Chapter 2
Mathematical Machinery

The purpose of this chapter is to introduce the necessary background from the semigroup theory, particularly, the Yosida approximations and their properties, analysis and probability in Banach spaces, including Itô stochastic calculus, stochastic convolution integrals, among others. As pointed out before, no attempt has been made to make the presentation self-contained as there are many excellent books available in the literature.

2.1 Semigroup Theory

Let $(X, ||\cdot||_X)$ be a Banach space.

Definition 2.1 A one parameter family $\{S(t) : 0 \leq t < \infty\}$ of bounded linear operators mapping X into X is a semigroup of bounded linear operators on X if

(i) $S(0) = I$, (I is the identity operator on X),
(ii) $S(t+s) = S(t)S(s)$ for every $t, s \geq 0$ (the semigroup property).

A semigroup of bounded linear operators, $\{S(t) : t \geq 0\}$, is uniformly continuous if

$$\lim_{t\downarrow 0} ||S(t) - I|| = 0.$$

The linear operator A defined by

$$D(A) = \{x \in X : \lim_{t\downarrow 0} \frac{S(t)x - x}{t} \text{ exists}\} \tag{2.1}$$

T. E. Govindan, *Yosida Approximations of Stochastic Differential Equations in Infinite Dimensions and Applications*, Probability Theory and Stochastic Modelling 79, DOI 10.1007/978-3-319-45684-3_2

and

$$Ax = \lim_{t \downarrow 0} \frac{S(t)x - x}{t} = \frac{d^+ S(t)x}{dt}\Big|_{t=0} \quad \text{for } x \in D(A), \tag{2.2}$$

is the infinitesimal generator of the semigroup $\{S(t) : t \geq 0\}$, where $D(A)$ is the domain of A.

Theorem 2.1 A linear operator A is the infinitesimal generator of a uniformly continuous semigroup if and only if A is a bounded linear operator.

Proof See Pazy [1, Theorem 1.2]. □

Definition 2.2 A semigroup $\{S(t) : t \geq 0\}$ of bounded linear operators on X is a strongly continuous semigroup of bounded linear operators if

$$\lim_{t \downarrow 0} S(t)x = x \quad \text{for every} \quad x \in X. \tag{2.3}$$

A strongly continuous semigroup of bounded linear operators on X will be called a C_0-semigroup. A C_0-semigroup $\{S(t) : t > 0\}$ is called compact if it is a compact operator.

Theorem 2.2 Let $\{S(t) : t \geq 0\}$ be a C_0-semigroup. There exist constants $\alpha \geq 0$ and $M \geq 1$ such that

$$||S(t)|| \leq Me^{\alpha t} \quad \text{for} \quad 0 \leq t < \infty. \tag{2.4}$$

Proof See Ahmed [1, Theorem 1.3.1]. □

Corollary 2.1 If $\{S(t) : t \geq 0\}$ is a C_0-semigroup then for every $x \in X$, $t \to S(t)x$ is a continuous function from R^+ into X.

Proof See Ahmed [1, Corollary 1.3.2]. □

Theorem 2.3 Let $\{S(t) : t \geq 0\}$ be a C_0-semigroup and let A be its infinitesimal generator. Then

(a) For $x \in X$,

$$\lim_{h \to 0} \frac{1}{h} \int_t^{t+h} S(s)xds = S(t)x.$$

(b) For $x \in X$,

$$\int_0^t S(t)xdx \in D(A) \quad \text{and} \quad A\left(\int_0^t S(t)xdx \right) = S(t)x - x.$$

(c) For $x \in D(A), S(t)x \in D(A)$ and

$$\frac{d}{dt} S(t)x = AS(t)x = S(t)Ax.$$

(d) For $x \in D(A)$,

$$S(t)x - S(s)x = \int_s^t S(\tau)Axd\tau = \int_s^t AS(\tau)xd\tau.$$

Proof See Pazy [1, Theorem 2.4]. □

Corollary 2.2 If A is the infinitesimal generator of a C_0-semigroup $\{S(t) : t \geq 0\}$, $D(A)$ is dense in X and A is a closed linear operator.

Proof See Pazy [1, Corollary 2.5]. □

2.1.1 The Hille-Yosida Theorem

Let $\{S(t) : t \geq 0\}$ be a C_0-semigroup. It follows from Theorem 2.2 that there exist constants $\alpha \geq 0$ and $M \geq 1$ such that $||S(t)|| \leq Me^{\alpha t}$ for $t \geq 0$. If $\alpha = 0$, $\{S(t) : t \geq 0\}$ is called uniformly bounded and if moreover $M = 1$ it is called a C_0-semigroup of contractions. If $M = 1$, $\{S(t) : t \geq 0\}$ is called a pseudo-contraction semigroup. A semigroup $\{S(t) : t \geq 0\}$ is said to be of negative type, or is exponentially stable if $||S(t)|| \leq Me^{-\alpha t}, t \geq 0$ for some constants $M > 0$ and $\alpha > 0$. This subsection is devoted to the characterization of the infinitesimal generators of C_0-semigroups of contractions. Conditions on the behavior of the resolvent of an operator A, which are necessary and sufficient for A to be the infinitesimal generator of a C_0-semigroup of contractions, are given.

Recall that if A is a linear, not necessarily bounded, operator in X, the resolvent set of A, $\rho(A)$, is the set of all complex numbers λ for which $\lambda I - A$ is invertible, i.e., $(\lambda I - A)^{-1}$ is a bounded linear operator in X. The family $R(\lambda, A) = (\lambda I - A)^{-1}, \lambda \in \rho(A)$ of bounded linear operators is called the resolvent of A.

Theorem 2.4 **(**Hille-Yosida**)** A linear (unbounded) operator A is the infinitesimal generator of a C_0-semigroup of contractions $\{S(t) : t \geq 0\}$ if and only if

(i) A is closed and $\overline{D(A)} = X$, and
(ii) the resolvent set $\rho(A)$ of A contains R^+ and for every $\lambda > 0$,

$$||R(\lambda, A)|| \leq \frac{1}{\lambda}. \tag{2.5}$$

*Proof (*Necessity*)* If A is the infinitesimal generator of a C_0-semigroup then it is closed and $\overline{D(A)} = X$ by Corollary 2.2. For $\lambda > 0$ and $x \in X$ let

$$R(\lambda)x = \int_0^\infty e^{-\lambda t} S(t)xdt \tag{2.6}$$

Since $t \to S(t)x$ is continuous and uniformly bounded, the integral in (2.6) exists as an improper Riemann integral and defines the bounded linear operator $R(\lambda)$ that satisfies

$$||R(\lambda)x|| \leq \int_0^\infty e^{-\lambda t}||S(t)x||dt \leq \frac{1}{\lambda}||x||. \tag{2.7}$$

Moreover, for $h > 0$,

$$\begin{aligned}\frac{S(h)-I}{h}R(\lambda)x &= \frac{1}{h}\int_0^\infty e^{-\lambda t}(S(t+h)x - S(t)x)dt \\ &= \frac{e^{\lambda h}-1}{h}\int_0^\infty e^{-\lambda t}S(t)xdt \\ &\quad -\frac{e^{\lambda h}}{h}\int_0^h e^{-\lambda t}S(t)xdt. \end{aligned}\tag{2.8}$$

As $h \downarrow 0$, the RHS of (2.8) converges to $\lambda R(\lambda)x - x$. This implies that for every $x \in X$ and $\lambda > 0$, $R(\lambda)x \in D(A)$ and $AR(\lambda) = \lambda R(\lambda) - I$, or

$$(\lambda I - A)R(\lambda) = I. \tag{2.9}$$

For $x \in D(A)$ we have

$$\begin{aligned}R(\lambda)Ax &= \int_0^\infty e^{-\lambda t}S(t)Axdt \\ &= \int_0^\infty e^{-\lambda t}AS(t)xdt \\ &= A\left(\int_0^\infty e^{-\lambda t}S(t)xdt\right) \\ &= AR(\lambda)x, \end{aligned}\tag{2.10}$$

where we used Theorem 2.3 (c) and the closedness of A. From (2.9) and (2.10) it follows that

$$R(\lambda)(\lambda I - A)x = x \quad \text{for} \quad x \in D(A). \tag{2.11}$$

Thus, $R(\lambda)$ is the inverse of $\lambda I - A$, it exists for all $\lambda > 0$ and satisfies the desired estimate (2.5). Conditions (i) and (ii) are therefore necessary. □

Next, in order to prove that the conditions (i) and (ii) are also sufficient for A to be the infinitesimal generator of a C_0-semigroup of contractions we will need some lemmas and Yosida approximations.

The proofs of the following two lemmas can be found in Pazy [1, pp. 9–10].

Lemma 2.1 Let A satisfy the hypothesis of Theorem 2.4 and let $R(\lambda, A) = (\lambda I - A)^{-1}$. Then

$$\lim_{\lambda \to \infty} \lambda R(\lambda, A)x = x \quad \text{for} \quad x \in X. \tag{2.12}$$

We now define, for every $\lambda > 0$, the Yosida approximation of A by

$$A_\lambda = \lambda A R(\lambda, A) = \lambda^2 R(\lambda, A) - \lambda I. \tag{2.13}$$

A_λ is an approximation of A in the following sense:

Lemma 2.2 Let A satisfy the hypothesis of Theorem 2.4. If A_λ is the Yosida approximation of A, then

$$\lim_{\lambda \to \infty} A_\lambda x = Ax \quad \text{for} \quad x \in D(A). \tag{2.14}$$

Lemma 2.3 Let A satisfy the hypothesis of Theorem 2.4. If A_λ is the Yosida approximation of A, then A_λ is the infinitesimal generator of a uniformly continuous semigroup of contractions $\{e^{tA_\lambda} : t \geq 0\}$. Furthermore, for every $x \in X$, $\lambda, \mu > 0$ we have

$$||e^{tA_\lambda}x - e^{tA_\mu}x|| \leq t||A_\lambda x - A_\mu x||. \tag{2.15}$$

Proof From (2.13) it is clear that A_λ is a bounded linear operator and hence is the infinitesimal generator of a uniformly continuous semigroup $\{e^{tA_\lambda} : t \geq 0\}$ of bounded linear operators (see Theorem 2.1). Moreover,

$$||e^{tA_\lambda}|| = e^{-t\lambda}||e^{t\lambda^2 R(\lambda,A)}|| \leq e^{-t\lambda}e^{t\lambda^2||R(\lambda,A)||} \leq 1 \tag{2.16}$$

and therefore $\{e^{tA_\lambda} : t \geq 0\}$ is a contraction semigroup. It is clear from the definitions that $e^{tA_\lambda}, e^{tA_\mu}, A_\lambda$ and A_μ commute with each other. Consequently,

$$\begin{aligned}
||e^{tA_\lambda}x - e^{tA_\mu}x|| &= ||\int_0^1 \frac{d}{ds}(e^{tsA_\lambda}e^{t(1-s)A_\mu}x)ds|| \\
&\leq \int_0^1 t||e^{tsA_\lambda}e^{t(1-s)A_\mu}(A_\lambda x - A_\mu x)||ds \\
&\leq t||A_\lambda x - A_\mu x||. \quad \square
\end{aligned}$$

Proof of Theorem 2.4 (Sufficiency) Let $x \in D(A)$. Then

$$\begin{aligned}
||e^{tA_\lambda}x - e^{tA_\mu}x|| &\leq t||A_\lambda x - A_\mu x|| \\
&\leq t||A_\lambda x - Ax|| + t||Ax - A_\mu x||.
\end{aligned} \tag{2.17}$$

From (2.17) and Lemma 2.2 it follows that for $x \in D(A), e^{tA_\lambda}x$ converges as $\lambda \to \infty$ and the convergence is uniform on bounded intervals. Since $D(A)$ is dense in X and $||e^{tA_\lambda}|| \leq 1$, it follows that

$$\lim_{\lambda\to\infty} e^{tA_\lambda}x = S(t)x \quad \text{for every} \quad x \in X. \tag{2.18}$$

The limit in (2.18) is again uniform on bounded intervals. From (2.18) it follows readily that the limit $S(t)$ satisfies the semigroup property, i.e., $S(0) = I$ and that $||S(t)|| \leq 1$. Also, $t \to S(t)x$ is continuous for $t \geq 0$ as a uniform limit of the continuous functions $t \to e^{tA_\lambda}x$. Thus $\{S(t) : t \geq 0\}$ is a C_0-semigroup of contractions on X. To conclude the proof we need to show that A is, in fact, the infinitesimal generator of $\{S(t) : t \geq 0\}$. Let $x \in D(A)$. Then using (2.18) and Theorem 2.3 we have

$$\begin{aligned} S(t)x - x &= \lim_{\lambda\to\infty} (e^{tA_\lambda}x - x) \\ &= \lim_{\lambda\to\infty} \int_0^t e^{sA_\lambda} A_\lambda x dx = \int_0^t S(s)Axds. \end{aligned} \tag{2.19}$$

The last equality follows from the uniform convergence of $e^{tA_\lambda}A_\lambda x$ to $S(t)Ax$ on bounded intervals. Let B be the infinitesimal generator of $\{S(t) : t \geq 0\}$ and let $x \in D(A)$. Dividing (2.19) by $t > 0$ and letting $t \downarrow 0$ we see that $x \in D(B)$ and that $Bx = Ax$. Thus $B \supseteq A$. Since B is the infinitesimal generator of $\{S(t) : t \geq 0\}$, it follows from the necessary conditions that $1 \in \rho(B)$. On the other hand, we assume (Hypothesis (ii)) that $1 \in \rho(A)$. Since $B \supseteq A$, $(I-B)D(A) = (I-A)D(A) = X$ which implies $D(B) = (I-B)^{-1}X = D(A)$ and therefore $A = B$. □

Hille-Yosida theorem has some simple consequences which are stated next.

Corollary 2.3 Let A be the infinitesimal generator of a C_0-semigroup of contractions $\{S(t) : t \geq 0\}$. If A_λ is the Yosida approximation of A, then

$$S(t)x = \lim_{\lambda\to\infty} e^{tA_\lambda}x \quad \text{for} \quad x \in X.$$

Proof See Pazy [1, Corollary 3.5]. □

Corollary 2.4 Let A be the infinitesimal generator of a C_0-semigroup of contractions $\{S(t) : t \geq 0\}$. The resolvent set of A contains the open right half-plane, i.e., $\rho(A) \supseteq \{\lambda : Re\lambda > 0\}$ and for such λ,

$$||R(\lambda, A)|| \leq \frac{1}{\mathrm{Re}\lambda}.$$

Proof See Pazy [1, Corollary 3.6]. □

Corollary 2.5 A linear operator A is the infinitesimal generator of a C_0-semigroup satisfying $||S(t)|| \leq e^{\alpha t}$ if and only if

(i) A is closed and $\overline{D(A)} = X$,
(ii) The resolvent set $\rho(A)$ of A contains the ray $\{\lambda : \mathrm{Im}\lambda = 0, \quad \lambda > \alpha\}$ and for such λ

$$||R(\lambda,A)|| \leq \frac{1}{\lambda - \alpha}.$$

Proof See Pazy [1, Corollary 3.8]. □

2.1.2 Yosida Approximations of Maximal Monotone Operators

Let X be a Banach space and X^* its dual space. Let $\mathcal{G}(A)$ denote the graph of the operator A.

Definition 2.3

(i) A multivalued operator $A : X \to 2^{X^*}$ is said to be monotone if

$$_{X^*}\langle y_1 - y_2, x_1 - x_2\rangle_X \geq 0, \quad \forall x_i, y_i \in \mathcal{G}(A), \quad i = 1,2.$$

(ii) A monotone operator $A : X \to 2^{X^*}$ is said to be maximal monotone if there exists no other proper monotone extension $\tilde{A}$ of A, i.e., $\mathcal{G}(A) \subsetneq \mathcal{G}(\tilde{A})$.

We now introduce Yosida approximation of a multivalued operator on Banach spaces. Let us assume that X is uniformly convex with uniformly convex dual X^*. Hence, by Theorem D.1, the duality mapping J is single-valued in view of Remark D.2.

For every $x \in X$ and $\lambda > 0$ let us consider the following resolvent equation:

$$0 \in J(x_\lambda - x) + \lambda A x_\lambda. \tag{2.20}$$

Proposition 2.1 For all $x \in X$, there exists a unique solution x_λ to (2.20).

Proof By Corollary D.1, λA is maximal monotone. By Proposition D.1 (i), J is monotone and demicontinuous (see Section 2.4). Further, let $\{x_n\}$ be a sequence such that $\lim_{n\to\infty} ||x_n|| = \infty$. Since

$$_{X^*}\langle J(x-y), x-y\rangle_X = ||x-y||^2 \quad \forall \quad x,y \in X,$$

we obtain

$$\lim_{n\to\infty} \frac{_{X^*}\langle J(x_n - \tilde{x}), x_n - \tilde{x}\rangle_X}{||x_n||} = \lim_{n\to\infty} \frac{||x_n - \tilde{x}||^2}{||x_n||} = \infty.$$

Therefore, the map $y \to J(y-\tilde{x})$ is coercive. Hence, applying Corollary C.1 it follows that the mapping $\tilde{A}: X \to 2^{X^*}$ defined by $x_\lambda \mapsto J(x_\lambda - x) + \lambda A x_\lambda$ is maximal monotone.

Claim For $x_0 \in D(A)$ the mapping $\bar{A}: x_\lambda \mapsto J(x_\lambda - x_0) + A x_\lambda$ is coercive.

Proof Take a sequence $\{x_n\} \subset D(A)$ such that $\lim_{n\to\infty} ||x_n|| = \infty$ and fix $y_n \in \bar{A}(x_n)$, i.e., $y_n = J(x_n - x_0) + \lambda v_n$ for some $v_n \in A(x_n)$. Then

$$\begin{aligned}
&\frac{{}_{X^*}\langle y_n, x_n - x_0\rangle_X}{||x_n||} \\
&= \frac{{}_{X^*}\langle J(x_n - x_0), x_n - x_0\rangle_X}{||x_n||} + \lambda \frac{{}_{X^*}\langle v_n, x_n - x_0\rangle_X}{||X_n||} \\
&= \frac{||x_n - x_0||^2}{||x_n||} + \lambda \frac{{}_{X^*}\langle v_n - w, x_n - x_0\rangle_X}{||x_n||} + \frac{{}_{X^*}\langle w, x_n - x_0\rangle_X}{||x_n||},
\end{aligned}$$

for $w \in A(x_0)$.

Clearly, $||x_n - x_0||^2/||x_n|| \to \infty$ as $n \to \infty$. By the monotonicity of A we get

$$\lambda \frac{{}_{X^*}\langle v_n - w, x_n - x_0\rangle_X}{||x_n||} \geq 0.$$

Further,

$$\frac{|{}_{X^*}\langle w, x_n - x_0\rangle_X|}{||x_n||} \leq \frac{||w||\,||x_n - x_0||}{||x_n||} < \infty.$$

Hence, $\lim_{n\to\infty} {}_{X^*}\langle y_n, x_n - x_0\rangle_X \, ||x_n||^{-1} = \infty$. By Proposition C.3, we obtain surjectivity of the map $x_\lambda \mapsto J(x_\lambda - \tilde{x}) + \lambda A x_\lambda$. Thus, there exists a solution x_λ to (2.20).

To show the uniqueness of the solution, let x_1, x_2 be two solutions of (2.20), i.e., $0 = J(x_i - \tilde{x}) + \lambda v_i$, for some $v_i \in A(x_i), i = 1,2$. Setting $\tilde{x}_i := x_i - \tilde{x}, i = 1,2$ by monotonicity of A and J we obtain

$$\begin{aligned}
0 &= {}_{X^*}\langle J(\tilde{x}_1) - J(\tilde{x}_2), \tilde{x}_1 - \tilde{x}_2\rangle_X + \lambda\, {}_{X^*}\langle v_1 - v_2, x_1 - x_2\rangle_X \\
&\geq {}_{X^*}\langle J(\tilde{x}_1) - J(\tilde{x}_2), \tilde{x}_1 - \tilde{x}_2\rangle_X \geq 0.
\end{aligned}$$

Hence ${}_{X^*}\langle J(\tilde{x}_1) - J(\tilde{x}_2), \tilde{x}_1 - \tilde{x}_2\rangle_X = 0$. Since J is strictly monotone (see Proposition D.1 (iii)), we conclude that $\tilde{x}_1 = \tilde{x}_2$ or equivalently, $x_1 = x_2$. □

Proposition 2.1 justifies the following definition.

Definition 2.4

(i) The resolvent $J_\lambda : X \to X$ of a maximal monotone operator A is defined by $J_\lambda x = x_\lambda$, where x_λ is the unique solution to (2.20).

(ii) The Yosida approximation $A_\lambda : X \to 2^{X^*}$ is given by

$$A_\lambda x = \frac{1}{\lambda} J(x - J_\lambda x), \quad \lambda > 0, \quad x \in X.$$

We have the following properties of the resolvent and the Yosida approximation.

Proposition 2.2

(i) A_λ is single-valued, maximal monotone, bounded on bounded subsets and demicontinuous from X to X^*.
(ii) $||A_\lambda x|| \leq ||A^0 x||$ for every $x \in D(A)$, $\quad \lambda > 0$.
(iii) J_λ is bounded on bounded subsets, demicontinuous and

$$\lim_{\lambda \to 0} J_\lambda x = x, \quad \forall x \in co\{D(A)\},$$

where $co\{\cdot\}$ denotes the closed convex hull of $\{\cdot\}$.
(iv) For $\lambda \to 0$, $A_\lambda x \to A^0 x$ for all $x \in D(A)$.
(v) For all $x \in X$, we have

$$A_\lambda(x) \in A(J_\lambda(x)).$$

(vi) If $\lambda_n \to 0, x_n \to x$ weakly, $A_{\lambda_n} x_n \to y$ weakly and

$$\limsup_{n,m\to\infty} {}_{X^*}\langle A_{\lambda_n} x_n - A_{\lambda_m} x_m, x_n - x_m \rangle_X \leq 0,$$

then $[x, y] \in \mathcal{G}(A)$ and

$$\lim_{n,m\to\infty} {}_{X^*}\langle A_{\lambda_n} x_n - A_{\lambda_m} x_m, x_n - x_m \rangle_X = 0.$$

Proof (i) According to Barbu [1, Section 2.1, Proposition 1.3], A_λ is single-valued, monotone, bounded on bounded subsets and demicontinuous. Applying Theorem C.1 it follows that A_λ is maximal monotone.

(ii)–(iv), (vi) See Barbu [1, Proposition 1.3].

(v) From (2.20) and the definition of J_λ, we conclude that

$$-J(J_\lambda(x) - x) \in \lambda A(J_\lambda(x)) \quad \forall x \in X.$$

Since J is odd, by the definition A_λ we obtain

$$A_\lambda(x) = \frac{1}{\lambda} J(x - J_\lambda(x)) = -\frac{1}{\lambda} J(J_\lambda(x) - x) \in A(J_\lambda(x)) \quad \forall x \in X. \quad \square$$

Instead of the implicit definition of the Yosida approximation as an operator depending on the resolvent which is implicitly defined via the resolvent equation (2.20), one can explicitly express the Yosida approximation in the following way.

Lemma 2.4 Let A_λ be the Yosida approximation of A. Then

$$A_\lambda(x) = \left(A^{-1} + \lambda J^{-1}\right)^{-1} x, \quad x \in X.$$

Proof Fix $x \in X$ and let $J_\lambda(x)$ be the resolvent of A defined by (2.20). Then, by the definition of the Yosida approximation and the homogeneity of the duality mapping J^{-1}, we have $J_\lambda(x) = x - \lambda J^{-1}(A_\lambda(x))$. Inserting this into the resolvent equation (2.20), we obtain $A_\lambda(x) \in A(x - \lambda J^{-1}(A_\lambda(x)))$ or equivalently,

$$x \in (A^{-1} + \lambda J^{-1})(A_\lambda(x)).$$

Since A_λ is single-valued, we conclude that $A_\lambda(x) = (A^{-1} + \lambda J^{-1})^{-1} x$. $\square$

The following lemma plays a fundamental role in the proof of existence and uniqueness of multivalued stochastic differential equations. It states that the coercivity of a maximal monotone operator is carried forward to its Yosida approximation.

Lemma 2.5 Let $\alpha \in (1,2]$, $A : X \to 2^{X^*}$ be a maximal monotone operator and A_λ its Yosida approximation. If for some constants $C_1 > 0$ and $C_2 \in R$,

$${}_{X*}\langle v, x\rangle_X \geq C_1 ||x||^\alpha + C_2 \quad \forall x \in D(A), \quad \forall v \in A(x),$$

then there exist $\lambda_0 > 0$ and $C > 0$ such that for all $0 < \lambda < \lambda_0$,

$${}_{X^*}\langle A_\lambda x, x\rangle_X \geq C_1 2^{-\alpha} ||x||^\alpha + C \quad \forall x \in X.$$

Proof Fix $x \in X$. By the definition of A_λ and a property of J we have

$$\begin{aligned} {}_{X^*}\langle A_\lambda x, x - J_\lambda x\rangle_X &= \frac{1}{\lambda}\, {}_{X^*}\langle J(x - J_\lambda x), x - J_\lambda x\rangle_X \\ &= \frac{1}{\lambda} ||x - J_\lambda x||^2. \end{aligned}$$

Since $A_\lambda(x) \in A(J_\lambda x)$ (see Proposition 2.2 (v)) and A is coercive we deduce that

$$\begin{aligned} {}_{X^*}\langle A_\lambda x, x\rangle_X &= {}_{X^*}\langle A_\lambda x, J_\lambda x\rangle_X + \frac{1}{\lambda} ||x - J_\lambda x||^2 \\ &\geq C_1 ||J_\lambda x||^\alpha + \frac{1}{\lambda} ||x - J_\lambda x||^2 + C_2 \\ &\geq C_1 ||J_\lambda x||^\alpha + \frac{1}{\lambda} ||x - J_\lambda x||^\alpha + C \end{aligned}$$

for some $C > 0$ since $\alpha \in (1,2]$. Further, for $\lambda_0 := \frac{1}{C_1}$ we have $(1/\lambda - C_1) \geq 0$ for all $0 < \lambda < \lambda_0$. Hence, we get

$$
\begin{aligned}
{}_{X^*}\langle A_\lambda x, x\rangle_X &= C_1||J_\lambda x||^\alpha + (\frac{1}{\lambda} - C_1)||x - J_\lambda x||^\alpha + C_1||x - J_\lambda x||^\alpha + C \\
&\geq C_1(||J_\lambda x||^\alpha + ||x - J_\lambda x||^\alpha) + C \\
&\geq C_1 2^{-\alpha+1}||x||^\alpha + C, \quad \forall \lambda < \lambda_0,
\end{aligned}
$$

by using $2^{\alpha-1}(a^\alpha + b^\alpha) \geq (a+b)^\alpha$ for $\alpha > 1$, $a, b \geq 0$. □

Note that in the Hilbert space case, the Yosida approximation is Lipschtiz continuous. However, in the Banach space case this is not necessarily true as the following example shows:

Example 2.1 Let $A := J$. Using Lemma 2.4, we derive its Yosida approximation:

$$
\begin{aligned}
A_\lambda(x) &= (J^{-1} + \lambda J^{-1})^{-1} x \\
&= \{y \in X^* | y = \left((1+\lambda)J^{-1}\right)^{-1} x\} \\
&= \{y \in X^* | (1+\lambda)J^{-1} y = x\} \\
&= \left\{y \in X^* | y = J\left(\frac{x}{1+\lambda}\right)\right\} = \frac{1}{1+\lambda} J(x).
\end{aligned}
$$

Since the duality map J is Lipschitz continuous, so is its Yosida approximation.

2.2 Yosida Approximations and The Central Limit Theorem

Paulauskas [1] proposed a new idea to obtain bounds for errors for some approximations of semigroups of operators using some methods and results of probability theory related to the central limit theorem. Bentkus [1] introduced a new approach for analysis of errors in central limit theorem and in approximations by accompanying laws. Bentkus and Paulauskas [1] demonstrated that this approach is also useful to get optimal convergence rates in some approximation formulas for operators. Vilkiene [1] used this method to obtain asymptotic expansions and optimal error bounds for Euler's approximations of semigroups.

In this section, we use this method to obtain optimal error bounds and asymptotic expansions for Yosida approximations of bounded holomorphic semigroups.

Our objective is to present here some recent results as an interesting connection between semigroup theory and probability theory for an interested reader. This section can be skipped without losing continuity from further reading of the book.

2.2.1 Optimal Convergence Rate for Yosida Approximations

Let $A_\lambda, \forall \lambda > 0$ be the Yosida approximation of A as defined earlier in (2.13). By Lemma 2.3, A_λ is the infinitesimal generator of a uniformly continuous semigroup of contractions $\{S_\lambda(t) : t \geq 0\}$. Moreover, by Corollary 2.3,

$$S(t)x = \lim_{\lambda \to \infty} S_\lambda(t)x, \quad \text{for} \quad x \in X. \tag{2.21}$$

We call $S_\lambda(t), \lambda > 0$ Yosida approximations of contraction semigroup $\{S(t) : t \geq 0\}$.

Definition 2.5 Let $\{S(t) : t \geq 0\}$ be a C_0-semigroup on a Banach space X. The semigroup $\{S(t) : t > 0\}$ is said to be differentiable if for every $x \in X$, the function $t \to S(t)x$ is differentiable for $t > 0$. A semigroup $S(t)$ is called differentiable if it is differentiable for $t > 0$.

One can show that the n-th derivative satisfies $S^{(n)}(t) = A^n S(t)$.

Definition 2.6 Let $\Sigma_\theta = \{z : |argz| < \theta\}$ be a sector in the complex plane for some $\theta > 0$ and for $z \in \Sigma_\theta$, let $S(z) \in L(X)$. The family $S(z), z \in \Sigma_\theta$ is called a holomorphic semigroup in Σ_θ if:

(i) the function $z \mapsto S(z)$ is analytic in Σ_θ,
(ii) $S(0) = I$ and $\lim_{z \to 0, z \in \Sigma_\theta} S(z)x = x$ for every $x \in X$, and
(iii) $S(z_1 + z_2) = S(z_1)S(z_2)$ for $z_1, z_2 \in \Sigma_\theta$.

A semigroup $\{S(t) : t \geq 0\}$ is called holomorphic if it is holomorphic in some sector Σ_θ containing the nonnegative real axis.

A semigroup $\{S(t) : t \geq 0\}$ is called bounded holomorphic semigroup in Σ_θ if it has a bounded holomorphic extension to $\Sigma_{\theta'}$ for each $\theta' \in (0, \theta)$. We call $\{S(t) : t \geq 0\}$ a bounded holomorphic semigroup if it is a bounded holomorphic semigroup in some sector $\Sigma_\theta, \theta > 0$. Note that if $S(t)$ is a bounded semigroup which is holomorphic, then it is not necessarily a bounded holomorphic semigroup (see W. Arendt, et al [1, p. 153]).

Assume that there exists a positive constant K independent of n, λ and t such that

$$||tAS(t)|| \leq K, \tag{2.22}$$

and

$$(n+1)||A\lambda^n(\lambda I - A)^{-n-1}|| \leq K, \quad n = 0, 1, 2, \ldots, \tag{2.23}$$

for all $\lambda > 0, t \geq 0$.

Note that bounded holomorphic semigroups satisfy (2.22) by Theorem 5.2 (see Pazy [1, p. 61]) and (2.23) by Theorem 5.5 (see Pazy [1, p. 65]).

Lemma 2.6 Let A be the infinitesimal generator of a contraction semigroup $\{S(t) : t \geq 0\}$. Suppose that the conditions (2.22) and (2.23) are satisfied. Then the Yosida approximations satisfy

$$||tA_\lambda S_\lambda(t)|| \leq K, \quad \forall\lambda > 0, \quad t \geq 0. \tag{2.24}$$

Proof We have

$$A_\lambda = \lambda A(\lambda I - A)^{-1} = \lambda^2(\lambda I - A)^{-1} - \lambda I.$$

Expanding $\exp\{t\lambda^2(\lambda I - A)^{-1}\}$ as a Taylor series, we obtain

$$\begin{aligned} tA_\lambda S_\lambda(t) &= tA_\lambda e^{tA_\lambda} \\ &= e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^{n+1}}{n!} A\lambda^n(\lambda I - A)^{-n-1}. \end{aligned}$$

From (2.23), we get

$$\begin{aligned} ||tA_\lambda S_\lambda(t)|| &\leq Ke^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^{n+1}}{(n+1)!} \\ &= k(1 - e^{-t\lambda}) \leq K, \end{aligned}$$

for all $\lambda > 0$ and $t \geq 0$. □

If (2.22) and (2.24) hold, then

$$\begin{aligned} ||(tA)^m S(t)|| &\leq m^m k^m \quad \text{and} \\ ||(tA_\lambda)^m S_\lambda(t)|| &\leq m^m k^m, \end{aligned} \tag{2.25}$$

for all $t \geq 0, \lambda > 0$ and, $m = 1, 2, \ldots$ see Lemma 2.1 (see Vilkiene [1]).

In the next subsection, we shall prove the integro-differential identity

$$S_\lambda(t)x = S(t)x + \frac{a_1}{\lambda} + \frac{a_2}{\lambda^2} + \cdots + \frac{a_k}{\lambda^k} + D_k, \tag{2.26}$$

where $\{S_\lambda(t) : t \geq 0\}$ is the Yosida approximation of the semigroup $\{S(t) : t \geq 0\}$ and the coefficients a_m do not depend on λ.

In what follows we obtain the optimal bound for the convergence rate $||S(t)\, x - S_\lambda(t)x||$.

Theorem 2.5 Let the semigroups $\{S(t) : t \geq 0\}$ and $\{S_\lambda(t) : t \geq 0\}$ satisfy the conditions (2.22) and (2.24). Then the following integro-differential identity holds:

$$D_0 = S_\lambda(t)x - S(t)x = \frac{1}{\lambda}\int_0^1 tAA_\lambda S_\lambda((1-\tau)t)S(\tau t)x d\tau, \tag{2.27}$$

for all $\lambda > 0$. Moreover, the following inequality holds

$$||S(t)x - S_\lambda(t)x|| \leq \frac{CK||Ax||}{\lambda}, \tag{2.28}$$

where C is some absolute positive constant.

Proof The proof is based on an application of Newton-Leibnitz formula along a smooth curve $\gamma(\tau)$ connecting two close objects a and b such that $b - a = \gamma(1) - \gamma(0) = \int_0^1 \gamma'(\tau)d\tau$. We choose γ in the following form

$$\gamma(\tau) = S_\lambda((1-\tau)t)S(\tau t). \tag{2.29}$$

Then $a = S_\lambda(t), b = S(t)$ and

$$\begin{aligned}
\gamma'(\tau) &= \big(S_\lambda((1-\tau)t)\big)' S(\tau t) + S_\lambda\big((1-\tau)t\big)(S(\tau t))' \\
&= -A_\lambda t S_\lambda\big((1-\tau)t\big)S(\tau t) + AtS_\lambda\big((1-\tau)t\big)S(\tau t) \\
&= t(A - A_\lambda)\gamma(\tau) = -\frac{1}{\lambda}tAA_\lambda\gamma(\tau).
\end{aligned}$$

So, we get

$$\begin{aligned}
D_0 &= S_\lambda(t)x - S(t)x = a - b \\
&= \frac{1}{\lambda}\int_0^1 tAA_\lambda\gamma(\tau)x d\tau.
\end{aligned} \tag{2.30}$$

Substituting (2.29) into (2.30), we obtain (2.27).

To obtain (2.28), we denote

$$\begin{aligned}
J_1 &= \int_0^{1/2} tAA_\lambda\gamma(\tau)x d\tau. \quad \text{and} \\
J_2 &= \int_{1/2}^1 tAA_\lambda\gamma(\tau)x d\tau.
\end{aligned}$$

Then the convergence rate

$$||D_0|| \leq \frac{1}{\lambda}(||J_1|| + ||J_2||).$$

Next, we estimate $||J_1||$ and $||J_2||$:

$$\begin{aligned} ||J_1|| &\leq \int_0^{1/2} ||tAA_\lambda \gamma(\tau)x||d\tau \\ &\leq \int_0^{1/2} \frac{\delta_1\delta_2}{1-\tau}d\tau, \end{aligned}$$

where $\delta_1 = ||AS(\tau t)x||$ and $\delta_2 = ||(1-\tau)tA_\lambda S_\lambda((1-\tau)t)||$. Since $\{S(t) : t \geq 0\}$ is a semigroup of contractions, we have $\delta_1 \leq ||Ax||$ and from (2.24) we also have $\delta_2 \leq K$. Thus

$$||J_1|| \leq K||Ax|| \int_0^{1/2} \frac{1}{1-\tau}d\tau = \log_e(2)K||Ax||, \tag{2.31}$$

and

$$\begin{aligned} ||J_2|| &\leq \int_{1/2}^{1} ||tAA_\lambda \gamma(\tau)x||d\tau \\ &\leq \int_{1/2}^{1} \frac{\delta_3\delta_4}{\tau}d\tau, \end{aligned}$$

where $\delta_3 = ||A_\lambda S_\lambda\big((1-\tau)t\big)x||$ and $\delta_4 = ||\tau tAS(\tau t)||$. By Hille-Yosida theorem $||\lambda(\lambda I - A)^{-1}|| \leq 1$ for all $\lambda > 0$. It follows that $||A_\lambda x|| = ||\lambda A(\lambda I - A)^{-1}x|| = ||\lambda(\lambda I - A)^{-1}Ax|| \leq ||Ax||$ and $\delta_3 \leq ||Ax||$. From condition (2.22) we have $\delta_4 \leq K$. Hence

$$||J_2|| \leq K||Ax|| \int_{1/2}^{1} \frac{1}{\tau}d\tau = \log_e(2)K||Ax||. \tag{2.32}$$

Substituting (2.31) and (2.32) into

$$||D_0|| \leq \frac{1}{\lambda}(||J_1|| + ||J_2||),$$

we obtain (2.28). □

Note that using the same approach as above we can obtain the inverse expansion, i.e., expansion of the semigroup $S(t)$ in terms of the Yosida approximations $S_\lambda(t)$, and also the optimal convergence rate.

Consider the integro-differential identity

$$S(t) = S_\lambda(t) + \frac{b_1}{\lambda} + \mathbb{L}_1, \tag{2.33}$$

where the coefficient b_1 is bounded with respect to λ.

First a bound for the optimal convergence rate $||S(t)x - S_\lambda(t)x||$ is obtained. The expansion (2.33) will be considered in the next subsection.

Theorem 2.6 Let the semigroups $\{S(t) : t \geq 0\}$ and $\{S_\lambda(t) : t \geq 0\}$ satisfy the conditions (2.22) and (2.24). Then the convergence rate in (2.21) satisfies

$$||S_\lambda(t) - S(t)|| \leq \frac{4K^2}{\lambda t} \tag{2.34}$$

for all $t > 0$ and $\lambda > 0$. Moreover, for all $x \in X$, we have the following inequality

$$||S_\lambda(t)x - S(t)x|| \leq \frac{K||Ax||}{\lambda} \tag{2.35}$$

for all $t \geq 0$ and $\lambda > 0$.

Proof Proceeding as in Theorem 2.5, we take

$$\begin{aligned}\gamma(\tau) := S_{\lambda/\tau}(t) &= \exp\left\{ tA\frac{\lambda}{\tau}\left(\frac{\lambda}{\tau}I - A\right)^{-1} \right\} \\ &= \exp\{tA\lambda(\lambda I - \tau A)^{-1}\}\end{aligned} \tag{2.36}$$

Then $\gamma(1) = S_\lambda(t)$ and $\gamma(0)x = \lim_{\tau\downarrow 0} S_{\lambda/\tau}(t)x = S(t)x$ for all $x \in X$. Differentiating, we get

$$\begin{aligned}\gamma'(\tau) &= tA^2\lambda(\lambda I - \tau A)^{-2} S_{\lambda/\tau}(t) \\ &= \frac{1}{\lambda} tA^2_{\lambda/\tau}(t) S_{\lambda/\tau}(t).\end{aligned}$$

So, we obtain

$$\begin{aligned}D_0 x := S_\lambda(t)x - S(t)x &= \int_0^1 \gamma'(\tau) x d\tau \\ &= \frac{1}{\lambda t}\int_0^1 (tA_{\lambda/\tau})^2 S_{\lambda/\tau}(t) x d\tau.\end{aligned} \tag{2.37}$$

From (2.25) we have

$$||(tA_{\lambda/\tau})^2 S_{\lambda/\tau}(t)x|| \leq 4K^2||x|| \quad \forall \tau \in (0,1).$$

Then

$$||D_0 x|| \leq \frac{4K^2||x||}{\lambda t} \quad \forall x \in X, \tag{2.38}$$

and hence

$$||D_0|| \le \frac{4K^2}{\lambda t}.$$

Further, from (2.24) we have

$$||tA_{\lambda/\tau}S_{\lambda/\tau}(t)|| \le K \quad \forall \tau \in (0,1).$$

From the definition of A_λ we obtain

$$||A_{\lambda/\tau}x|| \le \left|\left|\frac{\lambda}{\tau}\left(\frac{\lambda}{\tau}I - A\right)^{-1}\right|\right| ||Ax||.$$

By Hille-Yosida theorem, we have $||\lambda(\lambda I - A)^{-1}|| \le 1$ for $\lambda > 0 >$, so that

$$||A_{\lambda/\tau}x|| \le ||Ax|| \quad \text{for} \tau \in (0,1).$$

Hence

$$||D_0x|| \le \frac{1}{\lambda}\int_0^1 ||tA_{\lambda/\tau}S_{\lambda/\tau}(t)|| ||A_{\lambda/\tau}x||d\tau \le \frac{K||Ax||}{\lambda}, \quad \forall x \in X. \quad \square$$

2.2.2 *Asymptotic Expansions for Yosida Approximations*

In this subsection we consider the expansions (2.26) and (2.33).

Let us introduce some notations:

$$\begin{aligned} d_{m,1,1} &= 1, \quad m = 1,2,\ldots, \\ d_{m,m,j} &= \frac{1}{m!}, \quad m = 1,2,\ldots, j = 1,2,\ldots,m, \\ d_{m,k,j} &= \sum_{i=1}^{j} d_{m-1,k,i}, \quad m = 2,3,\ldots, k = 1,2,\ldots,m-1, \\ & j = 1,2,\ldots,k. \end{aligned} \tag{2.39}$$

Theorem 2.7 Let $\{S(t) : t \ge 0\}$ be a differentiable semigroup. Then the coefficients a_m in (2.26) are given by

$$a_m = \sum_{k=1}^{m} d_{m,k,k} t^k A^{m+k} S(t)x, \tag{2.40}$$

and the remainder terms D_m are

$$D_m = D_{m,1} + D_{m,2}, \tag{2.41}$$

where

$$D_{m,1} = \frac{1}{\lambda^{m+1}} \sum_{k=1}^{m} \sum_{j=1}^{k} d_{m,k,j} t^k A^{m+j} A_\lambda^{k+1-j} S(t)x,$$

and

$$D_{m,2} = \frac{1}{\lambda^{m+1}} \int_0^1 \frac{\tau^m}{m!} (tAA_\lambda)^{m+1} S_\lambda((1-\tau)t) S(\tau t) x d\tau,$$

where coefficients $d_{m,k,j}$ are given in (2.39).

Proof From (2.27) we have

$$S_\lambda(t)x = S(t)x + D_0,$$

where

$$D_0 = \frac{1}{\lambda} \int_0^1 tAA_\lambda \gamma(\tau) x d\tau.$$

Integrating D_0 by parts, we have

$$D_0 = \frac{1}{\lambda} tAA_\lambda S(t)x + \frac{1}{\lambda^2} \int_0^1 \tau (tAA_\lambda)^2 \gamma(\tau) x d\tau. \tag{2.42}$$

It is easy to prove the identity $A_\lambda = A + AA_\lambda/\lambda$. Substituting this into the first term on the RHS of (2.42), we obtain

$$\begin{aligned} D_0 &= \frac{tA^2}{\lambda} S(t)x + \frac{tA^2 A_\lambda}{\lambda^2} S(t)x \\ &\quad + \frac{1}{\lambda^2} \int_0^1 \tau (tAA_\lambda)^2 \gamma(\tau) x d\tau = \frac{a_1}{\lambda} + D_1. \end{aligned}$$

This proves (2.40) and (2.41) for $m = 1$. Using induction on m, we obtain the general result. □

For instance, the first three coefficients of the expansion are

$$\begin{aligned} a_1 &= tA^2 S(t)x, \\ a_2 &= tA^3 S(t)x + \frac{t^2 A^4}{2} S(t)x, \end{aligned}$$

$$a_3 = tA^4S(t)x + t^2A^5S(t)x + \frac{t^3A^6}{6}S(t)x.$$

Theorem 2.8 Let the semigroups $\{S(t) : t \geq 0\}$ and $\{S_\lambda(t) : t \geq 0\}$ satisfy the conditions (2.22) and (2.24). Then the remainder terms D_m in (2.26) satisfy

$$||D_m|| \leq \frac{C_m(1+K^{m+1})||A^{m+1}x||}{\lambda^{m+1}}, \quad m = 1,2,\ldots$$

for $\lambda > 0$ and some positive constant C_m depending only on m.

Proof From the definition of Yosida approximations (2.13) and (2.22), it follows that

$$||D_{m,1}|| \leq C_{m,1}K^m||A^{m+1}x||/\lambda^{m+1},$$

where $C_{m,1}$ is some positive constant depending only on m, The bound

$$||D_{m,2}|| \leq C_{m,2}K^{m+1}||A^{m+1}x||$$

can be obtained in a similar manner as the bound for $||D_0||$ in the proof of Theorem 2.5. $\square$

We now consider the asymptotic expansion (2.33).

Theorem 2.9 Let the semigroup $\{S(t), t \geq 0\}$ and $\{S_\lambda(t), t \geq 0\}$ satisfy the conditions (2.22) and (2.24), Then the coefficient b_1 in (2.33) is given by

$$b_1 = -tA_\lambda^2 S_\lambda(t), \tag{2.43}$$

and the remainder term $\mathbb{L}_1$ satisfies

$$||\mathbb{L}_1|| \leq \frac{CK^3(1+K)}{\lambda^2 t^2}, \quad t > 0, K > 0,$$

where C is a positive constant independent of λ and t.

Proof From the proof of Theorem 2.6, we have

$$S(t)x = S_\lambda(t)x - D_0x,$$

where $D_0x = \int_0^1 \gamma'(\tau)xd\tau$. Integrating by parts, we obtain

$$D_0x = \frac{b_1}{\lambda}x + \mathbb{L}_1x,$$

where $b_1 = tA_\lambda^2 S_\lambda(t)$, and

$$\begin{aligned}\mathbb{L}_1 x &= -\int_0^1 \gamma''(\tau) x d\tau \\ &= -\frac{1}{t^2\lambda^2}\int_0^1 \tau S_{\lambda/\tau}(t)\big((tA_{\lambda/\tau})^4 + 2(tA_{\lambda/\tau})^3\big) x d\tau.\end{aligned}$$

From (2.25) we have

$$||(tA_{\lambda/\tau})^4 S_{\lambda/\tau}(t)x|| \leq 4^4 K^4 ||x||,$$

and

$$||(tA_{\lambda/\tau})^3 S_{\lambda/\tau}(t)x|| \leq 3^3 K^3 ||x||$$

for all $\tau \in (0,1)$. Then

$$||\mathbb{L}_1 x|| \leq \frac{CK^3(1+K)||x||}{\lambda^2 t^2} \quad \forall x \in X,$$

and hence

$$||\mathbb{L}_1|| \leq \frac{CK^3(1+K)}{\lambda^2 t^2}. \quad \square$$

2.3 Almost Strong Evolution Operators

This section is needed to study time-varying stochastic evolution equations.

Definition 2.7 (Mild evolution operator) Let $\Delta(T) = \{(t,s) : 0 \leq s \leq t \leq T\}$, then $U(t,s) : \Delta(T) \to L(X)$ is a mild evolution operator if

(a) $U(t,t) = I, \quad t \in [0,T],$
(b) $U(t,r)U(r,s) = U(t,s), \quad 0 \leq s \leq r \leq t \leq T,$
(c) $U(\cdot,s)$ is strongly continuous on $[s,T]$ and $U(t,\cdot)$ is strongly continuous on $[0,T]$.

A consequence of (c) is that $\operatorname{ess\,sup}_{\Delta(T)} ||U(t,s)|| < \infty$. Clearly, if $\{S(t) : t \geq 0\}$ is a strongly continuous semigroup, then $S(t-s)$ is a mild evolution operator.

A mild evolution operator, if in addition, satisfies

(d) For every $T > 0$ there is a constant C_T such that

$$||U(t,s)||_{L(X)} \leq C_T, \quad 0 \leq s \leq t \leq T,$$

then $U(t,s)$ is an evolution operator.

Definition 2.8 (Quasi-evolution operators) A quasi-evolution operator $U(t,s)$ is a mild evolution operator such that there exists a nonzero $x_0 \in X$ and a closed linear operator $A(s)$ on X for almost all $s \in [0,T]$ satisfying

$$U(t,s)x_0 - x_0 = \int_s^t U(t,r)A(r)x_0 dr.$$

We denote the set of $x_0 \in X$ for which (a) is valid as $D(A(t))$ and we call $A(t)$ the quasi-generator of $U(t,s)$.

Those quasi evolution operators which are also differentiable in the first variable are also important in applications and so we define

Definition 2.9 (a) (Almost strong evolution operator) An almost strong evolution operator is a mild evolution operator on X for which there exists an associated closed linear operator $A(t)$ on X for almost all $t \in [0,T]$ such that

(i) $U(t,s) : D(A(s)) \to D(A(t))$ for all $t > s \in [0,T]$,
(ii) $\int_s^t A(r)U(r,s)x_0 dr = (U(t,s) - I)x_0$ for $x_0 \in D(A(s))$.

Note that (i) implies

$$\frac{\partial}{\partial t}U(t,s)x_0 = A(t)U(t,s)x_0 \quad a.e. \quad \text{for} \quad x_0 \in D(A(t)).$$

(b) (Strong evolution operator) A strong evolution operator is an evolution operator for which there exists a closed, linear, densely defined operator $A(t), t \geq 0$, with the domain $D(A(t))$, such that

(a) $U(t,s) : D(A(s)) \to D(A(t))$ for $t > s$,
(b) $\frac{\partial}{\partial t}U(t,s)h = A(t)U(t,s)h$ for $h \in D(A(s))$, $t > s$.

2.4 Basics from Analysis and Probability in Banach Spaces

Let $(X, ||\cdot||_X)$ be a real Banach space and $(X^*, ||\cdot||_{X^*})$ be its dual space. We mean by ${}_{X^*}\langle\cdot,\cdot\rangle_X$ the duality pairing between X and X^* and is defined by

$${}_{X^*}\langle x^*, x\rangle_X := x^*(x) \quad \text{for} \quad x^* \in X^*, x \in X.$$

If X is a Hilbert space, then $\langle\cdot,\cdot\rangle_X$ denotes the inner product in X, and 2^X stands for the family of all subsets of X.

In the following, let $\{x_n\}$ be any sequence in X and $\mathbb{T} : X \to Y$ is any operator, where Y is another real Banach space. Then

(a) $\mathbb{T}$ is said to be continuous at x_0 if

$$x_n \to x_0 \Rightarrow \mathbb{T}x_n \to \mathbb{T}x_0.$$

(b) $\mathbb{T}$ is said to be demicontinuous at x_0 if

$$x_n \to x_0 \Rightarrow \mathbb{T}x_n \to \mathbb{T}x_0 \quad \text{weakly.}$$

(c) $\mathbb{T}$ is said to be hemicontinuous at x_0 if for any sequence $\{x_n\}$ converging to x_0 along a line, the sequence $\{\mathbb{T}x_n\}$ converges weakly to $\mathbb{T}x_0$. That is,

$$\mathbb{T}x_n = \mathbb{T}(x_0 + t_n x) \to \mathbb{T}x_0 \quad \text{weakly as} \quad t_n \to 0 \quad \text{for all} \quad x \in X.$$

Note that demicontinuity implies hemicontinuity. Conversely, if a hemicontinuous operator is monotone, then it is demicontinuous.

An operator $\mathbb{T} : X \to Y$ is said to be Fréchet differentiable at x if there exists a continuous linear operator $\mathsf{A} : X \to Y$ such that

$$\mathbb{T}(x+h) - \mathbb{T}(x) = \mathsf{A}h + w(x,h)$$

where

$$\lim_{||h|| \to 0} ||w(x,h)||/||h|| = 0.$$

A is called the Fréchet derivative of $\mathbb{T}$ at x and is denoted by $\mathbb{T}'(x)$.

A Banach space X is said to be separable if it has a countable subset that is everywhere dense.

The following lemma will be crucial in the subsequent analysis.

Bellman-Gronwall's Lemma

(a) If $g \geq 0$ and h are integrable on $[t_0, T]$ $(0 < T < \infty)$ and if

$$g(t) \leq h(t) + \ell \int_{t_0}^{t} g(s)ds, \quad t_0 \leq t \leq T,$$

for $\ell > 0$, then

$$g(t) \leq h(t) + \ell \int_{t_0}^{t} e^{\ell(t-s)} h(s)ds, \quad t_0 \leq t \leq T.$$

(b) Let $g(t)$ and $h(t)$ be nonnegative functions and let k be a positive constant such that for $t \geq s$,

$$g(t) \leq k + \int_{s}^{t} h(\tau)g(\tau)d\tau.$$

Then for $t \geq s$,

$$g(t) \leq k\exp\left\{\int_s^t h(\tau)d\tau\right\}.$$

The following Cauchy's formula will be used in the sequel.

Cauchy's Formula

Let $g : [t_0, T] \to R$ be integrable. Then, for $t \in [t_0, T]$,

$$\int_{t_0}^{t}\int_{t_0}^{t_{n-1}}\cdots\int_{t_0}^{t_1} g(s)dsdt_1\ldots dt_{n-1} = \int_{t_0}^{t} g(s)\frac{(t-s)^{n-1}}{(n-1)!}ds, \quad n = 1,2,3,\cdots.$$

Let Ω be a nonempty abstract set, whose elements ω are termed elementary events. $\mathcal{F}$ is a σ-algebra of subsets of Ω; that is $\mathcal{F}$ is a nonempty class of subsets of Ω satisfying the following conditions: (i) $\Omega \in \mathcal{F}$, (ii) if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$, and (iii) if $A_n \in \mathcal{F}$, $n = 1, 2, \ldots$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$. The elements of $\mathcal{F}$ are called events. P is a probability measure on the measurable space $(\Omega, \mathcal{F})$; that is, P is a set function, with domain $\mathcal{F}$, which is nonnegative, countably additive, and such that $P(A) \in [0, 1]$ for all $A \in \mathcal{F}$, with $P(\Omega) = 1$. We call $(\Omega, \mathcal{F}, P)$ a probability measure space. Let us assume throughout this book that P is a complete probability measure; that is, P is such that the conditions $A \in \mathcal{F}$, $P(A) = 0$, and $A_0 \subseteq A$ imply $P(A_0) = 0$.

Let $(X, \mathcal{B}(X))$ be a measurable space, where $\mathcal{B}(X)$ is the σ-algebra of all Borel subsets of X. A sequence $\{x_n\}$ of elements in X converges strongly, or converges in the strong topology to an element x if $\lim_{n\to\infty} ||x_n - x||_X = 0$, x called the strong limit of $\{x_n\}$. A sequence $\{x_n\}$ of elements in X converges weakly, or converges in the weak topology, to an element x if (i) the norms $||x_n||$ are uniformly bounded, that is, $||x_n||_X \leq M$, and (ii) $\lim_{n\to\infty} x^*(x_n) = x^*(x)$ for every $x^* \in X^*$. If a sequence $\{x_n\}$ of elements in a Banach space X converges strongly to an element $x \in X$, then $\{x_n\}$ also converges weakly to x.

Definition 2.10 A mapping $x : \Omega \to X$ is said to be a random variable with values in X if the inverse image under the mapping x of every Borel set $B \in \mathcal{F}$; that is, $x^{-1}(B) \in \mathcal{F}$ for all $B \in \mathcal{B}$.

Definition 2.11 A mapping $x : \Omega \to X$ is said to be a finitely valued random variable if it is constant on each of a finite number of disjoint sets $A_i \in \mathcal{F}$ and equal to θ (the null element of X on $\Omega \setminus (\cup_{i=1}^{n} A_i)$, and a simple random variable if it is finitely valued and $P\{\omega : ||x(\omega)||_X > 0\} < \infty$. A mapping $x : \Omega \to X$ is said to be a countably valued random variable if it assumes at most a countable set of values in X, assuming each value different from θ on a set in $\mathcal{F}$.

Definition 2.12 A mapping $x : \Omega \to X$ is said to be a strong (or Bochner) random variable if there exists a sequence $\{x_n\}$ of countably valued random variables which converges to x, P-$a.s.$, that is, there exists a set $A_0 \in \mathcal{F}$, with $P(A_0) = 0$ such that

$$\lim_{n\to\infty} ||x_n(\omega) - x(\omega)||_X = 0 \quad \text{for every} \quad \omega \in \Omega \setminus A_0.$$

Since $P(\Omega) = 1$, we can replace countably valued in Definition 2.11 by simple.

Definition 2.13 A mapping $x : \Omega \to X$ is said to be a weak (or Pettis) random variable if the functions $x^*(x)$ are real-valued random variables for each $x^* \in X^*$.

The concepts of weak and strong random variables coincide in separable Banach spaces.

Definition 2.14 x is said to be a Bochner integrable if and only if there exists a sequence of simple random variables $\{x_n\}$ converging P-$a.s.$ to x such that

$$\lim_{n\to\infty} \int_\Omega ||x_n - x|| dP = 0.$$

By definition

$$\int_A x dp = \lim_{n\to\infty} \int_A x_n dP$$

for every $A \in \mathcal{F}$ and $A = \Omega$.

It is clear from the above definition that every Bochner integrable random variable is a strong random variable.

Let x be a strong random variable. The expectation of x, denoted by $E(x)$, or simply Ex, is defined as the Bochner integral of x over Ω; that is,

$$E(x) = \int_\Omega x dP.$$

For some properties of expectation we refer to Hille and Phillips [1, Section 3.7].

The variance of a Banach space-valued random variable is defined as

$$\begin{aligned} V(x) &= E||x - E(x)||_X^2 \\ &= \int_\Omega ||x - E(x)||_X^2 dP. \end{aligned}$$

Let $x : \Omega \to X$ be a square-integrable random variable, i.e., $x \in L^2(\Omega, \mathcal{F}, P; X)$, where X is a Hilbert space. The covariance operator of x is defined by

$$Cov(x) = E(x - E(x)) \otimes (x - E(x))$$

and $\otimes$ is the tensor product. $g \otimes h \in L(X)$ for any $g, h \in X$ is defined by

$$(g \otimes h)k = g\langle h, k\rangle, \quad k \in X.$$

$Cov(x)$ is a self-adjoint nonnegative trace class (or nuclear) operator and

$$\begin{aligned}\operatorname{tr} Cov(x) &= E\|x - E(x)\|_X^2 \\ &= E\|x\|_X^2 - \|E(x)\|_X^2,\end{aligned}$$

where tr denotes the trace. If $P_1 \in L(X)$, then

$$\begin{aligned}\operatorname{tr} P_1\, Cov(x) &= \operatorname{tr} Cov(P_1 x, x) \\ &= E\langle P_1(x - E(x)), x - E(x)\rangle,\end{aligned}$$

where $Cov(x,y) = E(x - E(x)) \otimes (y - E(y))$ is the joint covariance of x and y. A random variable $x \in L^2(\Omega, \mathcal{F}, P; X)$ is Gaussian if $\langle x, e_i\rangle$ is a real Gaussian random variable for all i, where $\{e_i\}$, $i = 1, 2, \ldots$, is a complete orthonormal basis for X.

The following result yields the definition of the conditional expectation.

Proposition 2.3 Let X be a separable Banach space and let x be a Bochner integrable X-valued random variable defined on $(\Omega, \mathcal{F}, P)$. Suppose that $\mathcal{A}$ is a σ-algebra contained in $\mathcal{F}$. There exists a unique, up to a set of probability zero, integrable X-valued random variable z, measurable with respect to $\mathcal{A}$ such that

$$\int_A x dP = \int_A z dP, \quad \forall A \in \mathcal{A}.$$

The random variable z will be denoted as $E(x|\mathcal{A})$ and called the conditional expectation of x given $\mathcal{A}$.

Proof See Da Prato and Zabczyk [1, Proposition 1.10]. □

We now give the definition of independence. Let $\{\mathcal{F}_i\}_{i\in I}$ be a family of sub-σ-algebras of $\mathcal{F}$. These σ-algebras are said to be independent if, for every finite subset $J \subset I$ and every family $\{A_i\}_{i\in J}$ such that $A_i \in \mathcal{F}_i$, $i \in J$,

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

Random variables $\{x_i\}_{i\in I}$ are independent if the σ-algebras $\{\sigma(x_i)\}_{i\in I}$ are independent, where $\sigma(x_i)$ is the smallest σ-algebra generated by x_i, $i \in I$.

Let I be a subinterval of $[0, \infty)$. Let X be a separable Banach space and $\mathcal{B}(X)$ its Borel σ-algebra. A stochastic process in X is a family of random variables $\{x(t), t \in I\}$ in X. Functions $x(\cdot, \omega)$ are called the trajectories or sample paths of $x(t)$. A stochastic process $\{x(t), t \in I\}$ is a modification or a version of $y(t)$ if for each $t \in I$, $x(t) = y(t)$ P-$a.s.$ If two processes are a modification of each other, we regard them as equivalent. The process $x(t)$ is measurable if x is measurable relative to $\mathcal{B}(I) \times \mathcal{F}$, where $\mathcal{B}(I)$ is the Borel σ-algebra of subsets of I. Let $\mathcal{F}_t, t \in I$, be a family of increasing sub σ-algebras of $\mathcal{F}$. A stochastic process $\{x(t), t \in I\}$ is adapted to $\mathcal{F}_t$ if $x(t)$ is $\mathcal{F}_t$-measurable for all $t \in I$. $\{x(t), t \in I\}$ is called progressively measurable with respect to $\mathcal{F}_t$, if, for every $t \in I$, the map $(s, \omega) \mapsto x(s, \omega)$ from $[0, t] \times \Omega$

into $(X, \mathcal{B}(X))$ is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable. A progressively measurable process is adapted. Conversely, any adapted process with right or left-continuous paths is progressively measurable. An X-valued right-continuous process $\{x(t), t \in I\}$ with paths having left limits is called c*à*dl*à*g. A nondecreasing process $\{N(t), t \geq 0\}$ is a real-valued process that is $\mathcal{F}_t$-adapted and has positive, nondecreasing and finite paths, P-$a.s.$

A stochastic process $\{x(t), t \in I\}$ is called a martingale with respect to $\{\mathcal{F}_t\}$ if it is adapted to $\mathcal{F}_t$ with properties:

(a) $E||x(t)|| < \infty \quad$ for all $t \in I$,
(b) $E(x(t)|\mathcal{F}_s) = x(s)$ P-$a.s.$

for all $s < t$, $s, t \in I$, where $E(\cdot|\mathcal{F}_s)$ denotes the conditional expectation with respect to $\mathcal{F}_s$.

In what follows, we state some fundamental results.

Proposition 2.4 If $x(t)$ is a martingale in X relative to $\mathcal{F}_t$, then $||x(t)||$ is a real submartingale, i.e.,

$$E(||x(t)|| \, |\mathcal{F}_s) \geq ||x(s)|| \quad P\text{-}a.s.$$

for all $s < t$, $s, t \in I$.

Proof See Ichikawa [3]. □

Theorem 2.10 The following statements hold:

(i) If $\{x(t), t \in I\}$ is an X-valued martingale, I a countable set and $p \geq 1$, then for arbitrary $\lambda > 0$,

$$P(\sup_{t \in I} ||x(t)|| \geq \lambda) \leq \frac{1}{\lambda^p} \sup_{t \in I} E||x(t)||^p.$$

(ii) If, in addition, $p > 1$, then,

$$E(\sup_{t \in I} ||x(t)||^p) \leq \left(\frac{p}{p-1}\right)^p \sup_{t \in I} E||x(t)||^p.$$

(iii) The above estimates remain true if the set I is uncountable and the martingale $x(t)$ is continuous.

Proof See Da Prato and Zabczyk [1, Theorem 3.8]. □

Let us fix a number $T > 0$ and denote by $M_T^2(X)$ the space of all X-valued continuous, square integrable martingales x.

Proposition 2.5 The space $M_T^2(X)$ equipped with the norm

$$||x||_{M_T^2(X)} = \left(E \sup_{t \in [0,T]} ||x(t)||^2 \right)^{1/2}$$

is a Banach space.

Proof See Da Prato and Zabczyk [1, Proposition 3.9]. □

If $x \in M_T^2(R)$ then there exists a unique, increasing, and adapted process $\langle\langle x(\cdot)\rangle\rangle$, starting from 0, such that the process $x^2(t) - \langle\langle x(t)\rangle\rangle$, $t \in [0,T]$, is a continuous martingale. The process $\langle\langle x(\cdot)\rangle\rangle$ is called the quadratic variation of x.

Proposition 2.6 (Lévy's Theorem) If $x \in M_T^2(R)$, $x(0) = 0$ and $\langle\langle x(t)\rangle\rangle = t$, $t \in [0,T]$, then $x(\cdot)$ is a standard Wiener process adapted to $\mathcal{F}_t$ and with increments $x(s) - x(t)$, $s > t$ independent of $\mathcal{F}_t$, for every $t \in [0,T]$.

Proof See Da Prato and Zabczyk [1, Proposition 3.10]. See also Ikeda and Watanabe [1]. □

Let $(X, \langle\cdot,\cdot\rangle_X)$ and $(Y, \langle\cdot,\cdot\rangle_Y)$ be two real separable Hilbert spaces.

Definition 2.15 A probability measure P on $(Y, \mathcal{B}(Y))$ is called Gaussian if for all $v \in Y$ the bounded linear mapping

$$v' : Y \to R$$

defined by

$$u \mapsto \langle u, v\rangle_Y, \quad u \in Y,$$

has a Gaussian law, i.e., for all $v \in Y$, there exists $m := m(v) \in R$ and $\sigma := \sigma(v) \in [0,\infty)$ such that, if $\sigma(v) > 0$,

$$\begin{aligned}(P \circ (v')^{-1})(\mathbb{A}) &= P(v' \in \mathbb{A}) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{A}} e^{\frac{-(x-m)^2}{2\sigma^2}} dx \quad \forall \mathbb{A} \in \mathcal{B}(R),\end{aligned}$$

and if $\sigma(v) = 0$,

$$P \circ (v')^{-1} = \delta_{m(v)}.$$

Theorem 2.11 A measure P on $(Y, B(Y))$ is Gaussian if and only if

$$\phi(u) := \int_Y e^{i\langle u,v\rangle_Y} P(dv)$$

$$= e^{i\langle m,u\rangle_Y} - \frac{1}{2}\langle Qu,u\rangle_Y, \quad u \in Y,$$

where $m \in Y$ and $Q \in L(Y)$ is nonnegative, symmetric, and with finite trace.

In this case P will be denoted by $N(m,Q)$ where m is called mean and Q is called the covariance operator. The measure P is uniquely determined by m and Q. Furthermore, for all $h,g \in Y$,

$$\begin{aligned} \int \langle x,h\rangle_Y P(dx) &= \langle m,h\rangle_Y, \\ \int (\langle x,h\rangle_Y - \langle m,h\rangle_Y)(\langle x,g\rangle_Y - \langle m,g\rangle_Y) P(dx) &= \langle Qh,g\rangle_Y, \\ \int ||x-m||_Y^2 P(dx) &= \mathrm{tr} Q. \end{aligned}$$

Proof See Prévôt and Röckner [1, Theorem 2.1.2]. □

2.4.1 Wiener Processes

We next define the standard Q-Wiener process. We fix an element $Q \in L(Y)$, nonnegative, symmetric, and with finite trace and a positive real number T.

Definition 2.16 A Y-valued stochastic process $\{w(t), t \in [0,T]\}$, on a probability space $(\Omega, \mathcal{F}, P)$ is called a standard Q-Wiener process if

(i) $w(0) = 0$,
(ii) $w(t)$ has a continuous trajectories P-*a.s.*,
(iii) $w(t)$ has independent increments,
(iv) the increments have the Gaussian laws:

$$\begin{aligned} P \circ (w(t) - w(s))^{-1} &= \mathcal{L}(w(t) - w(s)) \\ &= N(0, (t-s)Q), \quad t \geq s \geq 0. \end{aligned}$$

Proposition 2.7 **(**Representation of a Q-Wiener process**)** Assume that $w(t)$ is a Q-Wiener process with $\mathrm{tr}\, Q < \infty$. Then the following statements hold:

(i) $w(t)$ is a Gaussian process on Y and

$$Ew(t) = 0, \quad Cov(w(t)) = tQ, \quad t \geq 0.$$

(ii) For arbitrary t, $w(t)$ has the expansion

$$w(t) = \sum_{j=1}^{\infty} \sqrt{\lambda_j}\beta_j(t)e_j, \tag{2.44}$$

where

$$\beta_j(t) = \frac{1}{\sqrt{\lambda_j}}\langle w(t), e_j\rangle, \quad j = 1, 2, \ldots,$$

are real-valued Brownian motions mutually independent on $(\Omega, \mathcal{F}, P)$ and the series in (2.44) is convergent in $L^2(\Omega, \mathcal{F}, P)$.

Proof See Da Prato and Zabczyk [1, Proposition 4.1]. □

Definition 2.17 (Normal filtration) A filtration $\mathcal{F}_t, t \in [0, T]$, on a probability space $(\Omega, \mathcal{F}, P)$ is called normal if

(i) $\mathcal{F}_0$ contains all elements $A \in \mathcal{F}$ with $P(A) = 0$ and
(ii) $\mathcal{F}_t = \mathcal{F}_{t+} = \cap_{s>t}\mathcal{F}_s \quad \text{for all} \quad t \in [0, T]$.

Definition 2.18 (Q-Wiener process with respect to a filtration) A Q-Wiener process $\{w(t), t \in [0, T]\}$, is called a Q-Wiener process with respect to a filtration $\mathcal{F}_t$, $t \in [0, T]$, if:

(i) $w(t)$ is $\mathcal{F}_t$-measurable $t \in [0, T]$, and
(ii) $w(t) - w(s)$ is independent of $\mathcal{F}_s$ for all $0 \leq s \leq t \leq T$.

In fact, it is possible to define a Wiener process when Q is not necessarily of finite trace. This leads to the concept of a cylindrical Wiener process. In this case the convergence of the series (2.44) is lost.

It is useful, at this moment, to introduce the subspace $Y_0 = Q^{1/2}(Y)$ of Y which, endowed with the inner product

$$\begin{aligned}\langle u, v\rangle_0 &= \sum_{k=1}^{\infty} \frac{1}{\lambda_k}\langle u, e_k\rangle\langle v, e_k\rangle \\ &= \langle Q^{-1/2}u, Q^{-1/2}v\rangle,\end{aligned}$$

is a Hilbert space. We will need a further Hilbert space $(Y_1, \langle\cdot,\cdot\rangle_1)$ and a Hilbert-Schmidt embedding

$$J : (Y_0, \langle\cdot,\cdot\rangle_0) \to (Y_1, \langle\cdot,\cdot\rangle_1).$$

Remark 2.1 $(Y_1, \langle\cdot,\cdot\rangle_1)$ and J as above always exist, e.g., choose $Y_1 := U$ and $a_k \in (0,\infty)$, $k \in N$, such that $\sum_{k=1}^{\infty} a_k^2 < \infty$. Define $J : Y_0 \to Y$ by

$$J(u) := \sum_{k=1}^{\infty} a_k \langle u, e_k\rangle_0 e_k, \quad u \in Y_0.$$

Then J is one-to-one and Hilbert-Schmidt.

The process given by the following proposition is called a cylindrical Wiener process in Y.

Proposition 2.8 Let $\{e_k\}$ be an orthonormal basis of Y_0 and $\beta_k, k \in \mathbb{N}$, a family of independent real-valued Brownian motions. Define $Q := JJ^*$. Then $Q \in L(Y_1), Q_1$ is nonnegative definite and symmetric with finite trace and the series

$$w(t) = \sum_{k=1}^{\infty} \beta_k(t) J e_k, \quad t \in [0,T],$$

converges in $M_T^2(Y_1)$ and defines a Q_1-Wiener process on Y_1. Moreover, we have that $Q_1^{1/2}(Y_1) = J(Y_0)$ and for all $u_0 \in Y_0$,

$$||u_0||_0 = ||Q_1^{-1/2} J u_0||_1 = ||J(u_0)||_{Q^{1/2}Y_1},$$

i.e., $J : Y_0 \to Q_1^{1/2} Y_1$ is an isometry.

Proof See Prévôt and Röckner [1, Proposition 2.5.2]. □

2.4.2 Poisson Random Measures and Poisson Point Processes

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and $(\mathbf{S}, \mathcal{S})$ a measurable space. Let Z_+ be the set of nonnegative integers. Suppose that $\mathbb{M}$ is the space of $Z_+ \cup \{+\infty\}$-valued measures on $(\mathbf{S}, \mathcal{S})$ and

$$\mathcal{B}_{\mathbb{M}} := \sigma(\mathbb{M} \ni \mu \mapsto \mu(B) | B \in \mathcal{S}).$$

Definition 2.19 (Poisson random measure) A random variable $\mu : (\Omega, \mathcal{F}) \to (\mathbb{M}, \mathcal{B}(\mathbb{M}))$ is called Poisson random measure if the following conditions hold:

(i) For all $B \in \mathcal{S}, \mu(B) : \Omega \to Z_+ \cup \{+\infty\}$ is Poisson distributed with parameter $E[\mu(B)]$, i.e.,

$$P(\mu(B) = n) = e^{-E[\mu(B)]} \frac{(E[\mu(B)])^n}{n!}, \quad n = 0, 1, 2, 3, \cdots.$$

If $E[\mu(B)] = \infty$, then $\mu(B) = \infty$ P-$a.s.$

(ii) If $B_1, \dots, B_m \in \mathcal{S}$ are pairwise disjoint, then $\mu(B_1), \dots, \mu(B_m)$ are independent.

Let $(Z, \mathcal{Z})$ be another measurable space and set

$$(\mathbf{S}, \mathcal{S}) = ([0,\infty) \times Z, \mathcal{B}([0,\infty)) \otimes \mathcal{Z}).$$

Definition 2.20 A point function p on Z is a mapping $p : D_p \subset (0,\infty) \to Z$ where the domain D_p of p is countable.

Remark 2.2 The point function p induces a measure $\mu(dt,dy)$ on $([0,\infty) \times Z, \mathcal{B}([0,\infty)) \otimes \mathcal{Z})$ in the following way:

Define $\tilde{p} : D_p \to (0,\infty) \times Z, t \mapsto (t, p(t))$ and denote by c the counting measure on $(D_p, \mathcal{P}(D_p))$, i.e., $c(A) := \#A$ for all $A \in \mathcal{P}(D_p)$. Here, $\mathcal{P}(D_p)$ denotes the power set of D_p. For $(A \times B) \in \mathcal{B}([0,\infty)) \otimes \mathcal{Z}$, define the measure

$$\mu(A \times B) := c(\tilde{P}^{-1}(A \times B)).$$

Then, in particular, for all $A \in \mathcal{B}([0,\infty))$ and $B \in \mathcal{Z}$ we obtain

$$\mu(A \times B) = \#\{t \in D_p | t \in A, p(t) \in B\}.$$

For $t \geq 0$, $B \in \mathcal{Z}$ we write

$$\mu(t,B) := \mu((0,t] \times B).$$

Let $\mathcal{P}_Z$ be the space of all point functions on Z and

$$\mathcal{B}_{\mathcal{P}_Z} := \sigma(\mathcal{P}_Z \ni p \mapsto \mu(t,B) | t > 0, B \in \mathcal{Z}).$$

Definition 2.21

(i) A point process on Z and $(\Omega, \mathcal{F}, P)$ is a random variable $p : (\Omega, \mathcal{F}) \to (\mathcal{P}_Z, \mathcal{B}_{\mathcal{P}_Z})$.
(ii) A point process p is called stationary if for every $t > 0$, p and $\theta_t p$ have the same probability law. Here, θ_t is given by $\theta_t : (0,\infty) \to (0,\infty), s \mapsto s+t$.
(iii) A point process p is called σ-finite if there exists $\{B_n\}_{n\in\mathbb{N}} \in \mathcal{Z}$ such that $B_n \uparrow Z$ as $n \to \infty$ and $E[\mu(t,B_n)] < \infty$ for all $t > 0$ and $n \in \mathbb{N}$.
(iv) A point process p on Z is called Poisson point process if there exists a Poisson random measure $\tilde{\mu}$ on $((0,\infty) \otimes Z, \mathcal{B}((0,\infty) \otimes \mathcal{Z})$ such that there exists a P-zero set $N \in \mathcal{F}$ such that for all $\omega \in N^c$ and all $A \times B \in \mathcal{B}((0,\infty)) \otimes \mathcal{Z}$,

$$\mu(\omega)(A \times B) = \tilde{\mu}(\omega)(A \times B).$$

Proposition 2.9 Let p be a σ-finite Poisson point process on Z and $(\Omega, \mathcal{F}, P)$. Then, p is stationary if and only if there exists a σ-finite measure m on $(Z, \mathcal{Z})$ such that

$$E[\mu(dt, dy)] = dt \otimes m(dz)$$

where dt denotes the Lebesgue-measure on $(0, \infty)$. In that case, the measure m is uniquely determined.

Proof See Knoche [1, Proposition 2.10]. $\square$

The measure m in Proposition 2.9 is called the characteristic measure of μ.

Definition 2.22 Let $\mathcal{F}_t$, $t \geq 0$, be a filtration on $(\Omega, \mathcal{F}, P)$ and p a point process on Z and $(\Omega, \mathcal{F}, P)$.

(i) The process p is called $\mathcal{F}_t$-adapted if for every $t \geq 0$ and $B \in \mathcal{Z}$, $\mu(t, B)$ is $\mathcal{F}_t$-measurable.
(ii) The process p is called an $\mathcal{F}_t$-Poisson point process if it is an $\mathcal{F}_t$-adapted, σ-finite Poisson point process such that $\{\mu((t, t+h] \times B) | h > 0, B \in \mathcal{Z}\}$ is independent of $\mathcal{F}_t$ for all $t \geq 0$.

We define the set $\Gamma_\mu := \{B \in \mathcal{Z} | E[\mu(t, B)] < \infty, \forall t > 0\}$.

Definition 2.23 Let $\mathcal{F}_t$ be a right-continuous filtration on $(\Omega, \mathcal{F}, P)$ and p a point process on Z. The process p is said to be of class (QL) with respect to $\mathcal{F}_t$ if it is $\mathcal{F}_t$-adapted and σ-finite and for all $B \in \mathcal{Z}$ there exists a process $\hat{\mu}(t, B), t \geq 0$, such that

(i) for $B \in \Gamma_\mu, \hat{\mu}(t, B), t \geq 0$, is a continuous $\mathcal{F}_t$-adapted increasing process with $\hat{\mu}(0, B) = 0$ P-$a.s.$,
(ii) for all $t \geq 0$ and P-$a.s.$ $\omega \in \Omega, \hat{\mu}(\omega)(t, \cdot)$ is a σ-finite measure on $(Z, \mathcal{Z})$.
(iii) for $B \in \Gamma_\mu$,

$$\bar{\mu}(t, B) := \mu(t, B) - \hat{\mu}(t, B), \quad t \geq 0,$$

is an $\mathcal{F}_t$-martingale.

Here $\hat{\mu}$ is called compensator of μ and $\bar{\mu}$ is called compensated Poisson random measure of μ.

Proposition 2.10 Let $\mathcal{F}_t, t \geq 0$, be a right-continuous filtration on $(\Omega, \mathcal{F}, P)$ and let m be a σ-finite measure on $(Z, \mathcal{Z})$ and p a stationary $\mathcal{F}_t$-Poisson point process on Z with characteristic measure m. Then p is quasi-leftcontinuous with respect to $\mathcal{F}_t$ with compensator $\hat{\mu}(t, B) = t \cdot m(B), t \geq 0, B \in \mathcal{Z}$.

Proof See Knoche [1, Corollary 2.18]. $\square$

2.4.3 Lévy Processes

Definition 2.24 Let $\{\mathbb{X}(t), t \geq 0\}$ be a stochastic process with values in Y.

(i) The process $\mathbb{X}(t)$ is said to be stochastically continuous if for every $t \geq 0$ and $\varepsilon > 0$

$$\lim_{s \to t} P(||\mathbb{X}(s) - \mathbb{X}(t)||_Y > \varepsilon) = 0.$$

(ii) The process $\mathbb{X}(t)$ has independent increments if $\mathbb{X}(t) - \mathbb{X}(s)$ is independent of $\mathcal{F}_s$, for all $0 \leq s < t < \infty$.
(iii) If the distribution of $\mathbb{X}(t) - \mathbb{X}(s)$ depends only on the difference $t - s$ we say that $\mathbb{X}(t)$ has stationary increments.
(iv) The process $\mathbb{X}(t)$ is a called Lévy process, if it has stationary independent increments and is stochastically continuous and $\mathbb{X}(0) = 0$.

Theorem 2.12 (Lévy-Khinchine formula) Let $\mathbb{X}(t)$ be a c$\grave{a}$dl$\grave{a}$g Lévy process on Y and let μ_t be the law of $\mathbb{X}(t)$. Then, there exists a unique triple (γ, Q, v) where $\gamma \in Y$, $Q \in L_1^+(Y)$, v is a nonnegative measure satisfying $v(\{0\}) = 0$ and

$$\int_Y (||y||_Y^2 \wedge 1) v(dy) < \infty,$$

such that

$$\int_Y e^{i\langle x,y\rangle_Y} \mu_t(dy) = e^{-t\Psi(x)},$$

where

$$\begin{aligned}\Psi(x) :&= -i\langle \gamma, x\rangle_Y + \frac{1}{2}\langle Qx, x\rangle_Y \\ &\quad + \int_Y \left(1 - e^{i\langle x,y\rangle_Y} + 1_{\{||y||<1\}}(y) i\langle x, y, \rangle_Y\right) v(dy).\end{aligned}$$

Proof See Peszat and Zabczyk [1, Theorem 4.24]. $\square$

Definition 2.25 We call the operator Q appearing in Theorem 2.12 the covariance of $\mathbb{X}(t)$, the measure μ the jump intensity measure of $\mathbb{X}(t)$ and the triple (γ, Q, ν) the characteristics of $\mathbb{X}(t)$.

Defining

$$N(t, A) := \#\{s \in (0, t] | \Delta\mathbb{X}(s) \in A\}, \quad A \in \mathcal{B}(Y \setminus \{0\}),$$

where

$$\Delta\mathbb{X}(s) = \begin{cases} \mathbb{X}(s) - \mathbb{X}(s-), & s > 0, \\ \mathbb{X}(0), & s = 0, \end{cases}$$

the Lévy process $\mathbb{X}(t)$ induces a Poisson random measure. We define the corresponding compensated Poisson random measure $\tilde{N}(t,A) := N(t,A) - t\nu(A), A \in \mathcal{B}(Y \setminus \{0\})$, where ν is the intensity measure of $\mathbb{X}(t)$.

Theorem 2.13 (Lévy-Itô decomposition) Let $\mathbb{X}(t)$ be a Lévy process on Y with the characteristics (γ, Q, ν). Then, for every $t \geq 0$,

$$\mathbb{X}(t) = t\gamma + w(t) + \int_{\{||x||_Y < 1\}} x\tilde{N}(t, dx) + \int_{\{||x||_Y \geq 1\}} xN(t, dx),$$

where $w(t)$ is a Wiener process with covariance Q independent of $N(\cdot, A)$ for all $A \in \mathcal{B}(Y \setminus \{0\})$.

Proof See Albeverio and Rüdiger [1, Theorem 4.1]. □

Definition 2.26 A Y-valued *càdlàg* process $\mathbb{X}(t)$ is called quasi-left-continuous if for every increasing sequence of stopping times $\{\tau_n\}_{n\in\mathbb{N}}$

$$\lim_{n\to\infty} \mathbb{X}(\tau_n) = \mathbb{X}(\lim_{n\to\infty} \tau_n) \quad \text{on} \quad \{\lim_{n\to\infty} \tau_n < \infty\}.$$

Proposition 2.11 Every Lévy process is quasi-left-continuous.

Proof See Bichteler [1, Lemma 4.6.7, p. 258]. □

2.4.4 Random Operators

Definition 2.27 A mapping $T(\omega) : \Omega \times X \to Y$ is said to be a random operator if $\{\omega : T(\omega)x \in B\} \in \mathcal{F}$ for all $x \in X, B \in \mathcal{B}(Y)$, where $(Y, \mathcal{B}(Y))$ is a measurable space.

In other words, the above definition simply states that $T(\omega)$ is a random operator if $T(\omega)x = y(\omega)$, say is a Y-valued random variable for every $x \in X$.

Let X and Y be separable. Let $L(X,Y)$ denote the space of bounded linear operators mapping X into Y. Let $T(\omega)$ be a random operator with values in $L(X,Y)$. The inverse $T^{-1}(\omega)$ of $T(\omega)$ from $\Omega \times Y \to X$ is defined if and only if $T(\omega)$ is one to one P-*a.s.*, which is the case if and only if $T(\omega)x = \theta$, P-*a.s.* implies $x = \theta$, P-*a.s.*

Definition 2.28 If $T(\omega)$ is a random operator with values in $L(X,Y)$, then $T^{-1}(\omega)$ is the random operator with values in $L(Y,X)$ which maps $T(\omega)x$ into x, P-*a.s.* Hence, $T^{-1}(\omega)T(\omega)x = x$, P-*a.s.*, $x \in D(T(\omega))$, and $T(\omega)T^{-1}(\omega)y = y$, P-*a.s.*, $y \in R(T(\omega))$. $T(\omega)$ is said to be invertible if $T^{-1}(\omega)$ exists.

The following result is from Hans [1].

Theorem 2.14 Let $T(\omega)$ be an invertible random operator with values in $L(X,Y)$, where X and Y are separable. Then $T^{-1}(\omega)$ is a random operator with values in $L(Y,X)$.

2.4.5 The Gelfand Triple

Definition 2.29 Let $(H, \langle\cdot,\cdot\rangle_H)$ be a real separable Hilbert space identified with its dual space H^* via the Riesz isomorphism $\mathbb{R}$. Let V be a Banach space with dual V^* such that the embedding $V \subset H$ is continuous, i.e.,

$$||v||_H \leq C||v||_V \quad \text{for all} \quad v \in V$$

and V is dense in H. (V,H,V^*) is called the Gelfand triple.

It follows that $H^* \subset V^*$ continuously and densely (see Zeidler [2, Proposition 23.13]). Consequently,

$$V \subset H \overset{\mathbb{R}}{\equiv} H^* \subset V^*$$

continuously and densely and

$${}_{V^*}\langle z,v\rangle_V = \langle z,v\rangle_H \quad \text{for all} \quad z \in H, v \in V.$$

Note that V^* is separable since $H \subset V^*$ continuously and densely, hence this is true for V as well.

2.5 Stochastic Calculus

This section is devoted to introducing stochastic calculus, more precisely, Itô stochastic integral and Itô's formula. To begin with, we define the Itô stochastic integral

$$\int_0^t \Phi(s)dw(s), \quad t \in [0,T],$$

where $w(t)$ is a Q-Wiener process on Y and Φ is a process with values that are linear but not necessarily bounded operators from Y into X. Next, we define an Itô stochastic integral when $w(t)$ is a cylindrical Wiener process. Subsequently, we also define a stochastic integral with respect to a compensated Poisson measure of the form

$$\int_0^t \int_Z \Phi(s,z)\tilde{N}(ds,dz),$$

where $\tilde{N}(dt,du)$ is the compensated Poisson measure.

2.5.1 Itô Stochastic Integral with respect to a Q-Wiener process

Let us fix $0 < T < \infty$. An $L = L(Y,X)$-valued process $\Phi(t), t \in [0,T]$, taking on a finite number of values is said to be elementary if there exists a sequence $0 = t_0 < t_1 < \ldots < t_k = T$ and a sequence $\Phi_0, \Phi_1, \ldots, \Phi_{k-1}$, of L-valued random variables taking only a finite member of values such that Φ_m are $\mathcal{F}_{t_m}$-measurable and $\Phi(t) = \Phi_m$, for $t \in (t_m, t_{m+1}]$, $m = 0,1,\ldots,k-1$. For elementary processes Φ one defines the stochastic integral as

$$\int_0^t \Phi(s)dw(s) = \sum_{m=0}^{k-1} \Phi_m \left(w_{t_{m+1}\wedge t} - w_{t_m \wedge t}\right) \tag{2.45}$$

and it is denoted by $\Phi \cdot w(t)$, $t \in [0,T]$.

In the construction of the stochastic integral for more general processes an important role will be played by the space of all Hilbert-Schmidt operators $L_2^0 = L_2(Y_0,X)$ from Y_0 to X. The space L_2^0 is also a separable Hilbert space, equipped with the norm

$$\begin{aligned}
||\Psi||^2_{L_2^0} &= \sum_{h,k=1}^{\infty} |\langle \Psi g_h, f_k \rangle|^2 \\
&= \sum_{h,k=1}^{\infty} \lambda_h |\langle \Psi e_h, f_k \rangle|^2 \\
&= ||\Psi Q^{1/2}||^2 = \mathrm{tr}[\Psi Q \Psi^*],
\end{aligned}$$

where $\{g_j\}$, with $g_j = \sqrt{\lambda_j} e_j, j = 1,2,,$ $\{e_j\}$ and $\{f_j\}$ are complete orthonormal bases in Y_0, Y, and X, respectively. One can check that $L \subset L_2^0$, but not all operators from L_2^0 can be regarded as restrictions of operators from L. The space L_2^0 contains genuinely unbounded operators on Y.

Let $\Phi(t)$, $t \in [0,T]$, be a measurable L_2^0 -valued process. Let us define the norm

$$\begin{aligned} ||\Phi||_t &= \left(E\int_0^t ||\Phi(s)||^2_{L_2^0} ds \right)^{1/2} \\ &= \left(E\int_0^t \operatorname{tr}(\Phi(s)Q^{1/2})(\Phi(s)Q^{1/2})^* ds \right)^{1/2}, \quad t \in [0,T]. \end{aligned}$$

Proposition 2.12 If a process Φ is elementary and $||\Phi||_T < \infty$, then the process $\Phi \cdot w(t)$ is a continuous, square integrable X-valued martingale on [0, T] and

$$E||\Phi \cdot w(t)||^2 = ||\Phi||_t^2, \quad 0 \le t \le T. \tag{2.46}$$

Proof See Da Prato and Zabczyk [1, Proposition 4.5]. □

Remark 2.3 Note that the stochastic integral is an isometric transformation from the space of all elementary processes equipped with the norm $||\cdot||_T$ into the space $M_T^2(X)$ of X-valued martingales.

The following σ-algebra $\mathcal{P}_\infty$ of subsets of $[0,\infty) \times \Omega$ will play an important role in what follows. $\mathcal{P}_\infty$ is the σ-field generated by sets of the form:

$$(s,t] \times F, \quad 0 \le s \le t \le \infty, \quad F \in \mathcal{F}_s \quad \text{and} \quad \{0\} \times F, \quad F \in \mathcal{F}_0.$$

This σ-algebra is called predictable σ-algebra and its elements predictable sets. The restriction of the σ-algebra $\mathcal{P}_\infty$ to $[0,T] \times \Omega$ will be denoted by $\mathcal{P}_T$. An arbitrary measurable mapping from $([0,\infty) \times \Omega, \mathcal{P}_\infty)$ or $([0,T] \times \Omega, \mathcal{P}_T)$ into $(X;\mathcal{B}(X))$ is called a predictable process. A predictable process is necessarily an adapted one.

To extend the definition of the stochastic integral to more general processes it is convenient to regard integrands as predictable processes with values in L_2^0; more precisely, measurable mappings from $(\Omega_\infty, \mathcal{P}_\infty)$ (respectively, $(\Omega_T, \mathcal{P}_T)$) into $(L_2^0, \mathcal{B}(L_2^0))$.

Proposition 2.13 The following statements hold:

(i) If a mapping Φ from Ω_T into L is L-predictable then Φ is also L_2^0-predictable. In particular, elementary processes are L_2^0-predictable.
(ii) If Φ is a L_2^0-predictable process such that $||\Phi||_T < \infty$, then there exists a sequence $\{\Phi_n\}$ of elementary processes such that $||\Phi - \Phi_n||_T \to 0$ as $n \to \infty$.

Proof See Da Prato and Zabczyk [1, Proposition 4.7]. □

We shall now extend the definition of the stochastic integral to all L_2^0-predictable process Φ such that $||\Phi||_T < \infty$ denoted by $N_w^2(0,T;L_2^0)$ which is a Hilbert space. This space is also denoted by $N_w^2(0,T)$ for simplicity. By Proposition 2.13, elementary processes form a dense set in $N_w^2(0,T)$ while by Proposition 2.12 the stochastic integral $\Phi \cdot w$ is an isometric transformation from this dense set into $M_T^2(X)$. Hence the definition of the integral can be extended to the whole of $N_w^2(0,T)$. Moreover, (2.45) holds and $\Phi \cdot w$ is a continuous square integrable martingale.

As a last step, the definition of the stochastic integral can be extended to L_2^0-predictable processes satisfying a weaker condition given by

$$P\left\{\int_0^T ||\Phi(s)||^2_{L_2^0} ds < \infty\right\} = 1. \tag{2.47}$$

Such processes are called stochastically integrable on $[0,T]$. They form a linear space denoted by $N_w(0,T;L_2^0)$, or simply $N_w(0,T)$. This extension can be accomplished by the so-called localization procedure. To do so, we need the following lemma.

A nonnegative random variable τ defined on $(\Omega,\mathcal{F})$ is said to be an $\mathcal{F}_t$-stopping time if, for arbitrary $t \geq 0$, $\{\omega \in \Omega; \tau(\omega) \leq t\} \in \mathcal{F}_t$.

Lemma 2.7 Assume that $\Phi \in N_w^2(0,T;L_2^0)$ and τ is an $\mathcal{F}_t$-stopping time such that $P(\tau \leq T) = 1$. Then

$$\int_0^T I_{[0,\tau]}(s)\Phi(s)dw(s) = \Phi \cdot w(\tau \wedge t), \quad P\text{-}a.s., t \in [0,T]. \tag{2.48}$$

Proof See Da Prato and Zabczyk [1, Lemma 4.9]. □

Let us assume that the condition (2.47) hold. Define

$$\tau_n = \inf\left\{t \in [0,T] : \int_0^t ||\Phi(s)||^2_{L_2^0} ds \geq n\right\}$$

with the convention that the infimum of an empty set is T. Then τ_n is a sequence such that

$$E\int_0^t ||I_{[0,\tau_n]}(s)\Phi(s)||^2_{L_2^0} ds < \infty. \tag{2.49}$$

Consequently, stochastic integrals $I_{[0,\tau_n]}(s)\Phi \cdot w(t)$, $t \in [0,T]$ are well defined for all $n = 1,2,\ldots$. Further, if $n < m$, then P-$a.s.$

$$\begin{aligned} I_{[0,\tau_n]}\Phi \cdot w(t) &= (I_{[0,\tau_n]}(I_{[0,\tau_m]}\Phi) \cdot w(t)) \\ &= (I_{[0,\tau_m]}\Phi) \cdot w(\tau_n \wedge t)), \quad t \in [0,T]. \end{aligned} \tag{2.50}$$

Hence one can assume that (2.49) holds for all $\omega \in \Omega$, $n < m$. For arbitrary $t \in [0,T]$, define

$$\Phi \cdot w(t) = I_{[0,\tau_n]}\Phi \cdot w(t), \tag{2.51}$$

when n is an arbitrary natural number such that $\tau_n \geq t$. Moreover, if $\tau_m \geq t$ and $m > n$, then

$$(I_{[0,\tau_m]}\Phi)\cdot w(t) = (I_{[0,\tau_m]}\Phi)\cdot w(\tau_n > t)$$
$$= I_{[0,\tau_n]}\Phi)\cdot w(t).$$

Therefore the definition (2.51) is consistent. By analogous arguments if $\{\tau_n'\} \uparrow T$ is another sequence satisfying (2.49) then the definition (2.51) leads to a stochastic process identical P -$a.s.$ for all $t \in [0,T]$. Note that for arbitrary $n = 1,2,\ldots$, $\omega \in \Omega$, $t \in [0,T]$,

$$\begin{aligned} \Phi\cdot w(\tau_n \wedge t) &= I_{[0,\tau_n]}\Phi\cdot w(\tau_n \wedge t) \\ &= M_n(\tau_n \wedge t), \quad t \in [0,T], \end{aligned} \tag{2.52}$$

where M_n is a square integrable continuous X-valued martingale. This property is referred to as the local martingale property of the stochastic integral.

Remark 2.4 It follows from the above construction that Lemma 2.7 is valid for all $\Phi \in N_w(0,T;L_2^0)$.

We collect below some important properties for the stochastic integral.

Proposition 2.14 Let $E\int_0^T ||\Phi(r)||^2_{L_2^0}dr < \infty$. Then for some constant $c > 0$,

$$\begin{aligned} P\left[\sup_{0\leq t\leq T}\left|\left|\int_0^t \Phi(r)dw(r)\right|\right| > c\right] &\leq \frac{1}{c^2}E\left|\left|\int_0^T \Phi(r)dw(r)\right|\right|^2 \\ &\leq \frac{\operatorname{tr} Q}{c^2}\int_0^\tau E||\Phi(r)||^2_{L_2^0}dr, \\ E\left[\sup_{0\leq t\leq T}\left|\left|\int_0^t \Phi(r)dw(r)\right|\right|^2\right] &\leq 4E\left|\left|\int_0^T \Phi(r)dw(r)\right|\right|^2 \\ &\leq 4\operatorname{tr} Q\int_0^T E||\Phi(r)||^2_{L_2^0}dr, \\ E\left[\sup_{0\leq t\leq T}\left|\left|\int_0^t \Phi(r)dw(r)\right|\right|\right] &\leq 3E\left|\left|\int_0^T \operatorname{tr}\Phi(r)Q\Phi^*(r)dr\right|\right|^{1/2}. \end{aligned}$$

Proof See Ichikawa [3]. $\square$

Proposition 2.15 Let $\int_0^T E||\Phi(r)||^p_{L_2^0}dr < \infty$ for some integer $p \geq 2$, and let $y(t) = \int_0^t \Phi(r)dw(r)$. Then

$$E||y(t)||^p \leq \left[\frac{1}{2}p(p-1)\right]^{p/2}\left[\int_0^t [E(\operatorname{tr}\Phi(r)Q\Phi^*(r))^{p/2}]^{2/p}dr\right]^{p/2}$$

$$\leq \left[\frac{1}{2}p(p-1)\right]^{p/2} (\mathrm{tr}\, Q)^{p/2} t^{p/2-1} \int_0^t E||\Phi(r)||^p_{L_2^0} dr.$$

Proof See Ichikawa [3]. □

2.5.2 Itô Stochastic Integral with respect to a Cylindrical Wiener Process

Let us fix $Q \in L(Y)$ nonnegative, symmetric but not necessarily of finite trace. We now define a stochastic integral with respect to a cylindrical Wiener process, precisely with respect to the standard Y_1-valued Q_1-Wiener process given by Proposition 2.8. We consider a process $\Phi(t), t \in [0,T]$ that is integrable with respect to this Q_1-Wiener process if it takes values in $L_2(Q_1^{1/2}(Y_1), X)$, is predictable and if

$$P\left\{\int_0^T ||\Phi(s)||^2_{L_2(Q_1^{1/2}(Y_1),X)} ds < \infty\right\} = 1. \tag{2.53}$$

We have by Proposition 2.8 that $Q^{1/2}(Y_1) = J(Y_0)$ and that

$$\begin{aligned}\langle Ju_0, Jv_0\rangle_{Q^{1/2}(Y_1)} &= \langle Q^{-1/2}Ju_0, Q^{-1/2}Jv_0\rangle,\\ &= \langle u_0, v_0\rangle_0\end{aligned}$$

for all $u_0, v_0 \in Y_0$. In particular, it follows that Je_k, $k \in \mathbb{N}$ is an orthonormal basis of $Q^{1/2}(Y_1)$. Hence

$$\begin{aligned}&\Phi \in L_2^0 = L_2(Q^{1/2}(Y), X)\\ &\qquad \Longleftrightarrow \Phi \circ J^{-1} \in L_2(Q^{1/2}(Y), X)\end{aligned}$$

since

$$\begin{aligned}||\Phi||^2_{L_2^0} &= \sum_{k\in\mathbb{N}} \langle \Phi e_k, \Phi e_k\rangle\\ &= \sum_{k\in\mathbb{N}} \langle \Phi \circ J^{-1}(Je_k), \Phi \circ J^{-1}(Je_k)\rangle\\ &= ||\Phi \circ J^{-1}||^2_{L_2(Q_1^{1/2}(Y_1),X)}.\end{aligned}$$

Now define

$$\int_0^t \Phi(s) dW(s) := \int_0^t \Phi(s) \circ J^{-1} dw(s), \quad t \in [0,T], \tag{2.54}$$

where the class of all integrable processes is given by $N_w = \Big\{ \Phi : \Omega_T \to L_2^0 | \Phi$ predictable and $P\left(\int_0^t ||\Phi(s)||^2_{L_2^0} ds < \infty \right) = 1 \Big\}$ as in the case of a standard Q-Wiener process $w(t), t \in [0,T]$ in Y.

Remark 2.5

(i) The stochastic integral defined in (2.54) is independent of the choice of $(Y_1, \langle \cdot, \cdot \rangle_1)$ and J.

(ii) If $Q \in L(Y)$ is nonnegative, symmetric, and with finite trace the standard Q-Wiener can be considered as a cylindrical Wiener process by setting $J = I : Y_0 \to Y$, where I is the identity map. In this case the definition (2.54) coincides with the definition of stochastic integral given in Section 2.5.1.

2.5.3 Stochastic Integral with respect to a Compensated Poisson Measure

In this subsection, we shall define the stochastic integral with respect to a compensated Poisson measure induced by a Poisson point process.

Let $(X, \langle \cdot, \cdot \rangle_X)$ be a separable Hilbert space and $(Z, \mathcal{Z})$ be a measure space with a σ-finite measure ν. Further, let p be a stationary $\mathcal{F}_t$-Poisson point process Z with characteristic measure ν.

The Poisson point process p induces a Poisson random measure N on $[0,T] \times Z$ (see Remark 2.2) and by Proposition 2.10, the compensator of N is given by $dt \otimes \nu$. The measure $\tilde{N} := N - dt \otimes \nu$ is called the compensated Poisson measure of N.

Remark 2.6 The integration theory in Knoche [1] is developed with respect to an $\mathcal{F}_t$-Poisson point process of class (QL) (see Definition 2.23). However, by Proposition 2.10, a stationary process is automatically of class (QL) and therefore, all results of Knoche [1] apply to this special case. Throughout this book, we always assume p being a stationary $\mathcal{F}_t$-Poisson point process.

Set

$$\Gamma := \{ B \in Z \,|\, \nu(B) < \infty \}$$

and define the predictable σ-field

$$\begin{aligned} \mathcal{P}_T(Z) &:= \sigma(g : [0,T] \times \Omega \times Z \to R \,|\, g \,\text{is}\, \mathcal{F}_t \otimes Z - \text{adapted and left-continuous}) \\ &= \sigma(\{(s,t] \times F_s \times B \,|\, 0 \le s \le t \le T, F_s \in \mathcal{F}_s, B \in \mathcal{Z}\} \\ &\quad \cup \{\{0\} \times F_0 \times B \,|\, F_0 \in \mathcal{F}_0, B \in \mathcal{Z}\}). \end{aligned}$$

In the first step, we define the stochastic integral with respect to $\tilde{N}$ for elementary processes.

Definition 2.30

(i) An X-valued process $\Phi(t) : \Omega \times Z \to X$, $t \in [0,T]$ is said to be elementary if there exists a partition $0 = t_0 < t_1 < \ldots < t_k = T$ and for $m \in \{0,\ldots,k-1\}$ there exist $B_1^m,\ldots,B_n^m \in \Gamma$ pairwise disjoint, such that

$$\Phi = \sum_{m=0}^{k-1}\sum_{i=1}^{n} \Phi_i^m I_{(t_m,t_{m+1}]\times B_i^m},$$

where $\Phi_i^m \in L^2(\Omega,\mathcal{F}_{t_m},P;X)$, $1 \le i \le n$, $0 \le m \le k-1$.

(ii) The linear space of all elementary processes is denoted by $\mathcal{E}$.

For $\Phi \in \mathcal{E}$ and $t \in [0,T]$, we define the stochastic integral by

$$\begin{aligned}\mathrm{Int}(\Phi)(t) &:= \int_0^t \int_Z \Phi(s,z)\tilde{N}(ds,dz) \\ &:= \sum_{m=0}^{k-1}\sum_{i=1}^{n} \Phi_i^m(\tilde{N}(t_{m+1}\wedge t, B_i^m) - \tilde{N}(t_m \wedge t, B_i^m)).\end{aligned} \tag{2.55}$$

Then $\mathrm{Int}(\Phi)$ is P-$a.s.$ well defined and Int is linear in $\Phi \in \mathcal{E}$. For $\Phi \in \mathcal{E}$, define

$$||\Phi||_T^2 := E\left[\int_0^t \int_Z ||\Phi(s,z)||_X^2 \nu(dz)ds\right].$$

Proposition 2.16 If $\Phi \in \mathcal{E}$ then $\mathrm{Int}(\Phi) \in \mathcal{M}_T^2(X)$, $\mathrm{Int}(\Phi)(0) = 0$ P-$a.s.$ and for all $t \in [0,T]$

$$E||\mathrm{Int}(\Phi)(t)||_X^2 = E\left[\int_0^t \int_Z ||\Phi(s,z)||_X^2 \nu(dz)ds\right].$$

In particular, $\mathrm{Int} : (\mathcal{E}, ||\cdot||_T^2) \to (\mathcal{M}_T^2(X), ||\cdot||_{\mathcal{M}_T^2})$ is an isometry,

$$||\mathrm{Int}(\Phi)||_{\mathcal{M}_T^2} = ||\Phi||_T^2.$$

Proof See Knoche [1, Proposition 2.22]. □

In order to get a norm on $\mathcal{E}$ one has to consider the equivalence class of elementary processes with respect to $||\cdot||_T$. For simplicity, the space of equivalence classes is again denoted by $\mathcal{E}$. Since $\mathcal{E}$ is dense in the abstract completion $\bar{\mathcal{E}}^{||\cdot||_T}$ of $\mathcal{E}$ with respect to $||\cdot||_T$, there exists a unique isometric extension of Int to $\bar{\mathcal{E}}^{||\cdot||_T}$. In particular, the isometric formula in Proposition 2.16 does also hold for every process in $\bar{\mathcal{E}}^{||\cdot||_T}$.

The completion of $\mathcal{E}$ with respect to $||\cdot||_T$ can be characterized as follows:

Proposition 2.17 Let $\mathcal{P}_T(Z)$ be the predictable σ-field on $[0,T]\times\Omega\times Z$ and

$$\begin{aligned}\mathcal{N}^2_{\tilde{N}}(T,Z;X) &:= \Big\{\Phi:[0,T]\times\Omega\times Z\to X\,|\,\Phi\,\text{is}\,\mathcal{P}_T(Z)/\mathcal{B}(X)\text{-measurable}\\ &\qquad \text{and}\,||\cdot||_T = E\left[\int_0^T\int_Z ||\Phi(s,z)||_X^2\nu(dz)ds\right]^{1/2}<\infty\Big\}\\ &= L^2([0,T]\times\Omega\times Z,\mathcal{P}_T(Z),dt\otimes P\otimes\nu;X).\end{aligned}$$

Then

$$\bar{\mathcal{E}}^{||\cdot||_T} = \mathcal{N}^2_{\tilde{N}}(T,Z;X).$$

Proof See Knoche [1, Proposition 2.24]. □

The following are some important properties of the Poisson integral.

Proposition 2.18 Let $\Phi\in\mathcal{N}^2_{\tilde{N}}(T,Z;X)$. Let $\tilde{X}$ be another Hilbert space and $L\in L(X,\tilde{X})$. Then $L(\Phi)\in\mathcal{N}^2_{\tilde{N}}(T,Z;\tilde{X})$ and for all $t\in[0,T]$,

$$L\left(\int_0^t\int_Z\phi(s,z)\tilde{N}(ds,dz)\right) = \int_0^t\int_Z L(\phi(s,z))\tilde{N}(ds,dz)\quad P\text{-}a.s.$$

Proof See Knoche [1, Proposition 3.7]. □

Proposition 2.19 Let $\Phi\in\mathcal{N}^2_{\tilde{N}}(T,Z;X)$. Then for all $t\in[0,T]$,

$$E\left[\int_0^t\int_Z||\phi(s,z)||_X^2\tilde{N}(ds,dz)\right] = E\left[\int_0^t\int_Z||\phi(s,z)||_X^2\nu(dz)ds\right].$$

Proof See Knoche [1, Proposition 3.1]. □

Let us denote the square bracket of an X-valued process $X(t)$ by $[X]_t$.

Proposition 2.20 Let $\Phi\in\mathcal{N}^2_{\tilde{N}}(T,Z;R)$. Then, for $t\geq 0$,

$$I_\Phi(t) := \int_0^t\int_Z\phi(s,z)\tilde{N}(ds,dz)\in M_T^2(R),$$

and

$$[I_\Phi]_t = \int_0^t\int_Z||\phi(s,z)||_X^2 N(ds,dz).$$

Proof See Peszat and Zabczyk [1, Theorem 8.23]. □

2.5.4 Itô's Formula for the case of a Q-Wiener Process

In the rest of this section we give some basic Itô's formula in various settings. Some more such formulas are given later on as and when needed.

Theorem 2.15 Let $Q \in L(Y)$ be a symmetric nonnegative trace-class operator, and let $\{w(t),\, t \in [0,T]\}$ be a Y-valued Q-Wiener process on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \le t \le T}, P)$. Assume that a stochastic process $x(t)$, $t \in [0,T]$, is given by

$$x(t) = x_0 + \int_0^t \psi(s)ds + \int_0^t \Phi(s)dw(s),$$

where x_0 is an $\mathcal{F}_0$-measurable X-valued random variable, $\psi(s)$ is an X-valued predictable process Bochner-integrable process on $[0,T]$, and Φ is an L_2^0-valued process stochastically integrable on $[0,T]$.

Assume that a function $v \in C^{1,2}([0,T] \times X, R)$, i.e., $v : [0,T] \times X \to R$ is such that v is continuous and so also v_t, and its Fréchet partial derivatives v_x, v_{xx} are continuous and bounded on bounded subsets of $[0,T] \times X$. Then the following Itô's formula holds:

$$\begin{aligned} v(t,x(t)) &= v(0,x(0)) + \int_0^t \langle v_x(s,x(s)), \Phi(s)dw(s)\rangle_X \\ &\quad + \int_0^t \Big\{ v_t(s,x(s)) + \langle v_x(s,x(s)), \psi(s)\rangle_X \\ &\quad + \frac{1}{2}\mathrm{tr}\big[v_{xx}(x,x(s))(\Phi(s)Q^{1/2})(\Phi(s)Q^{1/2})^*\big] \Big\} ds, \end{aligned} \tag{2.56}$$

P-$a.s.$ for all $t \in [0,T]$.

Proof See Da Prato and Zabczyk [1, Theorem 4.17]. □

Let $M(Y,X)$ be the space of stochastic processes $\Phi(\cdot,\cdot) : [0,T] \times \Omega \to L(Y,X)$ which are strongly measurable, i.e., $\Phi(t,\cdot)y$ is a measurable stochastic process for all $y \in Y$. Define also $M_1(Y,X) = \Big\{ \Phi \in M(Y,X) : \int_0^T ||\Phi(t)||^2 dt < \infty, P\text{-}a.s. \Big\}$. Let $x(t)$, $t \in [0,T]$, have a stochastic differential:

$$x(t) = x_0 + \int_0^t \psi(s)ds + \int_0^t \Phi(s)dw(s), \tag{2.57}$$

where x_0 is an $\mathcal{F}_0$-measurable X-valued random variable, $\psi(s)$ is an X-valued and adapted to $\mathcal{F}_t$ with $\int_0^T ||\psi(t)||dt < \infty$ P-$a.s.$ and $\Phi \in M_1(Y,X)$. Let Z be a real separable Hilbert space and let $P(\cdot,\cdot) \in L(X \times X, Z)$ and $\Phi \in L(Y,X)$. We define

$$\operatorname{tr} P[\Phi; Q] = \sum_{i=1}^{\infty} \lambda_i P(\Phi e_i, \Phi e_i) \in Z.$$

We have an Itô's formula in a Hilbert space.

Theorem 2.16 Suppose that $v(t,x) : [0,T] \times X \to Z$ is continuous with properties:

(i) $v(t,x)$ is differentiable in t and $v_t(t,x)$ is continuous on $[0,T] \times X$,
(ii) $v(t,x)$ is twice Fréchet differentiable in y and $v_x(t,x)x_1 \in Z$, $v_{xx}(t,x)(x_1,x_2) \in Z$ are continuous on $[0,T] \times X$ for all x, x_1, $x_2 \in X$.

If $x(t)$ is given as in (2.57), then $z(t) = v(t,x(t))$ has the stochastic differential

$$\begin{aligned} dz(t) = & \Big\{ v_t(t,x(t)) + v_x(t,x(t))\psi(t) \\ & + \frac{1}{2} \operatorname{tr} v_{xx}(t,x(t))[\Phi(t);Q] \Big\} dt \\ & + v_x(t,x(t))\Phi(t)dw(t). \end{aligned} \tag{2.58}$$

Proof See Ichikawa [3]. □

In applications to stochastic evolution equations, we need the following:

Corollary 2.6 Let A be a closed linear operator with dense domain $D(A)$ in X. Let $v(t,x)$ satisfy the hypothesis of Theorem 2.16 except (i) which is replaced by

(a) $v(t,x)$ is differentiable in t for each $x \in D(A)$ and $v_t(t,x)$ is continuous on $[0,T] \times D(A)$, where $D(A)$ is equipped with the graph norm of A, i.e., $||x||^2_{D(A)} = ||x||^2 + ||Ax||^2$.

Let $x(t)$ be as given in (2.57) with $x_0 \in D(A)$, $\int_0^t ||A\psi(t)||dt < \infty$ P-$a.s.$ and $A\Phi \in M_1(Y,X)$. Then the conclusion of Theorem 2.16 holds.

2.5.5 Itô's Formula for the case of a Cylindrical Wiener Process

Let $M_2(Y,X)$ denote the class of $L_2(Y,X)$-valued stochastic processes adapted to the filtration $\{\mathcal{F}_t\}_{t \le T}$, measurable as mappings from $([0,T] \times \Omega, \mathcal{B}([0,T]) \otimes \mathcal{F}_T)$ to $(L_2(Y,X), \mathcal{B}(L_2(Y,X)))$ and

$$P\left[\int_0^T ||\Phi(t)||^2_{L_2(Y,X)} dt < \infty\right] = 1.$$

Theorem 2.17 Let $\{w(t), 0 \le t \le T\}$ be a Y-valued cylindrical Wiener process on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \le t \le T}, P)$. Assume that a stochastic process $\{x(t), 0 \le t \le T\}$ is given by

$$x(t) = x_0 + \int_0^t \psi(s)ds + \int_0^t \Phi(s)dw(s),$$

where x_0 is an $\mathcal{F}_0$-measurable X-valued random variable, $\psi(s)$ is an X-valued $\mathcal{F}_s$-measurable P-$a.s.$ Bochner-integrable process on $[0,T]$,

$$\int_0^t ||\psi(s)||_H ds < \infty, \quad P\text{-}a.s.,$$

and $\Phi \in M_2(Y,X)$. Assume that a function $v : [0,T] \times X \to R$ is such that v is continuous and its Fréchet partial derivatives v_t, v_x, v_{xx} are continuous and bounded on bounded subsets of $[0,T] \times X$. Then the following Itô's formula holds:

$$\begin{aligned} v(t,x(t)) &= v(0,x(0)) + \int_0^t \langle v_x(s,x(s)), \Phi(s)dw(s)\rangle_X \\ &\quad + \int_0^t \Big\{ v_t(s,x(s)) + \langle v_x(s,x(s)), \psi(s)\rangle_X \\ &\quad + \frac{1}{2}\mathrm{tr}\left[v_{xx}(s,x(s))\Phi(s)(\Phi(s))^*\right] \Big\} ds, \end{aligned} \tag{2.59}$$

P-$a.s.$ for all $t \in [0,T]$.

Proof See Gawarecki and Mandrekar [1, Theorem 2.10]. □

2.5.6 Itô's Formula for the case of a Compensated Poisson process

We give an the Itô's formula based on Mao and Yuan [1, Theorem 1.45, p. 48] and Peszat and Zabczyk [1, Theorem D.2, p. 392].

Let Z be a vector space with a norm $||\cdot||$. Let $\mathcal{B}(Z)$ be a Borel σ-algebra on Z and $\nu(dz)$, a σ-finite measure defined on $\mathcal{B}(Z)$. Let m be a positive integer. Let $\{r(t), t \in R^+\}$ be a right-continuous irreducible Markov chain on the probability space $(\Omega, \mathcal{F}, P)$ taking values in a finite state space $S = \{1, 2, \ldots, m\}$ with generator $\Gamma = (\gamma_{ij})_{m \times m}$ given by

$$P\{r(t+h) = j | r(t) = i\} = \begin{cases} \gamma_{ij}h + o(h), & if \quad i \ne j, \\ 1 + \gamma_{ii}h + o(h), & if \quad i = j, \end{cases}$$

for any $t \geq 0$ and small $h > 0$. Here $\gamma_{ij} \geq 0$ is the rate of transition from i to j, if $i \neq j$, while $\gamma_{ii} = -\sum_{i \neq j} \gamma_{ij}$.

Theorem 2.18 Let $\{w(t), 0 \leq t \leq T\}$ be a Y-valued cylindrical Wiener process on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, P)$. Assume that a stochastic process $\{x(t), 0 \leq t \leq T\}$ is given by

$$\begin{aligned} x(t) = x(0) &+ \int_0^t F(s, x(s), r(s))ds + \int_0^t G(s, x(s), r(s))dw(s) \\ &+ \int_0^t \int_Z \Phi(s, x(s-), r(s), u)\tilde{N}(ds, du), \end{aligned}$$

where $f : [0,T] \times X \times S \to X$, $g\colon [0,T] \times X \times S \to L_2(X,X)$, and $\Phi\colon [0,T] \times X \times S \times Z \to X$; $x(0) = x_0 \in X$ and $r(0) = r_0 \in S$ and $x(t-) = \lim_{s \uparrow t} x(s)$ and the integrals are all well defined. We assume further that the Wiener process $w(t)$, the compensated Poisson process $\tilde{N}(ds, du)$ and the Markov chain $r(t)$ are all independent.

Let $U : R^+ \times X \times S \to R^+$ be continuous and its Fréchet partial derivatives U_t, U_x, U_{xx} are continuous and bounded on bounded subsets of $[0,T] \times X$. For $t \geq 0$, $x \in D(A)$ and $i \in S$, define an operator

$$\begin{aligned} \mathscr{L}U(t,x,i) := \; & U_t(t,x,i) + \langle Ax + F(t,x,i), U_x(t,x,i)\rangle_X \\ & + \sum_{j=1}^m \gamma_{ij} U(t,x,j) + \frac{1}{2}\mathrm{tr}[U_{xx}(t,x,i)G(t,x,i)G^*(t,x,i)] \\ & + \int_{\mathbb{Z}} [U(t, x + \Phi(t,x,i,u), i) - U(t,x,i) \\ & - \langle U_x(t,x,i), \Phi(t,x,i,u)\rangle_X]\nu(du). \end{aligned}$$

Then the following Itô's formula holds:.

$$\begin{aligned} U(t, x(t), r(t)) = \; & U(0, x_0, r_0) + \int_0^t \mathscr{L}U(s, x(s), r(s))ds \\ & + \int_0^t \langle U_x(s, x(s), r(s)), G(s, x(s), r(s))dw(s)\rangle_X \\ & + \int_0^t \int_{\mathbb{Z}} [U(s, x(s-) + \Phi(s, x(s-), r(s), u), r(s)) \\ & - U(s, x(s-), r(s))]\tilde{N}(ds, du) \\ & + \int_0^t \int_R [U_x(s, x(s-), r_0 + h(r(s), \ell)) \\ & - U(s, x(s-), r(s))]N(ds, d\ell), \end{aligned} \tag{2.60}$$

where $N(ds,d\ell)$ is a Poisson random measure with intensity $ds \times \vartheta(d\ell)$ and ϑ is a Lebesgue measure on R.

For more details on the function h and the martingale measure $N(ds,d\ell)$, we refer to Mao and Yuan [1].

2.6 The Stochastic Fubini Theorem

Let us begin with a basic stochastic Fubini theorem from Ichikawa [3].

Proposition 2.21 Let $I = [0,T]$ and let $G : I \times I \times \Omega \to L(Y,X)$ be strongly measurable such that $G(s,t)$ is $\mathcal{F}_t$-measurable for each s and

$$\int_0^T \int_0^T ||G(t,s)||^2 ds\, dt < \infty \quad P\text{-}a.s.$$

Then

$$\int_0^T \int_0^T G(t,s)dw(s)dt = \int_0^T \int_0^T G(t,s)dtdw(s) \quad P\text{-}a.s., \tag{2.61}$$

where we interpret the right-hand side as $\sum_{i=1}^{\infty} \int_0^T \int_0^T G(t,s)e_i\, dt\, d\beta_i(s)$.

The following version is more general.

Let $(X,\mathcal{X})$ be a measurable space and let $\Phi : (t,\omega,x) \to \Phi(t,\omega,x)$ be a measurable mapping from

$$(\Omega_T \times X, \mathcal{P}_T \times \mathcal{B}(X)) \quad \text{into} \quad (L_2^0, \mathcal{B}(L_2^0)). \tag{2.62}$$

Thus, in particular, for arbitrary $x \in X$, $\Phi(\cdot,\cdot,x)$ is a predictable L_2^0-valued process. Let in addition μ be a finite positive measure on $(X,\mathcal{X})$.

Proposition 2.22 Assume (2.62) and that:

$$\int_X ||\Phi(\cdot,\cdot,x)||_T \mu(dx) < \infty \tag{2.63}$$

then P-$a.s.$

$$\int_X \left[\int_0^T \Phi(t,x)dw(t)\right]\mu(dx) = \int_X \left[\int_0^T \Phi(t,x)\mu(dx)\right]dw(t). \tag{2.64}$$

Proof See Da Prato and Zabczyk [1, Theorem 4.18]. □

The following stochastic Fubini theorem involving Poisson integral will also be needed in the sequel.

Let $\mathcal{P} = \mathcal{P}([0,T] \times \Omega)$ denote the predictable σ-algebra and $(Z, \mathcal{Z}, \mu)$ be a finite measure space. Let $\mathcal{O} \in \mathcal{B}(Y - \{0\})$ and $\mathcal{H}_2(T, \mathcal{O}, Z)$ be the real Hilbert space of all $\mathcal{P} \times \mathcal{B}(\mathcal{O}) \times \mathcal{Z}$-measurable functions G from $[0,T] \times \Omega \times \mathcal{O} \times Z \to X$ for which

$$\int_Z \int_0^T \int_{\mathcal{O}} E||G(s,y,z)||_X^2 \, \nu(dy)\, ds\, \mu(dz) < \infty.$$

The space $S(T, \mathcal{O}, Z)$ is dense in $\mathcal{H}_2(T, \mathcal{O}, Z)$, where $G \in S(T, \mathcal{O}, Z)$ if

$$G = \sum_{i=0}^{N_1} \sum_{j=0}^{N_2} \sum_{k=0}^{N_3} G_{ijk} \chi_{A_i} \chi_{(t_j, t_{j+1}]} \chi_{B_k},$$

where N_1, N_2, $N_3 \in \mathbb{N}$, $A_0, \ldots, A_{N_1}$ are disjoint sets in $\mathcal{B}(\mathcal{O})$, $0 = t_0 < t_1 < \cdots < t_{N_2+1} = T$, $B_0, \ldots, B_{N_3}$ is a partition of Z, wherein each $B_k \in \mathcal{Z}$ and each G_{ijk} is a bounded $\mathcal{F}_{i_j}$-measurable random variable with values in X.

Proposition 2.23 If $G \in \mathcal{H}_2(T, \mathcal{O}, Z)$, then for each $0 \leq t \leq T$,

$$\begin{aligned} &\int_Z \left(\int_0^t \int_{\mathcal{O}} G(s,y,z) \tilde{N}(ds, dy) \right) \mu(dz) \\ &= \int_0^t \int_{\mathcal{O}}^t \left(\int_Z G(s,y,z) \mu(dz) \right) \tilde{N}(ds, dy), \quad P\text{-}a.s. \end{aligned} \tag{2.65}$$

Proof See Luo and Liu [1]. □

2.7 Stochastic Convolution Integrals

In this section, we collect some properties of stochastic convolution integrals. In Section 2.7.1, we present another use of Yosida approximation to estimate such integrals.

The following lemma is from Da Prato and Zabczyk [2].

Lemma 2.8 Let $W_A^\Phi(t) = \int_0^t S(t-s)\Phi(s)dw(s)$, $t \in [0,T]$. For any arbitrary $p > 2$, there exists a constant $c(p,T) > 0$ such that for any $T \geq 0$ and a proper modification of the stochastic convolution W_A^Φ, one has

$$E \sup_{t \leq T} ||W_A^\Phi(t)||^p \leq c(p,T) \sup_{t \leq T} ||S(t)||^p E \int_0^t ||\Phi(s)||_{L_2^0}^p ds.$$

Moreover, if $E \int_0^T ||\Phi(s)||_{L_2^0}^p ds < \infty$, then there exists a continuous version of the process $\{W_A^\Phi, t \geq 0\}$.

Lemma 2.9 Suppose A generates a contraction semigroup. Then the process $W_A^{\Phi}(\cdot)$ has a continuous modification and there exists a constant $k > 0$ such that

$$E \sup_{s\in[0,T]} ||W_A^{\Phi}(s)||^2 \leq kE\int_0^t ||\Phi(s)||^2_{L_2^0} ds, \quad t \in [0,T].$$

Proof See Da Prato and Zabczyk [1, Theorem 6.10]. □

2.7.1 A Property using Yosida Approximations

Lemma 2.10 Let $r > 1$, $T > 0$ and let Φ be a L_2^0-valued predictable process such that $E\int_0^T ||\Phi(s)||^{2r}_{L_2^0} ds < \infty$. There exists a constant $C_T > 0$ such that

$$E \sup_{t\in[0,T]} \left|\left|\int_0^t S(t-s)\Phi(s)dw(s)\right|\right|^{2r} \leq C_T E\left(\int_0^T ||\Phi(s)||^{2r}_{L_2^0} ds\right). \tag{2.66}$$

Moreover

$$\lim_{n\to\infty} E \sup_{t\in[0,T]} ||W_A^{\Phi}(t) - W_{A,n}^{\Phi}(t)||^{2r} = 0, \tag{2.67}$$

where W_A^{Φ} and $W_{A,n}^{\Phi}$ are defined as

$$\begin{aligned} W_A^{\Phi}(t) &= \int_0^t S(t-s)\Phi(s)dw(s) \\ W_{A,n}^{\Phi}(t) &= \int_0^t e^{(t-s)A_n}\Phi(s)dw(s), \quad t \in [0,T], \end{aligned} \tag{2.68}$$

and A_n are the Yosida approximations of A. Finally, W_A^{Φ} has a continuous modification.

Proof We will use the factorization method, see the proof of Theorem 5.14 (see Da Prato and Zabczyk [1]). Let $\alpha \in (1/2r, 1/2)$, the stochastic Fubini theorem (see Proposition 2.22) implies that

$$W_A^{\Phi}(t) = \frac{\sin \pi\alpha}{\pi}\int_0^t (t-s)^{\alpha-1} S(t-s)Y(s)ds, \quad t \in [0,T],$$

where

$$Y(s) = \int_0^s (s-\sigma)^{-\alpha} S(s-\sigma)\Phi(\sigma)dw(\sigma), \quad s \in [0,T],$$

Since $\alpha > 1/2r$, applying Hölder's inequality one obtains that there exists a constant $C_{1,T} > 0$ such that

$$\sup_{t\in[0,T]} ||W_A^{\Phi}(t)||^{2r} \le C_{1,T}\int_0^T ||Y(s)||^{2r}ds. \tag{2.69}$$

Moreover, by Lemma 7.2 (see Da Prato and Zabczyk [1]), there exists a constant $C_{2,T} > 0$ such that

$$E||Y(s)||^{2r} \le C_{2,T}E\left(\int_0^s (s-\sigma)^{-2\alpha}||\Phi(\sigma)||^2_{L_2^0}d\sigma\right)^r, \tag{2.70}$$

from which using the Young's inequality,

$$\begin{aligned}\int_0^T E||Y(s)||^{2r}ds &\le C_{2,T}E\left(\int_0^T (s-\sigma)^{-2\alpha}d\sigma\right)^r\int_0^s ||\Phi(\sigma)||^{2r}_{L_2^0}d\sigma\\ &\le C_{3,T}E\left(\int_0^T ||\Phi(\sigma)||^{2r}_{L_2^0}d\sigma\right).\end{aligned}$$

This finishes the proof of (2.66) with $C_T = TC_{1,T}C_{3,T}$.

We now prove (2.67) we have:

$$W_{A,n}^{\Phi}(t) = \frac{\sin\pi\alpha}{\pi}\int_0^t e^{(t-s)A_n}(t-s)^{\alpha-1}Y_n(s)ds,$$

where

$$Y_n(s) = \int_0^s e^{(s-\sigma)A_n}(s-\sigma)^{-\alpha}\Phi(\sigma)dw(\sigma).$$

Thus, we can write

$$\begin{aligned}W_A^{\Phi}(t) - W_{A,n}^{\Phi}(t) &= \frac{\sin\pi\alpha}{\pi}\int_0^t [S(t-s)-e^{(t-s)A_n}](t-s)^{\alpha-1}Y(s)ds\\ &\quad+\frac{\sin\pi\alpha}{\pi}\int_0^t e^{(t-s)A_n}(t-s)^{\alpha-1}[Y(s)-Y_n(s)]ds,\\ &= I_n(t)+J_n(t).\end{aligned}$$

We proceed now in two steps.

Step 1 We show that

$$\lim_{n\to\infty} E\sup_{t\in[0,T]} ||I_n(t)||^{2r} = 0. \tag{2.71}$$

Since $\sum_n(t) = S(t) - e^{tA_n}$; then, by the Hölder's inequality, there exists $C_{4,T} > 0$ such that

$$\sup_{t\in[0,T]} ||I_n(t)||^{2r} \leq C_{4,T} \int_0^T \left|\left|\sum_n (t-s)Y(s)\right|\right|^{2r} ds.$$

So (2.71) follows from the dominated convergence theorem.

Step 2 We have

$$\lim_{n\to\infty} E \sup_{t\in[0,T]} ||J_n(t)||^{2r} = 0. \tag{2.72}$$

The following estimate is proved as (2.69):

$$\sup_{t\in[0,T]} ||J_n(t)||^{2r} \leq C_{2,T} \int_0^T ||Y(s) - Y_n(s)||^{2r} ds. \tag{2.73}$$

Now, by the Young's inequality

$$\int_0^T E||Y(s) - Y_n(s)||^{2r} ds \leq C_{3,T} E\left(\int_0^T ||\sum_n(\sigma)\Phi(\sigma)||^{2r}_{L_2^0} d\sigma\right), \tag{2.74}$$

and (2.72) follows letting n tend to infinity.

Finally, the existence of the continuous modification of W_A^Φ follows easily from (2.67). □

For more applications, we refer to Da Prato and Zabczyk [1, Chapter 6].

2.8 Burkholder Type Inequalities

The following lemma could be regarded as the stochastic convolution inequality of Burkholder type in infinite dimensions. Consider the process:

$$W_A^F(t) = \int_0^t S(t-s)F(s)dw(s) \tag{2.75}$$

defined for any fixed $t \in [0,T]$, where $\{S(t) : t \geq 0\}$ is a strongly continuous semigroup of bounded linear operators with infinitesimal generator A on X and $F(t)$ is some appropriate stochastic process. It is well known that for such an operator A, there exists a nonnegative number $\alpha \geq 0$ such that

$$\langle Ax, x\rangle \leq \alpha||x||^2, \quad \forall x \in D(A). \tag{2.76}$$

Lemma 2.11 Assume $T > 0$, $F(t) : \Omega \times R^+ \to L(Y,X)$, is a progressively measurable process, and for some $p \geq 2$,

$$E\left(\int_0^T \mathrm{tr}[F(s)QF(s)^*]ds\right)^{p/2} < +\infty, \tag{2.77}$$

then there exists a positive constant $C_p > 0$, depending on p and α, such that

$$E \sup_{0\leq s\leq T} ||W_A^F(t)||^p \leq C_p e^{p^2\alpha T} E\left(\int_0^T \mathrm{tr}\,[F(s)QF(s)^*]ds\right)^{p/2}. \tag{2.78}$$

Proof See Tubaro [1]. □

We next consider a Burkholder type inequality for the Poisson integral.

Lemma 2.12 Assume that Φ: $\Omega \times R^+ \times Z \to X$ is a progressively measurable process, and for $p \geq 2$,

$$E\int_0^T \int_Z ||\Phi(s,u)||_X^p \nu(du)ds < \infty.$$

If $\{S(t) : t \geq 0\}$ is a pseudo-contraction C_0-semigroup satisfying $||S(t)|| \leq e^{\alpha t}$, for some $\alpha \geq 0$ then

$$E \sup_{0\leq t\leq T}\left|\left|\int_0^t \int_Z S(t-s)\Phi(s,u)\tilde{N}(ds,du)\right|\right|_X^p \leq C_p E\int_0^T \int_Z ||\Phi(s,u)||_X^p \nu(du)ds,$$

where $C_p > 0$ is a constant dependent on p, α, T.

Proof See Marinelli, Prévôt, and Röckner [1, Proposition 3.3]. □

Let $\mathcal{O}_c = \{y \in Y - \{0\} : ||y||_Y < c\}$ and $M_\nu^p([0,T] \times \mathcal{O}_c; X), p \geq 2$ denote the space of X-valued mappings $J(t,y)$, progressively measurable with respect to $\mathcal{F}_t$ such that

$$E\int_0^T \int_{\{||y||_Y<c\}} ||J(t,y)||_X^p \nu(dy)dt < \infty. \tag{2.79}$$

We are interested in the stochastic convolution

$$Z(t) = \int_0^t \int_{\{||y||_Y<c\}} S(t-s)J(s,y)\tilde{N}(ds,dy),$$

defined for any fixed $t \in [0,T]$. In particular, we establish below a special case of Burkholder type of inequality for stochastic convolutions driven by the compensator $\tilde{N}(\cdot,\cdot)$ of the Poisson random measure $N(\cdot,\cdot)$.

Lemma 2.13 Suppose $J \in M_\nu^2([0,T] \times \mathcal{O}_c; X) \cap M_\nu^4([0,T] \times \mathcal{O}_c; X)$; then for any $t \in [0,T]$

$$E \sup_{0 \le s \le t} ||Z(s)||_X^2 \le C \Bigg\{ E\left(\int_0^t \int_{\{||y||_Y < c\}} ||J(s,y)||_X^2 \nu(dy) ds \right) + E\left(\int_0^t \int_{\{||y||_Y < c\}} ||J(s,y)||_X^4 \nu(dy) ds \right)^{1/2} \Bigg\} \tag{2.80}$$

for some number $C = C(T) > 0$. In particular, if $\alpha = 0$ then $C(T)$ can be chosen independent of T.

Proof See Luo and Liu [1]. □

Lemma 2.14 Let $p \ge 1$ and $\{M(t), t \ge 0\}$ be a real-valued square integrable c*à*dl*à*g martingale with $M(0) = 0$. Then, for any $T \ge 0$, there exists a positive constant C_p such that

$$C_p^{-1} E[M,M]_T^{p/2} \le E\left[\sup_{t \in [0,T]} |M(t)|^p \right] \le C_p E[M,M]_T^{p/2},$$

where $[M,M]_t$ is the quadratic variation process of $\{M(t), t \ge 0\}$.

Proof See Kallenberg [1, Theorem 26.12]. □

2.9 Bounded Stochastic Integral Contractors

The purpose of this section is to introduce the concept of a random (stochastic) contractor and motivate its applicability to Itô stochastic integral equations. The existence, uniqueness, measurability, stochastic stability, and approximation of random solutions of such equations may be established by using the notion of random contractors.

Let X and Y be separable Banach spaces, and let $U(\omega) : \Omega \times D(U) \to Y$ be a nonlinear random operator, where $D(U)$ denotes the domain of $U(\omega)$. Let $\Gamma(x, \cdot) : \Omega \times Y \to X$ be a bounded linear random contractor associated with $x \in X$. First, the definition of a random contractor of a nonlinear random operator is given.

The random operator $U(\omega)$ has a random contractor $\Gamma(x, \cdot)$ at $x \in D(U) \subset X$ if there exists a positive random variable $q(\omega)$, $0 < q(\omega) < 1$, *P-a.s.*, and a constant $\eta > 0$ such that

$$||U(\omega)[x + \Gamma(x,\omega)y] - U(\omega)x - y|| \le q(\omega)||y||, \quad P\text{-}a.s.,$$

where $y \in Y$ and $||y|| \le \eta$.

The random contractor is said to have a bounded random contractor $\Gamma(x,\cdot)$ if there exists a positive random variable $B(\omega)$ such that $||\Gamma(x,\cdot)|| \leq B(\omega)$, P-$a.s.$ for all $x \in D(U)$.

Example 2.2 The inverse of the Fréchet derivative of the random operator, that is, $[U'(\omega)(x)]^{-1}$ is a random contractor. More generally, an inverse derivative of $U(\omega)$ is a random contractor.

Example 2.3 Let $P(\omega) : \Omega \times X \to Y$ be a contraction mapping, that is, there exists a positive random variable $q(\omega)$, $0 < q(\omega) < 1$, P-$a.s.$ such that

$$||U(\omega)x_1 - U(\omega)x_2|| \leq q(\omega)||x_1 - x_2||, \quad P\text{-}a.s \quad \forall x_1, x_2 \in X.$$

Then the random operator $U(\omega)$ of the term $U(\omega)x = x - P(\omega)x$, $x \in X$ has $\Gamma(x,\omega) \equiv I$ (identity operator) as a contractor. In fact, we have

$$\begin{aligned} ||U(\omega)[x+\Gamma(x,\omega)y] - U(\omega)x - y|| &= ||x+\Gamma(x,\omega)y - P(\omega)[x+\Gamma(x,\omega)y] \\ &\quad -x + P(\omega)x - y|| \\ &= ||P(\omega)x - P(\omega)[x+y]|| \\ &\leq q(\omega)||y|| \quad \forall y \in Y. \end{aligned}$$

The concept of random contractor introduced above is useful in studying the existence of a solution of random operator equations of the first kind

$$U(\omega)x = \xi(\omega) \tag{2.81}$$

where $U(\omega) : \Omega \times X \to Y$, $x \in X$ and $\xi \in Y$.

Consider now the random operator equations of the second kind

$$x - U(\omega)x = \xi(\omega). \tag{2.82}$$

Or, more specifically, consider the following Itô stochastic integral equation

$$x(t,\omega) + \int_0^t F(t,x(t,\omega))dt + \int_0^t G(t,x(t,\omega))dw(t,\omega) = \xi(t,\omega), \tag{2.83}$$

where $F : [0,T] \times X \to X$, $G : [0,T] \times X \to L(Y,X)$ and $\xi(t,\omega)$ is a given stochastic process. For convenience (as in the literature), we suppress the random parameter ω henceforth. Define the random integral operators

$$[\tilde{F}x](t) = \int_0^t F(t,x(t))dt$$

and

$$[\tilde{G}x](t) = \int_0^t g(t,x(t))dw(t).$$

With these definitions, the equation (2.83) can be written in the form of (2.81) with $U = I + \tilde{F} + \tilde{G}$. Consider

$$\begin{aligned}
||U(x+\tilde{\Gamma}(x)y) - Ux - y|| &= ||x+\tilde{\Gamma}(x)y - \tilde{F}(x+\tilde{\Gamma}(x)y) + \tilde{G}(x+\tilde{\Gamma}(x)y) \\
&\quad -x - \tilde{F}(x) - \tilde{G}(x) - y|| \\
&= ||\tilde{F}(x+\tilde{\Gamma}(x)y) - \tilde{F}(x) - \tilde{G}(x+\tilde{\Gamma}(x)y) \\
&\quad -\tilde{G}(x) - (I - \tilde{\Gamma}(x))y||.
\end{aligned}$$

But

$$[\tilde{\Gamma}(x)y](t) = y(t) + \int_0^t \Gamma_1(s,x(s))y(s)ds + \int_0^t \Gamma_2(s,x(s))y(s)dw(s).$$

Defining the integral operators

$$[\tilde{\Gamma}_1 x](t) = \int_0^t \Gamma_1(s,x(s))y(s)ds$$

and

$$[\tilde{\Gamma}_2 x](t) = \int_0^t \Gamma_2(s,x(s))y(s)dw(s),$$

we have

$$\begin{aligned}
&||U(x+\tilde{\Gamma}(x)y) - Ux - y|| \\
&\le ||\tilde{F}(x+\tilde{\Gamma}(x)y) - \tilde{F}(x) - \tilde{\Gamma}_1(x)y|| \\
&+ ||\tilde{G}(x+\tilde{\Gamma}(x)y) - \tilde{G}(x) - \tilde{\Gamma}_2(x)y|| \\
&= \sup_t || \int_0^t [F(s,x(s)+y(s) + \int_0^s \Gamma_1(\tau,x(\tau))y(\tau)d\tau \\
&\quad + \int_0^s \Gamma_2(\tau,x(\tau))y(\tau)dw(\tau)) \\
&\quad -F(s,x(s)) - \Gamma_1(s,x(s))y(s)]ds|| \\
&+ \sup_t || \int_0^t [G(s,x(s)+y(s) + \int_0^s \Gamma_1(\tau,x(\tau))y(\tau)d\tau \\
&\quad + \int_0^s \Gamma_2(\tau,x(\tau))y(\tau)dw(\tau)) \\
&\quad -G(s,x(s)) - \Gamma_2(s,x(s))y(s)]dw(s)||
\end{aligned} \tag{2.84}$$

We now impose the following conditions on F and G:

There exist positive constants K_1 and K_2 such that the following inequalities hold P-$a.s.$:

$$\begin{aligned}
&||F(t,x(t)+y(t)+\int_0^t \Gamma_1(\tau,x(\tau))y(\tau)d\tau \\
&+\int_0^t \Gamma_2(\tau,x(\tau))y(\tau)dw(\tau)) \\
&-F(t,x(t))-\Gamma_1(t,x(t))y(t)|| \leq K_1||y(t)||,
\end{aligned} \tag{2.85}$$

$$\begin{aligned}
&||G(t,x(t)+y(t)+\int_0^t \Gamma_1(\tau,x(\tau))y(\tau)d\tau \\
&+\int_0^t \Gamma_2(\tau,x(\tau))y(\tau)dw(\tau)) \\
&-G(t,x(t))-\Gamma_2(t,x(t))y(t)|| \leq K_2||y(t)||,
\end{aligned} \tag{2.86}$$

for all $x, y \in C$ (space of continuous functions). The vector (K_1, K_2) is called the vector of contractor constants.

Under the conditions (2.85) and (2.86), (2.84) gives

$$||U(x+\tilde{\Gamma}(x)y)-Ux-y|| \leq K'||y||, \quad K'>0.$$

Therefore $\tilde{\Gamma}(x)$ is a contractor for U. Moreover, if there exists Γ_1 and Γ_2 that satisfy the conditions (2.85) and (2.86), then F and G are said to have bounded stochastic integral contractor.

Remark 2.7 Conditions (2.85) and (2.86) are weaker than the usual Lipschitz condition. However, if $\Gamma_1 = \Gamma_2 \equiv 0$, then these conditions reduce to the Lipschitz condition.

2.9.1 Volterra Series

Consider the special case of $G=0$. Then the bounded integral contractor is given by

$$[\tilde{\Gamma}_1(x)y](t) = y(t)+\int_0^t \Gamma_1(s,x(s))y(s)ds. \tag{2.87}$$

On the other hand, we have

$$U'(x)z(t) = (I+\tilde{F})'(x)z(t) = z(t)+\int_0^t F_x(s,x(s))z(s)ds,$$

if F_x exists, and

$$[U'(x)]^{-1}y(t) = [\tilde{\Gamma}_1(x)]^{-1}y(t) = y(t) + \int_0^t \sum_{n=1}^{\infty} F_x^n(s,x(s))y(s)ds. \tag{2.88}$$

In view of Example 2.2, we obtain from (2.87) and (2.88) that

$$\Gamma_1(t,x(t)) = \sum_{n=1}^{\infty} F_x^n(t,x(t)). \tag{2.89}$$

It is interesting to observe that Γ_1, in fact, exists and is precisely the Volterra series given in (2.89).

http://www.springer.com/978-3-319-27328-0

Imaging Convection and Magnetism in the Sun
Hanasoge, S.
2015, XVII, 131 p. 48 illus., 22 illus. in color., Softcover
ISBN: 978-3-319-27328-0