Chapter 1
Preliminaries on Stochastic Calculus in Infinite Dimension

1.1 Basic Probability

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

1.1.1 Probability Spaces, σ-Fields

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

(i) A non-empty class of subsets of Ω, $\mathscr{F} \subset \mathcal{P}(\Omega)$, is called a *$\pi$-system* if it is closed under finite intersections.
(ii) A class of subsets of Ω, $\mathscr{F} \subset \mathcal{P}(\Omega)$, is called a *$\sigma$-field in* Ω if $\Omega \in \mathscr{F}$ and $\mathscr{F}$ is closed under complements and countable unions.
(iii) A class of subsets of Ω, $\mathscr{F} \subset \mathcal{P}(\Omega)$, is called a *$\lambda$-system* if:

- $\Omega \in \mathscr{F}$;
- if $A, B \in \mathscr{F}$, $A \subset B$, then $B \setminus A \in \mathscr{F}$;
- if $A_i \in \mathscr{F}$, $i = 1, 2, ...$, $A_i \uparrow A$, then $A \in \mathscr{F}$.

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

If $\mathscr{C} \subset \mathcal{P}(\Omega)$ and $A \subset \Omega$ we define $\mathscr{C} \cap A := \{B \cap A : B \in \mathscr{C}\}$. We denote by $\sigma_A(\mathscr{C} \cap A)$ the σ-field of subsets of A generated by $\mathscr{C} \cap A$. It is easy to see that $\sigma_A(\mathscr{C} \cap A) = \sigma(\mathscr{C}) \cap A$ (see, for instance, [18], p. 5).

G. Fabbri et al., *Stochastic Optimal Control in Infinite Dimension*,
Probability Theory and Stochastic Modelling 82,
DOI 10.1007/978-3-319-53067-3_1

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

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

Proof See [370], Theorem 1.1, p. 2. □

Corollary 1.3 *Let $\mathscr{G}$ be a π-system and $\mathscr{F}$ be the smallest family of subsets of Ω such that:*

- *$\mathscr{G} \subset \mathscr{F}$;*
- *if $A \in \mathscr{F}$ then $A^c \in \mathscr{F}$;*
- *if $A_i \in \mathscr{F}$, $A_i \cap A_j = \emptyset$ for $i, j = 1, 2, ..., i \neq j$, then $\cup_{i=1}^{\infty} A_i \in \mathscr{F}$.*

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

Proof Since $\sigma(\mathscr{G})$ satisfies the three conditions for $\mathscr{F}$, we obviously have $\mathscr{F} \subset \sigma(\mathscr{G})$. For the opposite inclusion it remains to observe that $\mathscr{F}$ is a λ-system. (For a self-contained proof, see also [180], Proposition 1.4, p. 17.) □

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

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

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

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

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

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

Let μ, ν be two measures on a measurable space $(\Omega, \mathscr{F})$. We say that μ is *absolutely continuous* with respect to ν (we write $\mu << \nu$) if for every $A \in \mathscr{F}$ such that $\nu(A) = 0$ we have $\mu(A) = 0$. If $\mu << \nu$ and $\nu << \mu$, we say that the measures μ and ν are *equivalent* (we write $\mu \sim \nu$). If there exists a set $A \in \mathscr{F}$ such that for every $B \in \mathscr{F}$ we have $\mu(B) = \mu(A \cap B)$, we say that μ is *concentrated on the set* A. If μ and ν are concentrated on disjoint sets we say that μ and ν are (mutually) *singular* and we write $\mu \perp \nu$.

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

Proof See [370], Lemma 1.17, p. 9. □

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

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

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

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

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

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

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

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

Every bounded measure on $(S, \mathcal{B}(S))$ is regular (see [478], Chap. II, Theorem 1.2). A bounded measure μ on $(S, \mathcal{B}(S))$ is called *tight* if for every $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset S$ such that $\mu(S \setminus K_\varepsilon) < \varepsilon$. If S is a Polish space then every bounded measure on $(S, \mathcal{B}(S))$ is tight (see [478], Chap. II, Theorem 3.2).

We refer to [58, 61, 267, 370, 478] for more on the general theory of measure and probability.

1.1.2 Random Variables

Definition 1.7 (*Random variable*) A *measurable* map X between two measurable spaces $(\Omega, \mathscr{F})$ and $(\tilde{\Omega}, \mathscr{G})$ is a called a *random variable*. This means that X is a random variable if $X^{-1}(A) \in \mathscr{F}$ for every $A \in \mathscr{G}$. We write it shortly as $X^{-1}(\mathscr{G}) \subset \mathscr{F}$. Sometimes we will just say that X is $\mathscr{F}/\mathscr{G}$-measurable.

If $\tilde{\Omega} = \mathbb{R}$ (resp. $\mathbb{R}^+$) and $\mathscr{G}$ is the Borel σ-field $\mathcal{B}(\mathbb{R})$ (resp. $\mathcal{B}(\mathbb{R}^+)$) then X is said to be a *real random variable* (resp. *positive random variable*).

If $\Omega, \tilde{\Omega}$ are topological spaces and $\mathscr{F}, \mathscr{G}$ are the Borel σ-fields then X is said to be *Borel measurable*.

If $(\Omega, \mathscr{F}, \mu)$ is a measure space and $X, X_1 \colon \Omega \to \tilde{\Omega}$, we say that X_1 is a *version* of X if $X = X_1$ μ-a.e.

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

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

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

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

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

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

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

Proof See, for instance, [370], Lemmas 1.4, 1.9, 1.10, and [520], Theorem 7.5, p. 138. □

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

Proof See [370], Lemma 1.13, p. 7, or [575] Theorem 1.7, p. 5. □

We refer to [58, 267, 370, 520] for more on measurability and for the general theory of integration.

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

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

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

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

The following theorem collects results that can be found in [478] (Chap. I, Theorems 2.8 and 2.12).

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

In particular, we have the following result.

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

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

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

Proof See [180], Lemma 1.3, p. 16. □

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

Proof The proof follows the lines of the proof of Lemma 1.25, p. 13, in [370].
Step 1: Let us assume first that $f = x\mathbf{1}_A$ ($\mathbf{1}_A$ denotes the characteristic function of the set A) for some $A \in \mathscr{G}_2$ and $x \in S$. By hypothesis, we can find $B \in \mathscr{G}_1$ s.t. $\mathbb{P}(A \Delta B) = 0$ and then the claim is proved if we choose $g_n \equiv g = x\mathbf{1}_B$. The same argument holds for a simple function f.
Step 2: For the case of a general f, thanks to Lemma 1.15 we can find a sequence of simple, $\mathscr{G}_2$-measurable functions f_n such that $d(f(\omega), f_n(\omega))$ monotonically decreases to 0. By Step 1, we can find simple, $\mathscr{G}_1$-measurable functions g_n such that $f_n = g_n$, $\mathbb{P}$-a.e. Thus the claim follows by taking $g(\omega) := \lim g_n(\omega)$ if the limit exists and $g(\omega) = s$ (for some $s \in S$) otherwise. □

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

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

Proof The embedding of V into E is continuous, so $\mathcal{B}(E) \cap V \subset \mathcal{B}(V)$. Since the embedding is also one-to-one, it follows from [478], Theorem 3.9, p. 21, that $\mathcal{B}(V) \subset \mathcal{B}(E)$, which completes the proof of (i). Parts (ii) and (iii) are direct consequences of (i). $f(\Omega)$ is separable because E is separable, so Part (iv) is a particular case of the Pettis theorem, see [488] Theorem 1.1. □

Lemma 1.18 *Let $(\Omega, \mathscr{F})$ be a measurable space and (S_1, ρ_1), (S_2, ρ_2) be two metric spaces with S_1 separable. Let $f : \Omega \times S_1 \to S_2$ be such that*

(i) *for each $x \in S_1$, the function $f(\cdot, x) : \Omega \to S_2$ is $\mathscr{F}/\mathcal{B}(S_2)$-measurable;*
(ii) *for each $\omega \in \Omega$ the function $f(\omega, \cdot) : S_1 \to S_2$ is continuous.*

Then $f : \Omega \times S_1 \to S_2$ is $\mathscr{F} \otimes \mathcal{B}(S_1)/\mathcal{B}(S_2)$-measurable.

Proof See Lemma 4.51, p. 153 of [8]. □

Notation 1.19 If E is a Banach space we denote by $|\cdot|_E$ its norm. Given two Banach spaces E and F, we denote by $\mathcal{L}(E, F)$ the Banach space of all continuous linear operators from E to F. If $E = F$ we will usually write $\mathcal{L}(E)$ instead of $\mathcal{L}(E, F)$. If H is a Hilbert space we denote by $\langle \cdot, \cdot \rangle$ its inner product. We will always identify H with its dual via Riesz representation theorem. If V, H are two real separable Hilbert spaces, we denote by $\mathcal{L}_2(V, H)$ the space of Hilbert–Schmidt operators from V to H (see Appendix B.3). The space $\mathcal{L}_2(V, H)$ is a real separable Hilbert space with the inner product $\langle \cdot, \cdot \rangle_2$, see Proposition B.25. ■

Lemma 1.20 *Let $(\Omega, \mathscr{F})$ be a measurable space and V, H be real separable Hilbert spaces. Suppose that $F : \Omega \to \mathcal{L}_2(V, H)$ is a map such that for every $v \in V$, $F(\cdot)v$ is $\mathscr{F}/\mathcal{B}(H)$-measurable. Then F is $\mathscr{F}/\mathcal{B}(\mathcal{L}_2(V, H))$-measurable.*

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

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

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

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

where

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

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

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

Assume now that, for all $t \in I$ and for some $m \geq 0$, $f(t, \cdot) \in B_m(E, F)$ (the space of Borel measurable functions with polynomial growth m, see Appendix A.2 for the precise definition). It is not true in general that the function

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

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

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

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

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

is not measurable (see e.g. [180], Sect. 1.2). Since $\mathcal{L}(L^2(\mathbb{R})) \subset B_1(L^2(\mathbb{R}), L^2(\mathbb{R}))$ and the norm in $\mathcal{L}(L^2(\mathbb{R}))$ is equivalent to the one induced by $B_1(L^2(\mathbb{R}), L^2(\mathbb{R}))$, the claim follows in a straightforward way.

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

Lemma 1.21 *Let I and Λ be two Polish spaces. Let μ be a measure defined on the Borel σ-field $\mathcal{B}(I)$ and denote by $\overline{\mathcal{B}(I)}$ the completion of $\mathcal{B}(I)$ with respect to μ. Let $f : I \times \Lambda \to \mathbb{R}$ be Borel measurable and such that for every $t \in I$, $f(t, \cdot)$ is bounded from below (respectively, above). Then the function*

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

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

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

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

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

is Lebesgue measurable.

Proof The first part is Example 7.4.2 in Volume 2 of [61] (recall that Polish spaces are Souslin spaces, see [61], Definition 6.6.1, and so $I \times \Lambda$ is a Souslin space).

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

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

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

Definition 1.22 (*Independence*) Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let I be a set of indices, and $\mathscr{C}_i \subset \mathscr{F}$ for all $i \in I$. We say that the families $\mathscr{C}_i$, $i \in I$, are independent if, for every finite subset J of I and every choice of $A_i \in \mathscr{C}_i$, $(i \in J)$, we have

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

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

Lemma 1.23 *Consider a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$*. Let* $\mathscr{C}_i \subset \mathscr{F}$ *be a* π*-system for every* $i \in I$*. If* $\mathscr{C}_i$, $i \in I$*, are independent, then* $\sigma(\mathscr{C}_i)$, $i \in I$*, are independent.*

Proof See [370] Lemma 2.6, p. 27. □

[2] Note that $\underline{f}$ is not always Borel measurable, see [61] Volume 2, Exercise 6.10.42(ii), p. 59.

1.1.3 The Bochner Integral

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

Lemma 1.24 *Let $X: (\Omega, \mathscr{F}) \to E$ be a random variable. Then the real-valued function $|X|_E$ is measurable.*

Proof See [180] Lemma 1.2, p. 16. □

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

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

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

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

(see e.g. [191] Theorem 7.17 p. 104). We will often write $L^p(\Omega, \mu; E)$ or $L^p(\Omega; E)$ for $L^p(\Omega, \mathscr{F}, \mu; E)$ and denote the norm by $|X|_{L^p}$ when the context is clear. If H is a separable Hilbert space, then $L^2(\Omega, \mathscr{F}, \mu; H)$ is a Hilbert space as well, equipped with the scalar product $\langle X, Y \rangle_{L^2(\Omega, \mathscr{F}, \mu; H)} = \int_\Omega \langle X(\omega), Y(\omega) \rangle_H \, d\mu(\omega)$.

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

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

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

Lemma 1.25 *If $\mathscr{F}$ is countably generated apart from null sets then $L^p(\Omega, \mathscr{F}, \mu; E)$ is a separable Banach space.*

Proof See [194], p. 92. □

Definition 1.26 (*Bochner integral*) Let $X: (\Omega, \mathscr{F}, \mu) \to E$ be a simple random variable $X = \sum_{i=1}^N x_i \mathbf{1}_{A_i}$, where $x_i \in E$, $A_i \in \mathscr{F}$, $\mu(A_i) < +\infty$. The *Bochner integral* of X is defined as

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

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

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

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

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

Remark 1.27 It follows easily from Lemma 1.15 that, for $X \in L^1(\Omega, \mathscr{F}, \mu; E)$, there always exists a sequence of simple random variables $X_n : (\Omega, \mathscr{F}, \mu) \to E$ as in Definition 1.26, satisfying (1.3). ■

Proposition 1.28 *Let $X \in L^1(\Omega, \mathscr{F}, \mu; E)$. Then the Bochner integral of X is well defined and does not depend on the choice of the sequence. Moreover,*

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

Proof See [180] Sect. 1.1 (in particular inequality (1.6), p. 19, and the part below Lemma 1.5). The proof there is done for a probability measure μ, but the general case is identical. □

Proposition 1.29 *Assume that $(\Omega, \mathscr{F}, \mu)$ is a complete measure space, E and F are separable Banach spaces and $A : D(A) \subset E \to F$ is a closed operator (see Definition B.3). If $X \in L^1(\Omega, \mathscr{F}, \mu; E)$ and $X \in D(A)$ a.s., then AX is an F-valued random variable, and X is a $D(A)$-valued random variable, where $D(A)$ is endowed with the graph norm of A (see Definition B.3). If, moreover, $\int_\Omega |AX(\omega)|_F \, d\mu(\omega) < +\infty$, then*

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

Proof The facts that X is a $D(A)$-valued random variable and AX is an F-valued random variable follow from Lemma 1.17-(ii). For the last part, see the proof of Proposition 1.6, Chap. 1 of [180]. □

Corollary 1.30 *Assume that E and F are separable Banach spaces and $T : E \to F$ is a continuous linear operator. If $X \in L^1(\Omega, \mathscr{F}, \mu; E)$, then*

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

Proof This is a particular case of Proposition 1.29. □

Remark 1.31 In this subsection we assumed that the space E is separable. This was done for simplicity and since we will only need this case in the vast majority of the book. However, the Bochner integral of a random variable $X : (\Omega, \mathscr{F}, \mu) \to E$ can also be defined when E is non-separable, see Sect. II.2 of [190]. If E is non-separable the definition of measurability is different. The random variable X is called measurable if there exists a sequence of simple random variables $X_n : (\Omega, \mathscr{F}, \mu) \to E$ such that $\lim_{n\to+\infty} |X(\omega) - X_n(\omega)|_E = 0$ μ-a.e. When E is separable this definition of measurability is equivalent to ours. Most of the results on the Bochner integral still hold in the non-separable case. In particular, Proposition 1.29 (hence also Corollary 1.30) still holds in the following form, which we will use later in Chap. 4 (see, for example, the proof of Corollary 4.14 and of Theorem 4.80).

Let $(\Omega, \mathscr{F}, \mu)$ be a complete measure space, E and F be Banach spaces and $A : D(A) \subset E \to F$ be a closed operator. If $X \in L^1(\Omega, \mathscr{F}, \mu; E)$ and $AX \in L^1(\Omega, \mathscr{F}, \mu; F)$, then

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

This is Theorem 6, p. 47 of [190]. ■

Theorem 1.32 *Let $(\Omega_1, \mathscr{F}_1)$ and $(\Omega_2, \mathscr{F}_2)$ be two measurable spaces and μ_1 (respectively μ_2) be a σ-finite measure on $(\Omega_1, \mathscr{F}_1)$ (respectively on $(\Omega_2, \mathscr{F}_2)$). Then there exists a unique measure $\mu_1 \otimes \mu_2$ on $\mathscr{F}_1 \otimes \mathscr{F}_2$ such that, for every $A \in \mathscr{F}_1$ and $B \in \mathscr{F}_2$ with finite measure,*

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

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

Proof See Theorem 8.2, p. 160 in Chap. VI, Sect. 8 of [397]. □

Theorem 1.33 (Fubini's Theorem) *Let $(\Omega_1, \mathscr{F}_1)$ and $(\Omega_2, \mathscr{F}_2)$ be two measurable spaces and μ_1 (respectively μ_2) be a σ-finite measure on $(\Omega_1, \mathscr{F}_1)$ (respectively on $(\Omega_2, \mathscr{F}_2)$). Let E be a separable Banach space with norm $|\cdot|_E$.*

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

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

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

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

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

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

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

Proof See Theorems 8.4, p. 162, and 8.7, p. 165 in Chap. VI, Sect. 8 of [397]. □

Theorem 1.34 *Let* E *be a separable Banach space and* μ *be a bounded measure on* $(E, \mathcal{B}(E))$*. Then the set of uniformly continuous and bounded functions* $UC_b(E)$ *is dense in* $L^p(E, \mathcal{B}(E), \mu)$ *for* $1 \le p < +\infty$*.*

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

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

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

1.1.4 Expectation, Covariance and Correlation

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

Definition 1.35 (*Expectation*) Given X in $L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$, we denote by $\mathbb{E}[X]$ the (Bochner) integral $\int_\Omega X(\omega) d\mathbb{P}(\omega)$. $\mathbb{E}[X]$ is said to be the *expectation* (or the *mean*) of X.

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

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

Definition 1.36 (*Covariance operator, correlation*) Given a real, separable Hilbert space H and $X \in L^2(\Omega, \mathscr{F}, \mathbb{P}; H)$, the *covariance operator* of X is defined by

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

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

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

Remark 1.37 For $X \in L^2(\Omega, \mathscr{F}, \mathbb{P}; H)$, the operator $Cov(X)$ is positive, symmetric and nuclear (see [180], p. 26). ■

1.1.5 Conditional Expectation and Conditional Probability

Theorem 1.38 *Consider a separable Banach space E, a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a sub-σ-field $\mathscr{G} \subset \mathscr{F}$. There exists a unique contractive linear operator $\mathbb{E}[\cdot|\mathscr{G}]\colon L^1(\Omega, \mathscr{F}, \mathbb{P}; E) \to L^1(\Omega, \mathscr{G}, \mathbb{P}; E)$ such that*

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

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

Proof See [180] Proposition 1.10, p. 26, and [458] Proposition V-2-5, pp. 102–103. □

Definition 1.39 (*Conditional expectation*) Given $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$, the random variable $\mathbb{E}[X|\mathscr{G}] \in L^1(\Omega, \mathscr{G}, \mathbb{P}; E)$, defined by Theorem 1.38, is called the *conditional expectation* of X given $\mathscr{G}$.

Definition 1.40 Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let E be a separable Banach space. A family $\mathcal{H}$ of integrable random variables $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$ is called *uniformly integrable* if

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

The following proposition collects various properties of conditional expectation (see e.g. [487] Proposition 3.15, p. 25, see also [572] Sect. 9.7, p. 88, for similar properties for real-valued random variables).

Proposition 1.41 *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let E be a separable Banach space. The conditional expectation has the following properties:*

(i) If $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$ is $\mathscr{G}$-measurable, then $\mathbb{E}[X|\mathscr{G}] = X$ $\mathbb{P}$-a.s.

(ii) *Given* $X \in L^1(\Omega, \mathscr{F}, \mathbb{P}; E)$ *and two* σ*-fields* $\mathscr{G}_1$ *and* $\mathscr{G}_2$ *such that* $\mathscr{G}_1 \subset \mathscr{G}_2 \subset \mathscr{F}$,

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Proposition 1.42 *Let* $(\Omega, \mathscr{F}, \mathbb{P})$ *be a probability space. Then:*

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

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

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

$$\mathbb{E}[\liminf_{n\to\infty} X_n|\mathscr{G}] \leq \liminf_{n\to\infty} \mathbb{E}[X_n|\mathscr{G}] \quad \mathbb{P}\text{-}a.s.$$

Proof See [572], Sect. 9.7, p. 88. □

Proposition 1.43 *Let $(E_1, \mathscr{E}_1)$ and $(E_2, \mathscr{E}_2)$ be two measurable spaces and $\psi : E_1 \times E_2 \to \mathbb{R}$ be a bounded measurable function. Let X_1, X_2 be two random variables in a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with values in $(E_1, \mathscr{E}_1)$ and $(E_2, \mathscr{E}_2)$ respectively, and let $\mathscr{G} \subset \mathscr{F}$ be a σ-field. If X_1 is $\mathscr{G}$-measurable and X_2 is independent of $\mathscr{G}$, then*

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

where

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

Proof See Proposition 1.12, p. 28 of [180]. □

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

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

Definition 1.44 Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and $\mathscr{G}$ be a sub-σ-field of $\mathscr{F}$. A function $p\colon \Omega \times \mathscr{F} \to [0, 1]$ is called a *regular conditional probability* given $\mathscr{G}$ if it satisfies the following conditions:

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

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

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

Theorem 1.45 *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, where $(\Omega, \mathscr{F})$ is a standard measurable space. Then, for every sub-σ-field $\mathscr{G} \subset \mathscr{F}$, there exists a regular conditional probability $p(\cdot, \cdot)$ given $\mathscr{G}$. Moreover, if $p'(\cdot, \cdot)$ is another regular conditional probability given $\mathscr{G}$, then there exists a set $N \in \mathscr{G}$, $\mathbb{P}(N) = 0$, such that, if $\omega \notin N$ then $p(\omega, A) = p'(\omega, A)$ for all $A \in \mathscr{F}$.*

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

Proof See Theorem 8.1, p. 147 in [478], or Theorems 3.1, 3.2, and the corollary following them in [356] (see also [575] Proposition 1.9, p. 11). □

Notation 1.46 If the regular conditional probability exists, we will often write $\mathbb{P}(\cdot|\mathscr{G})(\omega)$ or $\mathbb{P}_\omega$ for $p(\omega, \cdot)$. ■

Definition 1.47 (*Law of a random variable*) Given a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, a measurable space $(\Omega_1, \mathscr{F}_1)$, and a random variable $X\colon (\Omega, \mathscr{F}) \to (\Omega_1, \mathscr{F}_1)$, the probability measure on $(\Omega_1, \mathscr{F}_1)$ defined by

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

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

Proposition 1.48 (Change of variables) *Given a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$*, a measurable space* $(\Omega_1, \mathscr{F}_1)$*, a random variable* $X\colon (\Omega, \mathscr{F}) \to (\Omega_1, \mathscr{F}_1)$*, and a bounded Borel function* $\varphi : \Omega_1 \to \mathbb{R}$ *we have*

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

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

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

Lemma 1.50 *Consider a probability space* $(\Omega, \mathscr{F}, \mathbb{P})$ *and a Polish space* (S, d) *endowed with the Borel* σ*-field. Let* $X_n\colon \Omega \to S$ *and* $X\colon \Omega \to S$ *be random variables.*

(i) If X_n *converges to* X $\mathbb{P}$-a.s. *then* X_n *converges to* X in probability.
(ii) If X_n *converges to* X in probability *then* X_n *converges to* X in law.
(iii) If X_n *converges to* X in probability *then it contains a subsequence* X_{n_k} *such that* X_{n_k} *converges to* X $\mathbb{P}$-a.s.
(iv) (Egoroff's theorem) *If* X_n *converges to* X $\mathbb{P}$-a.s. *then for every* $\varepsilon > 0$*, there exists an* $\tilde{\Omega} \in \mathscr{F}$ *such that* $\mathbb{P}(\Omega \setminus \tilde{\Omega}) < \varepsilon$*, and* X_n *converges uniformly to* X *on* $\tilde{\Omega}$.

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

(v) *Let* $X, X_n \in L^p(\Omega, \mathscr{F}, \mathbb{P}; E), n \in \mathbb{N}, p \geq 1$*, and* E *be a separable Banach space. If* X_n *converges to* X *in* $L^p(\Omega, \mathscr{F}, \mathbb{P}; E)$*, then* X_n *converges to* X in probability.

Proof For (i), (ii) and (iii) see, for instance, [370] Lemmas 4.2, p. 63 and 4.7, p. 66. Part (iv) can be found, for instance, in [73] Theorem 2, p. 170, Sect. 4.5.4. Property (v) is straightforward. □

Lemma 1.51 *Let* $p > 1$ *and* $X, X_n \in L^p(\Omega, \mathscr{F}, \mathbb{P}; E), n \in \mathbb{N}$*, for some separable Banach space* E*. Suppose that, for some* $M > 0$*,* $\mathbb{E}\left[|X_n|_E^p\right] \leq M$ *for all* $n \in \mathbb{N}$*. If* $X_n \to X$ *in probability, then* $\mathbb{E}\left[|X - X_n|_E\right] \to 0$.

Proof Since the sequence (X_n) is bounded in $L^p(\Omega, \mathscr{F}, \mathbb{P}; E)$, it is uniformly integrable (see e.g. [572], p. 127, Sect. 13.3). The claim follows, for example, from Theorem 13.7, p. 131 of [572]. □

1.1.6 Gaussian Measures on Hilbert Spaces and the Fourier Transform

In this section we recall the notions of Gaussian measure and the Fourier transform for Hilbert space-valued random variables. For an extensive treatment of the subject we refer to [180], Chap. 2, [153], Chap. 1 or [154], Chap. 1.

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

Proposition 1.52 *Consider a real, separable Hilbert space* H *with the Borel* σ*-field* $\mathcal{B}(H)$ *and a probability measure* $\mathbb{P}$ *on* $(H, \mathcal{B}(H))$*. If* $\int_H |y|\, d\mathbb{P}(y) < +\infty$*, then we can define*

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

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

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

Proof See [153], p. 7. □

Definition 1.53 (*Mean and covariance of a measure on* H) We call m and Q, defined by Proposition 1.52, respectively the *mean* and the *covariance* of $\mathbb{P}$. In other words, the mean (respectively covariance) of $\mathbb{P}$ is the mean (respectively covariance) of the identity random variable $I\colon (H, \mathcal{B}(H), \mathbb{P}) \to (H, \mathcal{B}(H))$.

Definition 1.54 (*Fourier transform of a measure*) Let H be a Hilbert space and $\mathcal{B}(H)$ be its Borel σ-field. Given a probability measure $\mathbb{P}$ on $(H, \mathcal{B}(H))$ we define, for $x \in H$,

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

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

Proposition 1.55 *Let H be a real, separable Hilbert space, $\mathcal{B}(H)$ be its Borel σ-field, and $\mathbb{P}_1$ and $\mathbb{P}_2$ be two probability measures on $(H, \mathcal{B}(H))$. If $\hat{\mathbb{P}}_1(x) = \hat{\mathbb{P}}_2(x)$ for all $x \in H$, then $\mathbb{P}_1 = \mathbb{P}_2$.*

Proof See [153] Proposition 1.7, p. 6, or [180], Proposition 2.5, p. 35. □

Theorem 1.56 *Let $X_1, \ldots, X_n$ be random variables in a real, separable Hilbert space H. The random variables are independent if and only if for every $y_1, \ldots, y_n \in H$*

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

Proof Obviously if $X_1, \ldots, X_n$ are independent then (1.7) holds. Also, Theorem 1.56 is well known if $H = \mathbb{R}^k$. Let now $k \in \mathbb{N}$ and $y_i^j \in H$, $i = 1, \ldots, n$, $j = 1, \ldots, k$, and consider random variables $X_i^k = (\langle X_i, y_i^1\rangle, \ldots, \langle X_i, y_i^k\rangle)$, $i = 1, \ldots, n$ in $\mathbb{R}^k$. Therefore, if (1.7) holds then X_i^k, $i = 1, \ldots, n$, are independent for every $k \in \mathbb{N}$ and $y_i^j \in H$, $j = 1, \ldots, k$. Since cylindrical sets of the form $\{x : (\langle x, y_i^1\rangle, \ldots, \langle x, y_i^k\rangle) \in A \in \mathcal{B}(\mathbb{R}^k)\}$ generate $\mathcal{B}(H)$ and are a π-system, the collection of sets $\{\omega : (\langle X_i, y_i^1\rangle, \ldots, \langle X_i, y_i^k\rangle) \in A \in \mathcal{B}(\mathbb{R}^k)\}$ over all $k \in \mathbb{N}$ and $y_i^j \in H$, $i = 1, \ldots, n$, $j = 1, \ldots, k$, $A \in \mathcal{B}(\mathbb{R}^k)$ is a π-system generating $\sigma(X_i)$. Thus, by Lemma 1.23, the sigma algebras $\sigma(X_1), \ldots, \sigma(X_n)$ are independent. □

Theorem 1.57 *Let H be a real, separable Hilbert space, $\mathcal{B}(H)$ be its Borel σ-field, $a \in H$, and $Q \in \mathcal{L}_1^+(H)$. Then there exists a unique probability measure $\mathbb{P}$ on $(H, \mathcal{B}(H))$ such that*

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

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

Proof See [153] Theorem 1.12, p. 12. □

Definition 1.58 (*Gaussian measure on H*) Let H be a real, separable Hilbert space, $\mathcal{B}(H)$ be its Borel σ-field, $a \in H$, and $Q \in \mathcal{L}_1^+(H)$. The unique probability measure $\mathbb{P}$ identified by Theorem 1.57 is called the *Gaussian measure* with mean a and covariance Q, and is denoted by $\mathcal{N}(a, Q)$. When $a = 0$ we will denote it by $\mathcal{N}_Q$ and call it a centered Gaussian measure.

We now provide two useful results about Gaussian measures.

Proposition 1.59 *Let $Q \in \mathcal{L}_1^+(H)$. Then for all $y, z \in H$*

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

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

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

where $Q^{-1/2}$ is the pseudoinverse of $Q^{1/2}$ (see Definition B.1). The map (called the "white noise function", see e.g. [154] Sect. 2.5)

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

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

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

Moreover, for all $m > 0$ we have

$$\int_H |x|^{2m}\mathcal{N}_Q(dx) \le K(m)[\mathrm{Tr}(Q)]^m \tag{1.10}$$

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

Proof Formula (1.8) follows from Proposition 1.2.4 in [179].

The second statement is proved, when $\ker Q = \{0\}$, in [154] Sect. 2.5.2 (see also Sect. 1.2.4 of [179]). Since here we do not assume $\ker Q = \{0\}$, we provide a proof. First we observe that $\ker Q = \ker Q^{1/2}$ and that $Q^{1/2}(H)$ is dense in $(\ker Q)^\perp$ since $Q^{1/2}$ is self-adjoint. Moreover, by Definition B.1, the pseudoinverse of $Q^{1/2}$ is the operator $Q^{-1/2} : Q^{1/2}(H) \to (\ker Q)^\perp$, hence the map $y \to \mathcal{Q}_y = \langle Q^{-1/2}y, x\rangle$ is well defined for all $y \in Q^{1/2}(H)$. Furthermore, thanks to formula (1.8), we have, for $y_1, y_2 \in Q^{1/2}(H)$

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

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

$$\int_H \mathcal{Q}_{y_1}(x)\mathcal{Q}_{y_2}(x)\mathcal{N}_Q(dx) = \langle y_1, y_2\rangle. \tag{1.11}$$

In view of the above the map $y \to \mathcal{Q}_y = \langle Q^{-1/2}y, x\rangle$ is an isometry and can be extended to $\overline{Q^{1/2}(H)} = (\ker Q)^\perp$ (endowed with the inner product inherited from H) and (1.11) extends to all $y_1, y_2 \in (\ker Q)^\perp$.

We remark that as pointed out in [154] Sect. 2.5.2, for a generic $y \in (\ker Q)^{\perp}$ the image $\mathcal{Q}_y$ is an element of $L^2(H, \mathcal{N}_Q)$, hence an equivalence class of random variables defined $\mathcal{N}_Q$-a.e.; in particular, writing $\mathcal{Q}_y(x) = \langle y, Q^{-1/2}x \rangle$, $\mathcal{N}_Q$-a.e., would be misleading since, as proved in [154] Proposition 2.22, $\mathcal{N}_Q(Q^{1/2}(H)) = 0$.

Concerning the third claim, by Proposition 2.19, p. 50, of [180], it holds for $m \in \mathbb{N}$. If $k-1 < m < k$ for $k = 1, 2, ...,$ we use

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

□

Theorem 1.60 (Cameron–Martin formula) *Let H be a real, separable Hilbert space. Let $a_1, a_2 \in H$ and $Q \in \mathcal{L}_1^+(H)$. Then:*

(1) The Gaussian measures $\mathcal{N}(a_1, Q)$ and $\mathcal{N}(a_2, Q)$ are either singular or equivalent.

(2) They are equivalent if and only if $a_1 - a_2 \in Q^{1/2}(H)$ and in this case

$$\frac{d\mathcal{N}(a_1, Q)}{d\mathcal{N}(a_2, Q)}(x) = \exp\left(\langle Q^{-1/2}(a_1 - a_2), Q^{-1/2}(x - a_2) \rangle - \frac{1}{2} \left| Q^{-1/2}(a_1 - a_2) \right|^2 \right)$$

for $\mathcal{N}(a_2, Q)$-a.e. $x \in H$.

Proof See Theorem 2.23, p. 53 of [180]. □

We now recall some results concerning compactness of a family of measures in $M_1(H)$ (see e.g. Sect. 2.1 in [180] or [219, 478] for more on this).

Definition 1.61

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

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

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

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

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

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

Theorem 1.62 *Let H be a real separable Hilbert space. A family $\Lambda \subset M_1(H)$ is relatively compact if and only if it is tight.*

Proof See [180], the proof of Theorem 2.3. □

The next theorem gives a useful sufficient condition for compactness.

Theorem 1.63 *Let H be a real separable Hilbert space and let $\{e_i\}_{i\in\mathbb{N}}$ be an orthonormal basis in H. A family $\Lambda\subset M_1(H)$ is relatively compact if*

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

Proof See [478], the proof of Theorem VI.2.2. □

Concerning Gaussian measures, we have the following result (see Proposition 1.1.5 of [493]).

Proposition 1.64 *Let $\mathcal{N}_{Q_n}$ ($n\in\mathbb{N}$) and $\mathcal{N}_Q$ be centered Gaussian measures on H. If $\lim_{n\to+\infty}\|Q_n-Q\|_{\mathcal{L}_1(H)}=0$, then the measures $\mathcal{N}_{Q_n}$ converge weakly to $\mathcal{N}_Q$.*

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

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

Since $\lim_{n\to+\infty}\|Q_n-Q\|_{\mathcal{L}_1(H)}=0$, the above formula implies in particular that Theorem 1.63 applies and thus the sequence $(\mathcal{N}_{Q_n})$ is relatively compact.

Moreover, from Theorem 1.57 and Definition 1.58 it is immediate that, as $n\to+\infty$,

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

Take now a subsequence $\mathcal{N}_{Q_{n_k}}$ weakly convergent to a probability measure $\mathbb{P}_0$. By Definition 1.54 we must have

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

This implies that $\widehat{\mathbb{P}_0}=\widehat{\mathcal{N}_Q}$ and hence, by Proposition 1.55, that $\mathbb{P}_0=\mathcal{N}_Q$. Since this is true for any convergent subsequence, the claim now follows by a standard contradiction argument. □

We conclude with a useful result about uniformity of weak convergence. The result is also true if H is a Polish space, see [478], Theorem II.6.8.

Theorem 1.65 *Let $\mathbb{P}_n$ be a sequence in $M_1(H)$ and $\mathbb{P}\in M_1(H)$. Then $\mathbb{P}_n$ is weakly convergent to $\mathbb{P}$ if and only if*

$$\lim_{n\to+\infty}\sup_{\phi\in\mathcal{C}_0}\left|\int_H\phi(x)\mathbb{P}_n(dx)-\int_H\phi(x)\mathbb{P}(dx)\right|=0$$

for every family $\mathcal{C}_0 \subset C_b(H)$ *which is equicontinuous at all points* $x \in H$ *and uniformly bounded, i.e., for some constant* $M > 0$, $|f(x)| \le M$ *for all* $x \in H$ *and* $f \in \mathcal{C}_0$.

Proof See [478], the proof of Theorem II.6.8. □

1.2 Stochastic Processes and Brownian Motion

1.2.1 Stochastic Processes

Definition 1.66 (*Filtration, usual conditions*) Let $t \ge 0$. A filtration $\left\{\mathscr{F}_s^t\right\}_{s\ge t}$ in a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is a family of σ-fields such that $\mathscr{F}_s^t \subset \mathscr{F}_r^t \subset \mathscr{F}$ whenever $t \le s \le r$.

(i) We say that $\left\{\mathscr{F}_s^t\right\}_{s\ge t}$ is *right-continuous* if, for all $s \ge t$, $\mathscr{F}_{s+}^t := \bigcap_{r>s} \mathscr{F}_r^t = \mathscr{F}_s^t$.

(ii) We say that $\left\{\mathscr{F}_s^t\right\}_{s\ge t}$ is *left-continuous* if, for all $s > t$, $\mathscr{F}_{s-}^t := \sigma\left(\bigcup_{r<s} \mathscr{F}_r^t\right) = \mathscr{F}_s^t$. We say that $\left\{\mathscr{F}_s^t\right\}_{s\ge t}$ is *continuous* if it is both left and right-continuous.

(iii) We say that $\left\{\mathscr{F}_s^t\right\}_{s\ge t}$ satisfies the *usual conditions* if it is right-continuous and complete, i.e. if $\mathscr{F}_s^t$ contains all $\mathbb{P}$-null sets of $\mathscr{F}$ for every $s \ge t$.

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

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

Definition 1.67 (*Filtered probability space*) Let $\mathscr{F}_s^t$ be a filtration satisfying the usual conditions on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$. The 4-tuple $\left(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}\right)$ is called a *filtered probability space*.

Notation 1.68 We use the following convention in this section. When we write $s \in [t, T]$ we mean that $s \in [t, T]$ if $T \in \mathbb{R}$, and $s \in [t, +\infty)$ if $T = +\infty$. So $[t, T]$ is understood to be $[t, +\infty)$ if $T = +\infty$. ■

Definition 1.69 (*Stochastic process*) Let $T \in (0, +\infty]$, $t \in [0, T)$ and $(\Omega, \mathscr{F})$ and $(\Omega_1, \mathscr{F}_1)$ be two measurable spaces. A family of random variables $X(\cdot) = \{X(s)\}_{s\in[t,T]}$, $X(s) \colon \Omega \to \Omega_1$, is called a *stochastic process* in $[t, T]$. If $(\Omega_1, \mathscr{F}_1) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ then $X(\cdot)$ is called a real stochastic process.

Definition 1.70 Let $\left(\Omega, \mathscr{F}, \left\{\mathscr{F}_s^t\right\}_{s\ge t}, \mathbb{P}\right)$ be a filtered probability space and $(\Omega_1, \mathscr{F}_1)$ be a measurable space. A stochastic process $\{X(s)\}_{s\in[t,T]} \colon [t, T] \times \Omega \to \Omega_1$ is said to be:

(i) *Measurable*, if the map $(s, \omega) \to X(s)(\omega)$ is $\mathcal{B}([t, T]) \otimes \mathscr{F}/\mathscr{F}_1$-measurable.
(ii) *Adapted*, if, for each $s \in [t, T]$, $X(s) \colon \Omega \to \Omega_1$ is an $\mathscr{F}_s^t/\mathscr{F}_1$-measurable random variable.
(iii) *Progressively measurable*, if for all $s \in (t, T]$, the restriction of $X(\cdot)$ to $[t, s] \times \Omega$ is $\mathcal{B}([t, s]) \otimes \mathscr{F}_s^t/\mathscr{F}_1$-measurable.
(iv) *Predictable*, if the map $(s, \omega) \to X(s)(\omega)$ is $\mathcal{P}_{[t,T]}/\mathscr{F}_1$-measurable, where $\mathcal{P}_{[t,T]}$ is the σ-field (the predictable σ-field) in $[t, T] \times \Omega$ generated by all sets of the form $(s, r] \times A$, $t \le s < r \le T$, $A \in \mathscr{F}_s^t$ and $\{t\} \times A$, $A \in \mathscr{F}_t^t$.
(v) If E is a separable Banach space (endowed with its Borel σ-field), the process $\{X(s)\}_{s\in[t,T]} \colon [t, T] \times \Omega \to E$ is called *stochastically continuous* at $s \in [t, T]$ if for every $\varepsilon, \delta > 0$ there exists $\rho > 0$ such that

$$\mathbb{P}\left(|X(r) - X(s)| \ge \varepsilon\right) \le \delta, \qquad \text{for all } r \in (s - \rho, s + \rho) \cap [t, T].$$

(vi) If (S, d) is a metric space (endowed with its Borel σ-field), the process $\{X(s)\}_{s\in[t,T]} \colon [t, T] \times \Omega \to S$ is called *continuous* (respectively, right-continuous, left-continuous), if for $\mathbb{P}$-a.e. $\omega \in \Omega$, the function $s \to X(s)(\omega)$ is continuous (respectively, right-continuous, left-continuous).
(vii) If E is a separable Banach space (endowed with its Borel σ-field), the process $\{X(s)\}_{s\in[t,T]} \colon [t, T] \times \Omega \to E$ is called *integrable* (respectively square-integrable) if $\mathbb{E}[|X(s)|] < +\infty$ (respectively $\mathbb{E}[|X(s)|^2] < +\infty$) for all $s \in [t, T]$. The process is called *uniformly integrable* if it is integrable and the family $\{X(s)\}_{s\in[t,T]}$ is uniformly integrable (see Definition 1.40).
(viii) If E is a separable Banach space (endowed with the Borel σ-field induced by the norm), the process $\{X(s)\}_{s\in[t,T]} \colon [t, T] \times \Omega \to E$ is said to be *mean square continuous* if $\mathbb{E}[|X(s)|^2] < +\infty$ for all $s \in [t, T]$ and $\lim_{r\to s} \mathbb{E}[|X(r) - X(s)|^2] = 0$ for all $s \in [t, T]$.

It is easy to see that if a process is mean square continuous then it is stochastically continuous.

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

Progressive measurability can also be defined using the concept of progressively measurable sets, see e.g. [447], p. 4, or [219], p. 71. We say that a set $A \subset [t, T] \times \Omega$ is $\mathscr{F}_s^t$-progressively measurable if the function $\mathbf{1}_A$ is a progressively measurable process. Equivalently this means that $A \cap ([t, s] \times \Omega) \in \mathcal{B}([t, s]) \otimes \mathscr{F}_s^t$ for every $s \in [t, T]$. It can be proved that the $\mathscr{F}_s^t$-progressively measurable sets form a σ-field and that a process $X(\cdot)$ is progressively measurable if and only if it is measurable with respect to the σ-field of $\mathscr{F}_s^t$-progressively measurable sets.

Definition 1.71 (*Stochastic equivalence, modification*) Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and $(\Omega_1, \mathscr{F}_1)$ be a measurable space. Processes $X(\cdot), Y(\cdot) \colon [t, T] \times \Omega \to \Omega_1$ are called *stochastically equivalent* if for all $s \in [t, T]$, $\mathbb{P}(X(s) = Y(s)) = 1$. In this case, $Y(\cdot)$ is said to be a *modification* or *version* of $X(\cdot)$. The processes $X(\cdot)$ and

$Y(\cdot)$ are called *indistinguishable* if $\mathbb{P}(X(s) = Y(s) : \forall s \in [t, T]) = 1$. We will also say that $Y(\cdot)$ is an *indistinguishable version* of $X(\cdot)$.

Lemma 1.72 *Let $\left(\Omega, \mathscr{F}, \left\{\mathscr{F}_s^t\right\}_{s\geq t}, \mathbb{P}\right)$ be a filtered probability space and let $\{X(s)\}_{s\geq t}$ be a process with values in a Polish space (S, d), endowed with the Borel σ-field induced by the distance.*

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

Proof Part (i): Since S is Borel isomorphic to a Borel subset A of $\mathbb{R}$, without loss of generality we can consider $X(\cdot)$ to be an $\mathbb{R}$-valued process with values in A. By [449], Theorem T46, p. 68, $X(\cdot)$ has an $\mathbb{R}$-valued, $\mathscr{F}_s^t$-progressively measurable modification $\tilde{X}(\cdot)$. Let $a \in A$. We define a process $Y(\cdot)$ by $Y(s) := \tilde{X}(s)\mathbf{1}_{\tilde{X}(s)\in A} + a\mathbf{1}_{\tilde{X}(s)\in(\mathbb{R}\setminus A)}$. The process $Y(\cdot)$ is $\mathscr{F}_s^t$-progressively measurable. Moreover, if $\tilde{X}(s) = X(s)$, then $Y(s) = X(s)$, so $Y(\cdot)$ is a modification of $X(\cdot)$. *Part (ii)*: See [449], Theorem T47, p. 70, or [372], Proposition 1.13, p. 5. □

Lemma 1.73 *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space and let $\{X(s)\}_{s\geq t}$ be a stochastic process with values in a separable Banach space E endowed with the Borel σ-field. If $X(\cdot)$ is stochastically continuous then it has a measurable modification.*

Proof See [180], Proposition 3.2. □

Lemma 1.74 *Let $\left(\Omega, \mathscr{F}, \left\{\mathscr{F}_s^t\right\}_{s\geq t}, \mathbb{P}\right)$ be a filtered probability space and let $\{X(s)\}_{s\geq t}$ be an adapted process with values in a separable Banach space E endowed with the Borel σ-field. If $X(\cdot)$ is stochastically continuous then it has an $\mathscr{F}_s^t$-progressively measurable modification.*

Proof See [180], Proposition 3.6. It is also a corollary of Lemmas 1.72-(i) and 1.73. □

1.2.2 Martingales

Notation 1.75 Unless specified otherwise, any Banach space E and any metric space (S, d) will be understood to be endowed with the Borel σ-field induced respectively by the norm and by the distance. ■

Definition 1.76 (*Martingale*) Let $\left(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}\right)$ be a filtered probability space, and let $M(\cdot)$ be an $\mathscr{F}_s^t$-adapted and integrable process with values in a separable Banach space E. Then $M(\cdot)$ is said to be a *martingale* if, for all $r, s \in [t, T]$, $s \leq r$,

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

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

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

Theorem 1.77 (Doob's maximal inequalities) *Let $T > 0$, $\left(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}\right)$ be a filtered probability space, and H be a separable Hilbert space. Let $M(\cdot)$ be a right-continuous H-valued martingale such that $M(s) \in L^p(\Omega, \mathscr{F}, \mathbb{P}; H)$ for all $s \in [t, T]$. Then:*

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

Proof We observe that, if $M(\cdot)$ is a right-continuous H-valued martingale such that $M(s) \in L^p(\Omega, \mathscr{F}, \mathbb{P}; H)$, $p \geq 1$, for all $s \in [t, T]$, then by Proposition 1.41-(vi), $|M(\cdot)|^p$ is a right-continuous $\mathbb{R}$-valued submartingale with $|M(s)| \in L^p(\Omega, \mathscr{F}, \mathbb{P}; \mathbb{R})$ for all $s \in [t, T]$. The claims now easily follow from [372] Theorem 3.8 (i) and (iii), pp. 13–14. □

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

Notation 1.78 (*Square-integrable martingales*) Let $T \in (0, +\infty)$, $t \in [0, T)$, let $\left(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}\right)$ be a filtered probability space, and E be a separable Banach space. The class of all continuous square-integrable martingales $M\colon [t, T] \times \Omega \to E$ is denoted by $\mathcal{M}_{t,T}^2(E)$. ■

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

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

is a Hilbert space (see [294], p. 22).

Theorem 1.79 (Angle bracket process, Quadratic variation process) *Let $T > 0$, $t \in [0, T)$, H be a separable Hilbert space, and $\left(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}\right)$ be a filtered probability space. For every $M \in \mathcal{M}_{t,T}^2(H)$ there exists a unique (real) increasing, adapted, continuous process starting from* 0 *at t, called the* angle bracket process, *and denoted by $\langle M\rangle_t$, such that $|M_s|^2 - \langle M\rangle_s$ is a continuous martingale. Moreover, there exists a unique $\mathcal{L}_1^+(H)$-valued continuous adapted process starting from* 0 *at t, called the* quadratic variation of M, *and denoted by $\langle\langle M\rangle\rangle_s$, such that, for all $x, y \in H$, the process*

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

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

Proof See [294], Definition 2.9 and Lemma 2.1, p. 22. □

Theorem 1.80 (Burkholder–Davis–Gundy inequality) *Let $T > 0$, $t \in [0, T)$, H be a separable Hilbert space, and $\left(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P}\right)$ be a filtered probability space. For every $p > 0$ there exists a $c_p > 0$ such that, for every $M \in \mathcal{M}_{t,T}^2(H)$ with $M(0) = 0$,*

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

Proof See [487], Theorem 3.49, p. 37. □

1.2.3 Stopping Times

Definition 1.81 (*Stopping time*) Consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a filtration $\left\{\mathscr{F}_s^t\right\}_{s\ge t}$ on Ω. A random variable $\tau : (\Omega, \mathscr{F}) \to [t, +\infty]$ is said to be an $\mathscr{F}_s^t$*-stopping time* if, for all $s \ge t$,

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

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

$$\mathscr{F}_\tau := \left\{A \in \mathscr{F} \ : \ A \cap \{\tau \le s\} \in \mathscr{F}_s^t \ \text{ for all } s \ge t\right\}.$$

Proposition 1.82 *Let $(\Omega, \mathscr{F}, \mathscr{F}_s^t, \mathbb{P})$ be a filtered probability space.*

(i) If τ and σ are $\mathscr{F}_s^t$-stopping times, so are $\tau \wedge \sigma$, $\tau \vee \sigma$ and $\tau + \sigma$.
(ii) If σ_n (for $n = 1, 2\ldots$) are $\mathscr{F}_s^t$-stopping times, then

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

are $\mathscr{F}_s^t$-stopping times.
(iii) For any $\mathscr{F}_s^t$-stopping time τ there exists a decreasing sequence of discrete-valued $\mathscr{F}_s^t$-stopping times τ_n, such that $\lim_{n\to\infty} \tau_n = \tau$.
(iv) Let (S, d) be a metric space (endowed with the Borel σ-field induced by the distance), and $X : [t, +\infty) \times \Omega \to S$ be a continuous and $\mathscr{F}_s^t$-adapted process. Let $A \subset S$ be an open or a closed set. Then the hitting time

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

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

Proof (i) and (ii) see [372], Lemmas 2.9 and 2.11, p. 7. (iii) see [370], Lemma 7.4, p. 122. (iv) see [575], Example 3.3, p. 24, or [452], Proposition 1.3.2, p. 12 (there $S = \mathbb{R}^n$, but the proofs are the same). □

Proposition 1.83 *Let $\left(\Omega, \mathscr{F}, \left\{\mathscr{F}_s^t\right\}_{s\geq t}, \mathbb{P}\right)$ be a filtered probability space, $(\Omega_1, \mathscr{F}_1)$ be a measurable space, $X\colon [t, +\infty) \times \Omega \to \Omega_1$ be an $\mathscr{F}_s^t$-progressively measurable process, and τ be an $\mathscr{F}_s^t$-stopping time. Then the random variable $X(\tau)$, (where $X(\tau)(\omega) := X(\tau(\omega), \omega)$), is $\mathscr{F}_\tau$-measurable and the process defined, for any $s \in [t, +\infty)$, by $X(s \wedge \tau)$ is $\mathscr{F}_s^t$-progressively measurable.*

Proof See [452], Proposition 1.3.5, p. 13, or [575], Proposition 3.5, p. 25. □

Theorem 1.84 (Doob's optional sampling theorem) *Let $\left(\Omega, \mathscr{F}, \left\{\mathscr{F}_s^t\right\}_{s\geq t}, \mathbb{P}\right)$ be a filtered probability space, $X\colon [t, +\infty) \times \Omega \to \mathbb{R}$ be a right-continuous $\mathscr{F}_s^t$-submartingale, and τ, σ be two $\mathscr{F}_s^t$-stopping times with τ bounded. Then X_τ is integrable and*

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

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

Proof See [370], Theorem 7.29, p. 135. □

Definition 1.85 (*Local martingale*) Let $\left(\Omega, \mathscr{F}, \left\{\mathscr{F}_s^t\right\}_{s\geq t}, \mathbb{P}\right)$ be a filtered probability space. An $\left\{\mathscr{F}_s^t\right\}_{s\geq t}$-adapted process $\{X(s)\}_{s\geq t}$ with values in a separable Banach space E is said to be a *local martingale* if there exists an increasing sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ with $\mathbb{P}(\tau_n \uparrow +\infty) = 1$, such that the process $\{X(s \wedge \tau_n)\}_{s\geq t}$ is a martingale for every $n \in \mathbb{N}$.

1.2.4 Q-Wiener Processes

Definition 1.86 (*Real Brownian motion*) Given $t \in \mathbb{R}$, a real stochastic process $\beta\colon [t, +\infty) \times \Omega \to \mathbb{R}$ on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is a *standard (one-dimensional) real Brownian motion* on $[t, +\infty)$ starting at 0, if

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

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

$Q_1 := JJ^*$ belongs to $\mathcal{L}_1^+(\Xi_1)$ and Ξ_0 is identical with the space $Q_1^{\frac{1}{2}}(\Xi_1)$ (see [180] Proposition 4.7, p. 85).

Theorem 1.87 *Consider the setting described above. Let $\{g_k\}_{k\in\mathbb{N}}$ be an orthonormal basis of Ξ_0 and $(\beta_k)_{k\in\mathbb{N}}$ be a sequence of mutually independent, standard one-dimensional Brownian motions $\beta_k\colon [t,+\infty)\times\Omega\to\mathbb{R}$ on $[t,+\infty)$ starting at 0. Then for every $s\in[t,+\infty)$ the series*

$$W_Q(s) := \sum_{k=1}^{\infty} g_k\beta_k(s) \tag{1.12}$$

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

Proof See [180] Propositions 4.3, p. 82, and 4.7, p. 85. □

Definition 1.88 (*Q-Wiener process*) The process W_Q defined by (1.12) is called a *Q-Wiener process* on $[t,+\infty)$ starting at 0.

Remark 1.89 We will use the notation W_Q to denote a Q-Wiener process. If Q is trace-class, $\Xi_1=\Xi$ is a canonical choice and it will be understood that W_Q is a Ξ-valued process. If Q is not trace-class, writing W_Q and calling it a Q-Wiener process is a slight abuse of notation as it would be more precise to write W_{Q_1} and call it a Q_1-Wiener process with values in Ξ_1. However, even though the construction we have described is not canonical if $\mathrm{Tr}(Q)=+\infty$, and the choice of Ξ_1 is not unique, the class of the integrable processes is independent of the choice of Ξ_1 (see [180] Sect. 4.1 and in particular Proposition 4.7). Moreover (see [180] Sect. 4.1.2), for arbitrary $a\in\Xi$ the stochastic process

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

is a real-valued Wiener process and

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

For these reasons, even when $\mathrm{Tr}(Q)=+\infty$, we will still use the notation W_Q. When Q is the identity on Ξ we will call it a *cylindrical Wiener process in* Ξ. ■

Proposition 1.90 *Let Ξ be a real, separable Hilbert space, $Q\in\mathcal{L}^+(\Xi)$ and let Ξ_0, Ξ_1 and J be as described above. Let $(\Omega,\mathscr{F},\mathbb{P})$ be a complete probability space and $B\colon [t,+\infty)\times\Omega\to\Xi_1$ be a stochastic process. Denote by $\mathscr{F}_s^{t,0}$ the filtration generated by B, i.e.*

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

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

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

Proof The "only if" part follows from [180], Proposition 4.7, p. 85 (observe that in [180] a Wiener process is in fact defined using the four properties (1)–(4)). The "if" part is proved in [180] Proposition 4.3-(ii), p. 81 (if $\mathrm{Tr}(Q) = +\infty$ we apply the proposition in the space Ξ_1). □

The existence of a process satisfying conditions (1)–(4) above can also be proved using the Kolmogorov extension theorem (see [180], Proposition 4.4).

Remark 1.91 If $W_Q(s) = \sum_{k=1}^{\infty} g_k \beta_k(s)$ for some orthonormal basis $\{g_k\}_{k\in\mathbb{N}}$ of Ξ_0, it is easy to see that regardless of the choice of Ξ_1, $\mathscr{F}^{t,0}_s = \sigma(\beta_k(r) : t \le r \le s, k \in \mathbb{N})$. Thus the filtration generated by W_Q does not depend on the choice of Ξ_1. ■

Definition 1.92 (*Translated* $\mathscr{G}^t_s$*-*Q*-Wiener process*) Let $0 \le t < T \le +\infty$. Let Ξ be a real, separable Hilbert space, $Q \in \mathcal{L}^+(\Xi)$ and let Ξ_0, Ξ_1 and J be as described above. Let $\left(\Omega, \mathscr{F}, \mathscr{G}^t_s, \mathbb{P}\right)$ be a filtered probability space. We say that a stochastic process $B\colon [t, T] \times \Omega \to \Xi_1$ is a *translated* $\mathscr{G}^t_s$*-*Q*-Wiener process* on $[t, T]$ if:

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

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

We remark that if B is a translated $\mathscr{G}^t_s$-Q-Wiener process, then it is also a translated $\mathscr{F}^t_s$-Q-Wiener process, where $\mathscr{F}^t_s$ is the augmented filtration generated by B. Moreover, if W_Q is a Q-Wiener process as in Definition 1.88 then it is also a $\mathscr{F}^t_s$-Q-Wiener process, where $\mathscr{F}^t_s$ is the augmented filtration generated by B.

Lemma 1.93 *Let* $0 \le t < T \le +\infty$. *Let* Ξ *be a real, separable Hilbert space,* $Q \in \mathcal{L}^+(\Xi)$ *and let* Ξ_0 *and* Ξ_1 *be as described above. Let* $(\Omega, \mathscr{F}, \mathbb{P})$ *be a complete probability space. Let* $B\colon [t, T] \times \Omega \to \Xi_1$ *be a continuous stochastic process such that* $B(t) = 0$. *Then* B *is a* Q*-Wiener process on* $[t, T]$ *if and only if, for all* $a \in \Xi_1$, $t \le t_1 \le t_2 \le T$, *we have*

$$\mathbb{E}\left[e^{i\langle a, B(t_2) - B(t_1)\rangle_{\Xi_1}} | \mathscr{F}^t_{t_1}\right] = e^{-\frac{\langle Q_1 a, a\rangle_{\Xi_1}}{2}(t_2 - t_1)}. \tag{1.13}$$

Proof (The proof uses the same arguments as in the finite-dimensional case, see Proposition 1.2.7 of [452].)

The "only if" part: if B is a Q-Wiener process then, by Proposition 1.90-(4), Theorem 1.57 and Definition 1.58,

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

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

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

The "if" part: We have to prove the four conditions in Proposition 1.90: (1) and (2) are already in the assumptions of the lemma. Condition (4) follows easily from (1.13), Theorem 1.57 and Definition 1.58. To prove condition (3), i.e. that $Y := B(t_2) - B(t_1)$ is independent of $\mathscr{F}^t_{t_1}$, observe that, for all $Z\colon \Omega \to \Xi_1$ which are $\mathscr{F}^t_{t_1}$-measurable, one has, for all $a, b \in \Xi_1$,

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

Since the above holds for all $Z\colon \Omega \to \Xi_1$ which are $\mathscr{F}^t_{t_1}$-measurable, and for all $a, b \in \Xi_1$, we conclude that Y is independent of $\mathscr{F}^t_{t_1}$ by Theorem 1.56. □

Lemma 1.94 *Let $\mathscr{F}^{t,0}_s$ and $\mathscr{F}^t_s$ be the filtrations defined in Proposition 1.90 for a Q-Wiener process W_Q. Then $\mathscr{F}^t_s$ is right-continuous. Moreover, for all $T > t$, $\mathscr{F}^{t,0}_T$, and consequently $\mathscr{F}^t_T$, are countably generated up to sets of measure zero. If the trajectories of W_Q are everywhere continuous then*

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

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

Proof The proof follows arguments from [513] and [372] (Sect. 2.7-A). Consider $\tau > s$ and $\varepsilon > 0$. Since $W_Q(\tau + \varepsilon) - W_Q(s + \varepsilon)$ is independent of $\mathscr{F}^{t,0}_{s+}$, for every $A \in \mathscr{F}^{t,0}_{s+}$ and $f \in C_b(\Xi_1)$

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

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

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

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

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

and since the sets $\{\{W_Q(\tau) - W_Q(s) \in B\} : B = \overline{B} \subset \Xi_1\}$ are a π-system generating $\sigma(W_Q(\tau) - W_Q(s))$, it follows from Lemma 1.23 that $\mathscr{F}^{t,0}_{s+}$ and $\sigma(W_Q(\tau) - W_Q(s))$ are independent.

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

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

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

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

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

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

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

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

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

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

Moreover,

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

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

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

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

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

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

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

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

1.2.5 Simple and Elementary Processes

Definition 1.95 ($\mathscr{F}_s^t$*-simple process*) Let E be a Banach space (endowed with the Borel σ-field) and let $(\Omega, \mathscr{F}, \{\mathscr{F}_s^t\}_{s \in [t,T]}, \mathbb{P})$ be a filtered probability space. A process $X : [t, T] \times (\Omega, \mathscr{F}, \mathbb{P}) \to E$ is called $\mathscr{F}_s^t$*-simple* if:

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

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

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

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

Definition 1.96 ($\mathscr{F}^t_s$*-elementary process*) Let $T \in (0, +\infty)$, $t \in [0, T)$. Let (S, d) be a complete metric space (endowed with the Borel σ-field), and $(\Omega, \mathscr{F}, \left\{\mathscr{F}^t_s\right\}_{s\in[t,T]}, \mathbb{P})$ be a filtered probability space. We say that a process $X \colon [t, T] \times (\Omega, \mathscr{F}, \mathbb{P}) \to S$ is $\mathscr{F}^t_s$*-elementary* if there exist S-valued random variables $\xi_0, \xi_1, ..,$ ξ_{N-1}, and a sequence $t = t_0 < t_1 < .. < t_N = T$, such that

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

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

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

Remark 1.97 In Definitions 1.14, 1.95 and 1.96 we introduced the concepts of a *simple* random variable, $\mathscr{F}^t_s$*-simple* process, and $\mathscr{F}^t_s$*-elementary* process. The reader should be aware that in the literature the use of these terms varies and the same word is often used by different authors to mean different things. ■

Lemma 1.98 *Let E be a separable Banach space endowed with the Borel σ-field, $(\Omega, \mathscr{F}, \mathscr{F}^t_s, \mathbb{P})$ be a filtered probability space and $X \colon [t, T] \times \Omega \to E$ be a bounded, measurable, $\mathscr{F}^t_s$-adapted process, where $T \in [t, +\infty) \cup \{+\infty\}$. There exists a sequence $X^m(\cdot)$ of $\mathscr{F}^t_s$-elementary E-valued processes on $[t, T]$ such that, for every $1 \leq p < +\infty$ and $R > t$,*

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

The same claim holds if, instead of the Banach space, we consider E to be an interval $[a, b] \subset \mathbb{R}$ or a countable closed subset of $[a, b]$. In these cases the norm $|\cdot|_E$ in (1.16) is replaced by $|\cdot|_{\mathbb{R}}$.

Proof It is enough to prove the result for a single $p \geq 1$. To obtain a sequence of $\mathscr{F}_s^t$-simple processes $X^m(\cdot)$ with the required properties, the proof follows exactly the proof of Lemma 3.2.4, p. 132, in [372] with obvious technical modifications as we now have to deal with Bochner integrals in E. We then use Lemma 1.16 to approximate the random variables ξ_i defining $X^m(\cdot)$ by simple random variables to obtain $\mathscr{F}_s^t$-elementary approximating processes.

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

Lemma 1.99 *Let $\mathscr{F}_s^{t,0}$ and $\mathscr{F}_s^t$ be as in Proposition 1.90, $T \in [t, +\infty) \cup \{+\infty\}$, and let $a(\cdot) : [t, T] \times \Omega \to S$ be an $\mathscr{F}_s^t$-progressively measurable process, where (S, d) is a Polish space endowed with the Borel σ-field. Then there exists an $\mathscr{F}_s^{t,0}$-progressively measurable and $\mathscr{F}_s^{t,0}$-predictable process $a_1(\cdot) : [t, T] \times \Omega \to S$, such that $a(\cdot) = a_1(\cdot)$, $dt \otimes \mathbb{P}$-a.e. on $[t, T] \times \Omega$.*

Proof In light of Theorems 1.12 and 1.13 we can assume that $S = [0, 1]$ or S is a countable closed subset of $[0, 1]$. Using Lemma 1.98, we can find approximating $\mathscr{F}_s^t$-elementary processes $a^n(\cdot)$ on $[t, T]$ of the form

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

such that

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

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

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

Obviously the processes $a_1^n(\cdot)$ are $\mathscr{F}_s^{t,0}$-progressively measurable. We can now extract a subsequence (still denoted by $a_1^n(\cdot)$) such that $a_1^n(\cdot) \to a(\cdot)$ $dt \otimes \mathbb{P}$-

a.e. on $[t, T] \times \Omega$, and define $a_1(\cdot) := \liminf_{n \to +\infty} a_1^n(\cdot)$. The process $a_1(\cdot)$ is $\mathscr{F}_s^{t,0}$-progressively measurable, $\mathscr{F}_s^{t,0}$-predictable, and $a(\cdot) = a_1(\cdot)$, $dt \otimes \mathbb{P}$-a.e. on $[t, T] \times \Omega$. □

1.3 The Stochastic Integral

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

Definition 1.100 A 5-tuple $\mu := \left(\Omega, \mathscr{F}, \left\{\mathscr{F}_s^t\right\}_{s \in [t,T]}, \mathbb{P}, W_Q\right)$ described above is called a *generalized reference probability space*.

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

1.3.1 Definition of the Stochastic Integral

In this section we will assume that $\mathrm{Tr}(Q) < +\infty$. If $\mathrm{Tr}(Q) = +\infty$, the construction of the stochastic integral is the same, we just have to consider W_Q as a Ξ_1-valued Wiener process with nuclear covariance Q_1 (see Sect. 1.2.4). This way W_Q is not uniquely determined but $Q_1^{1/2}(\Xi_1) = \Xi_0 = Q^{1/2}(\Xi)$, $|x|_{\Xi_0} = |Q_1^{-1/2}x|_{\Xi_1}$ for all possible extensions Ξ_1 and the class of integrands and the value of the integrals are independent of the choice of the space Ξ_1 (see [180], Proposition 4.7 and Sect. 4.1.2).

We recall that we denote by $\mathcal{L}_2(\Xi_0, H)$ the space of Hilbert–Schmidt operators from Ξ_0 to H (see Appendix B.3). It is equipped with its Borel σ-field $\mathcal{B}(\mathcal{L}_2(\Xi_0, H))$. $\mathcal{L}_2(\Xi_0, H)$ is a real, separable Hilbert space (see Proposition B.25), and $\mathcal{L}(\Xi, H)$ is dense in $\mathcal{L}_2(\Xi_0, H)$ (see e.g. [294], pp. 24–25).

Definition 1.101 (*The space* $\mathcal{N}_Q^p(t, T; H)$) Given $p \geq 1$, we denote by $\mathcal{N}_Q^p(t, T; H)$ the space of all $\mathcal{L}_2(\Xi_0, H)$-valued, progressively measurable processes $X(\cdot)$, such that

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

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

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

Remark 1.102 In several classical references (see e.g. [180] or [491]), the theory of stochastic integration is developed for predictable processes instead of progressively measurable ones like in our case. However, it follows for instance from Lemma 1.99, that for every $\mathcal{L}_2(\Xi_0, H)$-valued progressively measurable process X there exists a predictable process X_1 which is $\mathbb{P} \otimes dt$-a.e. equal to X. Thus, since we are working with stochastic integrals with respect to Wiener processes (which are continuous), the two concepts coincide. ■

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

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

Note that if we take Φ to be $\mathcal{L}_2(\Xi_0, H)$-valued, we cannot guarantee that the expression above is well defined, since $\mathcal{L}_2(\Xi_0, H)$ contains genuinely unbounded operators in Ξ (see e.g. [294], p. 25, Exercise 2.7).

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

Theorem 1.103 (Itô isometry) *For every $\mathcal{L}(\Xi, H)$-valued, $\mathscr{F}_s^t$-simple process Φ we have*

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

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

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

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

Proof See [294], Propositions 2.1, 2.2, and Definition 2.10. See also [180], Proposition 4.22 in the context of predictable processes. □

Proposition 1.104 *For $\Phi \in \mathcal{N}_Q^2(t, T; H)$, consider the process*

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

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

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

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

Proof See [294] Theorem 2.3, p. 34. □

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

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

Lemma 1.105 *Let $\{\Phi(s)\}_{s\in[t,T]}$ be an $\mathcal{L}_2(\Xi_0, H)$-valued progressively measurable process satisfying (1.17). Then there exists a sequence Φ_n of $\mathcal{L}(\Xi, H)$-valued $\mathscr{F}_s^t$-simple processes such that*

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

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

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

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

Proof See [294], Lemmas 2.3, p. 39, and 2.6, p. 41. □

The process $\mathcal{I}$ defined by Lemma 1.105 is called the stochastic integral of Φ with respect to W_Q, and is denoted by $\int_t^T \Phi(s) dW_Q(s)$. We also set $\int_t^r \Phi(s) dW_Q(s) := \int_t^T \Phi(s)\mathbf{1}_{[t,r]} dW_Q(s)$.

Proposition 1.106 *Let $\{\Phi(s)\}_{s\in[t,T]}$ be an $\mathcal{L}_2(\Xi_0, H)$-valued progressively measurable process satisfying (1.17). Then the process*

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

is a continuous local martingale.

Proof See [294], pp. 42–44. □

Finally, we may extend the definition of stochastic integral to all processes (not necessarily progressively measurable) that are $dt \otimes \mathbb{P}$-equivalent to progressively measurable processes satisfying (1.17) in the sense of the following definition (see also [372], p. 130).

Definition 1.107 We say that two processes Φ_1 and Φ_2 are $dt \otimes \mathbb{P}$-equivalent if $\Phi_1 = \Phi_2$, $dt \otimes \mathbb{P}$-a.e. If Φ belongs to the equivalence class of a progressively measurable process Φ_1 satisfying (1.17),[4] we set

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

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

Example 1.108 Every $\mathcal{L}_2(\Xi_0, H)$-valued, $\mathscr{F}_s^t$-adapted, and $\overline{\mathcal{B}([t,T]) \otimes \mathscr{F}}$-measurable process Φ satisfying (1.17) is stochastically integrable, where $\overline{\mathcal{B}([t,T]) \otimes \mathscr{F}}$ is the completion of $\mathcal{B}([t,T]) \otimes \mathscr{F}$ with respect to $dt \otimes \mathbb{P}$. To see this we need to find a progressively measurable process Φ_1 which is equivalent to Φ. First, let Φ_2 be a $\mathcal{B}([t,T]) \otimes \mathscr{F}$-measurable process equivalent to Φ (which exists by Lemma 1.16). Then, for a.e. $s \in [t,T]$, we have $\Phi_2(s,\cdot) = \Phi(s,\cdot)$ $\mathbb{P}$-a.s. and, since every $\mathscr{F}_s^t$ is complete, also $\Phi_2(s,\cdot)$ is $\mathscr{F}_s^t$-measurable for a.e. s. Thus there exists an $A \in \mathcal{B}([t,T])$ of full measure such that $\Phi_2(s,\cdot)$ is $\mathscr{F}_s^t$-measurable for $s \in A$. We then define $\Phi_3 = \Phi_2 \mathbf{1}_A$. Φ_3 is $\mathcal{B}([t,T]) \otimes \mathscr{F}$-measurable and $\mathscr{F}_s^t$-adapted, thanks to Lemma 1.72 it has a progressively measurable modification Φ_1 which is clearly equivalent to Φ. ■

Theorem 1.109 *Let $(E, \mathscr{G}, \mu)$ be a measure space with bounded measure. Let $\Phi : [t,T] \times \Omega \times E \to \mathcal{L}_2(\Xi_0, H)$ be $(\mathcal{B}([t,T]) \otimes \mathscr{F}_T^t \otimes \mathscr{G})/\mathcal{B}(\mathcal{L}_2(\Xi_0, H))$-measurable. Suppose that, for any $x \in E$, $\{\Phi(s,\cdot,x)\}_{s\in[t,T]}$ is progressively measurable and*

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

Then:

(i) $\int_t^T \Phi(s,\cdot,\cdot) dW_Q(s)$ *has an $\mathscr{F}_T^t \otimes \mathscr{G}/\mathcal{B}(H)$-measurable version.*
(ii) $\int_E \Phi(\cdot,\cdot,x) d\mu(x)$ *is progressively measurable.*
(iii) The following equality holds $\mathbb{P}$-a.s.:

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

[4] Note that if a process X is progressively measurable and satisfies (1.17) and Y is $dt \otimes \mathbb{P}$-equivalent to X, then Y must also satisfy (1.17) since for $\mathbb{P}$-a.s. ω, $X(\cdot,\omega) = Y(\cdot,\omega)$, a.e. on [t,T].

Proof See Theorem 2.8, Sect. 2.2.6, p. 57 of [294] and Theorem 4.33, Sect. 4.5, p. 110 of [180]. □

1.3.2 Basic Properties and Estimates

Lemma 1.110 *Let $T > 0$ and $t \in [0, T)$. Assume that Φ is in $\mathcal{N}_Q^2(t, T; H)$ and that τ is an $\mathscr{F}_s^t$-stopping time such that $\mathbb{P}(\tau \le T) = 1$. Then $\mathbb{P}$-a.s.*

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

Proof See [294], Lemma 2.7, p. 43 (also [180], Lemma 4.24, p. 99). □

As a consequence of Theorem 1.80 and Proposition 1.104 we obtain the following theorem (see also e.g. [177], Theorem 5.2.4, p. 58).

Theorem 1.111 (Burkholder–Davis–Gundy inequality for stochastic integrals) *Let $T > 0$ and $t \in [0, T)$. For every $p \ge 2$, there exists a constant c_p such that, for every Φ in $\mathcal{N}_Q^p(t, T; H)$,*

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

Proposition 1.112 *Let $T > 0$ and $t \in [0, T)$. Let A be the generator of a C_0-semigroup $\{e^{rA},\ r \ge 0\}$ on H such that $\|e^{rA}\| \le Me^{\alpha r}$ for every $r \ge 0$ for some $\alpha \in \mathbb{R}$, $M > 0$. Let $p > 2$ and $\Phi \in \mathcal{N}_Q^p(t, T; H)$. Let A_n be the Yosida approximation of A. Then the stochastic convolution process*

$$\Psi(s) := \int_t^s e^{(s-r)A}\Phi(r)dW_Q(r), \qquad s \in [t, T], \tag{1.19}$$

has a continuous modification,

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

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

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

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

Proof See [294], Lemma 3.3, p. 87. The claims for p=2 can be proved by repeating the arguments of the proof of Proposition 3.3 of [543], which uses the Unitary Dilation Theorem. □

Proposition 1.113 *Let A be the generator of a C_0-semigroup on H, $T > 0$, and $t \in [0, T)$. Assume that $\Phi : [t, T] \times \Omega \to \mathcal{L}_2(\Xi_0, H)$ is a progressively measurable process such that $\Phi(s) \in \mathcal{L}_2(\Xi_0, D(A))$ $\mathbb{P}$-a.s., for a.e. $s \in [t, T]$. Assume that*

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

Then

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

and

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

Proof We can assume without loss of generality that $Q \in \mathcal{L}_1^+(\Xi)$. The proof follows the proof of Proposition 3.1 (p. 76) of [294], however we present it here to clarify a measurability issue. Indeed, we first need to show that Φ is an $\mathcal{L}_2(\Xi_0, D(A))$-valued, progressively measurable process. To do this we take $\Psi_n = J_n\Phi$, where $J_n = n(nI - A)^{-1}$ (see Definition B.40). Since $J_n \in \mathcal{L}(H, D(A))$, Ψ_n is an $\mathcal{L}_2(\Xi_0, D(A))$-valued, progressively measurable process. Moreover, it is easy to see that if, for some $s \in [t, T]$ and $\omega \in \Omega$, $\Phi(s)(\omega) \in \mathcal{L}_2(\Xi_0, D(A))$, then $\Psi_n(s)(\omega) \to \Phi(s)(\omega)$ in $\mathcal{L}_2(\Xi_0, D(A))$. Therefore, defining $V := \{(s, \omega) : \Psi_n(s)(\omega)$ converges in $\mathcal{L}_2(\Xi_0, D(A))\}$, it follows from Lemma 1.8-(iii) that Φ is equivalent to a progressively measurable process $\lim_{n\to+\infty} \mathbf{1}_V \Psi_n$. The proof is now done in two steps.
Step 1: The claim is true for $\mathscr{F}_s^t$-simple $\mathcal{L}(\Xi, D(A))$-valued processes.
Step 2: If Φ is a $\mathcal{L}_2(\Xi_0, D(A))$-valued progressively measurable process satisfying the hypotheses of this proposition, we take a sequence of $\mathscr{F}_s^t$-simple $\mathcal{L}(\Xi, D(A))$-valued processes Φ_n approximating Φ in the sense of (1.18) so that

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

In particular we have

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

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

in probability, so the claim follows since A is a closed operator. □

In the rest of this section we explain how the factorization method is used to prove continuity of trajectories of stochastic convolution processes.

Lemma 1.114 (Factorization Lemma) *Let $T > 0$, $t \in [0, T)$, and $0 < \alpha < 1$. Let A be the generator of a C_0-semigroup $\{e^{rA}, r \geq 0\}$ on H. Consider a linear, densely defined, closed operator $A_1 : D(A_1) \subset H \to H$ such that, for any $r > 0$, $e^{rA} H \subset D(A_1)$, $A_1 e^{rA}$ is bounded and $A_1 e^{rA} = e^{rA} A_1$ on $D(A_1)$. Let $\Phi : [t, T] \times \Omega \to \mathcal{L}_2(\Xi_0, H)$ be progressively measurable and such that for every $s \in [t, T]$*

$$\mathbb{E} \int_t^s \left\| A_1 e^{(s-r)A} \Phi(r) \right\|^2_{\mathcal{L}_2(\Xi_0, H)} dr < +\infty.$$

Assume that, for all $s \in [t, T]$,

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

Then

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

for all $s \in [t, T]$, where $Y_\alpha^\Phi(\cdot)$ is a $\mathcal{B}([t, T]) \otimes \mathscr{F}_T^t / \mathcal{B}(H)$-measurable process which is $dt \otimes \mathbb{P}$-equivalent to

$$\int_t^r (r-h)^{-\alpha} A_1 e^{(r-h)A} \Phi(h) dW_Q(h).$$

Proof The statement is similar to [177], Theorem 5.2.5, p. 58, Sect. 5.2.1. We give the proof for completeness.

We use the identity

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

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

$$X(r, h) = \mathbf{1}_{[t,r]}(h)(r-h)^{-\alpha} A_1 e^{(r-h)A} \Phi(h).$$

Since (1.24) implies

$$\int_t^T \left(\mathbb{E} \int_t^T \|X(r,h)\|^2_{\mathcal{L}_2(\Xi_0,H)}\, dh \right)^{1/2} dr < +\infty,$$

by the stochastic Fubini Theorem 1.109 (see also Theorem 4.33, p. 110 of [180] or Theorem 2.8, p. 57 of [294]) there exists a $\mathcal{B}([t,T]) \otimes \mathscr{F}^t_T/\mathcal{B}(H)$-measurable process $Y^\Phi_\alpha \colon [t,T] \times \Omega \to H$ such that

$$\int_t^T X(r,h) dW_Q(h) = \int_t^r (r-h)^{-\alpha} A_1 e^{(r-h)A} \Phi(h) dW_Q(h) = Y^\Phi_\alpha(r), \quad dt \otimes \mathbb{P}\text{-a.e.}$$

Then for every $s \in [t,T]$ the process $Z^{\Phi,s}_\alpha(\cdot)$, defined for any $r \in [t,s]$ by $Z^{\Phi,s}_\alpha(r) = (s-r)^{\alpha-1} e^{(s-r)A} Y^\Phi_\alpha(r)$, is jointly measurable and $dt \otimes \mathbb{P}$-equivalent to

$$(s-r)^{\alpha-1} e^{(s-r)A} \int_t^r (r-h)^{-\alpha} A_1 e^{(r-h)A} \Phi(h) dW_Q(h)$$

on $[t,s] \times \Omega$. Thus fixing any $s \in [t,T]$ and applying the stochastic Fubini Theorem on $[t,s] \times [t,s] \times \Omega$ (whose assumptions are satisfied by (1.24)) and noticing that we can use the process $Z^{\Phi,s}_\alpha(\cdot)$ in place of a process provided by the stochastic Fubini Theorem (since it will give $\mathbb{P}$-a.e. the same integrals) we obtain for $\mathbb{P}$-a.e. ω

$$\begin{aligned} &\frac{\pi}{\sin(\pi\alpha)} \int_t^s A_1 e^{(s-h)A} \Phi(h) dW_Q(h) \\ &\quad = \int_t^s \int_t^s \mathbf{1}_{[h,s]}(r)(s-r)^{\alpha-1} e^{(s-r)A} (r-h)^{-\alpha} A_1 e^{(r-h)A} \Phi(h) dr dW_Q(h) \\ &\quad = \int_t^s (s-r)^{\alpha-1} e^{(s-r)A} Y^\Phi_\alpha(r) dr. \end{aligned}$$

□

Lemma 1.115 *Let A be the generator of a C_0-semigroup $\{e^{rA},\ r \geq 0\}$ on H, $T > 0$, $t \in [0,T)$ and $f \in L^p(t,T;H)$, $p \geq 1$. Then:*

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

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

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

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

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

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

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

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

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

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

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

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

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

Part (ii) follows from Proposition A.1.1 in Appendix A, p. 307 of [177]. □

Proposition 1.116 *Let* $T > 0$ *and* $t \in [0, T)$. *Let* A, A_1, Φ *satisfy the assumptions of Lemma 1.114 except (1.24). Assume that there exist* $0 < \alpha < 1$, $C > 0$ *and* $p > \frac{1}{\alpha}$, $p \ge 2$ *such that*

$$\int_t^T \mathbb{E} \left(\int_t^r \| (r - h)^{-\alpha} A_1 e^{(r-h)A} \Phi(h) \|^2_{\mathcal{L}_2(\Xi_0, H)} dh \right)^{p/2} dr < C. \tag{1.25}$$

Then

$$\Psi(s) := \int_t^s A_1 e^{(s-r)A} \Phi(r) dW_Q(r), \qquad s \in [t, T],$$

has a continuous modification.

Proof We follow the scheme of the proof of Theorem 5.2.6 in [177] (p. 59, Sect. 5.2.1). We give some details because our claim is slightly more general.

Observe that using Hölder's and Jensen's inequalities we obtain

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

where we used (1.25) and that $\frac{(1-\alpha)p}{p-1} < 1$, which follows from $p > 1/\alpha$. Therefore the hypotheses of Lemma 1.114 are satisfied and thus we have

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

for all $s \in [t,T]$, where $Y_\alpha^\Phi(\cdot)$ is defined in Lemma 1.114. The claim will follow from Lemma 1.115-(i) applied to a.e. trajectory. Thus we need to know that the process $Y_\alpha^\Phi(\cdot)$ has p-integrable trajectories a.s. This is guaranteed if

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

However, from Theorem 1.111, we have

$$\int_t^T \mathbb{E}\left(\left[\left|Y_\alpha^\Phi(s)\right|^p\right]\right) ds \leq c_p \int_t^T \mathbb{E}\left(\int_t^s \|(s-r)^{-\alpha}A_1 e^{(s-r)A}\Phi(r)\|^2_{\mathcal{L}_2(\Xi_0,H)}dr\right)^{p/2} ds, \tag{1.26}$$

which is bounded thanks to (1.25). □

The factorization method can also be used to show the continuity of deterministic convolution integrals. The following lemma deals with a case which arises in Sects. 1.5.2 and 1.5.3.

Lemma 1.117 *Let $T > 0$, $t \in [0,T)$, and $0 < \alpha < 1$. Let A be the generator of a C_0-semigroup $\{e^{rA},\ r \geq 0\}$ on H. Let ϕ be a function defined on $[t,T]$ such that, for every $s \in (0, T-t]$, $e^{sA}\phi : [t,T] \to H$ is well defined, measurable and*

$$|e^{sA}\phi(r)| \leq s^{-\beta}g(r) \quad \textit{for } r \in [t,T], \tag{1.27}$$

where $0 \leq \beta < 1$, $g \in L^q(t,T;H)$, $q > \frac{1}{1-\beta}$. Then the function

$$\psi(s) = \int_t^s e^{(s-r)A}\phi(r)dr$$

belongs to $C([t,T],H)$.

Proof Let $0 < \alpha$ be such that $\alpha+\beta < 1$ and $q > \frac{1}{1-(\alpha+\beta)}$. We have, by the Fubini Theorem 1.33,

$$\int_t^s e^{(s-r)A}\phi(r)dr = \frac{\sin(\pi\alpha)}{\pi}\int_t^s (s-r)^{\alpha-1}e^{(s-r)A}Y(r)dr,$$

where

$$Y(r) = \int_t^r (r-h)^{-\alpha}e^{(r-h)A}\phi(h)dh.$$

It remains to notice that, using (1.27) and Hölder's inequality, we have for $t \le r \le T$

$$|Y(r)| \le \int_t^r (r-h)^{-(\alpha+\beta)}g(h)dh \le C_T|g|_{L^q(t,T;H)}.$$

Thus the result follows from Lemma 1.115-(i). □

1.4 Stochastic Differential Equations

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

1.4.1 Mild and Strong Solutions

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

$$\begin{cases} dX(s) = (AX(s) + b(s, X(s)))ds + \sigma(s, X(s))dW_Q(s) & s \in (0, T] \\ X(0) = \xi, \end{cases} \tag{1.28}$$

where ξ is an H-valued $\mathscr{F}_0$-measurable random variable. To simplify the notation we dropped the ω variable in (1.28) and we use this convention throughout the section.

Definition 1.118 (*Strong solution of* (1.28)) An H-valued progressively measurable process $X(\cdot)$ is called a *strong solution* of (1.28) if:

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

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

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

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

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

Definition 1.119 (*Mild solution of* (1.28)) An H-valued progressively measurable process $X(\cdot)$ is called a *mild solution* of (1.28) if:

(i) For every $t \in [0, T]$

$$\mathbb{P}\left(\int_0^t \left(|X(s)| + |e^{(t-s)A}b(s, X(s))|\right) ds < +\infty\right) = 1$$

and

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

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

$$X(t) = e^{tA}\xi + \int_0^t e^{(t-s)A}b(s, X(s))ds + \int_0^t e^{(t-s)A}\sigma(s, X(s))dW_Q(s) \quad \mathbb{P}\text{-a.e.}$$

In order for the above definitions to be meaningful, all the processes involved must be well defined and have proper measurability properties so that the integrals that appear in the definitions make sense. We do not want to analyze here the required measurability properties in the most generality. Instead, we discuss one case which will frequently appear in applications to optimal control in Remark 1.123 below. Moreover, note that if A_n is the Yosida approximation of A, since by Lemma 1.17-(i) $D(A) \in \mathcal{B}(H)$, it follows that the processes $\mathbf{1}_{X(\cdot)\in D(A)} A_n X(\cdot)$ are progressively measurable and they converge as $n \to +\infty$ to $\mathbf{1}_{X(\cdot)\in D(A)} AX(\cdot)$ for every (s, ω). Thus the process $AX(\cdot)$ (understood as $\mathbf{1}_{X(\cdot)\in D(A)} AX(\cdot)$) is progressively measurable.

Remark 1.120 In the definition of a mild solution we assumed that $b\colon [0, T] \times H \times \Omega \to H$ and $\sigma\colon [0, T] \times H \times \Omega \to \mathcal{L}_2(\Xi_0, H)$. However, Definition 1.119 may still make sense even if b and σ do not have values in H and $\mathcal{L}_2(\Xi_0, H)$, provided that the terms $e^{(t-s)A}b(s, X(s))$ and $e^{(t-s)A}\sigma(s, X(s))$ have values in these spaces when they are interpreted properly (see, for instance, Sect. 1.5.1 and also Remark 1.123). Therefore in the future when we are dealing with such cases, we will not repeat the definition of a mild solution, instead we will just explain how to interpret the above terms. ■

Definition 1.121 (*Weak mild solution of* (1.28)) Assume that in (1.28) we have $b : [0, T] \times H \to H$ and $\sigma : [0, T] \times H \to \mathcal{L}_2(\Xi_0, H)$. A *weak mild solution* of (1.28) is defined to be any 6-tuple $(\Omega, \mathscr{F}, \mathscr{F}_s, W_Q, \mathbb{P}, X(\cdot))$, where $(\Omega, \mathscr{F}, \mathscr{F}_s, \mathbb{P})$ is a filtered probability space, W_Q is a translated $\mathscr{F}_s$-Q-Wiener process on Ω, and $X(\cdot)$ is a mild solution for (1.28) in the generalized reference probability space $(\Omega, \mathscr{F}, \mathscr{F}_s, W_Q, \mathbb{P})$.

Notation 1.122 In the existing literature, different authors often give different names to the same notion of solution, and the same name does not always correspond to the same definition. For instance, the *weak mild solution* introduced above is often called a weak solution and in [180] Chap. 8 it is called a *martingale solution*. ■

Remark 1.123 Let Λ be a Polish space. Suppose that $\sigma : [0, T] \times H \times \Lambda \to \mathcal{L}(\Xi_0, H)$ is such that for every $u \in \Xi_0$, the map $(t, x, a) \to \sigma(t, x, a)u$ is $\mathcal{B}([0, T]) \otimes \mathcal{B}(H) \otimes \mathcal{B}(\Lambda)/\mathcal{B}(H)$-measurable, and $e^{sA}\sigma(t, x, a) \in \mathcal{L}_2(\Xi_0, H)$ for every (t, x, a) and $s > 0$. It then follows from Lemma 1.20 that, after possibly redefining it at $s = 0$, the map $(s, t, x, a) \to e^{sA}\sigma(t, x, a)$ is $\mathcal{B}([0, T]) \otimes \mathcal{B}([0, T]) \otimes \mathcal{B}(H) \otimes \mathcal{B}(\Lambda)/\mathcal{B}(\mathcal{L}_2(\Xi_0, H))$-measurable. Now, if $X(\cdot) : [0, T] \times \Omega \to H$, $a(\cdot) : [0, T] \times \Omega \to \Lambda$ are $\mathscr{F}_s$-progressively measurable, then for every $t \in [0, T]$,

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

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

$$Z(t) = \int_0^t e^{(t-s)A}\sigma(s, X(s), a(s))dW_Q(s), \qquad t \in [0, T]$$

is an H-valued $\mathscr{F}_t$-adapted process. One way to argue that $Z(\cdot)$ has a progressively measurable modification is the following.

Suppose that there is a constant $K \geq 0$ such that

$$\mathbb{E}|Z(t)| \leq K \quad \text{for all } t \in [0, T]$$

and that for all $0 \leq t \leq h \leq T$

$$\mathbb{E}\int_t^h \left\| e^{(h-s)A}\sigma(s, X(s), a(s)) \right\|^2_{\mathcal{L}_2(\Xi_0, H)} ds \leq \rho(h - t)$$

for some modulus ρ. We have for $0 \leq t \leq h \leq T$

$$Z(h) - Z(t) = \left(e^{(h-t)A} - I\right) Z(t) + \int_t^h e^{(h-s)A}\sigma(s, X(s), a(s))dW_Q(s).$$

Let $\{e_n\}$ be an orthonormal basis of H. Then

$$\langle Z(h) - Z(t), e_n\rangle = \left\langle Z(t), e^{(h-t)A}e_n - e_n\right\rangle + \left\langle \int_t^h e^{(h-s)A}\sigma(s, X(s), a(s))dW_Q(s), e_n\right\rangle$$

and hence

$$\mathbb{E}\,|\langle Z(h) - Z(t), e_n\rangle| \leq K|e^{(h-t)A^*}e_n - e_n| + \sqrt{\rho(h - t)} \leq \rho_n(h - t)$$

for some modulus ρ_n. Therefore it is easy to see that the process $\langle Z(t), e_n\rangle$ is stochastically continuous and thus, by Lemma 1.74, it has a progressively measurable modification which we denote by $Z_n(\cdot)$. The process $\tilde{Z}(\cdot)$ defined, for $t \in [0, T]$, by

$$\tilde{Z}(t) = \begin{cases} \sum_{n=1}^{+\infty} Z_n(t)e_n & \text{if the limit exists,} \\ 0 & \text{otherwise} \end{cases}$$

is a progressively measurable modification of $Z(\cdot)$. ■

1.4.2 Existence and Uniqueness of Solutions

Definition 1.124 (*The space* $M^p_\mu(t, T; E)$) In this definition $T \in (0, +\infty) \cup \{+\infty\}$. Let $p \geq 1$ and $0 \leq t < T$. Given a Banach space E, we denote by $M^p_\mu(t, T; E)$ the space of all E-valued progressively measurable processes $X(\cdot)$ such that

$$|X(\cdot)|_{M^p_\mu(t,T;E)} := \left(\mathbb{E}\left(\int_t^T |X(s)|^p ds\right)\right)^{1/p} < +\infty. \tag{1.29}$$

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

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

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

$$\begin{cases} dX(s) = (AX(s) + b(s, X(s), a(s)))\, ds + \sigma(s, X(s), a(s))dW_Q(s) \\ X(0) = \xi. \end{cases} \tag{1.30}$$

This equation falls into the category of equations (1.28) with $b(s, x, \omega) := b(s, x, a(s, \omega))$ and $\sigma(s, x, \omega) := \sigma(s, x, a(s, \omega))$. Thus strong, mild and weak mild solutions of (1.30) are defined using the definitions for Eq. (1.28).

Hypothesis 1.125 The operator A is the generator of a strongly continuous semigroup e^{sA} on H. The function $b \colon [0, T] \times H \times \Lambda \to H$ is $\mathcal{B}([0, T]) \otimes \mathcal{B}(H) \otimes \mathcal{B}(\Lambda)/\mathcal{B}(H)$-measurable, $\sigma \colon [0, T] \times H \times \Lambda \to \mathcal{L}_2(\Xi_0, H)$ is $\mathcal{B}([0, T]) \otimes \mathcal{B}(H) \otimes \mathcal{B}(\Lambda)/\mathcal{B}(\mathcal{L}_2(\Xi_0, H))$-measurable, and there exists a constant $C > 0$ such that

$$|b(s,x,a) - b(s,y,a)| \le C|x-y| \qquad \forall x,y \in H, s \in [0,T], a \in \Lambda, \tag{1.31}$$

$$\|\sigma(s,x,a) - \sigma(s,y,a)\|_{\mathcal{L}_2(\Xi_0,H)} \le C|x-y| \quad \forall x,y \in H, s \in [0,T], a \in \Lambda, \tag{1.32}$$

$$|b(s,x,a)| \le C(1+|x|) \qquad \forall x \in H, s \in [0,T], a \in \Lambda, \tag{1.33}$$

$$\|\sigma(s,x,a)\|_{\mathcal{L}_2(\Xi_0,H)} \le C(1+|x|) \qquad \forall x \in H, s \in [0,T], a \in \Lambda. \tag{1.34}$$

Definition 1.126 (*The space* $\mathcal{H}_p^\mu(t,T;E)$) Let $p \ge 1$ and $0 \le t < T$. Given a Banach space E, we denote by $\mathcal{H}_p^\mu(t,T;E)$ the set of all progressively measurable processes $X\colon [t,T] \times \Omega \to E$ such that

$$|X(\cdot)|_{\mathcal{H}_p^\mu(t,T;E)} := \left(\sup_{s \in [t,T]} \mathbb{E}|X(s)|^p \right)^{1/p} < +\infty. \tag{1.35}$$

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

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

Mild solutions in $\mathcal{H}_p^\mu(0,T;E)$ (or $M_\mu^p(0,T;E)$) of various versions of (1.30) will be obtained as fixed points in these spaces of some maps. We point out that this will not imply that every representative of the equivalence class is a mild solution. Since a mild solution $X(\cdot)$ satisfies the integral equality in Definition 1.119-(ii) for every $t \in [0,T]$, $X(t)$ is prescribed by the right-hand side of this equality, which does not depend on the choice of a representative of the equivalence class. Thus there is a unique (up to a modification) representative of the equivalence class which is a mild solution. We will then always be able to evaluate $\mathbb{E}|X(t)|^p$ for the mild solution $X(\cdot)$ for every $t \in [0,T]$ (and in fact compute the $\mathcal{H}_p^\mu(0,T;E)$ norm of this representative by taking the *sup* over all $t \in [0,T]$ instead of the *esssup*).

Theorem 1.127 *Let $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ for some $p \ge 2$, and let A, b and σ satisfy Hypothesis 1.125. Let $a(\cdot)\colon [0,T] \to \Lambda$ be an $\mathscr{F}_s$-progressively measurable process. Then the SDE (1.30) has a unique, up to a modification, mild solution $X(\cdot) \in \mathcal{H}_p(0,T;H)$. The solution is in fact unique among all processes such that $\mathbb{P}\left(\int_0^T |X(s)|^2 ds < +\infty\right) = 1$, in particular among the processes in $M_\mu^2(0,T;H)$. $X(\cdot)$ has a continuous modification. Given two continuous versions $X_1(\cdot)$, $X_2(\cdot)$ of the solution, there exists a $\tilde{\Omega} \subset \Omega$ with $\mathbb{P}(\tilde{\Omega}) = 1$ s.t. $X_1(s) = X_2(s)$ for all $s \in [0,T]$ and $\omega \in \tilde{\Omega}$, i.e. they are indistinguishable.*

Proof The proof can be found, for instance, in [180], Theorem 7.2, p. 188 or [294], Theorems 3.3, p. 97, and 3.5, p. 105. For the last claim, we can take

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

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

We will denote the solution of (1.30) by $X(\cdot; \xi, a(\cdot))$ if we want to emphasize the dependence on the initial datum and the control.

Corollary 1.128 *Let $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ for some $p \geq 2$, let A, b and σ satisfy Hypothesis 1.125. If $a_1(\cdot), a_2(\cdot)\colon [0, T] \times \Omega \to \Lambda$ are two progressively measurable processes such that $a_1(\cdot) = a_2(\cdot)$, $dt \otimes \mathbb{P}$-a.e. on $[0, T] \times \Omega$, then, $\mathbb{P} - a.e.$,*

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

Proof Define $X_i(\cdot) := X\,(\cdot; \xi, a_i(\cdot))$. Using Theorem 1.103, Jensen's inequality, and $\sup_{s\in[0,T]} \|e^{sA}\| \leq C$ for some $C \geq 0$, it follows that, for suitable positive C_1 and C_2:

$$\begin{aligned}\mathbb{E}\left[|X_1(s) - X_2(s)|^2\right] \leq C_1\Bigg(&\int_0^s \mathbb{E}|b(r, X_1(r), a_1(r)) - b(r, X_2(r), a_2(r))|^2 dr\\ &+ \int_0^s \mathbb{E}\|\sigma(r, X_1(r), a_1(r)) - \sigma(r, X_2(r), a_2(r))\|^2_{\mathcal{L}_2(\Xi_0,H)} dr\Bigg)\\ &\leq C_2 \int_0^s \mathbb{E}|X_1(r) - X_2(r)|^2 dr, \qquad s \in [0, T],\end{aligned}$$

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

Remark 1.129 Above we assumed that the σ always takes values in $\mathcal{L}_2(\Xi_0, H)$. Existence and uniqueness results for SDEs with more general σ can be found, for instance, in [294] Theorem 3.15, p. 143, or in [180] Theorem 7.5, p. 197. To treat some specific examples we will also prove more general results in Sect. 1.5. ■

1.4.3 Properties of Solutions

Theorem 1.130 *Let $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ for some $p \geq 2$, $a\colon [0, T] \times \Omega \to \Lambda$ be $\mathscr{F}_s$-progressively measurable, and let A, b and σ satisfy Hypothesis 1.125.*

(i) Let $X(\cdot) = X(\cdot; \xi, a(\cdot))$ be the unique mild solution of (1.30) (provided by Theorem 1.127). Then, for any $s \in [0, T]$,

$$\sup_{s\in[0,T]} \mathbb{E}\left[|X(s)|^p\right] \le C_p(T)(1+\mathbb{E}|\xi|^p) \quad \textit{if } p \ge 2, \tag{1.36}$$

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

and

$$\mathbb{E}\left[\sup_{r\in[0,s]} |X(r)-\xi|^p\right] \le \omega_\xi(s) \quad \textit{if } p > 2, \tag{1.38}$$

where $C_p(T)$ is a constant depending on p, T, C (from Hypothesis 1.125) and M, α (where $\|e^{rA}\| \le Me^{r\alpha}$ for $r \ge 0$), and ω_ξ is a modulus depending on the same constants and on ξ (in particular they are independent of the process $a(\cdot)$ and of the generalized reference probability space).

(ii) If $\xi, \eta \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ for $p > 2$, and $X(\cdot) = X(\cdot;\xi, a(\cdot))$, $Y(\cdot) = Y(\cdot;\eta, a(\cdot))$ are the solutions of (1.30), then, for all $s \in [0,T]$,

$$\mathbb{E}\left[\sup_{s\in[0,T]} |X(s)-Y(s)|^2\right] \le C_T\left(\mathbb{E}\left[|\xi-\eta|^p\right]\right)^{\frac{2}{p}}, \tag{1.39}$$

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

Proof Part (i): For (1.36) and (1.37) we refer, for instance, to [180] Theorem 9.1, p. 235, or [294], Lemma 3.6, p. 102, and Corollary 3.3, p. 104. Regarding (1.38), we have that there is a constant c_1 depending only on p and $\sup_{t\in[0,T]}\|e^{tA}\|$, such that

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

Using Hypothesis 1.125, (1.37), Hölder's inequality, and Proposition 1.112, we see that

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

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

Part (ii): See [180] Theorem 9.1, p. 235. □

Theorem 1.131 *Let $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ for some $p > 2$, and let A, b and σ satisfy Hypothesis 1.125. Let $a\colon [0, T] \times \Omega \to \Lambda$ be a progressively measurable process. Let $X(\cdot)$ be the unique mild solution of (1.30). Consider the approximating equations*

$$\begin{cases} dX^n(s) = (A_n X^n(s) + b(s, X^n(s), a(s)))\, ds + \sigma(s, X^n(s), a(s)) dW_Q(s) \\ X^n(0) = \xi, \end{cases} \tag{1.40}$$

where A_n is the Yosida approximation of A. Let $X_n(\cdot)$ be the solution of (1.40). Then

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{s\in[0,T]} |X^n(s) - X(s)|^p \right] = 0. \tag{1.41}$$

Proof See [180] Proposition 7.4, p. 196, or [294], Proposition 3.2, p. 101. □

The next proposition is a simpler version of Theorem 1.131 which will be useful in the proofs of the results of Sect. 1.7.

Proposition 1.132 *Let $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$, $f \in M^p_\mu(0, T; H)$, and $\Phi \in \mathcal{N}^p_Q(0, T; H)$ for some $p \geq 2$. Let $X(\cdot)$ be the mild solution of*

$$\begin{cases} dX(s) = (AX(s) + f(s))\, ds + \Phi(s) dW_Q(s) \\ X(0) = \xi \end{cases} \tag{1.42}$$

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

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

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

$$\lim_{n\to\infty} \mathbb{E}\left[\sup_{s\in[0,T]} |X^n(s) - X(s)|^p \right] = 0. \tag{1.44}$$

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

$$\sup_{s\in[0,T]} \mathbb{E}\left[|X^n(s)|^p\right] \leq M, \qquad \sup_{s\in[0,T]} \mathbb{E}\left[|X(s)|^p\right] \leq M. \tag{1.45}$$

Proof Observe first that the mild solution of (1.42) is well defined thanks to the assumptions on ξ, f and Φ, and

$$X(s) = e^{sA}\xi + \int_0^s e^{(s-r)A} f(r) dr + \int_0^s e^{(s-r)A} \Phi(r) dW_Q(r), \qquad s \in [0, T].$$

The same is true for the mild solution of (1.43) (which is also a strong solution).

To prove (1.44), we write, for $s \in [0, T]$,

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

It is enough to show that $\lim_{n\to\infty} \mathbb{E}\left[\sup_{s\in[0,T]} |I_i^n(s)|^p\right] = 0$ for $i \in \{1, 2, 3\}$. For $i = 3$ this follows from (1.21). To prove it for $i = 2$, we observe that (B.15) implies that if

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

then $\psi_n(r) \xrightarrow{n\to\infty} 0$ a.e. on Ω. Moreover, thanks to (B.14), there exists a C_1 such that, for all $t \in [0, T]$ and all n, $\left\|e^{tA_n}\right\| \le C_1$, so $\psi_n(r) \le 2C_1|f(r)|$ for all n. Since $\int_t^T |f(r)|dr < +\infty$ for almost every $\omega \in \Omega$, by the Lebesgue dominated convergence theorem we have

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

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

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

and the last expression is integrable (on Ω), since $f \in M_\mu^p(0, T; H)$. Therefore we can apply the Lebesgue dominated convergence theorem, obtaining $\lim_{n\to\infty} \mathbb{E} \left[\sup_{s\in[0,T]} |I_2^n(s)|^p\right] = 0$. The claim for $i = 1$ follows again from (B.15) and the Lebesgue dominated convergence theorem.

Estimates (1.45) are easy consequences of (B.14) and the assumptions on ξ, f, Φ. □

1.4.4 Uniqueness in Law

Definition 1.133 (*Finite-dimensional distributions*) Let $T > 0$ and $t \in [0, T)$. Consider a measurable space $(\Omega, \mathscr{F})$, two probability spaces $(\Omega_i, \mathscr{F}_i, \mathbb{P}_i)$ for $i = 1, 2$, and two processes $\{X_i(s)\}_{s\in[t,T]} : (\Omega_i, \mathscr{F}_i, \mathbb{P}_i) \to (\Omega, \mathscr{F})$. We say that $X_1(\cdot)$ and $X_2(\cdot)$ have the same *finite-dimensional distributions* on $D \subset [t, T]$ if for any

$t \le t_1 < t_2 < ... < t_n \le T$, $t_i \in D$ and $A \in \underbrace{\mathscr{F} \otimes \mathscr{F} \otimes ... \otimes \mathscr{F}}_{n \text{ times}}$, we have

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

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

Theorem 1.134 *Let H be a separable Hilbert space. Let $(\Omega_i, \mathscr{F}_i, \mathbb{P}_i)$ for $i = 1, 2$ be two complete probability spaces, and $(\tilde{\Omega}, \tilde{\mathscr{F}})$ be a measurable space. Let $\xi_i : \Omega_i \to \tilde{\Omega}$, $i = 1, 2$ be two random variables, and $f_i : [t, T] \times \Omega_i \to H$, $i = 1, 2$, be two processes satisfying*

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

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

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

Then

$$\mathcal{L}_{\mathbb{P}_1} \left(\int_t^\cdot f_1(s) ds, \xi_1 \right) = \mathcal{L}_{\mathbb{P}_2} \left(\int_t^\cdot f_2(s) ds, \xi_2 \right) \quad \text{on } [t, T]. \tag{1.46}$$

Proof See [471] Theorem 8.3, where the theorem was proved for a more general case of Banach space-valued processes. □

Theorem 1.135 *Let $\left(\Omega_1, \mathscr{F}_1, \mathscr{F}_s^{1,t}, \mathbb{P}_1, W_{Q,1}\right)$ and $\left(\Omega_2, \mathscr{F}_2, \mathscr{F}_s^{2,t}, \mathbb{P}_2, W_{Q,2}\right)$ be two generalized reference probability spaces. Let $\Phi_i : [t, T] \times \Omega_i \to \mathcal{L}_2(\Xi_0, H)$, $i = 1, 2$, be two $\mathscr{F}_s^{i,t}$-progressively measurable processes satisfying*

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

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

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

Then

$$\mathcal{L}_{\mathbb{P}_1} \left(\int_t^\cdot \Phi_1(s) dW_{Q,1}(s), \xi_1 \right) = \mathcal{L}_{\mathbb{P}_2} \left(\int_t^\cdot \Phi_2(s) dW_{Q,2}(s), \xi_2 \right) \quad \text{on } [t, T]. \tag{1.47}$$

Proof See [471] Theorem 8.6. □

Consider now an operator A and mappings b, σ satisfying Hypothesis 1.125, and $x \in H$. Let $\left(\Omega_1, \mathscr{F}_1, \mathscr{F}_s^{1,t}, \mathbb{P}_1, W_{Q,1}\right)$ and $\left(\Omega_2, \mathscr{F}_2, \mathscr{F}_s^{2,t}, \mathbb{P}_2, W_{Q,2}\right)$ be as in Theorem 1.135. For $i = 1, 2$ consider an $\mathscr{F}_s^{i,t}$-progressively measurable process $a_i : [t, T] \times \Omega_i \to \Lambda$.

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

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

Let $\mathcal{K}_i : \mathcal{H}_{p,i} \to \mathcal{H}_{p,i}$ be the continuous map (see [180], p. 189) defined as

$$\mathcal{K}_i(Z_i(\cdot))(s) := e^{(s-t)A}\zeta_i + \int_t^s e^{(s-r)A} b(r, Z_i(r), a_i(r))dr + \int_t^s e^{(s-r)A}\sigma(r, Z_i(r), a_i(r))dW_{Q,i}(r). \tag{1.48}$$

Lemma 1.136 *Consider the setting described above, and let* $\theta_i : [t, T] \times \Omega_i \to H, i = 1, 2$, *be stochastic processes. If*

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

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

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

Proof Observe that, since we only have to check the finite-dimensional distributions, the claims of Theorems 1.134 and 1.135 hold even if ξ_1 and ξ_2 are stochastic processes, with (1.46) and (1.47) then being true on some set of full measure. Let us choose a partition $(t_1, .., t_n)$, with $t \le t_1 < t_2 < ... < t_n \le T, t_k \in D, k = 1, ..., n$. We need to show that

$$\mathcal{L}_{\mathbb{P}_1}(\mathcal{K}_1(Z_1(\cdot))(t_k), a_1(t_k), W_{Q,1}(t_k), \theta_1(t_k), \zeta_1 : k = 1, ..., n) = \mathcal{L}_{\mathbb{P}_2}(\mathcal{K}_2(Z_2(\cdot))(t_k), a_2(t_k), W_{Q,1}(t_k), \theta_2(t_k), \zeta_2 : k = 1, ..., n). \tag{1.49}$$

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

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

and thus, by Theorem 1.134 applied with

$$\xi_1(\cdot) = (f^1(\cdot), \Phi^1(\cdot), Z_1(\cdot), a_1(\cdot), W_{Q,1}(\cdot), \theta_1(\cdot), \zeta_1),$$

$$\xi_2(\cdot) = (f^2(\cdot), \Phi^2(\cdot), Z_2(\cdot), a_2(\cdot), W_{Q,2}(\cdot), \theta_2(\cdot), \zeta_2),$$

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

Now, applying Theorem 1.135 with

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

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

we obtain

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

on D (we recall that the stochastic convolution terms in (1.48) and the stochastic integrals above have continuous trajectories a.e.). In particular, this implies that

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

We now repeat the above procedure for t_2, ..., t_n which will yield (1.49) as its consequence. □

Proposition 1.137 *Let the operator A and the mappings b, σ satisfy Hypothesis 1.125. Let $\left(\Omega_1, \mathscr{F}_1, \mathscr{F}_s^{1,t}, \mathbb{P}_1, W_{Q,1}\right)$ and $\left(\Omega_2, \mathscr{F}_2, \mathscr{F}_s^{2,t}, \mathbb{P}_2, W_{Q,2}\right)$ be two generalized reference probability spaces. Let $a_i\colon [t,T] \times \Omega_i \to \Lambda$, $i = 1, 2$ be an $\mathscr{F}_s^{i,t}$-progressively measurable process, and let $\zeta_i \in L^p(\Omega_i, \mathscr{F}_t^{i,t}, \mathbb{P}_i)$, $i = 1, 2$, $p > 2$. Let $\mathcal{L}_{\mathbb{P}_1}(a_1(\cdot), W_{Q,1}(\cdot), \zeta_1) = \mathcal{L}_{\mathbb{P}_2}(a_2(\cdot), W_{Q,1}(\cdot), \zeta_2)$ on some subset $D \subset [0,T]$ of full measure. Denote by $X_i(\cdot)$, $i = 1, 2$, the unique mild solution of*

$$\begin{cases} dX_i(s) = (AX_i(s) + b(s, X_i(s), a_i(s)))\,ds + \sigma(s, X_i(s), a_i(s))dW_{Q,i}(s) \\ X_i(t) = \zeta_i \end{cases} \tag{1.50}$$

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

Proof It is known (see [180], proof of Theorem 7.2, pp. 188–193) that the map $\mathcal{K}_i$ is a contraction in $\mathcal{H}_{p,i}$ if $[t, T]$ is small enough. Thus if we divide $[t, T]$ into such small intervals $[t, T_1]$, ...$[T_k, T]$, $X_i(\cdot)$ on $[t, T_1]$ is obtained as the limit in $\mathcal{H}_{p,i}$ (restricted to $[t, T_1]$) of the iterates $(\mathcal{K}_i^n(x))(\cdot)$. Therefore, using Lemma 1.136 and passing to the limit as $n \to +\infty$ we obtain

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

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

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

Thus, again using Lemma 1.136 and passing to the limit as $n \to +\infty$, it follows that

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

We repeat the procedure to obtain the required claim. □

1.5 Further Existence and Uniqueness Results in Special Cases

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

1.5.1 SDEs Coming from Boundary Control Problems

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

$$\begin{cases} dX(s) = \big(AX(s) + b(s, X(s), a(s)) + (\lambda I - A)^{\beta} G a_b(s)\big)\, ds \\ \qquad\qquad\qquad\qquad + \sigma(s, X(s), a(s)) dW_Q(s), \qquad s \in (0, T] \\ X(0) = \xi. \end{cases} \tag{1.51}$$

Hypothesis 1.138

(i) A generates an analytic semigroup e^{tA} for $t \geq 0$ and λ is a real constant such that $(\lambda I - A)^{-1} \in \mathcal{L}(H)$.
(ii) $a\colon [0, T] \times \Omega \to \Lambda$ is progressively measurable, $b(\cdot, \cdot, \cdot)$ satisfies (1.31) and (1.33).
(iii) Λ_b is a Hilbert space and $a_b(\cdot)\colon [0, T] \times \Omega \to \Lambda_b$ is progressively measurable.
(iv) $G \in \mathcal{L}(\Lambda_b, H)$.
(v) $\beta \in [0, 1)$.
(vi) γ is a constant belonging to the interval $\left[0, \frac{1}{2}\right)$, σ is a mapping such that $(\lambda I - A)^{-\gamma}\sigma : [0, T] \times H \times \Lambda_b \to \mathcal{L}_2(\Xi_0, H)$ is continuous. There exists a constant $C > 0$ such that

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

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

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

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

Remark 1.139 Part (i) of Hypothesis 1.138 implies, thanks to (B.18), that for every $\theta \geq 0$ there exists an $M_\theta > 0$ such that

$$|(\lambda I - A)^{\theta} e^{tA} x| \leq \frac{M_\theta}{t^\theta}|x|, \quad \text{for every } t \in (0, T],\ x \in H. \tag{1.52}$$

■

Following Remark 1.120, the definition of a mild solution of (1.51) is given by Definition 1.119 in which the term

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

is interpreted as

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

and the term

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

as

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

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

Remark 1.140 SDEs of type (1.51) appear most frequently in optimal control problems of parabolic equations on a domain $\mathcal{O} \subset \mathbb{R}^n$ with boundary control/noise, see Sect. 2.6.2. More precisely, the cases $\beta \in \left(\frac{3}{4}, 1\right)$ and $\beta \in \left(\frac{1}{4}, \frac{1}{2}\right)$ are related respectively to the Dirichlet and Neumann boundary control problems when one takes $\Lambda_b = L^2(\partial\mathcal{O})$ (or some subset of it) and $H = L^2(\mathcal{O})$. $\gamma \in \left(\frac{1}{4}, \frac{1}{2}\right)$ arises when one treats problems with boundary noise of Neumann type where again $\Lambda_b = L^2(\partial\mathcal{O})$ and $H = L^2(\mathcal{O})$. $\gamma, \beta \in \left(\frac{1}{2} - \varepsilon, \frac{1}{2}\right)$ arise in some specific Dirichlet boundary control/noise problems when one considers $\Lambda_b = L^2(\partial\mathcal{O})$ and a suitable weighted L^2 space as H. ∎

Theorem 1.141 *Assume that Hypothesis 1.138 holds, $p \geq 2$, and let $\alpha := \frac{1}{2} - \gamma$. Suppose that*

$$p > \frac{1}{\alpha} \tag{1.53}$$

and $a_b(\cdot) \in M^q_\mu(0, T; \Lambda_b)$ for some $q \geq p$, $q > \frac{1}{1-\beta}$. Then, for every initial condition $\xi \in L^2(\Omega, \mathscr{F}_0, \mathbb{P})$, there exists a unique mild solution $X(\cdot) = X(\cdot; 0, \xi, a(\cdot), a_b(\cdot))$ of (1.51) in $\mathcal{H}_2(0, T; H)$ with continuous trajectories $\mathbb{P}$-a.s. If there exists a constant $C > 0$ such that

$$\|(\lambda I - A)^{-\gamma} \sigma(s, x, a)\|_{\mathcal{L}_2(\Xi_0, H)} \leq C \tag{1.54}$$

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

$$\sup_{s\in[0,T]} \mathbb{E}|X(s)|^p \leq C_{T,p}(1 + \mathbb{E}|\xi|^p). \tag{1.55}$$

Proof Assume first that $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ where $p \geq 2$ without the restriction (1.53). Similarly to the proof of Theorem 1.127, we will show that for some $T_0 \in (0, T]$ the map

$$\begin{cases} \mathcal{K}\colon \mathcal{H}_p(0,T_0) \to \mathcal{H}_p(0,T_0), \\ \mathcal{K}(Y)(s) = e^{sA}\xi + \int_0^s e^{(s-r)A} b(r,Y(r),a(r))dr + \int_0^s (\lambda I - A)^{\beta} e^{(s-r)A} G a_b(r) dr \\ \qquad\qquad + \int_0^s (\lambda I - A)^{\gamma} e^{(s-r)A} (\lambda I - A)^{-\gamma} \sigma(r,Y(r),a(r)) dW_Q(r) \end{cases} \tag{1.56}$$

is well defined and is a contraction. The only difference between our case here and that considered in Theorem 1.127 is the last two terms in (1.56).

First we prove that $\mathcal{K}$ maps $\mathcal{H}_p(0,T_0)$ into $\mathcal{H}_p(0,T_0)$. We only show how to deal with the non-standard terms. For the third term in (1.56) we can argue as follows. If M_β is the constant from (1.52) for $\theta = \beta$, using (1.52), Hölder and Jensen's inequalities, and $q \geq p$, $q > \frac{1}{1-\beta}$, we obtain

$$\begin{aligned} \sup_{s\in[0,T_0]} \mathbb{E} \left| \int_0^s (\lambda I - A)^{\beta} e^{(s-r)A} G a_b(r) dr \right|^p &\\ \leq \sup_{s\in[0,T_0]} M_\beta^p \|G\|^p \mathbb{E} \left(\int_0^s \frac{1}{(s-r)^\beta} |a_b(r)| dr \right)^p &\\ \leq M_\beta^p \|G\|^p \left(\int_0^{T_0} \frac{1}{(T_0-r)^{\frac{\beta q}{q-1}}} dr \right)^{\frac{p(q-1)}{q}} \mathbb{E} \left[\int_0^{T_0} |a_b(r)|^q dr \right]^{\frac{p}{q}} &\\ \leq C_1 \left(\mathbb{E} \left[\int_0^{T_0} |a_b(r)|^q dr \right] \right)^{\frac{p}{q}} < +\infty. & \end{aligned} \tag{1.57}$$

As regards the stochastic integral term, using Theorem 1.111, (1.52), and Hypothesis 1.138-(vi), we estimate

$$\begin{aligned} \sup_{s\in[0,T_0]} \mathbb{E} \left| \int_0^s (\lambda I - A)^{\gamma} e^{(s-r)A} (\lambda I - A)^{-\gamma} \sigma(r,Y(r),a(r)) dW_Q(r) \right|^p &\\ \leq \sup_{s\in[0,T_0]} C_1 \mathbb{E} \left| \int_0^s \frac{1}{(s-r)^{2\gamma}} \|(\lambda I - A)^{-\gamma} \sigma(r,Y(r),a(r))\|^2_{\mathcal{L}_2(\Xi_0,H)} dr \right|^{\frac{p}{2}} &\\ \leq \sup_{s\in[0,T_0]} C_2 \left(\int_0^{T_0} \frac{1}{(T_0-r)^{2\gamma}} dr \right)^{\frac{p}{2}-1} \int_0^s \frac{1}{(s-r)^{2\gamma}} \mathbb{E}[(1+|Y(r)|)^p] dr &\\ \leq C_3 \left(1 + |Y|^p_{\mathcal{H}_p(0,T_0)} \right) & \end{aligned} \tag{1.58}$$

for some constant C_3. Progressive measurability of all the terms appearing in the definition of $\mathcal{K}(Y)(\cdot)$ can be proved by using estimates similar to (1.57) and (1.58) and arguing as in Remark 1.123.

Regarding the proof that, for T_0 small enough, $\mathcal{K}$ is a contraction, the only non-standard term to check is the stochastic convolution term, since the third term in

(1.56) does not depend on X. Arguing as before we have that for $X, Y \in \mathcal{H}_p(0, T_0)$, thanks to Theorem 1.111, (1.52), Hypothesis 1.138-(vi), and Jensen's inequality,

$$\begin{aligned}
&\sup_{s\in[0,T_0]} \mathbb{E}\left|\int_0^s (\lambda I - A)^{\gamma} e^{(s-r)A} (\lambda I - A)^{-\gamma} \left[\sigma(r, X(r), a(r)) - \sigma(r, Y(r), a(r))\right] dW_Q(r)\right|^p \\
&\le \sup_{s\in[0,T_0]} C_1 \mathbb{E}\left(\int_0^s \frac{1}{(s-r)^{2\gamma}} \left\|(\lambda I - A)^{-\gamma} \left[\sigma(r, X(r), a(r)) - \sigma(r, Y(r), a(r))\right]\right\|^2_{\mathcal{L}_2(\Xi_0, H)} dr\right)^{\frac{p}{2}} \\
&\qquad \le \sup_{s\in[0,T_0]} C_2 \mathbb{E}\left(\int_0^s \frac{1}{(s-r)^{2\gamma}} |X(r) - Y(r)|^2 dr\right)^{\frac{p}{2}} \\
&\quad \le \sup_{s\in[0,T_0]} C_2 \left(\int_0^{T_0} \frac{1}{(T_0-r)^{2\gamma}} dr\right)^{\frac{p}{2}-1} \int_0^s \frac{1}{(s-r)^{2\gamma}} \mathbb{E}[|X(r) - Y(r)|^p] dr \\
&\qquad\qquad \le \omega(T_0) |X - Y|^p_{\mathcal{H}_p(0,T_0)},
\end{aligned} \tag{1.59}$$

where $\omega(r) \xrightarrow{r\to 0^+} 0$. So for T_0 small enough (which is independent of the initial condition) we can apply the Banach fixed point theorem in $\mathcal{H}_p(0, T_0)$ as in the proof of Theorem 1.127 (see also the proof of [180], Theorem 7.2, p. 188). The process can now be reapplied on intervals $[T_0, 2T_0]$, ..., $[kT_0, T]$, where $k = [T/T_0]$, to obtain the existence of a unique mild solution in $\mathcal{H}_p(0, T)$ in the sense of the integral equality being satisfied for a.e. $s \in [0, T]$.

Estimate (1.55) follows from similar arguments using the growth assumptions on b, σ in Hypothesis 1.138 and Gronwall's lemma in the form given in Proposition D.30.

We will now prove the continuity of the trajectories if condition (1.53) is satisfied. We will only prove the continuity of the stochastic convolution term in (1.56) since the continuity of the other terms is easier to show. In particular, the continuity of the trajectories of the third term in (1.56) follows from Lemma 1.115-(ii).

Let now $p > \frac{1}{\alpha}$. Hence there is an $0 < \alpha' < \alpha$ such that $p > \frac{1}{\alpha'}$. Then, for $r \in [t, T]$, using (1.52), (1.55), Hypothesis 1.138-(vi), and Jensen's inequality

$$\begin{aligned}
&\mathbb{E}\left(\int_0^r (r-h)^{-2\alpha'} \left\|(\lambda I - A)^{\gamma} e^{(r-h)A} (\lambda I - A)^{-\gamma} \sigma(h, X(h), a(h))\right\|^2_{\mathcal{L}_2(\Xi_0, H)} dh\right)^{\frac{p}{2}} \\
&\le \mathbb{E}\left(\int_0^r (r-h)^{-2\alpha'} \|(\lambda I - A)^{\gamma} e^{(r-h)A}\|^2_{\mathcal{L}(H)} \left\|(\lambda I - A)^{-\gamma} \sigma(h, X(h), a(h))\right\|^2_{\mathcal{L}_2(\Xi_0, H)} ds\right)^{\frac{p}{2}} \\
&\qquad \le C_1 \mathbb{E}\left(\int_0^r (r-h)^{-2\alpha'} (r-h)^{-2\gamma} (1 + |X(h)|)^2 dh\right)^{\frac{p}{2}} \\
&\le C_1 \left(\int_0^T (T-h)^{-2\alpha'} (T-h)^{-2\gamma} dh\right)^{\frac{p}{2}} \sup_{h\in[0,T]} \mathbb{E}[(1 + |X(h)|)^p] =: C_2 < +\infty.
\end{aligned} \tag{1.60}$$

Observe that C_2 does not depend on $r \in [0, T]$. This proves (1.25) and thus the claim follows from Proposition 1.116. When (1.54) holds, estimate (1.60) is easier and can be done for any exponent $p' > 1/\alpha$ in place of p, and thus (1.25) is always satisfied.

Finally, we need to discuss the continuity of the trajectories if $\xi \in L^2(\Omega, \mathscr{F}_0, \mathbb{P})$. We argue as in the proof of Theorem 7.2 of [180]. For $n \ge 1$ we define the random variables

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

The solutions $X(\cdot; 0, \xi, a(\cdot), a_b(\cdot))$ and $X(\cdot; 0, \xi_n, a(\cdot), a_b(\cdot))$ on $[0, T_0]$ are obtained as fixed points in $\mathcal{H}_2(0, T_0)$ and $\mathcal{H}_p(0, T_0)$, with p large enough, of the same contraction map (1.56) with the second map having the term $e^{sA}\xi_n$ in place of $e^{sA}\xi$. Therefore both solutions can be obtained as limits of successive iterations starting, say, from processes $e^{sA}\xi$ and $e^{sA}\xi_n$, respectively. It is then easy to see that we have $X(\cdot; 0, \xi, a(\cdot), a_b(\cdot)) = X(\cdot; 0, \xi_n, a(\cdot), a_b(\cdot))$, $\mathbb{P}$-a.s. on $\{\omega : |\xi(\omega)| \le n\}$. However, the solutions $X(\cdot; 0, \xi_n, a(\cdot), a_b(\cdot))$ have continuous trajectories. Thus $X(\cdot; 0, \xi, a(\cdot), a_b(\cdot))$ has continuous trajectories $\mathbb{P}$-a.s. on $[0, T_0]$ and we can then continue the argument on intervals $[T_0, 2T_0]$, ... □

Proposition 1.142 *Let the assumptions of Theorem 1.141 be satisfied. Denote the unique mild solution of (1.51) in $\mathcal{H}_p(0, T; H)$ by $X(\cdot) = X(\cdot; 0, \xi, a(\cdot), a_b(\cdot))$.*

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

(ii) Let $\left(\Omega_1, \mathscr{F}_1, \mathscr{F}_s^1, \mathbb{P}_1, W_{Q,1}\right)$ and $\left(\Omega_2, \mathscr{F}_2, \mathscr{F}_s^2, \mathbb{P}_2, W_{Q,2}\right)$ be two generalized reference probability spaces. Let $\zeta_i \in L^p(\Omega_i, \mathscr{F}_0^i, \mathbb{P}_i)$, $i = 1, 2$. Let (a^i, a_b^i): $[0, T] \times \Omega_i \to \Lambda \times \Lambda_b$, $i = 1, 2$ be $\mathscr{F}_s^i$-progressively measurable processes satisfying the assumptions of Theorem 1.141. Suppose that $\mathcal{L}_{\mathbb{P}_1}(a^1(\cdot), a_b^1(\cdot), W_{Q,1}(\cdot), \zeta_1) = \mathcal{L}_{\mathbb{P}_2}(a^2(\cdot), a_b^2(\cdot), W_{Q,1}(\cdot), \zeta_2)$ on some subset $D \subset [t, T]$ of full measure. Then $\mathcal{L}_{\mathbb{P}_1}(X(\cdot; 0, \zeta_1, a^1(\cdot), a_b^1(\cdot)), a^1(\cdot), a_b^1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(X(\cdot; 0, \zeta_2, a^2(\cdot), a_b^2(\cdot)), a^2(\cdot), a_b^2(\cdot))$ on D.

(iii) The solution of (1.51) is unique in $M_\mu^p(0, T; H)$ as well.

Proof (i) If $X_i(\cdot) := X\left(\cdot; 0, \xi^i, a^i(\cdot), a_b^i(\cdot)\right)$, arguing as in (1.59) and using Hölder's inequality, we obtain, for $s \in [0, T]$,

$$\mathbb{E}|X_1(s) - X_2(s)|^p \le C_T \int_0^s \mathbb{E}|X_1(r) - X_2(r)|^p dr,$$

and the claim follows by using Gronwall's lemma (Proposition D.29), and the continuity of the trajectories.

(ii) The argument is the same as the one used to prove Lemma 1.136 and Proposition 1.137, since in the current case the solution is also found by iterating the map $\mathcal{K}$.

(iii) The uniqueness in $M_\mu^p(0, T_0; H)$ follows from the estimate in Part (i) above and Proposition D.29. □

1.5.2 Semilinear SDEs with Additive Noise

In this section we give more precise results for some semilinear SDEs with additive noise, i.e. for Eq. (1.28) when the coefficient σ is constant and we have possible unboundedness in the drift.

Hypothesis 1.143

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

$$|e^{tA}x| \leq Me^{\omega t}|x|, \quad \forall t \geq 0,\ x \in H. \tag{1.61}$$

(ii) $Q \in \mathcal{L}^{+}(\Xi), \sigma \in \mathcal{L}(\Xi, H)$ and $e^{sA}\sigma Q\sigma^{*}e^{sA^{*}} \in \mathcal{L}_{1}(H)$ for all $s > 0$. Moreover, for all $t \geq 0$,

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

so the symmetric positive operator

$$Q_t : H \to H, \qquad Q_t := \int_0^t e^{sA}\sigma Q\sigma^{*}e^{sA^{*}} ds, \tag{1.62}$$

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

$$\mathrm{Tr}\,[Q_s] < +\infty. \tag{1.63}$$

Let W_Q be a Q-Wiener process in Ξ and consider the stochastic convolution process defined, for $s \geq 0$, as follows:

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

Proposition 1.144 *Suppose that Hypothesis 1.143 is satisfied. Then the process $W^A(\cdot)$ defined in (1.64) is a Gaussian process with mean 0 and covariance operator Q_s, is mean square continuous and $W^A(\cdot) \in \mathcal{H}_p^{\mu}(0, T; H)$ for every $p \geq 2$. Moreover, if there exists a $\gamma > 0$ such that*

$$\int_0^T s^{-\gamma}\mathrm{Tr}\left[e^{sA}\sigma Q\sigma^{*}e^{sA^{*}}\right] ds < \infty, \tag{1.65}$$

then $W^A(\cdot)$ has continuous trajectories[5] *and, for $p > 0$,*

[5]Without assuming (1.65) such continuity of trajectories may fail to hold, see e.g. [357].

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

Proof See [180] Chap. 5, Theorems 5.2 and 5.11. The fact that $W^A(\cdot) \in \mathcal{H}_p^\mu(0, T; H)$ for every $p \ge 2$ follows from Theorem 1.111. The last estimate can be found, for example, as a particular case of Proposition 3.2 in [284]. □

A completely analogous result holds for the stochastic convolution starting at a point $t \ge 0$, i.e.

$$W^A(t, s) := \int_t^s e^{(s-r)A}\sigma dW_Q(r), \qquad s \ge t. \tag{1.66}$$

Let $T > 0$. We consider the SDE

$$\begin{cases} dX(s) = (AX(s) + b(s, X(s)))\, ds + \sigma dW_Q(s), \quad s > 0 \\ X(0) = \xi. \end{cases} \tag{1.67}$$

Hypothesis 1.145 $p \ge 1$ and $b(s, x) = b_0(s, x, a_1(s)) + a_2(s)$, where:

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

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

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

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

$$|e^{tA}a_2(s, \omega)| \le t^{-\beta}g(s, \omega) \qquad \forall (t, s, \omega) \in [0, T] \times [0, T] \times \Omega, \tag{1.68}$$

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

Hypothesis 1.145 covers some cases which are not standard and for which a separate proof of existence and uniqueness of mild solutions of (1.67) is required.

Remark 1.146 Hypothesis 1.145-(ii) is satisfied, for example, when A is the generator of an analytic C_0-semigroup and the process $a_2(\cdot)$ is of the form $a_2(s) = (\lambda I - A)^\beta a_3(s)$, where $\lambda \in \mathbb{R}$ is such that $(\lambda I - A)$ is invertible, $\beta \in (0, 1)$, $a_3(\cdot) \in M^q_\mu(0, T; H)$, $q \ge p, q > \frac{1}{1-\beta}$. In such cases the definition of a mild solution of (1.67) is given by Definition 1.119 in which the formal term

$$\int_0^s e^{(s-r)A}a_2(r)dr = \int_0^s e^{(s-r)A}(\lambda I - A)^\beta a_3(r)dr$$

appearing in the definition of a mild solution is interpreted as

$$\int_0^s (\lambda I - A)^\beta e^{(s-r)A} a_3(r)dr.$$

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

Another more general case where Hypothesis 1.145-(ii) is satisfied is when $a_2(\cdot) : [0, T] \times \Omega \to V^*$ is progressively measurable, where V^* denotes the topological dual of $V = D(A^*)$. In such a case the semigroup e^{tA} may be extended, by a standard construction (see e.g. [232]), to the space V^*. Denoting this extension still by e^{tA}, the process $e^{tA}a_2(\cdot) : [0, T] \times \Omega \to V^*$ is well defined. If we further assume that $e^{tA}a_2(\cdot)$ takes values in H and satisfies (1.68) for some $\beta \in (0, 1)$, then Hypothesis 1.145-(ii) is satisfied. A similar and even slightly more general case has been studied in [232] in a deterministic context. ■

Proposition 1.147 *Let $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$ and Hypotheses 1.143 and 1.145 be satisfied. Then Eq. (1.67) has a unique mild solution $X(\cdot; 0, \xi) \in \mathcal{H}_p^\mu(0, T; H)$. The solution satisfies, for some $C_p(T) > 0$ independent of ξ,*

$$\sup_{s\in[0,T]} \mathbb{E}\left[|X(s; 0, \xi)|^p\right] \le C_p(T)(1 + \mathbb{E}[|\xi|^p]). \tag{1.69}$$

Moreover, if $\xi_1, \xi_2 \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$, we have, $\mathbb{P}$-a.s.,

$$|X(s; 0, \xi_1) - X(s; 0, \xi_2)| \le Me^{\omega T}|\xi_1 - \xi_2|e^{Me^{\omega T}\int_0^s f(r)dr}, \qquad s \in [0, T]. \tag{1.70}$$

Finally, if (1.65) also holds for some $\gamma > 0$, then the solution $X(\cdot; 0, \xi)$ has $\mathbb{P}$-a.s. continuous trajectories, and if $\xi = x \in H$ is deterministic we then have

$$\mathbb{E}(\sup_{s\in[0,T]} |X(s)|^p) \le C_p(T)(1 + |x|^p) \tag{1.71}$$

for a suitable constant $C_p(T) > 0$ independent of x. In particular, if g in Hypothesis 1.145-(ii) is in $M_\mu^q(0, T; \mathbb{R})$ for every $q \ge 1$, then estimate (1.69) holds for every $p > 0$ and the same is true for (1.71) if $\xi = x \in H$.

Proof The proof of existence and uniqueness uses the same techniques employed in the Lipschitz case (Theorem 1.127) but contains a small additional difficulty due the presence of the term $a_2(\cdot)$ and possible singularities in s of the Lipschitz norm of $b_0(s, \cdot)$. We will write $\mathcal{H}_p(0, T)$ for $\mathcal{H}_p^\mu(0, T; H)$. For $Y \in \mathcal{H}_p(0, T)$ we set

$$\mathcal{K}(Y)(s) = e^{(s-t)A}\xi + \int_0^s e^{(s-r)A} b_0(r, Y(r), a_1(r))dr + \int_0^s e^{(s-r)A} a_2(r)dr + W^A(s). \tag{1.72}$$

W^A belongs to $\mathcal{H}_p(0, T)$ thanks to Proposition 1.144. Hypotheses 1.145-(i) and 1.145-(ii) ensure, respectively, that the second and third term in the definition of the map $\mathcal{K}$ belong to $\mathcal{H}_p(0, T)$ as well (one can use the same arguments as these to obtain

(1.57) when $\beta \in (0, 1)$ and Hölder's inequality if $\beta = 0$). So $\mathcal{K}$ maps $\mathcal{H}_p(0, T)$ into itself. For $Y_1, Y_2 \in \mathcal{H}_p(0, T)$, $s \in [0, T]$,

$$|\mathcal{K}(Y_1)(s) - \mathcal{K}(Y_2)(s)| \le M e^{\omega T} \int_0^s f(r)|Y_1(r) - Y_2(r)|dr,$$

which yields, for $T_0 \in (0, T]$,

$$\begin{aligned} |\mathcal{K}(Y_1) - \mathcal{K}(Y_2)|^p_{\mathcal{H}_p(0,T_0)} &\le M e^{\omega T} \sup_{s\in[0,T_0]} \mathbb{E}\left[\int_0^s f(r)|Y_1(r) - Y_2(r)|dr\right]^p \\ &\le M e^{\omega T} \left[\int_0^{T_0} f(r)dr\right]^p \sup_{s\in[0,T_0]} \mathbb{E}|Y_1(s) - Y_2(s)|^p \\ &= M e^{\omega T} \left[\int_0^{T_0} f(r)dr\right]^p |Y_1 - Y_2|^p_{\mathcal{H}_p(0,T_0)}. \end{aligned} \tag{1.73}$$

Therefore, if T_0 is sufficiently small, we can apply the contraction mapping principle to find the unique mild solution of (1.67) in $\mathcal{H}_p(0, T_0)$. The existence and uniqueness of a solution on the whole interval $[0, T]$ follows, as usual, by repeating the procedure a finite number of times, since the estimate (1.73) does not depend on the initial data, and the number of steps does not blow up since f is integrable. Estimate (1.69) follows from (1.72) applied to the solution X if we perform estimates similar to those above and use Gronwall's Lemma.

To show (1.70) we observe that if $Z(s) = X(s; 0, \xi_1) - X(s; 0, \xi_2)$, then for $s \in [0, T]$

$$Z(s) = e^{sA}(\xi_1 - \xi_2) + \int_0^s e^{(s-r)A}[b_0(r, X(r; 0, \xi_1), a_1(r)) - b_0(r, X(r; 0, \xi_2), a_1(r))]dr.$$

By Hypothesis 1.145 we thus have

$$|Z(s)| \le M e^{\omega T}|\xi_1 - \xi_2| + M e^{\omega T} \int_0^s f(r)|Z(r)|dr, \quad s \in [0, T]$$

so that, by Gronwall's inequality (see Proposition D.29),

$$|Z(s)| \le M e^{\omega T}|\xi_1 - \xi_2| e^{M e^{\omega T} \int_0^s f(r)dr},$$

which gives the claim. The continuity of trajectories follows from Proposition 1.144, Hypothesis 1.145 and Lemma 1.115 for the second and fourth terms in (1.72), and from Lemma 1.117 for the $\int_0^s e^{(s-r)A}a_2(r)dr$ term.

The last estimate (1.71) follows by standard arguments (see the proof of (1.37) in Theorem 1.130) if we use Proposition 1.144. This implies that if $g \in M^q_\mu(0, T; \mathbb{R})$ for any $q > 0$, (1.71) holds for any $p \ge 2$. For $p \in (0, 2)$, defining $Z_r(s) :=$

$\sup_{s\in[0,T]} |X(s)|^r$, we have

$$\mathbb{E}(Z_p(s)) \leq [\mathbb{E}(Z_p(s)^{2/p})]^{p/2} \leq (C(1+|x|^2))^{p/2} \leq C_1(1+|x|^p).$$

□

Proposition 1.148 *Assume that Hypotheses 1.143, 1.145, together with (1.65), are satisfied, and let $a_2(\cdot)$ be as in Remark 1.146. Then:*

(i) Let $\xi_1, \xi_2 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P})$, $\xi_1 = \xi_2$ $\mathbb{P}$-a.s. Let $(a_1^1(\cdot), a_3^1(\cdot))$, $(a_1^2(\cdot), a_3^2(\cdot))$ be two processes satisfying Hypothesis 1.145, together with Remark 1.146, such that $(a_1^1(\cdot), a_3^1(\cdot)) = (a_1^2(\cdot), a_3^2(\cdot))$, $dt \otimes \mathbb{P}$-a.s. Then, denoting by $X^i(\cdot; 0, \xi_i)$ the solution of (1.67) for $b(s,x) = (\lambda - A)^\beta a_3^i(s) + b_0(s, x, a_1^i(s))$, we have $X^1(\cdot; 0, \xi_1) = X^2(\cdot; 0, \xi_2)$, $\mathbb{P}$-a.s. on $[0,T]$.

(ii) Let $(\Omega_1, \mathscr{F}_1, \mathscr{F}_s^1, \mathbb{P}_1, W_{Q,1})$ and $(\Omega_2, \mathscr{F}_2, \mathscr{F}_s^2, \mathbb{P}_2, W_{Q,2})$ be two generalized reference probability spaces. Let $\xi_i \in L^2(\Omega_i, \mathscr{F}_0^i, \mathbb{P}_i)$, $i = 1, 2$. Let $a_1^i(\cdot), a_3^i(\cdot)$, $i = 1, 2$, be processes on $[0,T] \times \Omega_i$ satisfying Hypothesis 1.145, together with Remark 1.146. Suppose that $\mathcal{L}_{\mathbb{P}_1}(a_1^1(\cdot), a_3^1(\cdot), W_{Q,1}(\cdot), \xi_1) = \mathcal{L}_{\mathbb{P}_2}(a_1^2(\cdot), a_3^2(\cdot), W_{Q,2}(\cdot), \xi_2)$. Then $\mathcal{L}_{\mathbb{P}_1}(X^1(\cdot; 0, \xi_1), a_1^1(\cdot), a_3^1(\cdot)) = \mathcal{L}_{\mathbb{P}_2}(X(\cdot; 0, \xi_2), a_1^2(\cdot), a_3^2(\cdot))$.

(iii) If $f \in L^2(0,T;\mathbb{R})$ then the solution of (1.67) ensured by Proposition 1.147 is unique in $M_\mu^2(0,T;H)$ as well.

Proof Parts (i) and (ii) are proved similarly as Proposition 1.142 (i)–(ii). Part (iii) follows from (1.70), which is also true in this case. We also point out that if $p = 2$, $f \in L^2(0,T;\mathbb{R})$ then $\mathcal{K}$ maps $M_\mu^2(0,T;H)$ into itself and is a contraction in $M_\mu^2(0,T_0;H)$ for small T_0. □

1.5.3 Semilinear SDEs with Multiplicative Noise

This section contains a result for a class of semilinear SDEs with multiplicative noise. Let $T > 0$, and let H, Ξ, Λ and a generalized reference probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s\in[0,T]}, \mathbb{P}, W)$ be as in Sect. 1.3, where $W(t)$, $t \in [0,T]$, is a cylindrical Wiener process (so here $\Xi_0 = \Xi$). We consider the following SDE in H for $s \in [0,T]$:

$$\begin{cases} dX(s) = AX(s)\,ds + b(s, X(s), a(s))\,ds + \sigma(s, X(s), a(s))\,dW(s), \\ X(0) = \xi. \end{cases} \tag{1.74}$$

Hypothesis 1.149

(i) The operator A generates a strongly continuous semigroup e^{tA} for $t \geq 0$ in H.
(ii) $a(\cdot)$ is a Λ-valued progressively measurable process.

(iii) b is a function such that, for all $s \in (0, T]$, $e^{sA}b : [0, T] \times H \times \Lambda \to H$ is measurable and there exist $L \geq 0$ and $\gamma_1 \in [0, 1)$ such that, with $f_1(s) = Ls^{-\gamma_1}$,

$$|e^{sA}b(t, x, a)| \leq f_1(s)(1 + |x|), \tag{1.75}$$

$$|e^{sA}(b(t, x, a) - b(t, y, a))| \leq f_1(s)|x - y|, \tag{1.76}$$

for any $s \in (0, T],\ t \in [0, T],\ x, y \in H,\ a \in \Lambda$.

(iv) The function $\sigma : [0, T] \times H \times \Lambda \to \mathcal{L}(\Xi, H)$ is such that, for every $v \in \Xi$, the map $\sigma(\cdot, \cdot, \cdot)v : [0, T] \times H \times \Lambda \to H$ is measurable and, for every $s > 0$, $t \in [0, T]$, $a \in \Lambda$ and $x \in H$, $e^{sA}\sigma(t, x, a)$ belongs to $\mathcal{L}_2(\Xi, H)$. Moreover, there exists a $\gamma_2 \in [0, 1/2)$ such that, with $f_2(s) = Ls^{-\gamma_2}$,

$$|e^{sA}\sigma(t, x, a)|_{\mathcal{L}_2(\Xi,H)} \leq f_2(s)(1 + |x|), \tag{1.77}$$
$$|e^{sA}\sigma(t, x, a) - e^{sA}\sigma(t, y, a)|_{\mathcal{L}_2(\Xi,H)} \leq f_2(s)|x - y|, \tag{1.78}$$

for every $s \in (0, T],\ t \in [0, T],\ x, y \in H,\ a \in \Lambda$.

Remark 1.150 Hypothesis 1.149-(iii) covers some cases where the term b is unbounded, which arise, for example, from a stochastic heat equation with a non-zero boundary condition which may also depend on the state variable x (see the last part of Example 4.222).

Moreover, Hypothesis 1.149-(iv) applies to cases, such as reaction-diffusion equations (see e.g. [177], Chap. 11 or, in our Chap. 2, Sect. 2.6.1 and, in particular, Eqs. (2.79) and (2.83), where the operator σ is a nonlinear Nemytskii type operator. Indeed, in such cases it is known that, when the underlying space is $L^2(\mathcal{O})$ ($\mathcal{O} \subset \mathbb{R}^n$, open), the operator $\sigma(t, \cdot) : H \to \mathcal{L}(H)$ is never Lipschitz continuous while $e^{sA}\sigma(t, \cdot) : H \to \mathcal{L}_2(H)$ is so (see e.g. [177], proof of Theorem 11.2.4 and Sect. 11.2.1, or [283], Remark 2.2). ■

Remark 1.151 If in Hypothesis 1.125 we set $W_Q = Q^{1/2}\tilde{W}$ for a suitable cylindrical Wiener process $\tilde{W}$ in $\tilde{\Xi} = R(Q^{-1/2})$ and we substitute σ with $\tilde{\sigma} = \sigma Q^{1/2}$, it is easy to see that Hypothesis 1.149 is more general. However, we need to replace Ξ by $\tilde{\Xi}$. A cylindrical Wiener process W in Ξ may not be adapted to the original filtration. Similarly, Hypothesis 1.149 is more general than Hypotheses 1.143 and 1.145, together with (1.65), if we take f bounded and $a_2(\cdot) \equiv 0$ there. ■

The solution of Eq. (1.74) is defined in the mild sense of Definition 1.119, where the convolution term

$$\int_0^s e^{(s-r)A}\sigma(r, X(r), a(r))\, dW(r), \qquad s \in [0, T],$$

makes sense thanks to (1.77) and Remark 1.123. Moreover, since $s \to e^{sA}b(t, x, a)$ is continuous on $(0, T]$ for every $t \in [0, T], x \in H, a \in \Lambda$, we have from Lemma 1.18 that $e^{\cdot A}b$ is $\mathcal{B}([0, T]) \otimes \mathcal{B}([0, T]) \otimes \mathcal{B}(H) \otimes \mathcal{B}(\Lambda)/\mathcal{B}(H)$-measurable.

Theorem 1.152 *Let Hypothesis 1.149 hold and let $a(\cdot)$ be a Λ-valued, progressively measurable process. Let $p \in [2, \infty)$. Then, for every initial condition $\xi \in L^p(\Omega, \mathscr{F}_0, \mathbb{P})$, the SDE (1.74) has a unique mild solution $X(\cdot)$ in $\mathcal{H}_p(0, T; H)$. The solution satisfies*

$$\sup_{s\in[0,T]} \mathbb{E}\left[|X(s)|^p\right] \le C_0(1 + \mathbb{E}[|\xi|^p]) \tag{1.79}$$

for some constant $C_0 > 0$ independent of ξ and $a(\cdot)$. The mild solution $X(\cdot)$ has continuous trajectories and, when $\xi \equiv x \in H$, we have

$$\mathbb{E}\left[\sup_{s\in[0,T]} |X(s)|^p\right] \le C(1 + |x|^p), \quad \text{for all } \ p > 0, \tag{1.80}$$

for some constant C depending only on $p, \gamma_1, \gamma_2, T, L$ and $M_T := \sup_{s\in[0,T]} |e^{sA}|$.

Finally, when b and σ do not depend on a, mild solutions of (1.74) defined on different generalized reference probability spaces have the same laws.

Proof Let $p \ge 2$. The existence of a unique solution is proved using the Banach contraction mapping theorem in $\mathcal{H}_p(0, T_0)$ for some $T_0 \in (0, T)$ small enough. We define $\mathcal{K}\colon \mathcal{H}_p(0, T) \to \mathcal{H}_p(0, T)$ by

$$\mathcal{K}(Y)(s) := e^{sA}\xi + \int_0^s e^{(s-r)A} b(r, Y(r), a(r))dr + \int_0^s e^{(s-r)A}\sigma(r, Y(r), a(r))dW(r). \tag{1.81}$$

We observe first that this expression belongs to $\mathcal{H}_p(0, T)$. Thanks to (1.75), (1.77) and Theorem 1.111, we have

$$\begin{aligned}
\mathbb{E}&\left|\int_0^s e^{(s-r)A} b(r, Y(r), a(r))dr + \int_0^s e^{(s-r)A}\sigma(r, Y(r), a(r))dW(r)\right|^p \\
&\le C_p\Bigg(\mathbb{E}\left|\int_0^s [f_1(s-r)(1 + |Y(r)|)]\, dr\right|^p \\
&\quad + \mathbb{E}\left|\int_0^s e^{(s-r)A}\sigma(r, Y(r), a(r))dW(r)\right|^p\Bigg) \\
&\le C_p\left[\int_0^T f_1(r)dr\right]^p \sup_{r\in[0,T]} \mathbb{E}(1 + |Y(r)|)^p \\
&\quad + C_p\left[\int_0^T f_2^2(r)dr\right]^{\frac{p}{2}} \sup_{r\in[0,T]} \mathbb{E}(1 + |Y(r)|)^p,
\end{aligned} \tag{1.82}$$

where the constant C_p depends only on p. Therefore, for any $Y \in \mathcal{H}_p(0, T)$, $\mathcal{K}(Y) \in \mathcal{H}_p(0, T)$. The estimates showing that $\mathcal{K}$ is a contraction on $\mathcal{H}_p(0, T_0)$ for $T_0 \in (0, T]$ small enough are essentially the same. Using (1.76) and (1.78) instead of (1.75) and (1.77) we obtain, for all $Y_1, Y_2 \in \mathcal{H}_p(0, T_0)$,

$$|\mathcal{K}(Y_1) - \mathcal{K}(Y_2)|^p_{\mathcal{H}_p(0,T_0)} \le C_p \left(\left[\int_0^{T_0} f_1(r)dr \right]^p + \left[\int_0^{T_0} f_2^2(r)dr \right]^{\frac{p}{2}} \right) \sup_{r\in[0,T_0]} \mathbb{E}(|Y_1(r) - Y_2(r)|^p),$$

and thus $\mathcal{K}$ is a contraction in $\mathcal{H}_p(0, T_0)$ if $T_0 \in (0, T]$ is small enough. The existence and uniqueness of solution in $\mathcal{H}_p(0, T)$ follows, as usual, by repeating the procedure a finite number of times, since the estimate does not depend on the initial data, and the number of steps does not blow up since f_1 and f_2^2 are integrable. Estimate (1.79) follows in a standard way by applying estimates like those in (1.82) to the fixed point of the map $\mathcal{K}$ and using Gronwall's lemma (see also the proof of Theorem 7.5 in [180]).

The continuity of the trajectories and (1.80) are proved using the factorization method similarly to the way it is done in the proof of Proposition 6.9 for $p > 2$. We extend (1.80) to $0 < p \le 2$ in the same way as in the proof of Proposition 1.147. Uniqueness in law is proved similarly as in Proposition 1.137. □

Proposition 1.153 *Assume that Hypothesis 1.149 holds. Let (t_1, x_1), $(t_2, x_2) \in [0, T] \times H$ with $t_1 \le t_2$. Denote by $X(\cdot; t_1, x_1, a(\cdot))$, $X(\cdot; t_2, x_2, a(\cdot))$ the corresponding mild solutions of (1.74) with the same progressively measurable process $a(\cdot)$ and initial conditions $X(t_i) = x_i \in H$, $i = 1, 2$. Then, for all $s \in [t_2, T]$ we have, setting $\gamma_3 := [2(1-\gamma_1)] \wedge [1 - 2\gamma_2]$,*

$$\begin{aligned}\mathbb{E}[|X(s; t_1, x_1, a(\cdot)) - X(s; t_2, x_2, a(\cdot))|^2] \le \\ \le C_2 \left[|x_1 - x_2|^2 + (1 + |x_1|^2)|t_2 - t_1|^{\gamma_3} + |e^{(t_2-t_1)A}x_1 - x_1|^2 \right]\end{aligned} \tag{1.83}$$

for some constant C_2 depending only on γ_1, γ_2, T, L and $M := \sup_{s\in[0,T]} |e^{sA}|$. Moreover, the term $|e^{(t_2-t_1)A}x_1 - x_1|^2$ can be replaced by $|e^{(t_2-t_1)A}x_2 - x_2|^2$.

Proof To simplify the notation we define $X_i(s) := X(s; t_i, x_i, a(\cdot))$, $b(r, X_i(r)) := b(r, X_i(r), a(r))$, $\sigma(r, X_i(r)) := \sigma(r, X_i(r), a(r))$, $i = 1, 2$. By the definition of a mild solution we have, for $s \in [t_i, T]$,

$$X_i(s) = e^{(s-t_i)A}x_i + \int_{t_i}^s e^{(s-r)A}b(r, X_i(r))dr + \int_{t_i}^s e^{(s-r)A}\sigma(r, X_i(r))dW(r),$$

hence

$$\begin{aligned}|X_1(s) - X_2(s)| &\le |e^{(s-t_1)A}x_1 - e^{(s-t_2)A}x_2| \\ &+ \left| \int_{t_1}^{t_2} e^{(s-r)A}b(r, X_1(r))dr \right| + \left| \int_{t_2}^s e^{(s-r)A} \left(b(r, X_1(r)) - b(r, X_2(r))\right) dr \right| \\ &+ \left| \int_{t_1}^{t_2} e^{(s-r)A}\sigma(r, X_1(r))dW(r) \right| + \left| \int_{t_2}^s e^{(s-r)A} \left(\sigma(r, X_1(r)) - \sigma(r, X_2(r))\right) dW(r) \right|.\end{aligned}$$

Therefore

$$
\begin{aligned}
&\mathbb{E}|X_1(s) - X_2(s)|^2 \le 5|e^{(s-t_1)A}x_1 - e^{(s-t_2)A}x_2|^2 \\
&+ 5\mathbb{E}\left|\int_{t_1}^{t_2} e^{(s-r)A} b(r, X_1(r))dr\right|^2 + 5\mathbb{E}\left|\int_{t_2}^{s} e^{(s-r)A} \left(b(r, X_1(r)) - b(r, X_2(r))\right) dr\right|^2 \\
&\qquad + 5\mathbb{E}\left|\int_{t_1}^{t_2} e^{(s-r)A}\sigma(r, X_1(r))dW(r)\right|^2 \\
&\qquad + 5\mathbb{E}\left|\int_{t_2}^{s} e^{(s-r)A} \left(\sigma(r, X_1(r)) - \sigma(r, X_2(r))\right) dW(r)\right|^2 .
\end{aligned}
\tag{1.84}
$$

To estimate the second and the third terms we use Jensen's inequality applied to the inner integral. Using Hypothesis 1.149-(ii) and (1.80) we then obtain

$$
\begin{aligned}
\mathbb{E}\left|\int_{t_1}^{t_2} e^{(s-r)A} b(r, X_1(r))dr\right|^2 &\le L^2\mathbb{E}\left|\int_{t_1}^{t_2} (s-r)^{-\gamma_1}(1 + |X_1(r)|)dr\right|^2 \\
&\le L^2 \left(\int_{t_1}^{t_2} (s-r)^{-\gamma_1} dr\right) \int_{t_1}^{t_2} (s-r)^{-\gamma_1}\mathbb{E}(1 + |X_1(r)|)^2 dr \\
&\le 2L^2[1 + C(1 + |x_1|^2)] \left(\int_{t_1}^{t_2} (s-r)^{-\gamma_1} dr\right)^2 \\
&\le 2L^2[1 + C(1 + |x_1|)^2)] \frac{1}{1-\gamma_1}(t_1 - t_2)^{2(1-\gamma_1)}.
\end{aligned}
$$

In the same way we estimate the third term obtaining, by Hypothesis 1.149-(ii),

$$
\begin{aligned}
&\mathbb{E}\left|\int_{t_2}^{s} e^{(s-r)A} \left(b(r, X_1(r)) - b(r, X_2(r))\right) dr\right|^2 \\
&\qquad \le L^2 \left(\int_{t_2}^{s} (s-r)^{-\gamma_1} dr\right) \int_{t_2}^{s} (s-r)^{-\gamma_1}\mathbb{E}|X_1(r) - X_2(r)|^2 dr \\
&\qquad \le \frac{L^2(s-t_2)^{1-\gamma_1}}{1-\gamma_1} \int_{t_2}^{s} (s-r)^{-\gamma_1}\mathbb{E}|X_1(r) - X_2(r)|^2 dr.
\end{aligned}
$$

The fourth and the fifth term of (1.84) are estimated using the isometry formula. We have

$$
\mathbb{E}\left|\int_{t_1}^{t_2} e^{(s-r)A}\sigma(r, X_1(r))dW(r)\right|^2 = \int_{t_1}^{t_2} \mathbb{E}|e^{(s-r)A}\sigma(r, X_1(r))|^2_{\mathcal{L}_2(\Xi,H)} dr
$$

$$
\le L^2 \int_{t_1}^{t_2} (s-r)^{-2\gamma_2}\mathbb{E}(1 + |X_1(r)|)^2 dr \le 2L^2[1 + C(1 + |x_1|^2)] \int_{t_1}^{t_2} (s-r)^{-2\gamma_2} dr
$$

$$\leq 2L^2[1 + C(1 + |x_1|^2)] \frac{1}{1 - 2\gamma_2} (t_1 - t_2)^{1-2\gamma_2}$$

and

$$\mathbb{E} \left| \int_{t_2}^{s} e^{(s-r)A} \left(\sigma(r, X_1(r)) - \sigma(r, X_2(r))\right) dW(r) \right|^2$$
$$= \int_{t_2}^{s} \mathbb{E} |e^{(s-r)A} \left(\sigma(r, X_1(r)) - \sigma(r, X_2(r))\right|^2_{\mathcal{L}_2(\Xi, H)} dr$$
$$\leq L^2 \int_{t_2}^{s} (s-r)^{-2\gamma_2} \mathbb{E}|X_1(r) - X_2(r)|^2 dr.$$

Using all these estimates in (1.84) we obtain, for a suitable constant $C_1 > 0$, for $\gamma_3 := [2(1 - \gamma_1)] \wedge [1 - 2\gamma_2]$ and $\gamma_4 := \gamma_1 \vee [2\gamma_2]$,

$$\mathbb{E}|X_1(s) - X_2(s)|^2 \leq 5|e^{(s-t_1)A}x_1 - e^{(s-t_2)A}x_2|^2 + C_1(1 + |x_1|^2)|t_2 - t_1|^{\gamma_3} +$$
$$+ C_1 \int_{t_2}^{s} (s-r)^{-\gamma_4} \mathbb{E}|X_1(r) - X_2(r)|^2 dr.$$

Observing that

$$|e^{(s-t_1)A}x_1 - e^{(s-t_2)A}x_2| \leq M|x_1 - x_2| + |e^{(s-t_2)A}(e^{(t_2-t_1)A}x_1 - x_1)|,$$

we can thus apply Gronwall's lemma in the form of Proposition D.30. It gives us

$$\mathbb{E}|X_1(s) - X_2(s)|^2 \leq C_2 \left[|x_1 - x_2|^2 + (1 + |x_1|^2)|t_2 - t_1|^{\gamma_3} + |e^{(t_2-t_1)A}x_1 - x_1|^2 \right]$$

for some $C_2 > 0$ with the required properties. □

Lemma 1.154 *Assume that Hypothesis 1.149 holds. Fix a Λ-valued progressively measurable process $a(\cdot)$. Let X be the unique mild solution of* (1.74) *described in Theorem 1.152 with initial condition $X(0) = x \in H$. Define, for $s \in [0, T]$, $\psi(s) = b(s, X(s), a(s))$, $\Phi(s) = \sigma(s, X(s), a(s))$. Let $\{e_i\}_{i \in \mathbb{N}}$ be an orthonormal basis of Ξ and, for any $k \in \mathbb{N}$, let $P^k : \Xi \to \Xi$ be the orthogonal projection onto* span$\{e_1, ..., e_k\}$. *Let X^k be the unique mild solution of*

$$\begin{cases} dX^k(s) = (AX^k(s) + e^{\frac{1}{k}A}\psi(s))ds + e^{\frac{1}{k}A}\Phi(s)P^k dW(s), \\ X^k(0) = x. \end{cases} \tag{1.85}$$

Then, for any $p > 0$, there exists an $M_p > 0$ such that

$$\sup_{k \in \mathbb{N}} \mathbb{E} \left[\sup_{s \in [0,T]} |X^k(s)|^p \right] \leq M_p. \tag{1.86}$$

Moreover, for every $s \in [0, T]$,

$$\lim_{k\to\infty} \mathbb{E}\left[|X^k(s) - X(s)|^2\right] = 0 \tag{1.87}$$

and, for every $\varphi \in C_m(H)$ $(m \geq 0)$,

$$\lim_{k\to\infty} \mathbb{E}\left[\varphi(X^k(s))\right] = \mathbb{E}\left[\varphi(X(s))\right], \qquad s \in [0, T]. \tag{1.88}$$

Proof It is easy to see, by using (1.80), that (1.86) is satisfied.

We now prove (1.87). We have, for $s \in [0, T]$,

$$\begin{aligned}\mathbb{E}\left|X(s) - X^k(s)\right|^2 \leq 2\mathbb{E}&\left|\int_0^s e^{(s-r)A}\left(\psi(r) - e^{\frac{1}{k}A}\psi(r)\right)dr\right|^2 \\ &+ 4\mathbb{E}\left|\int_0^s e^{(s-r)A}\Phi(r)(I - P^k)dW(r)\right|^2 \\ &+ 4\mathbb{E}\left|\int_0^s (I - e^{\frac{1}{k}A})e^{(s-r)A}\Phi(r)P^k dW(r)\right|^2 = I_1 + I_2 + I_3.\end{aligned}$$

We have for any k,

$$\left|e^{(s-r)A}\left(\psi(r) - e^{\frac{1}{k}A}\psi(r)\right)\right| \leq 2L(s-r)^{-\gamma_1}(1 + |X(r)|)$$

which is integrable on $[0, s]$ for a.e. ω. Moreover,

$$\left|e^{(s-r)A}\left(\psi(r) - e^{\frac{1}{k}A}\psi(r)\right)\right| \to 0 \quad \text{as } k \to +\infty$$

$dr \otimes \mathbb{P}$-a.s. Therefore it follows from the dominated convergence theorem that

$$\int_0^s e^{(s-r)A}\left(\psi(r) - e^{\frac{1}{k}A}\psi(r)\right)dr \to 0 \quad \text{as } k \to +\infty$$

$\mathbb{P}$-a.s. Now by Hölder's inequality

$$\begin{aligned}&\left|\int_0^s e^{(s-r)A}\left(\psi(r) - e^{\frac{1}{k}A}\psi(r)\right)dr\right|^2 \\ &\qquad\leq 4L^2\left(\int_0^s (s-r)^{-\gamma_1}dr\right)\left(\int_0^s (s-r)^{-\gamma_1}(1 + |X(r)|)^2 dr\right)\end{aligned}$$

which is integrable on Ω. Thus, using the dominated convergence theorem again we conclude that $\lim_{k\to\infty} I_1 = 0$.

Recall that $\Xi_0 = \Xi$. To estimate I_2, we set $Q^k := I - P^k$. We have

$$I_2 = 4\mathbb{E}\left|\int_0^s e^{(s-r)A}\Phi(r)(I-P^k)dW(r)\right|^2$$
$$= 4\int_0^s \mathbb{E}\left\|e^{(s-r)A}\Phi(r)Q^k\right\|^2_{\mathcal{L}_2(\Xi,H)}dr$$
$$= 4\int_0^s \mathbb{E}\sum_{i\in\mathbb{N}}\left\langle e^{(s-r)A}\Phi(r)Q^k e_i, e^{(s-r)A}\Phi(r)Q^k e_i\right\rangle dr =: \eta(k).$$

Observe that

$$\sum_{i\in\mathbb{N}}\left\langle e^{(s-r)A}\Phi(r)Q^k e_i, e^{(s-r)A}\Phi(r)Q^k e_i\right\rangle$$
$$= \sum_{i=k+1}^{+\infty}\left\langle e^{(s-r)A}\Phi(r)e_i, e^{(s-r)A}\Phi(r)e_i\right\rangle$$
$$\le \sum_{i\in\mathbb{N}}\left\langle e^{(s-r)A}\Phi(r)e_i, e^{(s-r)A}\Phi(r)e_i\right\rangle = \left\|e^{(s-r)A}\Phi(r)\right\|^2_{\mathcal{L}_2(\Xi,H)}.$$

Since the series above has nonnegative terms, we obtain

$$\lim_{k\to\infty}\left\|e^{(s-r)A}\Phi(r)Q^k\right\|^2_{\mathcal{L}_2(\Xi,H)} = 0 \quad dr\otimes\mathbb{P}\text{-a.s.}$$

Therefore, thanks to (1.80), Hypothesis 1.149 and the dominated convergence theorem, we obtain

$$\lim_{k\to\infty} I_2 \le \lim_{k\to\infty}\eta(k) = 0.$$

The term I_3 is estimated similarly.

Thanks to (1.87), for any subsequence of $X^k(s)$ we can extract a sub-subsequence converging to $X(s)$ almost everywhere and then, thanks to (1.86), (1.80) and the dominated convergence theorem, we obtain (1.88) along the sub-subsequence. This implies (1.88) for the whole sequence $X^k(s)$. □

Remark 1.155 Observe that if b and σ satisfy Hypothesis 1.149, the functions $e^{\frac{1}{k}A}b(s,x,a)$ and $e^{\frac{1}{k}A}\sigma(s,x,a)P_k$ satisfy Hypothesis 1.125. ■

The last lemma concerns the additive noise case of Sect. 1.5.2, however we included it here since its proof is similar to the proof of Lemma 1.154.

Let W_Q be from Sect. 1.5.2. We know (see (1.12)) that $W_Q(s) = \sum_{n=1}^{+\infty} g_n\beta_n(s)$, $s\ge 0$, where $\{g_n\}$ is an orthonormal basis of Ξ_0. Define $e_n = Q^{-1/2}g_n$, $n\in\mathbb{N}$. Then $\{e_n\}$ is an orthonormal basis of Ξ. Let $\tilde{P}^k$ be the orthogonal projection in Ξ_0 onto $\mathrm{span}\{g_1,\ldots,g_k\}$ and P^k be the orthogonal projection in Ξ onto $\mathrm{span}\{e_1,\ldots,e_k\}$, $k\in\mathbb{N}$. It is easy to see that $\tilde{P}^kQ^{1/2} = Q^{1/2}P^k$ as operators on Ξ.

Lemma 1.156 *Let Hypotheses 1.143 and 1.145 be satisfied and let $q\ge 2$. Let X be the unique mild solution of (1.67) described in Proposition 1.147 with initial condition $X(0)=x\in H$. Define for $k,m\in\mathbb{N}$, $B_k = \{(s,\omega) : |b_0(s,X(s),a_1(s))|\le k\}$,*

$D_m = \{(s, \omega) : |g(s, \omega)| \le m\}$. *There exists a sequence* m_k *such that the sequence* X^k *of the solutions of the SDE*

$$\begin{cases} dX^k(s) = \left(AX^k(s) + \psi_k(s)\right) ds + \sigma \tilde{P}^k dW_Q(s), \quad s > 0, \\ X^k(0) = x, \end{cases} \tag{1.89}$$

where $\psi_k(s) = b_0(s, X(s), a_1(s))\mathbf{1}_{B_k}(s, \omega) + e^{\frac{1}{k}A} a_2(s)\mathbf{1}_{D_{m_k}}(s, \omega)$, *satisfies the following.*

(i) For any $p \in [2, q]$ *there exists an* $M_p > 0$ *such that*

$$\sup_k \sup_{s\in[0,T]} \mathbb{E}\left[|X^k(s)|^p\right], \sup_{s\in[0,T]} \mathbb{E}\left[|X(s)|^p\right] \le M_p. \tag{1.90}$$

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

$$\lim_{k\to\infty} \mathbb{E}\left[|X^k(s) - X(s)|^2\right] = 0.$$

Proof Part (i). The moment estimates are uniform in k (regardless of the choice of m_k) thanks to the following facts:

(a) Define $W^{A,k}(s) := \int_0^s e^{(s-r)A}\sigma \tilde{P}^k dW_Q(r)$, $s \in [0, T]$. Given an orthonormal basis $\{w_n\}$ of H, for any $k \in \mathbb{N}$ and $s \in [0, T]$, we have

$$\begin{aligned} 0 &\le \mathrm{Tr}\left(\left(e^{sA}\sigma \tilde{P}^k Q^{1/2}\right)\left(e^{sA}\sigma \tilde{P}^k Q^{1/2}\right)^*\right) \\ &= \mathrm{Tr}\left(\left(e^{sA}\sigma Q^{1/2} P^k\right)\left(e^{sA}\sigma Q^{1/2} P^k\right)^*\right) \\ &= \sum_{n\in\mathbb{N}} |P_k Q^{1/2}\sigma^* e^{sA^*} w_n|^2 \le \sum_{n\in\mathbb{N}} |Q^{1/2}\sigma^* e^{sA^*} w_n|^2 = \sum_{n\in\mathbb{N}} \mathrm{Tr}\left(e^{sA}\sigma Q\sigma^* e^{sA^*}\right). \end{aligned} \tag{1.91}$$

Thus, by Theorem 1.111, it follows that for any $k \in \mathbb{N}$ and $p \ge 1$,

$$\sup_k \sup_{s\in[0,T]} \mathbb{E}\left[|W^{A,k}(s)|^p\right] < +\infty.$$

Using (1.91) we also have, by the Lebesgue dominated convergence theorem,

$$\int_0^T \|e^{sA}\sigma \tilde{P}^k - e^{sA}\sigma\|^2_{\mathcal{L}_2(\Xi_0,H)} ds = \int_0^T \sum_{n\in\mathbb{N}} |(P_k - I)Q^{1/2}\sigma^* e^{sA^*} w_n|^2 ds \to 0. \tag{1.92}$$

(b) By the definition

$$|e^{tA}\psi_k(s)| \le f(s)(1 + |X(s)|) + t^{-\beta} g(s, \omega) \quad \text{for } t, s \in [0, T], \omega \in \Omega.$$

Part (ii). The scheme of the proof is similar to that of (1.87). We choose m_k such that

$$\mathbb{E}\left|\int_0^T k^\beta g(r,\omega)|1-\mathbf{1}_{D_{m_k}}(r,\omega)|dr\right|^2 \le \frac{1}{k}. \tag{1.93}$$

We have for every $s \in [0,T]$,

$$\begin{aligned}\mathbb{E}\left|X(s)-X^k(s)\right|^2 \le 4\mathbb{E}\left|\int_0^s e^{(s-r)A} b_0(r,X(r),a_1(r))(1-\mathbf{1}_{B_k}(r,\omega))dr\right|^2 \\ +4\mathbb{E}\left|\int_0^s e^{(s-r)A}(a_2(r)-e^{\frac{1}{k}A}a_2(r))dr\right|^2 \\ +4\mathbb{E}\left|\int_0^s e^{(\frac{1}{k}+s-r)A}a_2(r)(1-\mathbf{1}_{D_{m_k}}(r,\omega))dr\right|^2 \\ +4\mathbb{E}\left|W^{A,k}(s)-W^A(s)\right|^2 = J_1+J_2+J_3+J_4.\end{aligned}$$

The term J_1 converges to 0 as $k\to+\infty$ by Hypothesis 1.145, Hölder's inequality, (1.69) for $p=2$ and the dominated convergence theorem. The term J_2 converges to 0 by the same arguments as for the term I_1 in the proof of Lemma 1.154. The term J_3 converges to 0 by (1.93) and finally $J_4\to 0$ by (1.92). □

1.6 Transition Semigroups

Let $T\in(0,+\infty]$ and recall that, as before, when $T=+\infty$ the notation $[0,T]$ and $[t,T]$ means $[0,+\infty)$ and $[t,+\infty)$. Let H, Ξ, Q, and the generalized reference probability space $\mu=(\Omega,\mathscr{F},\{\mathscr{F}_s\}_{s\in[0,T]},\mathbb{P},W_Q)$ be the same as in Sect. 1.3. Consider for $t\in[0,T]$ the following SDE with non-random coefficients

$$\begin{cases} dX(s)=(AX(s)+b(s,X(s)))\,ds+\sigma(s,X(s))dW_Q(s) \\ X(t)=x\in H, \end{cases} \tag{1.94}$$

where $b\colon[0,T]\times H\to H$ and $\sigma\colon[0,T]\times H\to\mathcal{L}_2(\Xi_0,H)$. If Hypothesis 1.125, where we drop the dependence on a in all conditions, (respectively, Hypotheses 1.143 and 1.145 with $a_2(\cdot)\equiv 0$ and with no dependence on a_1, respectively, Hypothesis 1.149 with no dependence on a) is satisfied, then Theorem 1.127 (respectively, Proposition 1.147, respectively, Theorem 1.152) ensures that (1.94) has a unique mild solution $X(\cdot;t,x)$. Moreover, we also have uniqueness in law of the solutions.

We will be using the spaces $B_b(H)$ of bounded Borel measurable functions on H and $B_m(H)$, $m>0$, of Borel measurable functions on H with at most polynomial growth of order m, defined in Appendix A.2.

For any $\phi\in B_b(H)$ and $t\ge 0$, $s\in[t,T]$, we define

$$\begin{cases} P_{t,s}[\phi]\colon H \to \mathbb{R} \\ P_{t,s}[\phi]\colon x \to \mathbb{E}[\phi(X(s;t,x))]. \end{cases} \tag{1.95}$$

It is not obvious that $P_{t,s}[\phi] \in B_b(H)$ and it has to be checked in each case. The general argument is the following and we illustrate it in the case when Hypothesis 1.149 is satisfied. First, using (1.83) it is easy to see that $P_{t,s}[\phi] \in C_b(H)$ if $\phi \in UC_b(H)$. Then, using the functions constructed in the proof of Theorem 1.34 and the dominated convergence theorem, we get that $P_{t,s}[\phi] \in B_b(H)$ for every $\phi = \mathbf{1}_A$, $A = \overline{A} \subset H$. This, together with Corollary 1.3 and the dominated convergence theorem, allows us to extend $P_{t,s}[\phi] \in B_b(H)$ to every $\phi = \mathbf{1}_A$, $A \in B(H)$. We can then use Lemma 1.15 to conclude that $P_{t,s}[\phi] \in B_b(H)$ for every $\phi \in B_b(H)$. Similar arguments can be applied in the cases when Hypotheses 1.143 and 1.145 hold or if Hypothesis 1.125 is satisfied. Moreover, thanks to estimates (1.36), (1.69) and (1.80), $P_{t,s}[\phi]$ is then also well defined for any $\phi \in B_m(H)$, $m > 0$.

Theorem 1.157 (Markov property) *Let $T \in (0, +\infty]$. Let Hypothesis 1.149 be satisfied with b and σ independent of a. Then for every $\phi \in B_m(H)$ $(m \geq 0)$ and $0 \leq t \leq s \leq r \leq T$ (with the last inequality strict when $T = +\infty$),*

$$\mathbb{E}\phi(X(r;t,x)|\mathscr{F}_s) = P_{s,r}[\phi](X(s;t,x)) \quad \mathbb{P} - \textit{almost surely},$$

and

$$P_{t,r}[\phi](x) = P_{t,s}\left[P_{s,r}[\phi]\right](x) \qquad \textit{for all } x \in H. \tag{1.96}$$

The same result is true if Hypotheses 1.143 and 1.145 hold without dependence on a_1 and with $a_2(\cdot) = 0$ or if Hypothesis 1.125 holds without the dependence on a in all conditions.

Proof See [180], Theorem 9.14, p. 248, and Corollary 9.15, p. 249. The hypotheses are a little different from these in [180], however the same arguments can be easily adapted using the proof of Proposition 1.153. The proof in [180] is given for $\phi \in B_b(H)$ but the argument is exactly the same when $\phi \in B_m(H)$ $(m > 0)$ simply recalling that the operator $P_{t,s}$ is well defined on such functions thanks to estimate (1.80). □

It follows from the uniqueness in law of the solutions of (1.94) that the operators $P_{t,s}$ do not depend on the choice of a generalized reference probability space μ. As a consequence of the uniqueness in law we also have the following corollary.

Corollary 1.158 *Let Hypothesis 1.149 be satisfied with b and σ independent of a and of the time variable s. Equation (1.94) then reduces to*

$$\begin{cases} dX(s) = (AX(s) + b(X(s)))\,ds + \sigma(X(s))dW_Q(s), \\ X(t) = x \in H. \end{cases} \tag{1.97}$$

Denote by $X(\cdot;t,x)$ the unique mild solution of this equation (defined on $[t,+\infty)$). In this case, for any $\phi \in B_m(H)$ $(m \geq 0)$ and $0 \leq t \leq s$, we have

$$P_{t,s}[\phi](x) = P_{0,s-t}[\phi]. \tag{1.98}$$

Hence, defining $P_s[\phi]$ as follows,

$$\begin{cases} P_s[\phi]\colon H \to \mathbb{R} \\ P_s[\phi]\colon x \to \mathbb{E}\phi(X(s;0,x)), \end{cases} \tag{1.99}$$

we have

$$P_{s+r}[\phi](x) = P_s\left[P_r[\phi]\right](x) \quad \text{for all } x \in H, s, r \geq 0. \tag{1.100}$$

The same result is true if Hypotheses 1.143 and 1.145 hold without dependence on a_1 and with $a_2(\cdot) = 0$ or if Hypothesis 1.125 holds without the dependence on a in all conditions.

Proof We only need to prove (1.98), which is an immediate consequence of the uniqueness in law of the mild solutions of (1.97). Indeed, by the uniqueness in law, for all $s \geq t \geq 0$ and $x \in H$, the random variables $X(s;t,x)$ and $X(s-t;0,x)$ have the same distributions, hence

$$P_{t,s}[\phi](x) = \mathbb{E}[\phi(X(s;t,x))] = \mathbb{E}[\phi(X(s-t;0,x))] = P_{0,s-t}[\phi](x). \qquad \square$$

Definition 1.159 (*Transition semigroup, (strong) Feller property*) If (1.96) (respectively, (1.100)) is satisfied we call $P_{t,s}$ (respectively, P_t) the two-parameter *transition semigroup* (respectively, one-parameter *transition semigroup*) associated to Eq. (1.94).

We say that $P_{t,s}$ (respectively, P_t) possesses the *Feller property* if

$$P_{t,s}(C_b(H)) \subset C_b(H) \text{ (respectively, } P_t(C_b(H)) \subset C_b(H))$$

and that $P_{t,s}$ (respectively, P_t) possesses the *strong Feller property* if

$$P_{t,s}(B_b(H)) \subset C_b(H) \text{ (respectively, } P_t(B_b(H)) \subset C_b(H))$$

for all $0 \leq t < s \leq T$ (respectively $t \in (0,T]$).

Lemma 1.160 *Assume that (1.94) has unique mild solutions $X(\cdot;t,x)$ which satisfy, for every $m \geq 0$, the estimate*

$$\mathbb{E}[|X(s;t,x)|^m] \leq C(m)(1+|x|^m), \qquad t \geq 0, \ s \in [t,T], \ x \in H, \tag{1.101}$$

for some constant $C(m)$. If the Feller property holds for the associated two-parameter transition semigroup $P_{t,s}$ ($t \geq 0$, $s \in [t,T]$), then we also have

$$P_{t,s}(C_m(H)) \subset C_m(H) \qquad \forall m \geq 0$$

while, if the strong Feller property holds, we also have

$$P_{t,s}(B_m(H)) \subset C_m(H) \qquad \forall m \geq 0.$$

Proof Let $\phi \in B_m(H)$ and define, for $k \in \mathbb{N}$,

$$\phi_k(x) = \phi(x)\mathbf{1}_{|x|\leq k} + \phi\left(k\frac{x}{|x|}\right)\mathbf{1}_{|x|>k}.$$

It is clear that $\phi_k \in B_b(H)$, it coincides with ϕ on $\{|x| \leq k\}$ and if ϕ is continuous so is ϕ_k. Moreover, when $k \to +\infty$, ϕ_k converges to ϕ uniformly on bounded sets. Assume now that the strong Feller property holds (the argument for the Feller property is exactly the same). In this case $P_{t,s}[\phi_k]$ is continuous, hence, to get the claim, it is enough to show that $P_{t,s}[\phi_k]$ converges to $P_{t,s}[\phi_k]$ uniformly on bounded sets. Indeed,

$$\begin{aligned} P_{t,s}[\phi_k - \phi](x) &= \mathbb{E}\left[(\phi_k - \phi)(X(s;t,x))\right] \\ &= \mathbb{E}\left[\left(\phi\left(k\frac{X(s;t,x)}{|X(s;t,x)|}\right) - \phi(X(s;t,x))\right)\mathbf{1}_{|X(s;t,x)|>k}\right] \\ &\leq 2\mathbb{E}\left[\|\phi\|_{B_m}(1+|X(s;t,x)|^m)\mathbf{1}_{|X(s;t,x)|\geq k}\right]. \end{aligned}$$

Hence, for any $p > 1$ we have by (1.101)

$$\begin{aligned} P_{t,s}[\phi_k - \phi](x) &\leq 2\|\phi\|_{B_m}\left[\mathbb{E}(1+|X(s;t,x)|^m)^p\right]^{1/p}\left[\mathbb{E}\mathbf{1}_{|X(s;t,x)|\geq k}\right]^{1-1/p} \\ &\leq C(1+|x|^m)\left[\frac{\mathbb{E}|X(s;t,x)|}{k}\right]^{1-1/p} \leq C(1+|x|^m)\left[\frac{1+|x|}{k}\right]^{1-1/p} \end{aligned}$$

which converges to 0 uniformly on bounded sets. □

Remark 1.161 Estimate (1.101) is satisfied in two important cases:

- when Hypothesis 1.149 is satisfied with b and σ independent of a;
- when Hypotheses 1.143 and 1.145 hold without dependence on a_1 and with $a_2(\cdot) = 0$.

This follows from the growth estimates of Theorem 1.152 and Proposition 1.147. ■

Theorem 1.162 *Assume that Hypothesis 1.149 is satisfied. Then for every $\phi \in C_m(H)$ ($m \geq 0$), the function $P_{t,s}[\phi]\colon H \to \mathbb{R}$ belongs to $C_m(H)$. The same holds if we assume that Hypotheses 1.143 and 1.145 hold without dependence on a_1 and with $a_2(\cdot) = 0$.*

Proof The result is a consequence of the continuous dependence and growth estimates of Theorem 1.152 and Propositions 1.153 and 1.147. □

1.7 Itô's and Dynkin's Formulae

In this section we assume that $T > 0$, H, Ξ, Q, and the generalized reference probability space $\mu = (\Omega, \mathscr{F}, \{\mathscr{F}_s\}_{s\in[0,T]}, \mathbb{P}, W_Q)$ are the same as in Sect. 1.3. The operator A is the generator of a C_0-semigroup on H, and Λ is a Polish space. The various Itô's and Dynkin's formulae presented in this section are used in proving existence of viscosity solutions (Chap. 3) and verification theorems (Chaps. 4 and 5).

Given a function $F\colon [0,T] \times H \to \mathbb{R}$, we denote by F_t the derivative of $F(t,x)$ with respect to t and by DF and D^2F the first and second-order Fréchet derivatives with respect to x.

Theorem 1.163 (Itô's Formula) *Assume that Φ is a process in $\mathcal{N}_Q^2(0,T;H)$, f is an H-valued progressively measurable ($\mathbb{P}$-a.s.) Bochner integrable process on $[0,T]$, and define, for $s \in [0,T]$,*

$$X(s) := X(0) + \int_0^s f(r)dr + \int_0^s \Phi(r)dW_Q(r),$$

where $X(0)$ is an $\mathscr{F}_0$-measurable H-valued random variable. Consider $F\colon [0,T] \times H \to \mathbb{R}$ and assume that F and its derivatives F_t, DF, D^2F are continuous and bounded on bounded subsets of $[0,T] \times H$. Let τ be an $\mathscr{F}_s$-stopping time. Then, for $\mathbb{P}$-a.e. ω,

$$\begin{aligned} F(s\wedge\tau, X(s\wedge\tau)) &= F(0,X(0)) + \int_0^{s\wedge\tau} F_t(r,X(r))dr \\ &+ \int_0^{s\wedge\tau} \langle DF(r,X(r)), f(r)\rangle\, dr + \int_0^{s\wedge\tau} \big\langle DF(r,X(r)), \Phi(r)dW_Q(r)\big\rangle \\ &+ \frac{1}{2}\int_0^{s\wedge\tau} \mathrm{Tr}\left[\left(\Phi(r)Q^{1/2}\right)\left(\Phi(r)Q^{1/2}\right)^* D^2F(r,X(r))\right]dr \quad \text{on } [0,T]. \end{aligned} \tag{1.102}$$

Proof See [294], Theorems 2.9 and 2.10. See also, under the assumption of uniform continuity on bounded sets of F and its derivatives, [180] Theorem 4.32, p. 106. □

Proposition 1.164 *Let $F\colon [0,T] \times H \to \mathbb{R}$ and $x \in H$. Assume that F and its derivatives F_t, DF, D^2F are continuous and bounded on bounded subsets of $[0,T] \times H$. Suppose that $DF\colon [0,T] \times H \to D(A^*)$ and that A^*DF is continuous and bounded on bounded subsets of $[0,T] \times H$. Let $f \in M_\mu^p(0,T;H)$, $\Phi \in \mathcal{N}_Q^p(0,T;H)$ for some $p > 2$. Let $X(\cdot)$ be the unique mild solution of (1.42) such that $X(0) = x$ and τ be an $\mathscr{F}_s$-stopping time. Then, for $\mathbb{P}$-a.e. ω,*

$$
\begin{aligned}
F(s\wedge\tau, X(s\wedge\tau)) &= F(0,x) + \int_0^{s\wedge\tau} F_t(r, X(r))dr \\
&+ \int_0^{s\wedge\tau} \left\langle A^* DF(r, X(r)), X(r)\right\rangle dr + \int_0^{s\wedge\tau} \langle DF(r, X(r)), f(r)\rangle\, dr \\
&+ \frac{1}{2}\int_0^{s\wedge\tau} \mathrm{Tr}\left[\left(\Phi(r)Q^{1/2}\right)\left(\Phi(r)Q^{1/2}\right)^* D^2F(r, X(r))\right] dr \\
&+ \int_0^{s\wedge\tau} \left\langle DF(r, X(r)), \Phi(r)dW_Q(r)\right\rangle \qquad on\ [0,T].
\end{aligned}
\tag{1.103}
$$

Proof Since both sides of (1.103) are continuous processes, it is enough to prove the formula for a single s. We approximate $X(\cdot)$ by the sequence $X^n(\cdot)$ introduced in Proposition 1.132. By definition $X^n(\cdot)$ solves the integral equation

$$
X^n(s) = \int_0^s \left(A_n X^n(r) + f(r)\right) dr + \int_0^s \Phi(r)dW_Q(r).
$$

For any $R > 0$ such that $|x| < R$ define the stopping times

$$
\hat{\tau}^R := \inf\{s\in[0,T] \ :\ |X(s)| > R\}, \quad \hat{\tau}^R_n := \inf\{s\in[0,T] \ :\ |X_n(s)| > R+1\}
$$

and denote by τ^R and τ^R_n, respectively,

$$
\tau^R := \min(\tau, \hat{\tau}^R), \quad \tau^R_n := \min(\tau, \hat{\tau}^R, \hat{\tau}^R_n).
$$

Observe that, thanks to (1.44), up to extracting a subsequence of X_n (still denoted by X_n), $\sup_{s\in[0,T]}|X^n(s) - X(s)|^p$ converges to 0 on some set $\tilde{\Omega}$ with $\mathbb{P}(\tilde{\Omega}) = 1$. It is then easy to see that on $\tilde{\Omega}$ we have

$$
\lim_{n\to\infty} \tau^R_n = \tau^R.
$$

We deduce that, for $\omega\in\tilde{\Omega}$,

$$
\lim_{n\to\infty} \mathbf{1}_{[0,s\wedge\tau^R_n]} = \mathbf{1}_{[0,s\wedge\tau^R]}, \quad \text{pointwise on } [0,T]. \tag{1.104}
$$

We can apply Itô's formula (1.102) to the approximating problem (A_n^* is the adjoint of A_n) obtaining, once we rewrite it using Lemma 1.110,

$$
\begin{aligned}
F(s\wedge\tau^R_n, X^n(s\wedge\tau^R_n)) &= F(0,x) + \int_0^s \mathbf{1}_{[0,s\wedge\tau^R_n]}(r) F_t(r, X^n(r))dr \\
&+ \int_0^s \mathbf{1}_{[0,s\wedge\tau^R_n]}(r)\left\langle A_n^* DF(r, X^n(r)), X^n(r)\right\rangle dr
\end{aligned}
$$

$$+ \int_0^s \mathbf{1}_{[0,s\wedge\tau_n^R]}(r) \left\langle DF(r, X^n(r)), f(r) \right\rangle dr$$
$$+ \frac{1}{2} \int_0^s \mathbf{1}_{[0,s\wedge\tau_n^R]}(r) \mathrm{Tr}\left[\left(\Phi(r) Q^{1/2}\right) \left(\Phi(r) Q^{1/2}\right)^* D^2 F(r, X^n(r)) \right] dr$$
$$+ \int_0^s \mathbf{1}_{[0,s\wedge\tau_n^R]}(r) \left\langle DF(r, X^n(r)), \Phi(r) dW_Q(r) \right\rangle . \quad (1.105)$$

By the local boundedness of F and its derivatives, it follows that for $\mathbb{P}$-a.e. ω all the integrands of the deterministic integrals in (1.105) are dominated for $n \in \mathbb{N}$ by integrable functions. Regarding the term containing $A_n^* DF(r, X^n(r))$, recall from (B.11) that $A_n = J_n A$ are uniformly bounded as linear operators from $D(A)$ (endowed with the graph norm) to H. Moreover, thanks to (1.104), (1.44) and the continuity of F and its derivatives, we know that these integrands converge to the corresponding ones in (1.103) (with τ_R instead of τ) on $[0, s]$, $\mathbb{P}$-a.s. We can thus conclude, by using the Lebesgue dominated convergence theorem, that the deterministic integrals in (1.105) converge to their counterparts in (1.103).

To justify the convergence of the stochastic integral we observe that, with

$$I_n := \int_0^s \mathbf{1}_{[0,s\wedge\tau_n^R]}(r) \left\langle DF(r, X^n(r)), \Phi(r) dW_Q(r) \right\rangle ,$$

$$I := \int_0^s \mathbf{1}_{[0,s\wedge\tau^R]}(r) \left\langle DF(r, X(r)), \Phi(r) dW_Q(r) \right\rangle ,$$

we have

$$\mathbb{E}\, |I_n - I|^2$$
$$\leq \int_0^s \mathbb{E} \|\Phi(r)\|^2_{\mathcal{L}_2(\Xi_0, H)} \left| \mathbf{1}_{[0,s\wedge\tau_n^R]}(r) DF(r, X^n(r)) - \mathbf{1}_{[0,s\wedge\tau^R]}(r) DF(r, X(r)) \right|^2 dr \to 0$$

as $n \to +\infty$ by the dominated convergence theorem. Therefore, up to a subsequence, we have $\lim_{n\to+\infty} I_n = I$, $\mathbb{P}$-a.s.

It now remains to let $R \to +\infty$ to obtain the claim. □

Proposition 1.165 *Let b and σ satisfy Hypothesis 1.125 and let $a\colon [t, T] \to \Lambda$ be a progressively measurable process. Let $X(\cdot)$ be the unique mild solution of (1.30) such that $X(0) = x \in H$. Consider $F\colon [0, T] \times H \to \mathbb{R}$. Assume that F and its derivatives F_t, DF, D^2F are continuous on $[0, T] \times H$. Suppose that $DF\colon [0, T] \times H \to D(A^*)$ and that $A^* DF$ is continuous on $[0, T] \times H$. Moreover, suppose that there exist $C \geq 0$, $N \geq 0$ such that*

$$|F(s, x)| + |DF(s, x)| + |F_t(s, x)| + \|D^2 F(s, x)\|$$
$$+ |A^* DF(s, x)| \leq C(1 + |x|)^N \quad (1.106)$$

for all $x \in H$, $s \in [0, T]$. Let τ be an $\mathscr{F}_s$-stopping time. Then:

(i) For $\mathbb{P}$-a.e. ω,

$$
\begin{aligned}
F(s\wedge\tau, X(s\wedge\tau)) &= F(0,x) + \int_0^{s\wedge\tau} F_t(r, X(r))dr \\
&+ \int_0^{s\wedge\tau} \left\langle A^* DF(r, X(r)), X(r)\right\rangle dr + \int_0^{s\wedge\tau} \langle DF(r, X(r)), b(r, X(r), a(r))\rangle\, dr \\
&+ \frac{1}{2}\int_0^{s\wedge\tau} \mathrm{Tr}\left[\left(\sigma(r, X(r), a(r))Q^{1/2}\right)\left(\sigma(r, X(r), a(r))Q^{1/2}\right)^* D^2F(r, X(r))\right] dr \\
&+ \int_0^{s\wedge\tau} \left\langle DF(r, X(r)), \sigma(r, X(r), a(r))dW_Q(r)\right\rangle \quad \textit{on } [0,T]. \qquad (1.107)
\end{aligned}
$$

(ii) Let η be a real process solving

$$
\begin{cases} d\eta(s) = \tilde{b}(s)ds \\ \eta(0) = \eta_0 \in \mathbb{R}, \end{cases}
$$

where $\tilde{b}\colon [0,T] \to \mathbb{R}$ is bounded and progressively measurable. Then, for $\mathbb{P}$-a.e. ω,

$$
\begin{aligned}
F(s\wedge\tau, X(s\wedge\tau))\eta(s\wedge\tau) &= F(0,x)\eta_0 + \int_0^{s\wedge\tau} (F_t(r, X(r))\eta(r) + F(r, X(r))\tilde{b}(r))dr \\
&+ \int_0^{s\wedge\tau} \left\langle A^* DF(r, X(r)), X(r)\right\rangle \eta(r)dr + \int_0^{s\wedge\tau} \langle DF(r, X(r)), b(r, X(r), a(r))\rangle\, \eta(r)dr \\
&+ \frac{1}{2}\int_0^{s\wedge\tau} \mathrm{Tr}\left[\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)^* D^2F(r, X(r))\right] \eta(r)dr \\
&+ \int_0^{s\wedge\tau} \left\langle DF(r, X(r))\eta(r), \sigma(r, X(r), a(r))dW_Q(r)\right\rangle \quad \textit{on } [0,T]. \qquad (1.108)
\end{aligned}
$$

In particular, for $s \in [0,T]$,

$$
\begin{aligned}
\mathbb{E}[F(s\wedge\tau, X(s\wedge\tau))\eta(s\wedge\tau)] &= F(0,x)\eta_0 + \mathbb{E}\int_0^{s\wedge\tau} (F_t(r, X(r))\eta(r) + F(r, X(r))\tilde{b}(r))dr \\
&+ \mathbb{E}\int_0^{s\wedge\tau} \left\langle A^* DF(r, X(r)), X(r)\right\rangle \eta(r)dr + \mathbb{E}\int_0^{s\wedge\tau} \langle DF(r, X(r)), b(r, X(r), a(r))\rangle\, \eta(r)dr \\
&+ \frac{1}{2}\mathbb{E}\int_0^{s\wedge\tau} \mathrm{Tr}\left[\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)^* D^2F(r, X(r))\right] \eta(r)dr.
\end{aligned} \qquad (1.109)
$$

Proof Part (i) follows directly from Proposition 1.164 applied with $f(s) := b(s, a(s), X(s))$ and $\Phi(s) := \sigma(s, a(s), X(s))$, $s \in [0,T]$, by noticing that, thanks to (1.33), (1.34) and (1.37), we have $f \in M^p_\mu(0,T;H)$ and $\Phi \in \mathcal{N}^p_Q(0,T;H)$ for every $p \geq 1$.

Part (ii) is a corollary of (i). We introduce the Hilbert space $\hat{H} := H \times \mathbb{R}$ (with the usual inner product), and set

$$
\hat{A} = \begin{pmatrix} A \\ 0 \end{pmatrix}, \quad \hat{b} = \begin{pmatrix} b \\ \tilde{b} \end{pmatrix}, \quad \hat{\sigma} = \begin{pmatrix} \sigma & 0 \\ 0 & 0 \end{pmatrix}.
$$

Then the process

$$\hat{X}(s) = \begin{pmatrix} X(s) \\ \eta(s) \end{pmatrix}$$

is the mild solution of the SDE

$$\begin{cases} d\hat{X}(s) = \left(\hat{A}\hat{X}(s) + \hat{b}(s, \hat{X}(s), a(s))\right) ds + \hat{\sigma}(s, \hat{X}(s), a(s)) dW_Q(s) \\ \hat{X}(0) = \begin{pmatrix} x \\ \eta_0 \end{pmatrix}. \end{cases}$$

Therefore, (1.108) follows from (1.107) applied to the function $\hat{F}(s, \hat{x}) = F(s, x)\eta_0$, where $\hat{x} = (x, \eta_0)$. Taking expectation in (1.108) we obtain (1.109). □

Proposition 1.166 *Let Hypothesis 1.125 be satisfied and A be maximal dissipative. Let $a\colon [t, T] \to \Lambda$ be a progressively measurable process. Let $X(\cdot)$ be the unique mild solution of (1.30) such that $X(0) = x \in H$. Let $F \in C^{1,2}([0, T] \times H)$ be of the form $F(t, x) = \varphi(t, |x|)$ for some $\varphi(t, r) \in C^{1,2}([0, T] \times \mathbb{R})$, where $\varphi(t, \cdot)$ is even and non-decreasing on $[0, +\infty)$. Moreover, suppose that there exist $C \geq 0$, $N \geq 0$ such that*

$$|F(s, x)| + |DF(s, x)| + |F_t(s, x)| + \|D^2 F(s, x)\| \leq C(1 + |x|)^N \tag{1.110}$$

for all $x \in H$, $s \in [0, T]$. Let τ be an $\mathscr{F}_s$-stopping time. Then:

(i) For $\mathbb{P}$-a.e. ω,

$$\begin{aligned} F(s \wedge \tau, X(s \wedge \tau)) \leq F(0, x) + \int_0^{s\wedge\tau} \Big[F_t(r, X(r)) + \langle b(r, X(r), a(r)), DF(r, X(r))\rangle \\ + \frac{1}{2}\mathrm{Tr}\left[\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)^* D^2F(r, X(r))\right]\Big] dr \\ + \int_0^{s\wedge\tau} \left\langle DF(r, X(r)), b(r, X(r), a(r)) dW_Q(r)\right\rangle \quad \text{on } [0, T]. \end{aligned} \tag{1.111}$$

(ii) If η is as in part (ii) of Proposition 1.165 and η is positive then, for $\mathbb{P}$-a.e. ω,

$$\begin{aligned} F(s \wedge \tau, X(s \wedge \tau))\eta(s \wedge \tau) \leq F(0, x)\eta_0 + \int_0^{s\wedge\tau} (F_t(r, X(r))\eta(r) + F(r, X(r))\tilde{b}(r)) dr \\ + \int_0^{s\wedge\tau} \langle DF(r, X(r)), b(r, X(r), a(r))\rangle\, \eta(r) dr \\ + \frac{1}{2}\int_0^{s\wedge\tau} \mathrm{Tr}\left[\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)\left(\sigma(r, X(r), a(r))Q^{\frac{1}{2}}\right)^* D^2F(r, X(r))\right] \eta(r) dr \\ + \int_0^{s\wedge\tau} \left\langle DF(r, X(r))\eta(r), \sigma(r, X(r), a(r)) dW_Q(r)\right\rangle \quad \text{on } [0, T]. \end{aligned} \tag{1.112}$$

In particular, for $s \in [0, T]$,

$$
\begin{aligned}
\mathbb{E}[F(s\wedge\tau, X(s\wedge\tau))\eta(s\wedge\tau)] \le F(0,x)\eta_0 \\
+\mathbb{E}\int_0^{s\wedge\tau}(F_t(r,X(r))\eta(r)+F(r,X(r))\tilde{b}(r))dr \\
+\mathbb{E}\int_0^{s\wedge\tau}\langle DF(r,X(r)), b(r,X(r),a(r))\rangle\,\eta(r)dr \\
+\frac{1}{2}\mathbb{E}\int_0^{s\wedge\tau}\mathrm{Tr}\left[\left(\sigma(r,X(r),a(r))Q^{\frac{1}{2}}\right)\left(\sigma(r,X(r),a(r))Q^{\frac{1}{2}}\right)^* D^2F(r,X(r))\right]\eta(r)dr.
\end{aligned}
\tag{1.113}
$$

Proof (i) We set, for $s\in[0,T]$, $f(s):=b(s,a(s),X(s))$ and $\Phi(s):=\sigma(s,a(s),X(s))$ and consider the approximation $X^n(\cdot)$ of $X(\cdot)$ as in Proposition 1.132. Observe that, thanks to (1.33), (1.34) and (1.37) we have $f\in M^p_\mu(0,T;H)$ and $\Phi\in\mathcal{N}^p_Q(0,T;H)$ for every $p\ge 1$ so the assumptions of Proposition 1.132 are satisfied.

We observe that $DF(s,x)=\frac{\partial\varphi}{\partial r}(s,|x|)\frac{x}{|x|}$ and, since $\varphi(s,\cdot)$ is non-decreasing on $[0,+\infty)$, $\frac{\partial\varphi}{\partial r}(s,r)\ge 0$. Therefore, since A, and thus A_n, is dissipative,

$$
\left\langle A_nX^n(s), DF(r,X^n(s))\right\rangle = \frac{\partial\varphi}{\partial r}(s,|X^n(s)|)\frac{1}{|X^n(s)|}\left\langle A_nX^n(s), X^n(s)\right\rangle \le 0
\tag{1.114}
$$

for every $s\ge 0$.

Hence, defining for any $R>|x|$ the stopping times τ_n^R as in Proposition 1.164, applying Itô's formula for $X^n(\cdot)$ and using (1.114), we obtain

$$
\begin{aligned}
F(s\wedge\tau_n^R, X^n(s\wedge\tau_n^R)) &= F(0,x)+\int_0^{s\wedge\tau_n^R}\Big[F_t(r,X^n(r))+\left\langle A_nX^n(r), DF(r,X^n(r))\right\rangle \\
&+\left\langle f(r), DF(r,X^n(r))\right\rangle+\frac{1}{2}\mathrm{Tr}\left[\left(\Phi(r)Q^{\frac{1}{2}}\right)\left(\Phi(r)Q^{\frac{1}{2}}\right)^* D^2F(r,X^n(r))\right]\Big]dr \\
&+\int_0^{s\wedge\tau_n^R}\left\langle DF(r,X^n(r)), b(r,X^n(r),a(r))dW_Q(r)\right\rangle \\
&\le F(0,x)+\int_0^{s\wedge\tau_n^R}\Big[F_t(r,X^n(r))+\left\langle f(r), DF(r,X^n(r))\right\rangle \\
&+\frac{1}{2}\mathrm{Tr}\left[\left(\Phi(r)Q^{\frac{1}{2}}\right)\left(\Phi(r)Q^{\frac{1}{2}}\right)^* D^2F(r,X^n(r))\right]\Big]dr \\
&+\int_0^{s\wedge\tau_n^R}\left\langle DF(r,X^n(r)), b(r,X^n(r),a(r))dW_Q(r)\right\rangle.
\end{aligned}
\tag{1.115}
$$

It remains to pass to the limit as $n\to+\infty$ and $R\to+\infty$ in (1.115). This is done following the same arguments as in the proof of Proposition 1.164.

(ii) The proof combines the proof of (i) with the arguments used in the proof of Proposition 1.165-(ii). □

Remark 1.167 Propositions 1.165 and 1.166 are used to work with viscosity solution test functions in Chap. 3. In particular, parts (ii) of them are useful when discount factors are present (see e.g. Lemma 3.65). ■

The next two non-standard versions of Dynkin's formula will be used to prove verification theorems in Chaps. 4 and 5.

Proposition 1.168 *Let $Q = I$. Assume that Hypothesis 1.149 is satisfied. Assume that there exists $\lambda \in \mathbb{R}$, $\lambda \in \varrho(A)$ such that $(\lambda I - A)^{-1}b \colon [0, T] \times H \times \Lambda \to H$ is measurable. Suppose moreover that there exists a $C > 0$ such that, for all $(t, x, a) \in [0, T] \times H \times \Lambda$,*

$$\begin{cases} |(\lambda I - A)^{-1}b(t, x, a)| \leq C(1 + |x|) \\ \|\sigma(t, x, a)\|_{\mathcal{L}(\Xi, H)} \leq C(1 + |x|). \end{cases} \tag{1.116}$$

Fix a Λ-valued progressively measurable process $a(\cdot)$. Let X be the unique mild solution of (1.74) described in Theorem 1.152 such that $X(0) = x \in H$. Let $F \colon [0, T] \times H \to \mathbb{R}$ be such that F and its derivatives F_t, DF, D^2F are continuous in $[0, T] \times H$. Suppose that $DF \colon [0, T] \times H \to D(A^*)$, that A^*DF is continuous in $[0, T] \times H$, that $D^2F : [0, T] \times H \to \mathcal{L}_1(H)$ is continuous, and that there exist $C > 0$ and $N \geq 1$ such that

$$|F(s, x)| + |DF(s, x)| + |F_t(s, x)| + \|D^2F(s, x)\|_{\mathcal{L}_1(H)} + |A^*DF(s, x)| \leq C(1 + |x|)^N. \tag{1.117}$$

Then, for any $s \in [0, T]$,

$$\begin{aligned} \mathbb{E}\left[F(s, X(s))\right] &= F(0, x) + \mathbb{E}\int_0^s F_t(r, X(r))dr + \mathbb{E}\int_0^s \left\langle A^*DF(r, X(r)), X(r)\right\rangle dr \\ &+ \mathbb{E}\int_0^s \left\langle(\lambda I - A^*)DF(r, X(r)), (\lambda I - A)^{-1}b(r, X(r), a(r))\right\rangle dr \\ &+ \frac{1}{2}\mathbb{E}\int_0^s \mathrm{Tr}\left[\sigma(r, X(r), a(r))\sigma(r, X(r), a(r))^* D^2F(r, X(r))\right] dr. \end{aligned} \tag{1.118}$$

Proof We approximate the process $X(\cdot)$ by the processes $X^k(\cdot)$ from Lemma 1.154.

Observe that, thanks to Hypothesis 1.149 and to (1.80), the processes $r \to e^{\frac{1}{k}A}b(r, X(r), a(r))$ and $r \to e^{\frac{1}{k}A}\sigma(r, X(r), a(r))$ belong respectively to $M^p_\mu(0, T; H)$ and $\mathcal{N}^p_I(0, T; H)$ for all $p \geq 1$. Thus we can apply Proposition 1.164 obtaining, for $s \in [0, T]$,

$$\begin{aligned} \mathbb{E}\left[F(s, X^k(s))\right] &= F(0, x) + \int_0^s \mathbb{E}\, F_t(r, X^k(r))dr \\ &+ \int_0^s \mathbb{E}\left\langle A^*DF(r, X^k(r)), X^k(r)\right\rangle dr + \int_0^s \mathbb{E}\left\langle DF(r, X^k(r)), e^{\frac{1}{k}A}b(r)\right\rangle dr \\ &+ \frac{1}{2}\int_0^s \mathbb{E}\,\mathrm{Tr}\left[\left(e^{\frac{1}{k}A}\sigma(r)P^k\right)\left(e^{\frac{1}{k}A}\sigma(r)P^k\right)^* D^2F(r, X^k(r))\right] dr, \end{aligned} \tag{1.119}$$

where we use the notation $b(r) := b(r, X(r), a(r))$, $\sigma(r) := \sigma(r, X(r), a(r))$. The claim will follow if we can pass to the limit as $k \to +\infty$ in each term of this expression. We will only show how to prove the convergence of the last two terms since the arguments for the other terms are similar and simpler.

Using (1.80), (1.86), (1.87) and the dominated convergence theorem it is easy to see that

$$\lim_{k\to\infty} |X(\cdot) - X^k(\cdot)|_{M^2_\mu(0,T;H)} = 0.$$

Therefore we can find a subsequence, still denoted by $X^k(\cdot)$, that converges to $X(\cdot)$ $dt \otimes \mathbb{P}$-a.e.

Using the assumptions it is obvious that

$$\begin{aligned}\left\langle DF(r, X^k(r)), e^{\frac{1}{k}A} b(r)\right\rangle &= \left\langle (\lambda I - A^*) DF(r, X^k(r)), e^{\frac{1}{k}A}(\lambda I - A)^{-1} b(r)\right\rangle \\ &\to \left\langle (\lambda I - A^*) DF(r, X(r)), (\lambda I - A)^{-1} b(r)\right\rangle \quad dt \otimes \mathbb{P} - a.e.\end{aligned}$$

as $k \to +\infty$. Moreover, thanks to (1.80), (1.86), (1.116) and (1.117),

$$\begin{aligned}&\int_0^s \mathbb{E}\left|\left\langle (\lambda I - A^*) DF(r, X^k(r)), e^{\frac{1}{k}A}(\lambda I - A)^{-1} b(r)\right\rangle\right|^2 dr \\ &\qquad\le C_1 \int_0^s \mathbb{E}\left[\left(1 + |X^k(r)|^{2N}\right)\left(1 + |X(r)|^2\right)\right] dr \le C_2\end{aligned}$$

for some C_1 and C_2 independent of k. Similarly we obtain

$$\int_0^s \mathbb{E}\left|\left\langle (\lambda I - A^*) DF(r, X(r)), (\lambda I - A)^{-1} b(r)\right\rangle\right|^2 dr \le C_3$$

for some C_3. Therefore it follows from Lemma 1.51 that

$$\begin{aligned}&\lim_{k\to+\infty} \int_0^s \mathbb{E}\left\langle DF(r, X^k(r)), e^{\frac{1}{k}A} b(r)\right\rangle dr \\ &\qquad= \int_0^s \mathbb{E}\left\langle (\lambda I - A^*) DF(r, X(r)), (\lambda I - A)^{-1} b(r)\right\rangle dr.\end{aligned}$$

Regarding the last term in (1.119),

$$\begin{aligned}&\mathrm{Tr}\left[e^{\frac{1}{k}A}\sigma(r)P^k(e^{\frac{1}{k}A}\sigma(r)P^k)^* D^2F(r, X^k(r))\right] - \mathrm{Tr}\left[\sigma(r)\sigma(r)^* D^2F(r, X(r))\right] \\ &= I_1 + I_2 := \mathrm{Tr}\left[e^{\frac{1}{k}A}\sigma(r)P^k(e^{\frac{1}{k}A}\sigma(r)P^k)^*\left(D^2F(r, X^k(r)) - D^2F(r, X(r))\right)\right] \\ &\qquad + \mathrm{Tr}\left[\left(e^{\frac{1}{k}A}\sigma(r)P^k(e^{\frac{1}{k}A}\sigma(r)P^k)^* - \sigma(r)\sigma(r)^*\right) D^2F(r, X(r))\right].\end{aligned}$$

By Proposition B.26, (1.116) and the assumptions for D^2F we have

$$|I_1| \le C_4(1+|X(r)|)^2 \|D^2F(r, X^k(r)) - D^2F(r, X(r))\|_{\mathcal{L}_1(H)} \to 0 \quad \text{as} \quad k \to +\infty$$

$dt \otimes \mathbb{P}$-a.e. Let $\{e_1, e_2, ...\}$ be an orthonormal basis of eigenvectors of $D^2F(r, X(r))$ and $\lambda_1, \lambda_2, ...$ be the corresponding eigenvalues. Then

$$\begin{aligned} \mathrm{Tr}\left[e^{\frac{1}{k}A}\sigma(r)P^k(e^{\frac{1}{k}A}\sigma(r)P^k)^* D^2F(r, X(r))\right] \\ = \sum_{n=1}^{\infty} \lambda_n \left| P^k\sigma(r)^* e^{\frac{1}{k}A^*} e_n \right|_{\Xi}^2 \to \sum_{n=1}^{\infty} \lambda_n \left|\sigma(r)^* e_n\right|_{\Xi}^2 \\ = \mathrm{Tr}\left[\sigma(r)\sigma(r)^* D^2F(r, X(r))\right] \quad \text{as} \quad k \to +\infty \end{aligned}$$

$dt \otimes \mathbb{P}$-a.e. Therefore $\lim_{k\to+\infty}(I_1 + I_2) = 0$ $dt \otimes \mathbb{P}$-a.e. Since, by (1.80), (1.86), (1.116) and (1.117), we also have

$$\int_0^s \mathbb{E}\, |I_1 + I_2|^2 dr \le C_5$$

for some constant C_5 independent of k, the convergence of the last term in (1.119) now follows from Lemma 1.51. □

Proposition 1.169 *Let Hypotheses 1.143 and 1.145 be satisfied and let $q \ge 2$. Consider $\lambda \in \mathbb{R}$ such that $(\lambda I - A)$ is invertible and $(\lambda I - A)^{-1} \in \mathcal{L}(H)$. Assume that $(\lambda I - A)^{-1} a_2(\cdot) \in M^1_\mu(0, T; H)$. Let X be the unique mild solution of (1.67) described in Proposition 1.147 such that $X(0) = x \in H$. Let $F : [0, T] \times H \to \mathbb{R}$ be such that F and its derivatives F_t, DF, D^2F are continuous in $[0, T] \times H$. Suppose that $DF : [0, T] \times H \to D(A^*)$, A^*DF is continuous in $[0, T] \times H$, $D^2F : [0, T] \times H \to \mathcal{L}_1(H)$ is continuous and there exists a $C > 0$ such that (1.117) holds with $N = 0$. Then, for any $s \in [0, T]$,*

$$\begin{aligned} \mathbb{E}\,[F(s, X(s))] = F(0, x) + \mathbb{E}\int_0^s F_t(r, X(r))dr \\ + \mathbb{E}\int_0^s \left\langle A^*DF(r, X(r)), X(r)\right\rangle dr + \mathbb{E}\int_0^s \langle DF(r, X(r)), b_0(r, X(r), a_1(r))\rangle\, dr \\ + \mathbb{E}\int_0^s \left\langle (\lambda I - A^*)DF(r, X(r)), (\lambda I - A)^{-1}a_2(r)\right\rangle dr \\ + \frac{1}{2}\mathbb{E}\int_0^s \mathrm{Tr}\left[\sigma Q \sigma^* D^2F(r, X(r))\right] dr. \end{aligned}$$

Proof We approximate X using the processes X^k defined in Lemma 1.156. It is immediate to see that $\psi_k \in M^p_\mu(0, T; H)$ and $\sigma\tilde{P}^k \in \mathcal{N}^p_Q(0, T; H)$ for all $p \ge 1$ so we can apply Proposition 1.164 obtaining for every $s \in [0, T]$,

$$\mathbb{E}\left[F(s, X^k(s))\right] = F(0, x) + \mathbb{E}\int_0^s F_t(r, X^k(r))dr + \mathbb{E}\int_0^s \left\langle A^* DF(r, X^k(r)), X^k(r)\right\rangle dr$$
$$+ \mathbb{E}\int_0^s \mathbf{1}_{B_k}(r,\omega)\left\langle DF(r, X^k(r)), b_0(r, X(r), a_1(r))\right\rangle dr$$
$$+ \mathbb{E}\int_0^s \mathbf{1}_{D_{m_k}}(r,\omega)\left\langle (\lambda I - A^*)DF(r, X^k(r)), e^{\frac{1}{k}A}(\lambda I - A)^{-1}a_2(r)\right\rangle dr$$
$$+ \frac{1}{2}\mathbb{E}\int_0^s \mathrm{Tr}\left[(\sigma Q^{1/2}P^k)(\sigma Q^{1/2}P^k)^* D^2F(r, X^k(r))\right] dr, \tag{1.120}$$

where B_k, D_{m_k} and P^k are introduced in Lemma 1.156 and in the paragraph before it.

We need to check the convergence of each term of this expression. Using parts (i) and (ii) of Lemma 1.156 we have

$$\lim_{k\to\infty} |X(\cdot) - X^k(\cdot)|_{M^1_\mu(0,T;H)} = 0.$$

Therefore we can find a subsequence of X^k, still denoted by X^k, that converges $dt \otimes \mathbb{P}$-a.e. to X. The proof proceeds using the same arguments (and even simpler) as those in the proof of Proposition 1.169. We only look at the two middle terms of the right-hand side of (1.120) that are a little different. We observe that

$$\left|\mathbb{E}\int_0^s \left\langle (\lambda I - A^*)DF(r, X^k(r)), \left(1 - \mathbf{1}_{D_{m_k}}(r,\omega)e^{\frac{1}{k}A}\right)(\lambda I - A)^{-1}a_2(r)\right\rangle dr\right|$$

converges to zero thanks to the dominated convergence (recall that, by assumption, (1.117) holds with $N = 0$). Regarding the fourth term observe that

$$\mathbf{1}_{B_k}(r,\omega)\left\langle DF(r, X^k(r)), b_0(r, X(r), a_1(r))\right\rangle$$

converges to

$$\langle DF(r, X(r)), b_0(r, X(r), a_1(r))\rangle$$

$dt \otimes \mathbb{P}$-a.e. as $k \to +\infty$. Moreover, since DF is bounded, Hypothesis 1.145-(i) implies

$$\left|\mathbf{1}_{B_k}(r,\omega)\left\langle DF(r, X^k(r)), b_0(r, X(r), a_1(r))\right\rangle\right| \le Cf(r)(1 + |X(r)|)$$

for all $k \in \mathbb{N}$. Thus the result follows by the dominated convergence theorem. □

1.8 Bibliographical Notes

Section 1.1 contains elements of basic probability and measure theory. Classical references include, for example, [18, 58, 61, 267, 370, 478, 520]. We refer in particular to [58, 61, 267, 370] for the general theory of measure and probability (Sect. 1.1.1) and to [58, 61, 267, 520] for results on measurability (Sects. 1.1.2 and 1.1.3). For the Bochner integral and the integration of Banach-valued functions (Sect. 1.1.3), the reader can consult [190, 191, 194, 397]; some results, useful from the stochastic calculus perspective are contained in [180]. For Sects. 1.1.4 and 1.1.5 conditional expectation for Banach-valued random variables the reader can refer to [180, 356, 370, 478, 572]. Gaussian measure in Hilbert spaces (Sect. 1.1.6) and Fourier transform are nicely introduced in [153, 180] and a more extended study of the subject is contained in [391].

Generalities about stochastic processes, martingales and stopping times in Sect. 1.2 can be found in many different books, e.g. [356, 372, 384, 447–449, 503, 508, 572], while for Hilbert-valued martingales (Sect. 1.2.2) the reader may consult [180, 294, 487]. For standard Wiener and Q-Wiener processes and related results we refer to [124, 180, 294, 372, 447, 448, 452]. The material of Sect. 1.2.4 is based on [180]. Definition 1.92, which not contained in the standard literature, is introduced here because it is useful to study stochastic control problems. The presentation of Lemma 1.94 is based on [372, 513]. The material of Sect. 1.2.5 is loosely based on [180, 294, 372].

The material of Sect. 1.3 is based on [177, 180, 294] (see also [124, 491]). These books present the theory in Hilbert spaces while [447, 448] (see also [192]) present the Banach space case.

The presentation of Sect. 1.4 on solutions of stochastic differential equations in Hilbert spaces is also based on [180, 294]. In particular, [180] is a standard reference in the theory. Other references on strong and mild solutions are, for example, in [124, 177, 413] while a good introduction to variational solutions is in [124, 387, 413, 491, 519]. The reader is also referred to [180] for more on weak mild solutions. Section 1.4.4, containing some results about uniqueness in law, uses the approach of [471]. For a different approach to weak uniqueness based on the theorem of Yamada–Watanabe, we refer the reader to [491], Appendix E.

Section 1.5 contains existence and uniqueness results for stochastic differential equations with special unbounded terms and cylindrical additive noise. They are more or less common knowledge, however we presented proofs since no complete references seem to be available in the literature.

Classical results on transition semigroups (Sect. 1.6) can be found in [180]. The statements here are a little modified and extended so that they may be used in our applications to optimal control, mainly in Chap. 4.

Section 1.7 contains various versions of Itô's and Dynkin's formulae (Propositions 1.164–1.166) in connection with mild solutions for functions that have properties of test functions used in the definition of a viscosity solution (Definition 3.32). Such results have been known and used in the viscosity solution literature, however complete proofs are available only in [374]. The statements here are slightly more general than those in [374] and we presented proofs for the reader's convenience. The last two results of Sect. 1.7 (Propositions 1.168 and 1.169) are used to prove the verification theorems of Sects. 4.8 and 5.5. They have been used in the literature (e.g. in [306]) but without complete proofs, hence we provide them for completeness. We finally recall that Itô's formula related to variational solutions of linear stochastic parabolic equations is proved in [467].