1
Preliminaries

We begin by recalling some basic definitions and results, especially those from functional analysis, partial differential equations (PDEs), probability, and stability theories. We review basic notions and notations from deterministic systems and recall important results from stochastic differential equations. We introduce two notions of solutions, mild and strong, for infinite-dimensional stochastic differential equations and consider the existence and uniqueness of solutions under suitable assumptions. We introduce and clarify various definitions of stochastic stability in Hilbert spaces, which are a natural generalization of deterministic stability concepts. To present the proofs of all the results here would require preparatory background material that would significantly increase both the size and scope of this book. Therefore, we adopt the approach of omitting those proofs, which are treated in detail in well-known standard textbooks such as Da Prato and Zabczyk [53], Pazy [187], and Yosida [224]. However, those proofs will be presented that are not available in the existing books and are to be found scattered in the literature, or that discuss ideas specially relevant to our purpose.

1.1 Linear Operators, Semigroups, and Examples

Throughout this book, the sets of nonnegative integers, positive integers, real numbers, and complex numbers are denoted by $\mathbb{N}$, $\mathbb{N}_+$, $\mathbb{R}$, and $\mathbb{C}$, respectively. Also, $\mathbb{R}_+$ denotes the set of all nonnegative real numbers and $\mathbb{R}^n$ denotes the n-dimensional real vector space equipped with the usual Euclidean norm $\|\cdot\|_{\mathbb{R}^n}$, $n \geq 1$. For any $\lambda \in \mathbb{C}$, the symbols $Re\,\lambda$ and $Im\,\lambda$ denote the real and imaginary parts of λ, respectively. Given a set E, the symbol $\mathbf{1}_E$ denotes the characteristic function of E, i.e., $\mathbf{1}_E(x) = 1$ if $x \in E$ and $\mathbf{1}_E(x) = 0$ if $x \notin E$.

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

A *Banach space* $(X, \|\cdot\|_X)$, real or complex, is a complete normed linear space over $\mathbb{R}$ or $\mathbb{C}$. If the norm $\|\cdot\|_X$ is induced by an inner product $\langle\cdot,\cdot\rangle_X$, then X is called a *Hilbert space*. In this book, we always take the inner product $\langle\cdot,\cdot\rangle_X$ of X to be linear in the first entry and conjugate-linear in the second. We say that a sequence $\{x_n\}_{n\geq 1} \subseteq X$ *(strongly) converges* to $x \in X$ if $\lim_{n\to\infty}\|x_n - x\|_X = 0$. If X contains n linearly independent vectors, but every system of $n+1$ vectors in X is linearly dependent, then X is called an *n-dimensional space*, denoted by $\dim X = n$. Otherwise, the space X is said to be infinite dimensional. We say that X is *separable* if there exists a countable set $S \subseteq X$ such that $\overline{S} = X$, where $\overline{S}$ is the closure of S in X. For a Hilbert space X, a collection $\{e_i\}_{i\geq 1}$ of elements in X is called an *orthonormal set* if $\langle e_i, e_i\rangle_X = 1$ for all i, and $\langle e_i, e_j\rangle_X = 0$ if $i \neq j$. If S is an orthonormal set and no other orthonormal set contains S as a proper subset, then S is called an *orthonormal basis* for X. A Hilbert space is separable if and only if it has a countable orthonormal basis $\{e_i\}$, $i = 1, 2, \ldots$

A typical example of Banach spaces is the so-called Sobolev space, which plays an important role in PDE theory. Let $\mathcal{O}$ be a nonempty domain of $\mathbb{R}^n$, and m be a positive integer. For $1 \leq p < \infty$ we denote by $W^{m,p}(\mathcal{O}; X)$ the set of all elements $y \in L^p(\mathcal{O}; X)$ such that y and its distributional derivatives $\partial^\alpha y$ of order $|\alpha| \leq m$ are in $L^p(\mathcal{O}; X)$, where

$$\partial^\alpha = \frac{\partial^{|\alpha|}}{\partial {x_1}^{\alpha_1}\cdots \partial {x_n}^{\alpha_n}} \qquad \text{and} \qquad |\alpha| = \sum_{i=1}^n \alpha_i.$$

Then $W^{m,p}(\mathcal{O}; X)$ is a Banach space under the norm

$$\|y\|_{m,p} = \left(\int_{\mathcal{O}} \sum_{|\alpha|\leq m} \|\partial^\alpha y(x)\|_X^p dx\right)^{1/p}, \qquad y \in W^{m,p}(\mathcal{O}; X).$$

On the other hand, we denote by $C^m(\mathcal{O}; X)$ the set of all m-times continuously differentiable vectors in $\mathcal{O}$, and by $C_0^m(\mathcal{O}; X)$ the subspace of $C^m(\mathcal{O}; X)$ consisting of those vectors that have compact supports in $\mathcal{O}$. Another important Banach space $W_0^{m,p}(\mathcal{O}; X)$ is defined as the completion of $C_0^\infty(\mathcal{O}; X)$ in the metric of $W^{m,p}(\mathcal{O}; X)$.

In general, the spaces $W^{m,p}(\mathcal{O}; X)$ and $W_0^{m,p}(\mathcal{O}; X)$ do not coincide for bounded $\mathcal{O}$. However, it is true that

$$W^{m,p}(\mathbb{R}^n, \mathbb{R}) = W_0^{m,p}(\mathbb{R}^n, \mathbb{R}).$$

The case $p = 2$ is special since the spaces $W^{m,2}(\mathcal{O}; X)$, $W_0^{m,2}(\mathcal{O}; X)$ (frequently written as $H^m(\mathcal{O}; X)$, $H_0^m(\mathcal{O}; X)$) are Hilbert spaces if X is a Hilbert space under the scalar product

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

$$\langle y, z\rangle_{m,2} = \int_\Omega \sum_{|\alpha|\le m} \langle \partial^\alpha y(x), \partial^\alpha z(x)\rangle_X dx.$$

Let X and Y be two Banach spaces and $\mathscr{D}(A)$ a subspace of X. A map $A\colon \mathscr{D}(A) \subseteq X \to Y$ is called a *linear operator* if the following relation holds:

$$A(\alpha x + \beta y) = \alpha Ax + \beta Ay \qquad \text{for any} \quad x,\ y \in \mathscr{D}(A),\ \ \alpha, \beta \in \mathbb{R} \text{ or } \mathbb{C}.$$

The subspace $\mathscr{D}(A)$ is called the *domain* of A. If A maps any bounded subsets of $\mathscr{D}(A)$ into bounded subsets of Y, we say that A is a *bounded linear operator*. We denote by $\mathscr{L}(X,Y)$ the set of all bounded linear operators A from X to Y with $\mathscr{D}(A) = X$. It may be shown that $\mathscr{L}(X,Y)$ is a Banach space under the operator norm $\|\cdot\|_{\mathscr{L}(X,Y)}$, or simply $\|\cdot\|$, given by

$$\|A\| := \sup_{\|x\|_X\le 1} \|Ax\|_Y = \sup_{\|x\|_X=1} \|Ax\|_Y \qquad \text{for any} \qquad A \in \mathscr{L}(X,Y).$$

For simplicity, we frequently write $\mathscr{L}(X)$ for $\mathscr{L}(X,X)$.

For any linear operator $A\colon \mathscr{D}(A) \subseteq X \to Y$, we define $\mathscr{K}(A) = \{x \in \mathscr{D}(A)\colon Ax = 0\}$ and $\mathscr{R}(A) = \{Ax\colon x \in \mathscr{D}(A)\}$. They are called the *kernel* and *range* spaces of A, respectively.

Theorem 1.1.1 *Let X and Y be two Banach spaces. Then the following results hold:*

(i) (Open Mapping Theorem) $A \in \mathscr{L}(X,Y)$ *and* $\mathscr{R}(A) = Y$ *imply that for any open set* $E \subseteq X$, *the set* $A(E)$ *is open in* Y.

(ii) (Inverse Mapping Theorem) $A \in \mathscr{L}(X,Y)$ *with* $\mathscr{R}(A) = Y$ *and* $\mathscr{K}(A) = \{0\}$ *imply that the inverse operator* A^{-1} *exists and* $A^{-1} \in \mathscr{L}(Y,X)$.

(iii) (Principle of Uniform Boundedness) $\Sigma \subseteq \mathscr{L}(X,Y)$ *and* $\sup_{A\in\Sigma} \|Ax\|_Y < \infty$ *for each* $x \in X$ *imply that* $\sup_{A\in\Sigma} \|A\| < \infty$.

Let $Y = K$ where $K = \mathbb{R}$ or $\mathbb{C}$. Any $f \in \mathscr{L}(X,K)$ is called a *bounded linear functional* on X. In the sequel, we put $X^* = \mathscr{L}(X,K)$, which is a Banach space under the norm $\|\cdot\|_{X^*}$ and call X^* the *dual space* of X. Quite often, we write $f(x)$ for any $f \in X^*$, $x \in X$ by $\langle\langle x, f\rangle\rangle_{X,X^*}$, and the symbol $\langle\langle\cdot,\cdot\rangle\rangle_{X,X^*}$ is referred to as the *duality pair* between X and X^*. The following theorem assures the existence of nontrivial bounded linear functionals on any Banach space.

Theorem 1.1.2 (Hahn–Banach Theorem) *Let X be a Banach space and X_0 a subspace of X. Let $f_0 \in X_0^*$, then there exists an extension $f \in X^*$ of f_0 such that $\|f\|_{X^*} = \|f_0\|_{X_0^*}$.*

Since X^* is a Banach space, we may also talk about the dual space of X^*, i.e., $X^{**} := (X^*)^*$. It is known that for any $x \in X$, by defining

$$x^{**}(f) = f(x) = \langle\!\langle x, f \rangle\!\rangle_{X,X^*} \qquad \text{for any} \qquad f \in X^*, \tag{1.1.1}$$

we have $x^{**} \in X^{**}$ and $\|x\|_X = \|x^{**}\|_{X^{**}}$. Thus, the map $x \to x^{**}$ from X into X^{**} is linear and injective and preserves the norm so that X is embeddable into X^{**}. If we regard x exactly the same as x^{**}, then $X \subset X^{**}$. In general, the strict inclusion may hold, a fact that naturally leads to the following definition.

Definition 1.1.3 A Banach space X is said to be *reflexive* if $X = X^{**}$. Precisely, for any $x^{**} \in X^{**}$, there exists an $x \in X$ such that (1.1.1) holds.

The most important class of reflexive spaces are Hilbert spaces, a fact that is justified by the following theorem.

Theorem 1.1.4 (Riesz Representation Theorem) *Let X be a Hilbert space, then $X^* = X$. More precisely, for any $f \in X^*$, there exists a unique element $y \in X$ such that*

$$f(x) = \langle x, y \rangle_X \qquad \textit{for any} \qquad x \in X, \tag{1.1.2}$$

and conversely, for any $y \in X$, by defining f as in (1.1.2), one has $f \in X^*$. It clearly makes sense to write $\langle \cdot, \cdot \rangle_X$ for $\langle\!\langle \cdot, \cdot \rangle\!\rangle_{X,X^*}$ on this occasion.

Closed linear operators, generally unbounded, frequently appear in applications, notably in connection with partial differential equations.

Definition 1.1.5 Let X and Y be two Banach spaces. A linear operator $A\colon \mathscr{D}(A) \subseteq X \to Y$ is said to be *closed* if whenever

$$x_n \in \mathscr{D}(A), \;\; n \geq 1, \;\; \text{and} \;\; \lim_{n\to\infty} x_n = x, \quad \lim_{n\to\infty} Ax_n = y,$$

it follows that $x \in \mathscr{D}(A)$ and $Ax = y$.

For a closed linear operator $A\colon \mathscr{D}(A) \subseteq X \to X$, it can be shown that the domain $\mathscr{D}(A)$ is a Banach space under the graph norm $\|x\|_{\mathscr{D}(A)} := \|x\|_X + \|Ax\|_X$, $x \in \mathscr{D}(A)$. It is easy to see that any bounded linear operator having a closed domain is closed. The converse statement can be true in the following sense.

Theorem 1.1.6 (Closed Graph Theorem) *Suppose that $A\colon \mathscr{D}(A) \subseteq X \to Y$ is a closed linear operator. If $\mathscr{D}(A)$ is closed in X, then operator A is bounded.*

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

In general, it is difficult to prove that an operator is closed. The next theorem states that if this operator is the algebraic inverse of a bounded linear operator, then it is closed.

Theorem 1.1.7 *Assume that X and Y are Banach spaces and let A be a linear operator from X to Y. If A is invertible with $A^{-1} \in \mathscr{L}(Y, X)$, then A is a closed linear operator.*

Let X and Y be two Banach spaces and a linear operator $A\colon \mathscr{D}(A) \subseteq X \to Y$ is called *densely defined* if $\overline{\mathscr{D}(A)} = X$. If A is densely defined, we may define *Banach space adjoint operator* $A'\colon \mathscr{D}(A') \subseteq Y^* \to X^*$ of A in the following manner. Let

$$\mathscr{D}(A') = \left\{y^* \in Y^* \colon\ y^* A \text{ is continuous on } \mathscr{D}(A)\right\}.$$

The linear operator $A'\colon \mathscr{D}(A') \subseteq Y^* \to X^*$ is defined by

$$\langle\!\langle x, A' y^* \rangle\!\rangle_{X, X^*} = \langle\!\langle Ax, y^* \rangle\!\rangle_{Y, Y^*} \qquad \text{for any} \qquad y^* \in \mathscr{D}(A'),\ \ x \in \mathscr{D}(A).$$

It turns out that A' is uniquely defined and closed and map $A \to A'$ is linear.

Now let us consider the case where A is a densely defined linear operator on a Hilbert space X. Then the Banach space adjoint A' of A is a mapping from X^* into itself. Let $\iota\, :\ X \to X^*$ be the map that assigns, for each $x \in X$, the bounded linear functional $\langle \cdot, x\rangle_X$ in X^*. Then ι is a linear isometry, which is surjective by the Riesz Representation Theorem. Now define a map $A^*\colon X \to X$ by

$$A^* = \iota^{-1} A' \iota.$$

Then $A^*\colon X \to X$ satisfies

$$\langle Ay, x\rangle_X = (\iota x)(Ay) = (A'\iota x)(y) = \langle y, \iota^{-1} A' \iota x\rangle_X = \langle y, A^* x\rangle_X$$

for any $y \in \mathscr{D}(A)$, $x \in \mathscr{D}(A^*)$, and A^* is called the *Hilbert space adjoint*, or simply *adjoint*, of A. In general, $A^* \neq A'$. However, if X is a real Hilbert space, then $A^* = A'$.

Definition 1.1.8 Let X be a Hilbert space. A densely defined linear operator $A\colon \mathscr{D}(A) \subseteq X \to X$ is *symmetric* if for all $x, y \in \mathscr{D}(A)$, $\langle Ax, y\rangle_X = \langle x, Ay\rangle_X$. A symmetric operator A is called *self-adjoint* if $\mathscr{D}(A^*) = \mathscr{D}(A)$.

All bounded and symmetric operators are self-adjoint. It may be shown that the adjoint of a densely defined linear operator on a Hilbert space X is closed, and so is every self-adjoint operator. A linear operator A on the Hilbert space X is called *nonnegative*, denoted by $A \geq 0$, if $\langle Ax, x\rangle_X \geq 0$ for all $x \in \mathscr{D}(A)$. It is called *positive* if $\langle Ax, x\rangle_X > 0$ for all non zero $x \in \mathscr{D}(A)$ and *coercive* if

$\langle Ax, x\rangle_X \geq c\|x\|_X^2$ for some $c > 0$ and all $x \in \mathscr{D}(A)$. We denote the spaces of all nonnegative, positive, and coercive operators by $\mathscr{L}^+(X)$, $\mathscr{L}_0^+(X)$, and $\mathscr{L}_c^+(X)$, respectively. A linear operator B is called the *square root* of A if $B^2 = A$.

Theorem 1.1.9 *Let A be a linear operator on the Hilbert space X. If A is self-adjoint and nonnegative, then it has a unique square root, denoted by $A^{1/2}$, which is self-adjoint and nonnegative such that $\mathscr{D}(A) \subset \mathscr{D}(A^{1/2})$. Furthermore, if A is positive, so is $A^{1/2}$.*

Theorem 1.1.10 *Suppose that A is self-adjoint and nonnegative on the Hilbert space X. Then A is coercive if and only if it has a bounded inverse $A^{-1} \in \mathscr{L}(X)$. In this case, A^{-1} is self-adjoint and nonnegative.*

In the family of all bounded linear operators, there is a subclass, called compact operators, which are in many ways analogous to linear operators in finite-dimensional spaces.

Definition 1.1.11 Let X and Y be two Banach spaces. An operator $A \in \mathscr{L}(X, Y)$ is *compact* if for any bounded sequence $\{x_n\}_{n\geq 1}$ in X, the sequence $\{Ax_n\}_{n\geq 1}$ has a convergent subsequence in Y.

Let X be a separable Hilbert space and $\{e_i\}_{i=1}^{\infty}$ an orthonormal basis. Then for any nonnegative operator $A \in \mathscr{L}(X)$, we define $Tr(A) = \sum_{i=1}^{\infty}\langle e_i, Ae_i\rangle_X$. The number $Tr(A)$ is called the *trace* of A and is independent of the orthonormal basis chosen. An operator $A \in \mathscr{L}(X)$ is called *trace class* if $Tr(|A|) < \infty$, where $|A| = (A^*A)^{1/2}$. If we endow the trace norm $\|A\|_1 := Tr(|A|)$ for any trace class operator A, then the associated family $\mathscr{L}_1(X)$ of all trace class operators forms a Banach space. An operator $A \in \mathscr{L}(X)$ is called *Hilbert–Schmidt* if $Tr(A^*A) < \infty$. The norm corresponding to a Hilbert–Schmidt inner product is $\|A\|_2 := (Tr(A^*A))^{1/2}$ under which all the Hilbert–Schmidt operators form a Hilbert space $\mathscr{L}_2(X)$. It is easy to show that the following inclusions hold and they are all proper when X is infinite dimensional:

$$\{\text{trace class}\} \subset \{\text{Hilbert–Schmidt}\} \subset \{\text{compact}\}.$$

An operator $A \in \mathscr{L}(X)$ is said to have *finite trace* if the series

$$\sum_{i=1}^{\infty}\langle e_i, Ae_i\rangle_X < \infty \tag{1.1.3}$$

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

for any orthonormal basis $\{e_i\}_{i\geq 1}$ in X. In general, it is not true that

$$\sum_{i=1}^{\infty} |\langle e_i, Ae_i\rangle_X| < \infty \tag{1.1.4}$$

for some orthonormal basis implies that $A \in \mathscr{L}_1(X)$. However, for a trace class operator A the sum in (1.1.3) is absolutely convergent and independent of the choice of the orthonormal basis. In particular, for a nonnegative operator $A \in \mathscr{L}(X)$, the concept of a trace class operator coincides with that of an operator having finite trace.

Let $A\colon \mathscr{D}(A) \subseteq X \to X$ be a linear operator on a Banach space X. The *resolvent set* $\rho(A)$ of A is the set of all complex numbers $\lambda \in \mathbb{C}$ such that $(\lambda I - A)^{-1}$ exists and $(\lambda I - A)^{-1} \in \mathscr{L}(X)$, where I is the identity operator on X. For $\lambda \in \rho(A)$, we write $R(\lambda, A) = (\lambda I - A)^{-1}$ and call it the *resolvent operator* of A. The *spectrum* of A is defined to be $\sigma(A) = \mathbb{C} \setminus \rho(A)$. It may be shown that the resolvent set $\rho(A)$ is open in $\mathbb{C}$.

Definition 1.1.12 Let A be a linear operator on Banach space X. Define

(i) $\sigma_p(A) = \{\lambda \in \mathbb{C}\colon \lambda I - A$ is not injective$\}$, and $\sigma_p(A)$ is called the *point spectrum* of A. Moreover, each $\lambda \in \sigma_p(A)$ is called the *eigenvalue*, and each nonzero $x \in \mathscr{D}(A)$ satisfying $(\lambda I - A)x = 0$ is called the *eigenvector* of A corresponding to λ.
(ii) $\sigma_c(A) = \{\lambda \in \mathbb{C}\colon \lambda I - A$ is injective, $\mathscr{R}(\lambda I - A) \neq X$ and $\overline{\mathscr{R}(\lambda I - A)} = X\}$, and $\sigma_c(A)$ is called the *continuous spectrum* of A.
(iii) $\sigma_r(A) = \{\lambda \in \mathbb{C}\colon \lambda I - A$ is injective and $\overline{\mathscr{R}(\lambda I - A)} \neq X\}$, and $\sigma_r(A)$ is called the *residual spectrum* of A.

From this definition, it is immediate that $\sigma_p(A)$, $\sigma_c(A)$, and $\sigma_r(A)$ are mutually exclusive and their union is $\sigma(A)$. If A is self-adjoint, we have $\sigma_r(A) = \emptyset$. Note that if $\dim X < \infty$, all the linear operators A on X are compact and in this case $\sigma(A) = \sigma_p(A)$, a fact that is extendable to any compact operators in infinite-dimensional spaces.

Theorem 1.1.13 *Let X be a Banach space with $\dim X = \infty$. If $A \in \mathscr{L}(X)$ is compact, then one and only one of the following cases holds:*

(i) $\sigma(A) = \{0\}$*;*
(ii) $\sigma(A) = \{0, \lambda_1, \ldots, \lambda_n\}$ *where for each* $1 \leq k \leq n$, $\lambda_k \neq 0$ *and* λ_k *is an eigenvalue of A;*
(iii) $\sigma(A) = \{0, \lambda_1, \lambda_2, \ldots\}$ *where for each* $k \geq 1$, $\lambda_k \neq 0$ *and* λ_k *is an eigenvalue of A with* $\lim_{k\to\infty} \lambda_k = 0$.

In this book, we shall employ the theory of linear semigroups, which usually allows a uniform treatment of many systems such as some parabolic, hyperbolic, and delay equations.

Definition 1.1.14 A *strongly continuous or C_0-semigroup* $S(t) \in \mathscr{L}(X)$, $t \geq 0$, on a Banach space X is a family of bounded linear operators $S(t)\colon X \to X$, $t \geq 0$, satisfying the following:

(i) $S(0)x = x$ for all $x \in X$;
(ii) $S(t+s) = S(t)S(s)$ for all $t, s \geq 0$;
(iii) $S(t)$ is strongly continuous, i.e., for any $x \in X$, $S(\cdot)x\colon [0,\infty) \to X$ is continuous.

For any C_0-semigroup $S(t)$ on X, there exist constants $M \geq 1$ and $\mu \in \mathbb{R}$ such that

$$\|S(t)\| \leq Me^{\mu t}, \qquad t \geq 0. \tag{1.1.5}$$

In particular, the semigroup $S(t)$ is called (*uniformly*) *bounded* if $\mu = 0$ and *exponentially stable* if $\mu < 0$. The semigroup $S(t)$, $t \geq 0$, is called *eventually norm continuous* if the map $t \to S(t)$ is continuous from (r,∞) to $\mathscr{L}(X)$ for some $r > 0$. In particular, $S(t)$, $t \geq 0$, is simply called *norm continuous* if the map $t \to S(t)$ is continuous from $(0,\infty)$ to $\mathscr{L}(X)$. If $M = 1$ in (1.1.5), the semigroup $S(t)$, $t \geq 0$, is called a *pseudo contraction* C_0-semigroup, and if further $\mu = 0$, it is called a *contraction* C_0-semigroup.

In association with the C_0-semigroup $S(t)$, we may define a linear operator $A\colon \mathscr{D}(A) \subseteq X \to X$ by

$$\mathscr{D}(A) = \left\{ x \in X\colon \lim_{t\downarrow 0} \frac{S(t)x - x}{t} \text{ exists in } X \right\},$$

$$Ax = \lim_{t\downarrow 0} \frac{S(t)x - x}{t}, \quad x \in \mathscr{D}(A).$$

The operator A is called the *infinitesimal generator*, or simply *generator*, of the semigroup $\{S(t)\}_{t\geq 0}$, which is frequently written as e^{tA}, $t \geq 0$, in this book. It may be shown that A is densely defined and closed.

For an arbitrary C_0-semigroup e^{tA}, $t \geq 0$, the following theorem gives a characterization of its generator A.

Theorem 1.1.15 (Hille–Yosida Theorem) *Let X be a Banach space and $A\colon \mathscr{D}(A) \subseteq X \to X$ be a linear operator. Then the following are equivalent:*

(i) A generates a C_0-semigroup e^{tA}, $t \geq 0$, on X such that (1.1.5) holds for some $M \geq 1$ and $\mu \in \mathbb{R}$.

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

(ii) A is densely defined, closed, and there exist constants $\mu \in \mathbb{R}$, $M \geq 1$ *such that* $\rho(A) \supset \{\lambda \in \mathbb{C} \colon Re\,\lambda > \mu\}$ *and*

$$\|R(\lambda, A)^n\| \leq \frac{M}{(Re\,\lambda - \mu)^n} \quad \text{for any} \quad n \in \mathbb{N}_+, \; Re\,\lambda > \mu. \tag{1.1.6}$$

In general, it is not easy to verify (1.1.6) for each $n \in \mathbb{N}_+$. We can give, however, a simple characterization of linear operators that generate pseudo contraction C_0-semigroups.

Definition 1.1.16 A linear operator $A \colon \mathscr{D}(A) \subset X \to X$ on a Banach space X is called *dissipative* if

$$\|(\lambda I - A)x\|_X \geq \lambda \|x\|_X \quad \text{for all} \quad x \in \mathscr{D}(A) \quad \text{and} \quad \lambda > 0.$$

Theorem 1.1.17 (Lumer and Phillips Theorem) *Let* $A \colon \mathscr{D}(A) \subset X \to X$ *be a linear operator defined on* X*. Then* A *is the generator of a contraction* C_0*-semigroup on* X *if and only if*

(i) A is a closed linear operator with dense domain in X;
(ii) A and its adjoint operator A' *are dissipative.*

If X is a Hilbert space, the conditions in Theorem 1.1.17 may be simplified. In particular, we have the following proposition, which is a consequence of Theorem 1.1.17.

Proposition 1.1.18 *Let* A *be a closed, densely defined linear operator on a Hilbert space* X*. There exists a real number* $\alpha \in \mathbb{R}$ *such that*

$$Re\,\langle x, Ax\rangle_X \leq \alpha \|x\|_X^2 \quad \text{for all} \quad x \in \mathscr{D}(A), \tag{1.1.7}$$

and

$$Re\,\langle x, A^* x\rangle_X \leq \alpha \|x\|_X^2 \quad \text{for all} \quad x \in \mathscr{D}(A^*), \tag{1.1.8}$$

if and only if A *generates a pseudo contraction* C_0*-semigroup* e^{tA}, $t \geq 0$, *satisfying*

$$\|e^{tA}\| \leq e^{\alpha t} \quad \text{for all} \quad t \geq 0. \tag{1.1.9}$$

We state some properties of C_0-semigroups and their generators.

Proposition 1.1.19 *Let* e^{tA}, $t \geq 0$, *be a* C_0*-semigroup on a Banach space* X *and* $A_n = nAR(n, A) \in \mathscr{L}(X)$, $n \in \rho(A)$, *called the Yosida approximation of* A*. Then*

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

$$\lim_{n\to\infty} \|A_n x - Ax\|_X = 0 \quad \textit{for any} \quad x \in \mathscr{D}(A),$$

and

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \|e^{tA_n}x - e^{tA}x\|_X = 0 \quad \textit{for any} \quad x \in X, \quad T \geq 0.$$

Proposition 1.1.20 *For the generator A of a C_0-semigroup e^{tA}, $t \geq 0$, on a Banach space X,*

(i) if $x \in \mathscr{D}(A)$, then $e^{tA}x \in \mathscr{D}(A)$ and

$$\frac{d}{dt}e^{tA}x = e^{tA}Ax = Ae^{tA}x \quad \textit{for all} \quad t \geq 0;$$

(ii) for every $t \geq 0$ and $x \in X$,

$$\int_0^t e^{sA}x\,ds \in \mathscr{D}(A) \quad \textit{and} \quad A\int_0^t e^{sA}x\,ds = e^{tA}x - x.$$

Let X be a Banach space and consider the following deterministic linear Cauchy problem on X,

$$\begin{cases} \dfrac{dy(t)}{dt} = Ay(t), & t \geq 0, \\ y(0) = y_0 \in X, \end{cases} \tag{1.1.10}$$

where A is a linear operator that generates a C_0-semigroup e^{tA}, $t \geq 0$, on X. If $y_0 \in \mathscr{D}(A)$, we have by Proposition 1.1.20 that $e^{tA}y_0 \in \mathscr{D}(A)$ and

$$\frac{d}{dt}(e^{tA}y_0) = Ae^{tA}y_0, \qquad t \geq 0. \tag{1.1.11}$$

Hence, $y(t) = e^{tA}y_0$, $t \geq 0$, is a solution of the differential equation (1.1.10). If $y_0 \notin \mathscr{D}(A)$, the equality (1.1.11) may not be meaningful. However, for any $y_0 \in X$ it does make sense to define $y(t) = e^{tA}y_0$, $t \geq 0$, which is called a *mild solution* of (1.1.10). Quite a few partial differential equations can be formulated in the form (1.1.10).

Example 1.1.21 Let $\{\lambda_i\}$ be a sequence of complex numbers and $\{e_i\}$, $i \in \mathbb{N}_+$, be an orthonormal basis in a separable Hilbert space H. We define on H an operator A by

$$Ax = \sum_{i=1}^{\infty} \lambda_i \langle x, e_i \rangle_H e_i, \qquad x \in \mathscr{D}(A),$$

with its domain

$$\mathscr{D}(A) = \left\{ x \in H : \sum_{i=1}^{\infty} |\lambda_i \langle x, e_i \rangle_H|^2 < \infty \right\}.$$

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

It can be shown that A is a closed, densely defined linear operator and $\lambda I - A$ is invertible if and only if $\inf_{i\geq 1} |\lambda_i - \lambda| > 0$. Moreover, it is true by virtue of the Hille–Yosida Theorem that A generates a C_0-semigroup e^{tA}, $t \geq 0$, if $\sup_{i\geq 1}\{Re\,\lambda_i\} < \infty$, and in this case, we have

$$e^{tA}x = \sum_{i=1}^{\infty} e^{\lambda_i t}\langle x, e_i\rangle_H e_i, \qquad x \in H, \qquad t \geq 0.$$

Moreover, if $\lambda_i \in \mathbb{R}$ for each $i \geq 1$, then A is a self-adjoint operator on H.

As a special case, we could take A to be the classical Laplace operator $\Delta = \partial^2/\partial x_1^2 + \cdots + \partial^2/\partial x_N^2$ on some open bounded set $\mathcal{O} \subset \mathbb{R}^N$ with zero boundary conditions on a smooth boundary $\partial\mathcal{O}$ such that

$$\mathscr{D}(A) = H^2(\mathcal{O}) \cap H_0^1(\mathcal{O}).$$

In particular, if $N = 1$ and $\mathcal{O} = (0,1)$, we may have $e_i(x) = \sqrt{2}\sin(i\pi x)$, $x \in (0,1)$, $\lambda_i = -i^2\pi^2$, $i \geq 1$. As soon as locating the eigenfunctions and eigenvalues of Δ, one can give in terms of semigroup $e^{t\Delta}$, $t \geq 0$, the solution of the partial differential equation

$$\begin{cases} \dfrac{\partial y(t,x)}{\partial t} = \Delta y(t,x) \quad \text{in } \mathcal{O}, \ \ t \geq 0, \\ y(t,x)|_{\partial\mathcal{O}} = 0, \quad t \geq 0; \quad y(0,x) = y_0(x) \in L^2(\mathcal{O}). \end{cases}$$

Example 1.1.22 Let A be a self-adjoint, nonnegative operator on a Hilbert space H such that the coercive condition holds:

$$\langle Ax, x\rangle_H \geq \beta\|x\|_H^2, \qquad \forall x \in \mathscr{D}(A), \quad \beta > 0.$$

Then, by virtue of Theorem 1.1.10, A has a bounded inverse A^{-1} that is self-adjoint and nonnegative. Moreover, we know by virtue of Theorem 1.1.9 that both the square root operators $A^{1/2}$ and $A^{-1/2}$ are well defined. Let $B \in \mathscr{L}(\mathscr{D}(A^{1/2}), H)$ be a self-adjoint operator on H with $\mathscr{D}(A^{1/2}) \subset \mathscr{D}(B)$. Assume that there exists a number $\alpha \in \mathbb{R}$ such that

$$\langle x, Bx\rangle_H \leq \alpha\|x\|_H^2, \qquad x \in \mathscr{D}(B).$$

We are interested in the following abstract wave equation on H,

$$\begin{cases} \dfrac{d^2u(t)}{dt^2} + Au(t) = B\dfrac{du(t)}{dt}, \qquad t \geq 0, \\ u(0) = u_0 \in H, \quad \dfrac{du}{dt}(0) = u_1 \in H. \end{cases} \tag{1.1.12}$$

To formulate (1.1.12) as a first-order equation in the form (1.1.10), we introduce a space $\mathcal{H} = \mathscr{D}(A^{1/2}) \times H$, equipped with a mapping $\langle\cdot,\cdot\rangle_{\mathcal{H}}\colon \mathcal{H} \times$

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

$\mathcal{H} \to \mathbb{C}$,

$$\langle y, \tilde{y} \rangle_{\mathcal{H}} := \langle A^{1/2} y_1, A^{1/2} \tilde{y}_1 \rangle_H + \langle y_2, \tilde{y}_2 \rangle_H,$$

where

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \qquad \tilde{y} = \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix} \in \mathcal{H}.$$

It turns out that $\mathcal{H}$ is a Hilbert space under the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$. Define two linear operators on $\mathcal{H}$

$$\mathcal{A} = \begin{pmatrix} 0 & I \\ -A & B \end{pmatrix} \quad \text{with domain} \quad \mathscr{D}(\mathcal{A}) = \mathscr{D}(A) \times \mathscr{D}(A^{1/2}),$$

and

$$\mathcal{B} = \begin{pmatrix} A^{-1}B & -A^{-1} \\ I & 0 \end{pmatrix},$$

then $\mathcal{B}$ is a bounded linear operator on $\mathcal{H}$ with the range $\mathscr{R}(\mathcal{B}) = \mathscr{D}(\mathcal{A})$ and $\mathcal{B}$ is the inverse of operator $\mathcal{A}$. This implies by Theorem 1.1.7 that $\mathcal{A}$ is a closed operator. Hence, (1.1.12) may be rewritten as a first-order differential equation on $\mathcal{H}$,

$$\begin{cases} \dfrac{dy(t)}{dt} = \mathcal{A}y(t), & t \geq 0, \\ y(0) = y_0 \in \mathcal{H}, \end{cases} \tag{1.1.13}$$

where

$$y(t) = \begin{pmatrix} u(t) \\ du(t)/dt \end{pmatrix}, \qquad y_0 = \begin{pmatrix} u_0 \\ u_1 \end{pmatrix}.$$

On the other hand, it is straightforward to show that

$$Re\,\langle \mathcal{A}y, y \rangle_{\mathcal{H}} = Re\,\langle Ay_1, y_2 \rangle_H + Re\,\langle -Ay_1 + By_2, y_2 \rangle_H \leq \alpha \|y_2\|_H^2 \leq |\alpha| \|y\|_{\mathcal{H}}^2$$

for any $y \in \mathscr{D}(A) \times \mathscr{D}(A^{1/2})$. Similarly, the adjoint of $\mathcal{A}$ is easily shown to be

$$\mathcal{A}^* \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 & -I \\ A & B \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \qquad \mathscr{D}(\mathcal{A}^*) = \mathscr{D}(\mathcal{A}),$$

which immediately yields $Re\,\langle \mathcal{A}^* y, y \rangle_{\mathcal{H}} \leq |\alpha| \|y\|_{\mathcal{A}}^2$ for any $y \in \mathscr{D}(\mathcal{A}^*)$. Therefore, by virtue of Proposition 1.1.18, it follows that $\mathcal{A}$ generates a C_0-semigroup $e^{t\mathcal{A}}$, $t \geq 0$, on $\mathcal{H}$.

As a typical example, we define for $t \geq 0$, $x \in (0,1)$,

$$Au(t,x) = -u''_{xx}(t,x), \qquad \mathscr{D}(A) = H^2(0,1) \cap H_0^1(0,1),$$
$$\text{and} \qquad Bu(t,x) = \alpha u'_x(t,x).$$

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

It is easy to verify that all the preceding conditions are satisfied, and the partial differential equation that gives the abstract version (1.1.13) is

$$\begin{cases} \dfrac{\partial^2 u(t,x)}{\partial t^2} - \dfrac{\partial^2 u(t,x)}{\partial x^2} = \alpha \dfrac{\partial^2 u(t,x)}{\partial x \partial t}, \qquad t \geq 0, \qquad \alpha \in \mathbb{R}, \\ u(t,0) = u(t,1) = 0, \qquad u(0,x) = u_0(x) \in H^2(0,1) \cap H_0^1(0,1), \qquad t \geq 0, \\ \dfrac{\partial u}{\partial t}(0,x) = u_1(x) \in H_0^1(0,1). \end{cases}$$

Example 1.1.23 Consider a retarded differential equation in $\mathbb{C}^n$ of the form

$$\begin{cases} dy(t) = A_0 y(t)dt + A_1 y(t-r)dt, \qquad t \geq 0, \\ y(0) = \phi_0 \in \mathbb{C}^n, \qquad y(t) = \phi_1(t) \in L^2([-r,0];\mathbb{C}^n), \qquad -r \leq t \leq 0, \end{cases} \tag{1.1.14}$$

where $A_0,\ A_1 \in \mathscr{L}(\mathbb{C}^n)$.

We wish to formulate (1.1.14) into an abstract linear differential equation on a proper Hilbert space. To this end, we introduce a product Hilbert space $\mathcal{H} = \mathbb{C}^n \times L^2([-r,0];\mathbb{C}^n)$, equipped with the usual inner product, and meanwhile a linear operator $\mathcal{A}$ on $\mathcal{H}$ by

$$\mathcal{A}\Phi = \left(A_0\phi_0 + A_1\phi_1(-r), \frac{d\phi_1}{d\theta}(\theta) \right) \quad \text{for} \quad \phi = (\phi_0, \phi_1) \in \mathscr{D}(\mathcal{A}),$$

with its domain

$$\mathscr{D}(\mathcal{A}) = \left\{ \phi = (\phi_0,\phi_1) \in \mathcal{H} \colon \phi_1 \in W^{1,2}([-r,0];\mathbb{C}^n),\ \phi_1(0) = \phi_0 \right\}.$$

It may be shown (see Appendix B) that $\mathcal{A}$ generates a C_0-semigroup $e^{t\mathcal{A}}$, $t \geq 0$, on $\mathcal{H}$ and the equation (1.1.14) becomes a Cauchy problem without delay on $\mathcal{H}$,

$$\begin{cases} dY(t) = \mathcal{A}Y(t)dt, \qquad t \geq 0, \\ Y(0) = (\phi_0,\phi_1) \in \mathcal{H}, \end{cases} \tag{1.1.15}$$

where $Y(t) := (y(t), y_t)$, $t \geq 0$, and $y_t(\theta) := y(t+\theta)$, $\theta \in [-r,0]$, is the so-called *lift-up* system of (1.1.14).

Last, we review some specific types of C_0-semigroups with delicate properties.

Definition 1.1.24 Let e^{tA}, $t \geq 0$, be a C_0-semigroup on a Banach space X with its generator $A\colon \mathscr{D}(A) \subset X \to X$.

(i) The semigroup e^{tA}, $t \geq 0$, is called *eventually compact* if there exists $r > 0$ such that $e^{tA} \in \mathscr{L}(X)$ is compact for any $t \in (r,\infty)$. Particularly, if e^{tA} is compact for all $t \in (0,\infty)$, this semigroup is simply called *compact*.

(ii) The semigroup e^{tA}, $t \geq 0$, is called *analytic* if it admits an extension e^{zA} on $z \in \Delta_\theta := \{z \in \mathbb{C}\colon |\arg z| < \theta\}$ for some $\theta \in (0,\pi]$, such that $z \to e^{zA}$ is analytic on Δ_θ and satisfies the following:

(a) $e^{(z_1+z_2)A} = e^{z_1 A} e^{z_2 A}$ for any $z_1, z_2 \in \Delta_\theta$;
(b) $\lim_{\Delta_{\bar\theta} \ni z \to 0} \|e^{zA}x - x\|_X = 0$ for all $x \in X$ and $0 < \bar\theta < \theta$.

Theorem 1.1.25 *Assume that A generates a C_0-semigroup e^{tA}, $t \geq 0$, on a Banach space X. For some $r \geq 0$, the semigroup e^{tA} is compact at any $t \in (r,\infty)$ if and only if e^{tA} is norm continuous on (r,∞) and the operator $R(\lambda, A)e^{rA}$ is compact for some (thus, all) $\lambda \in \rho(A)$.*

Theorem 1.1.26 *Assume that A generates a C_0-semigroup e^{tA}, $t \geq 0$, which is eventually compact on a Banach space X, then the spectrum of A consists of isolated eigenvalues, i.e., for any numbers m and M, there are only a finite number of eigenvalues of A in the strip*

$$\{\lambda \in \mathbb{C}\colon m \leq Re\,\lambda \leq M\}.$$

For analytic semigroups, we have the following characterization.

Theorem 1.1.27 *Let e^{tA}, $t \geq 0$, be a C_0-semigroup on a Banach space X with generator A. The following statements are equivalent.*

(i) The semigroup e^{tA}, $t \geq 0$, is analytic.
(ii) There exist constants $M > 0$ and $L \geq 0$ such that

$$\|AR(\lambda, A)^{n+1}\| \leq M/n\lambda^n \quad \textit{for all} \quad \lambda > nL,\ n = 1, 2, \ldots$$

(iii) The semigroup e^{tA} is differentiable for $t > 0$, i.e., for every $x \in X$, the mapping $t \to e^{tA}x$ is differentiable for $t > 0$, and there exist constants $M > 0$ and $\mu > 0$ such that

$$\|Ae^{tA}\| \leq \frac{M}{t} e^{\mu t} \quad \textit{for} \quad t > 0.$$

In general, it is hard to check (ii) for every $n \in \mathbb{N}_+$. The following theorem is much more easily verified and thus quite useful in application.

Theorem 1.1.28 *Let e^{tA}, $t \geq 0$, be a C_0-semigroup with generator A on X. The semigroup e^{tA}, $t \geq 0$, is analytic if and only if there exist $M > 0$ and $\mu \in \mathbb{R}$ such that*

$$\rho(A) \supset \{\lambda\colon Re\,\lambda \geq \mu\} \quad \textit{and} \quad \|R(\lambda, A)\| \leq \frac{M}{1+|\lambda|} \quad \textit{for all} \quad Re\,\lambda \geq \mu.$$

Assume that A generates an exponentially stable analytic semigroup e^{tA}, $t \geq 0$, on X. Then $i\mathbb{R} \subset \rho(A)$ and for any $\alpha \in (0, 1)$, the integral

$$(-A)^{-\alpha} := \frac{\sin \alpha\pi}{\pi} \int_0^\infty t^{-\alpha} R(t, A) dt$$

is well defined, which is a bounded linear operator $(-A)^{-\alpha} \in \mathscr{L}(X)$. It may be shown that this operator $(-A)^{-\alpha}$ is injective, a fact that leads to the following definition:

$$(-A)^{\alpha} := \begin{cases} [(-A)^{-\alpha}]^{-1} & \text{if} \quad 0 < \alpha < 1, \\ I & \text{if} \quad \alpha = 0. \end{cases}$$

The operator $(-A)^{\alpha}$ with domain $\mathscr{D}((-A)^{\alpha})$, $\alpha \in [0, 1)$, is called a *fractional power* of $-A$. Further, we have a relation

$$\mathscr{D}((-A)^{\beta}) \subset \mathscr{D}((-A)^{\alpha}) \subset X, \qquad 0 \leq \alpha \leq \beta < 1,$$

and there exists a number $C_\alpha > 0$ such that

$$\|(-A)^{\alpha} e^{tA}\| \leq C_\alpha t^{-\alpha} \quad \text{for each} \;\; t > 0.$$

In finite-dimensional spaces, it is well known that the spectrum relation $\sigma(e^{tA})\backslash\{0\} = e^{t\sigma(A)}$ holds for each $t \geq 0$ between C_0-semigroup e^{tA}, $t \geq 0$, and its generator A. A partial result remains valid in infinite dimensions, which is the content of the following *spectral mapping theorem*.

Theorem 1.1.29 *Let e^{tA}, $t \geq 0$, be a C_0-semigroup with generator A on a Banach space X. Then*

$$e^{t\sigma(A)} \subset \sigma(e^{tA})\backslash\{0\} \qquad \textit{for all} \qquad t \geq 0. \tag{1.1.16}$$

In general, the strict inclusion in (1.1.16) may hold, although this is not the case for norm continuous semigroups in which compact, differentiable, and analytic semigroups are typical examples.

Theorem 1.1.30 *Let e^{tA}, $t \geq 0$, be a C_0-semigroup that is (eventually) norm continuous on the Banach space X. Then*

$$\sigma(e^{tA})\backslash\{0\} = e^{t\sigma(A)} \qquad \textit{for each} \qquad t \geq 0.$$

For an arbitrary C_0-semigroup e^{tA}, $t \geq 0$, it is not generally true that the adjoint of e^{tA} is a C_0-semigroup since the mapping $e^{tA} \rightarrow (e^{tA})'$ does not necessarily preserve the strong continuity of e^{tA}. But this could be true if the underlying space X is a Hilbert space.

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

Proposition 1.1.31 *Suppose that X is a Hilbert space and e^{tA}, $t \geq 0$, is a C_0-semigroup on X. Then $(e^{tA})^*$, $t \geq 0$, is a C_0-semigroup on X with its infinitesimal generator A^*, i.e., $(e^{tA})^* = e^{tA^*}$ for $t \geq 0$.*

1.2 Stochastic Processes and Martingales

A *measurable space* is a pair $(\Omega, \mathscr{F})$ where Ω is a set and $\mathscr{F}$ is a σ-field, also called a σ-algebra, of subsets of Ω. This means that the family $\mathscr{F}$ contains Ω and is closed under the operation of taking complements and countable unions of its elements. If $(\Omega, \mathscr{F})$ and $(S, \mathscr{S})$ are two measurable spaces, then a mapping ξ from Ω into S such that the set $\{\omega \in \Omega\colon \xi(\omega) \in A\} = \{\xi \in A\}$ belongs to $\mathscr{F}$ for arbitrary $A \in \mathscr{S}$ is called *measurable* from $(\Omega, \mathscr{F})$ into $(S, \mathscr{S})$. In this book, we shall only be concerned with the case where S is a complete metric space. Thus, we always set $\mathscr{S} = \mathscr{B}(S)$, the Borel σ-field of S, which is the smallest σ-field containing all closed (or open) subsets of S.

A *probability measure* $\mathbb{P}$ on a measurable space $(\Omega, \mathscr{F})$ is a σ-additive function from $\mathscr{F}$ into $[0, 1]$ such that $\mathbb{P}(\Omega) = 1$. The triplet $(\Omega, \mathscr{F}, \mathbb{P})$ is called a *probability space*. If $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space, we set

$$\overline{\mathscr{F}} = \{A \subset \Omega\colon \exists B, C \in \mathscr{F};\ B \subset A \subset C,\ \mathbb{P}(B) = \mathbb{P}(C)\}.$$

Then it may be shown that $\overline{\mathscr{F}}$ is a σ-field, called the *completion* of $\mathscr{F}$. If $\mathscr{F} = \overline{\mathscr{F}}$, the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is said to be *complete.* Unless otherwise stated, completeness of $(\Omega, \mathscr{F}, \mathbb{P})$ will always be assumed in this book.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ denote a complete probability space. A family $\{\mathscr{F}_t\}$, $t \geq 0$, for which each $\mathscr{F}_t$ is a sub-σ-field of $\mathscr{F}$ and forms an increasing family of σ-fields, is called a *filtration* of $\mathscr{F}$. With this $\{\mathscr{F}_t\}_{t \geq 0}$, one can associate another filtration by setting σ-fields $\mathscr{F}_{t+} = \bigcap_{s>t} \mathscr{F}_s$ for $t \geq 0$. We say that the filtration $\{\mathscr{F}_t\}_{t \geq 0}$ is *normal* or satisfies *the usual conditions* if $\mathscr{F}_{t+} = \mathscr{F}_t$ for each $t \geq 0$ and $\mathscr{F}_0$ contains all $\mathbb{P}$-null sets in $\mathscr{F}$.

If ξ is a measurable mapping from $(\Omega, \mathscr{F})$ into $(S, \mathscr{B}(S))$ or an S-valued *random variable* and $\mathbb{P}$ a probability measure on $(\Omega, \mathscr{F})$, then we will denote by $\mathbb{D}_\xi(\cdot)$ the image of $\mathbb{P}$ under the mapping ξ:

$$\mathbb{D}_\xi(A) = \mathbb{P}\{\omega \in \Omega\colon \xi(\omega) \in A\}, \quad \forall A \in \mathscr{B}(S).$$

This is a probability measure on $(S, \mathscr{B}(S))$, which is called the *distribution* or *law* of ξ.

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

Definition 1.2.1 Let $\{\mathscr{F}_t\}_{t\geq 0}$ be a filtration of $\mathscr{F}$. A mapping $\tau: \Omega \to [0, \infty]$ is called the *stopping time* with respect to $\{\mathscr{F}_t\}$, $t \geq 0$, if $\{\omega: \tau(\omega) \leq t\} \in \mathscr{F}_t$ for each $t \geq 0$. The σ-field of events prior to τ, denoted by $\mathscr{F}_\tau$, is defined as

$$\mathscr{F}_\tau = \big\{A \in \mathscr{F}: A \cap \{\tau \leq t\} \in \mathscr{F}_t \ \text{ for every } \ t \geq 0\big\}.$$

Now assume that $S = H$, a separable Hilbert space with norm $\|\cdot\|_H$, and ξ is an H-valued random variable on $(\Omega, \mathscr{F}, \mathbb{P})$. By a standard limit argument, we can define the integral $\int_\Omega \xi(\omega)\mathbb{P}(d\omega)$ of ξ with respect to probability measure $\mathbb{P}$, often denote it by $\mathbb{E}(\xi)$. The integral defined in this way is a Bochner type of integral, which is frequently called the *expectation* or *mean* of ξ in this book. We denote by $L^p(\Omega, \mathscr{F}, \mathbb{P}; H)$, $p \in [1, \infty)$, the set of all equivalence classes of H-valued random variables with respect to the equivalence relation of almost sure equality. Then one can verify that $L^p(\Omega, \mathscr{F}, \mathbb{P};\ H)$, $p \in [1, \infty)$, equipped with the norm

$$\|\xi\|_p = (\mathbb{E}\|\xi\|_H^p)^{1/p}, \qquad p \in [1, \infty), \qquad \xi \in L^p(\Omega, \mathscr{F}, \mathbb{P}; H),$$

is a Banach space. If Ω is an interval $[0, T]$, $\mathscr{F} = \mathscr{B}([0, T])$, $0 \leq T < \infty$, and $\mathbb{P}$ is the standard Lebesgue measure on $[0, T]$, we also write $L^p([0, T]; H)$, or more simply $L^p(0, T)$ when no confusion is possible.

Let K, H be two separable Hilbert spaces. A mapping $\Phi(\cdot)$ from Ω into $\mathscr{L}(K, H)$ is said to be *measurable* if for arbitrary $k \in K$, $\Phi(\cdot)k$ is measurable as a mapping from $(\Omega, \mathscr{F})$ into $(H, \mathscr{B}(H))$. Let $\mathscr{F}(\mathscr{L}(K, H))$ be the smallest σ-field of subsets of $\mathscr{L}(K, H)$ containing all sets of the form

$$\{\Phi \in \mathscr{L}(K, H): \Phi k \in A\}, \quad k \in K, \quad A \in \mathscr{B}(H),$$

then $\Phi: \Omega \to \mathscr{L}(K, H)$ is a measurable mapping from $(\Omega, \mathscr{F})$ into the measurable space $(\mathscr{L}(K, H), \mathscr{F}(\mathscr{L}(K, H)))$. The mapping Φ is said to be *Bochner integrable* with respect to measure $\mathbb{P}$ if for arbitrary $k \in K$, the mapping $\Phi(\cdot)k$ is Bochner integrable and there exists a bounded linear operator $\Psi \in \mathscr{L}(K, H)$ such that

$$\int_\Omega \Phi(\omega)k\mathbb{P}(d\omega) = \Psi k, \quad \forall\, k \in K.$$

The operator Ψ is then denoted by $\Psi = \displaystyle\int_\Omega \Phi(\omega)\mathbb{P}(d\omega)$ and called the *Bochner integral* of Φ.

An arbitrary family $M = \{M(t)\}$, $t \geq 0$, of H-valued random variables defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is called a *stochastic* or *random process*. Sometimes, we also write $M(t, \omega)$ or M_t in place of $M(t)$ for all $t \geq 0$. In the study of stochastic processes, we usually need additional regularities of M to proceed with our program. Specially, a process M is called *measurable*

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

if the mapping $M(\cdot,\cdot)\colon \mathbb{R}_+ \times \Omega \to H$ is $\mathscr{B}(\mathbb{R}_+) \times \mathscr{F}$-measurable. Let $\{\mathscr{F}_t\}$, $t \geq 0$, be an increasing family of sub-σ-fields of $\mathscr{F}$. The process M is called $\{\mathscr{F}_t\}_{t\geq 0}$-*adapted* if each $M(t)$ is measurable with respect to $\mathscr{F}_t$, $t \geq 0$. Clearly, M is always $\{\mathscr{F}_t^M\}_{t\geq 0}$-adapted, where $\mathscr{F}_t^M := \sigma(M(s); 0 \leq s \leq t)$ is the family of the σ-fields generated by $M = \{M(t)\}_{t\geq 0}$. For any $\omega \in \Omega$, the function $M(\cdot,\omega)$ is called a *path* or *trajectory* of M. A stochastic process $N = \{N(t)\}$ is called a *modification* or *version* of $M = \{M(t)\}$ if

$$\mathbb{P}\{\omega \in \Omega\colon M(t,\omega) \neq N(t,\omega)\} = 0, \qquad \forall t \geq 0.$$

Given an H-valued process $M = \{M(t)\}$, $t \geq 0$, and a stopping time $\tau\colon \Omega \to \mathbb{R}_+$, it is desirable for many applications that the mapping $M_\tau\colon \Omega \to H$ defined by $M_{\tau(\omega)}(\omega) = M(\tau(\omega),\omega)$ is also measurable. This is generally not the case if M is only a measurable process. However, this could be true if we confine ourselves to a smaller class of stochastic processes, i.e., progressively measurable processes, defined as follows.

Definition 1.2.2 Suppose that $M = \{M(t)\}$, $t \geq 0$, is an H-valued process and $\{\mathscr{F}_t\}_{t\geq 0}$ is a filtration of $\mathscr{F}$. The process M is said to be *progressively measurable* with respect to $\{\mathscr{F}_t\}_{t\geq 0}$ if for every $t \geq 0$, the mapping

$$[0,t] \times \Omega \to H, \qquad (s,\omega) \to M(s,\omega),$$

is $\mathscr{B}([0,t]) \times \mathscr{F}_t$-measurable.

It is obvious that if M is progressively measurable with respect to $\{\mathscr{F}_t\}_{t\geq 0}$, then it must be both measurable and $\{\mathscr{F}_t\}_{t\geq 0}$-adapted. The following theorem provides the extent to which the converse is true.

Proposition 1.2.3 *Suppose that stochastic process $M = \{M(t)\}$, $t \geq 0$, is measurable and adapted to the filtration $\{\mathscr{F}_t\}_{t\geq 0}$. Then it has a progressively measurable modification.*

Theorem 1.2.4 *Let $M = \{M(t)\}$, $t \geq 0$, be an H-valued progressively measurable process with respect to $\{\mathscr{F}_t\}_{t\geq 0}$, and let τ be a finite stopping time. Then the random variable M_τ is $\mathscr{F}_\tau$-measurable.*

Let $\mathscr{G}$ be an arbitrary sub-σ-field of $\mathscr{F}$. We use $\mathbb{E}(\cdot \mid \mathscr{G})$ to denote the conditional expectation given $\mathscr{G}$. Let M be a stochastic process with state space H. Then it can be shown that there exists a function $\mathbb{P}(s, x, t, \Gamma)$ ($s < t$, $x \in H$, $\Gamma \in \mathscr{B}(H)$) associated with the process M such that

(a) for all (s, x, t), $\mathbb{P}(s, x, t, \cdot)$ is a probability measure on $\mathscr{B}(H)$;
(b) for each (s, t, Γ), $\mathbb{P}(s, \cdot, t, \Gamma)$ is $\mathscr{B}(H)$-measurable;
(c) $\mathbb{E}(\mathbf{1}_{\{M(t)\in\Gamma\}} \mid \mathscr{F}_s^M) = \mathbb{P}(s, x, t, \Gamma)|_{x=M(s)}$ almost surely.

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

An H-valued process $M = M(t)$, $t \geq 0$, defined on $(\Omega, \mathscr{F}, \mathbb{P})$ and adapted to the family $\{\mathscr{F}_t\}_{t\geq 0}$ is said to be a *Markov process* with respect to $\{\mathscr{F}_t\}_{t\geq 0}$ if the following property is satisfied: for all t, $s \geq 0$,

$$\mathbb{E}(f(M(t+s)) \mid \mathscr{F}_t) = \mathbb{E}(f(M(t+s)) \mid \sigma(M(t))) \qquad a.s. \tag{1.2.1}$$

for every bounded real-valued Borel function $f(\cdot)$ on H. In particular, if a relation of the form (1.2.1) continues to hold when the time t is replaced by a stopping time τ, we say that M has *strong Markov property* or M is a *strong Markov process*. If M is a Markov process with respect to $\mathscr{F}_t^M$, $t \geq 0$, we simply say that M is a Markov process. A function $\mathbb{P}(s, x, t, \Gamma)$ satisfying (a), (b), and (c) is called the *transition probability function* of the Markov process M if it further satisfies the following Chapman–Kolmogorov equation

$$\mathbb{P}(s, x, t, \Gamma) = \int_H \mathbb{P}(s, x, u, dy)\mathbb{P}(u, y, t, \Gamma) \tag{1.2.2}$$

for all $x \in H$, $\Gamma \in \mathscr{B}(H)$ and (s, u, t) such that $s \leq u \leq t$. The process $M(t)$, $t \geq 0$, is said to have *homogeneous* transition probability function if

$$\mathbb{P}(s,x,t,\Gamma) = \mathbb{P}(0,x,t-s,\Gamma) \qquad \text{for all} \qquad x \in H, \;\; \Gamma \in \mathscr{B}(H), \;\; s \leq t.$$

In this case, we write $\mathbb{P}(x,t,\Gamma)$ for $\mathbb{P}(0,x,t,\Gamma)$ and the Chapman–Kolmogorov equation (1.2.2) now reduces to

$$\mathbb{P}(x, s+t, \Gamma) = \int_H \mathbb{P}(x, s, dy)\mathbb{P}(y, t, \Gamma) \qquad \text{for every} \quad s, t \geq 0. \tag{1.2.3}$$

On the class of all bounded Borel-measurable functions $B_b(H)$ on H, we can define for any $t \geq 0$ that

$$\mathbb{P}_t f(x) = \int_H f(y)\mathbb{P}(t,x,dy), \qquad \forall\, f \in B_b(H). \tag{1.2.4}$$

Then by virtue of (1.2.3), we can establish the following semigroup property for the family $\mathbb{P}_t$, $t \geq 0$:

$$\mathbb{P}_{t+s} f = \mathbb{P}_t \mathbb{P}_s f \qquad \text{for any} \quad s, t \geq 0, \tag{1.2.5}$$

which is, in essence, a restatement of (1.2.3). Let $C_b(H)$ be the class of all real-valued, bounded continuous functions on H.

Definition 1.2.5 Semigroup $\mathbb{P}_t$, $t \geq 0$, is said to have the *Feller property* if for arbitrary $f \in C_b(H)$ and $t \geq 0$, function $\mathbb{P}_t f(\cdot)$ is continuous. Further, $\mathbb{P}_t$ is said to have the *strongly Feller property* if for arbitrary $f \in B_b(H)$ and $t \geq 0$, the function $\mathbb{P}_t f(\cdot)$ is continuous.

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

Let H be a Hilbert space and $M = \{M(t)\}, t \geq 0$, be an H-valued stochastic process defined on $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$. If $\mathbb{E}\|M(t)\|_H < \infty$ for all $t \geq 0$, then M is called *integrable*. An integrable and adapted H-valued process $M(t)$, $t \geq 0$, is said to be a *martingale* with respect to $\{\mathscr{F}_t\}_{t\geq 0}$ if

$$\mathbb{E}(M(t) \mid \mathscr{F}_s) = M(s) \qquad \mathbb{P} - a.s. \tag{1.2.6}$$

for arbitrary $t \geq s \geq 0$. By the definition of conditional expectations, relation (1.2.6) is equivalent to the following statement

$$\int_F M(t)\mathbb{P}(d\omega) = \int_F M(s)\mathbb{P}(d\omega), \quad \forall F \in \mathscr{F}_s, \; s \leq t.$$

We also recall that a real-valued integrable and adapted process $M(t)$, $t \geq 0$, is said to be a *submartingale* (resp. *supermartingale*) with respect to $\{\mathscr{F}_t\}_{t\geq 0}$ if

$$\mathbb{E}(M(t) \mid \mathscr{F}_s) \geq M(s), \;\; (\text{resp. } \mathbb{E}(M(t) \mid \mathscr{F}_s) \leq M(s)), \;\; \mathbb{P} - a.s.$$

for any $0 \leq s \leq t$. An H-valued stochastic process M is a *continuous martingale* if it is a martingale with almost surely continuous trajectories. An adapted process M is called a *local martingale* if there exists a sequence of stopping times τ_n such that $\tau_n \uparrow \infty$ and for each n, the stopped process $M(t \wedge \tau_n)$, $t \geq 0$, is a martingale.

If $M(t)$, $t \geq 0$, is an H-valued continuous martingale, then $\|M(t)\|_H^2$, $t \geq 0$, is a real-valued continuous submartingale. By the well-known Doob–Meyer decomposition, there exists a unique real-valued, nondecreasing process, denoted by $[M](t)$, with $[M](0) = 0$ such that $\|M(t)\|_H^2 - [M](t)$ is an $\mathscr{F}_t$-martingale. Recall the following strong law of large numbers for martingales, which is useful in stability analysis.

Proposition 1.2.6 *Let $M(t)$, $t \geq 0$, be an H-valued, continuous local martingale with $M(0) = 0$. If*

$$\overline{\lim_{t\to\infty}} \frac{[M](t)}{t} < \infty \qquad a.s.$$

then

$$\lim_{t\to\infty} \frac{M(t)}{t} = 0 \qquad a.s.$$

Let $[0, T]$, $0 \leq T < \infty$, be a subinterval of $[0, \infty)$. An H-valued stochastic process $M(t)$, $t \in [0, T]$, defined on $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\in[0,T]}, \mathbb{P})$, is a *continuous L^p-martingale*, $p \geq 1$, with respect to $\{\mathscr{F}_t\}_{t\in[0,T]}$ if it is a martingale with almost surely continuous trajectories and satisfies, in addition, $\mathbb{E}\sup_{t\in[0,T]} \|M(t)\|_H^p < \infty$. Let us denote by $\mathcal{M}_T^p(H)$ the space of all

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

H-valued continuous L^p-martingales on $[0,T]$. By using Theorem 1.2.9 it is possible to show the following result.

Theorem 1.2.7 *For $p \geq 1$, the space $\mathcal{M}_T^p(H)$, equipped with the norm*

$$\|M\|_{\mathcal{M}_T^p(H)} = \left(\mathbb{E} \sup_{t\in[0,T]} \|M(t)\|_H^p\right)^{1/p}, \qquad \forall\, M \in \mathcal{M}_T^p(H),$$

is a Banach space.

An $\mathscr{L}_1(H)$-valued process V is said to be *nondecreasing* if operator $V(t)$, $t \in [0,T]$, is nonnegative, so denote it by $V(t) \geq 0$, i.e., for any $x \in H$ and $t \in [0,T]$, $\langle V(t)x, x\rangle_H \geq 0$ and $V(t) - V(s) \geq 0$ if $0 \leq s \leq t \leq T$. For any $M \in \mathcal{M}_T^2(H)$, an $\mathscr{L}_1(H)$-valued continuous, adapted, and nondecreasing process $V(t)$ with $V(0) = 0$ is called a *quadratic variation process* of M if for arbitrary $a, b \in H$, the process

$$\langle M(t), a\rangle_H \langle M(t), b\rangle_H - \langle V(t)a, b\rangle_H, \quad t \in [0,T],$$

is a continuous $\mathscr{F}_t$-martingale, $t \in [0,T]$. One can show that such a process $V(t)$, $t \in [0,T]$, is uniquely determined, thus denote it by $\langle\langle M\rangle\rangle(t)$, $t \in [0,T]$.

Theorem 1.2.8 *For arbitrary $M \in \mathcal{M}_T^2(H)$, there exists a unique nonnegative symmetric process $Q_M(t) \in \mathscr{L}_1(H)$, $t \in [0,T]$, such that*

$$\langle\langle M\rangle\rangle(t) = \int_0^t Q_M(s)d[M](s) \quad \text{for all} \quad t \in [0,T].$$

This process $M(t)$, $t \in [0,T]$, is called a $Q_M(t)$-martingale process.

In a similar manner, one can define the so-called *cross quadratic variation* for any $M \in \mathcal{M}_T^2(H)$, $N \in \mathcal{M}_T^2(H)$ as a unique continuous process $\langle\langle M, N\rangle\rangle$ of operators on H such that for arbitrary $a, b \in H$, the process

$$\langle M(t), a\rangle_H \langle N(t), b\rangle_H - \langle \langle\langle M, N\rangle\rangle(t)a, b\rangle_H, \qquad t \in [0,T],$$

is a continuous $\mathscr{F}_t$-martingale, $t \in [0,T]$.

As an immediate consequence of the classic maximal inequalities for real-valued submartingales, we have the following Doob's type of inequalities in Hilbert spaces.

Theorem 1.2.9 *Let $M(t)$, $t \in [0,T]$, be a continuous H-valued, L^p-martingale, $p \geq 1$. Then the following statements hold.*

(i) For $p \geq 1$ and any $\lambda > 0$,

$$\mathbb{P}\left\{\sup_{0\leq t\leq T} \|M(t)\|_H \geq \lambda\right\} \leq \lambda^{-p} \sup_{0\leq t\leq T} \mathbb{E}(\|M(t)\|_H^p). \tag{1.2.7}$$

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

(ii) For $p > 1$,

$$\mathbb{E}\left(\sup_{0\le t\le T}\|M(t)\|_H^p\right) \le \left(\frac{p}{p-1}\right)^p \sup_{0\le t\le T}\mathbb{E}(\|M(t)\|_H^p). \tag{1.2.8}$$

(iii) For $p = 1$,

$$\mathbb{E}\left(\sup_{0\le t\le T}\|M(t)\|_H\right) \le 3\mathbb{E}\{Tr(\langle\!\langle M\rangle\!\rangle(t))\}^{1/2}. \tag{1.2.9}$$

1.3 Wiener Processes and Stochastic Integration

Let K be a separable Hilbert space with inner product $\langle\cdot,\cdot\rangle_K$. A probability measure $\mathcal{N}$ on $(K, \mathscr{B}(K))$ is called *Gaussian* if for arbitrary $u \in K$, there exist numbers $\mu \in \mathbb{R}$, $\sigma > 0$, such that

$$\mathcal{N}\{x \in K : \langle u, x\rangle_K \in A\} = N(\mu,\sigma)(A), \qquad A \in \mathscr{B}(\mathbb{R}),$$

where $N(\mu,\sigma)$ is the standard one-dimensional normal distribution with mean μ and variance σ. It is true that if $\mathcal{N}$ is Gaussian, there exist an element $m \in K$ and a nonnegative self-adjoint operator $Q \in \mathscr{L}_1(K)$ such that the characteristic function of $\mathcal{N}$ is given by

$$\int_K e^{i\langle\lambda,\, x\rangle_K}\mathcal{N}(dx) = e^{i\langle\lambda,\, m\rangle_K - \frac{1}{2}\langle Q\lambda,\, \lambda\rangle_K}, \qquad \lambda \in K.$$

Therefore, the measure $\mathcal{N}$ is uniquely determined by m and Q and denoted thus by $\mathcal{N}(m, Q)$. In particular, we call m the *mean* and Q the *covariance operator* of $\mathcal{N}(m, Q)$, respectively.

For a self-adjoint and nonnegative operator $Q \in \mathscr{L}(K)$, we assume, without loss of generality, that there exists an orthonormal basis $\{e_k\}_{k\ge1}$ in K, and a bounded sequence of positive numbers λ_k such that

$$Qe_k = \lambda_k e_k, \qquad k = 1, 2, \ldots$$

A stochastic process W_t or $W(t)$, $t \ge 0$, is called a *Q-Wiener process* in K if

(i) $W(0) = 0$;
(ii) $W(t)$ has continuous trajectories;
(iii) $W(t)$ has independent increments;
(iv) $\mathbb{D}_{W(t)-W(s)} = \mathcal{N}(0, (t-s)Q)$ for all $t \ge s \ge 0$.

If $Tr(Q) = \sum_{k=1}^{\infty}\lambda_k < \infty$, then W is a genuine Wiener process that has continuous paths in K. It is possible that $Tr(Q) = \infty$, e.g., $Q = I$, and in this case, we call W a *cylindrical* Wiener process in K, which, in general, has

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

continuous paths only in another Hilbert space larger than K. It is immediate that the quadratic variation of a Q-Wiener process with $Tr(Q) < \infty$ is given by $\langle\langle W\rangle\rangle(t) = tQ$, $t \geq 0$.

Assume that probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is equipped with a normal filtration $\{\mathscr{F}_t\}_{t\geq 0}$. Let $W(t)$, $t \geq 0$, be a Q-Wiener process in K that is assumed to be adapted to $\{\mathscr{F}_t\}_{t\geq 0}$, and for every $t > s \geq 0$ the increments $W(t) - W(s)$ are independent of $\mathscr{F}_s$. Then $W(t)$, $t \geq 0$, is a continuous martingale relative to $\{\mathscr{F}_t\}_{t\geq 0}$, and W has the following representation:

$$W(t) = \sum_{i=1}^{\infty} \sqrt{\lambda_i}\, w_i(t) e_i, \qquad t \geq 0, \tag{1.3.1}$$

where $(\lambda_i > 0, i \in \mathbb{N}_+)$ are the eigenvalues of Q with their corresponding eigenvectors $(e_i, i \in \mathbb{N}_+)$, and $(w_i(t), i \in \mathbb{N}_+)$ is a group of independent standard real-valued Brownian motions. We introduce a subspace $K_Q = \mathscr{R}(Q^{1/2})$ of K, which is a Hilbert space endowed with the inner product

$$\langle u, v\rangle_{K_Q} = \langle Q^{-1/2}u, Q^{-1/2}v\rangle_K \quad \text{for any} \quad u,\ v \in K_Q.$$

Let $\mathscr{L}_2(K_Q, H)$ denote the space of all Hilbert–Schmidt operators from K_Q into H. Then $\mathscr{L}_2(K_Q, H)$ turns out to be a separable Hilbert space under the inner product

$$\langle L, P\rangle_{\mathscr{L}_2(K_Q,H)} = Tr[LQ^{1/2}(PQ^{1/2})^*] \qquad \text{for any} \quad L,\ \ P \in \mathscr{L}_2(K_Q, H).$$

For arbitrarily given $T \geq 0$, let $B(t, \omega)$, $t \in [0, T]$, be an $\mathscr{L}_2(K_Q, H)$-valued process. We define the following norm for arbitrary $t \in [0, T]$,

$$|B|_t := \left\{\mathbb{E}\int_0^t Tr\big[B(s)Q^{1/2}(B(s)Q^{1/2})^*\big]ds\right\}^{\frac{1}{2}}. \tag{1.3.2}$$

In particular, we denote all $\mathscr{L}_2(K_Q, H)$-valued measurable processes B, adapted to the filtration $\{\mathscr{F}_t\}_{t\in[0,T]}$, satisfying $|B|_T < \infty$ by $\mathcal{U}^2\big([0, T] \times \Omega; \mathscr{L}_2(K_Q, H)\big)$. Recall (see Da Prato and Zabczyk [53]) that the stochastic integral $\int_0^t B(s)dW(s) \in H$, $t \geq 0$, may be defined for all $B \in \mathcal{U}^2([0, T] \times \Omega; \mathscr{L}_2(K_Q, H))$ by

$$\int_0^t B(s)dW(s) = L^2 - \lim_{n\to\infty} \sum_{i=1}^{n} \int_0^t \sqrt{\lambda_i} B(s) e_i dw_i(s), \qquad t \in [0, T]. \tag{1.3.3}$$

It is worth mentioning that stochastic integral (1.3.3) may be generalized, as in finite-dimensional cases, to any $\mathscr{L}_2(K_Q, H)$-valued adapted process $B(\cdot)$ satisfying

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

$$\mathbb{P}\left\{\int_0^T \|B(s)\|^2_{\mathscr{L}_2(K_Q,H)}ds < \infty\right\} = 1.$$

By employing the definition of stochastic integral and a standard limiting procedure, we may establish some useful properties of stochastic integrals.

Proposition 1.3.1 *For arbitrary $T \geq 0$, assume that $B(\cdot) \in \mathcal{U}^2([0,T] \times \Omega; \mathscr{L}_2(K_Q, H))$. Then*

(i) the stochastic integral $\int_0^t B(s)dW(s)$ is a continuous, square integrable H-valued martingale on $[0,T]$. Moreover,

$$\mathbb{E}\left\|\int_0^t B(s)dW(s)\right\|_H^2 = |B|_t^2, \qquad t \in [0,T]; \qquad (1.3.4)$$

(ii) the quadratic variation process of $\int_0^t B(s)dW(s)$ has the form

$$\left\langle\!\!\left\langle \int_0^\cdot B(s)dW(s)\right\rangle\!\!\right\rangle(t) = \int_0^t B(s)Q^{1/2}(B(s)Q^{1/2})^*ds, \qquad t \in [0,T].$$

Proposition 1.3.2 *Assume that B_1, $B_2 \in \mathcal{U}^2([0,T] \times \Omega; \mathscr{L}_2(K_Q, H))$. Then the covariance operators*

$$V(s,t) = Cov\left(\int_0^s B_1(u)dW(u), \int_0^t B_2(u)dW(u)\right), \quad s,\ t \in [0,T],$$

are given by

$$V(s,t) = \mathbb{E}\int_0^{s\wedge t} B_1(u)Q^{1/2}(B_2(u)Q^{1/2})^*du \quad s,\ t \in [0,T].$$

Moreover, for any s, $t \in [0,T]$,

$$\mathbb{E}\left\langle\int_0^s B_1(u)dW(u), \int_0^t B_2(u)dW(u)\right\rangle_H$$
$$= \mathbb{E}\int_0^{s\wedge t} Tr\big[B_1(u)Q^{1/2}(B_2(u)Q^{1/2})^*\big]du.$$

Assume that e^{tA}, $t \geq 0$, is a C_0-semigroup with its infinitesimal generator A on H. Suppose that $B \in \mathcal{U}^2([0,T] \times \Omega; \mathscr{L}_2(K_Q, H))$ is such a process that the stochastic integral

$$\int_0^t e^{(t-s)A}B(s)dW(s) =: W_A^B(t), \qquad t \in [0,T], \qquad (1.3.5)$$

is well defined. This process $W_A^B(t)$ is called the *stochastic convolution* of B. In general, $W_A^B(t)$, $t \in [0,T]$, is no longer a martingale, a fact that makes

W_A^B fail to have decent properties. However, one can still expect some useful results to be valid for this process. For instance, we have a useful version of the following Burkholder–Davis–Gundy type of inequality for $W_A^B(t)$, $t \in [0,T]$.

Theorem 1.3.3 *Let $p > 2$, $T \geq 0$ and assume that process $B \in \mathcal{U}^2([0,T] \times \Omega; \mathscr{L}_2(K_Q, H))$ satisfies*

$$\mathbb{E}\left(\int_0^T \|B(s)\|^p_{\mathscr{L}_2(K_Q,H)} ds\right) < \infty.$$

Then there exists a number $C_{p,T} > 0$, depending on p and T, such that

$$\mathbb{E} \sup_{t\in[0,T]} \left\| \int_0^t e^{(t-s)A} B(s) dW(s) \right\|_H^p \leq C_{p,T} \cdot \mathbb{E}\left(\int_0^T \|B(s)\|^p_{\mathscr{L}_2(K_Q,H)} ds\right). \tag{1.3.6}$$

Note that in Theorem 1.3.3, there is a weak point on the condition $p > 2$ to secure the validness of (1.3.6) for any C_0-semigroup e^{tA}, $t \geq 0$, on H. An alternative version of this theorem is possible to cover the case $p = 2$, although we have to restrict at this moment the C_0-semigroup e^{tA}, $t \geq 0$, to a pseudocontraction one.

Theorem 1.3.4 *Let $p \geq 2$ and $T \geq 0$. Assume that A generates a pseudocontraction C_0-semigroup e^{tA}, $t \geq 0$, and $B \in \mathcal{U}^2([0,T] \times \Omega; \mathscr{L}_2(K_Q,H))$. Then there exists a number $C_{p,T} > 0$, depending only on p and T, such that*

$$\mathbb{E}\left(\sup_{t\in[0,T]} \left\| \int_0^t e^{(t-s)A} B(s) dW(s) \right\|_H^p\right) \leq C_{p,T} \cdot \mathbb{E}\left(\int_0^T \|B(s)\|^2_{\mathscr{L}_2(K_Q,H)} ds\right)^{p/2}.$$

Moreover, if A generates a contraction C_0-semigroup, number $C_{p,T} > 0$ may be chosen to depend on p only.

The following stochastic version of the well-known Fubini theorem will be frequently used in this book.

Proposition 1.3.5 *Let $T \geq 0$ and*

$$B\colon [0,T] \times [0,T] \times \Omega \to \mathscr{L}_2(K_Q, H)$$

be measurable such that for each $s \in [0,T]$, $B(s,t)$ is $\{\mathscr{F}_t\}$-adapted, $t \in [0,T]$, and satisfies

$$\int_0^T \int_0^T \mathbb{E}\|B(s,t)\|^2_{\mathscr{L}_2(K_Q,H)} ds dt < \infty.$$

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

Then

$$\int_0^T \int_0^T B(s,t)dW(t)ds = \int_0^T \int_0^T B(s,t)dsdW(t) \qquad a.s. \tag{1.3.7}$$

Assume that $B \in \mathcal{U}^2([0,T]\times\Omega; \mathscr{L}_2(K_Q,H))$, and F is an H-valued, $\{\mathscr{F}_t\}$-adapted, Bochner integrable process on $[0,T]$. Then the following process

$$y(t) = y_0 + \int_0^t F(s)ds + \int_0^t B(s)dW(s), \qquad t \in [0,T], \qquad y_0 \in H, \tag{1.3.8}$$

is well defined. A function $\Lambda(t,x)\colon [0,T]\times H \to \mathbb{R}$ is called an *Itô functional* if Λ and its Fréchet partial derivatives Λ'_t, Λ'_x, Λ''_{xx} are continuous and bounded on any bounded subsets of $[0,T]\times H$.

Theorem 1.3.6 (Itô's formula) *Assume that $\Lambda\colon [0,T]\times H \to \mathbb{R}$ is an Itô functional. Then for all $t \in [0,T]$, $\Lambda(t,y(t))$ satisfies the following equality:*

$$\begin{aligned} d\Lambda(t,y(t)) = \Big\{ & \Lambda'_t(t,y(t)) + \langle \Lambda'_x(t,y(t)), F(t)\rangle_H \\ & + \frac{1}{2} Tr\big[\Lambda''_{xx}(t,y(t))B(t)Q^{1/2}(B(t)Q^{1/2})^*\big]\Big\} dt \\ & + \langle \Lambda'_x(t,y(t)), B(t)dW(t)\rangle_H. \end{aligned} \tag{1.3.9}$$

1.4 Stochastic Differential Equations

The theory of stochastic differential equations in Hilbert spaces is a natural generalization of finite-dimensional stochastic differential equations introduced by Itô and in a slightly different form by Gihman in the 1940s. The reader is referred to Da Prato and Zabczyk [53] for a systematic statement about this topic. On this occasion, we content ourselves with a presentation of how it is possible to formulate a standard stochastic partial differential equation as a stochastic differential equation in Hilbert spaces.

Let $\mathcal{O}$ be a bounded domain in $\mathbb{R}^n$, $n \in \mathbb{N}_+$, with smooth boundary $\partial\mathcal{O}$. Consider the following initial-boundary value problem for the randomly perturbed heat equation

$$\begin{cases} \dfrac{\partial y}{\partial t}(t,x) = \displaystyle\sum_{i=1}^n \frac{\partial^2 y}{\partial x_i^2}(t,x) + \frac{\partial}{\partial t}W(t,x), & t \geq 0, \quad x \in \mathcal{O}, \\ y(0,x) = y_0(x), \quad x \in \mathcal{O}; \quad y(t,x) = 0, & t \geq 0, \quad x \in \partial\mathcal{O}, \end{cases} \tag{1.4.1}$$

where $W(t,x)$ is a standard Wiener random field (see, e.g., [41]).

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

By analogy with partial differential equations, this stochastic partial differential equation (1.4.1) can be viewed in two different ways. One natural way is to consider its solution as a real-valued random field indexed by temporal and spatial variables t and x. In general, this approach uses complicated probability and calculus, and it will not be developed in this book. On the other hand, one can consider a solution of this equation as a stochastic process indexed by t with values in a proper space of functions of x, say, $L^2(\mathcal{O};\mathbb{R})$. In this manner, we can use advanced analysis to develop a stochastic process theory in an infinite-dimensional setting. For instance, we can write $\sum_{i=1}^n \partial^2/\partial x_i^2$ in (1.4.1) as an abstract operator, say A, from Sobolev space $H^2(\mathcal{O};\mathbb{R}) \cap H_0^1(\mathcal{O};\mathbb{R})$ into $L^2(\mathcal{O};\mathbb{R})$ or from $H_0^1(\mathcal{O};\mathbb{R})$ into $H^{-1}(\mathcal{O};\mathbb{R})$, the dual of $H^1(\mathcal{O};\mathbb{R})$, and let $W(t)$, $t \geq 0$, be a Wiener process in $L^2(\mathcal{O};\mathbb{R})$. Note that the Dirichlet boundary condition here is implicit in the fact that we look for solutions in $H_0^1(\mathcal{O};\mathbb{R})$. In other words, given an initial datum $y_0 \in L^2(\mathcal{O};\mathbb{R})$, we may reformulate (1.4.1) as a system in $L^2(\mathcal{O};\mathbb{R})$ or $H^{-1}(\mathcal{O};\mathbb{R})$ with the following form:

$$\begin{cases} dy(t) = Ay(t)dt + dW(t), & t \geq 0, \\ y(0) = y_0 \in L^2(\mathcal{O};\mathbb{R}). \end{cases}$$

In this book, we mainly adopt the latter viewpoint to establish a stochastic stability theory. We develop two formulations, i.e., semigroup and variational methods, to give a rigorous meaning to the solutions of abstract stochastic differential equations.

1.4.1 Semigroup Approach and Mild Solutions

Let $T \geq 0$ and consider the following semilinear stochastic system on a Hilbert space H,

$$\begin{cases} dy(t) = \big[Ay(t) + F(t, y(t))\big]dt + B(t, y(t))dW(t), & t \in [0, T], \\ y(0) = y_0 \in H, \end{cases} \tag{1.4.2}$$

where A is the infinitesimal generator of a C_0-semigroup e^{tA}, $t \geq 0$, of bounded linear operators on H. The coefficients $F(\cdot,\cdot)$ and $B(\cdot,\cdot)$ are two nonlinear measurable mappings from $[0, T] \times H$ into H and $\mathscr{L}_2(K_Q, H)$, respectively.

Definition 1.4.1 Let $T \geq 0$. An $\{\mathscr{F}_t\}_{t\geq 0}$-adapted stochastic process $y(t) \in H$, $t \in [0, T]$, defined on probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$ is called a *mild solution* of (1.4.2) if it satisfies that

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

$$\mathbb{P}\left\{\int_0^T \|y(t)\|_H^2 dt < \infty\right\} = 1, \tag{1.4.3}$$

$$\mathbb{P}\left\{\int_0^T \Big(\|F(t,y(t))\|_H + \|B(t,y(t))\|^2_{\mathscr{L}_2(K_Q,H)}\Big)dt < \infty\right\} = 1, \tag{1.4.4}$$

and

$$\begin{aligned} y(t) = e^{tA}y_0 + \int_0^t e^{(t-s)A}F(s,y(s))ds \\ + \int_0^t e^{(t-s)A}B(s,y(s))dW(s), \qquad t \in [0,T], \end{aligned} \tag{1.4.5}$$

for $y_0 \in H$ almost surely.

By the standard Picard iteration procedure or a probabilistic fixed-point theorem type of argument, one can establish an existence and uniqueness theorem of mild solutions to (1.4.2) in the case that for any $y, z \in H$ and $t \in [0,T]$,

$$\begin{aligned} &\|F(t,y) - F(t,z)\|_H + \|B(t,y) - B(t,z)\|_{\mathscr{L}_2(K_Q,H)} \\ &\qquad\qquad \le \alpha(T)\|y-z\|_H, \quad \alpha(T) > 0, \\ &\|F(t,y)\|_H + \|B(t,y)\|_{\mathscr{L}_2(K_Q,H)} \le \beta(T)(1+\|y\|_H), \quad \beta(T) > 0. \end{aligned} \tag{1.4.6}$$

Theorem 1.4.2 *Let $T \ge 0$, $p \ge 2$ and assume that condition (1.4.6) holds. Then there exists a unique mild solution $y \in C([0,T]; L^p(\Omega;H))$ to (1.4.2). If, in addition, $\mathbb{E}\|y_0\|_H^p < \infty$, $p > 2$, then the solution y satisfies*

$$\mathbb{E}\Big(\sup_{0\le t\le T} \|y(t,y_0)\|_H^p\Big) < \infty, \qquad p > 2.$$

As a direct application of semigroup theory, we have the following result straightaway.

Proposition 1.4.3 *For arbitrary $y_0 \in \mathscr{D}(A)$, assume that $y(t) \in \mathscr{D}(A)$, $t \in [0,T]$, is an $\{\mathscr{F}_t\}_{t\ge 0}$-adapted stochastic process satisfying (1.4.3), (1.4.4), and the equation*

$$y(t) = y_0 + \int_0^t (Ay(s) + F(s,y(s)))ds + \int_0^t B(s,y(s))dW(s), \qquad t \in [0,T], \tag{1.4.7}$$

then it is a mild solution of equation (1.4.2).

The process y satisfying (1.4.7) is called a solution of (1.4.2) in the strong sense, and a mild solution of (1.4.2) is not necessarily a solution in the strong sense. On the other hand, it is known that the stochastic convolution in (1.4.5) is no longer a martingale, implying that one cannot apply Itô's formula directly to mild solutions of (1.4.2), although it is possible to apply it to solutions of (1.4.2) in the strong sense. This consideration just suggests the usefulness of finding conditions under which a mild solution to (1.4.2) becomes a strong one.

Proposition 1.4.4 *Suppose that the following conditions hold:*

(1) $y_0 \in \mathscr{D}(A)$, $e^{(t-s)A}F(s,y) \in \mathscr{D}(A)$, $e^{(t-s)A}B(s,y)z \in \mathscr{D}(A)$ *for each* $y \in H$, $z \in K$, *and* $t \geq s$*;*
(2) $\left\|Ae^{(t-s)A}F(s,y)\right\|_H \leq f(t-s)\|y\|_H$, $y \in H$, *for some* $f \in L^1([0,T];\mathbb{R}_+)$*;*
(3) $\left\|Ae^{(t-s)A}B(s,y)\right\|_{\mathscr{L}_2(K_Q,H)} \leq g(t-s)\|y\|_H$, $y \in H$, *for some* $g \in L^2([0,T];\mathbb{R}_+)$.

Then for any mild solution $y(t)$, $t \in [0,T]$, *of (1.4.2), it is also a solution of (1.4.2) in the strong sense.*

Proof By the conditions (1), (2), and (3), it is easy to see that

$$\int_0^T \int_0^t \|Ae^{(t-r)A}F(r,y(r))\|_H drdt < \infty \qquad a.s.$$

$$\int_0^T \int_0^t \|Ae^{(t-r)A}B(r,y(r))\|^2_{\mathscr{L}_2(K_Q,H)} drdt < \infty \qquad a.s.$$

Thus by the classic Fubini's theorem and Proposition 1.1.20, we have

$$\begin{aligned}\int_0^t \int_0^s Ae^{(s-r)A}F(r,y(r))drds &= \int_0^t \int_r^t Ae^{(s-r)A}F(r,y(r))dsdr \\ &= \int_0^t e^{(t-r)A}F(r,y(r))dr - \int_0^t F(r,y(r))dr.\end{aligned} \tag{1.4.8}$$

Meanwhile, by virtue of Proposition 1.3.5,

$$\begin{aligned}\int_0^t \int_0^s Ae^{(s-r)A}B(r,y(r))dW(r)ds &= \int_0^t \int_r^t Ae^{(s-r)A}B(r,y(r))dsdW(r) \\ &= \int_0^t e^{(t-r)A}B(r,y(r))dW(r) \\ &\quad - \int_0^t B(r,y(r))dW(r).\end{aligned} \tag{1.4.9}$$

Hence, by the closedness of A, (1.4.8), and (1.4.9), it follows that $Ay(t) \in H$, $t \in [0,T]$, which is integrable almost surely and

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

$$\int_0^t Ay(s)ds = e^{tA}y_0 - y_0 + \int_0^t e^{(t-r)A}F(r, y(r))dr - \int_0^t F(r, y(r))dr$$
$$+ \int_0^t e^{(t-r)A}B(r, y(r))dW(r) - \int_0^t B(r, y(r))dW(r)$$
$$= y(t) - y_0 - \int_0^t F(r, y(r))dr - \int_0^t B(r, y(r))dW(r).$$

That is, y is also a solution of (1.4.2) in the strong sense. The proof is now complete. □

To employ Itô's formula in handling the mild solutions of (1.4.2), we introduce a Yosida approximating system of (1.4.2) in the following form:

$$\begin{cases} dy(t) = Ay(t)dt + R(n)F(t, y(t))dt + R(n)B(t, y(t))dW(t), \\ y(0) = R(n)y_0 \in \mathscr{D}(A), \end{cases} \tag{1.4.10}$$

where $n \in \rho(A)$, the resolvent set of A, $R(n) := nR(n, A)$ and $R(n, A) = (nI - A)^{-1}$ is the resolvent of A.

Proposition 1.4.5 *Let $T \geq 0$ and $p \geq 2$. Suppose that the nonlinear terms $F(\cdot,\cdot)$, $B(\cdot,\cdot)$ in (1.4.10) satisfy condition (1.4.6). Then, for each $n \in \rho(A)$, the equation (1.4.10) has a unique solution $y_n(t) \in \mathscr{D}(A)$ in the strong sense, which lies in $L^p(\Omega; C([0, T]; H))$. In addition, if $\mathbb{E}\|y_0\|_H^p < \infty$, $p > 2$, and we let y be the mild solution of (1.4.2), then we have*

$$\lim_{n\to\infty} \mathbb{E}\left(\sup_{0\leq t\leq T} \|y_n(t) - y(t)\|_H^p\right) = 0, \qquad p > 2. \tag{1.4.11}$$

Proof The existence of a unique mild solution $y_n \in C([0, T]; L^p(\Omega; H))$, $n \in \rho(A)$, of (1.4.10) is an immediate consequence of Theorem 1.4.2 through a probabilistic fixed-point theorem type of argument. The fact that $y_n \in L^p(\Omega; C([0, T]; H))$ and y_n is also a solution of (1.4.10) in the strong sense follows from Proposition 1.4.4 and the relation

$$AR(n) = nAR(n, A) = n - n^2R(n, A) \in \mathscr{L}(H), \qquad n \in \rho(A).$$

To prove the remainder of the proposition, let us suppose that $\mathbb{E}\|y_0\|_H^p < \infty$, $p > 2$, and consider for any $t \in [0, T]$,

$$\begin{aligned} y(t) - y_n(t) = e^{tA}&(y_0 - R(n)y_0) \\ &+ \int_0^t e^{(t-s)A}[F(s, y(s)) - R(n)F(s, y_n(s))]ds \\ &+ \int_0^t e^{(t-s)A}[B(s, y(s)) - R(n)B(s, y_n(s))]dW(s). \end{aligned} \tag{1.4.12}$$

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

Since $|a+b+c|^p \le 3^p(|a|^p+|b|^p+|c|^p)$ for any real numbers a, b, c, this yields, in addition to (1.4.12), that for any $T \ge 0$, $p > 2$,

$$
\begin{aligned}
&\mathbb{E}\sup_{0\le t\le T}\|y(t)-y_n(t)\|_H^p \\
&\quad\le 3^p\mathbb{E}\sup_{0\le t\le T}\left\|\int_0^t e^{(t-s)A}R(n)[F(s,y(s))-F(s,y_n(s))]ds\right\|_H^p \\
&\qquad + 3^p\mathbb{E}\sup_{0\le t\le T}\left\|\int_0^t e^{(t-s)A}R(n)[B(s,y(s))-B(s,y_n(s))]dW(s)\right\|_H^p \\
&\qquad + 3^p\Bigg\{\mathbb{E}\sup_{0\le t\le T}\Bigg\|e^{tA}(y_0-R(n)y_0)+\int_0^t e^{(t-s)A}[I-R(n)]F(s,y(s))ds \\
&\qquad + \int_0^t e^{(t-s)A}[I-R(n)]B(s,y(s))dW(s)\Bigg\|_H^p\Bigg\} \\
&:= 3^pI_1+3^pI_2+3^pI_3.
\end{aligned}
\tag{1.4.13}
$$

Note that by the Hille–Yosida theorem, $\|R(n)\| \le 2M$ for an $n \in \mathbb{N}_+$ large enough where $M \ge 1$ is the number given in (1.1.5). Condition (1.4.6) and Hölder's inequality imply that

$$
\begin{aligned}
I_1 &\le \mathbb{E}\sup_{0\le t\le T}\left(\int_0^t \left\|e^{(t-s)A}R(n)\big[F(s,y(s))-F(s,y_n(s))\big]\right\|_H ds\right)^p \\
&\le C_1(T)\mathbb{E}\sup_{0\le t\le T}\left\{\int_0^t \|F(s,y(s))-F(s,y_n(s))\|_H^p ds\right\} \\
&\le C_2(T)\mathbb{E}\int_0^T \sup_{0\le r\le s}\|y(r)-y_n(r)\|_H^p ds,
\end{aligned}
\tag{1.4.14}
$$

where $C_1(T)$, $C_2(T)$ are positive numbers, dependent on $T \ge 0$. In a similar way, by virtue of Theorem 1.3.3, for $n \in \mathbb{N}_+$ large enough there exists a real number $C_3(T) > 0$ such that

$$
\begin{aligned}
I_2 &\le \mathbb{E}\sup_{0\le t\le T}\left\|\int_0^t e^{(t-s)A}R(n)[B(s,y(s))-B(s,y_n(s))]dW(s)\right\|_H^p \\
&\le C_3(T)\mathbb{E}\int_0^T \sup_{0\le r\le s}\|y(r)-y_n(r)\|_H^p ds.
\end{aligned}
\tag{1.4.15}
$$

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

For the term I_3, it is easy to see that

$$I_3 \le 3^p \Big\{ \mathbb{E} \sup_{0\le t\le T} \|e^{tA}(y_0 - R(n)y_0)\|_H^p + \mathbb{E} \sup_{0\le t\le T} \left\| \int_0^t e^{(t-s)A}[I - R(n)]F(s, y(s))ds \right\|_H^p + \mathbb{E} \sup_{0\le t\le T} \left\| \int_0^t e^{(t-s)A}[I - R(n)]B(s, y(s))dW(s) \right\|_H^p \Big\}. \tag{1.4.16}$$

We now estimate each term at the right-hand side of (1.4.16). By the Dominated Convergence Theorem and the fact that $R(n) \to I$ strongly as $n \to \infty$, it is easy to see that

$$\mathbb{E} \sup_{0\le t\le T} \|e^{tA}(y_0 - R(n)y_0)\|_H^p \le C_4(T) \cdot \mathbb{E}\|y_0 - R(n)y_0\|_H^p \to 0, \quad n \to \infty,$$

where $C_4(T) > 0$ is some positive number. On the other hand, by using the Hölder inequality and Dominated Convergence Theorem, we get for some $C_5(T) > 0$ that

$$\begin{aligned} &\mathbb{E} \sup_{0\le t\le T} \left\| \int_0^t e^{(t-s)A}[I - R(n)]F(s, y(s))ds \right\|_H^p \\ &\le C_5(T) \int_0^T \mathbb{E}\|[I - R(n)]F(s, y(s))\|_H^p ds \\ &\to 0 \quad \text{as} \quad n \to \infty. \end{aligned} \tag{1.4.17}$$

In a similar manner, by using Theorem 1.3.3 and the Dominated Convergence Theorem again, we have that for some $C_6(T) > 0$,

$$\begin{aligned} &\mathbb{E} \sup_{0\le t\le T} \left\| \int_0^t e^{(t-s)A}[I - R(n)]B(s, y(s))dW(s) \right\|_H^p \\ &\le C_6(T) \int_0^T \mathbb{E}\|[I - R(n)]B(s, y(s))\|_{\mathscr{L}_2(K_Q, H)}^p ds \to 0 \quad \text{as} \quad n \to \infty. \end{aligned} \tag{1.4.18}$$

Combining (1.4.13) through (1.4.18), we thus have that there exist numbers $C(T) > 0$ and $\varepsilon(n) > 0$ such that

$$\mathbb{E} \sup_{0\le t\le T} \|y(t) - y_n(t)\|_H^p \le C(T) \int_0^T \mathbb{E} \sup_{0\le r\le s} \|y(r) - y_n(r)\|_H^p ds + \varepsilon(n),$$

where $\lim_{n\to\infty} \varepsilon(n) = 0$. By the well-known Gronwall's inequality, we further deduce that

$$\mathbb{E} \sup_{0\le t\le T} \|y(t) - y_n(t)\|_H^p \le \varepsilon(n)e^{C(T)T} \to 0, \quad \text{as} \quad n \to \infty. \tag{1.4.19}$$

The proof is thus complete. □

Corollary 1.4.6 *Let $y_0 \in H$ be an arbitrarily given nonrandom vector. Suppose that the nonlinear terms $F(\cdot,\cdot)$, $B(\cdot,\cdot)$ in (1.4.2) and (1.4.10) satisfy condition (1.4.6) for any $T \geq 0$. Then there exists a sequence $y_n(t) \in \mathscr{D}(A)$ of solutions to (1.4.10) in the strong sense, which lies in $L^p(\Omega; C([0,T]; H))$, $p > 2$, such that $y_n(t) \to y(t)$ almost surely as $n \to \infty$, uniformly on any compact set of* $[0,\infty)$.

Proof We may construct the desired sequence by a diagonal sequence trick. Indeed, by virtue of (1.4.11) there exists a positive integer sequence $\{n_1(i)\}$ in $\rho(A)$ such that $y_{n_1(i)}(t) \to y(t)$ almost surely as $i \to \infty$, uniformly with respect to $t \in [0,1]$. Now consider the sequence $y_{n_1(i)}(t)$. We can find a subsequence $y_{n_2(i)}(t)$ of $y_{n_1(i)}(t)$ such that $y_{n_2(i)}(t) \to y(t)$ almost surely as $i \to \infty$, uniformly with respect to $t \in [0,2]$. Proceeding inductively, we find successive subsequences $y_{n_m(i)}(t)$ such that (a) $y_{n_m(i)}(t)$ is a subsequence of $y_{n_{m-1}(i)}(t)$ and (b) $y_{n_m(i)} \to y(t)$ almost surely as $i \to \infty$, uniformly with respect to $t \in [0,m]$. To get a sequence converging for each m, one may take the diagonal sequence $\hat{n}(m) := \{n_m(m)\}$. Then the sequence $y_{\hat{n}(m)}(t)$, $y_{\hat{n}(m+1)}(t), \ldots$ is a subsequence of $y_{n_m(i)}(t)$ so that $y_{\hat{n}(i)}(t) \to y(t)$ almost surely as $i \to \infty$, uniformly with respect to $t \in [0,m]$ for each $m \in \mathbb{N}_+$. □

Remark 1.4.7 In general, it is not immediate to know from Theorem 1.4.2 that the mild solution of (1.4.2) has almost surely continuous paths. However, Corollary 1.4.6 permits a modification of any mild solution of (1.4.2) with continuous sample paths. Unless otherwise stated, we always assume that the mild solution of this kind of equation under investigation has continuous sample paths in the sequel.

1.4.2 Variational Approach and Strong Solutions

Let V be a reflexive Banach space which is densely and continuously embedded in a Hilbert space H. We identity H with its dual space H^* according to Theorem 1.1.4. Then we have the following relations:

$$V \hookrightarrow H \cong H^* \hookrightarrow V^*$$

where $\hookrightarrow$ denotes the injection. We denote the duality pair between V and V^* by $\langle\!\langle \cdot,\cdot \rangle\!\rangle_{V,V^*}$. Let K be a separable Hilbert space and assume that $W(t)$, $t \geq 0$, is a Q-Wiener process in K defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$, equipped with a normal filtration $\{\mathscr{F}_t\}_{t\geq 0}$ with respect to which $\{W(t)\}_{t\geq 0}$ is a continuous martingale.

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

Let $T \geq 0$ and consider the following nonlinear stochastic differential equation on V^*:

$$\begin{cases} y(t) = y(0) + \int_0^t A(s, y(s))ds + \int_0^t B(s, y(s))dW(s), \qquad t \in [0, T], \\ y(0) = y_0 \in H, \end{cases} \tag{1.4.20}$$

where $A\colon [0,T] \times V \to V^*$ and $B\colon [0,T] \times V \to \mathscr{L}_2(K_Q, H)$ are two families of nonlinear measurable functions (they may be random as well in an appropriate setting). For any $y_0 \in H$, an $\{\mathscr{F}_t\}_{t\geq 0}$-adapted, V-valued process y is said to be a *strong solution* of the equation (1.4.20) if $y \in L^p([0,T] \times \Omega; V)$ for some $p \geq 1$ and the equation (1.4.20) holds in V^* almost surely. In contrast with system (1.4.2), one remarkable feature of (1.4.20) is that both mappings A and B could be nonlinear here.

To obtain the existence and uniqueness of solutions to (1.4.20), we impose the following conditions on $A(\cdot,\cdot)$ and $B(\cdot,\cdot)$.

(a) (Coercivity) There exist numbers $p > 1, \alpha > 0, \lambda \in \mathbb{R}$ and function $\gamma \in L^1(0,T)$ such that for all $y \in V$ and $t \in [0,T]$,

$$2\langle\!\langle y, A(t,y)\rangle\!\rangle_{V,V^*} + \|B(t,y)\|^2_{\mathscr{L}_2(K_Q,H)} \leq -\alpha\|y\|^p_V + \lambda\|y\|^2_H + \gamma(t), \tag{1.4.21}$$

(b) (Boundedness) and there exists a function $\theta \in L^{\frac{p}{p-1}}(0,T)$ such that for all $y \in V, t \in [0,T]$,

$$\|A(t,y)\|_{V^*} \leq \theta(t) + c\|y\|^{p-1}_V \tag{1.4.22}$$

for some number $c > 0$.

(c) (Continuity) The map $s \in \mathbb{R} \to \langle\!\langle x,\ A(t, y+sz)\rangle\!\rangle_{V,V^*}$ is continuous for arbitrary $y, z, x \in V$ and $0 \leq t \leq T$.

(d) (Monotonicity) There exists a number $\mu \in \mathbb{R}$ such that for any $y, z \in V$, and $t \in [0,T]$,

$$\begin{aligned} &2\langle\!\langle y - z, A(t,y) - A(t,z)\rangle\!\rangle_{V,V^*} + \|B(t,y) - B(t,z)\|^2_{\mathscr{L}_2(K_Q,H)} \\ &\quad \leq \mu\|y - z\|^2_H. \end{aligned} \tag{1.4.23}$$

Theorem 1.4.8 *Assume that $y_0 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P}; H)$. Under the assumptions (a)–(d), equation (1.4.20) has a unique $\{\mathscr{F}_t\}_{t\geq 0}$-progressively measurable strong solution*

$$y \in L^2(\Omega; C([0,T]; H)) \cap L^2([0,T] \times \Omega; V) \qquad \textit{for any} \qquad T \geq 0,$$

which has strong Markov property and satisfies the following energy equation: for all $t \in [0,T]$,

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

$$\|y(t)\|_H^2 = \|y_0\|_H^2 + 2\int_0^t \langle\langle y(s), A(s,y(s))\rangle\rangle_{V,V^*} ds \\ + 2\int_0^t \langle y(s), B(s,y(s))dW(s)\rangle_H + \int_0^t \|B(s,y(s))\|^2_{\mathscr{L}_2(K_Q,H)} ds. \tag{1.4.24}$$

Equality (1.4.24) is the usual Itô's formula for quadratic function $\Lambda(\cdot) = \|\cdot\|_H^2$. To extend this formula to more general function Λ, one need impose stronger conditions on Λ. A function $\Lambda\colon [0,T] \times H \to \mathbb{R}$ is called an *Itô type of functional* if it satisfies:

(i) Λ has locally bounded partial derivatives $\partial_t \Lambda$, $\partial_x \Lambda$ and $\partial_{xx}^2 \Lambda$ on $[0,T] \times H$;

(ii) $\partial_t \Lambda$ and $\partial_x \Lambda$ are continuous in $[0,T] \times H$;

(iii) for any trace class operator P, the map $(t,x) \to Tr[\partial_{xx}^2 \Lambda(t,x)P]$ is continuous on $[0,T] \times H$;

(iv) if $x \in V$, then $\partial_x \Lambda(t,x) \in V$ for any $t \in [0,T]$ and $\langle\langle \partial_x \Lambda(t,x), v^*\rangle\rangle_{V,V^*}$ is continuous in $t \in [0,T]$ for any $v^* \in V^*$. Moreover, there exists a number $M > 0$ such that

$$\|\partial_x \Lambda(t,x)\|_V \le M(1 + \|x\|_V), \qquad (t,x) \in [0,T] \times V. \tag{1.4.25}$$

Theorem 1.4.9 *Let $y \in L^2(\Omega; C([0,T]; H)) \cap L^2([0,T] \times \Omega; V)$, $T \ge 0$, be the strong solution of (1.4.20) with $y_0 \in L^2(\Omega, \mathscr{F}_0, \mathbb{P}; H)$. For any Itô type of functional Λ satisfying (i), (ii), (iii) and (iv) on $[0,T] \times H$, the following Itô's formula holds: for $t \in [0,T]$,*

$$\Lambda(t, y(t)) = \Lambda(0, y_0) + \int_0^t (\mathcal{L}\Lambda)(s, y(s)) ds \\ + \int_0^t \langle \partial_x \Lambda(s, y(s)), B(s, y(s)) dW(s)\rangle_H,$$

where

$$(\mathcal{L}\Lambda)(s, y(s)) = \partial_s \Lambda(s, y(s)) + \langle\langle \partial_x \Lambda(s, y(s)), A(s, y(s))\rangle\rangle_{V,V^*} \\ + \frac{1}{2} Tr\Big[\partial_{xx}^2 \Lambda(s, y(s)) B(s, y(s)) Q^{1/2} (B(s, y(s)) Q^{1/2})^*\Big]. \tag{1.4.26}$$

1.5 Definitions and Methods of Stochastic Stability

The term stability is one that has a variety of different meanings within mathematics. One often says that a system is stable if it is "continuous" with

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

respect to initial conditions. Precisely, suppose that $y(t) = y(t, y_0)$, $t \geq 0$, is a solution to some differential equation on a Hilbert space H,

$$\begin{cases} dy(t) = f(t, y(t))dt, & t \geq 0, \\ y(0) = y_0 \in H, \end{cases} \tag{1.5.1}$$

where $f(\cdot,\cdot)$ is a properly given function. Let $\tilde{y}(t)$, $t \geq 0$, be a particular solution to (1.5.1) and the corresponding system is thought of as describing a process without perturbations. Those systems associated with other solutions $y(t)$ are regarded as perturbed ones. When we talk about stability of the solution $\tilde{y}(t)$, $t \geq 0$, it means that the norm $\|y(t) - \tilde{y}(t)\|_H$, $t \geq 0$, could be made smaller and smaller if the initial perturbation scale $\|y(0) - \tilde{y}(0)\|_H$ is sufficiently small.

Another notion of stability is that of asymptotic stability. Here we say an equation is stable if all of its solutions get close to some nice solution $\tilde{y}$, e.g., equilibrium solution, as time goes to infinity. In most situations, it is enough to consider asymptotic stability of the null solution for some relevant system. Indeed, let $z(t) = y(t) - \tilde{y}(t)$ in (1.5.1), then the equation (1.5.1) could be rewritten as

$$\begin{aligned} dz(t) &= dy(t) - d\tilde{y}(t) \\ &= [f(t, z(t) + \tilde{y}(t)) - f(t, \tilde{y}(t))]dt \\ &=: F(t, z(t))dt, \qquad t \geq 0, \end{aligned} \tag{1.5.2}$$

where $F(t, 0) = 0$, $t \geq 0$. Note that if $z(0) = z_0 = 0$, it is immediate that the null is the unique solution to system (1.5.2). Hence, this treatment could be thought of as considering asymptotic stability of this null solution.

Definition 1.5.1 The null solution of (1.5.2) is said to be *stable* if for arbitrarily given $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that the relation $\|z_0\|_H < \delta$ implies

$$\|z(t, z_0)\|_H < \varepsilon \qquad \text{for all} \qquad t \geq 0. \tag{1.5.3}$$

Definition 1.5.2 The null solution of (1.5.2) is said to be *asymptotically stable* if it is stable and there exists $\delta > 0$ such that the relation $\|z_0\|_H < \delta$ implies

$$\lim_{t\to\infty} \|z(t, z_0)\|_H = 0. \tag{1.5.4}$$

For any $z_0 \in H$, if there exists $T(z_0) \geq 0$ such that (1.5.3) or (1.5.4) remains valid for all $t \geq T(z_0)$, then the null solution of (1.5.2) is said to have *global stability*. In addition to asymptotic stability, one might also want to know the rate of convergence, which leads to the following notion.

Definition 1.5.3 The null solution of (1.5.2) is said to be *(asymptotic) exponentially stable* if it is asymptotically stable and there exist numbers $M > 0$ and $\mu > 0$ such that

$$\|z(t,z_0)\|_H \leq M\|z_0\|_H e^{-\mu t} \qquad \text{for all} \qquad t \geq 0. \tag{1.5.5}$$

There are at least three times as many definitions for the stability of stochastic systems as there are for deterministic ones. This is certainly because in a stochastic setting, there exist three basic types of convergence: convergence in probability, convergence in mean, and convergence in sample paths. The preceding deterministic stability definitions can be translated into a stochastic setting by properly interpreting the notion of convergence.

Consider the following stochastic differential equation on the Hilbert space H,

$$\begin{cases} dy(t) = A(t,y(t))dt + B(t,y(t))dW(t), & t \geq 0, \\ y(0) = y_0 \in H, \end{cases} \tag{1.5.6}$$

where $y_0 \in H$ is a nonrandom vector; W is an infinite-dimensional Q-Wiener process; and A, B are families of measurable mappings with $A(t,0) = 0$, $B(t,0) = 0$ for any $t \geq 0$.

Definition 1.5.4 (Stability in Probability) The null solution of (1.5.6) is said to be *stable or strongly stable in probability* if for arbitrarily given $\varepsilon_1, \varepsilon_2 > 0$, there exists $\delta = \delta(\varepsilon_1,\varepsilon_2) > 0$ such that the relation $\|y_0\|_H < \delta$ implies

$$\mathbb{P}\big\{\|y(t,y_0)\|_H > \varepsilon_1\big\} < \varepsilon_2 \qquad \text{for all} \qquad t \geq 0, \tag{1.5.7}$$

or

$$\mathbb{P}\left\{\sup_{t\geq 0}\|y(t,y_0)\|_H > \varepsilon_1\right\} < \varepsilon_2.$$

Definition 1.5.5 (Asymptotic Stability in Probability) The null solution of (1.5.6) is said to have *asymptotic stability or strongly asymptotic stability in probability* if it is stable or strongly stable in probability and for each $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that the relation $\|y_0\|_H < \delta$ implies

$$\lim_{t\to\infty} \mathbb{P}\big\{\|y(t,y_0)\|_H > \varepsilon\big\} = 0, \tag{1.5.8}$$

or

$$\lim_{T\to\infty} \mathbb{P}\left\{\sup_{t\geq T}\|y(t,y_0)\|_H > \varepsilon\right\} = 0.$$

Definition 1.5.6 (Stability in the pth Moment) The null solution of (1.5.6) is said to be *stable or strongly stable in the pth moment*, $p > 0$, if for arbitrarily

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

given $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that the relation $\|y_0\|_H < \delta$ implies

$$\mathbb{E}\|y(t, y_0)\|_H^p < \varepsilon \qquad \text{for all} \qquad t \geq 0, \tag{1.5.9}$$

or

$$\mathbb{E}\left\{\sup_{t\geq 0} \|y(t, y_0)\|_H^p\right\} < \varepsilon.$$

Definition 1.5.7 (Asymptotic Stability in the pth Moment) The null solution of (1.5.6) is said to have *asymptotic stability or strongly asymptotic stability in the pth moment*, $p > 0$, if it is stable or strongly stable in the pth moment and there exists $\delta > 0$ such that the relation $\|y_0\|_H < \delta$ implies

$$\lim_{t\to\infty} \mathbb{E}\|y(t, y_0)\|_H^p = 0, \tag{1.5.10}$$

or

$$\lim_{T\to\infty} \mathbb{E}\left\{\sup_{t\geq T} \|y(t, y_0)\|_H^p\right\} = 0.$$

Definition 1.5.8 (Pathwise Stability) The null solution of (1.5.6) is said to be *stable or strongly stable in sample paths* if for any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that the relation $\|y_0\|_H < \delta$ implies

$$\mathbb{P}\big\{\|y(t, y_0)\|_H > \varepsilon\big\} = 0 \qquad \text{for all} \qquad t \geq 0,$$

or

$$\mathbb{P}\left\{\sup_{t\geq 0} \|y(t, y_0)\|_H > \varepsilon\right\} = 0,$$

which means with probability one, all the paths of solutions are stable or strongly stable.

Definition 1.5.9 (Pathwise Asymptotic Stability) The null solution of (1.5.6) is said to have *asymptotic stability or strongly asymptotic stability in sample paths* if it is stable or strongly stable in probability and there exists $\delta > 0$ such that the relation $\|y_0\|_H < \delta$ implies

$$\mathbb{P}\left\{\lim_{t\to\infty} \|y(t, y_0)\|_H = 0\right\} = 1,$$

or

$$\mathbb{P}\left\{\lim_{T\to\infty} \sup_{t\geq T} \|y(t, y_0)\|_H = 0\right\} = 1.$$

For any $y_0 \in H$, if there exists $T(y_0) \geq 0$ such that the claims in Definitions 1.5.4 through 1.5.9 remain valid for all $t \geq T(y_0)$, then the system (1.5.6) is said to have its *global* stability, respectively.

In application, there exist various versions of stochastic stability that show explicitly the decay rate of systems. Most noteworthy is the pth moment or pathwise exponential stability. Let $\lambda\colon [0,\infty) \to (0,\infty)$ be a continuous function with $\lim_{t\to\infty} \lambda(t) = 0$.

Definition 1.5.10 For $p > 0$, the null solution of (1.5.6) is said to be *pth momently stable with rate* λ if for each $y_0 \in H$, there exists a number $M(y_0) > 0$ such that

$$\mathbb{E}\|y(t,y_0)\|_H^p \le M(y_0)\lambda(t) \qquad \text{for all} \qquad t \ge 0. \tag{1.5.11}$$

Definition 1.5.11 The null solution of (1.5.6) is said to have *almost sure stability with rate* λ if for each $y_0 \in H$, there exists a random variable $M(y_0) > 0$ such that

$$\|y(t,y_0)\|_H \le M(y_0)\lambda(t) \qquad \text{for all} \qquad t \ge 0 \quad \text{almost surely}. \tag{1.5.12}$$

In Definitions 1.5.10 and 1.5.11, if $\lambda(t) = e^{-\gamma t}$, $(1+t)^{-\gamma}$ or $(\ln(1+t))^{-\gamma}$, $t \ge 0$, for some constant $\gamma > 0$, the system (1.5.6) is said to have *exponential, polynomial, or logarithmic stability*, respectively. In general, if $\lambda\colon [0,\infty) \to (0,\infty)$ is a continuous function such that $\overline{\lim}_{t\to\infty} \lambda(t) < \infty$ and (1.5.11) or (1.5.12) holds, the system (1.5.6) is called *ultimately bounded* in the pth moment or almost sure sense.

Remark 1.5.12 The definition of the almost sure stability with rate $\lambda(t)$ can be equivalently stated in the following way: for each $y_0 \in H$, there exist a number $M(y_0) > 0$ and random time $T = T(y_0) \ge 0$ such that

$$\|y(t,y_0)\|_H \le M(y_0)\lambda(t) \qquad \text{for all} \qquad t \ge T(y_0) \quad \text{almost surely}.$$

Remark 1.5.13 All the preceding stability definitions remain meaningful if we remove the condition $A(t,0) = 0$, $B(t,0) = 0$ in (1.5.6). This fact leads to a natural generalization of all the stability concepts. That is, we say in this case that the solution of system (1.5.6) has a *decay*, e.g., exponential decay.

Remark 1.5.14 For stochastic stability, it is enough in most cases to consider a nonrandom initial $y_0 \in H$. To illustrate this, suppose for the moment that y_0 is random and the null solution of (1.5.6) with nonrandom initial $x \in H$ is stable in probability, i.e., for arbitrarily given ε_1, $\varepsilon_2 > 0$, there exist $\delta = \delta(\varepsilon_1,\varepsilon_2) > 0$, $T = T(\varepsilon_1,\varepsilon_2) \ge 0$ such that if $\|x\|_H < \delta$, then

$$\mathbb{P}\big\{\omega\colon \|y(t,x)\|_H > \varepsilon_1\big\} < \varepsilon_2 \qquad \text{for all} \qquad t \ge T.$$

Now suppose that $\|y_0(\omega)\|_H < \delta$ almost surely. Let $B_\delta = \{x \in H\colon \|x\|_H < \delta\}$ and define the law of $y_0(\omega)$ by

$$\mathbb{D}_{y_0}(A) = \mathbb{P}\{\omega \in \Omega\colon y_0(\omega) \in A\}, \qquad \forall A \in \mathscr{B}(H).$$

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

Then we have

$$\begin{aligned}\mathbb{P}\{\omega\colon \|y(t,y_0)\|_H > \varepsilon_1\} &= \int_{\{x\in H\colon \|x\|_H<\delta\}} \mathbb{P}\{\|y(t,x)\|_H > \varepsilon_1\}\mathbb{D}_{y_0}(dx)\\ &\le \int_{\{x\in H\colon \|x\|_H<\delta\}} \varepsilon_2 \mathbb{D}_{y_0}(dx)\\ &\le \varepsilon_2 \qquad \text{for all} \qquad t \ge T.\end{aligned}$$

Unless otherwise stated, we always assume in the sequel that the initial data of the system under consideration are nonrandom.

Remark 1.5.15 (Also see Example 6.7, pp. 225–226 in [103]) It is clear that strong stability of a stochastic system implies its stability. The converse statement is not true in general. To see this, let us consider a stochastic process on the circle with one unit radius,

$$\begin{cases} dy(t) = \left[-2\sin^2\dfrac{y(t)}{2} + \sin^3\dfrac{y(t)}{2}\cos\dfrac{y(t)}{2}\right]dt - 2\sin^2\dfrac{y(t)}{2}dw(t), \quad t \ge 0,\\ y(0) = y_0 < 0, \end{cases}$$

where $y(t)$ is the angle coordinate of a point on the circle and $w(t)$, $t \ge 0$, is a standard real Brownian motion.

It may be computed that the solution of this system is the process

$$y(t) = 2\text{arccot}\left(t + w(t) + \cot\frac{y_0}{2}\right), \qquad t \ge 0.$$

Since $t + w(t) \to \infty$ almost surely as $t \to \infty$, it is easy to see that the null solution is strongly unstable in sample paths. However, by a direct calculation one can show that the null solution is stable in the almost sure sense.

It is clear that (strong) stability in the pth moment of the null solution of (1.5.6) for any value of $p > 0$ implies its (strong) moment stability for every smaller value than p and (strong) stability in probability. On the other hand, one can easily show that the null solution could be the pth moment (strongly) stable for some $p > 0$ but not the qth moment (strongly) stable for $q > p$. The case most frequently discussed in the literature is (strong) moment stability with $p = 2$. We shall also refer to this case as (strong) *stability in mean square.*

Although some stability, for instance, Definition 1.5.7 or 1.5.10, does not appear to be as strong a restriction on systems as that given in Definition 1.5.4 or 1.5.5, there are significant implications in Definitions 1.5.7 and 1.5.10 for sample stability behavior. However, it is worth pointing out that stability of the moment alone does not always provide a satisfactory intuitive basis upon which to judge the stability characteristics of the systems of interest.

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

Example 1.5.16 Consider a simple one-dimensional linear Itô equation

$$dy(t) = ay(t)dt + by(t)dw(t), \qquad t \geq 0, \tag{1.5.13}$$

where $w(t)$, $t \geq 0$, is a standard one-dimensional Brownian motion, $y(0) = y_0 \in \mathbb{R}$, and a, b are real numbers.

A direct computation shows that the solution process $y(t)$, $t \geq 0$, is given by

$$y(t) = \exp\left\{bw(t) + (a - b^2/2)t\right\}y_0, \qquad t \geq 0. \tag{1.5.14}$$

Hence, by using the law of iterated logarithm for Brownian motion (cf. Revuz and Yor [197]), it is easy to deduce that the asymptotically exponential growth rate of solution y is given by

$$\overline{\lim_{t\to\infty}} \frac{\log|y(t)|}{t} = a - \frac{b^2}{2} \qquad a.s. \tag{1.5.15}$$

We then conclude that the null solution has global exponential stability in the almost sure sense if and only if $a < b^2/2$. On the other hand, using the standard exponential martingale properties for Brownian motion, it is also easy to see for any $n \in \mathbb{N}_+$ that

$$\mathbb{E}y(t)^n = y_0^n \cdot \exp\left\{(a - b^2/2)nt + \frac{b^2n^2}{2}t\right\}.$$

Hence, we conclude that the null solution has the global nth moment exponential stability if and only if $a < b^2(1-n)/2$. Therefore, unlike deterministic systems, for $a < 0$, the first moment is exponentially stable, but higher moments are probably unstable. For $a < -b^2/2$, the first and second moments are exponentially stable, and higher moments are probably unstable, etc. It seems difficult to associate a physical meaning to the behavior of a system, knowing only that the first nth moments are stable and all higher moments are unstable. On the other hand, it is clear from (1.5.15) that the stability of sample trajectories are determined by the algebraic sign of $a - b^2/2$ only. It is interesting to note that for $a < b^2/2$, the sample path possesses almost surely asymptotic stability, but it is possible that all moments will diverge exponentially. Hence, we see in this example that unlike deterministic systems, even though stability in mean square implies almost sure stability, almost sure stability need not imply the moment stability.

Remark 1.5.17 If a system is almost surely (strong) asymptotically stable, then it is also (strong) asymptotically stable in probability. From the analogy of deterministic stability, it seems reasonable to assume in Definition 1.5.9 almost sure (strong) stability rather than (strong) stability in probability. However, it

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

is worth pointing out this requirement is actually too strong. In fact, let us consider Example 1.5.16 again. By (1.5.14) and the properties of Brownian motion, it is easy to see that for no positive constant $\varepsilon > 0$ does there exist a number $\delta > 0$ such that almost all the sample trajectories of the solutions originating at $y_0 \neq 0$, $|y_0| < \delta$, remain in an ε-neighborhood of zero (i.e., not almost surely stable) even if the unperturbed term is very stable (i.e., $a < 0$) and $|b|$ is very small.

It is not always possible to get an explicit solution for a stochastic differential equation. Therefore, it is generally unrealistic to deal with stochastic stability problems in such a way as we did in Example 1.5.16. In the history of stability study, one of the most effective approaches is the so-called Lyapunov function method or Lyapunov's second (direct) method. To gain some insight into the main ideas of this method, let us analyze a simple situation.

Consider a nonnegative continuous function Λ on $\mathbb{R}^n$ with $\Lambda(0) = 0$ and $\Lambda(x) > 0$ for $x \neq 0$. Suppose that for some $\delta > 0$, the set $D_\delta = \{x \in \mathbb{R}^n : \Lambda(x) < \delta\}$ is bounded and $\Lambda(x)$ has continuous first-order derivatives in D_δ. Let $y(t) = y(t, y_0)$ be the unique solution of the initial value problem:

$$\begin{cases} dy(t) = f(y(t))dt, & t \geq 0, \\ y(0) = y_0 \in D_\delta \subset \mathbb{R}^n, \end{cases} \tag{1.5.16}$$

for a given function $f(\cdot) \in \mathbb{R}^n$ with $f(0) = 0$. Since $\Lambda(x)$ is continuous, the open set D_δ contains the origin and monotonically decreases to the singleton set $\{0\}$ as $\delta \downarrow 0$. If the total derivative $\dot{\Lambda}(y(t))$ of Λ, along the solution trajectory $y(t)$, satisfies

$$\dot{\Lambda}(y(t)) = f(y(t)) \cdot \left.\frac{d\Lambda(x)}{dx}\right|_{x=y(t)} = -k(y(t)) \leq 0, \qquad t \geq 0, \tag{1.5.17}$$

where $k(\cdot)$ is some nonnegative continuous function, then $\Lambda(y(t))$ is a nonincreasing function of t, i.e., $\Lambda(y_0) < \delta$ implies $\Lambda(y(t)) < \delta$ for all $t \geq 0$. In other words, $y_0 \in D_\delta$ implies that $y(t) \in D_\delta$ for all $t \geq 0$. This establishes the stability of the null solution to (1.5.16) in the sense of Lyapunov, and $\Lambda(x)$ is thus called a Lyapunov function of equation (1.5.16). If we further assume that $k(x) > 0$ for $x \in D_\delta \setminus \{0\}$, then $\Lambda(y(t))$, as a function of t, is strictly monotone decreasing. Moreover, we have from (1.5.17) that

$$0 < \Lambda(y_0) - \Lambda(y(t)) = \int_0^t k(y(s))ds < \infty \qquad \text{for all} \qquad t \in [0, \infty). \tag{1.5.18}$$

In this case, $\Lambda(y(t)) \to 0$ as $t \to \infty$ from (1.5.18) for sufficiently small $\delta > 0$ (otherwise, $\Lambda(y(t)) \geq \Lambda(y_0)$ for some sufficiently large $t > 0$). This further implies that $y(t) \to 0$ as $t \to \infty$, i.e., the null solution of system (1.5.16) is asymptotically stable.

It is possible to generalize the preceding Lyapunov function method to stochastic systems. For instance, let us consider a stochastic process $y(t) \in \mathbb{R}^n$, $t \geq 0$, on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$. At present, it is not realistic to require that $\dot{\Lambda}(y(t,\omega)) \leq 0$ for all $\omega \in \Omega$. What one can expect for stability is that the time derivative of the expectation of $\Lambda(y(t))$, denote it by $\mathcal{L}\Lambda(\cdot)$, is nonpositive, where

$$\mathcal{L}\Lambda(y_0) := \lim_{t\to 0^+} \frac{\mathbb{E}(\Lambda(y(t))) - \mathbb{E}\Lambda(y_0)}{t}, \qquad y_0 \in \mathbb{R}^n. \tag{1.5.19}$$

Here, the domain of $\mathcal{L}$ is defined as a family of those functions Λ for which (1.5.19) is well defined. This is a natural analogue of the total derivative of Λ along the process trajectory $y(t)$ to the deterministic case. Now suppose that there exists a Lyapunov function Λ satisfying the aforementioned conditions or

$$\mathcal{L}\Lambda(y_0) \leq 0, \quad y_0 \in \mathbb{R}^n,$$

then it is possible to show, usually under additional conditions such as a strong Markov property of y, that for any $t \geq s \geq 0$,

$$\mathbb{E}(\Lambda(y(t,y_0)) \mid \mathscr{F}_s^y) \leq \Lambda(y(s,y_0)) \qquad a.s.$$

This means that $\Lambda(y(t,y_0))$ is a nonnegative supermartingale, and by the well-known martingale convergence theorem, we may show that $\Lambda(y(t,y_0)) \to 0$, which further implies $y(t,y_0) \to 0$, almost surely as $t \to \infty$ and $\|y_0\|_{\mathbb{R}^n} \to 0$.

The Lyapunov function $\Lambda(\cdot)$ may be regarded as a generalized energy function of the system under investigation. The preceding argument illustrates the physical intuition that if the energy of a physical system is always decreasing near an equilibrium state, then the equilibrium state is stable.

Since Lyapunov's original work [159], the Lyapunov function method for stability has been extensively developed. The advantage of this method is that one can obtain considerable information about stability properties of a given system without being required to solve the system equation explicitly. The main drawback of this method is that there does not exist a general method to construct appropriate Lyapunov functions, especially for nonlinear systems. A stability criterion obtained in this manner, which usually provides only a sufficient condition, depends sensitively on the chosen Lyapunov function. In the remainder of this book, we shall mainly explore the Lyapunov function approach to establish a stochastic stability theory for infinite-dimensional stochastic differential equations.

1.6 Notes and Comments

All the material in Section 1.1 is standard, and the statement there is mainly based on Curtain and Zwart [49], Engel and Nagel [70], Kreyszig [112], Pazy [187], and Yosida [224]. The proof of Theorem 1.1.10 is sketched in Example A.4.2 in Curtain and Zwart [49]. A systematic presentation of the material in Sections 1.2 and 1.3 is given in Da Prato and Zabczyk [53]. Theorem 1.3.3 is presented in Da Prato and Zabczyk [53], and its version, Theorem 1.3.4, is established in Tubaro [212].

A systematic statement of the variational method for infinite-dimensional stochastic systems is presented by many authors such as Krylov and Rozovskii [113], Pardoux [184], and Prévôt and Röckner [190]. As for applications of semigroup approaches to infinite-dimensional stochastic systems, a comprehensive statement can be found in the existing literature such as Chow [41], Da Prato and Zabczyk [53], and Métivier [171]. Much material in Section 1.4.1 is taken from Ichikawa [91].

The stability of a real system is the ability of the system to resist an influence or disturbance unknown beforehand. The system is said to be stable if such a disturbance does not essentially change it. Indeed, an individual predictable process can be physically realized only if it is stable in the corresponding natural sense. For instance, we know from classical control theory that, before we can consider the design of a regulatory or tracking control system, we need to make sure that such a system is stable from input to output. For stability and the relevant Lyapunov function method of finite-dimensional deterministic systems, some systematic statements can be found in the literature, e.g., Hahn [82].

For a finite-dimensional stochastic system, there are two main techniques dealing with its stability properties. The first significantly extends Lyapunov's direct method for deterministic systems to a stochastic setting. The main ingredient here is a Lyapunov function, and as in the deterministic theory, a major difficulty in this method is to construct a suitable Lyapunov function to find the optimal stability condition for nonlinear stochastic systems. The earliest attempt to generalize the classic Lyapunov function method to stochastic stability goes back at least to Kats and Krasovskii [101]. During the initial development of the Lyapunov theory and method of stochastic stability, some confusion about the formulation of Lyapunov functions, their usefulness in application, and the relationship among the different concepts of stability existed. Kozin's survey [109] clarified some of the confusion and provided a good foundation for further development. Shortly, quite a few important works appeared, e.g., Kushner [114, 115] used martingale convergence techniques

https://doi.org/10.1017/9781108653039.002 Published online by Cambridge University Press

to develop a Lyapunov function theory for strong Markov processes and study related control problems, and Pinsky [188] introduced specific Lyapunov functions to handle stochastic stability of some Dirichlet problems in two dimensions. In the meanwhile, a comprehensive statement on stochastic stability theory was presented in Has'minskii [86] for diffusion processes given as the solutions of Itô's stochastic differential equations. For subsequent developments of this topic over the last several decades, the reader is referred to some informative monographs in this field such as Arnold [3], Khas'minskii [103], Kolmanovskii and Nosov [107], and Mao [165, 167], among others.

The other important development in finite-dimensional stochastic stability is the application of the so-called Lyapunov exponent method to stochastic systems. This is the stochastic counterpart of the notion of characteristic exponents introduced in Lyapunov's work on asymptotically exponential stability. Although the Lyapunov exponent method provides necessary and sufficient conditions for asymptotic (exponential) stability, this method needs to use sophisticated mathematical techniques, especially for nonlinear systems, and significant computational problems must be solved. Important studies of the Lyapunov exponent method, when applied to stochastic systems, have been made in such work as Arnold, Kliemann, and Oeljeklaus [7]; Arnold and Wihstutz [9]; Furstenberg [77, 78]; Khas'minskii [103]; Mohammed and Scheutzow [177, 178]; and Oseledec [182], among others.

Last, we mention some monographs that include chapters dealing with stability problems for finite-dimensional deterministic or stochastic systems with time delay: Hale [84], Kolmanovskii and Nosov [107], and Mao [165], among others.