# Invariant Measures for Stochastic Partial Differential Equations

Lectures by Szymon Peszat Notes taken by James Inglis

November 11, 2008

# Contents

1	Introduction					
2	Line	ar Equations with Gaussian Noise	7			
	2.1	$C_0$ -semigroups and well-posedness of deterministic linear problems	7			
	2.2	Nuclear and Hilbert-Schmidt operators	9			
	2.3	Wiener processes in Hilbert spaces	10			
		2.3.1 $Q$ -Wiener processes	10			
		2.3.2 Reproducing kernel Hilbert space	11			
		2.3.3 Cylindrical Wiener processes	14			
	2.4	Solving the SPDE	15			
	2.5	Invariant measures	17			
3	Orns	stein-Uhlenbeck Processes Driven By Lévy Processes	21			
-	3.1	Martingales on Hilbert spaces	21			
	3.2	The stochastic integral with respect to a square integrable martingale	25			
	3.3	Lévy processes and the Lévy-Khinchin formula	27			
	3.4	Existence of an invariant measure for OU process driven by a Lévy process	32			
4	Convergence to the Invariant Measure					
	4.1	Exponential mixing	41			
	4.2	Existence and exponential mixing: Regular case	42			
	4.3	An example: the Heath-Jarrow-Morton model	47			
5 Appendix						
•	5.1	Gaussian measures in Hilbert spaces	51			
	5.2	Lévy-Khinchin decomposition	52			
	5.3	Krylov-Bogoliubov Theorem	53			
	5.4	Yosida approximations	53			
	5.5	Itô formula for Hilbert space valued semimartingales	54			

# 1 Introduction

These notes are devoted to the study of invariant measures for stochastic evolution equations in infinite dimensions. Our main reference will be *Stochastic Partial Differential Equations with Lévy Noise* (Peszat and Zabczyk, 2007) [11], which includes most (but not all) the results contained here, and should be referred to for the details. However, our focus here will be more specific, in that we will not worry too much about existence and regularity of solutions, but try to describe the invariant measure in situations where this is known. The invariant measure can be thought of as describing the long-term behaviour of a dynamical system, and this is one of the reasons that the results here and in [11] have many important applications in, for example, lattice systems and financial mathematics.

We start by considering linear evolution equations driven by a Wiener process in a separable Hilbert space, covering also the background material needed to describe them rigourously. Then in Chapter 2 we try and generalise the results to the situation where the equations are driven by a general Lévy process L on a Hilbert space (which does not necessarily have continuous trajectories). In Chapter 3, in the case where L is a square integrable Lévy process, we describe conditions under which we have exponential convergence of the semigroup to the invariant measure. We finally apply the results obtained to an important model from mathematical finance: the so-called Heath-Jarrow-Morton model.

# 1 Introduction

### 2 Linear Equations with Gaussian Noise

In this chapter we will be interested in the following family of stochastic differential equations:

$$dX(t) = AX(t)dt + BdW(t)$$
(2.1)

where

- $(A, \mathcal{D}(A))$  is the generator of a  $C_0$  semigroup on a Hilbert space H;
- W(t) is a cylindrical Wiener process on a Hilbert space U;
- B is a bounded linear operator from U to H i.e  $B \in L(U, H)$ .

In the next few sections we will make this description precise.

### 2.1 $C_0$ -semigroups and well-posedness of deterministic linear problems

Deterministic linear evolution equations can often be formulated as an evolution equation in a Hilbert space H:

$$\frac{dX}{dt} = A_0 X, \qquad t \ge 0, \qquad X(0) = x \in \mathcal{D}(A_0) \subset H, \tag{2.2}$$

with  $A_0 : \mathcal{D}(A_0) \to H$  an operator (in general unbounded) defined on a dense linear subspace  $\mathcal{D}(A_0)$  of H. In (2.2) dX/dt stands for the strong derivative of X(t) i.e.

$$\frac{dX}{dt} = \lim_{h \to 0} \frac{X(t+h) - X(t)}{h}.$$

Problem (2.2) is called the initial value problem or the *Cauchy problem* relative to the operator  $A_0$ .

**Definition 2.1.1.** We say that the Cauchy problem (2.2) is well-posed if:

- (i) for arbitrary  $x \in \mathcal{D}(A_0)$  there exists exactly one strongly differentiable function  $X(t, x), t \geq 0$  satisfying (2.2) for all  $t \geq 0$ ,
- (ii) for  $\{x_n\} \subset \mathcal{D}(A_0)$  such that  $\lim_{n\to\infty} x_n = 0$ , we have

$$\lim_{n \to \infty} X(t, x_n) = 0,$$

for all  $t \geq 0$ .

If the limit in the above definition is uniform in t on compact subsets of  $[0, \infty)$  we say that the Cauchy problem (2.2) is *uniformly well-posed*. From now on we assume that the Cauchy problem is uniformly well-posed, and define operators  $S(t) : \mathcal{D}(A_0) \to H$  by the formula

$$S(t)x = X(t, x), \quad \forall x \in \mathcal{D}(A_0), \forall t \ge 0.$$

By the density of  $\mathcal{D}(A_0)$  in H and the well-posedness of the problem, for all  $t \geq 0$  the linear operator S(t) can be uniquely extended to a bounded linear operator on the whole of H, which we still denote by S(t). We have clearly that

$$S(0) = I.$$
 (2.3)

Moreover, by the uniqueness

$$S(t+s) = S(t)S(s), \qquad \forall t, s \ge 0.$$
(2.4)

Finally, by the uniform boundedness theorem, it follows that

$$S(\cdot)(x)$$
 is continuous in  $[0,\infty)$  for all  $x \in H$ . (2.5)

In this way we are led directly from the study of the uniformly well-posed Cauchy problem to the family  $(S(t), t \ge 0)$  of bounded linear operators on H satisfying (2.3), (2.4) and (2.5). We say that a family  $(S(t), t \ge 0)$  of bounded linear operators on Hsatisfying (2.3), (2.4) and (2.5) is a  $C_0$ -semigroup of linear operators. So the concept of a  $C_0$ -semigroup is in a sense equivalent to that of a uniformly well-posed Cauchy problem.

**Definition 2.1.2.** The generator of a  $C_0$ -semigroup  $S(\cdot)$  is a linear operator  $(A, \mathcal{D}(A))$ on H such that

$$\mathcal{D}(A) = \left\{ x \in H \colon \lim_{t \to 0} \frac{S(t)x - x}{t} \text{ exists} \right\},\$$

and for  $x \in \mathcal{D}(A)$ ,

$$Ax = \lim_{t \to 0} \frac{S(t)x - x}{t}$$

It is easy to see that A is an extension of  $A_0$  and moreover that the problem

$$\frac{dX}{dt} = AX, \qquad t \ge 0, \qquad X(0) = x \in H, \tag{2.6}$$

is also uniformly well-posed with the same associated semigroup  $S(\cdot)$ . Thus, in our investigations we will only consider the Cauchy problem (2.6) with A being the generator of a  $C_0$ -semigroup.

We can also consider the equation with added drift term i.e

$$\frac{dX}{dt} = AX + f, \qquad X(0) = x \in H, \tag{2.7}$$

where  $f: [0, \infty) \to H$  is a bounded measurable function. By the variation of constants formula we have

$$X(t) = S(t)x + \int_0^t S(t-s)f(s)ds,$$

where S(t) solves (2.6) above. We say that a process which satisfies this integral equation is a *mild solution* to (2.7). Moreover, if

$$< X(t), u >_{H} = < x, u >_{H} + \int_{0}^{t} \{ < X(s), A^{*}u >_{H} + < f(s), u >_{H} \} ds, \quad \forall u \in \mathcal{D}(A^{*}),$$

then we say that X(t) is a *weak solution* to (2.7). By direct calculation it can be shown that every mild solution is a weak solution.

### 2.2 Nuclear and Hilbert-Schmidt operators

Let E, G be Banach spaces and let L(E, G) be the Banach space of all bounded linear operators from E into G with the usual supremum norm. Denote by  $E^*$  and  $G^*$  the dual spaces of E and G respectively.

**Definition 2.2.1.**  $T \in L(E, G)$  is a nuclear operator if there exist two sequences  $\{a_j\} \subset G$  and  $\{\phi_j\} \subset E^*$  such that

$$\sum_{j=1}^{\infty} \|a_j\|_G \|\phi_j\|_{E^*} < \infty,$$

and such that T has the representation

$$Tx = \sum_{j=1}^{\infty} a_j \phi_j(x), \qquad \forall x \in E.$$

The space  $L_1(E,G)$  of all nuclear operators from E into G endowed with the norm

$$||T||_{L_1(E,G)} := \inf\left\{\sum_{j=1}^{\infty} ||a_j||_{-G} |\phi_j||_{E^*} : Tx = \sum_{j=1}^{\infty} a_j \phi_j(x)\right\}$$

is a Banach space. This space is interesting for us because of the following facts (see Appendix C of [12]). Let H be a separable Hilbert space, and let  $\{e_k\}$  be a complete orthonormal system in H. For  $T \in L_1(H, H) = L_1(H)$  define the trace of T to be

$$\operatorname{Tr} T = \sum_{j=1}^{\infty} \langle Te_j, e_j \rangle_H.$$

**Proposition 2.2.1.** For  $T \in L_1(H)$ , Tr T is well defined and independent of the choice of orthonormal basis  $\{e_j\}$ .

**Proposition 2.2.2.** A non-negative operator  $T \in L(H)$  is nuclear if and only if for some orthonormal basis  $\{e_j\}$  on H

Tr 
$$T = \sum_{j=1}^{\infty} \langle Te_j, e_j \rangle_H < \infty.$$

Moreover, in this case  $||T||_{L_1(H)} = \text{Tr } T$ .

**Definition 2.2.2.** Let H, F be two separable Hilbert spaces with complete orthonormal basis  $\{e_i\} \subset H$ . We say that  $T \in L(H, F)$  is Hilbert - Schmidt if

$$\sum_{j=1}^\infty |Te_j|_F^2 < \infty.$$

The space  $L_{HS}(H, F)$  of all Hilbert-Schmidt operators from H into F is a separable Hilbert space with the scalar product

$$< S, T >_{HS} = \sum_{j=1}^{\infty} < Se_j, Te_j >_F.$$

Now suppose  $(\Omega, \mathcal{F}, \mathbb{P})$  is a probability space and H is a separable Hilbert space. Let  $X : \Omega \to H$  be a square integrable H-valued random variable such that  $\mathbb{E}(X) = 0$ . Then we define the *covariance operator* of the random variable X to be the operator Q given by

$$\langle Qx, y \rangle_H = \mathbb{E}(\langle X, x \rangle_H \langle X, y \rangle_H), \quad \forall x, y \in H.$$

Then

$$|Qx|_{H}^{2} = \langle Qx, Qx \rangle_{H} = \mathbb{E}\left(\langle X, x \rangle_{H} \langle X, Qx \rangle_{H}\right)$$
  
$$\leq |x|_{H}|Qx|_{H}\mathbb{E}|X|_{H}^{2}, \qquad (2.8)$$

by Cauchy-Schwarz, so that  $Q \in L(H)$ . Moreover, if  $\{e_j\}$  is a complete orthonormal sequence for H we have

Tr 
$$Q = \sum_{j=1}^{\infty} \langle Qe_j, e_j \rangle_H$$
  
=  $\sum_{j=1}^{\infty} \int_{\Omega} |\langle X(\omega), e_j \rangle_H |^2 \mathbb{P}(d\omega)$   
=  $\mathbb{E}|X|_H^2 < \infty.$ 

Therefore by Proposition 2.2.2 we see that  $Q \in L_1(H)$ .

### 2.3 Wiener processes in Hilbert spaces

We first define Q-Wiener processes before constructing more general *cylindrical* Wiener processes.

### 2.3.1 *Q*-Wiener processes

Let U be a separable Hilbert space and let  $Q \in L(U)$  be non-neqative and such that Tr  $Q < \infty$ . Then there exists a complete orthonormal system  $\{e_k\}$  in U and a bounded sequence of non-negative real numbers  $\gamma_k$  such that

$$Qe_k = \gamma_k e_k, \qquad k = 1, 2, \dots$$

**Definition 2.3.1.** A U-valued stochastic process  $W(t), t \ge 0$  is called a Q-Wiener process (or just a Wiener process) if

- (i) W(0) = 0,
- (ii) W has continuous trajectories,
- (iii) W has independent increments,
- (iv)  $\mathcal{L}(W(t) W(s)) = \mathcal{N}(0, (t-s)Q), \forall t \ge s \ge 0.$

With the help of the Kolmogorov extension theorem (see for example Theorem 3.7 of [11]), it is fairly straightforward to show the existence of a Q-Wiener process for any symmetric non-negative operator Q on U such that Tr  $Q < \infty$ .

**Proposition 2.3.1.** Let W be a Q-Wiener process with Tr  $Q < \infty$ . Then W is a Gaussian process on U,  $\mathbb{E}(W(t)) = 0$  and Cov(W(t)) = tQ. Moreover, for  $t \ge 0$ ,

$$W(t) = \sum_{j=1}^{\infty} \sqrt{\gamma_j} \beta_j(t) e_j$$

where

$$\beta_j(t) = \frac{1}{\sqrt{\gamma_j}} < W(t), e_j >_U, \qquad j = 1, 2, \dots$$

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

It can be be shown that the series in Proposition 2.3.1 is in fact uniformly convergent on any [0, T] P-a.s. (see Theorem 4.3 of [12]).

### 2.3.2 Reproducing kernel Hilbert space

In view of the above, one can expect the covariance operator of the noise to play a fundamental role in the study of stochastic evolution equations. However, it turns out that it is much more convenient to study the reproducing kernel Hilbert space of the noise; see definition 2.3.2 below. Unlike the covariance operator, the RKHS is independent of the space on which the noise is considered.

Let Z be a square integrable random variable with mean zero in a Hilbert space U, and let  $\langle \cdot, \cdot \rangle : U^* \times U \to \mathbb{R}$  be the duality form. Let Q be the covariance operator of Z. Since Q is a nuclear self-adjoint operator on U, there is an orthonormal basis  $\{e_j\}$ of U consisting of eigenvectors of Q. Then  $Qe_j = \gamma_j e_j, j \in \mathbb{N}$ . Since Q is non-negative definite, we have  $\gamma_j \geq 0$ . The square root of Q is given by

$$Q^{1/2}x = \sum_{j} \langle x, e_{j} \rangle_{U} \gamma_{j}^{1/2}e_{j},$$

for  $x \in U$ . Generally Q and  $Q^{1/2}$  are not injective. However, we will denote by  $Q^{-1/2}$  the *pseudo-inverse operator*, which is defined by

$$Q^{-1/2}y = x$$
 if  $Q^{1/2}x = y$ , and  $|x|_U = \inf\{|y|_U : Q^{1/2}y = x\}$ .

**Definition 2.3.2.** Let Z be a square integrable zero-mean random variable taking values in a Hilbert space U, and let Q be the covariance of Z. Then  $U_0 = Q^{1/2}(U)$ , equipped with the inner product  $\langle x, y \rangle_{U_0} = \langle Q^{-1/2}x, Q^{-1/2}y \rangle_U$ , is called the reproducing kernel Hilbert space (RKHS) of Z.

If Z is as above and  $U_0$  is its RKHS, then since Q is nuclear,  $Q^{1/2}$  is Hilbert-Schimdt. Consequently the embedding  $U_0 \hookrightarrow U$  is Hilbert-Schmidt; that is for an arbitrary orthonormal basis  $\{f_j\}$  of  $U_0$  one has  $\sum_j |f_j|_U^2 < \infty$ .

The RKHS of a random variable Z is independent of the space upon which the random element Z is considered. More precisely, let  $U_0$  and  $\tilde{U}_0$  be the RKHSs of Z considered as a random variable on U and  $\tilde{U}$  respectively. Then  $U_0 = \tilde{U}_0$ . We also note that the concept of the reproducing kernel can be extended to a non-square integrable random variable Z taking values in a Hilbert space U, provided there is a bigger  $\tilde{U} \leftrightarrow U$  such that  $\mathbb{E}|Z|^2_{\tilde{U}} < \infty$ .

A useful result which characterises the RKHS of Z is the following (see Proposition 7.1 of [11]).

**Proposition 2.3.2.** Let  $(U_0, < \cdot, \cdot >_{U_0})$  be a Hilbert space continuously embedded into U. Then the following are equivalent:

- (i)  $U_0 = Q^{1/2}(U)$  and  $\langle x, y \rangle_{U_0} = \langle Q^{-1/2}x, Q^{-1/2}y \rangle_U$  for all  $x, y \in U_0$  (i.e  $U_0$  is the RKHS of Z).
- (ii) For any orthonormal basis  $\{f_j\}$  of  $U_0$ ,

$$\mathbb{E} < x, Z >_{U_0} < y, Z >_{U_0} = \sum_j < x, f_j >_{U_0} < y, f_j >_{U_0}, \qquad \forall x, y \in U.$$

(iii) For any orthonormal basis  $\{f_i\}$  of  $U_0$ ,

$$\mathbb{E}\langle x, Z \rangle \langle y, Z \rangle = \sum_{j} \langle x, f_j \rangle \langle y, f_j \rangle, \qquad \forall \ x, y \in U^*.$$

In the two propositions below,  $U_0$  and U are Hilbert spaces such that  $U_0$  is densely embedded into U. Then under the identification of  $U_0$  with  $U_0^*$ , we have that  $U^* \hookrightarrow U_0^* = U_0 \hookrightarrow U$  and we can treat  $U^*$  as a subspace of  $U_0$ . Recalling that  $\langle \cdot, \cdot \rangle$  is the duality form on  $U^* \times U$ , we clearly have that  $\langle x, y \rangle = \langle x, y \rangle_{U_0}$  for all  $x \in U^*, y \in U_0$ .

**Proposition 2.3.3.** Let Z be a square integrable zero-mean random variable in U. Assume that  $\mathbb{E}\langle x, Z \rangle \langle y, Z \rangle = \langle x, y \rangle_{U_0}$  for all  $x, y \in U^*$ . Then  $U_0$  is the RKHS of Z.

*Proof.* Let  $\{f_j\}$  be an arbitrary orthonormal basis of  $U_0$ . Then

$$\langle x, y \rangle_{U_0} = \sum_j \langle x, f_j \rangle_{U_0} \langle y, f_j \rangle_{U_0}$$

Since for  $x \in U^*$ ,  $\langle x, f_j \rangle_{U_0} = \langle x, f_j \rangle$  we have that

$$\mathbb{E}\langle x, Z \rangle \langle y, Z \rangle = \langle x, y \rangle_{U_0} = \sum_j \langle x, f_j \rangle \langle y, f_j \rangle,$$

which by Proposition 2.3.2 completes the proof.

**Proposition 2.3.4.** Assume that the embedding  $U_0 \hookrightarrow U$  is Hilbert-Schmidt. Let  $Z : U_0 \to L^2(\Omega, \mathcal{F}, \mathbb{P})$  be a linear operator such that  $\mathbb{E}(Zx)^2 = c|x|_{U_0}^2$  and  $\mathbb{E}Zx = 0$  for  $x \in U_0$ . Then there is a unique square integrable zero-mean random variable  $\tilde{Z}$  in U such that

$$Zx = \langle x, \tilde{Z} \rangle, \qquad \forall \ x \in U^*.$$

Moreover,  $U_0$  is the RKHS of  $\tilde{Z}$ .

*Proof.* Let  $\{f_j\}$  be an orthonormal basis of  $U_0$ . We assume that  $\{f_j\} \subset U^*$ . Such a basis exists since  $U^*$  is dense in  $U_0$ . Then

$$\mathbb{E} \left| \sum_{j=n}^{n+m} (Zf_j) f_j \right|_U^2 = c \sum_{j,k=n}^{n+m} (\mathbb{E} Zf_j Zf_k) < f_j, f_k >_U = \sum_{j,k=n}^{n+m} < f_j, f_k >_{U_0} < f_j, f_k >_U \\ = \sum_{j=n}^{n+m} |f_j|_U^2 \to 0$$

as  $n, m \to \infty$ . Thus the series  $\sum_j (Zf_j)f_j$  converges in  $L^2(\Omega, \mathcal{F}, \mathbb{P}; U)$ . We will show that its limit  $\tilde{Z}$  has the desired properties. To this end, note that for any  $x \in U^*$ ,

$$\langle x, \tilde{Z} \rangle = \sum_{j} (Zf_j) \langle x, f_j \rangle = \sum_{j} (Zf_j) \langle x, f_j \rangle_{U_0}$$

Since Z is a continuous linear operator on  $U_0$ ,

$$\sum_{j} (Zf_j) < x, f_j >_{U_0} = Z\left(\sum_{j} < f_j, x >_{U_0} f_j\right) = Zx$$

and hence by use of Proposition 2.3.3 the result follows.

We will identify  $\tilde{Z}$  with Z, and write Z instead of  $\tilde{Z}$ .

### 2.3.3 Cylindrical Wiener processes

Now let  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$  be a filtered probability space and let U be a separable Hilbert space as usual.

**Definition 2.3.3.** A cylindrical Wiener process (adapted to  $(\mathcal{F}_t)$ ) on U is a linear (in the second variable) mapping  $W : [0, \infty) \times U \to L^2(\Omega, \mathcal{F}, \mathbb{P})$  satisfying the following conditions:

- (i) for all  $t \ge 0$  and  $x \in U$ ,  $\mathbb{E}|W(t,x)|^2 = t|x|_U^2$ ,
- (ii) for each  $x \in U$ ,  $(W(t, x), t \ge 0)$  is a real-valued  $(\mathcal{F}_t)$ -adapted Wiener process.

**Lemma 2.3.5.** If W is a cylindrical Wiener process then, for all  $t \ge s \ge 0$  and  $x, y \in U$ ,  $\mathbb{E}W(t, x)W(s, y) = (t \land s) < x, y >_U$ .

*Proof.* Assume that  $t \ge s \ge 0$ . Then

$$\begin{split} \mathbb{E}W(t,x)W(s,y) &= \mathbb{E}\mathbb{E}\left((W(t,x) - W(s,y))W(s,y)|\mathcal{F}_s\right) + \mathbb{E}W(s,x)W(s,y) \\ &= \mathbb{E}W(s,x)W(s,y) \\ &= \frac{1}{4}\mathbb{E}\left((W(s,x) + W(s,y))^2 - (W(s,x) - W(s,y))^2\right) \\ &= \frac{1}{4}\mathbb{E}\left(W(s,x+y)^2 - W(s,x-y)^2\right) \\ &= \frac{1}{4}\left(|x+y|_U^2 - |x-y|_U^2\right) = s < x, y >_U. \end{split}$$

Now let  $U_1$  be a Hilbert space such that the embedding  $U \hookrightarrow U_1$  is dense and Hilbert-Schmidt. We identify  $U_1^*$  with a subspace of U and denote by  $\langle \cdot, \cdot \rangle$  the bilinear form on  $U_1^* \times U_1$  as above. As a simple consequence of Propositions 2.3.3 and 2.3.4 we have the following result.

**Theorem 2.3.6.** (i) If W is a cylindrical Wiener process on U then there is a  $U_1$  valued Q-Wiener process, which we will denote also by W such that

$$\langle x, W(t) \rangle = W(t, x), \qquad t \ge 0, x \in U^*.$$
(2.9)

Moreover, the RKHS of W is equal to U.

(ii) Conversely, if W is a Wiener process in  $U_1$  with RKHS equal to U then (2.9) defines a cylindrical Wiener process on U.

Assume that W is a cylindrical Wiener process in U. Let  $\{e_j\}$  be an orthonormal basis of U. Let  $W_j(t) := W(t, e_j)$ . Then  $(W_j)$  is a sequence of independent standard realvalued Wiener processes. Let  $U_1$  be a Hilbert space such that the embedding  $U \hookrightarrow U_1$ is Hilbert-Schmidt. Then

$$W(t) = \sum_{j} W_j(t)e_j, \qquad t \ge 0,$$

where the series converges in  $L^2(\Omega, \mathcal{F}, \mathbb{P}; U_1)$ .

**Remark 2.3.1.** The concept of a cylindrical Wiener process is closely related to that of space-time white noise. Loosely speaking, the latter is the time derivative of a cylindrical Wiener process.

Note that a cylindrical Wiener process W(t) on U has covariance equal to the identity operator (which is certainly not trace class), whereas every Q-Wiener process on Uhas covariance Q such that Tr  $Q < \infty$ . A cylindrical Wiener process is an example of square integrable cylindrical martingale with independent stationary increments, so that stochastic integration can be defined with respect to W(t) according to section 3.2 below. For more details specifically for the case of a Wiener process, see [12] Chapter 4. A cylindrical Wiener process is not uniquely determined, but it can be shown that the class of integrands and the spaces of stochastically integrable processes are independent of the space  $U_1$  chosen.

### 2.4 Solving the SPDE

Let H and U be separable Hilbert spaces. Let  $\{e_j\}$  be an orthonormal basis for U. Recall we are interested in an H-valued process X(t) which solves the equation (2.1)

$$dX(t) = AX(t)dt + BdW(t),$$

where  $(A, \mathcal{D}(A))$  is the generator of a  $C_0$ -semigroup S(t) on H,  $W(t) = \sum_{j=1}^{\infty} W_j(t) e_j$ is a cylindrical Wiener process on U and  $B \in L(U, H)$ . What should our solution look like?

In view of the previous sections, we will define our solution to be the *H*-valued process X(t) which satisfies

$$X(t) = S(t)X(0) + \sum_{j=1}^{\infty} \int_{0}^{t} S(t-s)Be_{j}dW_{j}(s).$$

We say X(t) is a *mild solution* to (2.1). However, to make this rigourous, we must say something about convergence. We would like to have convergence in some sense in H. We know that

$$\left\{\int_0^t Be_j dW_j(s) : j \in \mathbb{N}\right\}$$

forms a sequence of independent Gaussian random variables on H, since we are integrating something continuous with respect to independent 1-dimensional Brownian motions. We can therefore make use of the Itô-Nisio theorem:

**Theorem 2.4.1** (Itô-Nisio). Let  $\{X_k\}$  be a sequence of independent random variables with values in a Banach space E. Then the following are equivalent:

- (i)  $\sum_{k=1}^{\infty} X_k$  converges  $\mathbb{P}$ -a.s.
- (ii)  $\sum_{k=1}^{\infty} X_k$  converges in probability.

(iii)  $\sum_{k=1}^{\infty} X_k$  converges in distribution.

As convergence in  $L^2$  implies convergence in probability, by this theorem if we require

$$\sum_{j=1}^{\infty} \int_0^t S(t-s) Be_j dW_j(s)$$

to be convergent in  $L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$ , then the sum will converge  $\mathbb{P}$ -almost surely. Since

$$\sum_{j=1}^{\infty} \mathbb{E} \left( \int_0^t S(t-s) Be_j dW_j(s) \right)^2 = \sum_{j=1}^{\infty} \int_0^t |S(t-s) Be_j|_H^2 ds$$
$$= \int_0^t \sum_{j=1}^{\infty} |S(s) Be_j|_H^2 ds$$
$$= \int_0^t \|S(s) B\|_{L_{HS}(U,H)}^2 ds$$

where  $||S(s)B||_{HS}^2 = \sum_{j=1}^{\infty} |S(s)Be_j|_H^2$  is the Hilbert-Schmidt norm, this is equivalent to requiring that

$$\int_{0}^{t} \|S(s)B\|_{HS}^{2} ds < \infty.$$
(2.10)

If we add this to our assumptions, then

$$X(t) = S(t)X(0) + \sum_{j=1}^{\infty} \int_{0}^{t} S(t-s)Be_{j}dW_{j}(s)$$

is a well-defined mild solution to our equation (2.1). This is essentially how we define stochastic integration with respect to W(t), and condition 2.10 ensures that S(s)B is an integrable process. Moreover, we also have that the stochastic integral  $\int_0^t S(s)BdW(s)$ is distributed  $\mathcal{N}(0, Q_t)$  where

$$Q_t = \int_0^t S(s)BB^*S^*(s)ds.$$

All this is shown in more generality in section 3.2 below (see in particular Remark 3.2.2), and Chapter 8 of [11].

**Remark 2.4.1.** In [11] conditions for the existence and uniqueness of solutions to equations that are much more general than (2.1) are given. These general conditions, however, still entail a requirement similar to (2.10), so that the stochastic integral appearing in the definition of a mild solution is well defined.

### 2.5 Invariant measures

The following theorem of Zabczyk states an important result about the existence of an invariant measure for the solution X(t) to (2.1) as described above.

**Theorem 2.5.1** (Zabczyk). The following conditions are equivalent:

- (a)  $\sup_{t\geq 0} \operatorname{Tr} Q_t < \infty$  where  $Q_t = \int_0^t S(s)BB^*S^*(s)ds$ .
- (b) There is an invariant measure  $\mu$  for equation (2.1) i.e. if X(0) has a distribution  $\mu$  and is independent of W, then X(t) also has distribution  $\mu$  for any  $t \ge 0$ .

Moreover, if either (a) or (b) holds, then any invariant measure has the following form

$$\mu = \sigma * \mathcal{N}(0, Q)$$

where  $Q = \int_0^\infty S(s)BB^*S^*(s)ds$  and  $\sigma$  is any probability measure on H such that

$$S(t)\sigma = \sigma$$

for all  $t \geq 0$ .

**Remark 2.5.1.** In the above theorem  $\sigma * \mathcal{N}(0, Q)$  denotes the convolution of the measures  $\sigma$  and  $\mathcal{N}(0, Q)$  on H i.e.

$$\sigma * \mathcal{N}(0,Q)(A) = \sigma \times \mathcal{N}(0,Q) \left( \{ (h_1, h_2) \in H^2 : h_1 + h_2 \in A \} \right)$$

for all  $A \in \mathcal{B}(H)$ . We also denote by  $S(t)\sigma$  the measure defined by

$$S(t)\sigma(A) = \int_{H} S(t) \mathbf{1}_{A}(x)\sigma(dx), \qquad A \in \mathcal{B}(H).$$

Recall the definition of a characteristic function of a probability measure  $\mu$  on H:

$$\hat{\mu}(\lambda) = \int_{H} e^{i < \lambda, x >_{H}} \mu(dx)$$

for all  $\lambda \in H$ . Theorem 2.5.1 depends on the following well-known result of Bochner:

**Theorem 2.5.2** (Bochner). Let H be a separable Hilbert space and  $\varphi : H \to \mathbb{C}$ . Then the following are equivalent:

- (i)  $\varphi$  is the characteristic function of some probability measure  $\mu$  i.e.  $\varphi = \hat{\mu}$ .
- (ii)  $\varphi(0) = 1, \varphi$  is positive definite in the sense that

$$\sum_{i,j} \varphi(\xi_i - \xi_j) z_i \bar{z}_j \ge 0$$

for all  $\{\xi_i\} \subset H, \{z_i\} \subset \mathbb{C}$ , and  $\varphi$  is S-continuous, in the sense that  $\forall \varepsilon > 0$  there exists a nuclear operator  $S_{\varepsilon} \in L_1(H)$  such that

Re 
$$\varphi(\lambda) \ge 1 - \varepsilon$$
 whenever  $\langle S_{\varepsilon} \lambda, \lambda \rangle_H \le 1$ .

Proof of Theorem 2.5.1. We have that

$$X(t) = S(t)X(0) + \int_0^t S(t-s)BdW(s),$$

which is interpreted in the way described above. From this we can see immediately that  $\mu$  is an invariant measure for equation (2.1) if and only if

$$\mu = (S(t)\mu) * \mathcal{N}(0, Q_t)$$

for all  $t \ge 0$ , because the distribution of  $\int_0^t S(t-s)BdW(s)$  is  $\mathcal{N}(0,Q_t)$ . Taking characteristic functions we see that

$$\hat{\mu}(\lambda) = \widehat{S(t)\mu}(\lambda) \cdot \widehat{\mathcal{N}(0,Q_t)}(\lambda)$$

where, by definition of  $S(t)\mu$ , we have that

$$\widehat{S(t)\mu}(\lambda) = \int_{H} e^{i < \lambda, x >_{H}} (S(t)\mu)(dx)$$
$$= \int_{H} e^{i < \lambda, S(t)x >_{H}} \mu(dx)$$
$$= \int_{H} e^{i < S^{*}(t)\lambda, x >_{H}} \mu(dx) = \hat{\mu}(S^{*}(t)\lambda).$$

Hence  $\mu$  is an invariant measure for equation (2.1) if and only if

$$\hat{\mu}(\lambda) = \hat{\mu}(S^*(t)\lambda)e^{-\frac{1}{2} < Q_t \lambda, \lambda >_H}$$

for all  $\lambda \in H, t \geq 0$ .

(b)  $\Rightarrow$  (a): Let  $\mu$  be an invariant measure for (2.1). Then by above

$$e^{\frac{1}{2} < Q_t \lambda, \lambda >_H} \operatorname{Re} \hat{\mu}(\lambda) = \operatorname{Re} \hat{\mu}(S^*(t)\lambda)$$
$$= \operatorname{Re} \int e^{i < x, S^*(t)\lambda >_H} \mu(dx)$$
$$\leq 1$$
$$\Rightarrow < Q_t \lambda, \lambda >_H \leq 2 \log\left(\frac{1}{\operatorname{Re} \hat{\mu}(\lambda)}\right)$$

for all  $\lambda \in H, t \geq 0$ .

By Bochner's Theorem, for  $\varepsilon = \frac{1}{2}$  there exists  $S \in L_1(H)$  such that

$$\operatorname{Re}\hat{\mu}(\lambda) \geq \frac{1}{2}, \quad \forall \ \lambda \text{ such that } < S\lambda, \lambda >_H \leq 1.$$

Therefore

 $< Q_t \lambda, \lambda >_H \le 2 \log 2, \qquad \forall \ \lambda \in H \text{ such that } < S \lambda, \lambda >_H \le 1$  which yields

$$0 \le Q_t \le 2\log 2S.$$

Hence  $\sup_{t>0} \operatorname{Tr} Q_t \leq 2 \log 2 \operatorname{Tr} S < \infty$ .

(a)  $\Rightarrow$  (b): If (a) holds, then it is clear that  $Q = \int_0^\infty S(s)BB^*S^*(s)ds$  is well defined. We show that  $\mu = \mathcal{N}(0, Q)$  is invariant. Indeed, we then have that

$$\hat{\mu}(\lambda) = e^{-\frac{1}{2} < Q\lambda, \lambda >_H}$$

so that

$$\hat{\mu}(S^*(t)\lambda) = e^{-\frac{1}{2} \langle S(t)QS^*(t)\lambda,\lambda \rangle_H}.$$

Note that by the semigroup property

$$S(t)QS^{*}(t) = \int_{0}^{\infty} S(t+s)BB^{*}S^{*}(t+s)ds = \int_{t}^{\infty} S(u)BB^{*}S^{*}(u)du$$
$$= \int_{0}^{\infty} S(u)BB^{*}S^{*}(u)du - \int_{0}^{t} S(u)BB^{*}S^{*}(u)du = Q - Q_{t},$$

so that

$$\hat{\mu}(S^*(t)\lambda) = e^{-\frac{1}{2} < Q\lambda, \lambda >_H} e^{\frac{1}{2} < Q_t\lambda, \lambda >_H}.$$

This yields

$$\hat{\mu}(\lambda) = \hat{\mu}(S^*(t)\lambda)e^{-\frac{1}{2} < Q_t \lambda, \lambda >}$$

which, by above, shows that the measure  $\mu$  is indeed invariant.

For the last part of the result, using the fact that  $\hat{\mu}(\lambda) = \hat{\mu}(S^*(t)\lambda)e^{-\frac{1}{2} < Q_t \lambda, \lambda >}$  and  $e^{-\frac{1}{2} < Q_t \lambda, \lambda >} \to \widehat{\mathcal{N}(0, Q)}$  as  $t \to \infty$ , we see that

$$\hat{\mu}(S^*(t)\lambda) \to \psi(\lambda)$$

for some function  $\psi$ . If  $\psi$  is the characteristic function of some measure  $\sigma$  i.e.  $\psi(\lambda) = \hat{\sigma}(\lambda)$  then

$$\hat{\mu}(\lambda) = \hat{\sigma} \cdot \widehat{\mathcal{N}(0, Q)}$$

and  $\sigma$  is invariant for  $S(\cdot)$  since

$$\hat{\sigma}(S(s)\lambda) = \lim_{t \to \infty} \hat{\mu}(S(t+s)\lambda) = \hat{\sigma}(\lambda), \qquad \lambda \in H.$$

So we are done if we show that  $\psi$  is indeed the characteristic function of some measure. For this we use Bochner's Theorem once more. Firstly, since  $\psi(\lambda)$  is the limit of  $\hat{\mu}(S^*(t)\lambda)$  we have that  $\psi(0) = 1$  and  $\psi$  is positive definite. Therefore we just need to show it is S-continuous. For this we just note that

$$\psi(\lambda) = \hat{\mu}(\lambda)e^{\frac{1}{2} < Q\lambda, \lambda >}$$
  
$$\Rightarrow \operatorname{Re} \psi(\lambda) = e^{\frac{1}{2} < Q\lambda, \lambda >} \operatorname{Re} \hat{\mu}(\lambda)$$
  
$$\geq \operatorname{Re} \hat{\mu}(\lambda)$$

from which S-continuity follows easily.

**Example 2.5.1.** Let  $H = L^2(0, \infty)$  and  $U = \mathbb{R}$ . Define

$$S(t)\varphi(\xi) = e^{\lambda t}\varphi(\xi + t),$$
$$b(\xi) = e^{-\xi^2} \in H,$$

and

$$A = \frac{d}{d\xi} + \lambda.$$

One can check that these satisfy our assumptions. We would like to find a non-trivial invariant measure for the equation

$$dX(t) = AX(t)dt + bdW(t)$$

where W(t) is a 1-dimensional Brownian motion. To achieve this we will find  $\varphi \neq 0$ such that  $S(1)\varphi = \varphi$ , and use the above Theorem. Let  $\sigma$  be the distribution of  $S(\eta)\varphi$ where  $\eta$  has uniform distribution on [0,1]. Take  $\lambda = 1$ ,  $\varphi(\xi) = e^{-k}$  for  $\xi \in [k, k+1)$ ,  $k = 0, 1, \ldots$  Then

$$S(1)\varphi(\xi) = e^1\varphi(\xi+1) = e^1e^{-(k+1)}$$
$$= e^{-k} = \varphi(\xi)$$

for  $\xi \in [k, k+1)$ . So we know that there exists a non-trivial invariant measure  $\sigma$  for the semigroup  $S(\cdot)$ . To apply the above theorem, we check that  $\sup_{t>0} \operatorname{Tr} Q_t < \infty$ :

$$\operatorname{Tr} Q_t = \mathbb{E} \left| \int_0^t S(t-s)bdW(s) \right|^2$$
$$= \int_0^t ||S(s)b||_{L_{HS}(\mathbb{R},H)}^2 ds$$
$$= \int_0^t |S(s)b|_{L^2[0,\infty)}^2 ds$$
$$\Rightarrow \sup_{t \ge 0} \operatorname{Tr} Q_t = \int_0^\infty |S(s)b|_{L^2[0,\infty)}^2 ds$$
$$= \int_0^\infty \int_0^\infty e^{2s} e^{-2(\xi+s)^2} ds d\xi < \infty.$$

# 3 Ornstein-Uhlenbeck Processes Driven By Lévy Processes

In this section we extend the results of the previous chapter to the case when the evolution equation is driven not by a Wiener process with continuous trajectories, but by a Lévy process with jumps.

### 3.1 Martingales on Hilbert spaces

Let  $(U, \langle \cdot, \cdot \rangle_U)$  be a Hilbert space and  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$  a filtered probability space. Let  $\mathcal{M}^2(U)$  be the space of all square integrable U-valued martingales which are rightcontinuous with left limits. This technical condition is standard and ensures that the Doob-Meyer decomposition theorem holds. Indeed, let  $M, N \in \mathcal{M}^2(U)$ . Denote by

$$< M, N >_t, \qquad t \ge 0$$

the unique predictable process, with trajectories having locally bounded variation for which

$$< M(t), N(t) >_U - < M, N >_t, \quad t \ge 0$$

is a martingale. By the Doob-Meyer decomposition, the process  $\langle M, N \rangle$  always exists, and is called the *angle bracket* of M and N.

We can also introduce the operator angle bracket  $\ll M, N \gg$  in the following way. As before, let  $L_1(U)$  be the space of all nuclear operators on U. Define, for  $x, y, z \in U$ ,  $x \otimes y(z) := \langle y, z \rangle_U x$ . It is easy to show that  $x \otimes y \in L_1(U)$  and  $||x \otimes y||_{L_1(U)} = |x|_U |y|_U$ .

We also denote by  $L_1^+(U)$  the subspace of  $L_1(U)$  consisting of all self-adjoint nonnegative nuclear operators. If  $M \in \mathcal{M}^2(U)$  then the process  $(M(t) \otimes M(t), t \ge 0)$  is an  $L_1(U)$ -valued right-continuous process such that

$$\mathbb{E} \| M(t) \otimes M(t) \|_{L_1(U)} = \mathbb{E} \| M(t) \|_U^2.$$

We have the following basic result.

**Theorem 3.1.1.** Let  $M \in \mathcal{M}^2(U)$ . Then there is a unique right-continuous  $L_1^+(U)$ -valued increasing predictable process ( $\ll M, M \gg_t, t \ge 0$ ) such that  $\ll M, M \gg_0 = 0$  and the process

$$M(t) \otimes M(t) - \ll M, M \gg_t, \qquad t \ge 0$$

is an  $L_1(U)$ -valued martingale. Moreover, there exists a predictable  $L_1^+(U)$ -valued process  $(Q_t, t \ge 0)$  such that

$$\ll M, M \gg_t = \int_0^t Q_s d < M, M >_s, \quad t \ge 0.$$
 (3.1)

Sketch of Proof. For details see Theorem 8.2 in [11]. The idea of the proof is first to show that for any  $M, N \in \mathcal{M}^2(U)$  there exists a predictable process  $(q(s), s \ge 0)$  such that

$$< M, N >_t = \int_0^t q(s) d(< M, M >_s + < N, N >_s)$$

Then, letting  $\{e_j\}$  be an orthonormal basis of U and writing  $M^j(t) = \langle M(t), e_j \rangle_U \in \mathcal{M}^2(\mathbb{R})$ , we take as a candidate for the operator angle bracket

$$\ll M, M \gg_t := \sum_{j,k} e_k \otimes e_j < M^k, M^j >_t, \qquad t \ge 0.$$
(3.2)

We then prove that the series converges in  $L^1(\Omega, \mathcal{F}, \mathbb{P}; L_{HS}(U))$  by looking at the Hilbert-Schmidt norm of the right hand side. Indeed,

$$\| \ll M, M \gg_t \|_{L_{HS}(U)}^2 = \sum_l \left| \sum_{k,j} < M^k, M^j >_t < e_j, e_l >_U e_k \right|_U^2$$
$$= \sum_l \left| \sum_k < M^k, M^l >_t e_k \right|_U^2$$
$$= \sum_{l,k} < M^k, M^l >_t^2.$$

Since

 $0 \le < M^k + aM^l, M^k + aM^l >_t = a^2 < M^l, M^l >_t + 2a < M^l, M^k >_t + < M^k, M^k >_t,$  we see that

$$< M^k, M^l >_t^2 \le < M^l, M^l >_t < M^k, M^k >_t$$
.

Hence

$$\| \ll M, M \gg_t \|_{L_{HS}(U)}^2 \leq \sum_{k,l} \langle M^l, M^l \rangle_t \langle M^k, M^k \rangle_t$$
$$= \left(\sum_k \langle M^k, M^k \rangle_t\right)^2$$
$$\Rightarrow \mathbb{E} \| \ll M, M \gg_t \|_{L_{HS}(U)} \leq \mathbb{E} \sum_k \langle M^k, M^k \rangle_t = \mathbb{E} |M(t)|_U^2 < \infty$$

Thus  $(\ll M, M \gg_t, t \ge 0)$  is a well-defined process taking values in the space of Hilbert-Schmidt operators on U. It is symmetric and non-negative. Note that for  $0 \le s \le t < \infty$  the operator  $\ll M, M \gg_t - \ll M, M \gg_s$  is also non-negative. Consequently

$$\| \ll M, M \gg_t - \ll M, M \gg_s \|_{L_1(U)} = \text{Tr } \{ \ll M, M \gg_t - \ll M, M \gg_s \}$$
$$= \sum_j \{ < M^j, M^j >_t - < M^j, M^j >_s \}$$

and

$$\mathbb{E} \| \ll M, M \gg_t - \ll M, M \gg_s \|_{L_1(U)} = \mathbb{E} \left( |M(t)|_U^2 - |M(s)|_U^2 \right) < \infty.$$

This shows that  $\ll M, M \gg$  is an  $L_1^+(U)$ -valued predictable increasing process. Finally to show that it is right-continuous and can be represented in the form (3.1) we apply the first part of the proof. Namely, it follows that for any pair k, j there exists a predictable process  $(q^{k,j}(t), t \ge 0)$  such that

$$< M^k, M^j >_t = \int_0^t q^{k,j}(s)d < M, M >_s = \int_0^t q^{k,j}(s)d\sum_l < M^l, M^l >_s .$$

Thus

$$\ll M, M \gg_t = \int_0^t Q_s d < M, M >_s,$$

where

$$Q_s = \sum_{k,j} e_k \otimes e_j q^{k,j}(s), \qquad s \ge 0,$$

is a predictable process with values in  $L_1^+(U)$ .

**Remark 3.1.1.** An  $L_1^+(U)$ -valued process  $V(\cdot)$  is said to be increasing if the operators V(t) - V(s) are non-negative definite for all  $0 \le s \le t$ .

**Definition 3.1.1.** We call the  $L_1^+(U)$ -valued process  $(Q_t, t \ge 0)$  satisfying (3.1) the martingale covariance of M, and the process  $(\ll M, M \gg_t, t \ge 0)$  the operator angle bracket process.

**Proposition 3.1.2.** Let  $M \in \mathcal{M}^2(U)$ . Then for any vectors  $x, y \in U$  and any  $0 \le s \le t \le u \le v < \infty$ ,

$$\mathbb{E}\left(\langle M(t) - M(s), x \rangle_{U} \langle M(t) - M(s), y \rangle_{U} | \mathcal{F}_{s}\right)$$
$$= \mathbb{E}\left(\int_{s}^{t} \langle Q_{r}x, y \rangle_{U} d \langle M, M \rangle_{r} | \mathcal{F}_{s}\right)$$

and

$$\mathbb{E}(\langle M(t) - M(s), x \rangle_{U} \langle M(u) - M(v), y \rangle_{U} | \mathcal{F}_{u}) = 0.$$

The most important case for us will be when  $M \in \mathcal{M}^2(U)$  has zero mean and independent stationary increments. The following result gives an important characterisation of the angle bracket and operator angle bracket in this case.

**Proposition 3.1.3.** Let  $M \in \mathcal{M}^2(U)$  be of zero mean and such that M has independent stationary increments. Then there exists  $Q \in L_1^+(U)$  such that

$$|M(t)|_U^2 - t \operatorname{Tr} Q$$

and

$$M(t) \otimes M(t) - tQ$$

are real- and  $L_1^+(U)$ -valued martingales respectively. Note that Q is the covariance operator of M(1), and that according to the definitions above this means that  $\langle M, M \rangle_t =$  $t \operatorname{Tr} Q, \ll M, M \gg_t = tQ$ , and  $Q_t$  of Theorem 3.1.1 is equal to Q for all t.

*Proof.* First note that for  $s \ge 0$  we have

$$U \times U \ni (x, y) \mapsto \mathbb{E} < M(s), x >_U < M(s), y >_U$$

is a non-negative definite continuous bilinear form on U. Therefore there exists a symmetric non-negative definite continuous linear operator Q(s) such that  $\langle Q(s)x, y \rangle_U = \mathbb{E} \langle M(s), x \rangle_U \langle M(s), y \rangle_U$  for all  $x, y \in U$ , exactly as in (2.8). Since, for any orthonormal basis  $\{e_i\}$  of U,

$$\sum_{j} < Q(s)e_{j}, e_{j} >_{U} = \sum_{J} \mathbb{E} < M(s), e_{j} >_{U}^{2} = \mathbb{E} |M(s)|_{U}^{2} < \infty,$$

Q(s) is nuclear and Tr  $Q(s) = \mathbb{E} |M(s)|_U^2$ . Let  $0 \leq s < t$  and  $x, y \in U$ . Then by independence of increments

$$\mathbb{E} < M(t) - M(s), x >_U < M(s), y >_U = \mathbb{E} < M(t) - M(s), x >_U \mathbb{E} < M(s), y >_U = 0.$$

Hence

$$\mathbb{E} < M(t), x >_U < M(s), y >_U = \mathbb{E} < M(s), x >_U < M(s), y >_U = < Q(s)x, y >_U.$$

We show that Q(s) = sQ(1). Indeed

$$< Q(s+h)x, y >_{U} = \mathbb{E} \left( < M(s+h) - M(s) + M(s), x >_{U} \\ \times < M(s+h) - M(s) + M(s), y >_{U} \right) \\ = \mathbb{E} < M(s+h) - M(s), x >_{U} < M(s+h) - M(s), y >_{U} \\ + \mathbb{E} < M(s), x >_{U} < M(s), y >_{U} \\ = < Q(h)x, y >_{U} + < Q(s)x, y >_{U},$$

again using independent stationary increments, so that Q(s+h) = Q(s) + Q(h). Since the functions  $s \mapsto \langle Q(s)x, x \rangle_U, x \in U$ , are increasing, they are measurable. Then for all  $x, y \in U$ , the function

$$s \mapsto \langle Q(s)x, y \rangle_U = \frac{1}{4} \left( \langle Q(s)(x+y), x+y \rangle_U - \langle Q(s)(x-y), x-y \rangle_U \right)$$

is measurable. Consequently  $\langle Q(s)x, y \rangle_U = s \langle Q(1)x, y \rangle_U$  for all  $x, y \in U$ . Hence, by definition, the covariance of M(s) is sQ.

Doing very similar calculations as above, and using the fact that M(t) - M(s) is independent of  $\mathcal{F}_s$  for all  $0 \leq s < t$ , we have that

$$\mathbb{E}\left(\langle M(t), x \rangle_{U} \langle M(t), y \rangle_{U} - \langle M(s), x \rangle_{U} \langle M(s), y \rangle_{U} \middle| \mathcal{F}_{s}\right) \\ = \mathbb{E} \langle M(t) - M(s), x \rangle_{U} \langle M(t) - M(s), y \rangle_{U} \\ = (t - s) \langle Qx, y \rangle_{U},$$

using the above. Hence  $M(t) \otimes M(t) - tQ$  is a martingale. Moreover, if  $\{e_j\}$  is an orthonormal basis of U, then

$$|M(t)|_{U}^{2} = \sum_{j} < M(t), e_{j} >_{U}^{2}$$

For every  $j, < M(t), e_j >_U^2 - t < Qe_j, e_j >_U, t \ge 0$  is a martingale, and therefore so is

$$|M(t)|_U^2 - t \sum_j \langle Qe_j, e_j \rangle_U = |M(t)|_U^2 - t \operatorname{Tr} Q,$$

which completes the proof.

# 3.2 The stochastic integral with respect to a square integrable martingale

To deal with stochastic partial differential equations one needs the concept of the stochastic integral,  $I_t^M(\Psi) := \int_0^t \Psi(s) dM(s)$ , where  $M \in \mathcal{M}^2(U)$  and  $\Psi(s, \omega)$  are operators from U to another Hilbert space H. As for real martingales we first define the stochastic integral for simple processes, and then extend the class of integrands using the isometric formula (3.3) below. This appeared for the first time in [8], though here we follow [11].

Let U, H be separable Hilbert spaces,  $M \in \mathcal{M}^2(U)$ , and let  $(Q_t, t \ge 0)$  be the martingale covariance of M.

**Definition 3.2.1.** An L(U, H)-valued stochastic process  $\Psi$  is said to be simple if there exists a sequence of non-negative numbers  $t_0 = 0 < t_1 < \cdots < t_m$ , a sequence of operators  $\Psi_j \in L(U, H), j = 0, \ldots, m-1$ , and a sequence of events  $A_j \in \mathcal{F}_{t_j}, j = 0, \ldots, m-1$ , such that

$$\Psi(s) = \sum_{j=0}^{m-1} \mathbf{1}_{A_j} \mathbf{1}_{(t_j, t_{j+1}]}(s) \Psi_j, \qquad s \ge 0.$$

Let  $\mathcal{S}(U, H)$  denote the class of all simple processes with values in L(U, H).

For a simple process  $\Psi \in \mathcal{S}(U, H)$ , we set

$$I_t^M(\Psi) := \sum_{j=0}^{m-1} \mathbf{1}_{A_j} \Psi \left( M(t_{j+1} \wedge t) - M(t_j \wedge t) \right), \qquad t \ge 0.$$

As before, let  $L_{HS}(U, H)$  be the space of all Hilbert-Schmidt operators from U into H.

**Proposition 3.2.1.** For any simple process  $\Psi$ ,  $I_t^M(\Psi), t \ge 0$  is a square-integrable H-valued martingale and

$$\mathbb{E} \left| I_t^M(\Psi) \right|_H^2 = \mathbb{E} \int_0^t \left\| \Psi(s) Q_s^{\frac{1}{2}} \right\|_{L_{HS}(U,H)}^2 d < M, M >_s, \qquad t \ge 0.$$
(3.3)

25

*Proof.* By direct calculation. See Proposition 8.6 of [11] for details.

Now let  $T < \infty$ . We equip the class of all simple processes  $\mathcal{S}(U, H)$  with the seminorm

$$\|\Psi\|_{M,T} := \mathbb{E} \int_0^T \left\|\Psi(s)Q_s^{\frac{1}{2}}\right\|_{L_{HS}(U,H)}^2 d < M, M >_s.$$
(3.4)

We may identify  $\Psi$  with  $\Phi$  if  $\|\Psi - \Phi\|_{M,T} = 0$ . Let  $\mathcal{J}^2_{M,T}(H)$  be the completion of  $(\mathcal{S}(U,H), \|\cdot\|_{M,T})$ . The norm on  $\mathcal{J}^2_{M,T}(H)$  will also be denoted by  $\|\cdot\|_{M,T}$ .

The following theorem shows that we can extend the stochastic integral  $I_t^M(\Psi)$  for  $\Psi \in \mathcal{J}_{M,T}^2(H)$ .

**Theorem 3.2.2.** (i) For any  $t \in [0,T]$ , there exists a unique extension of  $I_t^M$  to a continuous linear operator, denoted also by  $I_t^M$ , from  $(\mathcal{J}_{M,T}^2(H), \|\cdot\|_{M,T})$  into  $L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$ . Moreover, for any  $\Psi \in \mathcal{J}_{M,T}^2(H)$  we have

$$\mathbb{E}\left|I_t^M(\Psi)\right|_H^2 = \|\Psi\|_{M,T}^2$$

(ii) For all  $\Psi \in \mathcal{J}^2_{M,T}(H)$  and  $0 \le s \le t \le T$ , we have  $\mathbf{1}_{(s,t]} \Psi \in \mathcal{J}^2_{M,T}(H)$  and

$$\mathbb{E} \left| I_t^M(\Psi) - I_s^M(\Psi) \right|_H^2 = \| \mathbf{1}_{(s,t]} \Psi \|_{M,T}^2 \le \| \Psi \|_{M,T}^2.$$

- (iii) For any  $\Psi \in \mathcal{J}^2_{M,T}(H)$ ,  $(I^M_t(\Psi), t \in [0,T])$  is an H-valued martingale. It is square integrable and mean-square continuous, and  $I^M_t(\Psi) = 0$ .
- (iv) For any  $\Psi, \Phi \in \mathcal{J}^2_{M,T}(H)$  and any  $t \in [0,T]$ .

$$\left\langle I^{M}(\Psi), I^{M}(\Phi) \right\rangle_{t} = \int_{0}^{t} \left\langle \Psi(s)Q_{s}^{1/2}, \Phi(s)Q_{s}^{1/2} \right\rangle_{L_{HS}(U,H)} d < M, M >_{s}$$

and

$$\ll I^{M}(\Psi), I^{M}(\Psi) \gg_{t} = \int_{0}^{t} \Psi(s)Q_{s}\Psi^{*}(s)d < M, M >_{s}.$$

(v) Let A be a bounded linear operator from H into a Hilbert space V. Then, for every  $\Phi \in \mathcal{J}^2_{M,T}(H)$ , we have  $A\Phi \in \mathcal{J}^2_{M,T}(V)$  and  $AI^M(\Phi) = I^M(A\Phi)$ .

*Proof.* The first two assertions follow from the linearity of  $I_t^M$  on  $\mathcal{S}(U, H)$  and from Proposition 3.2.1. In order to prove mean-square continuity we need to show that

$$\lim_{s \to t} \|\mathbf{1}_{(s,t]}\Psi\|_{M,T} = 0.$$
(3.5)

To do this, consider the family of linear operators  $T(s) : \Psi \mapsto \mathbf{1}_{(s,t]} \Psi$  from  $\mathcal{J}^2_{M,T}(H)$  into  $\mathcal{J}^2_{M,T}(H)$ . We have

$$\sup_{a} \|T(s)\|_{L\left(\mathcal{J}^{2}_{M,T}(H), \mathcal{J}^{2}_{M,T}(H)\right)} \leq 1.$$

Thus since (3.5) holds on a dense subspace, it holds on the whole space by a Banach Steinhaus argument.

It is enough to check the martingale property and the identities in (iv) for simple  $\Psi, \Phi$ . The last assertion of the theorem clearly holds for simple  $\Phi$  and therefore for all  $\Phi$  by standard limiting arguments. Let us again consider the case where  $M \in \mathcal{M}^2(U)$  has zero mean and stationary independent increments. Exactly as in section 2.3.2, it is convenient to introduce the reproducing kernel Hilbert space, and work with this space rather than the space in which M takes its values. First note that for  $Q \in L_1^+(U)$  as in Proposition 3.1.3 we have by above that for all  $t \geq 0$ ,

$$\mathbb{E}\left|\int_{0}^{t}\Psi(s)dM(s)\right|_{U}^{2} = \int_{0}^{t}\|\Psi(s)Q^{1/2}\|_{L_{HS}(U,H)}^{2}ds,$$

since  $Q_s = Q$  for all s. Then let  $U_0 := Q^{1/2}(U)$  equipped with the inner product

$$< u, v >_{U_0} := < Q^{-1/2}u, Q^{-1/2}v >_U,$$

where  $Q^{-1/2}$  is the pseudo-inverse of  $Q^{1/2}$ . We call  $U_0$  the reproducing Hilbert kernel space of M, and M a cylindrical martingale in  $U_0$ .

**Remark 3.2.1.** Exactly as in the case of a cylindrical Wiener process, M does not take values in  $U_0$  unless dim $U_0 < \infty$ . It does however take values in any Hilbert space V such that the embedding  $U_0 \hookrightarrow V$  is Hilbert-Schmidt.

It follows that the class of admissible integrands equals

$$L^{2}(\Omega \times [0,\infty), \mathcal{P}, d\mathbb{P}dt; L_{HS}(U_{0},U))$$

where  $\mathcal{P}$  is the  $\sigma$ -field of predictable sets. Moreover, for any  $\Psi$  in this space, by Theorem 3.2.2 we have that

$$\int_0^t \Psi(s) dM(s)$$

is a square integrable H-valued martingale,

$$\mathbb{E}\left|\int_0^t \Psi(s) dM(s)\right|_H^2 = \int_0^t \mathbb{E}\|\Psi(s)\|_{L_{HS}(U_0,H)}^2 ds$$

and  $\ll \int_0^t \Psi(s) dM(s) \gg_t$  as a process in  $L_1^+(U_0)$  is given by

$$\ll \int_0^t \Psi(s) dM(s) \gg_t = \int_0^t \Phi(s) \Phi^*(s) ds.$$

**Remark 3.2.2.** In the case of a cylindrical Wiener process, this shows that  $\int_0^t S(s)BdW(s)$  is distributed  $\mathcal{N}\left(0, \int_0^t S(s)BB^*S^*(s)ds\right)$ , as claimed at the end of section 2.4.

### 3.3 Lévy processes and the Lévy-Khinchin formula

As usual, let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space and U be a separable Hilbert space.

**Definition 3.3.1.** A stochastic process  $L: \Omega \times [0, \infty) \to U$  is a Lévy process in U if

(i) L(0) = 0,

(ii) L is stochastically continuous i.e.  $\forall \varepsilon > 0, \lim_{h \to 0} \mathbb{P}(|X_{t+h} - X_t| \ge \varepsilon) = 0,$ 

(ii) L has independent stationary increments.

**Remark 3.3.1.** A Lévy process L has a modification which is right-continuous with left limits (see [11], Theorem 4.3). We sometimes say that such a process is càdlàg.

**Remark 3.3.2.** Condition (ii) of the above definition does not in any way imply that the sample paths are continuous. Indeed a Poisson process is a Lévy process. It serves to exclude processes with jumps at fixed non-random times and means that for a given time t, the probability of seeing a jump at t is zero.

**Remark 3.3.3.** It is clear that if L is integrable with zero mean then L is a martingale.

Note that every Lévy process is also Markov with corresponding semigroup

$$P_t\varphi(x) = \int_E \varphi(x+y)\mathcal{L}(L(t))(dy),$$

where  $\mathcal{L}(L(t))$  is the law of L(t).

**Example 3.3.1** (Compound Poisson process). Any compound Poisson process on U is Lévy. Let  $\nu$  be a finite Borel measure on U. Recall that a compound Poisson process with *jump*, or equivalently *Lévy measure*, is given by

$$L(t) := \sum_{j=1}^{\Pi(t)} X_j,$$

where  $\Pi$  is a Poisson process with intensity  $\lambda = \nu(U) < \infty$ , and  $X_j$  are independently identically distributed random variables with distribution

$$\mathbb{P}(X_j \in \Gamma) = \frac{\nu(\Gamma)}{\nu(U)}, \qquad \Gamma \in \mathcal{B}(U).$$

One can think of a compound Poisson process as describing the position of a random walk with step size  $X_j$  after a random number of time steps, given by  $\Pi(t)$ . One can show that

$$\mathbb{E}e^{i < L(t), u > U} = e^{-t\Psi(u)},$$

where

$$\Psi(u) := \int_U \left(1 - e^{i \langle u, v \rangle_U}\right) \nu(dv).$$

Indeed

$$\mathbb{E}\left(e^{i < L(t), u > U}\right) = \mathbb{E}\left(\prod_{j=1}^{\Pi(t)} \mathbb{E}\left(e^{i < X_j, u > U}\right)\right)$$

using independence. Moreover, since  $X_j$  is identically distributed for all j, we have that the characteristic function of  $X_j$  is  $\mathbb{E}\left(e^{i\langle X_j,u\rangle_U}\right) = \frac{\hat{\nu}(u)}{\lambda}$  for all j. Therefore

$$\mathbb{E}\left(e^{i < L(t), u >_U}\right) = \mathbb{E}\left(\left(\frac{\hat{\nu}(u)}{\lambda}\right)^{\Pi(t)}\right)$$
$$= \sum_{n=0}^{\infty} \frac{e^{-\lambda t} (\lambda t)^n \left(\frac{\hat{\nu}(u)}{\lambda}\right)^n}{n!}$$
$$= e^{-t(\lambda - \hat{\nu}(u))}$$
$$= \exp\left\{-t \int_U \left(1 - e^{i < u, v >_U}\right) \nu(dv)\right\}.$$

One can also show that L is integrable if and only if

$$\int_U |u|_U \nu(du) < \infty,$$

and if this is the case, then

$$\mathbb{E}L(t) = t \int_U u\nu(du).$$

Finally, L is square integrable if and only if

$$\int_U |u|_U^2 \nu(du) < \infty$$

and if this is the case, then

$$\mathbb{E}\left\langle L(t) - \mathbb{E}L(t), u \right\rangle_U \left\langle L(t) - \mathbb{E}L(t), v \right\rangle_U = \int_U \langle z, u \rangle_U \langle z, v \rangle_U \nu(dz).$$
(3.6)

**Remark 3.3.4.** The concept of a compound Poisson process is crucial for understanding the characterisation of a general Lévy process. The Lévy-Khinchin Theorem (see below) says that an arbitrary Lévy process is the sum of a Wiener process, a uniform movement, and a "compound Poisson process L with infinite jump measure". How do we construct a compound Poisson process with infinite (but  $\sigma$ -finite) jump measure?

Using the definition of  $\sigma$ -finiteness, we can divide U into a countable sum  $U = \bigcup U_n$ of measurable sets  $U_n$  such that  $\nu(U_n) = 1$  and  $U_n \cap U_m = \emptyset$  for  $m \neq n$ . Then one may try and write  $L = \sum_n L_n$  where  $L_n$  are independent compound Poisson processes each with Lévy measure  $\nu_n$ , which is the restriction of  $\nu$  to  $U_n$ . However, usually this series does not converge in any reasonable sense! The idea is to write

$$U = U_0 \cup U_0^c,$$

where  $U_0 = \bigcup_{n \in I} U_n$  is such that

$$\int_{U_0} |u|_U^2 \nu(du) < \infty, \qquad \int_{U_n} |u|_U \nu(du) < \infty, \qquad \forall n \in I,$$

and  $I^c$  is finite. We can then define the Lévy process with intensity  $\nu$  as as

$$L(t) := \sum_{n \in I} \left( L_n(t) - \mathbb{E}L_n(t) \right) + \sum_{n \notin I} L_n(t),$$

where  $L_n$  are as above. The first sum is a sum of square integrable martingales with

$$\sum_{n\in I} \mathbb{E} \left| L(t) - \mathbb{E}L(t) \right|_U^2 = \int_{U_0} |u|_U^2 \nu(du) < \infty.$$

This follows from (3.6) with appropriate choices of u and v. Recall that the Doob submartingale inequality says that if X is a right-continuous submartingale then

$$r^2 \mathbb{P}\left(\sup_{t \in [0,T]} X(t) \ge r\right) \le \mathbb{E}X^+(T).$$

Using this in our case, we see that

$$\mathbb{P}\left(\sup_{t\in[0,T]}\sum_{n\leq N,n\in I}\left|L(t)-\mathbb{E}L(t)\right|_{U}^{2}\geq r\right)\leq\frac{\int_{\bigcup_{n\leq N}U_{n}}|u|_{U}^{2}\nu(du)}{r^{2}},$$

so that  $\sum_{n \in I} (L_n(t) - \mathbb{E}L_n(t))$  converges in probability (and  $\mathbb{P}$ -a.s.) uniformly in t on any bounded interval.

Assume now that L is a Lévy process on a Hilbert space U which is right-continuous with left limits. The following theorem provides a very useful decomposition.

**Theorem 3.3.1** (Lévy-Khinchin formula). (i) Given a non-negative nuclear operator  $Q_0 \in L_1^+(U), a \in U$ , and a measure  $\nu$  on  $U \setminus \{0\}$  satisfying

$$\int_U |x|_U^2 \wedge 1\nu(dy) < \infty,$$

there is a Lévy process L such that

$$\mathbb{E}e^{i\langle x,L(t)\rangle_U} = \int_U e^{i\langle x,y\rangle_U} \mathcal{L}(L(t))(dy) = e^{-t\psi(x)}$$

and

$$\begin{split} \psi(x) &= -i < a, x >_U + \frac{1}{2} < Q_0 x, x >_U \\ &+ \int_U \left\{ 1 - e^{|\langle x, y \rangle_U|} + \mathbf{1}_{\{|y| < 1\}}(y)i < x, y >_U \right\} \nu(dy) \end{split}$$

Note that the measure  $\nu$  is not necessarily finite, but the condition ensures that the integral is finite.

(ii) Conversely, if L is a Lévy process then there exist  $a, Q_0$  and  $\nu$  as in (i).

**Definition 3.3.2.** We call the measure  $\nu$  appearing in the above theorem the Lévy measure of L, and we call the triple  $(a, Q_0, \nu)$  the characteristics of L.

The Lévy-Khinchin formula follows directly from the Lévy-Khinchin decomposition, which states that any Lévy process L can be decomposed in the following way:

$$L(t) = at + W(t) + \xi(t),$$
(3.7)

where  $a \in U$ , W(t) is a U-valued Wiener process,  $\xi(t)$  is a compound Poisson process, and the processes W(t) and  $\xi(t)$  are independent. If  $Q_0$  is the covariance operator of the Wiener process W(t) and  $\nu$  is the Lévy measure of  $\xi(t)$  then the characteristics of L are  $(a, Q_0, \nu)$  (see [11] Theorem 4.23 or section 5.2 of the appendix for details). Note also that by Remark 3.3.4,  $\nu$  is not necessarily finite.

The Lévy-Khinchin formula gives the explicit form of the characteristic function of a Lévy process. It turns out that it is also useful for computing characteristic functionals of stochastic integrals, as we will see in Theorem 3.4.5 below.

The other major use of the Lévy-Khichin decomposition is that it facilitates the construction of the stochastic integral with respect to a general Lévy process. We will need this construction in section 3.4 below. The decomposition allows us to construct the integral as the sum of a Riemannian integral, an integral with respect to a Wiener process and an integral with respect to a compound Poisson process. To describe the class of operator valued process which are integrable with respect to a general Lévy process is quite technical, and we refer the reader to [11] section 8.6, or [1] Chapter 1 for details.

We finally consider square-integrable Lévy processes. We have (see Theorem 4.47 of [11])

**Theorem 3.3.2.** (i) A Lévy process L on a Hilbert space U is square integrable if and only if its Lévy measure satisfies

$$\int_U |y|_U^2 \nu(dy) < \infty.$$

(ii) Assume L is square integrable, and let L have the representation (3.7) i.e.

$$L(t) = at + W(t) + \xi(t).$$

Let  $Q_0$  be the covariance operator of the Wiener part of L and let  $Q_1$  be the covariance operator of the jump part. Then

$$< Q_1 x, z >_U = \int_U < x, y >_U < z, y >_U \nu(dy), \qquad x, z \in U,$$
  
 $\mathbb{E}L(t) = \left(a + \int_{\{|y|_U \ge 1\}} y\nu(dy)\right) t,$ 

and the covariance Q of L is equal to  $Q_0 + Q_1$ .

If L is a square integrable zero-mean Lévy process, then it clearly a martingale, and so we can use Proposition 3.1.3 to see that  $\langle L, L \rangle_t = t \operatorname{Tr} Q$  and  $\ll L, L \gg_t = tQ$  where Q is the covariance of L(1).

### 3.4 Existence of an invariant measure for OU process driven by a Lévy process

Let U be a separable Hilbert space. Consider the equation on U

$$dX(t) = AX(t)dt + dL(t)$$
(3.8)

where

- A generates a  $C_0$ -semigroup on U,
- L is a Lévy process in U.

We define a mild solution to this equation to be the U-valued process X(t) such that

$$X(t) = S(t)X(0) + \int_0^t S(t-s)dL(s).$$

In this situation it can be shown that a mild solution to (3.8) exists and is unique. Indeed, as mentioned at the end of section 3.3, by decomposing the process L into a deterministic process, a Wiener process and a compound Poisson processes, we can define the integral with respect L, and show that S(s) lies in the space of integrands. Then by Theorem 9.34 of [11] there is a unique weak solution to (3.8). Moreover, by Theorem 9.15 of the same book, we therefore have that there exists a unique mild solution to (3.8). Note that in [11] they deal with much more general equations, but in our simple case the Lipschitz conditions imposed on the constants are trivially satisfied.

We now try and describe conditions on A and the characteristics  $(a, Q_0, \nu)$  of L under which there is a stationary solution (i.e. invariant measure) to (3.8).

**Proposition 3.4.1.** If there is a stationary solution to (3.8) then

$$\sup_{t \ge 0} \operatorname{Tr} \int_0^t S(s) Q_0 S^*(s) ds < \infty.$$

*Proof.* If X(t) is a mild solution to (3.8), then we can write

$$X(t) = S(t)X(0) + \int_0^t S(t-s)dW(s) + \int_0^t S(t-s)d(L(s) - W(s)),$$

where W(s) is the Wiener process in U with covariance  $Q_0$  appearing in the Lévy-Khinchin decomposition. If  $\mu$  is an invariant measure for (3.8) then since everything is independent, we can just multiply characteristic functions to get that

$$\hat{\mu}(\lambda) = \hat{\mu}(S^*(t)\lambda) \cdot \widehat{\mathcal{N}(0,Q_t)} \cdot \hat{\gamma}_t(\lambda)$$

in a similar way to the proof of Theorem 2.5.1, where  $\gamma_t = \mathcal{L}\left(\int_0^t S(t-s)d(L(s)-W(s))\right)$ and  $Q_t = \int_0^t S(s)Q_0S^*(s)ds$ . Now

$$\widehat{\mathcal{N}(0,Q_t)} = e^{-\frac{1}{2} \langle Q_t \lambda, \lambda \rangle_U}$$

so that

$$\operatorname{Re} \hat{\mu}(\lambda) = e^{-\frac{1}{2} \langle Q_t \lambda, \lambda \rangle_U} \operatorname{Re} \left( \hat{\mu}(S^*(t)\lambda) \cdot \hat{\gamma}_t(\lambda) \right)$$
$$\leq e^{-\frac{1}{2} \langle Q_t \lambda, \lambda \rangle_U}$$
$$\Rightarrow \langle Q_t \lambda, \lambda \rangle_U \leq 2 \log \left( \frac{1}{\operatorname{Re} \hat{\mu}(\lambda)} \right).$$

The result follows by an application of Bochner's Theorem, exactly as in Theorem 2.5.1.  $\hfill \Box$ 

**Proposition 3.4.2** (Chojnowska-Michalik). If  $\mathcal{L}\left(\int_0^t S(t-s)dL(s)\right)$  converges weakly, then there is an invariant measure  $\mu$  for the equation (3.8). Moreover, any invariant measure is of the form

$$\mu = \sigma * \eta$$

where  $\sigma$  is any invariant measure for S i.e.  $S(t)\sigma = \sigma$  and

$$\eta = \lim_{t \to \infty} \mathcal{L}\left(\int_0^t S(t-s)dL(s)\right).$$

To prove this theorem we need the following useful lemma.

**Lemma 3.4.3.** Let  $\eta_t = \mathcal{L}\left(\int_0^t S(t-s)dL(s)\right)$ . Then  $\eta_t$  is weakly convergent as  $t \to \infty$  if and only if  $\int_0^\infty S(s)dL(s)$  exists. In this case

$$w - \lim_{t \to \infty} \eta_t = \mathcal{L}\left(\int_0^\infty S(s) dL(s)\right).$$

*Proof.* Let  $\overline{L}$  be the double-sided Lévy process:

$$\bar{L}(t) = \begin{cases} L(t), & t \ge 0\\ \tilde{L}(-t), & t < 0 \end{cases}$$
(3.9)

where  $\tilde{L}$  is an independent indentically distributed Lévy process. Then by properties of the stochastic integral

$$\int_0^t S(t-s)dL(s) \stackrel{d}{=} \int_{-t}^0 S(-u)d\bar{L}(u) \stackrel{d}{=} \int_0^t S(u)dL(u)$$

Since  $\left(\int_0^t S(u)dL(u)\right)_{t\geq 0}$  is a process with independent increments, it converges in distribution as  $t \to \infty$  if and only if it converges in probability. The last convergence means the existence of the integral  $\int_0^\infty S(u)dL(u)$  and the lemma follows.

Proof of Proposition 3.4.2. Let  $\eta_t = \mathcal{L}\left(\int_0^t S(t-s)dL(s)\right)$ . As above, we have that a measure  $\mu$  is invariant for (3.8) if and only if

$$\hat{\mu}(\lambda) = \hat{\mu}(S^*(t)\lambda) \cdot \hat{\eta}_t(\lambda),$$

for all  $\lambda \in U, t \geq 0$ . By assumption  $\eta_t \to \eta$  weakly as  $t \to \infty$ , so by the above lemma  $\int_0^\infty S(s)dL(s)$  exists and  $\eta = \mathcal{L}\left(\int_0^\infty S(s)dL(s)\right)$ . We first show that  $\eta$  is an invariant measure for the equation (3.8).

Let L be the double sided Lévy process as defined in (3.9). From the proof of Lemma 3.4.3 we obtain

$$\eta_t = \mathcal{L}\left(\int_{-t}^0 S(-u)d\bar{L}(u)\right),$$

so that  $\eta$  is also the distribution of  $\int_{-\infty}^{0} S(-u) d\bar{L}(s)$ . Then we have that

$$S(t)\left(\int_{-\infty}^{0} S(-u)d\bar{L}(u)\right) + \int_{0}^{t} S(t-u)d\bar{L}(u) = \int_{-\infty}^{t} S(t-u)d\bar{L}(u)$$
$$\stackrel{d}{=} \int_{-\infty}^{0} S(-u)d\bar{L}(u).$$

Hence, taking characteristic functions in the usual way, we have

$$\hat{\eta}(\lambda) = S(t)\eta(\lambda) \cdot \hat{\eta}_t(\lambda)$$
$$= \hat{\eta}(S(t)^*\lambda) \cdot \hat{\eta}_t(\lambda)$$

for all  $\lambda \in U, t \geq 0$ , which shows that  $\eta$  is indeed and invariant measure for (3.8), and we have shown existence.

For the second part of the proposition, let  $\mu$  be an invariant measure for (3.8). Then

$$\mu = S(t_n)\mu * \eta_{t_n}$$

for  $t_n \to \infty$ .

We now make use of the following general result (see Theorem 2.1 in [9]):

**Theorem 3.4.4.** Assume  $\{\lambda_n\}, \{\mu_n\}, \{\nu_n\}$  are sequences of probability measures on a complete separable metric group G. Let

$$\lambda_n = \mu_n * \nu_n.$$

If  $\{\lambda_n\}$  and  $\{\nu_n\}$  are relatively weakly compact, then so is  $\{\mu_n\}$ .

By assumption  $\{\eta_{t_n}\}$  is relatively weakly compact, and trivially so is  $\{\mu\}$ . Therefore the theorem shows that  $\{S(t_n)\mu\}$  is relatively weakly compact and we may assume  $S(t_n)\mu$  converges weakly to some probability measure  $\sigma$ . Therefore  $\mu = \sigma * \eta$ . Moreover, since  $\mu = S(t)\mu * \eta_t$  for all  $t \ge 0$ , we have that

$$\sigma * \eta = S(t)\sigma * S(t)\eta * \eta_t = S(t)\sigma * \eta$$

Hence we get that  $S(t)\sigma = \sigma$  for all  $t \ge 0$ .

The following is the main result of this chapter, and can be found in [1].

**Theorem 3.4.5.** Assume that the  $C_0$ -semigroup  $S(t), t \ge 0$  is exponentially stable, which means that there exist constants  $C, \alpha > 0$  such that

$$||S(t)||_{L(U)} \le Ce^{-\alpha t}.$$

Then the following are equivalent:

- (i) There is a unique invariant measure  $\eta$  for the Ornstein-Uhlenbeck equation (3.8).
- (ii)  $\eta_t = \mathcal{L}\left(\int_0^t S(t-s)dL(s)\right)$  converges weakly as  $t \to \infty$  to a probability measure  $\eta$ .
- (iii)  $\int_U \log^+ |x|_U \nu(dx) < \infty$ , where  $\nu$  is the Lévy measure for L.

*Proof.* We first show (ii) implies (i). So suppose  $\mathcal{L}\left(\int_0^t S(t-s)dL(s)\right)$  converges weakly to  $\eta$  as  $t \to \infty$ . By Proposition 3.4.2 we therefore know that there is an invariant measure  $\mu$  for (3.8) and  $\mu = \sigma * \eta$  where  $\sigma$  is an invariant measure for S. However, we also have that

$$\hat{\sigma}(\lambda) = \lim_{t \to \infty} \hat{\mu}(S^*(t)\lambda)$$

as usual, and  $S^*(t)\lambda \to 0$  since S is exponentially stable. Therefore  $\hat{\sigma}(\lambda) = \hat{\mu}(0) = 1$ , which implies  $\sigma = \delta_0$ , so that  $\mu = \eta$  as required.

For (i) implies (ii) if  $\mu$  is an invariant measure for (3.8), then we have that

$$\hat{\mu}(\lambda) = \hat{\mu}(S^*(t)\lambda) \cdot \hat{\eta}_t(\lambda)$$
$$\Leftrightarrow \mu = S(t)\mu * \eta_t.$$

 $\{S(t)\mu : t \geq 0\}$  is relatively weakly compact since by the exponential stability of S,  $S(t)\mu \to \delta_0$  as  $t \to \infty$ , and trivially so is  $\{\mu\}$ . Hence by Theorem 3.4.4 we have  $\{\eta_t : t \geq 0\}$  is relatively weakly compact i.e.  $\exists \{t_j\} \subset [0, \infty)$  such that  $t_j \to \infty$  and  $\eta_{t_j}$  converges weakly to some measure  $\eta$ . Finally we have that

$$\hat{\eta}_t(\lambda) = \frac{\hat{\mu}(\lambda)}{\hat{\mu}(S^*(t)\lambda)} \to \hat{\mu}(\lambda)$$

so by uniqueness of limits  $\eta_t \to \eta = \mu$ .

We give an outline of how to prove (iii)  $\Rightarrow$  (ii). For full details see [1]. The first thing to notice is that in the Lévy-Khinchin decomposition of the Lévy process L we may assume that the Wiener part is 0, since the law of the integral with respect to the Wiener part converges weakly irrespective of whether condition (iii) holds or not.

Our strategy will be the following. Let  $Z(t) = \int_0^t S(t-s)dL(s)$ . We will show that  $\mathcal{L}(Z(t)) = \mathcal{L}(L^t(1))$  where  $L^t$  is some Lévy process with characteristics  $a_t \in U$ , and  $\nu_t$  (where  $\nu_t$  the Lévy measure). We will then pass to the limit  $a_t \to a, \nu_t \to \nu$ . We must check the convergence, and that

$$\int |x|_U^2 \wedge 1\nu(dx) < \infty$$

to ensure that  $\nu$  is a Lévy measure.

We will need the following lemma:

**Lemma 3.4.6.** Let L be a Lévy process. Let  $\psi$  be the exponent of the characteristic function of L i.e.  $\psi$  is such that

$$\mathbb{E}e^{i \langle x, L(t) \rangle_U} = e^{-t\psi(x)}$$

(which exists by the Lévy-Khinchin formula). Let  $F:[0,T] \rightarrow U = U^*$ . Then

$$\mathbb{E}e^{i\alpha\int_0^T F(s)dL(s)} = e^{-\int_0^T \psi(\alpha F(s))ds}$$

provided the integral is well defined in the Riemannian sense.

Proof of Lemma. We have

$$\mathbb{E}e^{i\alpha\int_0^T F(s)dL(s)} = \lim \mathbb{E}e^{i\alpha\sum_i F(s_i)(L(s_{i+1}) - L(s_i))}$$

by definition of the integral, where the limit is over all finite partitions  $\{s_i\}$  of the interval [0, T]. Since L has independent and stationary increments by definition, we then have that

$$\mathbb{E}e^{i\alpha\int_0^T F(s)dL(s)} = \lim \prod_i \mathbb{E}e^{i\alpha F(s_i)(L(s_{i+1}) - L(s_i))}$$
$$= \lim \prod_i \mathbb{E}e^{i\alpha F(s_i)L(s_{i+1} - s_i)}.$$

Then by definition of  $\psi$  in the Lévy-Khinchin formula, this gives

$$\mathbb{E}e^{i\alpha\int_0^T F(s)dL(s)} = \lim \prod_i e^{-(s_{i+1}-s_i)\psi(\alpha F(s_i))}$$
$$= e^{-\int_0^T \psi(\alpha F(s))ds}.$$

Now	we	cal	leui	late
	w	Ca.	ւսս	auc

$$\mathcal{L}\left(\int_{0}^{t} \widetilde{S(t-s)} dL(s)\right)(\lambda) = \mathbb{E}e^{i < \int_{0}^{t} S(t-s) dL(s), \lambda > U}$$
$$= \mathbb{E}e^{i \int_{0}^{t} S^{*}(t-s) \lambda dL(s)}$$
$$= e^{-\psi_{t}(\lambda)}$$

where we have used the above lemma, since  $S^*(t-s)\lambda \in U^*$ , and where

$$\psi_t(\lambda) = \int_0^t \psi(S^*(t-s)\lambda) ds.$$

36

Now by a change of variables and using the explicit form of  $\psi$  given in the Lévy-Khinchin formula (Theorem 3.3.1)

$$\begin{split} \psi_t(\lambda) &= \int_0^t \psi(S^*(s)\lambda) ds \\ &= -i \int_0^t < S^*(s)\lambda, a >_U ds \\ &\quad -\int_0^t \int_U \left\{ e^{i < S^*(s)\lambda, y >_U} - 1 - \mathbf{1}_{\{|y| < 1\}}(y)i < S^*(s)\lambda, y >_U \right\} \nu(dy) ds \\ &= -i \left\langle \lambda, \int_0^t S(s) a ds \right\rangle_U \\ &\quad -\int_0^t \int_U \left\{ e^{i < \lambda, S(s)y >_U} - 1 - \mathbf{1}_{\{|y| < 1\}}(y)i < \lambda, S(s)y >_U \right\} \nu(dy) ds. \end{split}$$

Denote the second term by  $I_t(\lambda)$ . We would like to have

$$I_t(\lambda) = \int_U \left( e^{i < x, y > U} - 1 - \mathbf{1}_{\{|y| < 1\}}(y) i < x, y > U \right) \nu_t(dy).$$

If we set  $\nu_t = \int_0^t S(s)\nu dt$  then we have

$$\begin{split} I_{t}(\lambda) &= \int_{0}^{t} \int_{U} \left\{ e^{i < \lambda, S(s)y >_{U}} - 1 - \mathbf{1}_{\{|y| < 1\}}(S(s)y)i < \lambda, S(s)y >_{U} \right\} \nu(dy) ds \\ &+ \int_{0}^{t} \int_{U} \left[ \mathbf{1}_{\{|y| < 1\}}(S(s)y) - \mathbf{1}_{\{|y| < 1\}}(y) \right] i < \lambda, S(s)y >_{U} \nu(dy) ds \\ &= \int_{U} \left\{ e^{i < \lambda, y >_{U}} - 1 - \mathbf{1}_{\{|y| < 1\}}(y)i < \lambda, y >_{U} \right\} \nu_{t}(dy) \\ &+ i \left\langle \lambda, \int_{0}^{t} \int_{U} \left[ \mathbf{1}_{\{|y| < 1\}}(S(s)y) - \mathbf{1}_{\{|y| < 1\}}(y) \right] S(s)y\nu(dy) ds \right\rangle_{U}. \end{split}$$

Therefore

$$\begin{split} \psi_t(\lambda) &= -i \left\langle \lambda, \int_0^t S(s) a ds + \int_0^t \int_U \left[ \mathbf{1}_{\{|y|<1\}}(S(s)y) - \mathbf{1}_{\{|y|<1\}}(y) \right] S(s) y \nu(dy) ds \right\rangle_U \\ &+ \int_U \left\{ 1 - e^{i < \lambda, y >_U} + \mathbf{1}_{\{|y|<1\}}(y) i < \lambda, y >_U \right\} \nu_t(dy) \\ &= -i \left\langle \lambda, a_t \right\rangle_U \\ &+ \int_U \left\{ 1 - e^{i < \lambda, y >_U} + \mathbf{1}_{\{|y|<1\}}(y) i < \lambda, y >_U \right\} \nu_t(dy) \end{split}$$

for  $a_t = \int_0^t S(s) a ds + \int_0^t \int_U \left[ \mathbf{1}_{\{|y| < 1\}}(S(s)y) - \mathbf{1}_{\{|y| < 1\}}(y) \right] S(s) y \nu(dy) ds.$ 

Note that this is in the same form as the exponent in the Lévy-Khinchin formula, so we have calculated the element  $a_t \in U$  and Lévy measure  $\nu_t$  which characterise  $\mathcal{L}\left(\int_0^t S(t-s)dL(s)\right)$ . Hence,

$$\mathcal{L}\left(\int_0^t S(t-s)dL(s)\right) = \mathcal{L}(L^t(1)),$$

where  $L^{t}(1)$  has characteristics  $(a_{t}, \nu_{t})$  (recall we are assuming the Wiener process to be 0), and we can see that  $\mathcal{L}\left(\int_{0}^{t} S(t-s)dL(s)\right)$  converges weakly if and only if

 $a_t \to a$ 

in U, and  $\nu = \int_0^\infty S(s)\nu ds = \lim_{t\to\infty} \nu_t$  satisfies

$$\int_U |x|_U^2 \wedge 1\tilde{\nu}(dx) < \infty.$$

Using more detailed calculations, one can show that the above conditions are equivalent to (iii) (again see [1]). Here we will just present some calculations for the convergence of  $a_t$ . Note that

$$\begin{split} \int_0^t \int_U \left[ \mathbf{1}_{\{|y|<1\}}(S(s)y) - \mathbf{1}_{\{|y|<1\}}(y) \right] S(s)y\nu(dy)ds &= \int_0^t \int_{\substack{|y|>1\\|S(s)y|\leq 1}} S(s)y\nu(dy)ds \\ &\quad -\int_0^t \int_{\substack{|y|\leq 1\\|S(s)y|>1}} S(s)y\nu(dy)ds \\ &= a_t^1 - a_t^2. \end{split}$$

Now

$$\begin{aligned} |a_t^2| &\leq \int_0^t \int_{|y|_U \leq 1} |S(s)y|_U^2 \nu(dy) ds \\ &\leq C^2 \int_0^t e^{-2\alpha s} \int_{|y|_U \leq 1} |y|_U^2 \nu(dy) ds \\ &\leq C^2 \int_0^t e^{-2\alpha s} ds \int |y|_U^2 \wedge 1\nu(dy) \end{aligned}$$

which converges as  $t \to \infty$  since  $\nu$  is a Lévy measure. For  $a_t^1$ :

$$\begin{split} |a_t^1| &\leq \int_0^\infty \int_{\substack{|y|_U > 1 \\ |S(s)y|_U \leq 1}} |S(s)y|_U \nu(dy) ds \\ &\leq \int_0^\infty \int_{1 < |y|_U < e^{\alpha s/2}} |S(s)y|_U \nu(dy) ds \\ &\int_0^\infty \int_{\substack{|y|_U \geq e^{\alpha s/2} \\ |S(s)y|_U \leq 1}} |S(s)y|_U \nu(dy) ds \\ &\leq \int_0^\infty \int_{1 < |y|_U < e^{\alpha s/2}} |S(s)y|_U \nu(dy) ds \\ &\int_0^\infty \int_{|y|_U \geq e^{\alpha s/2}} \nu(dy) ds. \end{split}$$

Note the first term will converge in the same way as  $a_t^2$ , using the exponential stability of S. Finally the second term can be written as

$$\int_0^\infty \int_{\frac{2}{\alpha}\log^+ |y|_U \ge s} \nu(dy) ds = \int_U \int_0^{\frac{2}{\alpha}\log^+ |y|_U} ds \nu(dy)$$
$$= \int_U \frac{2}{\alpha}\log^+ |y|_U \nu(dy)$$

which is finite if and only if  $\int_U \log^+ |y|_U \nu(dy) < \infty$ . Hence  $a_t$  converges if condition (iii) holds.

### 4 Convergence to the Invariant Measure

### 4.1 Exponential mixing

Let  $P(t, x, \cdot)$  be a transition function on a Banach space E, with corresponding transition semigroup  $(P_t)_{t\geq 0}$ . We assume that  $P_t$  is Feller, i.e. that  $P_t$  maps the space  $C_b(E)$  of continuous bounded functions on E to itself.

Denote by  $(\varphi, \mu)$  the action of a probability measure  $\mu$  on a function  $\varphi$  i.e.  $(\varphi, \mu) = \mu(\varphi)$ .

**Definition 4.1.1.** We say that an invariant measure  $\mu$  for  $P_t$  is exponentially mixing with exponent  $\omega > 0$  and a function  $c : E \to (0, \infty)$  if

$$|P_t\varphi(x) - (\varphi, \mu)| \le c(x)e^{-\omega t} \|\varphi\|_{Lip}$$

for all  $x \in E, t \ge 0$  and bounded Lipschitz functions  $\varphi$  on E.

Let us equip the space of Borel probability measures on  ${\cal E}$  with the so-called Fortet-Mourier norm

$$\|\rho\|_{FM} = \sup\{|(\varphi, \rho)| : \|\varphi\|_{\infty} \le 1, \|\varphi\|_{Lip} \le 1\}$$

Then  $\mu$  is exponentially mixing with exponent  $\omega$  and function c if and only if

$$||P_t(x,\cdot) - \mu||_{FM} \le c(x)e^{-\omega t}, \qquad \forall t > 0, x \in E.$$

It is known (see e.g. [7]) that weak convergence of measures is equivalent to convergence in the Fortet-Mourier norm. Thus, if  $\mu$  is exponentially mixing then, for any  $x \in E$ ,  $P_t(x, \cdot)$  converges weakly to  $\mu$  as  $t \to \infty$ .

The following result provides useful conditions for the existence, uniqueness and exponential mixing of an invariant measure.

#### **Proposition 4.1.1.** Assume that

- (i) there exists  $x_0 \in E$  such that  $P_t(x_0, \cdot)$  converges weakly to a probability measure  $\mu$ ,
- (ii) there exist functions  $c : E \to (0, \infty)$ ,  $\tilde{c} : E \times E \to (0, \infty)$  and a constant  $\omega > 0$ such that for all  $s \ge t \ge 0, x, \tilde{x} \in E$  and  $\varphi \in Lip(E)$  we have

$$|P_t\varphi(x) - P_s\varphi(x)| \le c(x)e^{-\omega t} \|\varphi\|_{Lip}$$

and

$$|P_t\varphi(x) - P_t\varphi(\tilde{x})| \le \tilde{c}(x,\tilde{x})e^{-\omega t} \|\varphi\|_{Lip}$$

Then  $\mu$  is the unique invariant measure for  $(P_t)$ , and it is exponentially mixing with exponent  $\omega$  and function c.

*Proof.* By the Krylov-Bogoliubov existence theorem (see appendix)  $\mu$  is invariant for  $(P_t)_{t\geq 0}$ . Moreover, by the assumptions

$$\begin{aligned} |P_t\varphi(x) - (\varphi, \mu)| &= \lim_{s \to \infty} |P_t\varphi(x) - P_s\varphi(x_0)| \\ &\leq \lim_{s \to \infty} |P_t\varphi(x) - P_s\varphi(x)| + \lim_{s \to \infty} |P_s\varphi(x) - P_s\varphi(x_0)| \\ &\leq c(x)e^{-\omega t} \|\varphi\|_{Lip}. \end{aligned}$$

For the uniquenes, let  $\tilde{\mu}$  be another invariant measure. Let  $\psi \in Lip(E)$ . Then

$$(\psi, \tilde{\mu}) = (P_t \psi, \tilde{\mu}) \rightarrow \int_E (\psi, \mu) \tilde{\mu}(dx) = (\psi, \mu).$$

Therefore  $(\psi, \tilde{\mu}) = (\psi, \mu)$  for all Lipschitz  $\psi$ , which implies that  $\mu = \tilde{\mu}$ .

### 4.2 Existence and exponential mixing: Regular case

In this section we will let L be a square integrable mean-zero martingale with RKHS U and Lévy measure  $\nu$ . Since L is square-integrable, by Theorem 3.1.3 we can assume that  $\int_{U} |y|_{U}^{2} \nu(dy) = \kappa < \infty$ . We will consider the equation

$$dX(t) = (AX(t) + F(X(t)))dt + G(X(t))dL$$
(4.1)

where

- $(A, \mathcal{D}(A))$  generates a  $C_0$ -semigroup S on a Hilbert space H,
- $F: H \to H$  and  $G: H \to L_{HS}(U, H)$  are Lipschitz.

Once again, under these Lipschitz conditions we have that there exists a unique mild (or equivalently weak) solution to (4.1). See [11] Chapter 9 for details, as before.

We would like to know when an invariant measure for such a process is unique and exponentially mixing. We outline the strategy. Let X(t, x) be the value at time t of the solution X(t, x) starting at time 0 from x. Taking into account the Krylov-Bogoliubov theorem we would like to show the weak convergence of  $P_t(x, \cdot) = \mathcal{L}(X(t, x))$ . To do this one can ask whether X(t, x) converges in probability (or in  $L^2$ ) to a random variable. This, however, is not even true in the simplest case, that of the one-dimensional Ornstein-Uhlenbeck diffusion, for which

$$dX(t) = -\frac{1}{2}X(t)dt + dW(t), \qquad X(0) = 0, \tag{4.2}$$

where W(t) is a Brownian motion in  $\mathbb{R}$ . Indeed, in this case

$$\mathcal{L}(X(t,0)) = \mathcal{L}\left(\int_0^t e^{-(t-s)/2} dW(s)\right)$$
$$= \mathcal{L}\left(\int_0^t e^{s/2} dW(s)\right)$$
$$\to \mathcal{L}\left(\int_0^\infty e^{s/2} dW(s)\right) = \mathcal{N}(0,1)$$

However for every  $\varepsilon > 0, n \in \mathbb{N}$ ,

$$\lim_{n \to \infty} \mathbb{P}\left( |X(n+1,0) - X(n,0)| \ge \varepsilon \right) \neq 0$$
(4.3)

so X(n,0) is not a Cauchy sequence in probability. (4.3) follows from the fact that  $\mathcal{L}(X(n+1,0) - X(n,0)) = \mathcal{N}(0,\delta_n)$ , where

$$\delta_n = \int_0^n \left( e^{-(n+1-s)} - e^{-n-s} \right)^2 ds + \int_n^{n+1} e^{-2(n+1-s)} ds$$

The first term converges to some  $\frac{1}{2}e^{-2}$  while the second term is constant  $(\frac{1}{2}(1-e^{-2}))$ . Therefore X(t,0) does not converge in probability to any random variable (since  $\delta_n \neq 0$ ).

To get around this problem we consider the double-sided Lévy process  $\overline{L}$  (as defined in the previous section):

$$\bar{L}(t) = \begin{cases} L(t), & t \ge 0\\ \tilde{L}(-t), & t < 0 \end{cases}$$

where  $\tilde{L}$  is an independent identically distributed Lévy process. Given  $-\infty < t_0 \leq t < \infty$ and x, let  $X(t, t_0, x)$  be the value at time t of the (mild) solution to the equation

$$dX(t) = (AX(t) + F(X(t)))dt + G(X(t))d\bar{L}(t), \qquad X(t_0) = x.$$
(4.4)

From the uniqueness of the solution,  $\mathcal{L}(X(t,x)) = \mathcal{L}(X(t_0 + t, t_0, x))$ . We will show that under certain conditions on A, F and  $G, X(t_0, 0, x)$  converges in probability as  $t_0 \to -\infty$ . In this way we obtain the existence of an invariant measure. To show it's exponentially mixing we will use Proposition 4.1.1.

Below,  $A_n$  stands for the Yosida approximation of A; see appendix. The following result is from [13], which should be referred to for the details.

**Theorem 4.2.1.** Assume there exists  $\omega > 0$  such that for all  $x, y \in U, n \in \mathbb{N}$ ,

$$2 < A_n(x-y) + F(x) - F(y), x-y >_H + ||G(x) - G(y)||^2_{L_{HS}(U,H)} \le -\omega|x-y|^2_H$$

(where  $A_n$  are the Yosida approximations to the generator A). Then there exists a unique invariant measure  $\mu$  and it is exponentially mixing with exponent  $\omega/2$  and a function  $c: H \to (0, \infty)$  of linear growth.

We first prove a Lemma:

#### Lemma 4.2.2. Suppose

$$dY(t) = a(t)dt + \beta(t)d\bar{L}(t),$$

where  $a(t) \in H$ ,  $\beta(t) \in L_{HS}(U, H)$  for  $t \in \mathbb{R}$ . Then

$$\mathbb{E}|Y(t)|_{H}^{2} \leq \mathbb{E}|Y(t_{0})|_{H}^{2} + \mathbb{E}\int_{t_{0}}^{t} \left(2 < Y(s), a(s) >_{H} + \|b(s)\|_{L_{HS}(U,H)}^{2}\right) ds.$$

*Proof.* Applying Itô's formula (see section 5.5 of the appendix) to the function  $\psi(x) = |x|_{H}^{2}$ , we obtain

$$\begin{split} |Y(t)|_{H}^{2} &= |Y(t_{0})|_{H}^{2} + \int_{t_{0}}^{t} < D\psi(Y(s-)), dY(s) >_{H} + \frac{1}{2} \int_{t_{0}}^{t} D^{2}\psi(Y(s-))d[\![M,M]\!]_{s}^{c} \\ &+ \int_{t_{0}}^{t} \int_{H} \varphi(s,y)\pi_{Y}(ds,dy), \end{split}$$

where  $\varphi(s, y) = \psi(Y(s-)+y) - \psi(Y(s-)) - \langle D\psi(Y(s-)), y \rangle_H = |y|_H^2, \pi_Y((0,t], \Gamma) := \sum_{s \leq t} \mathbf{1}_{\Gamma}(Y(s) - Y(s-))$  is the measure of the jumps of Y and M is the matringale part of Y. Hence

$$|Y(t)|_{H}^{2} = |Y(t_{0})|_{H}^{2} + 2\int_{t_{0}}^{t} \langle Y(s-), dY(s) \rangle_{H} + \int_{t_{0}}^{t} d\llbracket M, M \rrbracket_{s}^{c} + \int_{t_{0}}^{t} \int_{H} |y|_{H}^{2} \pi_{Y}(ds, dy).$$

$$(4.5)$$

Now by definition (again see section 5.5 of the appendix),

$$\int_{t_0}^t d[\![M, M]\!]_s^c = (1 - \kappa) \int_{t_0}^t \|b(s)\|_{L_{HS}(U, H)} ds$$

and

$$\int_{t_0}^t \int_H |y|_H^2 \pi_Y(ds, dy) = \int_{t_0}^t \int_U |b(s)y|_H^2 \pi(ds, dy)$$

where  $\pi((0,t],\Gamma) = \sum_{s \le t} \mathbf{1}_{\Gamma}(L(s) - L(s-))$  is the measure of the jumps of L. We also clearly have that

$$\mathbb{E}\int_{t_0}^t \langle Y(s-), dY(s) \rangle_H = \mathbb{E}\int_{t_0}^t \langle Y(s-), a(s) \rangle_H \, ds$$

Let  $B_s := \{Y(s) \neq Y(s-)\}$ . Then

$$\mathbb{E} < Y(s) - Y(s-), a(s) >_{H} = \mathbb{E} \left( \mathbf{1}_{B_{s}} < Y(s) - Y(s-), a(s) >_{H} \right) = 0,$$

since  $\mathbb{P}(B_s) = 0$ . Then

$$\mathbb{E} \int_{t_0}^t \langle Y(s-), a(s) \rangle_H \, ds = \mathbb{E} \int_{t_0}^t \langle Y(s), a(s) \rangle_H \, ds.$$

44

Finally

$$\begin{split} \mathbb{E} \int_{t_0}^t \int_U |b(s)y|_H^2 \pi(ds, dy) &= \mathbb{E} \int_{t_0}^t \int_U |b(s)y|_H^2 \nu(dy) ds \\ &\leq \mathbb{E} \int_{t_0}^t \int |y|_U^2 \|b(s)\|_{L_{HS}(U,H)}^2 \nu(dy) ds \\ &= \kappa \mathbb{E} \int_{t_0}^t \|b(s)\|_{L_{HS}(U,H)}^2 ds, \end{split}$$

since we are assuming that  $\int_U |y|_U^2 \nu(dy) = \kappa$ . Putting all this together and taking expectations in 4.5 we see that

$$\mathbb{E}|Y(t)|_{H}^{2} \leq \mathbb{E}|Y(t_{0})|_{H}^{2} + 2\mathbb{E}\int_{t_{0}}^{t} \langle Y(s), a(s) \rangle_{H} \, ds + \mathbb{E}\int_{t_{0}}^{t} \|b(s)\|_{L_{HS}(U,H)} ds.$$

Proof of Theorem 4.2.1: Let  $\overline{L}$  be the double sided Lévy process corresponding to L. Given  $n \in \mathbb{N}, t_0 \in \mathbb{R}$  and  $x \in H$  we consider the regularised problem

$$dX(t) = (A_n X(t) + F(X(t))) dt + G(X(t)) d\bar{L}(t), \qquad t \ge t_0, \qquad X(t_0) = x,$$

with a straightforward generalisation of the stochastic integral with respect to  $\overline{L}$ . It is quite easy to show that the regularisation equation has a unique solution  $X_n(t) = X_n(t, t_0, x)$  by general existence theorems (see for example [11] chapter 9) and that, since  $A_n$  is a bounded linear operator,  $X_n$  is a strong solution. That is,

$$X_n(t) = x + \int_{t_0}^t \left( A_n X_n(s) + F(X_n(s)) \right) ds + \int_{t_0}^t G(X_n(s)) d\bar{L}(s) d\bar{L}$$

Moreover, for each  $t \ge t_0$ ,  $X_n(t)$  converges in  $L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$  to the unique (mild) solution  $X(t, t_0, x)$  of (4.4).

We will divide the proof into three steps:

Step 1: Here we prove that

$$\mathbb{E}|X(t,t_0,x)|_H^2 \le C(1+|x|_H^2) \tag{4.6}$$

for all  $x \in H, t > t_0, t_0 \in \mathbb{R}$ . Using Lemma 4.2.2 we get

$$\mathbb{E} |X_n(t)|_H^2 \le |x|_H^2 + \mathbb{E} \int_{t_0}^t \left\{ 2 \left\langle X_n(s), A_n X_n(s) + F(X_n(s)) \right\rangle_H + \|G(X_n)\|_{L_{HS}(U,H)}^2 \right\} ds.$$

Note that

$$2 \langle A_n x + F(x), x \rangle_H + \|G(x)\|^2_{L_{HS}(U,H)}$$
  

$$\leq 2 \langle A_n x + F(x) - F(0), x \rangle_H + \|G(x) - G(0)\|^2_{L_{HS}(U,H)} + I(x),$$

#### 4 Convergence to the Invariant Measure

where

$$I(x) := 2|F(0)|_{H}|x|_{H} + 2||G(x) - G(0)||_{L_{HS}(U,H)}||G(0)||_{L_{HS}(U,H)} + ||G(0)||^{2}_{L_{HS}(U,H)}.$$

Clearly for any  $\varepsilon > 0$  there is a constant  $C_{\varepsilon}$  such that

$$I(x) \le \varepsilon \left( \|G(x) - G(0)\|_{L_{HS}(U,H)}^2 + |x|_H^2 \right) + C_{\varepsilon}.$$

Thus, since G is Lipschitz continuous and by assumption, there is a constant  $C_1$  such that for all n and x, we have

$$2\langle A_n x + F(x), x \rangle_H + \|G(x)\|_{L_{HS}(U,H)}^2 \le -\frac{\omega}{2} |x|_H^2 + C_1.$$

Hence

$$\mathbb{E} |X_n(t)|_H^2 \le |x|_H^2 - \frac{\omega}{2} \mathbb{E} \int_{t_0}^t |X_n(s)|_H^2 \, ds + C_1(t - t_0),$$

and so by Gronwall's lemma

$$\mathbb{E} |X_n(t)|_H^2 \le e^{-\omega(t-t_0)/2} \left( |x|_H^2 + C_1(t-t_0) \right).$$

Letting  $n \to \infty$  we obtain (4.6).

**Step 2:** Recall that  $X(t, t_0, x)$  is the value at time t of the solution to (4.4). We will show that there is a constant K such that, for all  $x \in H, t_0 < 0$  and h > 0,

$$\mathbb{E}|X(0,t_0,x) - X(0,t_0-h,x)|_H^2 \le K e^{\omega t_0} (1+|x|_H^2), \tag{4.7}$$

and

$$\mathbb{E}|X(0,t_0,x) - X(0,t_0,\tilde{x})|_H^2 \le K e^{\omega t_0} |x - \tilde{x}|_H^2.$$
(4.8)

To do this, observe that  $X_n(t, t_0 - h, x) = X_n(t, t_0, X_n(t_0, t_0 - h, x))$ . Thus, by Itô's lemma

$$\Delta_n(t, t_0, h, x) := \mathbb{E} \left| X_n(t, t_0, x) - X_n(t, t_0 - h, x) \right|_H^2$$

satisfies

$$\Delta_n(t, t_0, h, x) \le \Delta_n(t_0, t_0, h, x) - \omega \int_{t_0}^t \Delta_n(s, t_0, h, x) ds$$

and hence by Gronwall's inequality,

$$\Delta_n(t, t_0, h, x) \le e^{-\omega(t-t_0)} \mathbb{E} |X_n(t_0, t_0 - h, x) - x|_H^2.$$

Since by step 1 there exists a constant C such that

$$\mathbb{E}|X(t,t_0,x)|_U^2 \le C(1+|x|_U^2)$$

for all  $t_0, h, x$ , we have

$$\Delta_n(t, t_0, h, x) \le e^{-\omega(t - t_0)} 2C(1 + |x|_U^2).$$

Letting  $n \to \infty$  and t = 0 we obtain (4.7). Similarly for (4.8).

Step 3: We will show that the assumptions of 4.1.1 are satisfied. It follows from (4.7) that  $X(0, t_0, x)$  converges in  $L^2(\Omega, \mathcal{F}, \mathbb{P}; H)$  as  $t_0 \to -\infty$  to a random variable  $\tilde{X}$ . Therefore  $\mathcal{L}(X(-t_0, x)) = \mathcal{L}(X(0, t_0, x))$  converges weakly to  $\mu := \mathcal{L}(\tilde{X})$ . Now, for  $\psi \in Lip(U), s \ge t \ge 0$  and  $x, \tilde{x} \in H$ 

$$|P_t\psi(x) - P_s\psi(x)|_H^2 = |\mathbb{E}(\psi(X(t,x)) - \psi(X(s,x)))|_H^2$$
  

$$\leq ||\psi||_{Lip}^2 \mathbb{E}|X(0,-t,x) - X(0,-s,x)|_H^2$$

and

$$|P_t\psi(x) - P_t\psi(\tilde{x})|_H^2 = |\mathbb{E}(\psi(X(t,x)) - \psi(X(t,\tilde{x})))|_H^2$$
  
$$\leq ||\psi||_{Lip}^2 \mathbb{E}|X(0,-t,x) - X(0,-t,\tilde{x})|_H^2.$$

Hence by (4.7)

$$|P_t\psi(x) - P_s\psi(x)|_H^2 \le Ke^{-\omega t}(1 + |x|_H^2) \|\psi\|_{Lip}^2$$

and, by (4.8)

$$P_t\psi(x) - P_t\psi(\tilde{x})|_H^2 \le Ke^{-\omega t}|x - \tilde{x}|_H^2 ||\psi||_{Lip}^2$$

### 4.3 An example: the Heath-Jarrow-Morton model

This is a well known model which is used in mathematical finance to price bonds. For an in-depth treatment see [11], or the original paper [4]. A basic concept in bond market theory is the forward rate function. Denote by

$$P(t,\theta), \qquad 0 \le t \le \theta$$

the price at time t of a bond paying the amount 1 at a time  $\theta$ . Denote also the shortrate process offered by a bank (i.e. the interest rate) by  $(R(t), t \ge 0)$ . A function  $f(t, \theta), 0 \le t \le \theta$  defined by the relation

$$P(t,\theta) = e^{-\int_t^\theta f(t,\eta)d\eta}, \qquad t \le \theta,$$

is called a forward rate function.

In Heath, Jarrow and Morton ([4]) it was assumed that

$$df(t,\theta) = \alpha(t,\theta)dt + \langle \sigma(t,\theta), dW(t) \rangle,$$

where W is a d-dimensional Wiener process with covariance Q. According to the observed data, the random function  $f(t, \theta)$  is regular in  $\theta$  for fixed t and chaotic in t for fixed  $\theta$ . The latter property is implied by the presence of W in the representation and the former is implied by the regular dependence of  $\alpha(t, \theta)$  and  $\sigma(t, \theta)$  on  $\theta$  for fixed t.

#### 4 Convergence to the Invariant Measure

For practical implementation of bond market models it is useful to replace the Wiener process W by a Lévy process L defined on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and taking values in a possibly infinite dimensional Hilbert space U. Thus we assume that the dynamics of the forward rate function are given by the equation

$$df(t,\theta) = \alpha(t,\theta)dt + \langle \sigma(t,\theta), dL(t) \rangle_U$$

for  $t \leq \theta$ .

Let

$$\hat{P}(t,\theta) := \exp\left\{-\int_0^t R(s)ds\right\} P(t,\theta), \qquad t \ge 0$$

be the discounted price of the bond. The fundamental theorem of asset pricing from [2] states that there are no arbitrage strategies (which is a key assumption in market models) if and only if there exists a probability measure  $\hat{\mathbb{P}}$  equivalent to the original one  $\mathbb{P}$ , such that  $\hat{P}(t,\theta), t \leq \theta$  is a local martingale on  $(\Omega, \mathcal{F}, \hat{\mathbb{P}})$ . In [5], [6], under mild assumptions a necessary and sufficient condition was given that ensures the discounted price process is in fact a local martingale with respect to the initial probability  $\mathbb{P}$ . This is the so-called HJM condition. In brief, under the assumption that the Lévy process has exponential moments, using the Lévy-Khinchin formula we may explicitly determine its Laplace transform. Indeed, for  $x \in U$ 

$$\mathbb{E}e^{-\langle x,L(t)\rangle_U} = e^{-t\tilde{\psi}(x)}$$

where  $\tilde{\psi}$  is explicitly given: if we define  $J := -\tilde{\psi}$  then

$$J(x) = -\langle a, x \rangle_U + \frac{1}{2} \langle Qx, x \rangle_U + J_0(x)$$

and

$$J_0(x) = \int_U \left( e^{-\langle x, y \rangle_U} - 1 + \langle x, y \rangle_U \, \mathbf{1}_{\{|y|_U \le 1\}} \right) \nu(dy).$$

The HJM condition requires that

$$\alpha(t,\theta) = \frac{d}{d\theta} J\left(\int_t^\theta \sigma(t,\eta) d\eta\right) = \left\langle DJ\left(\int_t^\theta \sigma(t,\eta) d\eta\right), \sigma(t,\theta) \right\rangle.$$
(4.9)

For more details about the HJM condition see also [11] section 20.2.

An important link between HJM modelling and stochastic partial differential equations is provided by the *Musiela parameterisation*. For  $t \ge 0, \xi \ge 0$  and  $u \in U$  define

$$r(t)(\xi) := f(t, t + \xi), a(t)(\xi) := \alpha(t, t + \xi), (b(t)u)(\xi) := < \sigma(t, t + \xi), u >_U.$$

We will call r the forward curve. Next let  $S(t)\varphi(\xi) = \varphi(\xi + t)$  be the shift semigroup. Then

$$\begin{aligned} r(t)(\xi) &= f(t,t+\xi) \\ &= f(0,t+\xi) + \int_0^t \alpha(s,t+\xi) ds + \int_0^t \langle \sigma(s,t+\xi), dL(s) \rangle_U \\ &= r(0)(t+\xi) + \int_0^t a(s)(t-s+\xi) ds + \int_0^t b(s)(t-s+\xi) dL(s) \\ &= S(t)r(0)(\xi) + \int_0^t S(t-s)a(s)(\xi) ds + \int_0^t S(t-s)b(s)(\xi) dL(s). \end{aligned}$$

Hence

$$r(t) = S(t)r(0) + \int_0^t S(t-s)a(s)ds + \int_0^t S(t-s)b(s)dL(s)$$

is a mild solution to the equation

$$dr(t) = \left(\frac{\partial}{\partial\xi}r(t) + a(t)\right)dt + b(t)dL(t),$$

where  $\frac{d}{d\xi}$  denotes the generator of  $(S(t), t \ge 0)$ . Identifying the  $L(U, \mathbb{R})$ -valued process  $b(\cdot)(\xi)$  with the corresponding U-valued process (denoted also by  $b(\cdot)(\xi)$ ) we have that under the HJM-condition (4.9)

$$dr(t)(\xi) = \left(\frac{\partial}{\partial\xi}r(t)(\xi) + \left\langle b(t)(\xi), DJ\left(\int_0^{\xi}b(t)(\zeta)d\zeta\right)\right\rangle_U\right)dt + b(t)(\xi)dL(t)$$
$$= \frac{\partial}{\partial\xi}\left(r(t)(\xi) + J\left(\int_0^{\xi}b(t)(\zeta)d\zeta\right)\right)dt + b(t)(\xi)dL(t).$$

Let the volatility b depend on the forward rate curve r according to, say,  $b(t)(\xi) = G(t, r(t))(\xi)$ , and let

$$F(t,r)(\xi) := \frac{\partial}{\partial \xi} J\left(\int_0^{\xi} G(t,r(t))(\zeta)d\zeta\right).$$

Then the forward curve process becomes a solution of the so-called *Heath-Jarrow-Morton-Musiela* (HJMM) equation

$$dr(t)(\xi) = \left(\frac{\partial}{\partial\xi}r(t)(\xi) + F(t,r(t))(\xi)\right)dt + G(t,r(t))(\xi)dL(t).$$
(4.10)

In [10], in the case where  $U = \mathbb{R}^d$  and where  $G(t, r(t))(\xi)[z] = \langle g(t, \xi, r(t)(\xi)), z \rangle$  with  $g : [0, \infty) \times [0, \infty) \times \mathbb{R} \to \mathbb{R}^d$  the following result was proven. We define  $\mathbf{H}_{\gamma} := H_{\gamma} \oplus \{\text{constant functions}\}$  where  $H_{\gamma} := L^2([0, \infty), \mathcal{B}([0, \infty)), e^{\gamma\xi}d\xi)$ . Note that  $\mathbf{H}_{\gamma}$  equipped with the scalar product  $\langle \psi + u, \varphi + v \rangle_{\mathbf{H}_{\gamma}} := \langle \psi, \varphi \rangle_{H_{\gamma}} + uv$  for  $\psi, \varphi \in H_{\gamma}, u, v \in \mathbb{R}$  is a real separable Hilbert space.

**Theorem 4.3.1.** Let L be an  $\mathbb{R}^d$ -valued square integrable zero-mean Lévy process with jump measure  $\nu$ , and let G be of the form described above. Assume that there exist functions  $\bar{g} \in H_{\gamma}$  and  $\bar{h} \in H_{\gamma} \cap L^{\infty}$  such that

- (i)  $\int_{\mathbb{R}} y^2 \exp\{|\bar{g}|_{L^1}|y|\}\nu(dy) < \infty,$
- (ii) for all  $t, \xi \in [0, \infty)$  and  $u, v \in \mathbb{R}$

$$|g(t,\xi,u)| \le \bar{g}(\xi), \qquad |g(t,\xi,u) - g(t,\xi,v)| \le \bar{h}(\xi)|u-v|.$$

Then, for for each  $r_0 \in \mathbf{H}_{\gamma}$  there is a unique solution to (4.10) in  $\mathbf{H}_{\gamma}$  satisfying  $r(0) = r_0$ . Moreover, if the coefficient g does not depend on t then (4.10) defines (time-homogeneous) Feller families on  $\mathbf{H}_{\gamma}$ .

We finally can say something about the invariant measure, using Theorem 4.2.1.

**Theorem 4.3.2.** Let G and F satisfy the assumptions of Theorem 4.3.1 for functions  $\bar{g}$  and  $\bar{h}$ . Define

$$K_1(J,\bar{g}) := \sup_{z:|z| \le |\bar{g}|_{L^1}} |DJ(z)|, \qquad K_2(J,\bar{g}) := \sup_{z:|z| \le |\bar{g}|_{L^1}} \left\| D^2 J(z) \right\|_{L(\mathbb{R}^d,\mathbb{R}^d)}$$

and let  $K := |\bar{h}|_{L^{\infty}} \left( 2K_2(J,\bar{g})|\bar{g}|_{H_{\gamma}}^2 + 2K_1(J,\bar{g}) \right)^{1/2}$ . Let  $\omega := \gamma - |\bar{h}|_{L^{\infty}}^2 - 2K^2 > 0$ . Then for any  $C \ge 0$ , there exists a unique invariant measure for (4.10) considered on  $H_{\gamma} + C$ , and it is exponentially mixing with exponent  $\omega/2$  and function c of linear growth.

*Proof.* Verify the conditions of Theorem 4.2.1. See [11] Chapter 20 for more details.  $\Box$ 

# 5 Appendix

### 5.1 Gaussian measures in Hilbert spaces

It is well known that if Q is a positive definite  $n \times n$  matrix with real entries, and  $m \in \mathbb{R}$  then the function

$$\frac{1}{((2\pi)^n \det Q)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}\langle Q^{-1}(x-m), x-m\rangle\right)$$

is the density of a Gaussian probability measure  $\mu$  on  $\mathbb{R}^n$  with mean m and covariance Q. Its characteristic function is given by

$$\hat{\mu}(\lambda) := \int_{\mathbb{R}^n} e^{i\langle\lambda,x\rangle} \mu(dx)$$
$$= e^{i\langle\lambda,m\rangle - \frac{1}{2}\langle Q\lambda,\lambda\rangle}.$$

We would like like to be able to extend this idea to an infinite dimensional separable Hilbert space, H.

**Definition 5.1.1.** A probability measure  $\mu$  on  $(H, \mathcal{B}(H))$  is Gaussian if for every  $h \in H$ there exists  $m \in \mathbb{R}$  and  $q \ge 0$  such that

$$\mu\{x \in H : \langle h, x \rangle_H \in A\} = \mathcal{N}(m, q)(A)$$

for all  $A \in \mathcal{B}(\mathbb{R})$ . A random variable X taking values in H is said to be Gaussian if its law is a Gaussian measure on H. A random process X taking values in H is Gaussian if for all  $t_1, \ldots, t_n$ ,  $(X(t_1), \ldots, X(t_n))$  is a Gaussian random element in  $H^n$ .

The definitions of the mean vector and the covariance matrix can be extended to the infinite dimensional case, thanks to the following theorem.

**Theorem 5.1.1.** Assume that X is a centred (i.e.  $\mathbb{E}X = 0$ ) Gaussian random variable with values in a Hilbert space H. Then  $\mathbb{E}|X|_{H}^{2} < \infty$ . Moreover,

$$\mathbb{E}e^{s|X|_H^2} \le \frac{1}{\sqrt{1 - 2s\mathbb{E}|X|_H^2}}, \qquad \forall s < \frac{1}{2\mathbb{E}|X|_H^2}$$

Proof. See Theorem 3.31 of [11].

It follows from this theorem, as in section 2.2 that for every centered Gaussian random variable X there exists a non-negative nuclear operator  $Q: H \to H$  called the *covariance* operator of X such that

$$\mathbb{E}\langle X, x \rangle_H \langle X, y \rangle_H = \langle Qx, y \rangle_H.$$

It is easy to see that Tr  $Q = \mathbb{E}|X|_{H}^{2}$ . More generally, if  $\mathbb{E}X = m$  then the covariance operator of X is the covariance operator of X - m.

It also follows that a Gaussian measure  $\mu$  on H with mean m and covariance Q has the following characteristic function

$$\hat{\mu}(\lambda) = e^{i\langle\lambda,m\rangle_H} e^{-\frac{1}{2}\langle Q\lambda,\lambda\rangle_H}$$

for  $\lambda \in H$ . It is therefore uniquely determined by m and Q, and is also denoted  $\mathcal{N}(m, Q)$ .

### 5.2 Lévy-Khinchin decomposition

Assume that L is a Lévy process which is right-continuous with left limits on a Hilbert space U. Let  $\Delta L(s) := L(t) - L(t-)$ . Given a Borel set A separated from 0, write

$$\pi_A(t) := \sum_{s \le t} \mathbf{1}_A(\Delta L(s)), \qquad t \ge 0.$$

Note that since L is right-continuous with left limits,  $\pi_A$  is  $\mathbb{Z}$ -valued. Clearly it is a Lévy process with jumps of size 1. Thus  $\pi_A$  is a Poisson process. Note also that  $\mathbb{E}\pi_A(t) = t\mathbb{E}\pi_A(1) = t\nu(A)$ , where  $\nu$  is a measure that is finite on sets separated from 0. Write

$$L_A(t) := \sum_{s \le t} \mathbf{1}_A \left( \Delta L(s) \right) \Delta L(s).$$

Then  $L_A$  is a well-defined Lévy process.

**Theorem 5.2.1** (Lévy-Khinchin decomposition). (i) If  $\nu$  is a jump intensity measure corresponding to a Lévy process then

$$\int_U \left( |y|_U^2 \wedge 1 \right) \nu(dy) < \infty.$$

(ii) Every Lévy process has the following representation:

$$L(t) = at + W(t) + \sum_{k=1}^{\infty} \left( L_{A_k}(t) - t \int_{A_k} y\nu(dy) \right) + L_{A_0}(t),$$

where  $A_0 := \{x : |x|_U \ge r_0\}$ ,  $A_k := \{x : r_k \le |x|_U < r_{k-1}\}$ ,  $(r_k)$  is an arbitrary sequence decreasing to 0, W is a Wiener process, all members of the representation are independent processes and the series converges  $\mathbb{P}$ -almost surely uniformly on each bounded subinterval of  $[0, \infty)$ .

It follows from the proof (see [11] Theorem 4.23 or [3]), that the processes

$$L_n(t) := L_{A_n}(t) - t \int_{A_n} y\nu(dy), \qquad t \ge 0,$$

are independent compensated coumpound Poisson processes. Hence we have the decomposition

$$L(t) = at + W(t) + \sum_{n=1}^{\infty} L_n(t) + L_0(t), \qquad t \ge 0,$$

where the processes  $W, L_n, n \ge 0$  and  $L_0$  are independent, W is a Wiener process,  $L_0$  is a compound Poisson process with jump intensity measure  $\mathbf{1}_{\{|y|_U \ge r_0\}}(y)\nu(dy)$  and each  $L_n$  is a compensated compound Poisson process with jump intensity measure

$$\mathbf{1}_{\{r_{n+1} \le |y|_U < r_n\}}(y)\nu(dy).$$

### 5.3 Krylov-Bogoliubov Theorem

Let  $(P_t)_{t\geq 0}$  be the transition function of a Markov process  $X = (X(t), t \geq 0)$  on a Banach space E.

**Definition 5.3.1.** A probability measure  $\mu$  is invariant with respect to the transition function  $(P_t)_{t>0}$  or invariant for X if, for any Borel set  $\Gamma \subset E$  and any  $t \ge 0$ ,

$$\mu(\Gamma) = \int_{H} \mu(dx) P_t(x, \Gamma).$$

If the initial position X(0) of X is a random variable with distribution  $\mu$  then the distribution of X(t) is equal to  $\mu$  for all  $t \ge 0$ . Thus one can expect that processes with invariant measures exhibit some kind of stability.

The following classical result provides a method of proving the existence of invariant measures for Feller semigroups.

**Theorem 5.3.1** (Krylov-Bogoliubov). Assume that  $(P_t)$  is a Feller transition semigroup on E (so that for all  $t \ge 0$ ,  $P_t$  maps bounded continuous functions to bounded continuous functions). Suppose also that there is an  $x \in E$  such that  $P_t(x, \cdot)$  converges weakly to a probability measure  $\mu$ . Then  $\mu$  is an invariant measure.

### 5.4 Yosida approximations

Let A be a closed, densely defined linear operator on a Banach space E. The resolvent set of A is  $\rho(A) = \{\alpha \in \mathbb{C} : \alpha I - A \text{ is invertible}\}$ . Denote the inverse  $(\alpha I - A)^{-1}$  by  $R(\alpha)$ . The family  $\{R(\alpha) : \alpha \in \rho(A)\}$  is called the resolvent of A.

**Theorem 5.4.1** (Hille-Yosida). (i) A densely defined closed operator A generates a  $C_0$ -semigroup S such that, for some  $\omega$  and M > 0,  $|S(t)z|_E \leq e^{\omega t}M|z|_E$  for all  $z \in E$  and  $t \geq 0$  if and only if  $(\omega, \infty) \subset \rho(A)$  and

$$||R^m(\alpha)||_{L(E,E)} \le \frac{M}{(\alpha - \omega)^m}, \qquad \forall \ m \in \mathbb{N}, \alpha > \omega.$$

Moreover if the above holds then

$$R(\alpha) = \int_0^\infty e^{-\alpha t} S(t) dt, \qquad \alpha > \omega,$$

and  $S(t)z = \lim_{\alpha \to \infty} e^{tA_{\alpha}}z$ , where  $A_{\alpha} = \alpha(\alpha R(\alpha) - I), \alpha > \omega$ .

(ii) If, for some  $z \in E$ ,  $A_{\alpha}z$  converges as  $\alpha \to \infty$ , then  $z \in D(A)$  and

$$\lim_{\alpha \to \infty} A_{\alpha} z = A z.$$

**Definition 5.4.1.** The operators  $(A_{\alpha})$  in the Hille-Yosida theorem are called the Yosida approximations of A.

### 5.5 Itô formula for Hilbert space valued semimartingales

First of all we need to define the quadratic variation process  $[M, M]_t$  of a general realvalued square integrable martingale M. We note that this is closely related the angle bracket process  $\langle M, M \rangle_t$  that appears in the Doob-Meyer decomposition (recall that this is defined as the unique increasing predictable process such that  $\langle M, M \rangle_0 = 0$ and  $M^2 - \langle M, M \rangle_t$  is a martingale), but is not equal to this process when the sample paths are not continuous. Its definition and properties are contained in the theorem below. For its proof we refer the reader to [8] Theorem 18.6.

**Theorem 5.5.1.** For every  $M \in \mathcal{M}^2$  there exists an increasing adapted process  $([M, M]_t, t \ge 0)$  which is right-continuous with left limits, called the quadratic variation of M, having the following properties.

(i) For every sequence  $\pi_n = (0 < t_0^n < t_1^n < \cdots)$  of partitions of  $[0, \infty)$  such that  $t_k^n \to \infty$  as  $k \to \infty$  and  $\lim_{n \to \infty} \sup_j (t_{j+1}^n - t_j^n) = 0$ , one has

$$[M, M]_t = \lim_{n \to \infty} \sum_j \left( M \left( t_{j+1}^n \wedge t_j^n \right) - M \left( t_j^n \wedge t \right) \right)^2,$$

where the limit is in  $L^1(\Omega, \mathcal{F}, \mathbb{P})$ .

- (ii)  $M^2 [M, M]$  is a martingale.
- (iii) If M has continuous trajectories then  $\langle M, M \rangle_t = [M, M]_t$ .

We can now state the Itô formula for a general Hilbert space valued valued semimartingale. This can be found in [8], or appendix D of [11]. First we will need some notation. Let H be a separable Hilbert space. For any process Y,  $\Delta Y(s) := Y(s) - Y(s)$ . Denote by  $H \otimes H$  the space  $H \otimes H$  completed with respect to the Hilbert-Schmidt norm. Clearly,  $L(H \otimes H, \mathbb{R}) \equiv L_{HS}(H, H)$ . Let  $\{e_n\}$  be an orthonormal basis of H and let Mbe a square integrable martingale taking values in H. Define  $M^k = \langle M, e_k \rangle_H$ , which is a real-valued square integrable martingale. We can then define

$$\llbracket M, M \rrbracket_s := \sum_{i,j} e_k \otimes e_j [M^k, M^j]_s,$$

which should be compared with the operator angle bracket defined in section 3.1. We also define the so-called *continuous part*  $[\![M, M]\!]_t^c$  of  $[\![M, M]\!]_t$  by

$$\llbracket M, M \rrbracket_t^c := \sum_{i,j} e_k \otimes e_j \left( [M^k, M^j]_t - \sum_{s \le t} \Delta M^k(s) M^j(s) \right).$$

**Theorem 5.5.2.** (Itô formula) Assume that X = M + A is a semimartingale taking values in a Hilbert space H. Let  $\psi : H \to \mathbb{R}$  be of class  $C^2$ . Assume that for each  $x \in$  $H, D^2\psi(x) \in L_{HS}(H, H)$  and the mapping  $x \mapsto D^2\psi(x) \in L_{HS}(H, H) \cong L(H \otimes H, \mathbb{R})$  is uniformly continuous on any bounded subset of H. Then  $\psi(X)$  is a local semimartingale and, for all  $t \ge 0$ ,  $\mathbb{P}$ -a.s.,

$$\begin{split} \psi(X(t)) &= \psi(X(0)) + \int_0^t \langle D\psi(X(s-)), dX(s) \rangle_H + \frac{1}{2} \int_0^t D^2 \psi(X(s-)) d[\![M, M]\!]_s \\ &+ \sum_{s \le t} \left\{ \Delta(\psi X)(s) - \langle D\psi(X(s-)), \Delta X(s) \rangle_H - Y(s) \right\} \end{split}$$

where

$$Y(s) := \frac{1}{2} D^2 \psi(X(s-)) \Delta X(s) \otimes \Delta X(s).$$

We also have

$$\begin{split} \psi(X(t)) &= \psi(X(0)) + \int_0^t \langle D\psi(X(s-)), dX(s) \rangle_H + \frac{1}{2} \int_0^t D^2 \psi(X(s-)) d\llbracket M, M \rrbracket_s^c \\ &+ \sum_{s \le t} \left\{ \Delta(\psi X)(s) - \langle D\psi(X(s-)), \Delta X(s) \rangle_H \right\}. \end{split}$$

If we let  $\pi_X((0,t],\Gamma) := \sum_{s \leq t} \mathbf{1}_{\Gamma}(\Delta X(s))$  be the jump measure of X, then the above formula becomes

$$\psi(X(t)) = \psi(X(0)) + \int_0^t \langle D\psi(X(s-)), dX(s) \rangle_H + \frac{1}{2} \int_0^t D^2 \psi(X(s-)) d\llbracket M, M \rrbracket_s^c + \int_0^t \int_H (\psi(X(s-)+y) - \psi(X(s-)) - \langle D\psi(X(s-)), y \rangle_H) \pi_X(ds, dy)$$

# 5 Appendix

# Bibliography

- A. CHOJNOWSKA-MICHALIK, On processes of Ornstein-Uhlenbeck type in Hilbert spaces, Stochastics, 21 (1987), pp. 251–286.
- [2] F. DELBAEN AND W. SCHACHERMAYER, A general version of the fundamental theorem of asset pricing, Math. Ann., 30 (1994), pp. 463–520.
- [3] I. I. GIKHMANN AND A. V. SKOROKHOD, *The Theory of Stochastic Processes*, vol. 2, Berin, Springer, 1974.
- [4] D. HEATH, R. JARROW, AND A. MORTON, Bond pricing and the term structure of interest rates: a new methodology, Econometrica, 61 (1992), pp. 77–105.
- [5] J. JAKUBOWSKI AND J. ZABCZYK, HJM condition for models with Lévy noise, IM PAN Preprint, 651, (2004).
- [6] <u>—</u>, *Exponential moments for HJM models with jumps*, Finance and Stochastics, to appear, (2007).
- [7] A. LASOTA AND J. A. YORKE, Lower bound technique for Markov operators and iterated function systems., Random Comput. Dynam., 2 (1994), pp. 41–77.
- [8] M. MÉTIVIER, Semimartingales: A Course on Stochastic Processes, De Gruyer, 1982.
- [9] K. R. PARTHASARATHY, Probability measures on metric spaces, Am. Math. Soc., 1967 (reprint 2005).
- [10] S. PESZAT, A. RUSINEK, AND J. ZABCZYK, On a stochastic partial differential equation of the bond market, Submitted, (2007).
- [11] S. PESZAT AND J. ZABCZYK, Stochastic Partial Differential Equations with Lévy Noise, Encyclopedia of Mathematics, Cambridge University Press, 2007.
- [12] G. D. PRATO AND J. ZABCZYK, Stochastic Equations in Infinite Dimensions, Encyclopedia of Mathematics, Cambridge University Press, 1992.
- [13] A. RUSINEK, Invariant measures for a class of stochastic evolution equations, Preprint IMPAN 667, Wardzawa, (2006).