases} (u_n)_t + \mathcal{A}u_n + F_0(t, x, D_Q u_n) + l_n(t, x) = 0 \\ u(T, x) = g_n(x). \end{cases} \tag{5.76}$$

Given the regularity of u_n, g_n and ψ_n, $\tilde{D}_Q u_n = D_Q u_n$ and then the fact that $\tilde{D}_Q u_n \in C_b([0, T] \times H, H)$, not directly stated in Proposition B.91, follows from $Du_n \in C_b([0, T] \times H, H)$ and the continuity of Q.

This, together with the convergences (5.69), (5.72), (5.73) and (5.75), shows that u is a strong solution in the sense of Definition 5.39. □

Theorem 5.35 shows that, under Hypotheses 5.1, 5.4 and 5.27, there exists a unique mild solution (u, U) of Eq. (5.49). Theorem 5.40 ensures that, under the same hypotheses, any strong solution is also a mild solution so, in particular there exists at most one strong solution of (5.49) and, whenever it exists, it can be identified with the mild solution. Theorem 5.41 proves, under the additional assumption $b = 0$, the reverse implication, ensuring in particular the existence of a (unique) strong solution in this case. This result was stated in [298] (see in particular Proposition 4.3) without the assumption $b = 0$ but the proof of the regularity of the u_n in the general case was not complete.

In Sect. 5.5, we work again under Hypotheses 5.1, 5.4 and 5.27 but we also suppose that a strong solution exists or, equivalently, that the mild solution of the equation is also strong. This is always the case if $b = 0$.

5.5 Application to Stochastic Optimal Control

We apply the results on abstract HJB equations from previous sections to study a family of optimal control problems.

5.5.1 The State Equation

We work, as usual, in a real separable Hilbert space H which will be both the state space and the noise space (see Sect. 1.2.4), that is we have $\Xi = H$. The control space Λ is a closed ball in a real separable Banach space E:

$$\Lambda = \overline{B_\varrho(0)}. \tag{5.77}$$

The linear operators A, Q and the function b satisfy Hypothesis 5.1. As in Chap. 2, the notation $\mu := \left(\Omega^\mu, \mathscr{F}^\mu, \left\{\mathscr{F}^t_{\mu,s}\right\}_{s\in[t,T]}, \mathbb{P}^\mu, W^\mu_Q\right)$ (or without the index μ if the context is clear) will be used to denote a generalized reference probability space (see Definition 1.100). We limit our attention here to the case where the σ-fields of the filtration $\mathscr{F}^t_{\mu,s}$ are countably generated up to sets of measure zero. This holds, for example, for filtrations generated by Wiener processes, see Lemma 1.94. We recall that the generalized reference probability spaces μ used in Sect. 5.5 may be different from μ_0 in Hypothesis 5.1.

We consider a stochastic controlled system governed by the state equation

$$\begin{cases} dX(s) = \left(AX(s) + b(X(s)) + Q^{\frac{1}{2}}R(s, X(s), a(s))\right) ds + dW_Q(s), \\ X(t) = x, \quad x \in H, \end{cases} \tag{5.78}$$

where R and a satisfy the following hypothesis.

Hypothesis 5.42 We assume that:

(i) $R\colon [0,T] \times H \times \Lambda \to H$ is Borel measurable and there exists an $M_R > 0$ such that

$$\sup_{(s,x,a)\in[0,T]\times H\times\Lambda} |R(s,x,a)| \le M_R < +\infty,$$

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

$$|R(s,x,a) - R(s,y,a)| \le M_R|x-y|.$$

(ii) For every $t \in [0,T]$ and a generalized reference probability space μ on $[t,T]$, the σ-fields of the filtration $\{\mathscr{F}^t_{\mu,s}\}_{s\in[t,T]}$ are countably generated up to sets of measure zero. Λ is as in (5.77) and the control processes $a(\cdot) : [t,T] \times \Omega \to \Lambda$ belong to the set

$$\mathcal{U}^\mu_t := \left\{a(\cdot)\colon [t,T] \times \Omega \to \Lambda \ : \ a(\cdot) \text{ is } \mathscr{F}^t_s - \text{progressively measurable}\right\}. \tag{5.79}$$

We recall that the control processes in $\mathcal{U}^\mu_t$ depend on the choice of the generalized reference probability space (Definition 1.100) μ because they are progressively measurable with respect to the filtration $\{\mathscr{F}^s_t\}_{s\in[t,T]}$ that depends on the choice of μ. See Sect. 2.1.1 for more on this.

Remark 5.43 The boundedness of R is imposed to be able to solve later, in Theorem 5.55, the closed loop equation using Girsanov's theorem. A similar approach is also used in Sect. 6.5. ∎

Proposition 5.44 *Let Hypotheses 5.1 and 5.42 be satisfied. Then, for any* $t \in [0,T]$, $x \in H$, $a(\cdot) \in \mathcal{U}^\mu_t$, *the state equation (5.78) has a unique solution* $X(\cdot; t, x, a(\cdot)) \in \mathcal{H}^\mu_p(t,T;H)$ *(see Definition 1.126) for all* $p \ge 1$. *In particular,* $X(\cdot; t, x, a(\cdot)) \in M^p_\mu(t,T;H)$ *(defined in (1.29)) for all* $p \ge 1$.

Moreover, there exists a $\mathcal{B}([t,T]) \otimes \mathcal{B}(H) \otimes \mathscr{F}/\mathcal{B}(H)$*-measurable function*

$$\begin{cases} [t,T] \times H \times \Omega \to H \\ (s,x,\omega) \to \tilde{X}(s;t,x,a(\cdot))(\omega) \end{cases} \tag{5.80}$$

such that, for every $x \in H$, $\tilde{X}(\cdot; t, x, a(\cdot))$ *is a version of the solution* $X(\cdot; t, x, a(\cdot))$. *Thus in the future we will not make a distinction between* $X(\cdot; t, x, a(\cdot))$ *and* $\tilde{X}(\cdot; t, x, a(\cdot))$.

Proof The result, except for the last claim, follows from Proposition 1.147. The whole term $\left[Q^{1/2}b(X(s)) + Q^{1/2}R(s, X(s), a(s))\right]$ corresponds to the term b_0 in Hypothesis 1.145, $a(\cdot)$ plays the role of $a_1(\cdot)$ and we have no $a_2(\cdot)$.

To prove the last claim, we consider a countable dense subset $\mathcal{S} := \{x_n\}_{n\in\mathbb{N}}$ of H. Thanks to (1.70) we can find $\Omega_2 \subset \Omega$ with $\mathbb{P}(\Omega_2) = 1$ such that (1.70) holds with $\xi_1 = x_1$ and $\xi_2 = x_2$ for any $s \in [t, T]$ and $\omega \in \Omega_2$. Similarly, for every $N > 2$ we can find a subset $\Omega_N \subset \Omega$ with $\mathbb{P}(\Omega_N) = 1$ such that (1.70) is satisfied for any choice $\xi_1 = x_i, \xi_2 = x_j, i, j = 1, \dots, N$, for all $s \in [t, T]$ and $\omega \in \Omega_N$. If we define $\Omega_\infty = \bigcap_{n\geq 1} \Omega_n$ we have again $\mathbb{P}(\Omega_\infty) = 1$. Given $s \in [t, T]$ and $\omega \in \Omega_\infty$, we define, for any $x \in H$,

$$\tilde{X}(s; t, x, a(\cdot))(\omega) := \lim_{n\to\infty} X(s; t, y_n, a(\cdot))(\omega), \tag{5.81}$$

where y_n is a sequence of elements of $\mathcal{S}$ such that $y_n \to x$ (the limit exists and it does not depend on the chosen sequence y_n, again thanks to (1.70) and the choice of Ω_∞). We define $\tilde{X}(s; t, x, a(\cdot))(\omega) = 0$ for $(s, x, \omega) \in [t, T] \times H \times (\Omega \setminus \Omega_\infty)$. The pointwise convergence (5.81) and the progressive measurability (and thus the $\mathcal{B}([t, T]) \otimes \mathscr{F}/\mathcal{B}(H)$-measurability) of X ensures that (see Lemma 1.8(iii)), for any $x \in H$, the restriction of $\tilde{X}(\cdot; t, x, a(\cdot))(\cdot)$ to $[t, T] \times \Omega_\infty$ is $\mathcal{B}([t, T]) \otimes (\mathscr{F} \cap \Omega_\infty)/\mathcal{B}(H)$-measurable. This fact, the completeness of $\mathscr{F}$ and the fact that $\tilde{X}$ is constant on $[t, T] \times H \times (\Omega \setminus \Omega_\infty)$ give easily the $\mathcal{B}([t, T]) \otimes \mathscr{F}/\mathcal{B}(H)$-measurability of $\tilde{X}(\cdot; t, x, a(\cdot))(\cdot)$ on $[t, T] \times \Omega$. Moreover, by construction, for any $s \in [t, T]$ and $\omega \in \Omega$, $x \to \tilde{X}(s; t, x, a(\cdot))(\omega)$ is continuous so that (see Lemma 1.18) the function defined in (5.80) is $\mathcal{B}([t, T]) \otimes \mathcal{B}(H) \otimes \mathscr{F}/\mathcal{B}(H)$-measurable. □

5.5.2 The Optimal Control Problem and the HJB Equation

Let Hypotheses 5.1, 5.4 and 5.42 be satisfied. We study an optimal control problem in its *strong* formulation (see Sect. 2.1.1 for details) so that the generalized reference probability space μ is fixed. We consider the following cost functional

$$J^\mu(t, x; a(\cdot)) = \mathbb{E}\left\{\int_t^T l(s, X(s; t, x, a(\cdot))) + h_2(a(s))ds + g(X(T; t, x, a(\cdot)))\right\} \tag{5.82}$$

which we want to minimize over the control set $\mathcal{U}_t^\mu$. In this expression $X(s; t, x, a(\cdot))$ represents the mild solution of (5.78) at time s which, as always, we will often denote by $X(s)$. The functions l, h_2 and g satisfy the following hypothesis.

Hypothesis 5.45 $l : [0, T] \times H \to \mathbb{R}$ and $g : H \to \mathbb{R}$ satisfy Hypothesis 5.27-(B) while $h_2 : \Lambda \to \mathbb{R}$ is Borel measurable and bounded.

The value function of the problem depends on μ and it is defined as in (2.4):

$$V_t^\mu(x) = \inf_{a(\cdot)\in\mathcal{U}_t^\mu} J^\mu(t,x;a(\cdot)). \tag{5.83}$$

The HJB equation corresponding to the described optimal control problem is

$$\begin{cases} v_t + \mathcal{A}v + F_0(t,x,D_Q v) + l(t,x) = 0 \\ v(T,x) = g(x), \end{cases} \tag{5.84}$$

where the operator $\mathcal{A}$ is defined in Sect. 5.2.3 and the Hamiltonian F_0 is given by

$$F_0(t,x,p) = \inf_{a\in\Lambda} \{\langle R(t,x,a), p\rangle + h_2(a)\} =: \inf_{a\in\Lambda} F_{0,CV}(t,x,p,a). \tag{5.85}$$

We will suppose that F_0 satisfies Hypothesis 5.27-(A). Indeed, thanks to Hypotheses 5.42 and 5.45, the Lipschitz continuity and growth conditions (5.50) are always satisfied but the $Leb \otimes \overline{\mathcal{B}} \otimes \overline{\mathcal{B}}/\mathcal{B}(\mathbb{R})$ measurability may not always be ensured. However, when R does not depend on t and x, the Hamiltonian F_0 is just a function from H to $\mathbb{R}$ and Lemma 1.21 then guarantees that it is $\overline{\mathcal{B}}/\mathcal{B}(\mathbb{R})$-measurable, so that Hypothesis 5.27-(A) is satisfied. Hypothesis 5.27-(A) is also always true if $R(t,x,\cdot)$ is continuous for every t and x due to the separability of Λ.

5.5.3 The Verification Theorem

We now show how to obtain a verification theorem and an explicit expression for optimal controls in feedback form.

Lemma 5.46 *Let $t\in[0,T]$, $x\in H$, $\mu = \left(\Omega, \mathscr{F}, \left\{\mathscr{F}_s^t\right\}_{s\in[t,T]}, \mathbb{P}, W_Q\right)$ be a generalized reference probability space on $[t,T]$ and let $a(\cdot)\in\mathcal{U}_t^\mu$. Assume that Hypotheses 5.1, 5.4, 5.42 and 5.45 hold. Define*

$$\rho_{a(\cdot)} = \exp\left(-\int_t^T \left\langle R(r, X(r;t,x,a(\cdot)), a(r)), dW_Q(r)\right\rangle \right.$$
$$\left. -\frac{1}{2}\int_t^T |R(r, X(r;t,x,a(\cdot)), a(r))|^2\, dr\right).$$

Then:

(i) The measure $\tilde{\mathbb{P}}$ on $(\Omega,\mathscr{F})$ defined by setting $d\tilde{\mathbb{P}}(A) := \rho_{a(\cdot)}(T)d\mathbb{P}(\omega)$, that is, for any $A\in\mathscr{F}$,

$$\tilde{\mathbb{P}}(A) := \int_A \rho_{a(\cdot)}(\omega)\, d\mathbb{P}(\omega),$$

is a probability measure on Ω, in particular $\mathbb{E}[\rho_{a(\cdot)}] = 1$.

(ii) There exists a positive constant $\tilde{c} < +\infty$ such that

$$\mathbb{E}\left[\left(\rho_{a(\cdot)}\right)^{-1}\right] \le \tilde{c}, \qquad \text{for any } x \in H \tag{5.86}$$

and we have

$$d\mathbb{P}(A) := \left(\rho_{a(\cdot)}\right)^{-1} d\tilde{\mathbb{P}}(\omega). \tag{5.87}$$

(iii) Denote by $X(\cdot; t, x)$ the solution of

$$\begin{cases} dX(s) = (AX(s) + b(X(s)))\, ds + dW_Q(s), \quad s \in [0, T], \\ X(t) = x \in H. \end{cases} \tag{5.88}$$

For any $s \in [t, T]$, $\mathcal{L}_{\mathbb{P}}(X(s; t, x)) = \mathcal{L}_{\tilde{\mathbb{P}}}(X(s; t, x, a(\cdot)))$.

(iv) For any nonnegative $w \in L^2(H, \mathcal{B}(H), m)$, for any $s \in [t, T]$,

$$\int_H \mathbb{E} w(X(s; t, x, a(\cdot)))dm(x) \le \sqrt{\tilde{c}} \left(\int_H \mathbb{E} w^2(X(s; t, x))dm(x)\right)^{1/2} = \sqrt{\tilde{c}}|w|_{L^2_m}, \tag{5.89}$$

where $\tilde{c}$ is the constant introduced in (5.86).

Proof Most of the statements of the lemma are corollaries of the Girsanov Theorem.

Part (i): Given the boundedness of R the claim follows from Proposition 10.17 and Theorem 10.14 of [180].

Part (ii): Observe first that if we replace $R(s, X(s; t, x, a(\cdot)), a(s))$ by $-R(s, X(s; t, x, a(\cdot)), a(s))$ we have again a bounded function so that the results of Part (i) hold: we get

$$\mathbb{E} \exp\left(\int_t^T \left\langle R(s, X(s; t, x, a(\cdot)), a(s)),\, Q^{-1/2} dW_Q(s)\right\rangle - \frac{1}{2}\int_t^T |R(s, X(s; t, x, a(\cdot)), a(s))|^2\, ds\right) = 1. \tag{5.90}$$

Since by Hypothesis 5.42 there exists an $M_R \in \mathbb{R}$ such that $|R(s, X(s; t, x, a(\cdot)), a(s))| \le M_R$ for any choice of $s \in [t, T]$, $x \in H$ and any $a(\cdot)$,

$$\mathbb{E}\left[\left(\rho_{a(\cdot)}\right)^{-1}\right] = \mathbb{E} \exp\left(\int_t^T \left\langle R(r, X(r; t, x, a(\cdot)), a(r)), dW_Q(r)\right\rangle + \frac{1}{2}\int_t^T |R(r, X(r; t, x, a(\cdot)), a(r))|^2\, dr\right) \le$$
$$e^{(T-t)M_R^2} \mathbb{E} \exp\left(\int_t^T \left\langle R(r, X(r; t, x, a(\cdot)), a(r)), dW_Q(r)\right\rangle\right.$$

$$-\frac{1}{2}\int_t^T |R(r, X(r; t, x, a(\cdot)), a(r))|^2 \, dr\Bigg) = e^{(T-t)M^2} =: \tilde{c},$$

where in the last step we used (5.90).

The second claim follows by the strict positivity of $\rho_{a(\cdot)}$ as a corollary of the Radon–Nikodym Theorem (see [18], p. 64).

Part (iii): Thanks to Theorem 10.14 of [180] we know that the process defined by

$$\tilde{W}_Q(s) = W_Q(s) - W_Q(t) + \int_t^s Q^{\frac{1}{2}} R(r, X(r; t, x, a(\cdot)), a(r))dr, \qquad s \in [t, T],$$

is a Q-Wiener process in H with respect to $\{\mathscr{F}_s^t\}_{s\geq t}$ and the probability measure $\tilde{\mathbb{P}}$. We have

$$\begin{aligned} X(s; t, x, a(\cdot)) &= e^{(s-t)A}x + \int_t^s e^{(s-r)A} b(X(r; t, x, a(\cdot)))dr \\ &\quad + \int_t^s e^{(s-r)A} Q^{\frac{1}{2}} R(r, X(r; t, x, a(\cdot)), a(r))dr + \int_t^s e^{(s-r)A} dW_Q(r) \\ &= e^{(s-t)A}x + \int_t^s e^{(s-r)A} b(X(r; t, x, a(\cdot)))dr + \int_t^s e^{(s-r)A} Q^{\frac{1}{2}} R(r, X(r; t, x, a(\cdot)), a(r))dr \\ &\quad + \int_t^s e^{(s-r)A} d\tilde{W}_Q(r) - \int_t^s e^{(s-r)A} Q^{\frac{1}{2}} R(r, X(r; t, x, a(\cdot)), a(r))dr \\ &= e^{(s-t)A}x + \int_t^s e^{(s-r)A} b(X(r; t, x, a(\cdot)))dr + \int_t^s e^{(s-r)A} d\tilde{W}_Q(r), \quad s \in [t, T], \end{aligned}$$

so $X(\cdot; t, x, a(\cdot))$ solves the same equation as $X(\cdot; t, x)$. The claim thus follows thanks to Proposition 1.148-(ii).

Part (iv): For any $s \in [t, T]$, the joint measurability of the function $(x, \omega) \to w(X(s; t, x, a(\cdot))(\omega))$ follows by the Borel measurability of w and by the measurability of X stated in Proposition 5.44.

Using first (5.87) and then the Cauchy–Schwarz inequality we have, for $s \in [t, T]$,

$$\begin{aligned} \int_H \mathbb{E} w(X(s; t, x, a(\cdot)))dm(x) &= \int_H \int_\Omega w(X(s; t, x, a(\cdot))(\omega))d\mathbb{P}(\omega)dm(x) \\ &= \int_\Omega w(X(s; t, x, a(\cdot))(\omega)) \left(\rho_{a(\cdot)}(\omega)\right)^{-1} d\tilde{\mathbb{P}}(\omega)dm(x) \\ &\leq \left(\int_H \int_\Omega \left(\rho_{a(\cdot)}(\omega)\right)^{-2} d\tilde{\mathbb{P}}(\omega)dm(x)\right)^{1/2} \left(\int_H \int_\Omega w^2(X(s; t, x, a(\cdot))(\omega))d\tilde{\mathbb{P}}(\omega)dm(x)\right)^{1/2} \\ &= \left(\int_H \int_\Omega \left(\rho_{a(\cdot)}(\omega)\right)^{-1} d\mathbb{P}(\omega)dm(x)\right)^{1/2} \left(\int_H \int_\Omega w^2(X(s; t, x)(\omega))d\mathbb{P}(\omega)dm(x)\right)^{1/2}, \end{aligned}$$

where in the last step we used, in the two terms, respectively Part (i) and Part (iii). Therefore, by (5.86) and then using the definition of the transition semigroup, the fact that it does not depend on a generalized reference probability space, and the property of the invariant measure (observe that w^2 belongs to $L^1(H, \mathcal{B}(H), m)$ so we refer to Proposition 5.9 for $p = 1$), we obtain

$$
\begin{aligned}
&\int_H \mathbb{E}w(X(s;t,x,a(\cdot)))dm(x) \le \sqrt{\tilde{c}}\left(\int_H \mathbb{E}w^2(X(s;t,x))dm(x)\right)^{1/2} \\
&= \sqrt{\tilde{c}}\left(\int_H P_{s-t}\left[|w(\cdot)|^2\right](x)dm(x)\right)^{1/2} = \sqrt{\tilde{c}}\left(\int_H |w(x)|^2 dm(x)\right)^{1/2} = \sqrt{\tilde{c}}|w|_{L^2_m},
\end{aligned}
$$

which gives the claim. □

The result of Part (iv) of Lemma 5.46 will be extended to a general $w \in L^2(H, \overline{\mathcal{B}}, m)$ in Corollary 5.48.

Lemma 5.47 *Let $t \in [0,T]$, $\mu = \left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s\in[t,T]}, \mathbb{P}, W_Q\right)$ be a generalized reference probability space on $[t,T]$ and $a(\cdot) \in \mathcal{U}_t^\mu$. Assume that Hypotheses 5.1, 5.4, 5.42 and 5.45 hold.*

Consider a $\overline{\mathcal{B}}/\mathcal{B}(\mathbb{R})$-measurable function $\phi\colon H \to \mathbb{R}$ (respectively, a $\overline{\mathcal{B}}/\mathcal{B}(H)$-measurable function $\phi\colon H \to H$) and $s \in [t,T]$. Then the function

$$
\begin{cases} H \times \Omega \to \mathbb{R} \\ (x,\omega) \to \phi(X(s;t,x,a(\cdot))(\omega) \end{cases}
$$

is $\overline{\mathcal{B}(H)\otimes\mathscr{F}}/\mathcal{B}(\mathbb{R})$-measurable (respectively, $\overline{\mathcal{B}(H)\otimes\mathscr{F}}/\mathcal{B}(H)$-measurable), where $\overline{\mathcal{B}(H)\otimes\mathscr{F}}$ is the completion of the σ-field $\mathcal{B}(H)\otimes\mathscr{F}$ w.r.t. the measure $m\otimes\mathbb{P}$.

Similarly, given a $\mathcal{B}([t,T])\otimes\overline{\mathcal{B}}/\mathcal{B}(\mathbb{R})$-measurable function $\phi\colon [t,T]\times H \to \mathbb{R}$, the function

$$
\begin{cases} [t,T]\times H \times \Omega \to \mathbb{R} \\ (s,x,\omega) \to \phi(s, X(s;t,x,a(\cdot))(\omega) \end{cases}
$$

is $\overline{\mathcal{B}([t,T])\otimes\mathcal{B}(H)\otimes\mathscr{F}}/\mathcal{B}(\mathbb{R})$-measurable, where $\overline{\mathcal{B}([t,T])\otimes\mathcal{B}(H)\otimes\mathscr{F}}$ is the completion of the σ-field $\mathcal{B}([t,T])\otimes\mathcal{B}(H)\otimes\mathscr{F}$ w.r.t. the measure $ds\otimes m\otimes\mathbb{P}$.

Proof The proof follows the same arguments as those used in the proof of Lemma 5.10. We give it for completeness.

If $\phi\colon H \to \mathbb{R}$ is Borel-measurable the statement follows from the measurability of the solutions of (5.78) stated in Proposition 5.44. If $\phi\colon H \to \mathbb{R}$ is $\overline{\mathcal{B}}/\mathcal{B}(\mathbb{R})$-measurable, let $\tilde{\phi}\colon H \to \mathbb{R}$ be a $\mathcal{B}/\mathcal{B}(\mathbb{R})$-measurable function and $V \in \mathcal{B}(H)$, $m(V) = 0$ be such that $\phi(x) = \tilde{\phi}(x)$ for all $x \in H \setminus V$. Then

$$
\begin{aligned}
0 &\le \int_H \mathbb{P}\{X(s;t,x,a(\cdot))(\omega) \in V\}\, dm(x) \\
&= \int_H \mathbb{E}\left[\mathbf{1}_V(X(s;t,x,a(\cdot)))\right] dm(x) \le \sqrt{\tilde{c}}\left(\int_H \mathbb{E}\left[\mathbf{1}_V^2(X(s;t,x))\right] dm(x)\right)^{1/2} \\
&= \sqrt{\tilde{c}}\left(\int_H |\mathbf{1}_V(x)|^2\, dm(x)\right)^{1/2} = \sqrt{\tilde{c}}\,(m(V))^{1/2} = 0, \quad (5.91)
\end{aligned}
$$

where we used (5.89) and then the property of the invariant measure. This fact shows that the functions $(x,\omega)\to\phi(X(t,x)(\omega)$ and $(x,\omega)\to\tilde{\phi}(X(t,x)(\omega)$ are $m\otimes\mathbb{P}$-e.e. equal on $H\times\Omega$. Thus, since $(x,\omega)\to\tilde{\phi}(X(t,x)(\omega)$ is $\mathscr{F}\otimes\mathcal{B}/\mathcal{B}(\mathbb{R})$-measurable, $(x,\omega)\to\phi(X(t,x)(\omega)$ is $\overline{\mathscr{F}\otimes\mathcal{B}}/\mathcal{B}(\mathbb{R})$-measurable.

The same proof applies if $\phi\colon H\to H$ is a $\overline{\mathcal{B}}/\mathcal{B}(H)$-measurable function.

Similarly, if $\phi\colon [t,T]\times H\to\mathbb{R}$ is $\mathcal{B}([t,T])\otimes\overline{\mathcal{B}}/\mathcal{B}(\mathbb{R})$-measurable we can find (again by Lemma 1.16, recalling that $\mathcal{B}([t,T])\otimes\overline{\mathcal{B}}\subset\overline{\mathcal{B}([t,T])\otimes\mathcal{B}}$)) a $\mathcal{B}([t,T])\otimes\mathcal{B}/\mathcal{B}(\mathbb{R})$-measurable function $\tilde{\phi}\colon [t,T]\times H\to\mathbb{R}$ and $V\in\mathcal{B}([t,T])\otimes\mathcal{B}(H)$ such that $(ds\otimes m)(V)=0$ and $\phi(s,x)=\tilde{\phi}(s,x)$ for all $(s,x)\in[t,T]\times H\setminus V$. If we define $V_s:=\{x\in H\ :\ (s,x)\in V\}$ then $V_s\in\mathcal{B}(H)$ and $m(V_s)=0$ for almost every $s\in[0,T]$. Instead of (5.91) we now have

$$
\begin{aligned}
0 \le \int_t^T\int_H \mathbb{P}\{(s,X(s;t,x,a(\cdot))(\omega))\in V\}\,dm(x)ds
&= \int_t^T\int_H \mathbb{P}\{X(s;t,x,a(\cdot))(\omega)\in V_s\}\,dm(x)ds\\
&= \int_t^T\int_H \mathbb{E}\left[\mathbf{1}_{V_s}(X(s;t,x,a(\cdot)))\right]dm(x)\\
&\le \sqrt{\tilde{c}}\int_t^T\left(\int_H\left|\mathbf{1}_{V_s}(x)\right|^2 dm(x)\right)^{1/2} ds = 0
\end{aligned}
$$

and the proof ends as before. □

Corollary 5.48 *Let $t\in[0,T]$, $x\in H$, $\mu=\left(\Omega,\mathscr{F},\left\{\mathscr{F}_s^t\right\}_{s\in[t,T]},\mathbb{P},W_Q\right)$ be a generalized reference probability space on $[t,T]$ and let $a(\cdot)\in\mathcal{U}_t^\mu$. Assume that Hypotheses 5.1, 5.4, 5.42 and 5.45 hold. Then, for any $w\in L^2(H,\overline{\mathcal{B}},m)$, the map $x\to\mathbb{E}w\left(X(s;t,x,a(\cdot))\right)$ belongs to $L^1\left(H,\overline{\mathcal{B}},m\right)$ and for almost every $s\in[t,T]$,*

$$
\int_H \mathbb{E}w(X(s;t,x,a(\cdot)))dm(x)\le\sqrt{\tilde{c}}\left(\int_H\mathbb{E}w^2(X(s;t,x))dm(x)\right)^{1/2}=\sqrt{\tilde{c}}|w|_{L^2_m},
$$

where $\tilde{c}$ is the constant introduced in (5.86).

Proof The statements about the joint measurability proved in Lemma 5.47 allow us, in particular, to ensure the measurability in s and x of integrals with respect to ω and then to extend Lemma 5.46-(iv) to any $w\in L^2\left(H,\overline{\mathcal{B}},m\right)$. □

Lemma 5.49 *Assume that Hypotheses 5.1, 5.4, 5.42, 5.45 hold and let $a(\cdot)\in\mathcal{U}_t^\mu$. Then, for every $s\in[t,T]$ and $w\in L^2\left(H,\overline{\mathcal{B}},m\right)$ (respectively, $L^2\left(H,\overline{\mathcal{B}},m;H\right)$), the map $x\to\mathbb{E}w\left(X(s;t,x,a(\cdot))\right)$ belongs to $L^1\left(H,\overline{\mathcal{B}},m\right)$ (respectively, $L^1\left(H,\overline{\mathcal{B}},m;H\right)$). Moreover, given a sequence w_n converging to w in $L^2\left(H,\overline{\mathcal{B}},m\right)$ (respectively in $L^2\left(H,\overline{\mathcal{B}},m;H\right)$), the sequence $\mathbb{E}w_n\left(X(s;t,x,a(\cdot))\right)$ converges to $\mathbb{E}w\left(X(s;t,x,a(\cdot))\right)$ in $L^1\left(H,\overline{\mathcal{B}},m\right)$ (respectively, $L^1\left(H,\overline{\mathcal{B}},m;H\right)$).*

Similarly, given $w \in L^2\left(t, T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right)$ *(respectively,* $L^2\left(t, T; L^2\left(H, \overline{\mathcal{B}}, m; H\right)\right)$*), the map* $(s, x) \to \mathbb{E}w\,(s, X(s; t, x, a(\cdot)))$ *belongs to* $L^1\Big((t, T) \times H, \overline{\mathcal{B}([t, T]) \otimes \mathcal{B}}, ds \otimes m\Big)$ *(respectively,* $L^1\Big((t, T) \times H, \overline{\mathcal{B}([t, T]) \otimes \mathcal{B}}, ds \otimes m; H\Big)$*), where* ds *is the Lebesgue measure on* $[t, T]$ *and* $\overline{\mathcal{B}([t, T]) \otimes \mathcal{B}}$ *is the completion of the* σ*-field* $\mathcal{B}([t, T]) \otimes \mathcal{B}$ *w.r.t.* $ds \otimes m$*. Moreover, given a sequence* w_n *converging to* w *in* $L^2\left(t, T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right)$ *(respectively,* $L^2\left(t, T; L^2\left(H, \overline{\mathcal{B}}, m; H\right)\right)$*), the sequence* $\mathbb{E}w_n\,(s, X(s; t, x, a(\cdot)))$ *converges to* $\mathbb{E}w\,(s, X(s; t, x, a(\cdot)))$ *in* $L^1\Big((t, T) \times H, \overline{\mathcal{B}([t, T]) \otimes \mathcal{B}}, ds \otimes m\Big)$ *(respectively,* $L^1\Big((t, T) \times H, \overline{\mathcal{B}([t, T]) \otimes \mathcal{B}}, ds \otimes m; H\Big)$*).*

Proof The statements about joint measurability of the various functions involved follow from Lemma 5.47, Corollary 5.48 or can be proved by similar arguments. Recall, for the case when $w \in L^2\left(t, T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right)$, that there exists (see Theorem 11.47, p. 427 of [8]) a $\tilde{w} \in L^2\left([t, T] \times H, \mathcal{B}([t, T]) \otimes \overline{\mathcal{B}}, ds \otimes m\right)$, uniquely determined up to a $ds \otimes m$-null set, such that, for a.e. $s \in [t, T]$, $\tilde{w}(s, \cdot) = w(s)(\cdot)$ m-a.e.

We only prove the remaining statements related to $w \in L^2\left(t, T; L^2\left(H, \overline{\mathcal{B}}, m\right)\right)$, the others being similar. Invoking Corollary 5.48 and Hölder's inequality, we obtain

$$
\begin{aligned}
&\int_t^T \int_H \mathbb{E}\,|w\,(s, X(s; t, x, a(\cdot)))|\; dm(x)\; ds \\
&\qquad\qquad \le C_T \left(\int_t^T \int_H \mathbb{E}\,|w\,(s, X(s; t, x))|^2\; dm(x)\; ds\right)^{1/2} \\
&\qquad\qquad\qquad = C_T \left(\int_t^T \int_H |w\,(s, \cdot)|^2\,(x)\, dm(x)\; ds\right)^{1/2} < +\infty
\end{aligned}
$$

and the first claim follows. The statements about the convergence follow using the same arguments as indeed we have

$$
\begin{aligned}
&\int_t^T \int_H |\mathbb{E}w_n\,(s, X(s; t, x, a(\cdot))) - \mathbb{E}w\,(s, X(s; t, x, a(\cdot)))|\; dm(x)\; ds \\
&\qquad\qquad \le C_T \left(\int_t^T \int_H |w_n\,(s, \cdot) - w\,(s, \cdot)|^2\,(x)\, dm(x)\; ds\right)^{1/2} \xrightarrow{n\to\infty} 0.
\end{aligned}
$$

Similar estimates give the other claims. □

We are now ready to prove the fundamental identity.

Lemma 5.50 *Let* $t \in [0, T]$*,* $\mu = \left(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s\in[t,T]}, \mathbb{P}, W_Q\right)$ *be a generalized reference probability space on* $[t, T]$ *and let* $a(\cdot) \in \mathcal{U}_t^\mu$*. Assume that Hypotheses 5.1, 5.4, 5.27, 5.42, 5.45 hold. Suppose that the mild solution* $(u, U) \in L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ *of (5.84) is also a strong solution. Then the following identity holds for* m*-a.e.* $x \in H$*:*

$$
\begin{aligned}
u(t, x) + \mathbb{E}\int_t^T F_{0,CV}\left(s, X(s), \tilde{D}_Q u(s, X(s)), a(s)\right) - F_0\left(s, X(s), \tilde{D}_Q u(s, X(s))\right) ds \\
= \mathbb{E}\left\{\int_t^T [l(s, X(s)) + h_2(a(s))]ds + g(X(T))\right\} = J^{\mu}(t, x; a(\cdot)), \qquad (5.92)
\end{aligned}
$$

where $X(\cdot) := X(\cdot; t, x, a(\cdot))$ *denotes the mild solution of (5.78).*

Proof We denote by g_n and ψ_n the approximating sequences of g and $F_0 + l$ characterized in (5.66), (5.67), (5.68), (5.69) and (5.70). We set

$$
u_n(t, x) = P_{T-t}[g_n] + \int_t^T P_{s-t}[\psi_n(s)]ds.
$$

We know that u_n satisfies in the classical sense the approximating HJB equation

$$
\begin{cases}
(u_n)_t + \mathcal{A}u_n + F_0(t, x, \tilde{D}_Q u_n) + l_n(t, x) = 0 \\
u_n(T, x) = g_n(x), \ x \in H,
\end{cases} \qquad (5.93)
$$

where

$$
l_n(t, x) := \psi_n(t, x) - F_0(t, x, \tilde{D}_Q u_n) \overset{n\to+\infty}{\longrightarrow} l \qquad \text{in } L^2\left(0, T; L^2(H, \overline{\mathcal{B}}, m; H)\right).
$$

By Dynkin's formula (see Proposition 1.169) and (5.78) we obtain

$$
\begin{aligned}
&\mathbb{E}u_n(T, X(T)) - u_n(t, x) \\
&= \mathbb{E}\int_t^T \left[(u_n)_s(s, X(s)) + \langle X(s), A^* Du_n(s, X(s))\rangle + \frac{1}{2}\mathrm{Tr}\left[QD^2 u_n(s, X(s))\right]\right] ds \\
&\qquad + \mathbb{E}\int_t^T \left[\left\langle Du_n(s, X(s)), b(X(s)) + Q^{\frac{1}{2}} R(s, X(s), a(s))\right\rangle\right] ds. \qquad (5.94)
\end{aligned}
$$

Then, using (5.93) and the notation $F_{0,CV}$ introduced in (5.85), we get

$$
\begin{aligned}
\mathbb{E}g_n(X(T)) - u_n(t, x) = \mathbb{E}\int_t^T \Big[F_{0,CV}\left(s, X(s), \tilde{D}_Q u_n(s, X(s)), a(s)\right) \\
- F_0(s, X(s), \tilde{D}_Q u_n(s, X(s))) - l_n(s, X(s)) - h_2(a(s))\Big] ds. \\
(5.95)
\end{aligned}
$$

We now pass to the limit as $n \to +\infty$ in (5.95). We use (5.69), (5.70) and the convergences of the sequences u_n and $\tilde{D}_Q u_n$ prescribed by Definition 5.39 (indeed they are proved explicitly in our context in (5.72) and (5.73)). Thanks to Lemma 5.49 it thus follows that, for m-a.e. $x \in H$,

$$\mathbb{E}g(X(T)) - u(t,x) = \mathbb{E}\int_t^T \Big[F_{0,CV}(s, X(s), \tilde{D}_Q u(s, X(s)), a(s)) \\ - F_0(s, X(s), \tilde{D}_Q u(s, X(s))) - l(s, X(s)) - h_2(a(s)) \Big] ds,$$

which gives (5.92) after rearranging the terms. □

Lemma 5.51 *Let $t \in [0,T]$, $\mu = \left(\Omega, \mathscr{F}, \left\{\mathscr{F}_s^t\right\}_{s\in[t,T]}, \mathbb{P}, W_Q\right)$ be a generalized reference probability space on $[t,T]$ satisfying Hypothesis 5.42 and let Λ be as in (5.77). For any $p \geq 1$ there exists a countable subset NU_t^μ of $\mathcal{U}_t^\mu$ dense in $\mathcal{U}_t^\mu$ endowed with the $M_\mu^p(t,T;E)$ norm.*

Proof A possible choice for NU_t^μ is a set of elementary processes (see Definition 1.96). Indeed, in the construction of Lemma 1.98 we can clearly limit the choice of the times t_i appearing in Definition 1.96 to those of a dense and countable subset of $[t,T]$ and the choice of the $\mathscr{F}_{t_i}^t$-random variables to those of a dense and countable subset $\{\xi_j^{t_i}\}_{j\in\mathbb{N}}$ of $L^p(\Omega, \mathscr{F}_{t_i}^t, \mathbb{P}; E)$ (this subset exists thanks to Lemma 1.25). Since we look for processes belonging to $\mathcal{U}_t^\mu$ (and thus having images in $\overline{B_\varrho(0)}$), instead of $\{\xi_j^{t_i}\}_{j\in\mathbb{N}}$ we consider the random variables $\tilde{\xi}_j^{t_i} := \left(\max\left\{\frac{|\xi_j^{t_i}|}{\varrho}, 1\right\}\right)^{-1} \xi_j^{t_i}$. They create a required dense set of $\overline{B_\varrho(0)}$-valued processes. This can be seen by observing that if $x, y \in E$, $|x|_E \leq \varrho$ and $|y|_E > \varrho$, if $\tilde{y} := \left(\max\left\{\frac{|y|}{\varrho}, 1\right\}\right)^{-1} y$, we have

$$|x - \tilde{y}|_E \leq |x - y|_E + |y - \tilde{y}|_E \leq 2|x - y|_E,$$

where the last inequality follows from the fact that $\tilde{y}$ is among the elements of $B_\varrho(0)$ nearest to y, so $|x - y|_E \geq |y - \tilde{y}|_E$. □

In the following lemma we give a sufficient condition to ensure that the functional $J^\mu(t,x;\cdot)$ is continuous with respect to the $M_\mu^p(t,T;E)$ norm.

Lemma 5.52 *Suppose that Hypotheses 5.42-(i) and 5.45 hold and that R, l, g and h_2 satisfy the following additional conditions:*

(i) There exists an $M_R > 0$ such that

$$|R(s,x,a_1) - R(s,y,a_2)| \leq M_R(|x-y| + |a_1 - a_2|)\,\forall s \in [0,T], x, y \in H, a_1, a_2 \in \Lambda.$$

(ii) For some $C, q > 0$,

$$|l(t,x)| \leq C(1 + |x|^q), \qquad \textit{for all } t \in [0,T], x \in H,$$

$$|g(x)| \leq C(1 + |x|^q), \qquad \textit{for all } x \in H.$$

(iii) $h_2 : \Lambda \to \mathbb{R}$ is continuous.

Then for every $t \in [0, T]$, $x \in H$ and every generalized reference probability space μ on $[t, T]$, the functional $J^\mu(t, x; \cdot)$ is continuous with respect to the $M^p_\mu(t, T; E)$ norm, for any $p > q$. In other words, for any sequence of controls $a_n(\cdot)$ in $\mathcal{U}^\mu_t$ converging to $a(\cdot) \in \mathcal{U}^\mu_t$ such that

$$\lim_{n\to\infty} |a_n(\cdot) - a(\cdot)|^p_{M^p_\mu} = \lim_{n\to\infty} \mathbb{E} \int_t^T |a_n(s) - a(s)|^p_E ds = 0, \tag{5.96}$$

we have

$$\lim_{n\to\infty} J(t, x; a_n(\cdot)) = J(t, x; a(\cdot)). \tag{5.97}$$

Moreover,

$$\lim_{n\to\infty} \sup_{s\in[t,T]} \left[\mathbb{E}|X(s; t, x, a_n(\cdot)) - X(s; t, x, a(\cdot))|^p\right] = 0. \tag{5.98}$$

Proof We denote, for $s \in [t, T]$, $X(s; t, x, a_n(\cdot))$ by $X_n(s)$ and $X(s; t, x, a(\cdot))$ by $X(s)$ and also denote by N a positive constant such that $\|e^{tA}\| \le N$ for any $t \in [0, T]$ and $\sup_{x\in H} |\nabla b(x)| \le N$. We have, for any $s \in [t, T]$,

$$\begin{aligned} |X_n(s) - X(s)| &\le \left|\int_t^s e^{(t-r)A}(b(X_n(r)) - b(X(r)))dr\right| \\ &\quad + \left|\int_t^s e^{(t-r)A}(R(r, X_n(r), a_n(r)) - R(r, X(r), a(r)))dr\right| \\ &\le N^2 \int_t^s |X_n(r) - X(r)|dr + N M_R \int_t^s (|a_n(r) - a(r)| + |X_n(r) - X(r)|)dr \end{aligned}$$

and then, for $s \in [t, T]$,

$$\begin{aligned} &\mathbb{E}\left[|X_n(s) - X(s)|^p\right] \\ &\le 3^{p-1}(N^{2p} + N^p M_R^p)T^{1/p} \int_t^s \mathbb{E}|X_n(\tau) - X(\tau)|^p dr + 3^{p-1}T^{1/p}|a_n(\cdot) - a(\cdot)|^p_{M^p_\mu} \end{aligned}$$

and we obtain (5.98) using (5.96) and Gronwall's Lemma (Proposition D.29).

It follows from (5.96) and an easy application of the Lebesgue dominated convergence theorem that $\mathbb{E}\int_t^T h_2(a_n(s))ds$ converges to $\mathbb{E}\int_t^T h_2(a(s))ds$.

So to show (5.97) it remains to prove the convergence of the term $\mathbb{E}\left\{\int_t^T l(s, X_n(s))ds + g(X_n(T))\right\}$. We define the following linear operators:

$$\begin{cases} S_n \colon L^2(H, \overline{\mathcal{B}}, m) \times L^2(t, T; L^2(H, \overline{\mathcal{B}}, m)) \to \mathbb{R} \\ S_n(g, l) := \mathbb{E}\left\{\int_t^T l(s, X_n(s))ds + g(X_n(T))\right\} \end{cases}$$

and

$$\begin{cases} S\colon L^2(H,\overline{\mathcal{B}},m)\times L^2(t,T;L^2(H,\overline{\mathcal{B}},m))\to\mathbb{R} \\ S(g,l):=\mathbb{E}\left\{\int_t^T l(s,X(s))ds+g(X(T))\right\}. \end{cases}$$

Since the constant $\tilde{c}$ appearing in (5.89) only depends on M_R (introduced in Hypothesis 5.42(i)) we know from (5.89) that the family $\{S_n\}$ is equi-continuous. Using the Lebesgue dominated convergence theorem and (5.98) it is easy to see that, for any $(g,l)\in C_b(H)\times C_b([t,T]\times H)$, we have $S_n(g,l)\xrightarrow{n\to\infty}S(g,l)$. Since $C_b(H)\times C_b([t,T]\times H)$ is dense in $L^2(H,\overline{\mathcal{B}},m)\times L^2(t,T;L^2(H,\overline{\mathcal{B}},m))$ and $\{S_n\}$ is equi-continuous, we can conclude that $S(g,l)=\lim_{n\to\infty}S_n(g,l)$ for any $(g,l)\in L^2(H,\overline{\mathcal{B}},m)\times L^2(t,T;L^2(H,\overline{\mathcal{B}},m))$, which completes the proof of (5.97). □

Theorem 5.53 (Verification Theorem, Sufficient Condition) *Let $p\geq 1$ and let Hypotheses 5.1, 5.4, 5.27, 5.42, 5.45 hold. Suppose that the mild solution $(u,U)\in L^2\left(0,T;W_Q^{1,2}(H,m)\right)$ of (5.84) is also a strong solution. Then the following are true:*

(i) For any $t\in[0,T]$ and any generalized reference probability space $\mu=\left(\Omega^\mu,\mathscr{F}^\mu,\left\{\mathscr{F}_t^{\mu,s}\right\}_{s\in[t,T]},\mathbb{P}^\mu\right)$ satisfying Hypothesis 5.42, if $J^\mu(t,x;\cdot)$ is continuous with respect to the $M_\mu^p(t,T;E)$ norm, then there exists a set Z_t^μ with $m(Z_t^\mu)=1$ such that, for all $x\in Z_t^\mu$ and all $a(\cdot)\in\mathcal{U}_t^\mu$ we have

$$u(t,x)\leq V_t^\mu(x)\leq J^\mu(t,x;a(\cdot)). \tag{5.99}$$

(ii) Choose $t\in[0,T]$. Let $\hat{\mu}$ be a generalized reference probability space satisfying Hypothesis 5.42 such that $J^{\hat{\mu}}(t,x;\cdot)$ is continuous with respect to the $M_{\hat{\mu}}^p(t,T;E)$ norm. Let x be in $Z_t^{\hat{\mu}}$. Let $a^*(\cdot)\in\mathcal{U}_t^{\hat{\mu}}$ be such that, denoting by $X^*(\cdot)$ the corresponding state, we have

$$a^*(s)\in\arg\min_{a\in\Lambda}F_{0,CV}(s,X^*(s),\tilde{D}_Qu(s,X^*(s)),a), \tag{5.100}$$

for almost every $s\in[t,T]$ and $\mathbb{P}$-almost surely. Then, the pair $(a^*(\cdot),X^*(\cdot))$ is $\hat{\mu}$-optimal at (t,x) and $u(t,x)=V_t^{\hat{\mu}}(x)=J^{\hat{\mu}}(t,x;a^*(\cdot))$.

Proof Part (i): We fix $t\in[0,T]$. By definition, for every $a\in\Lambda$, $F_{0,CV}(\cdot,a)-F_0(\cdot)\geq 0$ everywhere so for any $a(\cdot)\in\mathcal{U}_t^\mu$, by (5.92), $v(t,x)\leq J^\mu(t,x;a(\cdot))$ for m-a.e. $x\in H$. Thanks to Lemma 5.51 we can then choose a countable subset NU_t^μ dense in $\mathcal{U}_t^\mu$ in the M_μ^p norm containing minimizing sequences for any $x\in H$ (observe that the set of the controls depends on t but it does not depend on the initial datum x). By taking the infimum over $a(\cdot)$ in NU_t^μ in the right-hand side of (5.92) we obtain (i).

Part (ii): Since

$$\mathbb{E}\int_t^T \Big[F_{0,CV}\left(s, X^*(s), \tilde{D}_Q u(s, X^*(s)), a^*(s)\right) \\ - F_0\left(s, X^*(s), \tilde{D}_Q u(s, X^*(s))\right)\Big] ds = 0,$$

by (5.92) we thus get

$$u(t,x) = J^{\hat{\mu}}(t, x; a^*(\cdot)). \tag{5.101}$$

Since (5.99) is satisfied at (t,x) because $x \in Z_t^{\hat{\mu}}$, it follows that $(a^*(\cdot), X^*(\cdot))$ is $\hat{\mu}$-optimal at (t,x) and $u(t,x) = V_t^{\hat{\mu}}(x)$. □

5.5.4 Optimal Feedbacks

Similarly to what we observed in Sect. 2.5.1 for the regular case and in Sect. 4.8 for mild solutions in spaces of continuous functions, we use the fundamental identity and the verification theorem to characterize optimal feedbacks in the L^2 framework.

We consider the hypotheses of Theorem 5.53 and we look at the, possibly multi-valued (and not always defined), function

$$\begin{cases} \Phi\colon (0,T)\times H \to \mathcal{P}(\Lambda) \\ \Phi\colon (s,x) \to \arg\min_{a\in\Lambda} F_{0,CV}(s, x, \tilde{D}_Q u(s,x), a), \end{cases} \tag{5.102}$$

where $(u, U) \in L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ is the mild solution of (5.84). The corresponding Closed Loop Equation is

$$\begin{cases} dX(s) \in \left(AX(s) + b(X(s)) + Q^{\frac{1}{2}} R(s, X(s), \Phi(s, X(s)))\right) ds + dW_Q(s), \\ X(t) = x, \quad x \in H. \end{cases} \tag{5.103}$$

Similarly to Sect. 4.8 we have the following corollary of Theorem 5.53.

Corollary 5.54 *Let $p \geq 1$ and let Hypotheses 5.1, 5.4, 5.27, 5.42, 5.45 hold. Suppose that the mild solution $(u, U) \in L^2\left(0, T; W_Q^{1,2}(H, m)\right)$ of (5.84) is also a strong solution.*

Choose $t \in [0, T]$ and $x \in H$. Assume that, on $[t, T) \times H$, the feedback map Φ defined in (5.103) admits a measurable selection $\phi_t : [t, T) \times H \to \Lambda$. Then:

(i) The Closed Loop Equation

$$\begin{cases} dX(s) = \left(AX(s) + b(X(s)) + Q^{\frac{1}{2}} R(s, X(s), \phi_t(s, X(s)))\right) ds + dW_Q(s), \\ X(t) = x, \end{cases} \tag{5.104}$$

has a weak mild solution (see Definition 1.121) $X_{\phi_t}(\cdot; t, x)$ in a suitable generalized reference probability space $\overline{\mu}$ (and unique in such a space); the elements of the filtration $\mathscr{F}^t_{\bar{\mu},s}$ are countably generated up to sets of measure zero.

(ii) Suppose that the generalized reference probability space $\bar{\mu}$ from part (i) is such that $J^{\bar{\mu}}(t, x; \cdot)$ is continuous with respect to the $M^p_{\bar{\mu}}(t, T; E)$ norm and that x in $Z^{\bar{\mu}}_t$. Define, for $s \in [t, T)$, $a_{\phi_t}(s) = \phi_t(s, X_{\phi_t}(s; t, x))$. Then the pair $(a_{\phi_t}(\cdot), X_{\phi_t}(\cdot; t, x))$ is $\overline{\mu}$-optimal at (t, x) and $u(t, x) = V^{\overline{\mu}}_t(x)$. If, finally, $\Phi(s, x)$ is a singleton for any $(s, x) \in (t, T) \times H$, then $a_{\phi_t}(\cdot)$ is the unique $\bar{\mu}$-optimal control.

Proof Part (i) follows from Theorem 6.36. We can always take the filtration to be the one generated by the Wiener process to ensure that the elements of the filtration are countably generated up to sets of measure zero.

All the statements of part (ii) follow immediately from Theorem 5.53-(ii) except for the uniqueness of optimal controls. If $(\hat{a}(\cdot), \hat{X}(\cdot))$ is another optimal pair at (t, x) with generalized reference probability space $\overline{\mu}$, we immediately have, by Lemma 5.50 and the fact that $u(t, x) = V^{\overline{\mu}}_t(x)$,

$$\mathbb{E}\int_t^T \left[F_{0,CV}\left(s, \hat{X}(s), \tilde{D}_Q u(s, \hat{X}(s)), \hat{a}(s)\right) - F_0\left(s, \hat{X}(s), \tilde{D}_Q u(s, \hat{X}(s))\right)\right] ds = 0.$$

This implies that, for a.e. $s \in [t, T]$ and $\mathbb{P}$-a.s., we have $\hat{a}(s) = \phi_t(s, \hat{X}(s))$. Uniqueness of solutions of (5.104) in $\overline{\mu}$ gives the claim. □

We conclude with a result in a specific case.

Theorem 5.55 *Let $p \geq 1$ and let Hypotheses 5.1, 5.4, 5.27, 5.42, 5.45 hold. Suppose that the mild solution $(u, U) \in L^2\left(0, T; W^{1,2}_Q(H, m)\right)$ of (5.84) is also a strong solution. Suppose also that:*

(i) $E = H$ and $R(t, x, a) \equiv a$, hence $F_{0,CV}$ does not depend on t and x and it is given by

$$F_{0,CV}(p, a) = \langle a, p\rangle + h_2(a).$$

(ii) $h_2 \colon \Lambda \to \mathbb{R}$ is strictly convex and lower semicontinuous.

(iii) $F_0(p) := \inf_{a\in\Lambda}(\langle a, p\rangle + h_2(a))$ is differentiable.

Then, for any $t \in [0, T]$ and $x \in H$, there exists a generalized reference probability space μ (where the elements of the filtration $\mathscr{F}^t_{\mu,s}$ are countably generated up to sets of measure zero) and a control $a^*(\cdot) \in \mathcal{U}^\mu_t$ which satisfies, together with the corresponding trajectory $X^*(\cdot) := X(\cdot; t, x, a^*(\cdot))$, the relation

$$a^*(s) = D_p F_0(\tilde{D}_Q u(s, X^*(s))), \qquad s \in [t, T]. \tag{5.105}$$

If $x \in Z^\mu_t$ and $J^\mu(t, x; \cdot)$ is continuous with respect to the $M^p_\mu(t, T; H)$ norm, then the control $a^*(\cdot)$ is μ-optimal.

Proof We extend the function $h_2 \colon \Lambda \to \mathbb{R}$ to a function $\tilde{h}_2 \colon H \to \mathbb{R} \cup \{+\infty\}$ by defining $\tilde{h}_2(a) = +\infty$ for any $a \notin \overline{B_\varrho(0)}$. One can easily see that $a \to \tilde{h}_2(a)$ is strictly convex and lower semicontinuous on H. Moreover (see e.g. Proposition 2.19, p. 77 of [39]), the function

$$\tilde{h}_2^* \colon H \to \mathbb{R}, \qquad \tilde{h}_2^*(p) := \sup_{a \in H} \Big(\langle a, p \rangle - \tilde{h}_2(a) \Big)$$

is convex and lower semicontinuous on H. Thanks to the way we extended h_2, we necessarily have $\sup_{a \in H} \big(\langle a, -p \rangle - \tilde{h}_2(a) \big) = \sup_{a \in \Lambda} \big(\langle a, -p \rangle - h_2(a) \big)$ and thus $\tilde{h}_2^*(-p) = -F_0(p)$ for any $p \in H$.

Let now $p \in H$. It follows from the lower semi-continuity of $\tilde{h}_2$, its convexity and the fact that its value is $+\infty$ on $H \setminus \overline{B_\varrho(0)}$, that $\arg\min_{a \in H} \Big(\langle a, p \rangle + \tilde{h}_2(a) \Big)$ is non-empty (Theorem 2.11 page 72 of [39]). Since $\tilde{h}_2$ is strictly convex it is single-valued (see p. 84 of [39]). Thanks to the way we extended h_2, this unique point a^* where the minimum is attained belongs to Λ so it is also the unique minimizer of the problem

$$\inf_{a \in \Lambda} \left(\langle a, p \rangle + h_2(a) \right).$$

Moreover (see [39], Proposition 2.33, p. 84), a^* must be in the sub-differential (Definition 2.30, p. 82 of [39]) of $\tilde{h}_2^*(\cdot)$ at $-p$ which is equal to the super-differential of $F_0(\cdot)$ at p. Since by hypothesis F_0 is differentiable, we must have $a^* = D_p F_0(p)$ (see Proposition 2.40, p. 87 of [39]).

We now define the feedback control by

$$a(t) = D_p F_0(\tilde{D}_Q u(t, X(t))). \tag{5.106}$$

Consider, for $s \in [t, T]$, the closed loop equation in the mild form

$$X(s) = e^{(s-t)A} x + \int_t^s e^{(s-r)A} \Big[b(X(r)) + Q^{\frac{1}{2}} D_p F_0(\tilde{D}_Q u(s, X(s))) \Big] dr \\ + \int_t^s e^{(r-s)A} dW_Q(r). \tag{5.107}$$

There exists (Theorem 6.36, where the selection is given by (5.106)) a generalized reference probability space μ where this equation has a mild solution $X^*(\cdot)$. We then take

$$a^*(s) = D_p F_0(\tilde{D}_Q u(s, X^*(s))), \qquad s \in [t, T],$$

and we conclude thanks to Theorem 5.53. $\square$

5.5.5 *Continuity of the Value Function and Non-degeneracy of the Invariant Measure*

The results we have described so far show one of the intrinsic limitations of the L^2 approach. Indeed, they can only describe the behavior of the value function in the support of the invariant measure. Such a support can be, in principle, very small. Also the verification theorem and construction of optimal feedbacks hold only on sets of full measure which may change with the generalized reference probability space. To remedy this we are going to introduce a non-degeneracy hypothesis. The non-degeneracy hypothesis, coupled with some continuity assumptions, will help us refine previous results and prove a number of propositions concerning the weak formulation of the optimal control problem (see Sect. 2.1.2).

Hypothesis 5.56 The invariant measure in Hypothesis 5.4 is non-degenerate. In other words, for any non-void open set $O \subset H$, $m(O) > 0$.

Recall that in the weak formulation of the optimal control problem the generalized reference probability space μ varies with the controls so that the set of admissible controls becomes

$$\overline{\mathcal{U}}_t := \bigcup_{\mu} \mathcal{U}_t^{\mu},$$

where $\mathcal{U}_t^{\mu}$ is the set of admissible controls for a given generalized reference probability space μ defined in (5.79). The value function for the optimal control problem in the weak formulation is then

$$\overline{V}(t, x) = \inf_{a(\cdot) \in \overline{\mathcal{U}}_t} J^{\mu}(t, x; a(\cdot)).$$

Corollary 5.57 *Let the hypotheses of Lemma 5.50 and Hypothesis 5.56 be satisfied. Suppose moreover that, for any choice of t, μ and $a(\cdot)$, the functions $u(t, x)$ and $J^{\mu}(t, x, a(\cdot))$ are continuous in the x variable. Then, for every $(t, x) \in [0, T] \times H$ and any generalized reference probability space μ on $[t, T]$, we have*

$$u(t, x) \le \overline{V}(t, x) \le V_t^{\mu}(x).$$

Proof Lemma 5.50 ensures that, for any choice of t, μ and $a(\cdot)$, $u(t, x) \le J^{\mu}(t, x, a(\cdot))$ for m-almost every $x \in H$. For any $y \in H$ we consider the sequence of balls $B_{1/n}(y)$, where $n \in \mathbb{N}$. Given the non-degeneracy of m, $m\left(B_{1/n}(y)\right) > 0$ and then $u(t, \cdot)$ cannot be strictly bigger than $J^{\mu}(t, \cdot, a(\cdot))$ on $B_{1/n}(y)$. We can thus obtain a sequence y_n converging to y such that $u(t, y_n) \le J^{\mu}(t, y_n, a(\cdot))$. By continuity we get $u(t, y) \le J^{\mu}(t, y, a(\cdot))$. Taking the infimum over $a(\cdot)$ and μ we have the claim. □

More precise results can be obtained under stronger continuity assumptions.

Corollary 5.58 *Let the assumptions of Corollary 5.57 be satisfied. Suppose that, for any choice of t, μ and $a \in \Lambda$, the functions $\tilde{D}_Q u(t, \cdot)$, $R(t, \cdot, a)$, $l(t, \cdot)$ and $F_0(t, \cdot, \cdot)$ are continuous. Suppose that there exist $C > 0$ and $N \in \mathbb{N}$ such that, for all (t, x), $|\tilde{D}_Q u(t, x)|, |l(t, x)| \le C(1 + |x|^N)$. Then the fundamental identity (5.92) holds for any $(t, x) \in [0, T] \times H$, any generalized reference probability space μ and any $a(\cdot) \in \mathcal{U}_t^{\mu}$.*

Proof Lemma 5.50 ensures that, for any choice of t, μ and $a(\cdot)$, we have, for m-a.e. $x \in H$,

$$u(t, x) + \mathbb{E} \int_t^T F_{0,CV} \Big(s, X(s), \tilde{D}_Q u(s, X(s)), a(s)\Big) \\ - F_0 \Big(s, X(s), \tilde{D}_Q u(s, X(s))\Big)\, ds = J^{\mu}(t, x; a(\cdot)), \tag{5.108}$$

where $X(s) := X(s; t, x, a(\cdot))$, for $s \in [t, T]$, is the mild solution of (5.78). Thus, as we did in the proof of Lemma 5.57, thanks to the non-degeneracy of m, for every $x \in H$ we can find a sequence y_n converging to x in H such that

$$u(t, y_n) + \mathbb{E} \int_t^T F_{0,CV} \Big(s, X(s; t, y_n, a(\cdot)), \tilde{D}_Q u(s, X(s; t, y_n, a(\cdot))), a(s)\Big) \\ - F_0 \Big(s, X(s; t, y_n, a(\cdot)), \tilde{D}_Q u(s, X(s; t, y_n, a(\cdot)))\Big)\, ds = J^{\mu}(t, y_n; a(\cdot)). \tag{5.109}$$

We need to show that, taking the limit $n \to \infty$, every term of (5.109) converges to the respective term in (5.108). The convergence of $J^{\mu}(t, y_n; a(\cdot))$ and $u(t, y_n)$ follows from their continuity in the x variable.

The terms inside the integral converge pointwise to the respective terms in (5.108) $\mathbb{P}$-a.s. and for almost any s thanks to (1.70) and the various continuity hypotheses. The convergence of the integral thus follows from Lemma 1.51, the uniform moment bounds from (1.69), the polynomial growth of $|\tilde{D}_Q u(t, \cdot)|$ and $l(t, \cdot)$, the boundedness of R and the bounds on the growth of b and F_0. □

Using this result we find the counterparts of Theorem 5.53, Corollary 5.54 and Theorem 5.55 as follows.

Theorem 5.59 (Verification Theorem, Sufficient Condition) *Let the assumptions of Corollary 5.58 be satisfied. Choose $(t, x) \in [0, T] \times H$ and denote by $\hat{\mu}$ a generalized reference probability space. Let $a^*(\cdot) \in \mathcal{U}_t^{\hat{\mu}}$ be such that, denoting by $X^*(\cdot)$ the corresponding state, we have*

$$a^*(s) \in \arg\min_{a \in \Lambda} F_{0,CV}(s, X^*(s), \tilde{D}_Q u(s, X^*(s)), a) \tag{5.110}$$

for almost every $s \in [t, T]$ and $\mathbb{P}$-almost surely. Then the pair $(a^*(\cdot), X^*(\cdot))$ is optimal at (t, x) for the weak formulation (and so in the $\hat{\mu}$-strong formulation) and $u(t, x) = \overline{V}(t, x) = V_t^{\hat{\mu}}(x) = J^{\hat{\mu}}(t, x; a^*(\cdot))$.

Proof The proof is identical to the proof of Theorem 4.197 if we use Corollary 5.58. □

Corollary 5.60 *Let the assumptions of Corollary 5.58 be satisfied. Choose $(t, x) \in [0, T] \times H$. Assume, moreover, that on $[t, T) \times H$ the feedback map Φ defined in (5.102) admits a measurable selection $\phi_t : [t, T) \times H \to \Lambda$. Then:*

(i) The Closed Loop Equation

$$\begin{cases} dX(s) = \Big(AX(s) + b(X(s)) + Q^{\frac{1}{2}} R(s, X(s), \phi_t(s, X(s)))\Big)\, ds + dW_Q(s), \\ X(t) = x, \end{cases} \tag{5.111}$$

has a weak mild solution (see Definition 1.121) $X_{\phi_t}(\cdot; t, x)$ in a suitable generalized reference probability space $\overline{\mu}$ and it is unique in this space if (5.111) is considered as an equation with the control process $a_{\phi_t}(s) := \phi(s, X_{\phi_t}(s; t, x))$, $s \in [t, T)$.

(ii) The pair $(a_{\phi_t}(\cdot), X_{\phi_t}(\cdot; t, x))$ is optimal for the weak formulation (and a fortiori $\overline{\mu}$-optimal) at (t, x) and $u(t, x) = \overline{V}(t, x) = V_t^{\overline{\mu}}(x) = J^{\bar{\mu}}(t, x; a_{\phi_t}(\cdot))$. If, finally, $\Phi(s, x)$ a singleton for any $(s, x) \in (t, T) \times H$, then a_{ϕ_t} is the unique $\overline{\mu}$-optimal control.

Proof The proof is the same as that of Corollary 5.54 but we have to use Corollary 5.58 instead of Lemma 5.50. □

Observe that, in the above corollary, if the uniqueness of solutions of (5.111) is not guaranteed, the optimality of the pair $(a_{\phi_t}(\cdot), X_{\phi_t}(\cdot; t, x))$ needs to be understood in terms of the extended weak formulation introduced in Remark 2.6.

Theorem 5.61 *Let the assumptions of Corollary 5.58 be satisfied.*
Suppose also that:

(i) $E = H$ and $R(t, x, a) \equiv a$, hence $F_{0,CV}$ does not depend on t and x and it is given by

$$F_{0,CV}(p, a) = \langle a, p\rangle + h_2(a).$$

(ii) $h_2 \colon \Lambda \to \mathbb{R}$ is strictly convex and lower semicontinuous.

(iii) $F_0(p) := \inf_{a\in\Lambda} (\langle a, p\rangle + h_2(a))$ is differentiable.

Then, for any $t \in [0, T]$ and $x \in H$, there exists a generalized reference probability space μ (where the elements of the filtration $\mathscr{F}^t_{\mu,s}$ are countably generated up to sets of measure zero) and a control $a^*(\cdot) \in \mathcal{U}_t^\mu$ which satisfies, together with the corresponding trajectory $X^*(\cdot) := X(\cdot; t, x, a^*(\cdot))$, the relation

$$a^*(s) = D_p F_0(\tilde{D}_Q u(s, X^*(s))), \qquad s \in [t, T].$$

$a^*(\cdot)$ is an optimal control for the weak formulation at (t, x) and the unique $\overline{\mu}$-optimal control at (t, x). For any $t \in [0, T]$ and $x \in H$, $u(t, x)$ equals the value function $\overline{V}(t, x)$.

Proof The proof follows the same arguments as these used in the proof of Theorem 5.55. In the very last step we use Corollary 5.60 instead of Theorem 5.53. □

5.6 Examples

We show how the L^2-theory we have developed so far can be used to treat some specific optimal control problems.

5.6.1 *Optimal Control of Delay Equations*

Let us consider a simple controlled one-dimensional linear stochastic differential equation with a delay $r > 0$:

$$\begin{cases} dy(s) = (\beta_0 y(s) + \beta_1 y(s-r) + \alpha(s))\, ds + \sigma dW_0(s), \\ y(t) = x_0, \\ y(t+\theta) = x_1(\theta),\ \theta \in [-r, 0), \end{cases} \tag{5.112}$$

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 $\left\{\mathscr{F}_s^t\right\}_{s\in[t,T]}$ is the augmented filtration generated by W_0. The control $\alpha(\cdot)$ is an $\mathscr{F}_s^t$-progressively measurable process with values in the interval $\Lambda = [0, R]$ for some $R > 0$. We assume that $x_1(\cdot) \in L^2(-r, 0)$.

As recalled in Sect. 2.6.8, Eq. (5.112) can be rewritten as a linear evolution equation in the Hilbert space $H = \mathbb{R} \times L^2(-r, 0)$ of the following form:

$$\begin{cases} dX(s) = (A_1 X(s) + B_1 a(s))\, dt + G dW_0(s), \\ X(t) = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} := \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} \in H, \end{cases} \tag{5.113}$$

where $a(\cdot) = \alpha(\cdot)$, A_1 is a suitable generator of a C_0-semigroup on H; $B_1 : \mathbb{R} \to H$ and $G : \mathbb{R} \to H$ are continuous operators $B_1 w_0 = \begin{pmatrix} w_0 \\ 0 \end{pmatrix}$ and $G w_0 = \begin{pmatrix} \sigma w_0 \\ 0 \end{pmatrix}$ (further details can be found in Sect. 2.6.8). Finally, considering $Q \in \mathcal{L}^+(H) = \mathcal{L}^+(\mathbb{R} \times L^2(-\tau, 0))$ defined as $Q := \begin{pmatrix} \sigma^2 & 0 \\ 0 & 0 \end{pmatrix}$, we can rewrite the equation once more obtaining

$$\begin{cases} dX(s) = \left(A_1 X(s) + Q^{1/2} \frac{1}{\sigma} B_1 a(s)\right) ds + dW_Q(s), \\ X(t) = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} := \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} \in H, \end{cases} \tag{5.114}$$

which is the form required by (5.78).

Proposition 5.62 *Assume that $\sigma \neq 0$, that $\beta_0 < 1$ and denote by γ a real number in $(0, \pi)$ such that $\gamma \coth \gamma = \beta_0$. Assume that*

$$\beta_0 < -\beta_1 < \sqrt{\gamma^2 + \beta_0^2}. \tag{5.115}$$

Then Eq. (5.113) and (5.114) have a unique invariant measure m which is non-degenerate.

Proof See Remark 10.2.6(i), Chap. 10 of [177]. □

Proposition 5.63 *Consider the operator $D_Q := Q^{1/2}D$ defined on $C_b^1(H) \subset L^2(H, \overline{\mathcal{B}}, m)$. Then:*

(i) D_Q is not closable in $L^2(H, \overline{\mathcal{B}}, m)$.
(ii) Hypothesis 5.1 holds.

Proof Part (i) is proved in [299], Sect. 7.2, pp. 15–16. The second statement can easily be verified. □

Thanks to Part (ii) of Proposition 5.63, the whole theory developed so far in this chapter can be applied even if the operator D_Q is not closable in the classical sense.

Remark 5.64 We considered a simple one-dimensional case of controlled stochastic delay equations for simplicity of presentation. In fact, this framework can be applied to more general cases like semilinear d-dimensional equations presented in Sect. 2.6.8. Conditions to guarantee the existence of a nontrivial invariant measure for the multidimensional case can be found in Sect. 10.3 of [177] (see, in particular, Theorem 10.2.5(i)). Using the same methodology, problems with cost functions f_0 and g_0 depending also on the history of the state y can be treated as well. ■

5.6.2 *Control of Stochastic PDEs of First Order*

The second example is an optimal control problem driven by a first-order stochastic PDE similar to the one considered in Sect. 2.6.7. This kind of equation is important in financial modeling since it provides a description of the time evolution of forward rates under the non-arbitrage assumption; we refer the reader to Sect. 2.6.7 and [303].

Fix $\kappa > 0$. The state space H we consider here is given by the following weighted L^2 space of real-valued functions defined on $[0, +\infty)$:

$$H := \left\{ f : [0, +\infty) \to \mathbb{R} \text{ measurable } : \int_0^{+\infty} f^2(\xi) e^{-\kappa\xi} d\xi < +\infty \right\}.$$

In particular, if $\kappa = 0$, $H = L^2(\mathbb{R})$. The inner product on H is given by

$$\langle f, g\rangle_H := \int_0^{+\infty} f(\xi)g(\xi)e^{-\kappa\xi}d\xi$$

and the induced norm will be denoted by $|\cdot|_H$.

The following result can be easily proved.

Proposition 5.65 *The semigroup $S(t)$ defined as*

$$S(t)f(\xi) := f(t+\xi), \qquad \xi \geq 0$$

is a C_0-semigroup on H. Its generator is given by

$$\begin{cases} D(A) = H^1_\kappa(0,\infty) := \left\{ f \in L^2_\rho \ : \ \frac{df}{d\xi} \in L^2_\rho \right\} \\ A = \frac{d}{d\xi} \end{cases}$$

(where $\frac{df}{d\xi}$ denotes the distribution derivative of f here). Moreover,

$$\|S(t)\|_{\mathcal{L}(H)} \leq e^{-\kappa t}.$$

We consider the following equation, studied for instance in [303],

$$dX(t) = (AX(t) + b(X(t)) + Bh_1(a(t)))\,dt + \tau dW_0(t), \tag{5.116}$$

where W_0 is a one-dimensional Brownian motion; $\tau \in H \cap B_b([0,+\infty),\mathbb{R})$; $B \in \mathcal{L}(H)$ and $h_1 \colon \Lambda \to \mathbb{R}$; $a(t) = a(t,\cdot) \in H$ is a control process and b is an operator defined on H as follows

$$b(x)(\xi) = -\tau(x(\xi))\int_0^\xi \frac{1}{1+e^{x(r)}}\tau(r)dr - \frac{1}{2}|\tau(\xi)|^2\frac{1}{1+e^{x(\xi)}}\tau(\xi)\int_0^\xi \tau(r)dr.$$

In order to apply the L^2 theory we need to ensure the existence of an invariant measure for the uncontrolled version of (5.116). This is the content of the following lemma.

Lemma 5.66 *If*

$$\|\tau\|_0 + |\tau|_H\,|\tau e^{\kappa\cdot}|_H \leq \kappa,$$

then there exists a non-degenerate invariant measure m for

$$dX(t) = (AX(t) + b(X(t)))\,dt + \tau dW_0(t).$$

Proof See Proposition 3.2 in [303]. □

Observe that $\tau dW_0(t)$ is of the form $dW_Q(t)$ prescribed by Hypothesis 5.1-(B) if we consider, for instance, the operator $Qx = \tau\langle\tau, x\rangle$. In this case one can easily see that Hypothesis 5.1-(C) is satisfied as well. To verify Hypothesis 5.42 we need $Bh_1(a(t))$ to be of the form $Q^{1/2}R$ for some R satisfying Hypothesis 5.42-(i). This is the case if we take $B = Q^{\frac{1}{2}}$ and $h_1\colon \Lambda \to H$ some bounded Borel measurable function. Λ needs to be specified, as in (5.77), as a closed ball of a real separable Banach space.

Remark 5.67 The operator $\left(D_Q, C_b^1(H)\right)$ is not always closable in $L^2(H, \overline{\mathcal{B}}, m)$ (see, e.g., Paragraph 7.1, pp. 13–14 of [299]). ■

5.6.3 Second-Order SPDEs in the Whole Space

The third example regards a stochastic controlled parabolic equation in the whole space (see Sects. 2.6.1 and 2.6.2 for stochastic controlled parabolic equations in bounded domains). We consider the problem using a weighted L^2 space as the underlying Hilbert space. For simplicity we limit our observations to the one-dimensional case.

We denote by H the weighted $L^2(\mathbb{R})$ space $L^2\,(\mathbb{R}, \rho_\kappa(\xi)dz)$, where the weight $\rho_\kappa(\xi) = e^{-\kappa|\xi|}$ with $\kappa > 0$.

The inner product and the norm in H are denoted by $\langle \cdot, \cdot \rangle_H$ and $|\cdot|_H$, respectively. Fix $\lambda > 0$ and define $A^{(0)} = \Delta - \lambda I$, where $\Delta\colon D(\Delta) \subset L^2\,(\mathbb{R}) \to L^2\,(\mathbb{R})$ is the Laplacian with domain $D(\Delta)$, which is the Sobolev space $H^2(\mathbb{R})$. Let $S^{(0)}(t)$ denote the C_0-semigroup on $L^2\,(\mathbb{R})$ generated by $A^{(0)}$. The semigroup $S^{(0)}(t)$ is self-adjoint on $L^2\,(\mathbb{R})$ and

$$\left\| S^{(0)}(t) \right\| \le e^{-\lambda t}. \tag{5.117}$$

Proposition 5.68 *$\left\{S^{(0)}(t),\ t \ge 0\right\}$ can be uniquely extended to a C_0-semigroup $\left\{S^{(\kappa)}(t),\ t \ge 0\right\}$ on H. Moreover,*

$$\left\| S^{(\kappa)}(t) \right\|_{\mathcal{L}(H)} \le e^{\left(\frac{1}{2}\kappa^2 - \lambda\right)t}, \quad t \ge 0. \tag{5.118}$$

Proof See Proposition 9.4.1, p. 187 of [177]. □

We denote by $A^{(\kappa)}$ the generator of $\left\{S^{(\kappa)}(t),\ t \ge 0\right\}$.

Consider the controlled equation

$$dX(t) = \left(AX^{(\kappa)}(t) + JR(X(t)) - Ja(t)\right) dt + JdW(t), \tag{5.119}$$

where W is a standard cylindrical Wiener process on $L^2\,(\mathbb{R})$; J is the embedding $L^2\,(\mathbb{R}) \hookrightarrow H$ and $a(\cdot)$ is a control process taking values in $L^2\,(\mathbb{R})$. Assume that the Lipschitz continuous map $R : L^2\,(\mathbb{R}) \to L^2\,(\mathbb{R})$ extends to a map $H \to H$ which satisfies Hypothesis 5.42-(i).

The following equation is the uncontrolled counterpart of (5.119)

$$dX(t) = A^{(\kappa)}X(t)dt + JdW(t). \tag{5.120}$$

Proposition 5.69 *For any $\kappa > 0$ and $\lambda > 0$ the solution of (5.120) is well defined in H and it admits a non-degenerate invariant measure m.*

Proof For the existence of the invariant measure, see Proposition 9.4.6, page 191 of [177]. In [119], Sect. 4.3, it is proved that the invariant measure can be chosen to be non-degenerate. □

It can be shown that the transition semigroup for this process is not strongly Feller, hence it violates the smoothing property required, for example, in Hypothesis 4.76. Thus the theory of the HJB equations developed in Chap. 4 does not apply in this case. Nevertheless, we can study the problem using the results of this chapter.

Remark 5.70 We observe that the family of optimal controls described by the state equation (5.78) needs to satisfy the *structural condition* described in Chap. 2: the image of the drift is always contained in the image of $Q^{1/2}$. The same kind of structure is also present in the state equation of the parabolic problem studied in [225] and described in (2.104). In that case the same operator B acts on the drift and on the diffusion but it is unbounded, so the theory described in this chapter cannot be used. Still, such a similarity in the structure suggests that some further development of the theory will probably be able to treat such a case. ■

5.7 Results in Special Cases

In this section we present further results about existence and uniqueness of solutions of HJB equations when a certain "commutative assumption" for the operators A and Q is satisfied. We will indeed suppose (see Hypothesis 5.71-(D) for a more precise statement) that there exists an orthonormal basis of H made of eigenvectors of both A and Q.

The problem was studied in [3, 4, 123, 125] in this case. In this section we recall some results, mainly from [4, 123]. We omit the proofs. An element of interest of the approaches developed in [4, 125] is the use of variational solutions of the HJB equations. In this kind of approach the solution is defined via the duality pairing of the candidate solution with regular functions. Since the duality is obtained by extending an L^2 inner product on H, the use of this scheme is strictly linked to the identification of a reference measure on H.

5.7.1 Parabolic HJB Equations

We consider the following set of assumptions (similar to Hypothesis 5.1).

Hypothesis 5.71 (A) A is the generator of a strongly continuous semigroup $\{e^{tA},\ t \geq 0\}$ on a real separable Hilbert space H and there exist constants $M \geq 1$ and $\omega > 0$ such that

$$\left\| e^{tA} \right\| \le M e^{-\omega t}, \qquad \forall t \ge 0.$$

(B) $Q \in \mathcal{L}^+(H)$, $T > 0$ and $\mu := \left(\Omega^\mu, \mathscr{F}^\mu, \left\{\mathscr{F}_{\mu,s}\right\}_{s\in[0,T]}, \mathbb{P}^\mu, W_Q^\mu\right)$ is a generalized reference probability space.

(C) $e^{sA} Q e^{sA^*} \in \mathcal{L}_1(H)$ for all $s > 0$. Moreover, for all $t \ge 0$,

$$\int_0^t \mathrm{Tr}\left[e^{sA} Q e^{sA^*}\right] ds < +\infty,$$

so the symmetric positive operator

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

is of trace class for every $t \ge 0$.

(D) There exists an orthonormal basis $\{e_1, e_2, \ldots\}$ of H made of elements of $D(A)$ such that

$$Ax = \sum_{n=1}^{+\infty} -\alpha_n \langle e_n, x\rangle e_n, \quad x \in D(A),$$

for some eigenvalues $0 < \alpha_1 \le \alpha_2 \le \alpha_3 \ldots$ and

$$Qx = \sum_{i=n}^{+\infty} q_n \langle e_n, x\rangle e_n, \quad x \in H,$$

for a sequence of nonnegative eigenvalues q_n.

If Hypothesis 5.71 holds, the existence of an invariant measure m associated with the following Ornstein–Uhlenbeck process

$$\begin{cases} dX(s) = AX(s)ds + dW_Q(s), \quad 0 \le s \le T, \\ X(0) = x \in H \end{cases} \tag{5.121}$$

is proved, for example, in [180], Theorem 11.30, page 325. Observe that, differently from what we did in previous sections, here the reference measure is the invariant measure of the homogeneous Cauchy problem (which coincides with that of previous sections if $b = 0$ in (5.3)). For any $\phi \in C_b(H)$, the notation[6] $P_t[\phi](x)$ will be used

[6]In Chap. 4 and in Appendix B, when the transition semigroup reduces to the Ornstein–Uhlenbeck case, the notation R_t is used. In this section, and in the proof of Theorem 5.41, we keep the notation P_t even for the Ornstein–Uhlenbeck case because the semigroup plays exactly the same role, from the perspective of the L^2 approach to the HJB equation, as the semigroup P_t in Sect. 5.3 and, differently from Chap. 4 and Appendix B, the two semigroups never appear at the same time, so there is no possibility of confusion.

to denote the transition semigroup P_t for (5.121):

$$P_t[\phi](x) = \mathbb{E}\phi(X(t,x)).$$

Denoting by $\overline{\mathcal{B}}$ the completion of the Borel σ-field $\mathcal{B}(H)$ with respect to m, P_t extends to a strongly continuous semigroup of contractions on $L^2(H, \overline{\mathcal{B}}, m)$ with the generator

$$\begin{cases} \mathcal{A}\colon D(\mathcal{A}) \subset L^2(H,\overline{\mathcal{B}},m) \to L^2(H,\overline{\mathcal{B}},m) \\ \mathcal{A}\colon \phi \to \mathcal{A}\phi, \end{cases}$$

whose explicit expression on regular functions is

$$\mathcal{A}\phi(x) = \frac{1}{2}\,\mathrm{Tr}\,[QD^2\phi] + \langle Ax, D\phi\rangle\,. \tag{5.122}$$

When Hypothesis 5.71, and in particular its part (D), is satisfied, Remark 5.12 ensures that the operator D_Q introduced in Definition 5.11 is closable so that the closability problem we mentioned in Sect. 5.2.4 is no longer an issue. Therefore we work here with more conventional Sobolev spaces. We introduce them now together with some notations that will be useful in the variational approach to the solution of the HJB equation described below. Denote by $\mathcal{H}$ the space $L^2(H, \overline{\mathcal{B}}, m)$, by $\mathcal{V}$ the Sobolev space $W^{1,2}(H,m)$ made of all functions f of $L^2(H,\overline{\mathcal{B}},m)$ such that $Df \in L^2(H,\overline{\mathcal{B}},m)$, and by $\mathcal{V}^*$ its dual. Identifying $\mathcal{H}$ with its dual, one gets the following Gelfand triple

$$\mathcal{V} \subset \mathcal{H} \subset \mathcal{V}^*.$$

Given $T > 0$ we introduce

$$\mathcal{W}_T := \left\{ f \;:\; f \in L^2(0,T;\mathcal{V}),\ \frac{d}{dt}f \in L^2(0,T;\mathcal{V}^*) \right\}.$$

It follows, for instance, from Theorem 1.2.15 of [5] that $\mathcal{W}_T \subset C([0,T],\mathcal{H})$. In particular, given $f \in \mathcal{W}_T$, $f(T)$ is a well-defined element of $\mathcal{H}$ and thus an m-a.e. defined function from H to $\mathbb{R}$. We will use this fact in the following, in particular in the statements of Theorems 5.78 and 5.79.

Lemma 5.72 *Let Hypothesis 5.71 be satisfied. The operator $\mathcal{A}\colon D(\mathcal{A}) \subset \mathcal{H} \to \mathcal{H}$ extends uniquely to a linear operator $\tilde{\mathcal{A}} \in \mathcal{L}(\mathcal{V},\mathcal{V}^*)$ such that, for any $\phi,\psi \in \mathcal{V}$,*

$$\left\langle \tilde{\mathcal{A}}\phi, \psi \right\rangle_{\langle \mathcal{V}^*,\mathcal{V}\rangle} = \left\langle \phi, \tilde{\mathcal{A}}\psi \right\rangle_{\langle \mathcal{V},\mathcal{V}^*\rangle} = \frac{1}{2}\int_H \left\langle \sqrt{Q}D\phi, \sqrt{Q}D\psi \right\rangle_H dm(x).$$

Finally, $\tilde{\mathcal{A}}$ satisfies the following coercivity estimate: there exist $\alpha, \beta > 0$ such that, for any $\phi \in \mathcal{V}$,

$$-\left\langle \tilde{\mathcal{A}}\phi, \phi \right\rangle_{\langle \mathcal{V}^*, \mathcal{V} \rangle} \geq \alpha |\phi|^2_{\mathcal{V}} - \beta |\phi|^2_{\mathcal{H}} \tag{5.123}$$

and (5.123) holds in particular if one considers $\alpha = 1/2$ and $\beta = 1/2$.

Proof See [3], Lemma 4.2, p. 111. For the last statement, see [4], p. 503. □

Given a measurable map $G: \mathcal{V} \to \mathcal{V}^*$, a function $f \in L^2(0, T; \mathcal{V}^*)$ and $g \in \mathcal{H}$ we consider the equation

$$\begin{cases} u_t + \mathcal{A}u + G(u) + f(t, x) = 0, \\ u(T, x) = g(x). \end{cases} \tag{5.124}$$

Definition 5.73 A function $u \in \mathcal{W}_T$ is a solution of (5.124) in the variational sense, if for any $\psi \in \mathcal{V}$ and any $t \in [0, T]$,

$$\langle u(t), \psi \rangle = \langle g, \psi \rangle + \int_t^T \left\langle \tilde{\mathcal{A}} u(s), \psi \right\rangle_{\langle \mathcal{V}^*, \mathcal{V} \rangle} ds + \int_t^T \langle G u(s), \psi \rangle_{\langle \mathcal{V}^*, \mathcal{V} \rangle} ds + \int_t^T \langle f(s), \psi \rangle_{\langle \mathcal{V}^*, \mathcal{V} \rangle} ds. \tag{5.125}$$

Theorem 5.74 *Assume that Hypothesis 5.71 is satisfied. Assume that $G : \mathcal{V} \to \mathcal{V}^*$ and there exists a positive constant $K < \alpha$ (where α is the constant from (5.123)) such that:*

(G1) $|G(\xi)|_{\mathcal{V}^*} \leq K(1 + |\xi|_{\mathcal{V}})$ *for all* $\xi \in \mathcal{V}$,
(G2) $|G(\xi) - G(\eta)|_{\mathcal{V}^*} \leq K|\xi - \eta|_{\mathcal{V}}$ *for all* $\xi, \eta \in \mathcal{V}$.

Then, for every $g \in \mathcal{H}$ and $f \in L^2(0, T; \mathcal{V}^*)$ the evolution equation (5.124) has a unique solution in $\mathcal{W}_T$ in the sense of Definition 5.73.

Proof See Theorem 5.2 in [3]. □

One can remove the restriction $K < \alpha$ assuming a stronger regularity of the function G.

Theorem 5.75 *Assume that Hypothesis 5.71 is satisfied. Assume that $G : \mathcal{V} \to \mathcal{H}$ and there exists a positive constant K such that:*

(G1) $|G(\xi)|_{\mathcal{H}} \leq K(1 + |\xi|_{\mathcal{V}})$ *for all* $\xi \in \mathcal{V}$,
(G2) $|G(\xi) - G(\eta)|_{\mathcal{H}} \leq K|\xi - \eta|_{\mathcal{V}}$ *for all* $\xi, \eta \in \mathcal{V}$.

Then, for every $g \in \mathcal{H}$ and $f \in L^2(0, T; \mathcal{V}^*)$ the evolution equation (5.124) has a unique solution in $\mathcal{W}_T$ in the sense of Definition 5.73.

Proof See Theorem 5.3 in [3]. □

5.7.2 *Applications to Finite Horizon Optimal Control Problems*

Let Hypothesis 5.71 be satisfied. We denote by Λ the closed ball $\overline{B_\varrho}(0)$ of radius ϱ in H. Given some generalized reference probability space $\mu := \left(\Omega^\mu, \mathscr{F}^\mu, \left\{\mathscr{F}_{\mu,s}\right\}_{s\in[0,T]}, \mathbb{P}^\mu, W_Q^\mu\right)$ we consider the class of admissible controls given by

$$\mathcal{U}_0^\mu = \left\{a(\cdot)\colon [0,T] \to \Lambda \ : \ a(\cdot) \text{ is } \mathscr{F}_{\mu,s} - \text{progressively measurable}\right\}. \quad (5.126)$$

We consider the optimal control problem, in the weak formulation, characterized by the state equation

$$\begin{cases} dX(s) = (AX(s) + b(X(s)) + B(X(s))a(s))\, ds + dW_Q^\mu(s), \quad 0 \le s \le T \\ X(0) = x, \quad x \in H, \end{cases} \quad (5.127)$$

and the target functional

$$J^\mu(x; a(\cdot)) = \mathbb{E}^\mu \left\{ \int_0^T [f(s, X(s; 0, x, a(\cdot))) + h(a(s))]ds + g(X(T; 0, x, a(\cdot))) \right\}. \quad (5.128)$$

The hypotheses on the functions $b\colon H \to H$, $B\colon H \to \mathcal{L}(H)$, f, h and g are specified below.

Since we are interested in the weak formulation of the problem, we let the generalized reference probability space μ vary and we consider the set of controls given by

$$\overline{\mathcal{U}}_0 := \bigcup_\mu \mathcal{U}_0^\mu, \quad (5.129)$$

where $\mathcal{U}_0^\mu$ is defined in (5.126). The value function of the problem is

$$\overline{V}_0(x) = \inf_{a(\cdot)\in\overline{\mathcal{U}}_0} J^\mu(x; a(\cdot)). \quad (5.130)$$

The corresponding HJB equation is

$$\begin{cases} v_t + \mathcal{A}v + \langle b(x), Dv\rangle + F(x, Dv) + f(t, x) = 0, \\ \\ v(T, x) = g(x), \ x \in H, \end{cases} \quad (5.131)$$

where the Hamiltonian F is given by

$$F(x, p) = \inf_{a\in\Lambda} \left\{\langle B(x)a, p\rangle + h(a)\right\}. \quad (5.132)$$

If we introduce

$$G(v)(x) := \langle b(x), Dv(x)\rangle + F(x, Dv(x)),$$

equation (5.131) can be rewritten in the form (5.124),

$$\begin{cases} v_t + \mathcal{A}v + G(v) + f(t,x) = 0, \\ v(T,x) = g(x) \end{cases} \tag{5.133}$$

and Theorems 5.74 and 5.75 can be applied. One gets the following propositions, as corollaries.

Proposition 5.76 *Assume that Hypothesis 5.71 is satisfied. Suppose that b and $x \to B(x)a$, for any $a \in \Lambda$, are Borel measurable maps from H to H, have images in $\sqrt{Q}(H)$ and there exist two positive constants k_1 and k_2 such that*

$$|Q^{-1/2}b(x)| \le k_1(1+|x|) \quad \text{for all } x \in H \tag{5.134}$$

and

$$\|B^*(x)Q^{-1/2}\|_{\mathcal{L}(H)} \le k_2(1+|x|) \quad \text{for all } x \in H, \tag{5.135}$$

where $Q^{-1/2}$ denotes the pseudoinverse of $Q^{1/2}$. Moreover, assume that $h : \Lambda \to \mathbb{R}$ is measurable and bounded. Then, for any $g \in \mathcal{H}$ and $f \in L^2(0,T;\mathcal{V}^*)$, (5.133) has a unique solution $v \in \mathcal{W}_T$, provided that k_1 and k_2 are sufficiently small.

Proof See Corollary 4.3 in [4]. □

One can remove the restrictions on k_1 and k_2 if the regularity of b and B is stronger.

Proposition 5.77 *Assume that Hypothesis 5.71 is satisfied. Suppose that b and $x \to B(x)a$, for any $a \in \Lambda$, are Borel measurable maps from H to H, have images in $\sqrt{Q}(H)$, and that*

$$\sup_{x\in H} |Q^{-1/2}b(x)| < +\infty \tag{5.136}$$

and

$$\sup_{x\in H} \|B^*(x)Q^{-1/2}\|_{\mathcal{L}(H)} < +\infty, \tag{5.137}$$

where $Q^{-1/2}$ denotes the pseudoinverse of $Q^{1/2}$. Moreover, assume that $h : \Lambda \to \mathbb{R}$ is measurable and bounded. Then, for any $g \in \mathcal{H}$ and $f \in L^2(0,T;\mathcal{V}^*)$, (5.133) has a unique solution $v \in \mathcal{W}_T$.

Proof See Corollary 4.4 in [4]. □

We now state two results that ensure the existence of an optimal control and characterize the value function as the unique variational solution of the HJB equation.

Theorem 5.78 *Assume that the hypotheses of Proposition 5.76 are satisfied. Moreover, assume that:*

(i) $f \in L^2(0, T; \mathcal{V}^*)$.
(ii) $b\colon H \to H$ *and* $B\colon H \to \mathcal{L}(H)$ *are Lipschitz-continuous.*
(iii) $h : \Lambda \to \mathbb{R}$ *is lower semicontinuous.*

Then, for each initial datum $x \in H$*, there exists an optimal control for the optimal control problem (5.127)–(5.129). Moreover, if* $v \in \mathcal{W}_T \subset C([0, T], \mathcal{H})$ *is the unique solution of (5.133) and* $\overline{V}_0$ *is the value function defined in (5.130), we have* $v(0, x) = \overline{V}_0(x)$ *for* m*-a.e.* $x \in H$.

Proof See Theorem 5.4 in [4]. □

Theorem 5.79 *Assume that the hypotheses of Proposition 5.77 are satisfied. Moreover, assume that:*

(i) $f \in L^2(0, T; \mathcal{H})$.
(ii) $b\colon H \to H$ *and* $B\colon H \to \mathcal{L}(H)$ *are Lipschitz-continuous.*
(iii) $h : \Lambda \to \mathbb{R}$ *is lower semicontinuous.*

Then, for each initial datum $x \in H$*, there exists an optimal control for the optimal control problem (5.127)–(5.129) and the unique solution of (5.133) is given by the value function defined in (5.130). Moreover, if* $v \in \mathcal{W}_T \subset C([0, T], \mathcal{H})$ *is the unique solution of (5.133) and* $\overline{V}_0$ *the value function defined in (5.130), we have* $v(0, x) = \overline{V}_0(x)$ *for* m*-a.e.* $x \in H$.

Proof See Theorem 5.2 in [4]. □

Remark 5.80 We can compare the results and the assumptions of this last section with those obtained in the previous parts of the chapter. We observe that:

(i) In this section, differently from Sects. 5.2–5.4, the "commutative" Hypothesis 5.71-(D) is needed.
(ii) The Gâteaux differentiability of b, which was demanded in part (D) of Hypothesis 5.1 and then required in Sects. 5.2–5.4, is not needed here.
(iii) In the formulation of the state equation (5.78) we find $Q^{1/2}$ in front of the coefficient B. Even if in this respect the state equation (5.127) seems more general, the situation is not much different since Hypotheses (5.134)–(5.135) or (5.136)–(5.137) are needed.
(iv) While in Sect. 5.2 we consider the invariant measure m related to the nonhomogeneous Cauchy problem (5.3) (see Hypothesis 5.4), here m represents the invariant measure associated with the homogeneous stochastic equation (5.121). Still, as discussed after Theorem 5.41, in Sect. 5.4 the mild solution of the HJB equation can be characterized as a strong solution only if $b = 0$ and the properties of strong solutions are needed (see Sect. 5.5) to identify the solution of the HJB equation and the value function of the optimal control problem.
(v) The results in Sects. 5.2–5.4 refer to the case where the operator D_Q can be nonclosable. Conversely, as observed in Remark 5.12, Hypothesis 5.71, in particular Hypothesis 5.71-(D), implies the closability of the operator D_Q.

■

5.7.3 Elliptic HJB Equations

In this section we present some results regarding the use of L^2 theory for the elliptic equation (5.2). They are mainly taken from [125] which, to the best of our knowledge, is the only article where an L^2-approach for HJB equations arising from optimal control problems with infinite horizon is developed. A variational solution of the HJB equation, different from the one given in Definition 5.73, is used. The identification of the solution with the value function is not provided.

We introduce the following set of assumptions.

Hypothesis 5.81 (A) A is the generator of a strongly continuous semigroup $\{e^{tA},\ t \geq 0\}$ on a real separable Hilbert space H and there exist constants $M \geq 1$ and $\omega > 0$ such that

$$\left\| e^{tA} \right\| \leq M e^{-\omega t}, \qquad \forall t \geq 0.$$

Moreover, A is self-adjoint and $A^{-1} \in \mathcal{L}(H)$.

(B) $Q \in \mathcal{L}^+(H)$ and $\mathrm{Tr}[A^{-1}Q] < +\infty$.

(C) There exists a reflexive Banach space V with $D(A) \subset V \subset H$ having the following property: A extends to a continuous operator $A\colon V \to V^*$ (where V^* is the dual of V).

(D) $\mu := \left(\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s\in[0,+\infty)}, \mathbb{P}, W_Q\right)$ is a generalized reference probability space.

(E) There exists an orthonormal basis $\{e_1, e_2, \ldots\}$ of H made of elements of $D(A)$ such that

$$Ax = \sum_{n=1}^{+\infty} -\alpha_n \langle e_n, x\rangle e_n, \quad x \in D(A)$$

for some eigenvalues $0 < \alpha_1 < \alpha_2 < \alpha_3 \ldots$ and

$$Qx = \sum_{i=n}^{+\infty} q_n \langle e_n, x\rangle e_n, \quad x \in H,$$

for a sequence of nonnegative eigenvalues q_n.

We consider the following SDE

$$\begin{cases} dX(s) = AX(s)ds + dW_Q(s), & s > 0, \\ X(0) = x \in H \end{cases} \tag{5.138}$$

and denote by $X(\cdot; x)$ its mild solution at time t (the existence and the uniqueness of the solution are provided, for instance, by Theorem 1.147).

Proposition 5.82 *Suppose that Hypothesis 5.81 is satisfied. Then there exists a unique invariant measure m for (5.138). The measure m is a centered Gaussian measure supported in V with covariance operator* $\Gamma := -\frac{1}{2}A^{-1}Q$.

Proof See Theorem 6.2.1, p. 97 of [177]. □

We denote by $\overline{\mathcal{B}}$ the completion of the Borel σ-field $\mathcal{B}(H)$ with respect to m and by $\mathcal{H}$ the Hilbert space $L^2(H, \overline{\mathcal{B}}, m)$. We also denote by P_t, $t \geq 0$, the transition semigroup (indeed the Ornstein–Uhlenbeck semigroup) associated to (5.138). For any $\phi \in C_b(H)$ it is given by

$$P_t[\phi](x) = \mathbb{E}\phi(X(t, x)).$$

Proposition 5.83 *Suppose that Hypothesis 5.81 is satisfied. Then P_t extends to a strongly continuous semigroup of contractions on $L^2(H, \overline{\mathcal{B}}, m)$. Its generator $\mathcal{A}\colon D(\mathcal{A}) \subset L^2(H, \overline{\mathcal{B}}, m) \to L^2(H, \overline{\mathcal{B}}, m)$ is self-adjoint.*

Proof The first part of the proposition is a particular case of Proposition 5.9. The last claim is part of Lemma 2.4 of [125]. □

Notation 5.84 Denote by $\mathcal{I}$ the set of all sequences $\ell = (\ell_1, \ell_2, \dots) \in \mathbb{N}^{\mathbb{N}}$ such that $\ell_i = 0$, except for a finite number of indices. ■

Definition 5.85 Let $\{e_n\}$ be the orthonormal basis of H introduced in Hypothesis 5.81-(E). For $j = 0, 1, 2 \dots$, denote by h_j the standard j-th one-dimensional Hermite polynomials

$$h_j(\xi) := \frac{(-1)^j}{\sqrt{n!}} e^{\frac{\xi^2}{2}} \frac{d^j\left(e^{\frac{-\xi^2}{2}}\right)}{d\xi^j}, \quad \xi \in \mathbb{R}.$$

Given $\ell \in \mathcal{I}$ we define

$$K_\ell(x) := \prod_{i \in \mathbb{N}} h_{\ell_i}\left(\left\langle x, \Gamma^{-1/2} e_i \right\rangle_H\right), \qquad x \in H,$$

the *Hermite polynomial on H of index ℓ*.

Proposition 5.86 *Suppose that Hypothesis 5.81 is satisfied. The set of the Hermite polynomials K_ℓ is an orthonormal basis in $L^2(H, \overline{\mathcal{B}}, m)$. Moreover, for any $\ell \in \mathcal{I}$, $K_\ell \in D(\mathcal{A})$ and*

$$\mathcal{A}(K_\ell) = \Lambda_\ell K_\ell,$$

where $\Lambda_\ell := -\sum_i \ell_i \alpha_i$ (it is a finite sum), and the α_i are from Hypothesis 5.81-(E).

Proof See Theorem 9.1.5, p. 191 of [179] and Lemma 2.2 of [125]. □

Definition 5.87 We define the following function spaces:

(i) The *Gauss–Sobolev space of order k*, for $k = 1, 2, \ldots$, is the space $\mathcal{H}_k$ defined by

$$\mathcal{H}_k := \left\{ \phi \in \mathcal{H} \ : \ \left(\sum_{\ell \in \mathcal{I}} (1 - \Lambda_\ell)^k \langle \phi, K_\ell \rangle_{\mathcal{H}}^2 \right)^{1/2} = |(I - \mathcal{A})^{k/2} \phi|_{\mathcal{H}} < +\infty \right\}$$

(observe that the expression is well defined since all α_i and ℓ_i are nonnegative and then $\Lambda_\ell \leq 0$).

(ii) We denote by $\mathcal{H}_k^*$ the dual of $\mathcal{H}_k$.

(iii) Given the weight $\rho_n(x) := \left(1 + |x|^2\right)^n$ for $x \in H$, we denote by $\mathcal{H}_{0,n}$ the space

$$\mathcal{H}_{0,n} := \left\{ f \in \mathcal{H} \ : \ \int_H f^2(x) \rho_n(x) dm(x) \right\}$$

endowed with the usual L^2-weighted Hilbert space structure.

(iv) Given $k = 1, 2, \ldots$ and $n = 0, 1, \ldots$, we denote by $\mathcal{H}_{k,n}$ the space

$$\mathcal{H}_{k,n} := \mathcal{H}_k \cap \mathcal{H}_{0,n},$$

and by $\mathcal{H}_{k,n}^*$ its dual.

Observe that, for any $\phi \in D(\mathcal{A})$, we have

$$\sum_{\ell \in \mathcal{I}} |\Lambda_\ell|^2 \langle \phi, K_\ell \rangle_{\mathcal{H}}^2 = \sum_{\ell \in \mathcal{I}} \langle \mathcal{A}\phi, K_\ell \rangle_{\mathcal{H}}^2 = |\mathcal{A}\phi|_{\mathcal{H}}^2 < +\infty$$

so one can easily see that $D(\mathcal{A}) \subset \mathcal{H}_1$. $\mathcal{A}$ can be extended to the whole space $\mathcal{H}_1$ as is shown in the next lemma.

Lemma 5.88 *Suppose that Hypothesis 5.81 is satisfied. Then $\mathcal{A}$ extends to a continuous linear operator from $\mathcal{H}_1$ to $\mathcal{H}_1^*$.*

Proof See Lemma 2.4 of [125]. □

Hypothesis 5.89 (i) Λ is a Polish space.

(ii) $\tilde{R} \colon V \times \Lambda \to H$ is Borel measurable and such that, for some $n \geq 0$ and $R_0 > 0$,

$$|\tilde{R}(x, a)| \leq R_0 (1 + |x|^2)^{n/2} \qquad \text{for all } (x, a) \in V \times \Lambda.$$

We denote by $R \colon V \times \Lambda \to Q^{1/2}(H)$ the function $R := Q^{\frac{1}{2}} \tilde{R}$.

(iii) $\lambda \colon H \to \mathbb{R}^+$ is Borel measurable and there exist two real constants $\lambda_0, \lambda_1 > 0$ such that

$$\lambda_0 (1 + |x|^2)^n \leq \lambda(x) \leq \lambda_1 (1 + |x|^2)^n \qquad \text{for all } x \in H.$$

(iv) $l\colon V \times \Lambda \to \mathbb{R}$ is Borel measurable and there exists a $c_0 > 0$ such that

$$|l(x,a)| \le c_0(1+|x|^2)^{n/2} \qquad \text{for all } (x,a) \in V \times \Lambda.$$

We are interested in studying the HJB equation

$$(\lambda(x)I - \mathcal{A})\, v - F(v) = 0, \tag{5.139}$$

where

$$F(v)(x) := \inf_{a\in\Lambda} \left\{ \langle R(x,a), Dv(x)\rangle + l(x,a) \right\}.$$

Remark 5.90 The HJB equation (5.139) is associated with the optimal control problem characterized by:

(i) The state equation

$$\begin{cases} dX(s) = (AX(s) + R(X(s), a(s)))\, ds + dW_Q(s), & s > 0, \\ X(0) = x, \quad x \in H. \end{cases}$$

(ii) The cost functional

$$\int_0^{+\infty} e^{\int_0^t -\lambda(X(s))ds}\, l(X(t), a(t))dt.$$

(iii) The set of admissible controls

$$\mathcal{U}_0 = \{a(\cdot)\colon [0,+\infty) \to \Lambda \ : \ a(\cdot) \text{ is } \mathscr{F}_s\text{-progressively measurable}\}.$$

■

In order to define and study the solution of (5.139) we introduce the nonlinear operator

$$\mathscr{M}(v) := (\lambda(x)I - \mathcal{A})\, v - F(v)$$

which, thanks to Lemma 5.88, can be defined for any $v \in \mathcal{H}_{1,n}$. We have the following regularity result for $\mathscr{M}$.

Lemma 5.91 *Under Hypotheses 5.81 and 5.89 the operator $\mathscr{M}$ is locally bounded and Lipschitz continuous from $\mathcal{H}_{1,n}$ to $\mathcal{H}_{1,n}^*$. Moreover, if $\lambda_0 > R_0^2/2$, then there exists a $\delta > 0$ such that, for any $f, g \in \mathcal{H}_{1,n}$,*

$$\langle \mathscr{M}(f) - \mathscr{M}(g), f - g\rangle_{\langle \mathcal{H}_{1,n}^*, \mathcal{H}_{1,n}\rangle} \ge \delta |f-g|^2_{\mathcal{H}_{1,n}}.$$

Proof See Lemmas 4.1 and 4.2 of [125].

Definition 5.92 The function $v \in \mathcal{H}_{1,n}$ is a solution of (5.139) if

$$\langle \mathscr{M}(v), f \rangle_{\langle \mathcal{H}_{1,n}^*, \mathcal{H}_{1,n} \rangle} = 0$$

for any $f \in \mathcal{H}_{1,n}$.

Theorem 5.93 *If Hypotheses 5.81 and 5.89 are satisfied and $\lambda_0 > R_0^2/2$ then Eq. (5.139) has a unique solution v in the sense of the Definition 5.92. Moreover, $v \in \mathcal{H}_{2,n}$.*

Proof See Theorem 4.3 of [125]. □

5.8 Bibliographical Notes

In this chapter we focused our attention on HJB equations in L^2 spaces with respect to the invariant measure of an SDE with addictive noise and globally Lipschitz continuous drift independent of time. A number of existence results for various abstract classes of SDEs of this form can be found in the literature, for instance: for linear systems in [164, 354, 355], Sect. 6.2 of [177] and Sect. 11.5 of [180]; for the dissipative case in [164, 174, 426, 427, 533], Sects. 6.3 and 6.4 of [177] and Sect. 11.6 of [180]; for the case of a compact semigroup in [56, 164] and Sect. 11.7 of [180]; for equations with additive noise and weakly continuous drift in [120].[7]

Some approximation lemmas are presented in Sect. 5.2.2. Lemma 5.6 is a standard approximation result for uniformly continuous functions. Observe that in fact we do no need the approximating sequence to be in $\mathcal{E}_A(H)$, a weaker regularity would be enough for our purposes. The technique of mollification over finite-dimensional subspaces used to prove the pointwise convergences of Lemma 5.8 is well known (see e.g. Lemma 1.2, page 164 of [486] or [410]); we also use this kind of approach in the proof of Lemma B.78. The approximation result of Lemma 5.7 (especially its part (iii)) is *ad hoc* for the approximation of HJB equations in L^2 spaces. Even if we are not able to quote directly a specific published result, the proof uses completely standard arguments. Observe that the claim holds for any L^2 space on H w.r.t. any bounded measure, so the fact that we are working with an invariant measure of (5.8) plays no role. Obviously this specific measure is essential in Proposition 5.9. The claim of Proposition 5.9 is proved for the Ornstein–Uhlenbeck case (the proof is exactly the same), together with some characterization of the domain of the generator (the operator $\mathcal{A}$ defined at the beginning of Sect. 5.2.3), in [148, 149, 176], see also [121, 122, 152, 153, 184, 270, 297], Chap. 7 of [294] and Chap. 10 of [179]. We also mention, respectively, [417, 446] and [19] for the finite-dimensional and Banach space cases.

Lemma 5.37 provides a way to approximate elements of $D(\mathcal{A})$ even when its explicit characterization is missing. The space $\mathcal{F}C_0^{2,A^*}(H)$ is used because we can

[7] For uniqueness results the reader is referred to the review [443] and the references there.

explicitly compute the operator $\mathcal{A}$ in it (as well as other operators that will be defined later) and it is dense in $L^2(H, \overline{\mathcal{B}}, m)$. Other possible choices can be found in the literature, for example in Chap. 9 of [179] or in Chap. 8 of [153], the authors use, for the Gaussian case, a space of exponential functions. Using $\mathcal{F}C_0^{2,A^*}(H)$ is consistent with other similar approximations employed in the book, in particular in Chap. 4, see e.g. Hypotheses 4.133 and 4.141.

In Definition 5.11 we introduce a notion of Sobolev space for the case when the derivative operator $Q^{1/2}D$ is non-closable. Sobolev spaces in infinite dimension with respect to Gaussian measures are studied, for example, in [153, 484], Chap. 10 and [179], Chaps. 9 and 10. Sobolev spaces with respect to Gibbs measures are studied in [150, 151, 171, 172], Chap. 11 of [153] and Chap. 12 of [179]. In all of these cases the derivative operator is closable. Regarding the non-closable case needed here (see, in particular, Sect. 5.2.4) there is much less in the literature, the readers may consult [298, 299]. The closability of D_Q is related to the closability of the associated Dirichlet form, see [270, 509] for more on this and [422] for a general introduction to Dirichlet forms.

For some comments about the results of Lemma 5.14 and a discussion of the related literature, the readers may check the proof of Proposition 4.61 and Remark 4.62. The proofs of Lemmas 5.15 and 5.17 are standard but we could not find precise references. Results similar to Lemma 5.18 are often used in the literature as a step to prove Bismut–Elworthy–Li formulae, see for instance [486, 582] or [180], Sect. 9.4 (original results for the finite-dimensional case are, for example, in [60, 216]). In its proof, which expands the ideas contained in Step 1 of the proof of Proposition 2.4 of [298], the claim of Lemma 6.11 of [582], originally proved there for $b \in UC_b^2(H, H)$ and $\varphi \in UC_b^2(H)$, is extended. Results similar to Proposition 5.20 are given in [179] (they follow as corollaries of the proofs of Propositions 10.5.2 and 11.2.17) or in [184] (see p. 241); we follow here the arguments of [298]. More details and references about the claims of Remarks 5.21 are given in Sect. 4.3.1.3 and in the bibliographical notes of Chap. 4.

Sections 5.3 and 5.4 contain the main results of the chapter. We generalize the theorems contained in [298] to take into account Hamiltonians dependent on $x \in H$ and $t \in [0, T]$. In [298] only Hamiltonians of the form $F_0(D_Q u)$ were studied. Apart from this the setting is the same, beginning with Definition 5.28 of a mild solution. The main arguments used to prove the key result of Sect. 5.3, i.e. Theorem 5.35, are the same as those used in the proof of Theorem 3.7 of [298]. The proofs of Theorems 5.40 and 5.41 follow the lines of the proof of Proposition 4.3 of [298]. The literature on solutions of HJB equations in L^2 spaces is not very extensive and this chapter contains most of the published results (in Sects. 5.3, 5.4 and then in Sect. 5.7), so we cannot present a long genealogy of the results. However, many ideas and techniques have been used before to study HJB equations in spaces of regular functions discussed in Chap. 4. Thus we refer the reader to Sects. 4.4 and 4.5 and to the bibliographical notes of Chap. 4 for more.

The structure of Sect. 5.5 follows the structure of Sect. 4.8, starting from the proof of the fundamental identity (Lemma 5.50) and its use to obtain a verification theorem and optimal feedbacks (Theorem 5.53, Corollary 5.54, Theorem 5.55); the counterparts in Sect. 4.8 are Lemma 4.196, Theorem 4.197, Corollary 4.198 and Theorem 4.201. We refer the reader to the bibliographical notes of Chap. 4 for references on the subject. Compared to [298], the generalization of the Hamiltonian studied in Sects. 5.3 and 5.4 allows us to consider in Sect. 5.5 a more general optimal control problem, where the function R appearing in (5.78) also depends on s and $X(s)$ in addition to $a(s)$. Lemmas 5.46 and 5.49 are similar to results in [298], other proofs of the section are new. Proposition 5.44 is a standard existence and uniqueness result for solutions of stochastic evolution equations in Hilbert spaces, see the references mentioned in Chap. 1. Lemma 5.46 is a corollary of Girsanov's Theorem, the reader is referred, for example, to [44, 180, 382, 383, 448, 483, 580] for more on its Hilbert space formulations and various consequences. Because of the L^2 context, the result of Lemma 5.50 holds only m-almost everywhere. This is the main reason for introducing additional hypotheses (namely the boundedness of Λ used in Lemma 5.51 and the continuity of $J^{\mu}(t, x; \cdot)$) that we need in the proofs of Theorems 5.53 and 5.55. The formulations of the results of Sect. 5.5.5 are new even if the use of the non-degeneracy hypothesis, together with some continuity assumptions, was already suggested in Remark 3.10 of [298].

In Sect. 5.6 we show how some of the examples from Sect. 2.6 can be treated using the approach introduced in this chapter. We focus in particular on the existence of a (possibly non-degenerate) invariant measure, which is the key assumption needed here. For material on invariant measures for stochastic delay differential equations, besides Chap. 10 of [177, 299] which were already mentioned in Sect. 5.6.1, we refer the reader to [56, 338, 562]; for first-order stochastic equations, especially those connected to financial problems, results can be found in [299, 303, 430, 522, 553, 565] and Chap. 20 of [487].

The material of Sect. 5.7 essentially comes from [3, 4, 125]. More precisely, the results described in Sect. 5.7.1 (in particular Theorems 5.74 and 5.75) are proved in [3] (the two mentioned theorems correspond to Theorems 5.2 and 5.3 of [3]) while the content of Sect. 5.7.2 comes from [4]. Theorems 5.78 and 5.79 are Theorems 5.4 and 5.2, respectively, in [4]. Section 5.7.3 is based on the results obtained in [125] and the main theorem (Theorem 5.93) is Theorem 4.3 of [125]. In [123] the author uses a similar technique to deal with the Kolmogorov equation while in [125], Sect. 3, the authors study the related unbounded case. Even if we use in various parts of the book the variational solution of the state equation, this is the only section where we use the notion of a variational solution of the HJB equation (see Definitions 5.73 and 5.92). Indeed, it naturally needs some reference measure on the Hilbert state space and it is then linked to the study of HJB equations in the L^2 space. As far as we know, the above mentioned papers are the only ones that use this kind of notion of solution in the context of optimal control but, in the same spirit, a characterization of the value function for optimal stopping time problems, in terms of variational inequalities, is given in [38, 116], see also [125, 581, 583].

We also mention the recent paper [574] where the L^2 theory for HJB equations in Hilbert spaces, employing the ideas discussed in Sects. 5.1–5.5, is used to study an infinite horizon optimal control problem with boundary noise and boundary control.

The key prerequisite for the approach developed in this chapter is clearly the theory of invariant measures for infinite-dimensional PDEs. The results we use in this chapter concern invariant measures for SDEs with addictive noise, but the existing generalizations can be employed to develop applications to optimal control theory for other classes of stochastic partial differential equations in the spirit of the theory described here. In particular, the existence results for invariant measures for SPDEs with multiplicative noise (see, e.g., [218], Chap. 6 and Sect. 11.2 of [177] and Sect. 11.4 of [180]) and extensions to stochastic Burgers, Euler and Navier–Stokes equations (e.g. [7, 59, 81, 82, 159, 161, 253, 256, 336, 337, 389, 390, 515, 570, 571] and Chaps. 14 and 15 of [177]), stochastic reaction-diffusion equations (see for instance [109, 110]), stochastic porous media equations (as in [32, 169]) and stochastic nonlinear damped wave equations [31] can be a starting point in the study of optimal control problems driven by such state equations.

Results about invariant measures for transition semigroups for stochastic evolution equations in Banach spaces (such as those contained in [83, 292]) can be exploited to extend the techniques presented in this chapter to the Banach space case. Similarly the studies of SPDEs in domains/half-spaces and related invariant measures (see, e.g., [19, 165, 166, 494, 495, 497, 498, 546]) can be used as a first step to try to apply the methods to problems with state constraints. Another possible extension of the results presented here is the case of locally Lipschitz continuous Hamiltonians, following the results and the techniques introduced in [105, 307, 438].