# Stochastic Differential Equations Based on Lévy Processes and Stochastic Flows of Diffeomorphisms

Hiroshi Kunita

Department of Mathematical Sciences, Nanzan University, Seirei-cho, Seto, 489-0863, Japan
kunita@ms.nanzan-u.ac.jp

## Introduction

Continuous stochastic differential equations (SDE) based on Brownian motions have been studied a lot. Among them, pathwise properties of the solution such as the continuity, the differentiability and the diffeomorphic properties of the solution with respect to the initial state were studied in detail in the past two decades. Some of these results can be found in the author's book [13].

In the mean time, similar problems have been studied for SDEs of jump type based on Lévy processes or semimartingales with jumps. Results are not parallel to those of continuous SDEs. In some cases the diffeomorphic property of the solution may fail.

The purpose of this chapter is to expose basic facts about the solution of a certain SDE with jumps. It will be shown in Section 3 that the solution is differentiable with respect to the initial state if coefficients of the equation are smooth. However, the homeomorphic property or the diffeomorphic property is not always satisfied owing to the behavior of jumps. In Section 3.4, we will show that the solution defines a stochastic flow of homeomorphisms, if it makes a "homeomorphic" jump.

For the study of SDE with jumps, we need stochastic analysis of semimartingales with jumps. In Section 1, we discuss briefly stochastic integrals based on semimartingales and establish Itô's formula for semimartingales with jumps. These could be considered as a basis for stochastic analysis of processes with jumps. For more details, see Meyer [13], Jacod-Shiryaev [12] and Protter [20]. In Section 2, we study Lévy processes by applying results of Section 1. Among Lévy processes, Brownian motions and Poisson random measures play important roles in this work. We study these two with details in Section 2. For related problems, we refer to Ikeda–Watanabe [10] and Sato [21].

Section 3 is the main part of this chapter. We introduce a SDE with jumps. In order to make the discussion simple, we will restrict our attention to a SDE based on a Brownian motion and a Poisson random measure, though more general SDEs based on semimartingales (with spatial parameter) are studied in the literature (e.g., Fujiwara–Kunita [7,8,9], Carmona–Nualart [3], Applebaum–Tang [2]). We study the pathwise properties of the solutions such as the differentiability and the diffeomorphic property

M. M. Rao (ed.), *Real and Stochastic Analysis*

of the solution with respect to the initial state. For this purpose we obtain various types of $L^p$ estimates of the solution by applying Burkholder's inequality for stochastic integrals and then we apply Kolmogorov's criterion on the continuity of random fields. In Section 4 (Appendix), we discuss the Kolmogorov criterion or Kolmogorov–Totoki's theorem.

Most material of Section 3 is chosen from the joint works with Fujiwara, though some improvements are given here. The author expresses his gratitude to T. Fujiwara for his cooperative work on stochastic flows with jumps. Also he thanks D. Applebaum for pointing out errors in the first version of this article.

# 1 Stochastic integrals for semimartingales

## 1.1 Martingales, localmartingales and semimartingales

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. Suppose that we are given a family of sub $\sigma$-fields $\{\mathcal{F}_t\}, t \in [0, T]$ of $\mathcal{F}$ satisfying the following properties.
1) (Increasing) $\mathcal{F}_s \subset \mathcal{F}_t$ for any $s < t$.
2) (Right continuous) $\cap_{h>0}\mathcal{F}_{t+h} = \mathcal{F}_t$ holds for any $t$.
3) (Complete) Each $\mathcal{F}_t$ contains null sets of $\mathcal{F}$.
Then $\{\mathcal{F}_t\}, t \in [0, T]$ is called a (standard) filtration.

Let $X(t), t \in [0, T]$ be a real stochastic process. If $X(t)$ is $\mathcal{F}_t$-measurable for any $t \in [0, T]$, the process is called *adapted*. An adapted process $X(t)$ is called a *martingale* if $X(t)$ is integrable for any $t$ and equalities

$$E[X(t)|\mathcal{F}_s] = X(s), \quad a.s. \ \forall s < t$$

hold. If equality signs are replaced by $\geq$ in the above, it is called a *submartingale*. Further, if $-X(t)$ is a submartingale $X(t)$ is called a *supermartingale*. It is known that any submartingale $X(t)$ has a modification $\tilde{X}(t)$ whose sample paths are cadlag (right continuous with the left hand limits) a.s. Further if its sample paths are continuous a.s., it is called a *continuous martingale*. In the following we always consider cadlag martingales or submartingales.

Here we quote a useful inequality for martingales, which may be found in text books discussing martingale theory. See, e.g., Dellacherie–Meyer [4], Ikeda–Watanabe [10].

**Theorem 1.1 (Doob's inequality).** *Let $p > 1$ be any number. Let $X(t)$ be a martingale such that $E[|X(t)|^p] < \infty$. Then it holds*

$$E[\sup_{r\leq t} |X(r)|^p] \leq q^p E[|X(t)|^p], \quad \forall t \tag{1.1}$$

*where $q$ is a positive number such that $q^{-1} = 1 - p^{-1}$.*

A random variable $\tau$ with values in $[0, T]$ is called a *stopping time* if $\{\omega; \tau(\omega) \leq t\} \in \mathcal{F}_t$ holds for any $t$. We denote by $\mathcal{T}$ the set of all stopping times.

An adapted cadlag process $X(t)$ is called a *localmartingale* if there exists an increasing sequence of stopping times $\tau_n$ such that $P(\tau_n < T) \to 0$ as $n \to \infty$ and each stopped process $X^n(t) = X(t \wedge \tau_n)$ is a martingale. The corresponding sequence of stopping times $\{\tau_n\}$ is called the *reducing stopping times*. Further if each $X^n(t)$ is a square integrable martingale, it is called a *locally square integrable martingale*.

An adapted cadlag process $A(t)$ with $A(0) = 0$ is called an *increasing process* if it is increasing with respect to $t$ a.s. Further a process $A(t)$ is called a *process of finite variation* if it is written as the difference of two increasing processes. An adapted cadlag process $X(t)$ is called a *semimartingale* if it is written as the sum of a locally square integrable martingale and a process of finite variation.

The *predictable* $\sigma$*-field* $\mathcal{P}$ is the $\sigma$-field on $\Omega \times [0, T]$ generated by left continuous adapted processes. A $\mathcal{P}$-measurable stochastic process is called a *predictable process*.

**Theorem 1.2 (Doob–Meyer decomposition).** *Let $X(t)$ be a supermartingale such that the class of random variables $\{X(\tau); \tau \in \mathcal{T}\}$ is uniformly integrable. Then there exists a unique martingale $M(t)$ and an integrable predictable increasing process $A(t)$ such that*

$$X(t) = M(t) - A(t), \quad A(0) = 0.$$

For the proof, see Dellacherie–Meyer [4], Ikeda–Watanabe [10], Protter [20].

## 1.2 Stochastic integrals

Let $X(t)$ be a square integrable martingale. Then $X(t)^2$ is a nonnegative submartingale. Further, $\sup_{0 \le t \le T} X(t)^2$ is integrable by Doob's inequality. Then the class of random variables $\{X(\tau)^2; \tau \in \mathcal{T}\}$ is uniformly integrable. Therefore there exists a unique integrable predictable increasing process $A(t)$ such that $X(t)^2 - A(t)$ is a martingale by Doob–Meyer's decomposition theorem. We denote $A(t)$ by $\langle X \rangle_t$. Thus $X(t)^2 - \langle X \rangle_t$ is a martingale. Then it holds for any $s < t$,

$$E[(X(t) - X(s))^2 | \mathcal{F}_s] = E[\langle X \rangle_t - \langle X \rangle_s | \mathcal{F}_s], \quad a.s. \tag{1.2}$$

Conversely, suppose that for a given square integrable martingale $X(t)$, there exists a predictable increasing process $A(t)$ such that

$$E[(X(t) - X(s))^2 | \mathcal{F}_s] = E[A(t) - A(s) | \mathcal{F}_s], \quad a.s. \ \forall s < t.$$

Then $X(t)^2 - A(t)$ is a martingale, so that $A(t) = \langle X \rangle_t$ holds by the uniqueness of the Doob–Meyer decomposition of a supermartingale.

We set

$$L^2(\langle X \rangle) = \left\{ f(s); \text{predictable and } E\left[ \int_0^T |f(s)|^2 d\langle X \rangle_s \right] < \infty \right\}.$$

It is a Hilbert space with the norm $\|f\| = E\left[ \int_0^T |f(s)|^2 d\langle X \rangle_s \right]^{1/2}$. Let $f(s)$ be an adapted process. It is called a *simple predictable process* if it is written as

$$f(s) = \sum_i f_i 1_{(s_i, s_{i+1}]}(s),$$

where $0 = s_0 < s_1 < \cdots < s_n = T$ with $f_i$ bounded and $\mathcal{F}_{s_i}$-measurable. The set of all simple predictable processes $\mathcal{S}$ is dense in $L^2(\langle X\rangle)$.

We shall define the stochastic integral based on a square integrable martingale $X(t)$. Let $f(s)$ be a simple predictable process. We define

$$M(t) = \sum_{i=1}^{n} f_i(X(s_{i+1} \wedge t) - X(s_i \wedge t)),$$

and call it the stochastic integral of $f$ by $X$ and denote it by $\int_0^t f(s)dX(s)$.

**Lemma 1.3.** *The stochastic integral is a square integrable martingale. Further*

$$\Big\langle \int_0^{\cdot} f(s)dX(s)\Big\rangle_t = \int_0^t f(s)^2 d\langle X\rangle_s, \tag{1.3}$$

$$E\Big[\Big|\int_0^t f(s)dX(s)\Big|^2\Big] = E\Big[\int_0^t |f(s)|^2 d\langle X\rangle_s\Big]. \tag{1.4}$$

*Proof.* It can be verified directly that the above $M(t)$ is a square integrable martingale. If $s_{n_0} = s$ and $s_{n_1} = t$, we have the equality

$$E[M(t)^2|\mathcal{F}_s] - M(s)^2 = \sum_{i=n_0}^{n_1-1} E[E[(M(s_{i+1}) - M(s_i))^2|\mathcal{F}_{s_i}]|\mathcal{F}_s]$$
$$+2 \sum_{n_0\le i<j\le n_1-1} E[E[f_i f_j (X(s_{j+1}) - X(s_j))(X(s_{i+1}) - X(s_i))|\mathcal{F}_{s_i}]|\mathcal{F}_s].$$

It holds

$$\begin{aligned} E[(M(s_{i+1}) - M(s_i))^2|\mathcal{F}_{s_i}] &= f_i^2 E[(X(s_{i+1}) - X(s_i))^2|\mathcal{F}_{s_i}] \\ &= f_i^2 E[\langle X\rangle_{s_{i+1}} - \langle X\rangle_{s_i}|\mathcal{F}_{s_i}], \end{aligned}$$

and for $i < j$

$$\begin{aligned} &E[f_i f_j (X(s_{j+1}) - X(s_j))(X(s_{i+1}) - X(s_i))|\mathcal{F}_{s_i}] = \\ &\quad E[f_i f_j E[X(s_{j+1}) - X(s_j)|\mathcal{F}_{s_j}](X(s_{i+1}) - X(s_i))|\mathcal{F}_{s_i}] = 0. \end{aligned}$$

Therefore,

$$\begin{aligned} E[(M(t) - M(s))^2|\mathcal{F}_s] &= \sum_{n_0\le i<n_1-1} E[f_i^2(\langle X\rangle_{s_{i+1}} - \langle X\rangle_{s_i})|\mathcal{F}_s] \\ &= E[\int_s^t |f(u)|^2 d\langle X\rangle_u|\mathcal{F}_s], \end{aligned}$$

for any $s < t$. This implies $\langle M\rangle_t = \int_0^t f(u)^2 d\langle X\rangle_u$, proving the first equality of the lemma. The second equality is immediate from the first equality.

Next consider any $f \in L^2(\langle X \rangle)$. We can choose a sequence of simple predictable processes $\{f_n(s)\}$ such that $E[\int_0^T |f(s) - f_n(s)|^2 d\langle X \rangle_s] \to 0$. Let $M_n(t)$ and $M_m(t)$ be stochastic integrals of $f_n$ and $f_m$, respectively. Then we have by Doob's inequality

$$E[\sup_{t \le T} |M_n(t) - M_m(t)|^2] \le 4E[|M_n(T) - M_m(T)|^2]$$
$$= 4E[\int_0^T |f_n(s) - f_m(s)|^2 d\langle X \rangle_s].$$

It converges to 0 as $n, m \to \infty$. Then the sequence of martingales $\{M_n(t)\}$ converges uniformly in $t$ in $L^2$-sense. The $L^2$-limit $M(t)$ is a cadlag martingale. It holds

$$E[(M(t) - M(s))^2 | \mathcal{F}_s] = \lim_{n \to \infty} E[(M_n(t) - M_n(s))^2 | \mathcal{F}_s]$$
$$= \lim_{n \to \infty} E[\int_s^t f_n(u)^2 d\langle X \rangle_u | \mathcal{F}_s] = E[\int_s^t f(u)^2 d\langle X \rangle_u | \mathcal{F}_s].$$

Therefore we have $\langle M \rangle_t = \int_0^t f(u)^2 d\langle X \rangle_u$.

We denote the above $M(t)$ by $\int_0^t f(s) dX(s)$ and call it the *stochastic integral of* $f(s)$ *based on* $X(t)$. We have isometric properties (1.3) and (1.4).

Next suppose that $X(t)$ is a locally square integrable martingale and $\{\tau_n\}$ is a sequence of reducing stopping times. Then for each $X^n(t) = X(t \wedge \tau_n)$, there exists a unique predictable increasing process $\langle X^n \rangle_t$ such that $X^n(t)^2 - \langle X^n \rangle_t$ is a martingale. If $m < n$, we have $X^n(t \wedge \tau_m) = X^m(t)$. Then we have $\langle X^n \rangle_{t \wedge \tau_m} = \langle X^m \rangle_t$. Therefore there exists a predictable increasing process $\langle X \rangle_t$ such that $\langle X \rangle_{t \wedge \tau_n} = \langle X^n \rangle_t$. It follows that the process $X(t)^2 - \langle X \rangle_t$ is a localmartingale.

We shall extend the definition of the stochastic integrals to locally square integrable martingales. Let $X(t)$ be a locally square integrable martingale with a sequence of reducing stopping times $\{\tau_n\}$. Let $f(s)$ be a predictable process with the square integrability condition $\int_0^T |f(s)|^2 d\langle X \rangle_s < \infty$, a.s. Define

$$\sigma_n = \inf\{t \in [0, T]; \int_0^t |f(s)|^2 ds \ge n\} \wedge \tau_n,$$

and set $f^n(s) = f(s) 1_{\{s < \sigma_n\}}$. Then it holds $P(\sigma_n < T) \to 0$ as $n \to \infty$ and $E[\int_0^T |f^n(s)|^2 d\langle X^n \rangle_s] < \infty$ for any $n$. Now we can define stochastic integrals $M^n(t) := \int_0^t f^n(s) dX^n(s)$ as square integrable martingales. Further it holds for $m < n$ $M^n(t \wedge \sigma_m) = M^m(t)$. Then there exists a locally square integrable martingale $M(t)$ such that $M(t \wedge \sigma_n) = M^n(t)$. We denote the localmartingale $M(t)$ by $\int_0^t f(s) dX(s)$ and call it the stochastic integral of $f(s)$ by $X(t)$.

Now let $X(t)$ be a semimartingale decomposed as $X(t) = M(t) + A(t)$, where $M(t)$ is a locally square integrable martingale and $A(t)$ is a process of finite variation. The total variation process of $A(t)$ is denoted by $|A|(t)$. Suppose that $f(t)$ is a predictable process satisfying the integrability condition, $\int_0^T |f(s)|^2 d\langle M \rangle_s + \int_0^T |f(s)| d|A|(s) < \infty$. We may define the stochastic integral of $f$ by $X(t)$ by

$$\int_0^t f(s)dX(s) := \int_0^t f(s)dM(s) + \int_0^t f(s)dA(s),$$

where the last integral $\int_0^t f(s)dA(s)$ is the usual Stieltjes integral by the function of finite variation. Then the stochastic integral $\int_0^t f(s)dX(s)$ is a semimartingale for any $f$ with the above integrability condition.

Now let $X(t)$ be a semimartingale decomposed as $X(t) = M(t) + A(t)$, where $M(t)$ is a locally square integrable martingale and $A(t)$ is a process of finite variation, and let $f(t)$ be a cadlag adapted process. Then the left limit $f(t-) = \lim_{h\downarrow 0} f(t-h)$ is a predictable process and it satisfies the integrability condition mentioned above. Therefore the stochastic integral $\int_0^t f(s-)dX(s)$ is well defined. We shall approximate it by a sequence of finite sums. Let us denote by $\Pi$ a partition of $[0,T]$; $\Pi = \{0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T\}$. We set $|\Pi| = \max_i |t_{i+1} - t_i|$, and define

$$Y^{\Pi}(t) = \sum_i f(t_i \wedge t)(X(t_{i+1} \wedge t) - X(t_i \wedge t)). \tag{1.5}$$

**Theorem 1.4.** *Let $\{\Pi_k\}$ be a sequence of partitions such that $|\Pi_k| \to 0$. Then $\{Y^{\Pi_k}(t)\}$ converges to $\int_0^t f(s-)dX(s)$ uniformly in $t$ in probability, i.e.,*

$$\lim_{k\to\infty} P\left(\sup_{0\le t\le T}\left|Y^{\Pi_k}(t) - \int_0^t f(s-)dX(s)\right| > \epsilon\right) = 0$$

*is valid for any $\epsilon > 0$.*

*Proof.* It is sufficient to prove the theorem in the case where $X(t)$ is a locally square integrable martingale and the case where $X(t)$ is a process of finite variation. In the latter case, the stochastic integral is just equal to the pathwise Stieltjes integral. Then $Y^{\Pi_k}(t)$ converges to the stochastic integral $\int_0^t f(s-)dX(s)$ uniformly in $t$ a.s. as $k \to \infty$. So we consider the case where $X(t)$ is a locally square integrable martingale. Let $\{\tau_n\}$ be a sequence of reducing stopping times of $X(t)$. Set $\sigma_n = \inf\{t;\ |f(t)|^2 > n\} \wedge \tau_n$. Then $\{\sigma_n\}$ is a sequence of the reducing stopping times of $\int_0^t f(s-)dX(s)$. It holds

$$E\left[\sup_{0\le t\le T}\left|Y^{\Pi_k}(t\wedge\sigma_n) - \int_0^{t\wedge\sigma_n} f(s-)dX(s)\right|^2\right]$$
$$\le 4E\left[\int_0^{T\wedge\sigma_n} |f^{\Pi_k}(s-) - f(s-)|^2 d\langle X\rangle_s\right],$$

where $f^{\Pi_k}(s) = \sum_i f(s_i)1_{[s_i,s_{i+1})}(s)$. Then the above expectation converges to 0 as $k \to \infty$. On the other hand, it holds $P(\sigma_n < T) \to 0$ as $n \to \infty$. For any $\delta > 0$, choose $n$ such that $P(\sigma_n < T) < \delta$. Hence we have

$$P\left(\sup_{0\le t\le T}\left|Y^{\Pi_k}(t) - \int_0^t f(s-)dX(s)\right| > \epsilon\right)$$
$$\le P\left(\sup_{0\le t\le T}\left|Y^{\Pi_k}(t\wedge\sigma_n) - \int_0^{t\wedge\sigma_n} f(s-)dX(s)\right| > \epsilon, \sigma_n = T\right) + \delta.$$

The first term of the right hand side converges to 0 as $k \to \infty$ by Chebychev's inequality. Since $\delta$ is arbitrary, we get the assertion of the theorem.

## 1.3 Orthogonal martingales

For two square integrable martingales $X(t)$ and $Y(t)$, we define

$$\langle X, Y\rangle_t = \frac{1}{4}\Big\{\langle X+Y\rangle_t - \langle X-Y\rangle_t\Big\}.$$

Then $X(t)Y(t) - \langle X, Y\rangle_t$ is a martingale. We have obviously $\langle X, X\rangle_t = \langle X\rangle_t$. The process $\langle X, Y\rangle_t$ is called the bracket of $X$ and $Y$. The bracket satisfies

$$E[(X(t)-X(s))(Y(t)-Y(s))|\mathcal{F}_s] = E[\langle X, Y\rangle_t - \langle X, Y\rangle_s|\mathcal{F}_s], \quad \forall\, s < t.$$

We give a characterization of stochastic integrals by means of the bracket. We first show

**Lemma 1.5.** *Let $X, Y$ be square integrable martingales. Let $f, g$ be predictable processes belonging to $L^2(\langle X\rangle)$ and $L^2(\langle Y\rangle)$, respectively. Then $fg$ is integrable with respect to $\langle X, Y\rangle$. Further, we have*

$$\left|\int_0^t f(s)g(s)d\langle X, Y\rangle_s\right| \le \left(\int_0^t |f(s)|^2 d\langle X\rangle_s\right)^{1/2}\left(\int_0^t |g(s)|^2 d\langle Y\rangle_s\right)^{1/2}. \quad (1.6)$$

*Proof.* We first observe that for any $t > s$, the bracket $\langle X, Y\rangle_t - \langle X, Y\rangle_s$ is a positive bilinear form a.s. Then we have by Schwarz's inequality,

$$|\langle X, Y\rangle_t - \langle X, Y\rangle_s| \le (\langle X\rangle_t - \langle X\rangle_s)^{1/2}(\langle Y\rangle_t - \langle Y\rangle_s)^{1/2}.$$

Now suppose that both $f, g$ are simple predictable processes of the form $\sum_i f_i 1_{(s_i, s_{i+1}]}(s)$ and $\sum_i g_i 1_{(s_i, s_{i+1}]}(s)$, respectively, where $f_i, g_i$ are bounded $\mathcal{F}_{s_i}$-measurable random variables. Then we have

$$\begin{aligned}
&\left|\sum_i f_i g_i(\langle X, Y\rangle_{s_{i+1}} - \langle X, Y\rangle_{s_i})\right| \\
&\le \sum_i (f_i^2(\langle X\rangle_{s_{i+1}} - \langle X\rangle_{s_i}))^{1/2}(g_i^2(\langle Y\rangle_{s_{i+1}} - \langle Y\rangle_{s_i}))^{1/2} \\
&\le \{\sum_i f_i^2(\langle X\rangle_{s_{i+1}} - \langle X\rangle_{s_i})\}^{1/2}\{\sum_i g_i^2(\langle Y\rangle_{s_{i+1}} - \langle Y\rangle_{s_i})\}^{1/2}.
\end{aligned}$$

This proves the inequality (1.6) in the case where $f, g$ are simple predictable processes. We can show the inequality for general $f, g$ by approximating them by sequences of simple predictable processes.

**Theorem 1.6.** *Let $X, Y$ be square integrable martingales and let $f \in L^2(\langle X\rangle)$. Then we have*

$$\left\langle \int f dX, Y \right\rangle_t = \int_0^t f(s) d\langle X, Y\rangle_s. \tag{1.7}$$

*Conversely suppose that a square integrable martingale $N(t)$ satisfies*

$$\langle N, Y\rangle_t = \int_0^t f(s) d\langle X, Y\rangle_s$$

*for any square integrable martingale $Y$. Then $N(t) = N(0) + \int_0^t f(s)dX(s)$ holds valid.*

*Proof.* Suppose that $f$ is a simple process given by $f = \sum_i f_i 1_{(s_i, s_{i+1}]}$, where $s_{n_0} = s$ and $s_{n_1} = t$. Set $M(t) = \int_0^t f(s)dX(s)$. Then we have

$$\begin{aligned}
&E[(M(t) - M(s))(Y(t) - Y(s))|\mathcal{F}_s] \\
&= \sum_{n_0 \le k < n_1} E[E[(M(s_{k+1}) - M(s_k))(Y(s_{k+1}) - Y(s_k))|\mathcal{F}_{s_k}]|\mathcal{F}_s] \\
&= \sum_{n_0 \le k < n_1} E[f_k E[(X(s_{k+1}) - X(s_k))(Y(s_{k+1}) - Y(s_k))|\mathcal{F}_{s_k}]|\mathcal{F}_s] \\
&= \sum_{n_0 \le k < n_1} E[f_k E[\langle X, Y\rangle_{s_{k+1}} - \langle X, Y\rangle_{s_k}|\mathcal{F}_{s_k}]|\mathcal{F}_s] \\
&= E[\sum_{n_0 \le k < n_1} f_k(\langle X, Y\rangle_{s_{k+1}} - \langle X, Y\rangle_{s_k})|\mathcal{F}_s] \\
&= E[\int_s^t f(u) d\langle X, Y\rangle_u |\mathcal{F}_s].
\end{aligned}$$

This proves (1.7) in the case where $f$ is a simple predictable process. The equality can be extended to any $f \in L^2(\langle X\rangle)$ by using Lemma 1.5, since simple predictable processes are dense in $L^2(\langle X\rangle)$.

For the proof of the latter assertion, observe that $M - N$ satisfies $\langle M - N, Y\rangle_s = 0$, for all square integrable martingales $Y$. Then setting $Y = M - N$, we find that $\langle M - N\rangle = 0$, proving $M - N =$ constant. The proof is complete.

We denote by $\mathcal{M}$ the set of all square integrable martingales $X(t)$ such that $X(0) = 0$. For each $X \in \mathcal{M}$, we define an $L^2$ norm by $\|X\| = E[|X(T)|^2]^{1/2}$. It is a Hilbert space.

Let $\mathcal{N}$ be a subset of $\mathcal{M}$. It is called a *stable subspace* of $\mathcal{M}$ if it satisfies the following.

1) $\mathcal{N}$ is a closed subspace of $\mathcal{M}$ as a vector space.
2) For any $X \in \mathcal{N}$ and $f \in L^2(\langle X\rangle)$, the stochastic integral $\int f dX$ belongs to $\mathcal{N}$.

For a given $X \in \mathcal{M}$, we set

$$\mathcal{L}(X) = \left\{ \int_0^t f(s)dX(s) : \ f \in L^2(\langle X\rangle) \right\}.$$

Then it is a stable subspace of $\mathcal{M}$. The set $\mathcal{L}(X)$ is called the *stable subspace generated by* $X$.

Let $X, Y$ be two square integrable martingales. These are called *orthogonal* if $\langle X, Y\rangle_t \equiv 0$ or equivalently the product $X(t)Y(t)$ is a martingale. For a given stable subspace $\mathcal{N}$, we set

$$\mathcal{N}^{\perp} = \{Y \in \mathcal{M} : \quad \langle Y, X\rangle = 0 \ \ \forall X \in \mathcal{N}\}.$$

It is clearly a closed vector space. Further it is also stable. Indeed, if $Y \in \mathcal{N}^{\perp}$ and $g \in L^2(\langle Y\rangle)$, then $\langle \int gdY, X\rangle = \int g(s)d\langle Y, X\rangle_s = 0$ holds for any $X \in \mathcal{N}$. Therefore $\int gdY \in \mathcal{N}^{\perp}$.

We will prove that an arbitrary given element $Y$ of $\mathcal{M}$ is decomposed uniquely to the sum of $Y_1$ and $Y_2$, where $Y_1 \in \mathcal{L}(X)$ and $Y_2$ is orthogonal to $\mathcal{L}(X)$. We first prepare a lemma:

**Lemma 1.7.** *For any given $X, Y \in \mathcal{M}$, there exists a unique $f \in L^2(\langle X\rangle)$ satisfying*

$$\langle X, Y\rangle = \int f d\langle X\rangle.$$

*Proof.* Let $A$ be a set in $\mathcal{B}([0, T]) \times \mathcal{F}$ such that the indicator function $1_A$ is predictable. Suppose that $\int 1_A d\langle X\rangle = 0$ a.s. Then we have $\int 1_A d\langle X, Y\rangle = 0$ by Lemma 1.5. Therefore, the process of finite variation $\langle X, Y\rangle_t(\omega)$ is absolutely continuous with respect to the increasing function $\langle X\rangle_t(\omega)$, for almost all $\omega$. The Radon–Nikodym density $f(t, \omega)$ can be chosen as a predictable process. We will prove that the above $f(t)$ belongs to $L^2(\langle X\rangle)$. Let $c$ be a positive constant. We set $f^c(s) = f(s)$ if $|f(s)| \le c$ and $f^c(s) = 0$ if $|f(s)| > c$. Since $|f^c(s)|^2 = f^c(s)f(s)$ holds, we have

$$\int |f^c(s)|^2 d\langle X\rangle_s = \int f^c(s)\langle X, Y\rangle_s \le \left(\int |f^c(s)|^2 d\langle X\rangle_s\right)^{1/2} \langle Y\rangle^{1/2}.$$

Therefore we get $\int |f^c(s)|^2 d\langle X\rangle_s \le \langle Y\rangle$. Since $c$ is arbitrary, we have $\int |f(s)|^2 d\langle X\rangle_s \le \langle Y\rangle$, showing $f \in L^2(\langle X\rangle)$. The uniqueness of $f$ will be obvious.

**Proposition 1.8.** *Let $X, Y$ be any elements of $\mathcal{M}$. Then $Y$ is decomposed uniquely to the sum of $Y_1 \in \mathcal{L}(X)$ and $Y_2$ which is orthogonal to $\mathcal{L}(X)$.*

*Proof.* Let $f(s)$ be the predictable process of Lemma 1.7. Set $Y_1 = \int_0^t f(s)dX(s)$ and $Y_2 = Y - Y_1$. Then $Y_1 \in \mathcal{L}(X)$. Further, $\langle Y_1, X\rangle = \int f(s)d\langle X\rangle = \langle Y, X\rangle$. Therefore we have $\langle Y_2, X\rangle = 0$. We have thus shown the existence of the orthogonal decomposition. The uniqueness will be obvious.

We denote $Y_1$ of the proposition by $P_{\mathcal{L}(X)}Y$ and call it the *orthogonal projection* of $Y$ to $\mathcal{L}(X)$.

Now suppose that we are given a sequence of martingales $X_1, \ldots, X_n$ of $\mathcal{M}$. We define a sequence of the orthogonal martingales by the Gram–Schmidt's orthogonalization method;

$$Y_1(t) = X_1(t), \quad Y_2(t) = X_2(t) - P_{\mathcal{L}(Y_1)}X_2, \ldots,$$
$$Y_n(t) = X_n(t) - \sum_{i=1}^{n-1} P_{\mathcal{L}(Y_i)}X_n.$$

Then these $Y_i, i = 1, \ldots, n$ are orthogonal martingales.

An orthogonal system $\{Y_n, n = 1, 2, \ldots\}$ is called an orthogonal base of $\mathcal{M}$ if $Y_n \neq 0$ for any $n$ and every $X$ is represented by $X = \sum_n P_{\mathcal{L}(Y_n)}X$.

For an arbitrary stable subspace $\mathcal{N}$, there exists a (at most) countable orthogonal basis $\{Y_n\}$. This can be shown similarly to the existence of orthogonal bases of $\mathcal{M}$. Now let $X$ be an arbitrary element of $\mathcal{M}$. We set $X_1 = \sum_n P_{\mathcal{L}(Y_n)}X$ and $X_2 = X - X_1$. Then $X_1 \in \mathcal{N}$ and $X_2 \in \mathcal{N}^\perp$. Therefore $X$ has the orthogonal decomposition $X = X_1 + X_2$. We can write the orthogonal decomposition as $\mathcal{M} = \mathcal{N} \oplus \mathcal{N}^\perp$.

Let $\mathcal{M}_c$ be the set of all continuous martingales $X(t)$ in $\mathcal{M}$. Then stochastic integrals $\int_0^t f(s)dX(s)$ for $X \in \mathcal{M}_c$ and $f \in L^2(\langle X \rangle)$ are continuous martingales. Therefore $\mathcal{M}_c$ is a stable subspace of $\mathcal{M}$. We set $\mathcal{M}_d = \mathcal{M}_c^\perp$ and call elements of $\mathcal{M}_d$ *purely discontinuous martingales*. Thus we have the orthogonal decomposition

$$\mathcal{M} = \mathcal{M}_c \oplus \mathcal{M}_d,$$

and any element $X$ of $\mathcal{M}$ is written as the sum of a continuous martingale $X_c \in \mathcal{M}_c$ and a purely discontinuous one $X_d \in \mathcal{M}_d$. Such a decomposition is unique.

### 1.4 Quadratic variations and Stratonovich integrals

Let $X(t)$ be a semimartingale. Then $X(t)$ is a cadlag process and $X(s-) = \lim_{h\downarrow 0} X(s-h)$ exists. Then the stochastic integral $\int_0^t X(s-)dX(s)$ is well defined as a semimartingale. We define the quadratic variation of the semimartingale $X(t)$ by

$$[X]_t = X(t)^2 - 2\int_0^t X(s-)dX(s) - X(0)^2. \tag{1.8}$$

Co-quadratic variation of two semimartingales $X, Y$ is defined by

$$[X, Y]_t = \frac{1}{4}\Big\{[X+Y] - [X-Y]\Big\}. \tag{1.9}$$

For a partition $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = T\}$, we set

$$[X, Y]_t^\Pi = \sum_{i=1}^{n-1}(X(t_{i+1} \wedge t) - X(t_i \wedge t))(Y(t_{i+1} \wedge t) - Y(t_i \wedge t)). \tag{1.10}$$

**Lemma 1.9.** *Let $\{\Pi_n\}$ be a sequence of partitions such that $|\Pi_n| \to 0$. Then we have*

$$\lim_{n\to\infty} [X, Y]_t^{\Pi_n} = [X, Y]_t$$

*uniformly in $t \in [0, T]$ in probability.*

*Proof.* We give the proof in the case where $X = Y$ only. We will apply Theorem 1.4. Let $\{Y^{\Pi_n}(t)\}$ be a sequence of stochastic integrals which approximates the integral $\int_0^t X(s-)dX(s)$. It holds $X(t)^2 - X(0)^2 - 2Y^{\Pi_n}(t) = [X]_t^{\Pi_n}$. Let $n$ tend to infinity. Then we find that $[X]_t^{\Pi_n}$ converges uniformly in probability and the limit is equal to $X(t)^2 - X(0)^2 - 2\int_0^t X(s-)dX(s)$. Therefore, $\lim_{n\to\infty}[X]_t^{\Pi_n}$ exists in probability and it coincides with $[X]$.

From the above lemma, the quadratic variation $[X]_t$ is an increasing process and the quadratic co-variation $[X, Y]$ is a process of finite variation. Then the latter is written as the sum of a continuous process of finite variation denoted by $[X, Y]_t^c$ and a purely discontinuous process of finite variation denoted by $[X, Y]_t^d$. The decomposition is unique. $[X, Y]_t^c$ is called the *continuous part* of $[X, Y]_t$.

**Proposition 1.10.** *Let $X(t)$ be a square integrable martingale. If $X(t)$ is a continuous martingale, then we have $[X]_t = \langle X\rangle_t$. If $X(t)$ is purely discontinuous, then we have $[X]_t = \sum_{s\le t}(\Delta X(s))^2$. Generally, let $X(t) = X_c(t) + X_d(t)$ be the orthogonal decomposition such that $X_c \in \mathcal{M}_c$ and $X_d \in \mathcal{M}_d$. Then it holds*

$$[X]_t = \langle X_c\rangle_t + \sum_{s\le t}(\Delta X(s))^2. \tag{1.11}$$

*Proof.* If $X(t)$ is a continuous martingale, then the stochastic integral $\int_0^t X(s-)dX(s)$ is a continuous martingale. Therefore the quadratic variation $[X]_t$ is a continuous increasing process. Further, since $X(t)^2 - [X]_t$ is a martingale, $\langle X\rangle_t - [X]_t$ is also a martingale. However a continuous martingale with finite variation is a constant, proving $\langle X\rangle_t - [X]_t \equiv 0$. We omit the case where $X(t)$ is purely discontinuous. See Lemma I.4.51 in Jacod–Shiryaev [12].

We will prove (1.11). Since $\langle X_c, X_d\rangle_t = 0$, we have $[X_c, X_d]_t = 0$. Then we get

$$[X]_t = [X_c]_t + [X_d]_t = \langle X_c\rangle_t + \sum_{s\le t}(\Delta X(s))^2.$$

**Remark.** If $X(t)$ and $Y(t)$ are continuous locally square integrable martingales, then these two are orthogonal if and only if the quadratic co-variation $[X, Y]$ is 0. However it is not the case if both of $X, Y$ have jumps. Indeed, if both $X$ and $Y$ are purely discontinuous, $[X, Y] = 0$ holds if and only if $X, Y$ have no common jumps, i.e., $\Delta X(s)\Delta Y(s) = 0$ a.s. for all $s$. In this case $X$ and $Y$ are orthogonal. However the orthogonality of $X, Y$ does not imply that they have no common jumps.

Let $X(t)$ and $f(t)$ be semimartingales. We define the *Stratonovich integral* of $f(t)$ based on $X(t)$ by

$$\int_0^t f(s)\circ dX(s) := \int_0^t f(s-)dX(s) + \frac{1}{2}[f, X]_t.$$

For a partition $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = T\}$, set

$$Z^{\Pi}(t) = \sum_i \frac{1}{2}\Big(f(t_{i+1}\wedge t) + f(t_{t_i}\wedge t)\Big)(X(t_{i+1}\wedge t) - X(t_i\wedge t)).$$

**Theorem 1.11.** *Let $\{\Pi_n\}$ be a sequence of partitions of $[0, T]$ such that $|\Pi_n| \to 0$. Then it holds*

$$\exists \lim_{n\to\infty} Z^{\Pi_n}(t) = \int_0^t f(s) \circ dX(s).$$

*Proof.* Let $Y^{\Pi_n}(t)$ be the process defined by (1.5). Then we have the relation $Z^{\Pi_n}(t) = Y^{\Pi_n}(t) + \frac{1}{2}[f, X]_t^{\Pi_n}$. Let $|\Pi_n| \to 0$. Then we obtain

$$\lim_{n\to\infty} Z^{\Pi_n}(t) = \int_0^t f(s-)dX(s) + \frac{1}{2}[f, X]_t.$$

by Theorem 1.4 and Lemma 1.9.

## 1.5 Itô's formula I

**Theorem 1.12 (Itô's formula).** *Let $X(t) = (X^1(t), ..., X^d(t))$ be a $d$-dimensional semimartingale and let $F(x_1, ..., x_d)$ be a $C^2$ function. Then $F(X^1(t), ..., X^d(t))$ is again a semimartingale, and the following formula holds.*

$$\begin{aligned} &F(X(t)) - F(X(0)) = \qquad\qquad (1.12)\\ &\sum_i \int_0^t \frac{\partial F}{\partial x_i}(X(s-))dX_s^i + \frac{1}{2}\sum_{i,j}\int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}(X(s-))d[X^i, X^j]_s^c \\ &\quad + \sum_{0<s\le t}\left\{F(X(s)) - F(X(s-)) - \sum_i \frac{\partial F}{\partial x_i}(X(s-))\Delta X_s^i\right\}. \end{aligned}$$

*Here $[X^i, X^j]_t^c$ is the continuous part of $[X^i, X^j]_t$.*

**Remark.** 1) The infinite sum of the last term of Itô's formula is absolutely convergent, because

$$\begin{aligned} &\sum_{0<s\le t}\left|F(X(s)) - F(X(s-)) - \sum_i \frac{\partial F}{\partial x_i}(X(s-))\Delta X_s^i\right| \\ &\le \frac{1}{2}\sum_{ij}\sum_{0\le s\le t}\left|\int_0^1 (1-\theta)\frac{\partial^2 F}{\partial x_i \partial x_j}(X(s-) + \theta\Delta X(s)d\theta)\right| \\ &\quad |\Delta X^i(s)\Delta X^j(s)| < \infty. \end{aligned}$$

2) Here is another expression of Itô's formula:

$$\begin{aligned} &F(X(t)) - F(X(0)) = \qquad\qquad (1.13)\\ &\sum_i \int_0^t \frac{\partial F}{\partial x_i}(X(s-))dX_s^i + \frac{1}{2}\sum_{i,j}\int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}(X(s-))d[X^i, X^j]_s \\ &\quad + \sum_{0<s\le t}\left\{F(X(s)) - F(X(s-)) - \sum_i \frac{\partial F}{\partial x_i}(X(s-))\Delta X_s^i \right.\\ &\quad \left. -\frac{1}{2}\sum_{i,j}\sum_{s\le t}\frac{\partial^2 F}{\partial x_i \partial x_j}(X(s-))\Delta X^i(s)\Delta X^j(s)\right\}. \end{aligned}$$

In fact, the last infinite sum of (1.13) can be divided into two parts. One is the same one as in (1.12). The other is the term involving $\Delta X^i(s)\Delta X^j(s)$. Sum up this term with the second term involving $[X^i, X^j]_t$. Then we get the second term on the right hand side of (1.12).

*Proof.* We give the proof in the case $d = 1$. We will prove (1.13). We assume $F''$ is bounded and uniformly continuous in $\mathbf{R}$. (If $F''$ is not bounded, consider the semi-martingales $X(t)1_{[0,\tau_n)}(t)$, where $\tau_n = \inf\{t; |X(t)| \geq n\}$.)

Let $\{\Pi_n = \{0 = t_1^n < \cdots < t_{k_n}^n = t\}\}$ be a sequence of partitions of the interval $[0, t]$ such that $|\Pi_n| \to 0$ as $n \to \infty$. We have

$$\begin{aligned} F(X(t)) - F(X(0)) &= \sum_i \{F(X(t_{i+1}^n) - F(X(t_i^n))\} \\ &= \sum_i F'(X(t_i^n))(X(t_{i+1}^n) - X(t_i^n)) \\ &\quad + \frac{1}{2}\sum_i F''(X(t_i^n))(X(t_{i+1}^n) - X(t_i^n))^2 \\ &\quad + \sum_i R(X(t_{i+1}^n), X(t_i^n)), \end{aligned}$$

where we have used Taylor's formula

$$\begin{aligned} F(y) - F(x) &= F'(x)(y-x) + \frac{1}{2}F''(x)(y-x)^2 + R(x, y), \\ R(x, y) &= \left(\int_0^1 (1-\theta)F''(x + \theta(y-x))d\theta\right)(y-x)^2 \\ &\quad - \frac{1}{2}F''(x)(y-x)^2. \end{aligned}$$

It holds

$$\lim_{n\to\infty} \sum_i F'(X(t_i^n))(X(t_{i+1}^n) - X(t_i^n)) = \int_0^t F'(X(s-))dX(s),$$

by Theorem 1.4 and

$$\lim_{n\to\infty} \frac{1}{2}\sum_i F''(X(t_i^n))(X(t_{i+1}^n) - X(t_i^n))^2 = \frac{1}{2}\int_0^t F''(X(s-))d[X]_s,$$

by Lemma 1.9. We will prove

$$\lim_{n\to\infty} \sum_i R(X(t_{i+1}^n), X(t_{t_i}^n)) = \sum_{0\leq s\leq t} r(X(s), X(s-))(\Delta X(s))^2, \tag{1.14}$$

where $r(x, y) = \frac{1}{2}\{\int_0^1 F''(x + \theta(y-x))d\theta - F''(x)\}$. Given $\epsilon > 0$, we set $J(\epsilon) = \{s \in [0, t]; |\Delta X(s)| > \epsilon\}$. It is a finite set a.s. Then we have

$$\lim_{n\to\infty} \sum_{i;J(\epsilon)\cap(t_i^n,t_{i+1}^n]\neq\phi} R(X(t_{i+1}^n), X(t_{t_i}^n)) = \sum_{s\in J(\epsilon)} r(X(s), X(s-))\Delta X(s)^2.$$

On the other hand, observe the inequality

$$\left| \sum_{i;J(\epsilon)\cap(t_i^n,t_{i+1}^n]=\phi} R(X(t_{i+1}^n), X(t_{t_i}^n)) \right|$$
$$\leq \left( \sup_{i;J(\epsilon)\cap(t_i^n,t_{i+1}^n]=\phi} |r(X(t_{i+1}^n), X(t_i^n))| \right) \sum_i (X(t_{i+1}^n) - X(t_i^n))^2.$$

Therefore,

$$\limsup_{n\to\infty} \left| \sum_{i;J(\epsilon)\cap(t_i^n,t_{i+1}^n]=\phi} R(X(t_{i+1}^n), X(t_{t_i}^n)) \right| \leq \sup_{|x-y|\leq\epsilon} |r(x,y)|[X]_t.$$

It tends to 0 as $\epsilon \to 0$. Therefore we get (1.14).

Now we have

$$r(x,y)(y-x)^2 = F(y) - F(x) - F'(x)(y-x) - \frac{1}{2}F''(x)(y-x)^2.$$

Then we get

$$\sum_{0<s\leq t} (X(s), X(s-))(\Delta X(s))^2$$
$$= \sum_{0<s\leq t} \left\{ F(X(s)) - F(X(s-)) - F'(X(s-))\Delta X(s) - \right.$$
$$\left. \frac{1}{2}F''(X(s-))(\Delta X(s))^2 \right\}.$$

Therefore we get the formula (1.13).

**Complex martingales.** Finally in this section we shall briefly discuss complex martingales. Let $Z(t)$ be an adapted cadlag process with values in **C**, the set of complex numbers. Let $X(t)$ and $Y(t)$ be the real part and the imaginary part of $Z(t)$, respectively. These are real adapted cadlag processes such that $Z(t) = X(t) + iY(t)$. $Z(t)$ is called a complex martingale if both of $X(t)$ and $Y(t)$ are martingales. Further $Z(t)$ is called a square integrable martingale if $|Z(t)|^2 = X(t)^2 + Y(t)^2$ is integrable. Complex localmartingales and complex semimartingales are defined similarly.

Now let $Z(t)$ and $\tilde{Z}(t) = \tilde{X}(t) + i\tilde{Y}(t)$ be square integrable complex martingales. The bracket of $Z(t)$ and $\tilde{Z}(t)$ is defined by

$$\langle Z, \tilde{Z}\rangle_t = \langle X, \tilde{X}\rangle_t + i\langle X, \tilde{Y}\rangle_t + i\langle Y, \tilde{X}\rangle_t - \langle Y, \tilde{Y}\rangle_t.$$

Then $Z(t)\tilde{Z}(t) - \langle Z, \tilde{Z}\rangle_t$ is a complex martingale. Note that $\langle Z, \bar{Z}\rangle_t = \langle X\rangle_t + \langle Y\rangle_t$ holds, where $\bar{Z}_t = X(t) - iY(t)$ is the complex conjugate of $Z(t)$. Let $f(t)$ be a complex predictable process such that $\int_0^T |f(s)|^2 d\langle Z, \bar{Z}\rangle_t < \infty$. Then we can define the stochastic integral $\int_0^t f(s)dZ(s)$ as a complex localmartingale. Theorem 1.6 is valid for complex stochastic integrals.

We can define the quadratic co-variation $[Z, \tilde{Z}]_t$ of two complex semimartingales $Z(t)$ and $\tilde{Z}(t)$ similarly to the bracket $\langle Z, \tilde{Z}\rangle_t$.

We can extend Itô's formula to complex $C^2$ functions and to complex semimartingales. The formula (1.12) is valid for a complex $C^2$ function $F$ and for a complex semimartingale $X(t)$.

## 2 Stochastic analysis of Lévy processes

### 2.1 Lévy processes and Poisson random measures

Let $Z(t), t \in [0, T]$ be an $m$-dimensional cadlag process defined on a probability space $(\Omega, \mathcal{F}, P)$. It is called a *Lévy process* if it has the following three properties.
1) (*Independent increments*) For any $0 \le t_0 < t_1 < \cdots < t_n \le T$, random variables $Z(t_i) - Z(t_{i-1}), i = 1, ..., n$, are independent.
2) (*Time homogeneous*) Laws of $Z(t + h) - Z(s + h)$ do not depend on $h$.
3) (*Continuous in probability*) For any $t$ it holds $\lim_{h\to 0} P(|Z(t+h) - Z(t)| > \epsilon) = 0$ for any $\epsilon > 0$.

In this work, we assume $Z(0) = 0$ for simplicity.

Given a Lévy process $Z(t)$, we define a family of sub $\sigma$-fields of $\mathcal{F}$ by

$$\mathcal{F}'_t := \cap_{\epsilon>0}\sigma(Z(s) : 0 \le s \le t + \epsilon).$$

It is called the *filtration generated by the Lévy process* $Z(t)$.

Suppose that a Lévy process $Z(t)$ is integrable for any $t$. Its mean vector is proportional to $t$, which we denote by $tm$. Then $M(t) = Z(t) - tm$ is a (vector) martingale with respect to the filtration $\{\mathcal{F}'_t\}$. In fact, $M(t) - M(s)$ is independent of $\mathcal{F}'_s$ and its expectation is 0. Suppose further that $Z(t)$ is square integrable. Its covariance matrix is proportional to $t$. It is written by $tV$, where $V = (v_{ij})$ is a nonnegative definite symmetric matrix. In this case $M(t) = (M^1(t), ..., M^m(t))$ is a square integrable martingale. The bracket of $M^i(t)$ and $M^j(t)$ is given by $\langle M^i, M^j\rangle_t = tv_{ij}$. In fact, $(M^i(t) - M^i(s))(M^j(t) - M^j(s))$ is independent of $\mathcal{F}'_s$ and its expectation is $(t-s)v_{ij}$. Generally, however, Lévy processes are not always integrable. We will see in Section 2.4 that any Lévy process is a semimartingale.

In some applications, it is convenient to deal with another filtration $\{\mathcal{F}_t\}$, larger than the filtration $\{\mathcal{F}'_t\}$ generated by the Lévy process. Suppose that we are given a filtration $\{\mathcal{F}_t\}$, an $\{\mathcal{F}_t\}$ adapted cadlag process $Z(t)$ is called a *$\{\mathcal{F}_t\}$-Lévy process* if it is time homogeneous, continuous in probability and $Z(t) - Z(s)$ is independent of $\mathcal{F}_s$ for any $s$. Any $\{\mathcal{F}_t\}$-Lévy process is a Lévy process. Conversely any Lévy process is an $\{\mathcal{F}'_t\}$-Lévy process.

One of the most important Lévy processes is a Brownian motion. A continuous $\{\mathcal{F}_t\}$-Lévy process $Z(t)$ is called a *$\{\mathcal{F}_t\}$-Brownian motion*. It will be shown in Section

2.4 that any $\{\mathcal{F}_t\}$-Brownian motion $Z(t)$ is square integrable and in fact $Z(t)$ is Gaussian distributed with mean $tm$ and covariance $tV$. In particular if $m = 0$ and $V$ is the identity matrix $I$, the process is called a standard $\{\mathcal{F}_t\}$-Brownian motion.

Another important $\{\mathcal{F}_t\}$-Lévy process is a Poisson random measure. Let $(\mathcal{Z}, \mathcal{B})$ be a measurable space. A mapping $p : \mathbf{D}_p \to \mathcal{Z}$ is said to be a *point function* if its domain $\mathbf{D}_p$ is a countable subset of the time interval $(0, T]$. $p$ defines a counting measure on $(0, T] \times \mathcal{Z}$ by

$$N_p((s,t] \times U) = \sharp\{r \in \mathbf{D}_p; s < r \le t, p(r) \in U\}, \quad U \in \mathcal{B}. \tag{2.1}$$

Let $\Pi_{\mathcal{Z}}$ be the totality of point functions. A point process is a measurable mapping $\Omega \to \Pi_{\mathcal{Z}}$. It induces a random counting measure $N(dtdz) \equiv N_{p(\omega)}(dtdz)$. A point process $p$ is called a *$\{\mathcal{F}_t\}$-Poisson point process* if for any $E_1, ..., E_n \in \mathcal{B}$ such that $N((0,T] \times E_i) < \infty$ a.s. $i = 1, 2, ...,$ the process $X(t) = (N((0,t] \times E_1), ..., N((0,t] \times E_n))$ is a $\{\mathcal{F}_t\}$-Lévy process. The random measure $N$ is called a *$\{\mathcal{F}_t\}$-Poisson random measure.*

**Theorem 2.1.** *Let $N(dtdz)$ be a $\{\mathcal{F}_t\}$-Poisson random measure. Set $N_t(E_0) = N((0,t] \times E_0)$. If $N_t(E_0) < \infty$ holds for any $t$, it is integrable and is Poisson distributed with intensity $\nu(E_0) = E[N_1(E_0)]$.*

*Let $E_1, ..., E_n$ be disjoint measurable subsets of $\mathcal{Z}$ such that $\nu(E_i) < \infty, i = 1, ..., n$. Then $N_t(E_1), ..., N_t(E_n)$ are independent $\{\mathcal{F}_t\}$-Lévy processes.*

*Proof.* Suppose that $N_t(E_0) < \infty$ a.s. We shall compute its characteristic function $\varphi_t(\alpha) = E[e^{i\alpha N_t(E_0)}]$, $\alpha \in \mathbf{R}$. Since the distribution of $N_t(E_0)$ is infinitely divisible and is time homogeneous, it is known that the characteristic function is written as $\varphi_t(\alpha) = e^{t\psi(\alpha)}$ (Sato [21]). Set $M^\alpha(t) = e^{i\alpha X(t)}e^{-t\psi(\alpha)}$. Since $M^\alpha(t)M^\alpha(s)^{-1}$ is independent of $\mathcal{F}_s$, we have

$$E[M^\alpha(t)M^\alpha(s)^{-1}|\mathcal{F}_s] = E[M^\alpha(t)M^\alpha(s)^{-1}] = 1.$$

Therefore $M^\alpha(t)$ is a bounded complex martingale. We consider the stochastic integral

$$Y^\alpha(t) = \int_0^t \frac{1}{M^\alpha(s-)} dM^\alpha(s). \tag{2.2}$$

It is a square integrable complex martingale, since the integrand is a bounded process. It holds

$$Y_t^\alpha = \sum_{0<s\le t} \frac{\Delta M^\alpha(s)}{M^\alpha(s-)} + \int_0^t (-\psi(\alpha))ds = (e^{i\alpha} - 1)N_t(E_0) - t\psi(\alpha).$$

Taking its expectation, we obtain $(e^{i\alpha} - 1)E(N_t(E_0)) - t\psi(\alpha) = 0$. Therefore $E(N_t(E_0)) = \nu_t(E_0)$ is finite and $t\psi(\alpha) = (e^{i\alpha} - 1)\nu_t(E_0)$. This yields $\nu_t(E_0) = t\nu(E_0)$ and

$$E[e^{i\alpha N_t(E_0)}] = \exp\{t(e^{i\alpha} - 1)\nu(E_0)\}.$$

The above formula is valid for any $\alpha \in \mathbf{R}$. It indicates that $N_t(E)$ is Poisson distributed with intensity $t\nu(E)$.

We next consider the $n$ vector process $X(t) = (N_t(E_1), ..., N_t(E_n))$. Consider its characteristic function $\varphi_t(\alpha) = E[\exp(i \sum_k \alpha_k N_t(E_k))]$, where $\alpha = (\alpha_1, ..., \alpha_n)$. It is represented by $\exp(t\psi(\alpha))$ as before. We define a complex martingale with parameter $\alpha$ by

$$M^\alpha(t) = \exp(i \sum_k \alpha_k N_t(E_k)) \exp(-t\psi(\alpha))$$

and we define $Y^\alpha(t)$ by (2.2). It is a square integrable complex martingale. Note that $N_t^\alpha(E_1), ..., N_t^\alpha(E_n)$ have no common jumps, since $E_1, ..., E_n$ are disjoint. Then we have

$$Y^\alpha(t) = \sum_{k=1}^n (e^{i\alpha_k} - 1)N_t(E_k) - t\psi(\alpha).$$

Taking its expectation in the case $t = 1$, we have $\sum_{k=1}^n (e^{i\alpha} - 1)\nu(E_k) - \psi(\alpha) = 0$. Therefore we have

$$\varphi_t(\alpha) = \exp(t \sum_k (e^{i\alpha_k} - 1)\nu(E_k)) = \prod_k \exp(t(e^{i\alpha_k} - 1)\nu(E_k)).$$

This proves

$$E\Big[\prod_{k=1}^n e^{i\alpha_k N_t(E_k)}\Big] = \prod_{k=1}^n E[e^{i\alpha_k N_t(E_k)}].$$

The above formula is valid for any $\alpha_1, ..., \alpha_n$. Thus, $N_t(E_1), ..., N_t(E_n)$ are independent.

Now consider $N_t(E_k) - N_s(E_k)$, $k = 1, ..., n$, in place of $N_t(E_k)$, $k = 1, ..., n$. Then we can show that these random variables are also independent if $E_k$, $k = 1, ..., n$, are disjoint. Further these are independent of $\mathcal{F}_s$, since the $N_t(E_k)$ are $\{\mathcal{F}_t\}$-*Lévy processes*. Consequently the stochastic processes $N_t(E_k)$, $k = 1, ..., n$, are independent of each other.

We set

$$\tilde{N}(dtdz) = N(dtdz) - ds\nu(dz), \tag{2.3}$$

and call it the *compensated* Poisson random measure.

**Corollary 2.2.** *If $\nu(E) < \infty$, then $\tilde{N}_t(E) := \tilde{N}((0,t] \times E)$ is a square integrable martingale with $\langle \tilde{N}(E)\rangle_t = t\nu(E)$.*

*Let $F$ be a set such that $\nu(F) < \infty$. Then we have*

$$\langle \tilde{N}(E), \tilde{N}(F)\rangle_t = t\nu(E \cap F).$$

*Let $\phi(z)$, $\psi(z)$ be measurable on $\mathcal{Z}$ such that $\int [|\phi(z)|^2 + |\psi(z)|^2]\nu(dz) < \infty$. Then both $X(t) = \int \phi(z)\tilde{N}_t(dz)$ and $Y(t) = \int \psi(z)\tilde{N}_t(dz)$ are square integrable martingales and satisfy*

$$\langle X, Y\rangle_t = t \int \phi(z)\psi(z)\nu(dz).$$

*Proof.* Set $M(t) = \tilde{N}_t(E)$. It is a martingale such that $\Delta M(s)$ is 0 or 1. Then we have $[M]_t = N_t(E)$. Therefore $[M]_t$ and $M(t)^2$ are integrable. Further, since $M(t)^2 - [M]_t$ and $[M]_t - t\nu(E)$ are martingales, $M(t)^2 - t\nu(E)$ is a martingale. This proves $\langle M\rangle_t = t\nu(E)$.

Let $E, F$ be measurable subsets of $\mathcal{Z}$ such that $\nu(E) < \infty$ and $\nu(F) < \infty$. We split $E$ and $F$ as $E = (E \cap F) \cup (E - F)$ and $F = (E \cap F) \cup (F - E)$. Then we have

$$\begin{aligned}\langle \tilde{N}(E), \tilde{N}(F)\rangle_t &= \langle \tilde{N}(E \cap F) + \tilde{N}(E - F), \tilde{N}(E \cap F) + \tilde{N}(F - E)\rangle_t \\ &= \langle \tilde{N}(E \cap F), \tilde{N}(E \cap F)\rangle_t = t\nu(E \cap E).\end{aligned}$$

In fact, the pair of $\tilde{N}_t(E - F)$ and $\tilde{N}_t(E \cap F)$ or the pair of $\tilde{N}_t(F - E)$ and $\tilde{N}_t(E \cap F)$ are orthogonal since they have no common jumps.

Suppose that $\phi, \psi$ are simple functions of the form $\phi = \sum_i c_i 1_{E_i}$ and $\psi = \sum_j d_j 1_{E_j}$, where $E_1, ..., E_n$ are disjoint subsets of $\mathcal{Z}$ such that $\nu(E_i) < \infty$ for any $i$. Then we have

$$\begin{aligned}\langle X, Y\rangle_t &= t\sum_{ij} c_i d_j \nu(E_i \cap E_j) = t\sum_i c_i d_i \nu(E_i) \\ &= t\int \phi(z)\psi(z)\nu(dz).\end{aligned}$$

It will be obvious that the above equality can be extended to any square integrable functions $\phi, \psi$ with respect to $\nu$.

We set $\hat{N}(dsdz) = ds\nu(dz)$ and call it the *compensator* of the Poisson random measure $N(dsdz)$.

## 2.2 Stochastic integrals based on compensated Poisson random measure

We shall define the stochastic integral based on the compensated Poisson random measure. Let $g(s, z)$ be a simple process written as $g(s, z) = \sum_{i=1}^n \psi_i(z) 1_{(s_i, s_{i+1}]}(s)$, where $\psi_i(z)$ is a $\mathcal{F}_{s_i} \times \mathcal{B}$-measurable function with the square integrability condition $E[\int |\psi_i(z)|^2 \nu(dz)] < \infty$. It is called a simple predictable process with square integrability condition.

A functional $g(s, z, \omega)$, $(s, z, \omega) \in [0, T] \times \mathcal{Z} \times \Omega$ is called a predictable process if it is $\mathcal{P} \times \mathcal{B}$-measurable, where $\mathcal{P}$ is the predictable $\sigma$-field on $[0, T] \times \Omega$. We denote by $L^2(\hat{N})$ the set of all predictable functionals $g(s, z)$ such that

$$E[\int_0^T \int_{\mathcal{Z}} |g(s, z)|^2 \hat{N}(dsdz)] < \infty.$$

The following fact can be verified by a standard argument.

**Lemma 2.3.** *Simple predictable processes $g$ with the square integrability conditions are dense in $L^2(\hat{N})$.*

We will define the stochastic integral of the form $\int\int g(s,z)\tilde{N}(dsdz)$. We first remark that

$$E[\tilde{N}((s,t]\times E_1)\tilde{N}((s,t]\times E_2)|\mathcal{F}_s] = \hat{N}((s,t]\times(E_1\cap E_2)) = (t-s)\nu(E_1\cap E_2)$$

is valid because of Proposition 2.1. Then if $\psi(z)$ is $\mathcal{F}_s$-measurable random variable such that $E[\int|\psi|^2\nu(dz)]<\infty$, then

$$E\Big[\Big(\int\psi(z)\tilde{N}((s,t],dz)\Big)^2|\mathcal{F}_s\Big] = \int\psi(z)^2\hat{N}((s,t],dz) = (t-s)\int\psi(z)^2\nu(dz).$$

Let $g(t,z)$ be a simple predictable process defined by $\sum\psi_i(z)1_{[s_i,s_{i+1})}(t)$. We set

$$X(t) = \sum_i\psi_i(z)\Big(\tilde{N}(s_{i+1}\wedge t,dz) - \tilde{N}(s_i\wedge t,dz)\Big).$$

We denote it by $\int_0^t\int_{\mathcal{Z}} g(s,z)\tilde{N}(dsdz)$ and call it the *stochastic integral of g by the compensated Poisson random measure.*

**Lemma 2.4.** *The stochastic integral is a square integrable martingale. Further it holds*

$$\Big\langle\int\int_{\mathcal{Z}} g(s,z)\tilde{N}(dsdz)\Big\rangle_t = \int_0^t\int_{\mathcal{Z}} g(s,z)^2\hat{N}(dsdz), \tag{2.4}$$

$$E\Big[\Big|\int_0^t\int_{\mathcal{Z}} g(s,z)\tilde{N}(dsdz)\Big|^2\Big] = E\Big[\int_0^t\int_{\mathcal{Z}} g(s,z)^2\hat{N}(dsdz)\Big]. \tag{2.5}$$

*Proof.* For simplicity we only consider the case $s=s_{n_0}$ and $t=s_{n_1}$. Since $X(t)$ is a square integrable martingale, we have

$$\begin{aligned}
E[(X(t)-X(s))^2|\mathcal{F}_s] &= \sum_{n_0\le i<n_1} E[(X(s_{i+1})-X(s_i))^2|\mathcal{F}_s]\\
&= \sum_i E\Big[\Big(\int_{\mathcal{Z}}\psi_i(z)\tilde{N}((s_i,s_{i+1}],dz)\Big)^2|\mathcal{F}_s\Big]\\
&= \sum_i E\Big[E\Big[\Big(\int_{\mathcal{Z}}\psi_i(z)\tilde{N}((s_i,s_{i+1}],dz)\Big)^2|\mathcal{F}_{s_i}|\mathcal{F}_s\Big]\\
&= \sum_i E\Big[\int_{\mathcal{Z}}\psi_i(z)^2\hat{N}((s_i,s_{i+1}],dz)|\mathcal{F}_s\Big]\Big]\\
&= E\Big[\int_s^t\int_{\mathcal{Z}}|g(r,z)|^2\hat{N}(drdz)|\mathcal{F}_s\Big].
\end{aligned}$$

This proves the first equality of the lemma. The second equality follows immediately if we take the expectation of the first one.

Now let $g$ be any element of $L^2(\hat{N})$ and let $\{g_n\}$ be a sequence of simple predictable processes converging to $g$ with respect to the norm of $L^2(\hat{N})$. Then the sequence of stochastic integrals $\{\int_0^t g_n(s,z)\tilde{N}(dsdz)\}$ converges in $L^2$ in view of (2.5). The limit is denoted by $\int_0^t \int_{\mathcal{Z}} g(s,z)\tilde{N}(dsdz)$. It is called the stochastic integral of $g(s,z)$ by the compensated Poisson random measure.

**Remark.** Suppose that $E[\int_0^T \int_{\mathcal{Z}} |g(r,z)|dr\nu(dz)] < \infty$. We may define

$$\int_0^t \int_{\mathcal{Z}} g(s,z)N(dsdz), \qquad \int_0^t \int_{\mathcal{Z}} g(s,z)ds\nu(dz),$$

as integrable processes of finite variation. Further,

$$\int_0^t \int_{\mathcal{Z}} g(s,z)\tilde{N}(dsdz) := \int_0^t \int_{\mathcal{Z}} g(s,z)N(dsdz) - \int_0^t \int_{\mathcal{Z}} g(s,z)dr\nu(dz)$$

is a martingale.

### 2.3 Itô's formula II

We shall rewrite Itô's formula introduced in Section 1.

**Theorem 2.5 (Kunita–Watanabe).** [16] *Let $X(t) = (X^1(t), ..., X^d(t))$ be a $d$-dimensional semimartingale represented by*

$$\begin{aligned} X^i(t) &= X^i(0) + A^i(t) + M^i(t) \\ &\quad + \int_0^t \int_{\mathcal{Z}} g^i(s,z)\tilde{N}(dsdz) + \int_0^t \int_{\mathcal{Z}} h^i(s,z)N(dsdz), \quad i = 1, ..., d. \end{aligned} \tag{2.6}$$

*Here $A^i(t)$ are continuous adapted processes of finite variation, $M^i(t)$ are continuous localmartingales and $h, g$ satisfy $|g||h| = 0$. Let $F(x_1, ..., x_d)$ be a $C^2$ function. Then we have*

$$\begin{aligned} &F(X^1(t), ..., X^d(t)) \\ &= F(X^1(0), ..., X^d(0)) + \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X(s-))dA^i(s) \\ &\quad + \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X(s-))dM^i(s) \\ &\quad + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}(X(s))d\langle M^i, M^j\rangle_s \\ &\quad + \int_0^t \int_{\mathcal{Z}} \{F(X(s-) + g(s,z)) - F(X(s-))\}\tilde{N}(dsdz) \\ &\quad + \int_0^t \int_{\mathcal{Z}} \{F(X(s) + h(s,x)) - F(X(s-))\}N(dsdx) \end{aligned} \tag{2.7}$$

$$+\int_0^t \int_{\mathcal{Z}} \Big\{ F(X(s-)+g(s,z)) - F(X(s-)) \\ - \sum_{i=1}^d g^i(s,z) \frac{\partial F}{\partial x_i}(X(s-)) \Big\} \hat{N}(dsdz).$$

*Proof.* We will use Itô's formula I. Denote the sum of the two discontinuous processes in (2.6) by $X_d(t)$. Then the semimartingale $X(t)$ is decomposed as $X(t) = X(0) + A(t) + M(t) + X_d(t)$. In Itô's formula I, we split the integral $\int \frac{\partial F}{\partial x_i}(X(s))dX^i(s)$ as the sum of the integral based on $A^i(t) + M^i(t)$ and that based on $X^i_d(t)$. We have

$$\int_0^t \frac{\partial F}{\partial x_i}(X(s-))(dA^i(s) + dM^i(s)) \\ = \int_0^t \frac{\partial F}{\partial x_i}(X(s-))dA^i(s) + \int_0^t \frac{\partial F}{\partial x_i}(X(s-))dM^i(s).$$

It holds $[X^i, X^j]^c_t = \langle M^i, M^j \rangle_t$. The remaining parts of Itô's formula I are the two jump parts. These are

$$\sum_i \int_0^t \frac{\partial F}{\partial x_i}(X(s-))dX^i_d(s) \\ = \sum_i \int_0^t \int_{\mathcal{Z}} \frac{\partial F}{\partial x_i}(X(s-))g^i(s,z)\tilde{N}(dsdz) \\ + \sum_i \int_0^t \int_{\mathcal{Z}} \frac{\partial F}{\partial x_i}(X(s-))h^i(s,z)N(dsdz),$$

and

$$\sum_{0<s\le t} \Big\{ F(X(s)) - F(X(s-)) - \sum_i \frac{\partial F}{\partial x_i} \Delta X^i_s \Big\} \\ = \int_0^t \int_{\mathcal{Z}} \{F(X(s-)+g(s,z)) - F(X(s-)) \\ - \sum_i \frac{\partial F}{\partial x_i}(X(s-))g^i(s,z)\} N(dsdx) \\ + \int_0^t \int_{\mathcal{Z}} \{F(X(s-)+h(s,z)) - F(X(s-)) \\ - \sum_i \frac{\partial F}{\partial x_i}(X(s-))h^i(s,z)\} N(dsdx).$$

The sum of the above two is written as

$$\int_0^t \int_{\mathcal{Z}} \{F(X(s-) + g(s,z)) - F(X(s-))\}\tilde{N}(dsdz)$$
$$+ \int_0^t \int_{\mathcal{Z}} \{F(X(s-) + g(s,z)) - F(X(s-))$$
$$- \sum_{i=1}^d g^i(s,z) \frac{\partial F}{\partial x_i}(X(s-)) \Bigg\} \hat{N}(dsdx)$$
$$+ \int_0^t \int_{\mathcal{Z}} \{F(X(s-) + h(s,z)) - F(X(s-))\}N(dsdz).$$

Therefore we get Itô's formula II.

**Remark.** The condition $|g(t,z)||h(t,z)| = 0$ means that two processes $\tilde{N}_t(g)$ and $N_t(h)$ do not jump simultaneously. In fact, setting $\Delta\tilde{N}_t(g) = \tilde{N}_t(g) - \tilde{N}_{t-}(g)$, we have

$$\sum_{s \le t} |\Delta\tilde{N}_s(g)||\Delta N_s(h)| = N_t(|g||h|).$$

The above is 0 if and only if $|g||h| = 0$ a.e. $(t, x, \omega)$.

As an application of Itô's formula, we will give a proposition concerning distributions of $\{\mathcal{F}_t\}$-Brownian motion.

**Proposition 2.6.** *Let $Y(t)$ be a square integrable $\{\mathcal{F}_t\}$-Brownian motion. Then it is Gaussian distributed. Further it is independent of any $\{\mathcal{F}_t\}$-Poisson random measure.*

*Proof.* We consider the case where $Y(t)$ is a one-dimensional $\{\mathcal{F}_t\}$-Brownian motion with mean 0. Set $\sigma^2 t = E[Y(t)^2]$. Then it holds $\langle Y\rangle_t = \sigma^2 t$. Let $X(t) = N_t(E_0)$, where $\nu(E_0) < \infty$. We have by Itô's formula II,

$$\begin{aligned} U(t) &:= e^{i\alpha Y(t)} e^{i\beta X(t)} \\ &= i\alpha \int_0^t U(s-)Y(ds) - \frac{1}{2}\alpha^2 \int_0^t U(s-)\sigma^2 ds \\ &\quad + \int_0^t \int_{\mathcal{Z}} U(s-)(e^{i\beta} - 1)1_{E_0}(z)\tilde{N}(dsdz) \\ &\quad + \int_0^t \int_{\mathcal{Z}} (e^{i\beta} - 1)1_{E_0}(z) ds\nu(dz). \end{aligned}$$

Now taking its expectation, since $E[U(s)] = E[U(s-)]$, we have

$$E[U(t)] = -\frac{1}{2}\alpha^2\sigma^2 \int_0^t E[U(s)]ds + (e^{i\beta} - 1)\nu(E_0) \int_0^t E[U(s)]ds.$$

Therefore we have

$$E[U(t)] = \exp\left[ t\left\{ -\frac{1}{2}\sigma^2\alpha^2 + (e^{i\beta} - 1)\nu(E_0) \right\} \right].$$

This means

$$E[e^{i\alpha Y(t)}e^{i\beta X(t)}] = \exp\left\{-\frac{1}{2}t\sigma^2\alpha^2\right\}\exp\left[t\left\{(e^{i\beta-1}-1)\nu(E_0)\right\}\right]$$
$$= E[e^{i\alpha Y(t)}]E[e^{i\beta X(t)}].$$

Consequently the processes $Y(t)$ and $X(t)$ are independent. Further, $Y(t)$ is Gaussian distributed.

## 2.4 Lévy–Itô's decomposition and representation of martingales

Let $Z(t)$ be an $m$-dimensional $\{\mathcal{F}_t\}$-Lévy process. We study jumps of the Lévy process $Z(t)$. Let $\mathcal{B}(\mathbf{R}^m - \{0\})$ be the Borel field of $\mathbf{R}^m - \{0\}$. We define a random measure $N((s,t]\times E)$ on $\mathcal{B}$ by

$$N((s,t]\times E) = \sum_{r\in(s,t]} \sharp\{r;\, \Delta Z(r)\in E\}.$$

Then $\{N((s,t]\times E_0);\, E_0\in\mathcal{B}(\mathbf{R}^m-\{0\})\}$ is independent of $\mathcal{F}_s$. Also it is continuous in probability and is time homogeneous. Therefore it is an $\{\mathcal{F}_t\}$-Poisson random measure on $\mathcal{B}(\mathbf{R}^m - \{0\})$ with the intensity $\nu(E_0) = E[N((0,1]\times E_0)]$. Note that $N_t(E_0)$ is finite for any $t$ a.s. if $E_0\cap\{z;\, |z|<\epsilon\}$ is empty for some $\epsilon>0$. Then $\nu(E_0)<\infty$ holds by Theorem 2.1. Consequently it holds $\nu(\{z : |z|\geq\epsilon\})<\infty$ for any $\epsilon>0$. However it may occur that $\nu(\mathbf{R}^m-\{0\})=\infty$.

**Theorem 2.7.** *1) (Lévy–Itô's decomposition) The $\{\mathcal{F}_t\}$-Lévy process $Z(t)$ is represented as*

$$Z(t) = W(t) + bt + \int_{(0,t]}\int_{|z|>1} xN(dsdz) + \int_{(0,t]}\int_{0<|z|\leq 1} x\tilde{N}(dsdz)\}, \qquad (2.8)$$

*where $W(t) = (W^1(t),\ldots,W^m(t))$ is an $m$-dimensional square integrable $\{\mathcal{F}_t\}$-Brownian motion with mean $0$ and covariance $tV$. $N(dsdz)$ is a $\{\mathcal{F}_t\}$-Poisson random measure on $[0,T]\times(\mathbf{R}^m-\{0\})$ with intensity measure $\hat{N}(dsdz) = ds d\nu(z)$, which is independent of $W(t)$. Further the Lévy measure satisfies $\int\frac{|z|^2}{1+|z|^2}\nu(dz)<\infty$.*

*The representation is unique, i.e., suppose there exists another $\{\mathcal{F}_t\}$-Brownian motion $W'(t)$ and another Poisson random measure $N'$ and $Z(t)$ is represented by (2.8). Then we have $W=W'$ and $N=N'$.*

*2) (Lévy–Khintchin) The characteristic function $\varphi_t(\alpha) = E[e^{i(\alpha,Z(t))}]$ admits the representation*

$$\varphi_t(\alpha) = \exp\left[t\left\{i(b,\alpha) - \frac{1}{2}(\alpha,V\alpha) + \int_{\mathbf{R}^m-\{0\}}(e^{i(\alpha,z)} - 1 - i(\alpha,z)1_{|z|\leq 1})\nu(dz)\right\}\right]. \qquad (2.9)$$

*Proof.* We give the proof in the one dimensional case only. We adopt a method used in the proof of Theorem 2.1. The characteristic function of $Z(t)$ can be written as $e^{t\psi(\alpha)}$, since the law of $Z(t)$ is infinitely divisible. Set $M^\alpha(t) = e^{i(\alpha, Z(t))}e^{-t\psi(\alpha)}$. Then for any $\alpha$, $M^\alpha(t)$ is a bounded complex martingale. We define a process $Y^\alpha(t)$ by (2.2). It is a square integrable complex martingale. We will show that it is a complex $\{\mathcal{F}_t\}$-Lévy process. For a partition $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = T\}$, consider a stochastic process

$$\begin{aligned} Y^\Pi(t) :&= \sum_k \frac{1}{M^\alpha(t_{k-1} \wedge t)}(M^\alpha(t_k \wedge t) - M^\alpha(t_{k-1} \wedge t)) \\ &= \sum_k e^{i\alpha(Z(t_k \wedge t) - Z(t_{k-1} \wedge t))}e^{-(t_k \wedge t - t_{k-1} \wedge t)\psi(\alpha)}. \end{aligned}$$

Then $Y^\Pi(t_k) - Y^\Pi(t_{k-1})$ are independent of $\mathcal{F}_{t_k - 1}$ for $k = 1, 2, ..., n$, since $Z(t)$ is a $\mathcal{F}_t$*-Lévy process*. Let $\{\Pi_n\}$ be a sequence of partitions such that $\Pi_n \subset \Pi_{n+1} \subset \cdots$ and $|\Pi_n| \to 0$. Then the sequence $\{Y^{\Pi_n}(t)\}$ converges to $Y^\alpha(t)$, and the limit $Y^\alpha(t)$ is a process such that $Y(t) - Y(s)$ are independent of $\mathcal{F}_s$ for any $s < t$ such that $s, t \in \cup_n \Pi$. The process is continuous in probability and is time homogeneous. Therefore it is a complex $\{\mathcal{F}_t\}$-Lévy process.

The process $Y^\alpha(t)$ is then decomposed as the sum of a continuous (complex) $\{\mathcal{F}_t\}$-Brownian motion $Y_c^\alpha(t)$ and a purely discontinuous $\{\mathcal{F}_t\}$-Lévy process $Y_d^\alpha(t)$. The jumps of $Y^\alpha(t)$ are $\Delta Y^\alpha(t) = \frac{1}{M_{t-}^\alpha}\Delta M^\alpha(t) = e^{i\alpha\Delta Y(t)} - 1$. Therefore $Y_d^\alpha(t)$ is represented by

$$Y_d^\alpha(t) = \int_0^t \int_{0<|z|<\infty} (e^{i\alpha z} - 1)\tilde{N}(dsdz).$$

We have

$$\langle Y_d^\alpha, \bar{Y}_d^\alpha \rangle_t = t \int_{0<|z|<\infty} |e^{i\alpha z} - 1|^2 \nu(dz) < \infty.$$

Then the Lévy measure $\nu$ should satisfy $\int \frac{|z|^2}{1+|z|^2}\nu(dz) < \infty$.

Now the complex martingale $M^\alpha(t)$ can be regarded as the solution of a linear SDE

$$M^\alpha(t) = 1 + \int_0^t M_{s-}^\alpha dY^\alpha(s). \tag{2.10}$$

Then, $M^\alpha(t)$ is written as

$$\begin{aligned} M^\alpha(t) = \exp\Bigg(& Y_c^\alpha(t) + i\alpha\Big(\int_0^t \int_{0<|z|\le 1} z\tilde{N}(drdz) + \int_0^t \int_{|z|>1} zN(drdz)\Big) \\ & -\frac{1}{2}\langle Y_c^\alpha, Y_c^\alpha \rangle_t - t\int_{0<|z|\le 1} (e^{i\alpha z} - 1 - i\alpha z)\nu(dz) \\ & -t\int_{|z|>1} (e^{i\alpha z} - 1)\nu(dz)\Bigg). \end{aligned}$$

In fact, denote the right hand side of the above by $M'(t)$. Then applying Itô's formula II to the exponential function $F(x) = e^x$, we can show that $M'(t)$ satisfies the linear SDE (2.10). Since the solution of SDE (2.10) is unique, we get $M^\alpha(t) = M'(t)$.

We define a (complex) $\{\mathcal{F}_t\}$-Lévy process $Y(t)$ by

$$Y(t) = (i\alpha)^{-1}Y_c^{\alpha}(t) + \int_0^t \int_{0<|x|\le 1} x\tilde{N}(dsdz) + \int_0^t \int_{|x|>1} xN(dsdz).$$

Since $M^{\alpha}(t) = e^{i\alpha Z(t)}e^{-t\psi(\alpha)}$, the above equation is rewritten as

$$e^{i\alpha(Z(t)-Y(t))} = e^{t\psi(\alpha)}\exp\Big(-\frac{1}{2}\langle Y_c^{\alpha}, Y_c^{\alpha}\rangle_t - t\int (e^{i\alpha z} - 1 - i\alpha z 1_{|z|\le 1})\nu(dz)\Big).$$

Then $Z(t) - Y(t)$ is a deterministic process. It is differentiable with respect to $t$. The derivative does not depend on $t$, since $Z(t) - Y(t)$ is time homogeneous. We denote it by $b$. Then we get $Y(t) = Z(t) - bt$, and the process $Y(t)$ is a real process and $b$ is a real constant.

Now $W(t) = (i\alpha)^{-1}Y_c^{\alpha}(t)$ is a real continuous square integrable $\{\mathcal{F}_t\}$-Lévy process. Hence it is an $\{\mathcal{F}_t\}$-Brownian motion. Further the $\{\mathcal{F}_t\}$-Brownian motion $W(t)$ and $\{\mathcal{F}_t\}$-Poisson random measure are independent by Proposition 2.6. Consequently we get the Lévy–Itô decomposition of $Z(t)$.

Finally we have $\langle Y_c^{\alpha}, Y_c^{\alpha}\rangle_t = -t\alpha^2\sigma^2$. Therefore $\exp(t\psi(\alpha))$ is represented by (2.9). Thus the Lévy–Khintchin formula is established.

**Corollary 2.8.** *Any $\{\mathcal{F}_t\}$-Brownian motion is square integrable and is Gaussian distributed.*

**Remark.** If $\bar{b} := \int_{|x|\le 1} |x|\nu(dx) < \infty$ (this includes the case where $\nu$ is a finite measure), the last integral of equation (2.8) can be decomposed as the difference of two integrals based on $N(dsdx)$ and $\hat{N}(dsdx)$. Then equation (2.8) can be written as

$$Z(t) = W(t) + b't + \int_{(0,t]} \int_{0<|z|<\infty} zN(dsdz), \tag{2.11}$$

where $b' = b - \int_{|z|\le 1} x\nu(dz)$. However if $\bar{b} = \infty$, we cannot decompose the last integral in (2.8) as the difference of two integrals. It may happen that both integrals $\int_0^t \int_{0<|z|\le 1} xN(dsdz)$ and $\int_0^t \int_{0<|z|\le 1} x\hat{N}(dsdz)$ may diverge, but the last integral of (2.8) can converge.

Now we shall return to $\{\mathcal{F}_t'\}$-martingale, where $\{\mathcal{F}_t'\}$ is the filtration generated by a given Lévy process $Z(t)$.

**Theorem 2.9 (Kunita–Watanabe).** [16] *Let $M(t)$ be a square integrable $\{\mathcal{F}_t'\}$-martingale. Then there exist a constant $h$, predictable processes $f(s) = (f_1(s), ..., f_m(s))$ and $g(s, x)$ such that*

$$\int_0^T |f(s)|^2 ds < \infty, \quad \int_0^T \int_{\mathbf{R}^m-\{0\}} |g(s, z)|^2 ds\nu(dz) < \infty, \quad a.s. \tag{2.12}$$

*and $M(t)$ is represented by*

$$M(t) = h + \sum_{i=1}^{m} \int_0^t f_i(s)dW^i(s) + \int_0^t \int_{\mathbf{R}^m - \{0\}} g(s,z)\tilde{N}(dsdz). \tag{2.13}$$

*The triple $(h, f(s), g(s,z))$ is uniquely determined from $M(t)$, i.e., if $M(t)$ is represented by (2.13) with another $(h', f'(s), g'(s,z))$ satisfying (2.12), then we have $h = h'$, $f(s) = f'(s)$ a.e. $\lambda \otimes P$ and $g(s,z) = g'(s,z)$ a.e. $\lambda \otimes \nu \otimes P$, where $\lambda$ is the Lebesgue measure on $[0,T]$ and $\nu$ is the Lévy measure on $\mathbf{R}^m - \{0\}$.*

*Proof.* (Taken from [15]) For simplicity we prove the theorem in the case $m = 1$ only. Let $\mathbf{Z} = (Z(t))$ be a one dimensional Lévy process and let (2.8) be the Lévy–Itô decomposition. We introduce a random measure $M(E_1)$ on $[0,T] \times \mathbf{R}$ by

$$M(E_1) = \int_{E_1(0)} dW(t) + \int_{E_1 - E_1(0)} \frac{z}{1+|z|} \tilde{N}(dtdz), \tag{2.14}$$

where $E_1(0) = \{(t,0) : (t,0) \in E_1\}$. Then $\int E[M(E_1)M(E_2)]dP = \mu(E_1 \cap E_2)$, where

$$\mu(E) = |E(0)| + \int_{E-E(0)} \left(\frac{z}{1+|z|}\right)^2 dt\nu(dz).$$

For each positive integer $p$, we define the multiple Wiener integral by

$$I_p(f) = \int \cdots \int f(\xi_1, ..., \xi_p) dM(\xi_1) \cdots dM(\xi_p). \tag{2.15}$$

Let $\mathbf{H_Z}$ be the $L^2$ space over $(\Omega, \mathcal{F}_T^{\prime}, P)$ and let $\mathbf{H_Z}^{(p)}$ be the closed linear manifold of $\{I_p(f); f \in L_p^2\}$, where $L_p^2$ is the $L^2$ space on $\mathbf{R}^p$ with the product measure of $\mu$. Then it is shown in Itô [11] that one has the direct sum expansion: $\mathbf{H_Z} = \sum_{p \geq 0} \oplus \mathbf{H_Z}^{(p)}$. Note that each $I_p(f)$ is written as the sum of the following terms

$$\begin{aligned} \int \cdots \int_{0 \leq t_1 < \cdots < t_p \leq T, (z_1, ..., z_p) \in \mathbf{R}^p} & f((t_1, z_1), ..., (t_p, z_p)) dM(t_1 z_1) \cdots dM(t_p z_p) \\ &= \int_0^T \int_{\mathbf{R}} \varphi(t_p, z_p) dM(t_p z_p), \end{aligned} \tag{2.16}$$

where

$$\varphi(t_p, z_p) = \int \cdots \int_{\Lambda(t_p, z_p)} f((t_1, z_1), ..., (t_p, z_p)) dM(t_1 z_1) \cdots dM(t_{p-1} z_{p-1})$$

and
$\Lambda(t_p, z_p) = \{0 < t_1 < \cdots < t_{p-1} < t_p, (z_1, ..., z_{p-1}, z_p) \in \mathbf{R}^p\}$. Setting $\phi(t) = \varphi(t, 0)$ and $\psi(t,z) = \varphi(t,z)\frac{1+|z|}{z}$ $(|z| > 0)$, we find that the above is written as

$$\int_0^T \phi(t)dW(t) + \int_0^T \int_{\mathbf{R}-\{0\}} \psi(t,z)\tilde{N}(dtdz). \tag{2.17}$$

Therefore any element of $\mathbf{H_Z}^{(p)}$ and hence any element $X$ of $\mathbf{H_Z}$ with mean 0 is written as the above. Now taking the conditional expectation of (2.17), we obtain the representation (2.13) for the square integrable martingale $M(t) = E[X|\mathcal{F}_t]$.

**Corollary 2.10.** *The spaces of continuous martingales $\mathcal{M}_c$ and that of purely discontinuous martingales $\mathcal{M}_d$ are characterized as follows.*

$$\mathcal{M}_c = \left\{ \sum_i \int_0^t f^i(s)dW^i(s) : E\left[ \int_0^T \sum_{i,j} f^i(s)f^j(s)v_{ij}ds \right] < \infty \right\}, \quad (2.18)$$

$$\mathcal{M}_d = \left\{ \int_0^t \int_{\mathbf{R}^m - \{0\}} g(s,z)\tilde{N}(dsdz) : \ g \in L^2(\hat{N}) \right\}. \quad (2.19)$$

Let $W(t) = (W^1(t), ..., W^m(t))$ be a standard $\{\mathcal{F}_t\}$-Brownian motion. Let $f(s) = (f_1(s), ..., f_m(s))$ be a predictable process with the square integrability condition. Then the stochastic integrals $\sum_i \int_0^t f_i(s)dW^i(s)$ are continuous square integrable martingales. These martingales are orthogonal to martingales $\int\int g\tilde{N}(dsdz)$. Indeed, the quadratic co-variation of the Brownian motion $W(t)$ and Poisson random measure $N(dsdz)$ is 0, since the former is a continuous martingale and the latter is a process of finite variation. Then their stochastic integrals are orthogonal.

Let $M$, $N$ be square integrable martingales represented by

$$M(t) = \sum_{i=1}^m \int_0^t f_i(s)dW^i(s) + \int_0^t \int_{\mathcal{Z}} g(s,z)\tilde{N}(dsdz),$$

$$N(t) = \sum_{i=1}^m \int_0^t \phi_i(s)dW^i(s) + \int_0^t \int_{\mathcal{Z}} \psi(s,z)\tilde{N}(dsdz),$$

then we have the formula

$$\langle M, N\rangle_t = \sum_{i=1}^m \int_0^t f_i(s)\phi_i(s)ds + \int_0^t \int_{\mathcal{Z}} g(s,z)\psi(s,z)ds\nu(dz),$$

$$[M, N]_t = \sum_{i=1}^m \int_0^t f_i(s)\phi_i(s)ds + \int_0^t \int_{\mathcal{Z}} g(s,z)\psi(s,z)N(dsdz).$$

## 2.5 $L^p$ estimates of stochastic integrals

In this section we are concerned with $L^p$ estimates of semimartingales $X(t)$. The $L^p$ estimate plays an important role for the study of the solution of SDEs. In particular, for the study of regularity of solutions with respect to the initial conditions, we will apply Kolmogorov's criterion for the continuity of random fields, where the $L^p$ estimates of solutions are required (see Appendix). In this section we study $L^p$ estimates for general semimartingales and in the next section we apply these estimates to the solution of a SDE.

We assume that we are given an $m$ dimensional standard $\{\mathcal{F}_t\}$-Brownian motion $W(t) = (W^1(t), ..., W^m(t))$ and a $\{\mathcal{F}_t\}$-Poisson random measure $N$ on a measurable space $\mathcal{Z}$. Let us consider a $d$-dimensional semimartingale $X(t) = (X^1(t), ..., X^d(t))$ represented by.

$$X^i(t) = x_i + \int_0^t b^i(r)dr + \sum_{j=1}^m \int_0^t f^{ij}(r)dW^j(r) + \int_0^t \int_{\mathcal{Z}} g^i(r,z)\tilde{N}(drdz), \quad (2.20)$$

for $i = 1, ..., d$. Set $b(r) = (b^1(r), ..., b^d(r))$, $f(r) = (f^{ij}(r))$, $g(x,r) = (g^1(x,r), ..., g^d(x,r))$ and write $|f(r)|^2 = \sum_{ij} |f^{ij}(r)|^2$.

**Theorem 2.11.** *For any $p \geq 2$, there exists a positive constant $C_p$ such that*

$$\begin{aligned} &E\Big[\sup_{0<s\leq t} |X(s)|^p\Big] \leq \\ &C_p\Big\{|x|^p + E\Big[\Big(\int_0^t |b(r)|dr\Big)^p\Big] + E\Big[\Big(\int_0^t |f(r)|^2 dr\Big)^{p/2}\Big] \\ &+ E\Big[\Big(\int_0^t \int_{\mathcal{Z}} |g(r,z)|^2 dr\nu(dz)\Big)^{p/2}\Big] + E\Big[\int_0^t \int_{\mathcal{Z}} |g(r,z)|^p dr\nu(dz)\Big]\Big\} \end{aligned} \quad (2.21)$$

*holds for a semimartingale $X(t)$ represented by (2.20).*

**Remark.** If $X(t) = \int_0^t f(s)dW(s)$, the above inequality is known as a special case of Burkholder's inequality, which states that for any $p > 0$, there exist positive constants $c_p, C_p$ such that

$$c_p E\Big[\sup_{0\leq s\leq t} |X(s)|^p\Big] \leq E\Big[\langle X, X\rangle_t\Big] \leq C_p E\Big[\sup_{0\leq s\leq t} |X(s)|^p\Big]$$

holds for any continuous martingale $X(t)$ with $X(0) = 0$. See Ikeda–Watanabe [10].

*Proof.* We give the proof in the case where $d = 1$ only. The inequality of the proposition is obvious if the right hand side is infinite. So we assume that the right hand side is finite. We shall obtain $L^p$ estimate for each term of the right hand side of (2.20). In the following arguments, constants $c_i, i = 1, 2, \ldots$ will be determined depending on $p$ only.

The inequality

$$E\Big[\sup_{0<s\leq t} \Big|\int_0^s b(r)dr\Big|^p\Big] \leq E\Big[\Big(\int_0^t |b(r)|dr\Big)^p\Big]$$

is obvious. We shall consider the martingale $Y(t) = \sum_j \int_0^t f^j(r)dW^j(r)$. Set $F(x) = |x|^p$ where $p \geq 2$. It is a $C^2$ function with $F'(x) = p|x|^{p-2}x$ and $F''(x) = p(p-1)|x|^{p-2}$. Then by Itô's formula for continuous martingales, we have

$$|Y(t)|^p = p\int_0^t |Y(r)|^{p-2}Y(r)dY(r) + \frac{1}{2}p(p-1)\int_0^t |Y(r)|^{p-2}|f(r)|^2 ds.$$

Choose an increasing sequence of stopping times $\tau_n$ such that $P(\tau_n < T) \to 0$ as $n \to \infty$ and the stopped precess of $\int_0^t |Y(r)|^{p-2} Y(r) dY(r)$ are all bounded martingales with mean 0. Then taking the expectation of $|Y(t \wedge \tau_n)|^p$ and using Hölder's inequality, we obtain

$$
\begin{aligned}
E[|Y(t \wedge \tau_n)|^p] &= \frac{1}{2} p(p-1) E\left[\int_0^{t \wedge \tau_n} |Y(r)|^{p-2} |f(r)|^2 ds\right] \\
&\leq \frac{1}{2} p(p-1) E\Big[\sup_{0<r\leq t\wedge\tau_n} |Y(r)|^p\Big]^{(p-2)/p} \\
&\quad \times E\left[\left(\int_0^{t\wedge\tau_n} |f(r)|^2 dr\right)^{p/2}\right]^{2/p}.
\end{aligned}
$$

Now, observe Doob's inequality $E[\sup_{0<r\leq t\wedge\tau_n} |Y(r)|^p] \leq p^q E[|Y(t \wedge \tau_n)|^p]$. Then we have

$$
\begin{aligned}
E\Big[\sup_{0<r\leq t\wedge\tau_n} |Y(r)|^p\Big] &\leq \frac{1}{2} p(p-1) p^q E\left[\sup_{0<r\leq t\wedge\tau_n} |Y(r)|^p\right]^{(p-2)/p} \\
&\quad \times E\left[\left(\int_0^{t\wedge\tau_n} |f(r)|^2 dr\right)^{p/2}\right]^{2/p}.
\end{aligned}
$$

Then we have

$$
E\Big[\sup_{0<r\leq t\wedge\tau_n} |Y(r)|^p]^{2/p}\Big] \leq \frac{1}{2} p(p-1) p^q E\left[\left(\int_0^{t\wedge\tau_n} |f(r)|^2 dr\right)^{p/2}\right]^{2/p}.
$$

Thus we obtain

$$
E\Big[\sup_{0<s\leq t} |Y(s)|^p\Big] \leq (p(p-1)p^q/2)^{p/2} E\left[\left(\int_0^t |f(r)|^2 dr\right)^{p/2}\right].
$$

We next consider $Y(t) = \int_0^t \int g(r,z) \tilde{N}(drdz)$. We apply Itô's formula II for $F(x) = |x|^p$. Since $Y(t)$ is purely discontinuous, we have

$$
\begin{aligned}
|Y(t)|^p &= Z(t) \qquad (2.22) \\
&+ \int_0^t \int_{\mathcal{Z}} \{|Y(r-) + g|^p - |Y(r-)|^p - p|Y(r-)|^{p-2} Y(r-) g\} \hat{N}(drdz),
\end{aligned}
$$

where $Z(t)$ is a localmartingale. We may choose an increasing sequence of stopping times $\tau_n$ such that $P(\tau_n < T) \to 0$ as $n \to \infty$, $Z(t \wedge \tau_n)$ are martingales with mean 0 and the last term in (2.22) stopped at $\tau_n$ are integrable. To make the notation simple, we denote the stopped processes $Y(t \wedge \tau_n)$ etc by the same notations $Y(t)$, etc. Since,

$$
\begin{aligned}
&|Y(r-)+g|^p - |Y(r-)|^p - p|Y(r-)|^{p-2}Y(r-)g \\
&= \frac{1}{2}p(p-1)|Y(r-)+\theta g|^{p-2}g^2 \\
&\le c_3|Y(r-)|^{p-2}g^2 + c_4|g|^p
\end{aligned}
$$

holds for some $|\theta| < 1$, we have from (2.22)

$$
\begin{aligned}
E[|Y(t)|^p] \le c_3 E\left[\int_0^t \int_Z |Y(r-)|^{p-2}g^2 dr\nu(dz)\right] \\
+c_4 E\left[\int_0^t \int_Z |g|^p dr\nu(dz)\right].
\end{aligned}
\tag{2.23}
$$

We calculate the first term of the right hand side. It holds by Hölder's inequality

$$
\begin{aligned}
&E\left[\int_0^t \int_Z |Y(r-)|^{p-2}g^2 dr\nu(dz)\right] \\
&\le E\left[\sup_{0<r\le t}|Y(r-)|^p\right]^{1-2/p} E\left[\left(\int_0^t \int_Z g^2 dr\nu(dz)\right)^{p/2}\right]^{2/p} \\
&\le c_5 E\left[\sup_{0<r\le t}|Y(r-)|^p\right] + c_6 E\left[\left(\int_0^t \int_Z g^2 dr\nu(dz)\right)^{p/2}\right].
\end{aligned}
$$

In the last inequality, we used the inequality $ab \le \frac{a^{p'}}{p'} + \frac{b^{q'}}{q'}$, where $a, b > 0$, $p', q' > 1$ and $\frac{1}{p'} + \frac{1}{q'} = 1$. Therefore, using Doob's inequality, we get from (2.23),

$$
\begin{aligned}
E\Big[\sup_{0<s\le t}|Y(s)|^p\Big] \le q^p c_3 c_5 E\Big[\sup_{0<r\le t}|Y(r-)|^p\Big] \\
+q^p c_3 c_6 E\left[\left(\int_0^t \int_Z |g|^2 dr\nu(dz)\right)^{p/2}\right] \\
+q^p c_4 E\left[\int_0^t \int_Z |g|^p dr\nu(dz)\right].
\end{aligned}
$$

Choose $c_5$ small enough such that $q^p c_3 c_5 < 1$ and move the term to the left. Note that $\sup_{0<s\le t}|Y(s)| = \sup_{0<r\le t}|Y(r-)|$ holds a.s. for any $t$. Then we get the inequality

$$
\begin{aligned}
E\Big[\sup_{0<s\le t}|Y(s)|^p\Big] \le c_7 E\left[\left(\int_0^t \int_Z |g|^2 dr\nu(dz)\right)^{p/2}\right] \\
+c_8 E\left[\int_0^t \int_Z |g|^p dr\nu(dz)\right].
\end{aligned}
\tag{2.24}
$$

So far we have chosen constants $c_7, c_8$ not depending on the form of martingales $Y(t)$ and stopping times $\tau_n$. Therefore the inequality holds up to stopping time $t \wedge \tau_n$ in (2.24) with the common constants $c_7, c_8$. Now let $n$ tend to infinity. Then we find that (2.24) holds for any $t$. The proof is complete.

**Corollary 2.12.** *For any $p \geq 2$, there exists a positive constant $C'_p$ such that*

$$E\Big[\sup_{0<s\leq t}|X(s)|^p\Big] \leq C'_p\Big\{|x|^p + E\Big[\int_0^t |b(r)|^p dr\Big] + E\Big[\int_0^t |f(r)|^p dr\Big] \tag{2.25}$$
$$+E\Big[\int_0^t \Big(\int_{\mathcal{Z}} |g(r,z)|^2\nu(dz)\Big)^{p/2} dr\Big] + E\Big[\int_0^t \Big(\int_{\mathcal{Z}} |g(r,z)|^p\nu(dz)\Big)dr\Big]\Big\}$$

*holds for any semimartingale represented by (2.20).*

We shall combine the above estimate and Itô's formula II. The following estimate will provide a useful tool for the study of stochastic flows in the next section.

**Corollary 2.13.** *For any $p \geq 2$, there exists a positive constant $C_p$ such that*

$$E\Big[\sup_{0<s\leq t}|F(X(s))|^p\Big] \tag{2.26}$$
$$\leq C_p\Big\{|F(x)|^p + E\Big[\int_0^t \Big|\sum_i F'_{x_i}(X(r-))b^i(r)\Big|^p dr\Big]$$
$$+E\Big[\int_0^t \Big(\sum_k |\sum_i F'_{x_i}(X(r-))f^{i,k}(r)|^2\Big)^{p/2} dr\Big]$$
$$+\frac{1}{2}E\Big[\int_0^t \Big|\sum_{i,j,k} F''_{x_ix_j}(X(r-))f^{ik}(r)f^{jk}(r)\Big|^p dr\Big]$$
$$+E\Big[\int_0^t \Big(\int_{\mathcal{Z}} |F(X(r-)+g(r,z)) - F(X(r-))|^2\nu(dz)\Big)^{p/2} dr\Big]$$
$$+E\Big[\int_0^t \Big(\int_{\mathcal{Z}} |F(X(r-)+g(r,z)) - F(X(r-))|^p\nu(dz)\Big)dr\Big]$$
$$+E\Big[\int_0^t \Big(\int_{\mathcal{Z}} |F(X(r-)+g(r,z)) - F(X(r-))$$
$$-\sum_i F'_{x_i}(X(r-))g^i(r,z)|\nu(dz)\Big)^p dr\Big]\Big\},$$

*where $X(t)$ is a $d$-dimensional semimartingale represented by (2.20) and $F(x_1, ..., x_d)$ is a $C^2$-function.*

Finally we shall study the uniform convergence of a sequence of stochastic integrals with respect to a parameter. It is convenient to introduce a "$p$-Lipschitz norm"for a $p$-th integrable random field $\{X(x); x \in \mathbf{R}^d\}$

$$\|X\|_p = \sup_{x\in\mathbf{R}^d} E[|X(x)|^p]^{1/p} + \sup_{x\neq y, x,y\in\mathbf{R}^d} \frac{E[|X(x)-X(y)|^p]^{1/p}}{|x-y|}.$$

Let $b(x,r)$, $f(x,r)$, $g(x,r,z)$ and $b_n(x,r)$, $f_n(x,r)$, $g_n(x,r,s)$, $n = 1, 2, \ldots$ be a sequence of predictable processes. We consider a sequence of semimartingales with parameter $x$

$$X(x,t) = X(x) + \int_0^t b(x,r)dr + \int_0^t f(x,r)dW(r) + \int_0^t \int_{\mathcal{Z}} g(x,r,z)\tilde{N}(drdz).$$

$$X_n(x,t) = X_n(x) + \int_0^t b_n(x,r)dr + \int_0^t f_n(x,r)dW(r) \\ + \int_0^t \int_{\mathcal{Z}} g_n(x,r,z)\tilde{N}(drdz).$$

**Theorem 2.14.** *Assume that the sequence $\{b_n, f_n, g_n\}$ satisfies*

$$\lim_{n\to\infty} \sup_{0<s\le T} \|b_n(s) - b(s)\|_p = 0,$$

$$\lim_{n\to\infty} \sup_{0<s\le T} \|f_n(s) - f(s)\|_p = 0,$$

$$\lim_{n\to\infty} \sup_{0<s\le T} \left\| \left( \int_{\mathcal{Z}} |g_n(\cdot,s,z) - g(\cdot,s,z)|^{p'} \nu(dz) \right)^{1/p'} \right\|_p = 0.$$

*for $p' = 2$ and $p' = p$, where $p > d$. Then if $\|X_n - X\|_p \to 0$, we have for any positive number $N$,*

$$\lim_{n\to\infty} E\Big[ \sup_{|x|\le N} \sup_{0<s\le T} |X_n(x,s) - X(x,s)|^p \Big] = 0. \tag{2.27}$$

*Proof.* We shall apply Theorem 4.4 in Appendix. Observe first that (2.25) implies

$$\lim_{n\to\infty} \sup_{0<s\le T} \|X_n(s) - X(s)\|_p = 0.$$

Then we get (2.27) by Theorem 4.4.

# 3 Stochastic differential equation and stochastic flow

## 3.1 Semimartingale with spatial parameter and a SDE based on it

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space equipped with a filtration $\{\mathcal{F}_t\}_{0\le t\le T}$. We assume that on the probability space an $m$-dimensional standard $\{\mathcal{F}_t\}$-Brownian motion $W(t) = (W^1(t), ..., W^m(t))$ and a $\{\mathcal{F}_t\}$-Poisson random measure $N((s,t] \times U)$ on a measurable space $(\mathcal{Z}, \mathcal{B})$ are defined.

Let $b(x,t) = (b^1(x,t), ..., b^d(x,t))$ and $f(x,t) = (f^{ij}(x,t))_{i=1,...,d,j=1,...,m}$ be predictable processes with spatial parameter $x \in \mathbf{R}^d$. Let $g(x,t,z) = (g^1(x,t,z), ..., g^d(x,t,z))$, $x \in \mathbf{R}^d$, $z \in \mathcal{Z}$ be a predictable process with parameters $x, z$. We will assume that these processes are of linear growth and are Lipschitz continuous, i.e., there exist positive constants $K$, $K(z)$ and $L$, $L(z)$ such that

$$\frac{|b(x,t)|}{1+|x|} \le K, \quad |b(x,t)-b(y,t)| \le L|x-y|, \tag{3.1}$$
$$\frac{|f(x,t)|}{1+|x|} \le K, \quad |f(x,t)-f(y,t)| \le L|x-y|,$$
$$\frac{|g(x,t,z)|}{1+|x|} \le K(z), \quad |g(x,t,z)-g(y,t,z)| \le L(z)|x-y|,$$

hold for all $x, y \in \mathbf{R}^d$ and $z \in \mathcal{Z}$, a.e. $\Lambda \times P$, where $\Lambda$ is the Lebesgue measure on $[0, T]$. Further constants $K(z)$, $L(z)$ satisfy

$$\int_{\mathcal{Z}} (K(z)^p + L(z)^p)\nu(dz) < \infty \quad \forall p \ge 2. \tag{3.2}$$

We define an $\mathbf{R}^d$-valued process $X(x,t) = (X^1(x,t), ..., X^d(x,t))$ with parameter $x$ by

$$X^i(x,t) = \int_0^t b^i(x,r)dr + \sum_j \int_0^t f^{ij}(x,r)dW^j(r) + \int_0^t \int_{\mathcal{Z}} g^i(x,r,z)\tilde{N}(drdz). \tag{3.3}$$

With vector–matrix notations, the above is written as

$$X(x,t) = \int_0^t b(x,r)dr + \int_0^t f(x,r)dW(r) + \int_0^t \int_{\mathcal{Z}} g(x,r,z)\tilde{N}(drdz).$$

It is a semimartingale with spatial parameter $x$.

Let $\eta_t$ be an $\mathbf{R}^d$-valued adapted cadlag process. Then the stochastic processes $b(\eta_{t-}, t)$, $f(\eta_{t-}, t)$ and $g(\eta_{t-}, t, z)$ are predictable processes. We set

$$\int_{t_0}^t X(\eta_{r-}, dr) := \int_{t_0}^t b(\eta_{r-}, r)dr + \int_{t_0}^t f(\eta_{r-}, r)dW(r) + \int_{t_0}^t \int_{\mathcal{Z}} g(\eta_{r-}, r, z)\tilde{N}(drdz).$$

It is well defined as an $\mathbf{R}^d$-valued semimartingale.

Now we shall consider a SDE based on $X(x,t)$. Suppose that we are given an $\mathcal{F}_{t_0}$-measurable $\mathbf{R}^d$-valued random variable $\xi_0$. Let $\xi_t = (\xi_t^1, ..., \xi_t^d)$ be an $\mathbf{R}^d$-valued $\{\mathcal{F}_t\}$-adapted cadlag process satisfying the equation

$$\xi_t = \xi_0 + \int_{t_0}^t X(\xi_{s-}, ds). \tag{3.4}$$

Then the process $\xi_t$ is called a *solution of the stochastic differential equation based on $X(x,t)$, starting from $\xi_0$ at time $t_0$*. The process $X(x,t)$ is called the *infinitesimal generator* of the solution $\xi_t$. Functionals $b(x,t)$, $f(x,t)$ and $g(x,t,z)$ are called *drift coefficients, diffusion coefficients* and *jump coefficients*, respectively.

**Theorem 3.1.** *Suppose that $b(x,t)$, $f(x,t)$, $g(x,t,z)$ satisfy (3.1) and (3.2). Then, if the initial data $\xi_0$ is p-th integrable, the equation has a unique solution in $L^p$.*

*Proof.* The existence of the solution can be verified by a standard method of successive approximation. We will find solutions in $L^p$ space. Given an $\mathcal{F}_{t_0}$-measurable random variable $\xi_0$, we define a sequence of $\mathbf{R}^d$-valued semimartingales by $\xi_t^0 = \xi_0$ and

$$\xi_t^n = \xi_0 + \int_{t_0}^t X(\xi_{s-}^{n-1}, ds), \quad n = 1, 2, ... \tag{3.5}$$

We shall compute the $L^p$-norm of the above. It holds

$$\xi_t^1 = \xi_0 + \int_{t_0}^t b(\xi_0, r)dr + \int_{t_0}^t f(\xi_0, r)dW(r) + \int_{t_0}^t \int_{\mathcal{Z}} g(\xi_0, r, z)\tilde{N}(drdz).$$

By applying Theorem 2.11, we have

$$\begin{aligned} E\Big[\sup_{t_0\le r\le t} |\xi_r^1 - \xi_r^0|^p\Big] \le C_p\Big\{ & E\Big[\Big(\int_{t_0}^t |b(\xi_0, r)|dr\Big)^p\Big] \\ & + E\Big[\Big(\int_{t_0}^t |f(\xi_0, r)|^2 dr\Big)^{p/2}\Big] \\ & + E\Big[\Big(\int_{t_0}^t \int_{\mathcal{Z}} |g(\xi_0, z)|^2 dr\nu(dz)\Big)^{p/2}\Big] \\ & + E\Big[\int_{t_0}^t \int_{\mathcal{Z}} |g(\xi_0, z)|^p dr\nu(dz)\Big]\Big\}. \end{aligned}$$

Here, we set $|f|^2 = \sum_{i,j} |f^{ij}|^2$. By the linear growth properties of coefficients $b$, $f$, $g$, the above is dominated by $C_p' E[(1 + |\xi_0|)^p] < \infty$.

We have further

$$\begin{aligned} E\Big[\sup_{t_0\le r\le t} |\xi_r^{n+1} - \xi_r^n|^p\Big] \le C_p\Big\{ & E\Big[\Big(\int_{t_0}^t |b(\xi_{r-}^n, r) - b(\xi_{r-}^{n-1}, r)|dr\Big)^p\Big] \\ & + E\Big[\Big(\int_{t_0}^t |f(\xi_{r-}^n, r) - f(\xi_{r-}^{n-1}, r)|^2 dr\Big)^{p/2}\Big] \\ & + E\Big[\Big(\int_{t_0}^t \int_{\mathcal{Z}} |g(\xi_{r-}^n, r, z) - g(\xi_{r-}^{n-1}, r, z)|^2 dr\nu(dz)\Big)^{p/2}\Big] \\ & + E\Big[\int_{t_0}^t \int_{\mathcal{Z}} |g(\xi_{r-}^n) - g(\xi_{r-}^{n-1})|^p dr\nu(dz))\Big]\Big\}. \end{aligned}$$

Using Lipschitz conditions for $b$, $f$, $g$, we get

$$\begin{aligned} E\Big[\sup_{t_0\le r\le t} |\xi_r^{n+1} - \xi_r^n|^p\Big] & \le C\int_{t_0}^t E[|\xi_{r-}^n - \xi_{r-}^{n-1}|^p]dr \\ & \le C\int_{t_0}^t E\Big[\sup_{t_0\le r\le s} |\xi_{r-}^n - \xi_{r-}^{n-1}|^p\Big]ds. \end{aligned}$$

Then we get the inequality

$$E\Big[\sup_{t_0 \le r \le t} |\xi_{r-}^{n+1} - \xi_{r-}^n|^p\Big] \le \frac{C^n}{n!} E\Big[\sup_{t_0 \le r \le t} |\xi_{r-}^1 - \xi_{r-}^0|^p\Big].$$

Then the sequence $\{\xi_r^n, t_0 \le r \le T\}$ converges in $L^p$ uniformly in $r$. Denote the limit as $\xi_r, r \le T$. It is a cadlag adapted process. Let $n$ tend to infinity in (3.5). Then $\xi_t$ satisfies (3.4). It is a desired solution of the SDE.

For the uniqueness of the solution, suppose that $\xi_t'$ is another solution belonging to $L^p$. Then we have $\xi_t - \xi_t' = \int_{t_0}^t X(\xi_{s-}, ds) - \int_{t_0}^t X(\xi_{s-}', ds)$. Then we get the estimate

$$E[|\xi_t - \xi_t'|^p] \le C \int_{t_0}^t E[|\xi_{s-} - \xi_{s-}'|^p] ds,$$

similarly as the above argument. This implies $E[|\xi_t - \xi_t'|^p] = 0$.

**Remark 1.** We can show that any solution of equation (3.4) belongs to $L^p$. Therefore the solution obtained by the successive approximation is the unique solution. Indeed, let $\xi_t$ be an arbitrary solution. Define a sequence of stopping times $\tau_n$ by $\tau_n = \inf\{t \in [t_0, T];\ \ |\xi_t| \ge n\}$. Then $P(\tau_n < T) \to 0$ as $n \to \infty$. We can show that there exists a positive constant $C$ not depending on $n$ such that

$$E\Big[\sup_{t_0 \le r \le t \wedge \tau_n} |\xi_r|^p\Big] \le C\Big(1 + E\Big[\int_{t_0}^{t \wedge \tau_n} |\xi_{r-}|^p dr\Big]\Big).$$

(Cf. the proof of (3.6) in the next section.) Since $|\xi_r| \le n$ holds for $r < \tau_n$, the right hand side is finite a.s. Then we obtain $E[\sup_{t \le r \le \tau_n} |\xi_r|^p] \le Ce$, from the above functional inequality. Let $n$ tend do infinity. Since the constant $C$ does not depend on $n$, we find that $\xi_t$ is $p$-th integrable.

**Remark 2.** It is often considered a SDE with 'big jumps'. Let $U$ be a Borel subset of $\mathcal{Z}$ such that $\nu(U^c) < \infty$. Consider a semimartingale with spatial parameter;

$$\begin{aligned}\tilde{X}(x,t) = &\int_0^t b(x,r)dr + \int_0^t f(x,r)dW(r) + \int_0^t \int_U g(x,r,z)\tilde{N}(drdz)\\ &+ \int_0^t \int_{U^c} g(x,r,z)N(drdz).\end{aligned}$$

where coefficients $b, f, g$ are Lipschitz continuous and of the linear growth in the above sense. However, $K(z), L(z)$ do not necessarily satisfy (3.2) but they satisfy

$$\int_U (K(z)^p + L(z)^p)\nu(dz) < \infty.$$

The solution of the SDE based on the above $\tilde{X}(x,t)$ exists uniquely. However, it does not belong to $L^p$ in general. Such a SDE will be discussed in Section 3.5.

**Remark 3.** Let $v_1(x), ..., v_m(x)$ be Lipschitz continuous $\mathbf{R}^d$-valued functions and let $(Z^1(t), ..., Z^m(t))$ be an $m$-dimensional Lévy process. A SDE of the form

$$\xi_t = \xi_0 + \sum_{j=1}^{m} \int_{t_0}^{t} v_j(\xi_{s-}) dZ^j(r)$$

where $Z^j$ are even general semimartingales, is studied in detail (see Protter [20]). Setting $X(x,t) = \sum_j v^j(x) Z^j(t)$, the above equation is written as (3.4). The equation is said to be of *separating type*.

Instead of the above SDE of the separating type, we will consider a "canonical SDE" written as

$$\xi_t = \xi_0 + \sum_{j=1}^{m} \int_{t_0}^{t} v_j(\xi_{s-}) \diamond dZ^j(r),$$

where the symbol $\diamond$ means the *canonical stochastic integral*. It will be discussed in Section 3.7.

## 3.2 Continuity

We denote by $\xi_t(x)$ the solution starting from $x$ at time $t_0$. We study the continuity of the solution $\xi_t(x)$ with respect to the initial condition $x$. Our idea is to apply Kolmogorov's criterion for the continuity of random field. See Theorem 4.1 in the Appendix. For this purpose, we claim the following.

**Theorem 3.2.** *Assume the same condition as in Theorem 3.1 for coefficients $b, f, g$. Then for any $p \geq 2$, there exists a positive constant $C_p$ such that*

$$E\Big[\sup_{t_0 \leq s \leq t} (1 + |\xi_s(x)|)^p\Big] \leq C_p (1 + |x|)^p, \quad \forall x \in \mathbf{R}^d \tag{3.6}$$

$$E\Big[\sup_{t_0 \leq s \leq t} |\xi_s(x) - \xi_s(y)|^p\Big] \leq C_p |x - y|^p, \quad \forall x, y \in \mathbf{R}^d \tag{3.7}$$

*hold for any* $0 \leq t_0 < t \leq T$.

*Proof.* We set $X(t) = \xi_t(x)$. It satisfies

$$\begin{aligned} X(t) = x &+ \int_{t_0}^{t} b(X(r-), r) dr \\ &+ \int_{t_0}^{t} f(X(r-), r) dW(r) + \int_{t_0}^{t} \int_{\mathcal{Z}} g(X(r-), r, z) \tilde{N}(drdz). \end{aligned}$$

We apply Corollary 2.12. Then we have

$$\begin{aligned} &E\Big[\sup_{t_0 \leq s \leq t} |X(s)|^p\Big] \\ &\leq C_p' \Big\{ |x|^p + E\Big[\int_{t_0}^{t} |b(X(r-), r)|^p dr\Big] + E\Big[\int_{t_0}^{t} |f(X(r-), r)|^p dr\Big] \\ &+ E\Big[\int_{t_0}^{t} \Big(\int_{\mathcal{Z}} |g(X(r-), r, z)|^2 \nu(dz)\Big)^{p/2} dr\Big] \\ &+ E\Big[\int_{t_0}^{t} \int_{\mathcal{Z}} |g(X(r-), r, z)|^p \nu(dz) dr\Big] \Big\}. \end{aligned}$$

Note the linear growth properties of $b$, $f$, $g$ in (3.1) and (3.2). Then we get

$$E\Big[\sup_{t_0\le s\le t}|X(s)|^p\Big] \le C'_p\Big\{|x|^p + (2K + K_2 + K_p)\int_{t_0}^t E[(1+|X(r-)|)^p]dr\Big\},$$

where $K_p = \int K(z)^p \nu(dz)$. Therefore,

$$E\Big[\sup_{t_0\le s\le t}(1+|X(s)|)^p\Big] \le C''_p\Big\{(1+|x|)^p$$
$$+ (2K + K_2 + K_p)\int_{t_0}^t E\Big[\sup_{t_0\le s\le r}(1+|X(s)|)^p\Big]dr\Big\}.$$

The above inequality implies (3.6).

We next set $Y(t) = \xi_t(x) - \xi_t(y)$. It satisfies

$$\begin{aligned} Y(t) = (x-y) &+ \int_{t_0}^t (b(\xi_{r-}(x), r) - b(\xi_{r-}(y), r))dr \\ &+ \int_{t_0}^t (f(\xi_{r-}(x), r) - f(\xi_{r-}(y), r))dW(r) \\ &+ \int_{t_0}^t \int_{\mathcal{Z}} (g(\xi_{r-}(x), r, z) - g(\xi_{r-}(y), r, z))\tilde{N}(drdz). \end{aligned}$$

Therefore we have by Corollary 2.12,

$$\begin{aligned} E[\sup_{t_0\le s\le t}|Y(s)|^p] \le C'_p\Big\{|x-y|^p &+ E\Big[\int_{t_0}^t |b(\xi_{r-}(x), r) - b(\xi_{r-}(y), r)|^p dr\Big] \\ &+E\Big[\int_{t_0}^t |f(\xi_{r-}(x), r) - f(\xi_{r-}(y), r)|^p dr\Big] \\ &+E\Big[\int_{t_0}^t \Big(\int_{\mathcal{Z}} |g(\xi_{r-}(x), r, z) - g(\xi_{r-}(y), r, z)|^2 \nu(dz)\Big)^{p/2} dr\Big] \\ &+E\Big[\int_{t_0}^t \int_{\mathcal{Z}} |g(\xi_{r-}(x), r, z) - g(\xi_{r-}(y), r, z)|^p \nu(dz)dr\Big]\Big\}. \end{aligned}$$

Using the Lipschitz conditions for $b$, $f$, $g$, we get the inequality

$$\begin{aligned} E\Big[\sup_{t_0\le s\le t}|Y(s)|^p\Big] &\le C'_p\Big\{|x-y|^p + (2L + L_2 + L_p)\int_{t_0}^t E\Big[|Y(r-)|^p\Big]dr\Big\} \\ &\le C''_p\Big\{|x-y|^p + (2L + L_2 + L_p)\int_{t_0}^t E\Big[\sup_{t_0\le s\le r}|Y(s)|^p\Big]dr\Big\}, \end{aligned}$$

where $L_p = \int L(z)^p \nu(dz)$. Therefore we get the desired inequality (3.7).

Now we apply Kolmogorov–Totoki's theorem (Theorem 4.1 in Appendix) by setting $\gamma = \alpha = p > d$ and $X(x) = \{\xi_s(x), s \in [t_0, t]\}$, where the norm is the supremum norm with respect to $s \in [t_0, t]$. Then the random field $\xi_t(x)$ has a modification $\xi'_t(x)$ such that for any $x$ it is cadlag with respect to $t$ and for any $t$ it is continuous in $x$, a.s.

From now we denote the above modification by the same notation $\xi_t(x)$. Let $C = C(\mathbf{R}^d; \mathbf{R}^d)$ be the space of continuous maps from $\mathbf{R}^d$ into itself. Then $\xi_t$ may be considered as a $C$-valued cadlag proceess.

### 3.3 Differentiability

We will study the differentiability of $\xi_t(x)$ with respect to $x$. Let $e_i = (0, ..., 0, 1, 0, ..., 0)$ (1 is at the $i$-th component) be a unit vector in $\mathbf{R}^d$ and let $\lambda$ be a real number such that $\lambda \neq 0$. Set

$$N_t(x, \lambda) = \frac{1}{\lambda}\Big(\xi_t(x + \lambda e_i) - \xi_t(x)\Big).$$

It is continuous in $(x, \lambda) \in D$ for any $t$, where $D = \mathbf{R}^d \times (\mathbf{R} - \{0\})$. If we can show that the limit exists for all $t, x$ a.s. as $\lambda \to 0$, then $\xi_t(x)$ is differentiable with respect to $x$ at any $(t, x)$ a.s. We will prove this fact by applying Kolmogorov's criterion. Indeed, by taking $p\delta > d + 1$ in the following theorem (Theorem 3.3), we will find that the random field $N_t(x, \lambda)$, $(x, \lambda) \in D$ is uniformly continuous in $D$ and in fact it can be extended continuously to the closure of $D$. Then the extended random field $N_t(x, \lambda)$ is continuous in $\bar{D} = \mathbf{R}^{d+1}$. This means in particular that $N_t(x, 0) = \exists \lim_{\lambda \to 0} N_t(x, \lambda)$ a.s. and it is continuous in $x$, where the convergence takes place uniformly on bounded sets of $[0, T] \times \mathbf{R}^d$. This shows that $\xi_t(x)$ is continuously partial differentiable with respect to $x_i$ ($i$-th component of $x$) and $\nabla_{x_i} \xi_t(x) = N_t(x, 0)$ holds for any $(t, x)$ a.s. This argument holds for any component $x_i$ of $x = (x_1, ..., x_n)$. Thus $\xi_t(x)$ is continuously differentiable.

In this section we assume that coefficients $b, f, g$ of the SDE are differentiable and their derivatives $\nabla b, \nabla f, \nabla g$ are bounded and $\delta$-Hölder continuous: there exists positive constants $K', K'(z)$ and $L', L'(z)$ such that

$$\begin{aligned} |\nabla b(x, t)| \leq K', \quad |\nabla b(x, t) - \nabla b(y, t)| &\leq L'|x - y|^\delta, \\ |\nabla f(x, t)| \leq K', \quad |\nabla f(x, t) - \nabla f(y, t)| &\leq L'|x - y|^\delta, \\ |\nabla g(x, t, z)| \leq K'(z), \quad |\nabla g(x, t, z) - \nabla g(y, t, z)|^\delta &\leq L'(z)|x - y|^\delta, \end{aligned} \tag{3.8}$$

holds for all $x, y$ a.e. $\Lambda \times P$ and

$$\int_{\mathcal{Z}} (K'(z)^p + L'(z)^p)\nu(dz) < \infty, \quad \forall p \geq 2. \tag{3.9}$$

**Theorem 3.3.** *Assume that coefficients of the SDE defining (3.3) are differentiable with respect to x and satisfy (3.1), (3.2), (3.8) and (3.9). Then for any $p \geq 2$, there exists a positive constant $C_p$ such that [E denoting expectation as usual]*

$$E\Big[\sup_{t_0 \leq s \leq t} |N_s(x, \lambda)|^p\Big] \leq C_p, \quad \forall (\lambda, x) \in D, \tag{3.10}$$

$$E\Big[\sup_{t_0 \leq s \leq t} |N_s(x, \lambda) - N_s(x', \lambda')|^p\Big] \leq C_p\{|x - x'|^{\delta p} + |\lambda - \lambda'|^{\delta p}\}, \\ \forall ((\lambda, x), (\lambda', x')) \in D^2 \tag{3.11}$$

*holds for any* $t_0 \le t \le T$.

*Proof.* Set $X'(t) = N_t(x, \lambda)$. It is written as

$$X'(t) = e_i + \int_{t_0}^{t} b'_\lambda(r)dr + \int_{t_0}^{t} f'_\lambda(r)dW(r) + \int_{t_0}^{t}\int_{\mathcal{Z}} g'_\lambda(r,z)\tilde{N}(drdz), \tag{3.12}$$

where

$$b'_\lambda(r) = \frac{1}{\lambda}\Big(b(\xi_{r-}(x+\lambda e_i), r) - b(\xi_{r-}(x), r)\Big),$$
$$f'_\lambda(r) = \frac{1}{\lambda}\Big(f(\xi_{r-}(x+\lambda e_i), r) - f(\xi_{r-}(x), r)\Big),$$
$$g'_\lambda(r,z) = \frac{1}{\lambda}\Big(g(\xi_{r-}(x+\lambda e_i), r, z) - g(\xi_{r-}(x), r, z)\Big).$$

We shall apply Corollary 2.12. We first consider $b'_\lambda(r)$. Setting $\int_0^1 \nabla b(x+\theta y, r)d\theta = \bar{b}(x,y,r)$, we have

$$b'_\lambda(r) = \left(\int_0^1 \nabla b(\xi_{r-}(x) + \theta\lambda N_{r-}(x,\lambda), r)d\theta\right) N_{r-}(x,\lambda)$$
$$= \bar{b}(\xi_{r-}(x), \lambda N_{r-}(x,\lambda), r)N_{r-}(x,\lambda).$$

Since $|\bar{b}| \le c$ (bounded), we have

$$E\Big[\int_{t_0}^{t} |b'_\lambda(r)|^p dr\Big] \le cE\Big[\int_{t_0}^{t} |X'(r-)|^p dr\Big].$$

We have similarly

$$E\Big[\int_{t_0}^{t} |f'_\lambda(r)|^p dr\Big] \le cE\Big[\int_{t_0}^{t} |X'(r-)|^p dr\Big].$$

Next consider $g'$. Define $\bar{g}(x,y,r,z)$ from $g(x,r,z)$ as before. Then we have

$$g'_\lambda(r,z) = \bar{g}(\xi_{r-}(x), \lambda N_{r-}(x,\lambda), r, z)N_{r-}(x,\lambda).$$

Since $\int_{\mathcal{Z}} |\bar{g}|^p \nu(dz) \le c_p$ (bounded) for $p \ge 2$, we can show similarly

$$E\left[\int_{t_0}^{t}\Big(\int_{\mathcal{Z}} |g'_\lambda|^2 \nu(dz)\Big)^{p/2} dr\right] \le cE\left[\int_{t_0}^{t} |X'(r-)|^p dr\right],$$

and

$$E\left[\int_{t_0}^{t}\int_{\mathcal{Z}} |g'_\lambda|^p \nu(dz)dr\right] \le cE\left[\int_{t_0}^{t} |X'(r-)|^p dr\right].$$

Consequently we have by Corollary 2.12,

$$E\Big[\sup_{t_0\le s\le t}|X'(s)|^p\Big]\le c\left(1+\int_{t_0}^t E\Big[\sup_{t_0\le s\le r}|X'(s)|^p\Big]dr\right).$$

Therefore we get the $L^p$ estimate (3.10) of $N_t(x,\lambda)$.

We shall next study the $L^p$ estimate of $\tilde{X}(t)=N_t(x,\lambda)-N_t(x',\lambda')$. It is written as

$$\tilde{X}(t)=\int_{t_0}^t\Big(b'_\lambda(r)-b'_{\lambda'}(r)\Big)dr+\int_{t_0}^t\Big(f'_\lambda(r)-f'_{\lambda'}(r)\Big)dW(r)$$
$$+\int_{t_0}^t\int_{\mathcal{Z}}\Big(g'_\lambda(r,z)-g'_{\lambda'}(r,z)\Big)\tilde{N}(drdz).$$

We shall consider the drift term of $\tilde{X}(t)$. It is written as

$$\begin{aligned}
&b'_\lambda(r)-b'_{\lambda'}(r)\\
&=\bar{b}(\xi_{r-}(x),\lambda N_{r-}(x,\lambda),r)N_{r-}(x,\lambda)\\
&\quad-\bar{b}(\xi_{r-}(x'),\lambda' N_{r-}(x',\lambda'),r)N_{r-}(x',\lambda')\\
&=\Big\{\bar{b}(\xi_{r-}(x),\lambda N_{r-}(x,\lambda),r)-\bar{b}(\xi_{r-}(x'),\lambda' N_{r-}(x',\lambda),r)\Big\}N_{r-}(x,\lambda)\\
&\quad+\bar{b}(\xi_{r-}(x'),\lambda' N_{r-}(x',\lambda'),r)\Big\{N_{r-}(x,\lambda)-N_{r-}(x',\lambda')\Big\}.
\end{aligned}$$

Observe that $\bar{b}(x,y,r)$ is $\delta$-Hölder continuous with respect to $x$ and $y$. Then we have

$$\begin{aligned}
&\Big|\bar{b}(\xi_{r-}(x),\lambda N_{r-}(x,\lambda),r)-\bar{b}(\xi_{r-}(x'),\lambda' N_{r-}(x',\lambda),r)\Big|\\
&\le L'\{|\xi_{r-}(x)-\xi_{r-}(x')|^\delta+|\xi_{r-}(x+\lambda e_i)-\xi_{r-}(x)-\\
&\qquad(\xi_{r-}(x'+\lambda' e_i)-\xi_{r-}(x'))|^\delta\}\\
&\le L'\{|\xi_{r-}(x)-\xi_{r-}(x')|^\delta+|\xi_{r-}(x+\lambda e_i)-\xi_{r-}(x'+\lambda' e_i)|^\delta\}.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
&|b'_\lambda(r)-b'_{\lambda'}(r)|\\
&\le L'\{|\xi_{r-}(x)-\xi_{r-}(x')|^\delta+|\xi_{r-}(x+\lambda e_i)-\xi_{r-}(x'+\lambda' e_i)|^\delta\}|N_{r-}(x,\lambda)|\\
&\quad+K'|\tilde{X}(r-)|,
\end{aligned}$$

where we used $|\bar{b}|\le K'$. Note that the $L^{2p}$ norm of $N_{r-}(x,\lambda)$ is uniformly bounded. Then we obtain the estimate

$$E\Big[\int_{t_0}^{t} |b'_\lambda(r) - b'_{\lambda'}(r)|^p dr\Big] \le c_1 \int_{t_0}^{t} E[|\xi_{r-}(x) - \xi_{r-}(x')|^{2\delta p}]^{1/2} dr$$
$$+c_2 \int_{t_0}^{t} E[|\xi_{r-}(x+\lambda e_i) - \xi_{r-}(x'+\lambda' e_i)|^{2\delta p}]^{1/2} dr$$
$$+c_3 \int_{t_0}^{t} E[|\tilde{X}(r-)|^p] dr$$
$$\le c'_1 |x - x'|^{\delta p} + c'_2 |\lambda - \lambda'|^{\delta p} + c_3 \int_{t_0}^{t} E[|\tilde{X}(r-)|^p] dr.$$

In the last inequality, we applied Theorem 3.2.

As to the diffusion coefficient, we have

$$E\left[\int_{t_0}^{t} \left|f'_\lambda(r) - f'_{\lambda'}(r)\right|^p dr\right] \le c'_1 |x - x'|^{\delta p} + c'_2 |\lambda - \lambda'|^{\delta p} + c'_3 \int_{t_0}^{t} E[|\tilde{X}(r-)|^p] dr,$$

similarly for the case of drift.

Next we will estimate the coefficient of jumps. Again we have

$$\left|g'_\lambda(r,z) - g'_{\lambda'}(r,z)\right|$$
$$\le L'(z)\{|\xi_{r-}(x) - \xi_{r-}(x')|^\delta + |\xi_{r-}(x+\lambda e_i) - (\xi_{r-}(x'+\lambda' e_i)|^\delta\} N_{r-}(x,\lambda)$$
$$+K'(z)|\tilde{X}(r-)|,$$

where $\int (L'(z)^p + K'(z)^p)\nu(dz) < \infty$ holds for any $p \ge 2$. Then we have

$$E\left[\int_{t_0}^{t} (\int_{\mathcal{Z}} \left|g'_\lambda(r,z) - g'_{\lambda'}(r,z)\right|^2 \nu(dz))^{p/2} dr\right]$$
$$\le c_4\{|x - x'|^{\delta p} + |\lambda - \lambda'|^{\delta p}\} + c_5 E\left[\int_{t_0}^{t} |\tilde{X}(r-)|^p dr\right].$$

Similarly we obtain

$$E\left[\int_{t_0}^{t} \int_{\mathcal{Z}} \left|g'_\lambda(r,z) - g'_{\lambda'}(r,z)\right|^p \nu(dz) dr\right]$$
$$\le c_6\{|x - x'|^{\delta p} + |\lambda - \lambda'|^{\delta p}\} + c_7 E\left[\int_{t_0}^{t} |\tilde{X}(r-)|^p dr\right].$$

Summing the three $L^p$ estimates, we obtain from Corollary 2.12,

$$E\Big[\sup_{t_0 \le s \le t} |\tilde{X}(s)|^p\Big] \le c_8\{|x - x'|^{\delta p} + |\lambda - \lambda'|^{\delta p}\} + c_9 E\left[\int_{t_0}^{t} |\tilde{X}(r-)|^p dr\right].$$

Therefore we get the inequality

$$E\Big[\sup_{t_0 \le s \le t} |\tilde{X}(s)|^p\Big] \le c_{10}\{|x - x'|^{\delta p} + |\lambda - \lambda'|^{\delta p}\}.$$

This proves the second assertion of the theorem.

**Theorem 3.4.** *Assume the same conditions as in Theorem 3.3 for the coefficients. Then the solution $\xi_t(x)$ is differentiable with respect to $x$ for any $t \geq t_0$ a.s. Further the derivative $\nabla \xi_t$ satisfies*

$$\nabla \xi_t(x) = I + \int_{t_0}^t \nabla X(\xi_{r-}(x), dr)\nabla \xi_{r-}(x), \tag{3.13}$$

*where $I$ is the identity matrix and*

$$\nabla X(x,t) = \int_0^t \nabla b(x,r)dr + \int_0^t \nabla f(x,r)dW(r) + \int_0^t \int_{\mathcal{Z}} \nabla g(x,r,z)\tilde{N}(drdz).$$

*Proof.* We have already shown the differentiability of $\xi_t(x)$.

For the proof of (3.13), observe the formula (3.12). Let $\lambda$ tend to 0 in the coefficients $b'_\lambda(r)$, $f'_\lambda(r)$ and $g'_\lambda(r,z)$. Since $\lambda N_{r-}(x,\lambda)$ converges to 0, we have

$$\lim_{\lambda\to 0} b'_\lambda(r) = \bar{b}(\xi_{r-}(x), 0, r)\nabla \xi_{r-}(x) = \nabla b(\xi_{r-}(x), r)\nabla \xi_{r-}(x).$$

Similarly we have

$$\lim_{\lambda\to 0} f'_\lambda(r) = \bar{f}(\xi_{r-}(x), 0, r)\nabla \xi_{r-}(x) = \nabla f(\xi_{r-}(x), r)\nabla \xi_{r-}(x),$$
$$\lim_{\lambda\to 0} g'_\lambda(z,r) = \nabla g(\xi_{r-}(x), r, z)\nabla \xi_{r-}(x).$$

Therefore we get the equality

$$\begin{aligned}\nabla \xi_t(x) = I &+ \int_s^t \nabla b(\xi_{r-}(x), r)\nabla \xi_{r-}(x)dr \\ &+ \int_s^t \nabla f(\xi_{r-}(x), r)\nabla \xi_{r-}(x)dW(r) \\ &+ \int_s^t \int_{\mathcal{Z}} \nabla g(\xi_{r-}(x), z, r)\nabla \xi_{r-}(x)\tilde{N}(drdz).\end{aligned}$$

This proves (3.13).

### 3.4 Homeomorphic property

The solution of a jump SDE does not admit the homeomorphic property in general. Indeed, at the jump time $\tau$, the solution jumps from $\xi_{\tau-}(x)$ to $\xi_\tau(x) + g(\xi_\tau(x), \tau)$. In order that the homeomorphic property is preserved at the jump time, it is necessary that the map $\phi_{\tau,z} : y \to y + g(y, \tau, z)$ should be homeomorphic, i.e., the map $\phi_{\tau,z}$ should be one to one and onto. In the next section we will show conversely that if $\phi_{\tau,z}$ are homeomorphisms a.e. $\nu \otimes P$, then the solution $\xi_t$ admits the homeomorphic property (Theorem 3.10).

In this section we prove the homeomorphic property under a stronger condition. Let $\psi_{r,z}$ be the inverse map of $\phi_{r,z}$. We assume that the inverses $\{\psi_{r,z}\}$ are uniformly

Lipschitz continuous and of uniformly linear growth; i.e., there exist positive constants $\hat{L}$ and $\hat{K}$ such that

$$|\psi_{r,z}(x)| \leq \hat{K}(1+|x|), \quad |\psi_{r,z}(x) - \psi_{r,z}(y)| \leq \hat{L}|x-y|, \quad \forall (r,z) \ a.e. \tag{3.14}$$

**Theorem 3.5.** *Assume the same condition for coefficients $b, f, g$ as in Theorem 3.3. Assume further that the maps $\phi_{r,z} : x \to x + g(x,r,z); \mathbf{R^d} \to \mathbf{R}^d$ are homeomorphic and the inverse maps $\psi_{r,z}$ are uniformly Lipschitz continuous and of uniformly linear growth. Then the maps $\xi_t : C \to C$ are onto homeomorphisms for any $t$ a.s.*

**Remark.** If we can take Lipschitz constants $L(z)$ and linear growth constants $K(z)$ for jump coefficients $g(x,r,z)$ uniformly less than 1, say $L(z) < \epsilon$ and $K(z) < \epsilon$ for some $0 < \epsilon < 1$, then the inverse maps $\psi_{r,z}(x)$ are uniformly Lipschitz continuous and of uniformly linear growth. Indeed, since $|x - y - (\phi_{r,z}(x) - \phi_{r,z}(x))| \leq L(z)|x-y|$ holds valid, we have

$$|\psi_{r,z}(x) - \psi_{r,z}(y) - (x-y)| \leq L(z)|\psi_{r,z}(x) - \psi_{r,z}(y)|.$$

This implies

$$(1 - L(z))|\psi_{r,z}(x) - \psi_{r,z}(y)| \leq |x-y|.$$

Then $\psi_{r,z}$ is Lipschitz continuous with Lipschitz constant $(1-L(z))^{-1} \leq (1-\epsilon)^{-1}$. Similarly if $K(Z) \leq 1-\epsilon$, then we have $|\psi_{r,z}(x)| \leq (1-\epsilon)^{-1}(1+|x|)$.

For the proof, we need some lemmas. In the following Lemmas 3.6-3.9, we assume that the coefficients $b, f, g$ satisfy the same conditions as in Theorem 3.5.

**Lemma 3.6.** *Set $F(x) = (1+|x|^2)^{-1}$. Then we have the following.*
*1) $F(x+g(x,r,z)) \leq \hat{K}^2 F(x)$ holds for any $x$ a.e. $\Lambda \times \nu \times P$.*
*2) For any $p \geq 2$, there exists a positive constant $c_p$ such that*

$$\int_{\mathcal{Z}} |F(x+g(x,r,z)) - F(x)|^p \nu(dz) \leq c_p F(x)^p, \quad \forall x \ a.e. \ \Lambda \times P. \tag{3.15}$$

*3) There exists a positive constant $c$ such that*

$$\int_{\mathcal{Z}} |F(x+g(x,r,z)) - F(x) - \sum_i F'_{x_i}(x) g^i(x,r,z)|\nu(dz) \tag{3.16}$$
$$\leq cF(x), \quad \forall x a.e. \Lambda \times P.$$

*Proof.* By condition (3.14), it holds $(1+|\psi_{r,z}(y)|^2) \leq (1+\hat{K}^2)(1+|y|)^2$ for any $y$. Substitute $y = \phi_{r,z}(x)$. Since $\psi_{r,z}(\phi_{r,z}(x)) = x$, we have $1+|x|^2 \leq (1+\hat{K}^2)(1+|\phi_{r,z}(x)|^2)$. Then we have

$$\begin{aligned} F(x+g(x,r,z)) &= (1+|\phi_{r,z}(x)|^2)^{-1} \\ &\leq (1+\hat{K}^2)(1+|x|^2)^{-1} \leq (1+\hat{K}^2)F(x). \end{aligned}$$

We next prove (3.15). It holds $F(x+g)-F(x) = -(|g|^2+2(g,x))F(x+g)F(x)$. Since $|x| \le F(x)^{-1/2}$, we have the inequality

$$|F(x+g)-F(x)| \le |g|^2F(x+g)F(x) + 2|g|F(x+g)F(x)^{1/2} \le |g|^2(1+\hat{K}^2)F(x)^2 + 2|g|(1+\hat{K}^2)F(x)^{3/2}. \tag{3.17}$$

Now integrate $p$-th power of both sides with respect to the Lévy measure $\nu$. Then we have

$$\int_{\mathcal{Z}} |F(x+g(x,r,z)) - F(x)|^p \nu(dz) \le c_p F(x)^p.$$

This proves inequality (3.15).

We will consider the third inequality. Since $\sum_i F'_{x_i}(x)g^i = -2(x,g)F(x)^2$, we have

$$F(x+g) - F(x) - \sum_i F'_{x_i}(x)g^i$$
$$= -|g|^2F(x)F(x+g) + 2(x,g)F(x)(F(x)-F(x+g)).$$

Therefore,

$$|F(x+g) - F(x) - \sum_i F'_{x_i}(x)g^i|$$
$$\le |g|^2F(x)F(x+g) + 2|x||g|F(x)|F(x+g)-F(x)|$$
$$\le 5|g|^2(1+\hat{K}^2)F(x)^2 + 2|g|^3(1+\hat{K}^2)F(x)^{5/2}.$$

Now integrate both sides of the above by the Lévy measure $\nu$. Then we obtain (3.16).

**Lemma 3.7.** *For any $p \ge 2$ there exists a positive constant $C_p$ such that*

$$E\Big[\sup_{t_0\le s\le t}(1+|\xi_s(x)|^2)^{-p}\Big] \le C_p(1+|x|^2)^{-p}, \quad \forall x \tag{3.18}$$

*holds for any $t_0 \le t \le T$.*

*Proof.* Set $F(x) = (1+|x|^2)^{-1}$ and $X(t) = \xi_t(x)$. Then,

$$X(t) = x + \int_{t_0}^t b(X(r-),r)dr + \int_{t_0}^t f(X(r-),r)dW(r)$$
$$+ \int_{t_0}^t \int_{\mathcal{Z}} g(X(r-),r,z)\tilde{N}(drdz).$$

We shall apply Corollary 2.13. We want to estimate each term of the right hand side of (2.27). We first consider the drift term. Since $\sum_i F'_{x_i}(x)b^i(x,r) = -2(x,b(x,r))F(x)^2$, we have

$$|\sum_i F'_{x_i}(x)b^i(x,r)| \le 2|x||b(x,r)|F(x)^2 \le 2KF(x).$$

Here we used the property $|b(x,r)| \le K(1+|x|^2)^{1/2}$ (The linear growth of $b(x,r)$). We have similarly,

$$\sum_k |\sum_i F'_{x_i}(x) f^{ik}(x)|^2 \le K' F(x).$$

Further, we have

$$\sum_{i,j,k} F''_{x_i x_j}(x) f^{ik}(x,r) f^{jk}(x,r) = 4F(x)^3 \sum_k (x, f^k(x,r))^2$$
$$-2F(x)^2 \sum_{ik} (f^{ik}(x,r))^2.$$

Using the linear growth property of $f^{jk}$ etc, we have

$$|\sum_{i,j,k} F''_{x_i x_j}(x) f^{ik}(x,r) f^{jk}(x,r)| \le cF(x).$$

Therefore

$$\begin{aligned} &E\Big[\int_{t_0}^t |\sum_i F'_{x_i}(X(r-)) b^i(X(r-),r)|^p dr\Big] \\ &+E\Big[\int_{t_0}^t \Big(\sum_k |\sum_i F'_{x_i}(X(r-)) f^{ik}(X(r-),r)|^2\Big)^{p/2} dr\Big] \\ &+\frac{1}{2} E\Big[\int_{t_0}^t \Big|\sum_{i,j,k} F''_{x_i x_j}(X(r-)) f^{ik}(X(r-),r) f^{jk}(X(r-),r)\Big|^p dr\Big] \\ &\le cE\Big[\int_{t_0}^t (1+|X(r-)|^2)^{-p} dr\Big]. \end{aligned}$$

We shall next consider the jump parts. We get from Lemma 3.6 (2),

$$E\Big[\int_{t_0}^t \Big(\int_{\mathcal{Z}} |F(X(r-)+g(X(r-),r,z)) - F(X(r-))|^{p'} \nu(dz)\Big)^{p/p'} dr\Big]$$
$$\le CE\Big[\int_{t_0}^t (1+|X(r-)|^2)^{-p} dr\Big],$$

for $p'=2$ and $p'=p$. Next, we have from Lemma 3.6 (3),

$$\begin{aligned} &E\Big[\int_{t_0}^t \Big(\int_{\mathcal{Z}} |F(X(r-)+g(X(r-),r,z)) \\ &\quad -F(X(r-)) - \sum_i F'_{x_i}(X(r-)) g^i(X(r-),r,z)|\nu(dz)\Big)^p dr\Big] \\ &\le C\int_{t_0}^t E[(1+|X(r-)|^2)^{-p}] dr. \end{aligned}$$

Summing up all these estimates, we get a functional inequality

$$E\Big[\sup_{t_0\le s\le t}(1+|X(s)|^2)^{-p}\Big]\le C\left\{(1+|x|^2)^{-p}+\int_{t_0}^t E[(1+|X(r-)|^2)^{-p}]dr\right\}.$$

This implies

$$E\Big[\sup_{t_0<s\le t}(1+|\xi_s(x)|^2)^{-p}\Big]\le C_p(1+|x|^2)^{-p},$$

where the constant $C_p$ does not depend on $x$.

**Lemma 3.8.** *Set $F(x)=(\delta+|x|^2)^{-1}$. Then we have the following.*
*1) $F(x-x'+g(x,r,z)-g(x',r,z))\le(1+\hat{L}^2)F(x-x')$ holds for all $x,x'$, a.e. $\Lambda\times\nu\times P$.*
*2) For any $p\ge 2$, there exists a positive constant $c'_p$ not depending on $\delta$ such that*

$$\int_{\mathcal{Z}}|F(x-x'+g(x,r,z)-g(x',r,z))-F(x-x')|^p\nu(dz)\qquad(3.19)$$
$$\le c'_pF(x-x')^p,\quad\forall x,x'\text{ a.e. }\Lambda\times P.$$

*3) There exists a positive constant $c'$ not depending on $\delta$ such that*

$$\int_{\mathcal{Z}}|F(x-x'+g(x,r,z)-g(x',r,z))-F(x-x')\qquad(3.20)$$
$$-\sum_i F'_{x_i}(x-x')(g^i(x,r,z)-g^i(x',r,z))|\nu(dz)\le c'F(x-x'),$$
$$\forall x,x'\text{ a.e. }\Lambda\times P.$$

*Proof.* We shall prove the first assertion. Since $\psi_{r,z}(y)$ is Lipschitz continuous, we have

$$|\psi_{r,z}(y)-\psi_{r,z}(y')|\le\hat{L}|y-y'|,\quad\forall y,y',$$

Now substitute $y=\phi_{r,z}(x)$ and $y'=\phi_{r,z}(x')$ in the above inequality and note the identities $x=\psi_{r,z}(\phi_{r,z}(x))$ and $x'=\psi_{r,z}(\phi_{r,z}(x'))$. Then we obtain $|x-x'|\le\hat{L}|\phi_{r,z}(x)-\phi_{r,z}(x')|$ for any $x,x'$. Then it holds $\delta+|x-x'|^2\le(1+\hat{L}^2)(\delta+|\phi_{r,z}(x)-\phi_{r,z}(x')|^2)$ for any $\delta>0$, which implies

$$\begin{aligned}F(x-x'+g(x,z)-g(x',z))&=(\delta+|\phi_{r,z}(x)-\phi_{r,z}(x')|^2)^{-1}\\&\le(1+\hat{L}^2)(\delta+|x-x'|^2)^{-1}\\&=(1+\hat{L}^2)F(x-x'),\quad\forall x,x'.\end{aligned}$$

Next we will prove (3.19). Set $w=x-x'$ and $k=g(x,r,z)-g(x',r,z)$. Then we have $F(w+k)\le(1+\hat{L}(z)^2)F(w)$ by the first assertion. Also we have the inequality

$$|F(w+k)-F(w)|\le|k|^2(1+\hat{L}^2)F(w)^2+2|k|(1+\hat{L}^2)F(w)^{3/2}$$

as in Lemma 3.6. Now integrate the $p$-th power of both sides of the above with respect to the Lévy measure $\nu$. Then we have

$$\begin{aligned}
&\int_{\mathcal{Z}} |F(w+k) - F(w)|^p \nu(dz) \\
&\le 2^p \int_{\mathcal{Z}} |g(x,r,z) - g(x',r,z)|^{2p} (1+\hat{L}^2)^p \nu(dz) F(w)^{2p} \\
&\quad +4^p \int_{\mathcal{Z}} |g(x,r,z) - g(x',r,z)|^p (1+\hat{L}^2)^p \nu(dz) F(w)^{3p/2} \\
&\le c'_p F(w)^p,
\end{aligned}$$

because

$$\int_{\mathcal{Z}} |g(x,r,z) - g(x',r,z)|^p (1+\hat{L}^2)^{p'} \nu(dz) \le c''_p |x-x'|^p \le c''_p F(y)^{-p/2}.$$

The above inequality proves (3.19). Inequality (3.20) can be verified similarly as in the proof of Lemma 3.6.

**Lemma 3.9.** *For any $p \ge 2$ there exists a positive constant $C'_p$ such that*

$$E\Big[\sup_{t_0 \le s \le t} |\xi_s(x) - \xi_s(y)|^{-2p}\Big] \le C'_p |x-y|^{-2p}, \quad \forall x, y \tag{3.21}$$

*holds for any $t_0 \le t \le T$.*

*Proof.* Set $F(x) = (\delta + |x|^2)^{-1}$ and $Y(t) = \xi_t(x) - \xi_t(y)$. In the following arguments, all constants $c_i$ will be chosen independently of $\delta > 0$ and $x, y \in \mathbf{R}^d$. Using Lipschitz conditions for $b$, $f$ we can show

$$\begin{aligned}
&E\Big[\int_{t_0}^t |\sum_i F'_{x_i}(Y(r-))(b^i(\xi_{r-}(x), r) - b^i(\xi_{r-}(y), r))|^p dr\Big] \\
&\quad +E\Big[\int_{t_0}^t |\sum_k |\sum_i F'_{x_i}(Y(r-))(f^{ik}(\xi_{r-}(x), r) - f^{ik}(\xi_{r-}(y), r))|^2|^{p/2} dr\Big] \\
&\quad +\frac{1}{2} E\Big[\int_{t_0}^t |\sum_{i,j,k} F''_{x_i x_j}(Y(r-))(f^{ik}(\xi_{r-}(x), r) - f^{ik}(\xi_{r-}(y), r)) \\
&\qquad \times (f^{jk}(\xi_{r-}(x), r) - f^{jk}(\xi_{r-}(y), r))|^p dr\Big] \\
&\le c_1 E\Big[\int_{t_0}^t (\delta + |Y(r-)|^2)^{-p} dr\Big].
\end{aligned}$$

We next consider jump parts. Using Lemma 3.8 (2), we can similarly show

$$E\Big[\int_{t_0}^t \Big(\int_{\mathcal{Z}} |F\Big(Y(r-) + g(\xi_{r-}(x), r, z) - g(\xi_{r-}(y), r, z)\Big) \\ - F\Big(Y(r-)\Big)|^{p'}\nu(dz)\Big)^{p/p'} dr\Big] \\ \le c_2 \int_{t_0}^t E[(\delta + |Y(r-)|^2)^{-p}]dr,$$

for $p' = p$ or $p' = 2$. Next, using Lemma 3.8 (3), we get

$$E\Big[\int_{t_0}^t \Big(\int_{\mathcal{Z}} \Big|F\Big(Y(r-) + g(\xi_{r-}(x), r, z) - g(\xi_{r-}(y), r, z)\Big) - F\Big(Y(r-)\Big) \\ -\Big(\sum_i F'_{x_i}(Y(r-))(g^i(\xi_{r-}(x), r, z) - g^i(\xi_{r-}(y), r, z))\Big)\Big|\nu(dz)\Big)^p dr\Big]$$

$$\le c_3 \int_{t_0}^t E[(\delta + |Y(r-)|^2)^{-p}]dr.$$

These computations yield

$$E\Big[\sup_{t_0 \le s \le t} (\delta + |Y(s-)|^2)^{-p}\Big] \le c_4 \Big\{(\delta + |x - y|^2)^{-p} \\ + \int_{t_0}^t E\Big[(\delta + |Y(r-)|^2))^{-p}\Big]dr\Big\}.$$

This implies

$$E\Big[\sup_{t_0 \le s \le t} (\delta + |Y(s)|^2)^{-p}\Big] \le c_5(\delta + |x - y|^2)^{-p}.$$

Now let $\delta$ tend to 0 and observe $Y(s) = \xi_s(x) - \xi_s(y)$. Then we get the inequality

$$E\Big[\sup_{t_0 \le s \le t} |\xi_s(x) - \xi_s(y)|^{-2p}\Big] \le c_5|x - y|^{-2p}.$$

Therefore we get the lemma.

*Proof of Theorem 3.5.* We first show that maps $\xi_t : \mathbf{R}^d \to \mathbf{R}^d$ are one to one for any $t$ a.s. Consider the random field

$$\eta_t(x, y) = \frac{1}{|\xi_t(x) - \xi_t(y)|}.$$

We can show that for any $p \ge 2$, there exists a positive constant $C = C(p)$ such that for any $\delta > 0$,

$$E\Big[\sup_{t_0 < s \le t} |\eta_s(x, y) - \eta_s(x', y')|^{2p}\Big] \le C\delta^{-4p}\{|x - x'|^{2p} + |y - y'|^{2p}\}$$

holds for any $x, x', y, y'$ such that $|x - y| \ge \delta$ and $|x' - y'| \ge \delta$. In fact, a simple computation yields

$$\begin{aligned}&|\eta_t(x, y) - \eta_t(x', y)|^{2p}\\&\le \eta_t(x, y)^{2p}\eta_t(x', y')^{2p} \times \{|\xi_t(x) - \xi_t(x')| + |\xi_t(y) - \xi_t(y')|\}^{2p}.\end{aligned}$$

Take expectations for both sides and use Hölder's inequality. Then we obtain the inequality.

Now, by Kolmogorov's criterion, the random field $\eta_t(x, y)$ is continuous in the domain $D_\delta = \{(x, y); |x - y| \ge \delta\}$. Since this is valid for any $\delta$, we find that the random field $\eta_t(x, y)$ is continuous in the domain $\{(x, y) : x \ne y\}$. This proves that the maps $\xi_t; \mathbf{R}^d \to \mathbf{R}^d$ are one to one for any $t$, a.s.

We will next prove that the maps $\xi_t : \mathbf{R}^d \to \mathbf{R}^d$ are onto for any $t$ a.s. Set $\hat{x} = |x|^{-2}x$ if $x \ne 0$ and define

$$\eta_t(\hat{x}) = \frac{1}{1 + |\xi_t(x)|} \quad \text{if } \hat{x} \ne 0.$$

We set $\eta_t(0) = 0$. Then for each $p > 1$, there exists a positive constant $C = C(p)$ such that

$$E\Big[\sup_{t_0\le s\le t} |\eta_s(x) - \eta_s(y)|^{2p}\Big] \le C|\hat{x} - \hat{y}|^{2p}.$$

In fact we have

$$|\eta_s(\hat{x}) - \eta_s(\hat{y})|^{2p} \le \eta_s(\hat{x})^{2p}\eta_s(\hat{y})^{2p}|\xi_s(x) - \xi_s(y)|^{2p}.$$

Take the supremum with respect to $s$ and then take expectations for both sides. Then we have

$$E\Big[\sup_{t_0\le s\le t} |\eta_s(\hat{x}) - \eta_s(\hat{y})|^{2p}\Big] \le C(1 + |x|)^{-2p}(1 + |y|)^{-2p}|x - y|^{2p} \le C|\hat{x} - \hat{y}|^{2p}.$$

By Kolmogorov's criterion, $\eta_t(\hat{x})$ can be extended continuously to 0. This means that $\eta_t(\hat{x})$ converges to 0 for any $t$ a.s. as $\hat{x} \to 0$. But this implies $\lim_{|x|\to\infty} |\xi_t(x)| = \infty$ exists for any $t$ a.s. This establishes the onto property of the maps.

Let $t_0 < s < t$. Let $\xi'_t(x)$ be the solution starting from $x$ at time $s$. We define $\eta_r, t_0 < r < T$ by $\eta_r = \xi_r(x)$ if $r \le s$ and $\eta_r = \xi'_r(\xi_s(x))$ if $r > s$. Substitute $y = \xi_s(x)$ in the equality $\xi'_t(y) = y + \int_s^t X(\xi'_{r-}(y), dr)$. Then $\eta_r$ satisfies

$$\eta_t = \xi_s(x) + \int_s^t X(\eta_{r-}, dr) = x + \int_{t_0}^t X(\eta_{r-}, dr).$$

Therefore $\eta_t = \xi_t$ holds. Consequently we have the cocycle property $\xi'_t \circ \xi_s = \xi_t$ for any $t_0 < s < t$ a.s.

Now for $s, t \in [t_0, T]$, we will define $\xi_{s,t}$ as follows. Let $\xi_t$ be the $C$-valued cadlag process with the initial condition $\xi_0(x) = x$. Let $\xi_t^{-1}$ be the inverse map. It is again a $C$-valued cadlag process. We set

$$\begin{aligned}\xi_{s,t}(x) &= x \quad \text{if } t \le s,\\&= \xi_t \circ \xi_s^{-1}(x), \quad \text{if } t \ge s.\end{aligned}$$

Then it has the following properties:
1) $\xi_{s,t}$ is a cadlag process with respect to $t$ and also is a cadlag process with respect to $s$
2) $\xi_{s,t} : C \to C$ are onto homeomorphisms for all $s < t$ a.s.
3) $\xi_{s,t}(x)$ is the solution of the SDE based on $X(x,t)$ starting from $x$ at time $s$.

The third property follows from the fact that $\xi_{s,t} = \xi_t^{\prime}$ holds a.s. The family of $C$-valued random variables $\{\xi_{s,t}; 0 \le t_0 \le s < t < T\}$ is called a *stochastic flow of homeomorphisms generated by* $X(x,t)$. Further if the map $\xi_{s,t} : \mathbf{R}^d \to \mathbf{R}^d$ are $C^k$-diffeomorphisms for all $s < t$, then it is called a *stochastic flow of $C^k$-diffeomorphisms.*

### 3.5 Stochastic flow of diffeomorphisms

So far we have studied the solution of a SDE through $L^p$ estimates. Conditions needed for this framework are that coefficients $b, f, g$ satisfy (3.1) and (3.2). In this section we will relax the latter condition (3.2). For $0 < \epsilon < 1$, set $U_\epsilon = \{z; K(z) < \epsilon, L(z) < \epsilon\}$. Instead of (3.2), we introduce a condition

$$\int_{U_\epsilon} (K(z)^2 + L(z)^2)\nu(dz) < \infty, \quad \text{and } \nu(U_\epsilon^c) < \infty. \tag{3.22}$$

We define

$$\begin{aligned} X'(x,t) = & \int_0^t b(x,r)dr + \int_0^t f(x,r)dW(r) \\ & + \int_0^t \int_{U_\epsilon} g(x,r,z)\tilde{N}(drdz), \end{aligned} \tag{3.23}$$

$$X(x,t) = X'(x,t) + \int_0^t \int_{U_\epsilon^c} g(x,r,z)N(drdz), \tag{3.24}$$

and consider a SDE based on $X(x,t)$ and $X'(x,t)$. Observe that $X'(x,t)$ satisfies all the conditions required in Theorems 3.1 and 3.2. Further assuming that the maps $\phi_{r,z}$ are homeomorphisms, the solution of the SDE based on $X'(x,t)$ defines a stochastic flow $\xi_{s,t}^{\prime}$ of homeomorphisms. (See Remark after Theorem 3.5.) We shall consider a SDE based on $X(x,t)$. Using the notation of the point process $p(t)$, the equation is written as

$$\begin{aligned} \xi_t &= x + \int_{t_0}^t X(\xi_{r-}, dr) \\ &= x + \int_{t_0}^t X'(\xi_{r-}, dr) + \sum_{t_0 < r \le t, r \in \mathbf{D}_p} g(\xi_{r-}, r, p(r)) 1_{U_\epsilon^c}(p(r)). \end{aligned} \tag{3.25}$$

The solution of the above equation can be constructed by means of the stochastic flow of homeomorphisms $\xi_{s,t}^{\prime}(x)$ generated by $X'(x,t)$ and jumps $g(\xi_{r-}(x), r, p(r))1_{U_\epsilon^c}(p(r))$. Let $0 = \sigma_0 < \sigma_1 < \cdots$ be a sequence of jump times of the Poisson process $N(t) = N((0,t] \times U_\epsilon^c), t \in [0,T]$. Then $P(\sigma_k < T) \to 0$ as $k \to \infty$. We let $\phi_{r,z}(x) = x + g(x,r,z)$ and define $\xi_{t_0,t}$ by

$$\xi_{t_0,t}(x) = \xi'_{\sigma_m,t} \circ \phi_{\sigma_m,p(\sigma_m)} \circ \cdots \circ \xi'_{\sigma_{n+1},\sigma_{n+2}} \circ \phi_{\sigma_{n+1},p(\sigma_{n+1})} \circ \xi'_{t_0,\sigma_{n+1}}(x), \tag{3.26}$$

on the set $A_{n,m}(t_0, t) = \{\omega; \sigma_n \le t_0 < \sigma_{n+1}, \sigma_m \le t < \sigma_{m+1}\}$.

Then $\xi_{t_0,t}$ is a cadlag process with values in $C$ with respect to $t_0$ and $t$. Further, $\xi_{t_0,t} : \mathbf{R}^d \to \mathbf{R}^d$ are homeomorphisms for any $t_0$, $t$, since each of them are compositions of homeomorphisms $\xi'_{r,u}$ and $\phi_{r,z}$. Further the above $\xi_{t_0,t}(x)$ is a solution of (3.25). We will check it on the set $A_{n,n+1}(t_0, t)$. Since $dX = dX'$ holds on $(\sigma_n, \sigma_{n+1})$, we see that

$$x + \int_{t_0}^{\sigma_{n+1}-} X(\xi_{t_0,r-}(x), dr) = x + \int_{t_0}^{\sigma_{n+1}} X'(\xi'_{t_0,r-}(x), dr) = \xi'_{t_0,\sigma_{n+1}}(x).$$

Moreover, since

$$\int_{\{\sigma_{n+1}\}} X(\xi'_{t_0,r-}(x), dr) = \phi_{\sigma_{n+1},p(\sigma_{n+1})}(\xi'_{t_0,\sigma_{n+1}}(x)) - \xi'_{t_0,\sigma_{n+1}}(x),$$

we have

$$x + \int_{t_0}^{\sigma_{n+1}+} X(\xi_{t_0,r-}(x), dr) = \xi_{t_0,\sigma_{n+1}}(x).$$

Also, observe that $dX = dX'$ on $(\sigma_{n+1}, t]$. Then

$$\begin{aligned}\xi_{t_0,\sigma_{n+1}}(x) &+ \int_{\sigma_{n+1}}^{t} X(\xi_{t_0,r-}(x), dr) \\ &= \xi_{t_0,\sigma_{n+1}}(x) + \int_{\sigma_{n+1}}^{t} X(\xi_{\sigma_{n+1},r-}(\xi_{t_0,\sigma_{n+1}}(x)), dr) \\ &= \{y + \int_{\sigma_{n+1}}^{t} X'(\xi'_{\sigma_{n+1},r-}(y), dr)\}|_{y=\xi_{t_0,\sigma_{n+1}}(x)}.\end{aligned}$$

Furthermore,

$$y + \int_{\sigma_{n+1}}^{t} X'(\xi'_{\sigma_{n+1},r-}(y), dr) = \xi'_{\sigma_{n+1},t}(y).$$

Consequently we obtain

$$\begin{aligned}x + \int_{t_0}^{t} X(\xi_{t_0,r-}(x), dr) &= \xi_{t_0,\sigma_{n+1}}(x) + \int_{\sigma_{n+1}}^{t} X(\xi_{t_0,r-}(x), dr) \\ &= \xi'_{\sigma_{n+1},t}(\xi_{t_0,\sigma_{n+1}}(x)) = \xi_{t_0,t}(x).\end{aligned}$$

We have thus proved the following theorem.

**Theorem 3.10.** *Assume that coefficients $b$, $f$, $g$ satisfy (3.1) and (3.22). Assume further that $\phi_{r,z}(x) = x + g(x, r, z)$ are homeomorphic for any $r$, $z$. Then the solution defined by (3.26) is a stochastic flow of homeomorphisms.*

Next we shall study the diffeomorphic property of the maps $\xi_{t_0,t}$. We introduce

$$\int_{U'_\epsilon} (K'(z)^2 + L'(z)^2)\nu(dz) < \infty, \quad \text{and } \nu(U_\epsilon^{'c}) < \infty, \tag{3.27}$$

where $U_\epsilon = \{z; K'(z) < \epsilon, L'(z) < \epsilon\}$. Then, under conditions (3.1), (3.22), (3.8) and (3.27), $\xi_{t_0,t}$ of (3.26) is a $C^1$-valued process, cadlag with respect to $s$ and $t$. Further, we can show that the Jacobian matrix $\nabla\xi_{t_0,t}$ satisfies (3.13).

**Theorem 3.11.** *Assume that coefficients of the equation satisfy (3.1), (3.22), (3.8) and (3.27). If $\phi_{r,z)}(x) := x + g(x, r, z)$ are homeomorphic and Jacobian matrix $I + \nabla g(x, r, z)$ is invertible for any $x$ a.e. $(r, z)$, then the solution defines a stochastic flow of $C^1$-diffeomorphisms.*

By Theorem 3.10, the solution defines a stochastic flow of homeomorphisms. Therefore, it is sufficient to show that the Jacobian matrix $\nabla\xi_{t_0,t}$ is invertible (locally diffeomorphic).

We define a matrix valued semimartingale $U(t)$ (with parameter $x$) by

$$U(t) = \int_{t_0}^t \nabla b(\xi_{t_0,r-}(x), r)dr + \int_{t_0}^t \nabla f(\xi_{t_0,r-}(x), r)dW(r) + \int_{t_0}^t \int_{\mathcal{Z}} \nabla g(\xi_{t_0,r-}(x), r, z)\tilde{N}(drdz).$$

Then, in view of Theorem 3.4, $\Phi_t := \nabla\xi_{t_0,t}(x)$ satisfies the linear SDE

$$\Phi_t = I + \int_{t_0}^t dU(r)\Phi_{r-}.$$

Observe that $\Delta U(r) = \nabla g(\xi_{r-}(x), r, p(r))$. Now the Jacobian matrix of $\phi_{r,z}(x) := x + g(x, r, z)$ (with respect to $x$) is given by $I + \nabla g(x, r, z)$. It is invertible for any $x$ a.e. $(r, z)$ by the assumption of the theorem. Therefore the matrices $(I + \Delta U(r))$ are invertible a.s.

Associated with $U(t)$, we define a matrix valued process by

$$V(t) = -U(t) + [U, U]_t^c + \sum_{t_0<r\le t} (I + \Delta U(r))^{-1}(\Delta U(r))^2.$$

The invertability of the matrix $\Phi_t \equiv \nabla\xi_{t_0,t}$ follows from the next lemma.

**Lemma 3.12 (Protter).** [20] *Let $\Psi_t$ be a matrix valued process satisfying*

$$\Psi_t = I + \int_{t_0}^t \Psi_{r-}dV(r).$$

*Then we have $\Psi_t\Phi_t = I$. In particular, $\Phi_t$ is invertible for any $t$.*

*Proof.* Apply Itô's formula I to the product of $\Psi_t$ and $\Phi_t$. Then we have

$$\Psi_t\Phi_t = I + \int_{t_0}^t \Psi_{r-}d\Phi_r + \int_{t_0}^t d\Psi_r\Phi_{r-} + [\Psi, \Phi]_t,$$

where

$$\int_{t_0}^{t} \Psi_{r-} d\Phi_r = \int_{t_0}^{t} \Psi_{r-} dU(r)\Phi_{r-},$$

$$\begin{aligned}\int_{t_0}^{t} d\Psi_r \Phi_{r-} &= -\int_{t_0}^{t} \Psi_{r-} dU(r)\Phi_{r-} \\ &\quad + \int_{t_0}^{t} \Psi_{r-} d[U,U]_r^c \Phi_{r-} \\ &\quad + \sum_{t_0<r\le t} \Psi_{r-}(I+\Delta U(r))^{-1}(\Delta U(r))^2 \Phi_{r-}.\end{aligned}$$

Note that

$$\begin{aligned}[V,U]_t &= -[U,U]_t + \sum_{t_0<r\le t} (I+\Delta U(r))^{-1}(\Delta U(r))^3 \\ &= -[U,U]_t^c - \sum_{t_0<r\le t} (\Delta U(r))^2 + \sum_{t_0<r\le t} (I+\Delta U(r))^{-1}(\Delta U(r))^3 \\ &= -[U,U]_t^c - \sum_{t_0<r\le t} (I+\Delta U(r))^{-1}(\Delta U(r))^2,\end{aligned}$$

because $(\Delta U(r))^2 = \{(\Delta U(r))^2 + (\Delta U(r))^3\}(I+\Delta U(r))^{-1}$. Then,

$$\begin{aligned}[\Psi,\Phi]_t &= \int_{t_0}^{t} \Psi_{r-} d[V,U]_r \Phi_{r-} \\ &= -\int_{t_0}^{t} \Psi_{r-} d[U,U]_r^c \Phi_{r-} \\ &\quad - \sum_{t_0<r\le t} \Psi_{r-}(I+\Delta U(r))^{-1}(\Delta U(r))^2 \Phi_{r-}.\end{aligned}$$

Then all terms of $\int_{t_0}^{t} \Psi_{r-} d\Phi_{r-} + \int_{t_0}^{t} d\Psi_r \Phi_{r-} + [\Psi,\Phi]_t$ are cancelled to yield 0. Then we get $\Psi_t \Phi_t = I$.

### 3.6 Inverse flow and backward SDE

Let $\xi_{s,t}(x)$ be the stochastic flow of diffeomorphisms generated by the SDE

$$\xi_{s,t}(x) = x + \int_{s}^{t} X(\xi_{s,r-}(x), dr), \tag{3.28}$$

where

$$X(x,t) = \int_{t_0}^{t} b(x,r)dr + \int_{t_0}^{t} f(x,r)dW(r) + \int_{t_0}^{t}\int_{\mathcal{Z}} g(x,r,z)\tilde{N}(drdz). \tag{3.29}$$

In this section we assume that the coefficients $b$, $f$, $g$ are deterministic functions. Then the flow $\xi_{s,t}$ has independent increments, i.e., $\xi_{t_i,t_{i+1}}$, $i = 0, \ldots, n-1$, are independent whenever $t_0 < t_1 < \cdots < t_n$. We call $\xi_{s,t}$ a *Lévy flow*.

Let $\xi_{s,t}^{-1}$ be the inverse map of $\xi_{s,t}$. It has the backward flow property: $\xi_{s,t}^{-1} = \xi_{s,r}^{-1} \circ \xi_{r,t}^{-1}$. In this section we fix the time $t$ and consider $\xi_{s,t}^{-1}(x)$ as a stochastic process with time parameter $s$. It is a cadlag process, because it holds $\xi_{s,t}^{-1} = \xi_{t_0,s} \circ \xi_{t_0,t}^{-1}$ and $\xi_{t_0,s}$ is a cadlag process with respect to $s$. The object of this section is to find a backward SDE governing the inverse flow $\xi_{s,t}^{-1}$ or to find the backward infinitesimal generator $\hat{X}(x,t)$ of the inverse flow.

We need to define a backward integral. Let $Z(t)$ be the Lévy process and let $\mathcal{F}_{s,t}$ be the complete sub $\sigma$-field generated by $\{Z(u) - Z(v); s \le u \le v \le t\}$. Then it is a filtration to the backward direction, i.e, $\mathcal{F}_{s,t} \subset \mathcal{F}_{s',t'}$ if $(s,t] \subset (s',t']$ and is left continuous with respect to $s$.

In the following we will fix the time $t$. A stochastic process $f(r)$, $0 < r < t$ is called *backward adapted* if $f(r)$ is $\mathcal{F}_{r,t}$-measurable for any $r < t$. The backward predictable process can be defined similarly. If $f(r)$ is right continuous and is backward adapted, then it is backward predictable. A left continuous backward adapted process $Z(r)$ is called a *backward martingale* if it is integrable and satisfies

$$X(r) = E[X(s)|\mathcal{F}_{r,t}], \quad \forall s < r < t.$$

The *backward localmartingale* and *backward semimartingale* can be defined similarly.

We shall define a backward Itô integral. Let $W(t)$ be a standard Brownian motion of the previous section. Let $f(r)$ be a right continuous backward adapted process. Then the *backward Itô integral* is defined by

$$\int_s^t f(r)\hat{d}W(r) = \lim_{|\Pi|\to 0} \sum_k f(t_{k+1})(W(t_{k+1}) - W(t_k)),$$

where $\Pi = \{s = t_0 < t_1 < \cdots < t_n = t\}$ are partitions of the interval $[s,t]$. It is a continuous backward localmartingale. Next let $N(drdz)$ be a Poisson random measure. If $g(r,z)$ is a right continuous backward predictable process, we can define the backward Poisson integral by

$$\int_s^t \int_{\mathcal{Z}} g(r,z)\tilde{N}(\hat{d}r, dz).$$

It is a left continuous backward localmartingale.

Now if $\xi_{s,t}$ is a Lévy flow, $\xi_{r,t}$ and its inverse $\xi_{r,t}^{-1}$ are backward adapted. Then we may define the backward integrals such as $\int_s^t f(\xi_{r,t}^{-1}(y), r)\hat{d}W(r)$ etc. We use the notation

$$\begin{aligned}\int_s^t X(\xi_{r,t}^{-1}(y), \hat{d}r) &:= \int_s^t b(\xi_{r,t}^{-1}(y), r)dr \\ &+ \int_s^t f(\xi_{r,t}^{-1}(y), r)\hat{d}W(r) + \int_s^t \int_{\mathcal{Z}} g(\xi_{r,t}^{-1}(y), r, z)\tilde{N}(\hat{d}r, dz).\end{aligned} \tag{3.30}$$

It is left continuous with respect to $s$.

Now let $\psi_{r,z}$ be the inverse map of $\phi_{r,z} : x \to x + g(x,r,z)$, and set

$$h(x,r,z) = -\psi_{r,z}(x) + x. \tag{3.31}$$

Then we have

$$\begin{aligned} g(x,r,z) - g(\psi_{r,z}(x), r, z) &= g(x,r,z) - \{\phi_{r,z}(\psi_{r,z}(x)) - \psi_{r,z}(x)\} \\ &= g(x,r,z) - h(x,r,z). \end{aligned}$$

**Theorem 3.13.** *Assume the same condition as in Theorem 3.5 for coefficients $b, f, g$. Assume further that the diffusion coefficient $f(x,r)$ is $C^{2,1}$ with respect to $(x,t)$ and the integral $j(x,r) := \int_U |g(x,r,z) - h(x,r,z)|\nu(dz)$ is bounded for some $U \subset \mathcal{Z}$ such that $\nu(U^c) < \infty$. Then the inverse flow satisfies the following backward SDE*

$$\xi_{s,t}^{-1}(y) = y - \int_s^t \hat{X}(\xi_{r,t}^{-1}(y), \hat{d}r) \quad a.s., \tag{3.32}$$

*for any $s < t$, where*

$$\hat{X}(x,t) = X(x,t) - 2\int_{t_0}^t c(x,r)dr - \int_{t_0}^t \int_{\mathcal{Z}} \{g(x,r,z) - h(x,r,z)\} N(\hat{d}rdz), \tag{3.33}$$

*and*

$$c(x,r) = \frac{1}{2}\sum_{ij} \frac{\partial f^j(x,r)}{\partial x_i} f^{ij}(x,r). \tag{3.34}$$

**Remark.** 1) The stochastic integral in the right hand side of (3.32) is defined as a left continuous backward semimartingale with respect to $s$, while the process $\xi_{s,t}^{-1}(y)$ is right continuous with respect to $s$. Hence the equality holds a.s. for each fixed $s$. We may take a modification of the stochastic integral in the right hand side in such a way that it is left continuous with respect to $t$. Then the equality

$$\xi_{s-,t-}^{-1}(y) = y - \int_s^t \hat{X}(\xi_{r,t}^{-1}(y), \hat{d}r), \tag{3.35}$$

holds for any $s < t$ a.s.

2) Assume further that $h(x,r,z)$ satisfies $\int_{\mathcal{Z}} |g(x,r,z) - h(x,r,z)|\nu(dz) < \infty$ and $\int_{\mathcal{Z}} |h(x,r,z)|^2 \nu(dz) < \infty$ for any $(x,r)$. Then $\hat{X}(x,t)$ is rewritten as

$$\begin{aligned} \hat{X}(x,t) = &\int_{t_0}^t b(x,r)dr + \int_{t_0}^t f(x,r)dW(r) \\ &+ \int_{t_0}^t \int_{\mathcal{Z}} h(x,r,z)\tilde{N}(drdz) \\ &-2\int_{t_0}^t c(x,r)dr - \int_{t_0}^t \int_{\mathcal{Z}} (g(x,r,z) - h(x,r,z))\hat{N}(drdz). \end{aligned} \tag{3.36}$$

Therefore in the inverse flow, the diffusion coefficient $f$ is not changed, while the rule of jumps is changed from $g(x,r,z)$ to $h(x,r,z)$. Further the drift coefficient is changed from $b$ to $b - 2c - \int_{\mathcal{Z}}(g-h)\nu(dz)$.

3) In the case where $\int_{\mathcal{Z}} |g|\nu(dz) < \infty$, the integrals $\int_{t_0}^t \int_{\mathcal{Z}} g\tilde{N}(drdz)$ and $\int_{t_0}^t \int_{\mathcal{Z}} h\tilde{N}(drdz)$ are split into integrals with respect to $N(drdz)$ and $dr\nu(dz)$. Then the above $X$ and $\hat{X}$ are rewritten as

$$X(x,t) = \int_{t_0}^t b'(x,r)dr + \int_{t_0}^t f(x,r)dW(r) + \int_{t_0}^t \int_{\mathcal{Z}} g(x,r,z)N(drdz),$$

$$\begin{aligned}\hat{X}(x,t) &= \int_{t_0}^t b'(x,r)dr + \int_{t_0}^t f(x,r)dW(r) \\ &\quad + \int_{t_0}^t \int_{\mathcal{Z}} h(x,r,z)N(drdz) - 2\int_{t_0}^t c(x,r)dr,\end{aligned}$$

where we set $b' = b - \int_{\mathcal{Z}} g(z)\nu(dz)$. In this case the drift coefficients are changed from $b'$ to $b' - 2c$. We may observe that the correction term $2c$ appears because of Itô's SDE. Indeed, the flow $\xi_{s,t}$ satisfies the following Stratonovich SDE

$$\xi_t = \xi_0 + \int_{t_0}^t b''(\xi_r, r)dr + \int_{t_0}^t f(\xi_r, r) \circ dW(r) + \int_{t_0}^t \int_{\mathcal{Z}} g(\xi_r, r, z)N(drdz),$$

where $b'' = b' - c$ and $\circ$ denotes the Stratonovich integral. Then the inverse flow satisfies

$$\begin{aligned}\xi_{s,t}^{-1}(y) &= y - \int_s^t b''(\xi_{r,t}^{-1}(y), r)dr \\ &\quad - \int_s^t f(\xi_{r,t}^{-1}(y), r) \circ \hat{d}W(r) - \int_s^t \int_{\mathcal{Z}} h(\xi_{r,t}^{-1}(y), r)N(\hat{d}rdz).\end{aligned}$$

It will be shown in the next section that if we consider "canonical" SDE, the infinitesimal generator (in the canonical sense, to be stated in the next section ) of the inverse flow coincides with the infinitesimal generator of the original flow.

Let us come back to the proof of the theorem. Set $x = \xi_{s,t}^{-1}(y)$. Then equation (3.28) implies

$$y = \xi_{s,t}^{-1}(y) + \int_s^t X(\xi_{s,r}(x), dr)\Big|_{x=\xi_{s,t}^{-1}(y)}$$

Therefore the equation of the theorem is verified if we can show that

$$\int_s^t X(\xi_{s,r}(x), dr)\Big|_{x=\xi_{s,t}^{-1}(y)} = \int_s^t \hat{X}(\xi_{r,t}^{-1}(y), \hat{d}r), \quad a.s., \tag{3.37}$$

for any $s < t$. In the rest of the section, we give a proof of the above formula.

Let $\Pi_n = \{s = t_0 < t_1 < \cdots < t_{k_n} = t\}, n = 1, 2, ..$ be a sequence of partitions such that $|\Pi_n| \to 0$. For each $\Pi = \Pi_n$ we define a simple forward predictable process by

$$\xi^{\Pi}_{s,t}(x) = \sum_k \xi_{s,t_k}(x) 1_{(t_k,t_{k+1}]}(t).$$

By applying Theorem 2.14, we can show that the sequence $\{\int_s^t X(\xi^{\Pi_n}_{s,r}(x), dr)\}$ converges to $\int_s^t X(\xi_{s,r}(x), dr)$ in probability uniformly on compact sets with respect to $x$. Then we have

$$\int_s^t X(\xi_{s,r}(x), dr)\Big|_{x=\xi^{-1}_{s,t}(y)} = \lim_{n\to\infty} \int_s^t X(\xi^{\Pi_n}_{s,r}(x), dr)\Big|_{x=\xi^{-1}_{s,t}(y)}.$$

By the cocycle property, we have

$$\begin{aligned}
&\int_s^t X(\xi^{\Pi_n}_{s,r}(x), dr)\Big|_{x=\xi^{-1}_{s,t}(y)} \\
&= \sum_{k=1}^{k_n} \{X(\xi_{s,t_k}(x), t_{k+1}) - X(\xi_{s,t_k}(x), t_k)\}\Big|_{x=\xi^{-1}_{s,t}(y)} \\
&= \sum_{k=1}^{k_n} \{(X(\xi^{-1}_{t_k,t}(y), t_{k+1}) - X(\xi^{-1}_{t_k,t}(y), t_k)\} \\
&= \sum_{k=1}^{k_n} \{(X(\xi^{-1}_{t_{k+1},t}(y), t_{k+1}) - X(\xi^{-1}_{t_{k+1},t}(y), t_k)\} \\
&\quad - \Bigg[ \sum_{k=1}^{k_n} \{X(\xi^{-1}_{t_{k+1},t}(y), t_{k+1}) - X(\xi^{-1}_{t_{k+1},t}(y), t_k)\} \\
&\qquad - \sum_{k=1}^{k_n} \{(X(\xi^{-1}_{t_k,t}(y), t_{k+1}) - X(\xi^{-1}_{t_k,t}(y), t_k)\} \Bigg] \\
&= I_1^{\Pi_n}(X) - I_2^{\Pi_n}(X).
\end{aligned}$$

The convergence

$$\exists \lim_{n\to\infty} I_1^{\Pi_n}(X) = \int_s^t X(\xi^{-1}_{r,t}(y), \hat{d}r)$$

is obvious from the definition of the backward Itô integral. We want to prove

**Lemma 3.14.**

$$\begin{aligned}
\lim_{n\to\infty} I_2^{\Pi_n}(X) &= 2\int_s^t c(\xi^{-1}_{r,t}(y), r) dr \\
&+ \int_s^t \int_{\mathcal{Z}} \{g(\xi^{-1}_{r,t}(y), r, z) - h(\xi^{-1}_{r,t}(y), r, z)\} N(\hat{d}r dz).
\end{aligned}$$

*Proof.* Set

$$B = \int_0^t b(x,r)dr, \quad M_c = \int_0^t f(x,r)dW(r), \quad M_d = \int_0^t \int_{\mathcal{Z}} g(x,r,z)\tilde{N}(drdz).$$

We shall consider the limits of $\{I_2^{\Pi_n}(B)\}$, $\{I_2^{\Pi_n}(M_c)\}$ and $\{I_2^{\Pi_n}(M_d)\}$ as $n \to \infty$. We have

$$\begin{aligned} |I_2^{\Pi_n}(B)| &\to |\int_s^t b(\xi_{r-,t}^{-1}(y), r)dr - \int_s^t b(\xi_{r,t}^{-1}(y), r)dr| \\ &= 0. \end{aligned}$$

As to the second term $I^{\Pi_n}(M_c)$, we have

$$\begin{aligned} I_2^{\Pi_n}(M_c) &= \sum_j \sum_k (f^j(\xi_{t_{k+1},t}^{-1}(y), t_{k+1}) - f^j(\xi_{t_k,t}^{-1}(y), t_k)) \\ &\qquad (W^j(t_{k+1}) - W^j(t_k)) \\ &= \sum_j \sum_k (f^j(\xi_{s,t_{k+1}}(x), t_{k+1}) - f^j(\xi_{s,t_k}(x), t_k)) \\ &\qquad (W^j(t_{k+1}) - W^j(t_k))\Big|_{x=\xi_{s,t}^{-1}(y)}. \end{aligned}$$

where $f^j(x,t) = (f^{1j}(x,t), ..., f^{dj}(x,t))$. Since $f^j$ are $C^{2,1}$ functions, $f^j(\xi_{s,t}, t)$ is a (forward) semimartingale by Itô's formula. Therefore we have

$$\begin{aligned} &\sum_k (f^j(\xi_{s,t_{k+1}}(x), t_{k+1}) - f^j(\xi_{s,t_k}(x), t_k))(W^j(t_{k+1}) - W^j(t_k)) \\ &\qquad \to [f^j(\xi_{s,t}(x), t), W^j(t)]. \end{aligned}$$

By Itô's formula II, $f^j(\xi_{s,t}(x), t)$ is written as

$$\begin{aligned} f^j(\xi_{s,t}(x), t) = f^j(x,s) &+ \sum_{i,k} \int_s^t \frac{\partial f^j}{\partial x_i}(\xi_{s,r}(x), r) f^{ik}(\xi_{s,r}(x), r)dW^k(r) \\ &+ W_d(t) + A(t), \end{aligned}$$

where $W_d(t)$ is a purely discontinuous localmartingale and $A(t)$ is a process of finite variation. We know $[W_d, W^j] = 0$ since $W^j$ and $W_d$ are orthogonal and $[A, W^j] = 0$ since $A(t)$ is a process of finite variation. Then we get

$$\begin{aligned} [f^j(\xi_{s,t}(x), t), W^j(t)] &= \sum_{i,k} [\int_s^t \Big(\frac{\partial f^j}{\partial x_i} f^{ik}\Big)(\xi_{s,r}(x), r)dW^k(r), W^j(t)] \\ &= \sum_i \int_s^t \Big(\frac{\partial f^j}{\partial x_i} f^{ij}\Big)(\xi_{s,r}(x), r)dr. \end{aligned}$$

Here we used $[W^k, W^j]_t = \langle W^k, W^i \rangle_t = \delta_{ik} t$. Therefore we have

$$\lim_{n\to\infty} I_2^{\Pi_n}(M_c) = \int_s^t \Big(\sum_{i,j} \frac{\partial f^j}{\partial x_i} f^{ij}(\xi_{s,r}(x), r)\Big) dr \Big|_{x=\xi_{s,t}^{-1}(y)}$$
$$= 2\int_s^t c(\xi_{r,t}^{-1}(y), r) dr.$$

We next consider $I_2^{\Pi_n}(M_d)$. We split $M_d$ into the sum of $M_\epsilon$ and $N_\epsilon$, where

$$M_\epsilon(t) = \int_0^t \int_{V_\epsilon} g(x, r, z)\tilde{N}(drdz), \qquad N_\epsilon = M_d - M_\epsilon.$$

where $V_\epsilon = \{z : K(z) < \epsilon\}$. Then $I_2^{\Pi_n}(M_\epsilon)$ is uniformly small if $\epsilon$ is sufficiently small. In fact we have for any $\delta > 0$,

$$\lim_{\epsilon\to 0} \sup_n P(|I_2^{\Pi_n}(M_\epsilon)| > \delta) = 0.$$

On the other hand, $N_\epsilon$ is a process of finite variation and the number of jumps are finite a.s. Then we have

$$\exists \lim_{n\to\infty} I_2^{\Pi_n}(N_\epsilon) = \sum_{r\in(s,t]} (\Delta X(\xi_{r,t}^{-1}(y), r) - \Delta X(\xi_{r-,t}^{-1}(y), r)) 1_{\{p(r)\in V_\epsilon^c\}} \tag{3.38}$$
$$= \sum_{r\in(s,t]} (\Delta X(\xi_{r,t}^{-1}(y), r) - \Delta X \circ \xi_{r-,r}^{-1}(\xi_{r,t}^{-1}(y), r)) 1_{\{p(r)\in V_\epsilon\}}$$
$$= \int_s^t \int_{V_\epsilon^c} (g(\xi_{r,t}^{-1}(y), r, z) - g \circ \psi_{r,z}(\xi_{r,t}^{-1}(y), r, z)) N(\hat{d}r, dz).$$

Note that $g(r, z) - g(r.z) \circ \psi_{r,z} = g(r, z) - h(r, z)$. Then we obtain

$$\lim_{n\to\infty} I_2^{\Pi_n}(N_\epsilon) = \int_s^t \int_{V_\epsilon^c} \{g(\xi_{r,t}^{-1}(y), r, z) - h(\xi_{r,t}^{-1}(x), r, z)\} N(\hat{d}rdz).$$

It converges to

$$\int_s^t \int_{\mathcal{Z}} \{g(\xi_{r,t}^{-1}(y), r, z) - h(\xi_{r,t}^{-1}(y), r, z)\} N(\hat{d}rdz) < \infty,$$

as $\epsilon \to 0$. Then we obtain that $\lim_{n\to\infty} I_2^{\Pi_n}(M_d)$ exists and is equal to the above. Therefore we get the assertion of the lemma.

Now equality (3.37) holds valid by Lemma 3.14. Then we have the formula (3.32). The proof of Theorem 3.13 is completed.

## 3.7 Canonical SDE

The canonical SDE was introduced by S.I. Marcus [18]. It may be written as

$$d\xi_t = \sum_{j=1}^{m} v_j(\xi_r) \diamond dZ^j(r), \tag{3.39}$$

where $v_1, ..., v_m$ are vector fields on $\mathbf{R}^d$ and $Z(t)$ is an $m$-dimensional Lévy process. The precise definition is as follows.

$$\begin{aligned} d\xi_t &= \sum_{j=1}^{m} v_j(\xi_t) \circ dZ_c^j(t) + \sum_{j=1}^{m} v_j(\xi_{t-}) dZ_d^j(t) \\ &\quad + \Big\{ \mathrm{Exp}(\sum_j \Delta Z^j(t) v_j)(\xi_{t-}) - \xi_{t-} - \sum_j \Delta Z^j(t) v_j(\xi_{t-}) \Big\}, \end{aligned} \tag{3.40}$$

where $Z_c(t)$ and $Z_d(t)$ are continuous and purely discontinuous parts of the Lévy processes, respectively and $\mathrm{Exp}(tv) = \varphi(t, x)$ is the solution flow of the ordinary differential equation generated by the vector field $v$

$$\frac{d\varphi(t)}{dt} = v(\varphi(t)), \quad \varphi(0) = x.$$

The above equation looks complicated. But the probabilistic meaning is clear. At the jump time of the driving process $Z(t)$, the solution flow flies from the state $\xi_{t-}(x)$ along with the integral curve $\mathrm{Exp}(tv), 0 \le t \le 1$ with infinite speed, where $v = \sum_j \Delta Z^j(t) v_j$ and lands at the position at $t = 1$, i.e.,

$$\xi_t(x) = \mathrm{Exp}(\sum_j \Delta Z_j(t) v_j)(\xi_{t-}(x)).$$

Therefore if the map $\xi_{t-} : \mathbf{R}^d \to \mathbf{R}^d$ is a homeomorphism, then the map $\xi_t$ should also be a homeomorphism, since the map $\mathrm{Exp}(v)$ is a homeomorphism.

If there is no jump part in the Lévy process $Z(t)$, i.e., if $Z(t)$ is a Brownian motion, then the canonical SDE coincides with the Stratonovich SDE. It is known that the Stratonovich SDE is coordinate free and we can define the Stratonovich SDE on any manifold. We may define the canonical SDE on any manifold $M$ provided that we are given vector fields $v_1, ..., v_m$ on $M$ and an $m$-dimensional Lévy process $Z(t)$. However we will not discuss the problem here. We refer to Fujiwara [6] and Applebaum–Kunita [1].

In this section we will study the property of the canonical SDE on Euclidean space, by applying results obtained in previous sections. For this purpose we shall transform the canonical equation to an Itô equation. Let us recall that the Lévy process $Z(t)$ admits the Lévy–Itô decomposition of Section 2.4 (Theorem 2.7). For simplicity, we assume that $W(t)$ in Lévy–Itô decomposition (2.8) is a standard Brownian motion. We assume that coefficients $v_1, ..., v_m$ are $C^2$ functions with bounded derivatives. Thus they are Lipschitz continuous. The term involving the Stratonovich integral is written as

$$
\begin{aligned}
\sum_j \int_{t_0}^t v_j(\xi_{r-}) \circ dZ_c^j(r) &= \sum_j \int_{t_0}^t v_j(\xi_{r-}) dW^j(r) + \sum_j \int_0^t v_j(\xi_{r-}) b^j dr \\
&\quad + \frac{1}{2} \sum_{l,k} \langle \int \frac{\partial v_j}{\partial x_k}(\xi_{r-}) v_l^k(\xi_{r-}) dW^l, W^j \rangle_t \\
&= \sum_j \int_{t_0}^t v_j(\xi_{r-}) dW^j(r) + \sum_j \int_0^t v_j(\xi_{r-}) b^j dr \\
&\quad + \frac{1}{2} \sum_{j,k} \int_{t_0}^t \frac{\partial v_j}{\partial x_k}(\xi_{r-}) v_j^k(\xi_{r-}) dr.
\end{aligned}
$$

The jump part is written as

$$
\begin{aligned}
&\sum_{j=1}^m \int_{t_0}^t v_j(\xi_{r-}) dZ_d^j(r) + \sum_{t_0 < r \le t} \Big\{ \mathrm{Exp}(\sum_j \Delta Z^j(r) v_j)(\xi_{r-}) - \xi_{r-} \\
&\qquad - \sum_j \Delta Z^j(r) v_j(\xi_{r-})) \Big\} \\
&= \int_{t_0}^t \int_{|z| \le 1} \Big\{ \mathrm{Exp}(\sum_j z^j v_j)(\xi_{r-}) - \xi_{r-} \Big\} \tilde{N}(drdz) \\
&\quad + \int_{t_0}^t \int_{|z| \le 1} \Big\{ \mathrm{Exp}(\sum_j z^j v_j)(\xi_{r-}) - \xi_{r-} - \sum_j z^j v_j(\xi_{r-}) \Big\} \hat{N}(drdz) \\
&\quad + \int_{t_0}^t \int_{|z| > 1} \Big\{ \mathrm{Exp}(\sum_j z^j v_j)(\xi_{r-}) - \xi_{r-} \Big\} N(drdz).
\end{aligned}
$$

Therefore the canonical SDE is written as the Itô SDE:

$$
\xi_t = x + \int_{t_0}^t X(\xi_{r-}, dr), \tag{3.41}
$$

where

$$
X(x,t) = X'(x,t) + \int_{t_0}^t \int_{|x|>1} \Big\{ \mathrm{Exp}(\sum_j z^j v_j)(x) - x \Big\} N(dsdz), \tag{3.42}
$$

$$
\begin{aligned}
X'(x,t) &= \sum_j v_j(x) W^j(t) + \int_{t_0}^t \int_{|x| \le 1} \Big\{ \mathrm{Exp}(\sum_j z^j v_j)(x) - x \Big\} \tilde{N}(dsdz) \\
&\quad + t \Big[ \sum_j b^j v_j(x) + \frac{1}{2} \sum_{j,k} \frac{\partial v_j}{\partial x_k}(x) v_j^k(x) \\
&\qquad + \int_{|x| \le 1} \Big\{ \mathrm{Exp}(\sum_j z^j v_j)(x) - x - \sum_j z^j v_j(x) \Big\} \nu(dz) \Big].
\end{aligned}
$$

We first consider the case where $\nu$ is supported by a bounded set, say $\{|z| \le c\}$. Then $X = X'$ holds. We show that the coefficients of the above equation satisfy Lipschitz conditions. The following is easily verified.

**Lemma 3.15.** *(1)* $|\mathrm{Exp}\, v(x) - x| \le \|v\|_{lg} e^{\|v\|_{lg}}(1+|x|)$.
*(2)* $|\mathrm{Exp}\, v(x) - x - (\mathrm{Exp}\, v(y) - y)| \le \|v\|_{Lip} e^{\|v\|_{Lip}}$.
*(3)* $|\mathrm{Exp}\, v(x) - x - v(x) - (\mathrm{Exp}\, v(y) - y - v(y)| \le \|\nabla v \cdot v\|_{Lip} e^{\|v\|_{Lip}}|x-y|$,
*where*

$$\|v\|_{lg} = \sup_x \frac{|v(x)|}{1+|x|}, \quad \|v\|_{Lip} = \sup_{x \ne y} \frac{|v(x)-v(y)|}{|x-y|}$$

*and*

$$(\nabla v \cdot w)^i = \sum_{j=1}^d \frac{\partial v^i}{\partial x^j}(x) w^j(x).$$

The diffusion coefficient $f(x) = (v_j^i(x))$ is clearly Lipschitz continuous, since $v_j, j = 1, 2, \ldots$ are Lipschitz continuous. We consider the jump coefficient. From the above lemma, we have

$$\left|\mathrm{Exp}(\sum_j z^j v_j)(x) - x - (\mathrm{Exp}(\sum_j z^j v_j)(y) - y)\right| \le L(z)|x-y|,$$

where $L(z) = \sum_j |z^j| \|v_j\|_{Lip} e^{\sum_j |z^j| \|v_j\|_{Lip}}$. Therefore if the Lévy measure $\nu$ is supported by a bounded set, then $\int_{\mathbf{R}^m - \{0\}} L(z)^p \nu(dz) < \infty$. We have similarly $\int K(z)^p \nu(dz) < \infty$. We can show also that the drift coefficient

$$b = \sum_j b^j v_j + \frac{1}{2}\sum_{jk} \frac{\partial v_j}{\partial x_k} v_j^k + \int_{\mathbf{R}^m - \{0\}} \left\{\mathrm{Exp}(\sum_j z^j v_j) - I - \sum_j z^j v_j\right\} \nu(dz) \quad (3.43)$$

is also Lipschitz continuous. Then the solution $\xi_t(x)$ of canonical SDE defines a flow of continuous maps.

Further, it holds $g(x,z) + x = \phi_z(x) = \mathrm{Exp}(\sum_j z^j v_j)(x)$ and $\phi_z$ are homeomorphisms. Its inverse map is given by $\psi_z(x) = \mathrm{Exp}(-\sum_j z^j v_j)(x)$. It is uniformly Lipschitz continuous and uniformly of linear growth, since $z$ is restricted to $\{z : |z| \le c\}$. Consequently the solution $\xi_{s,t}$ of the canonical SDE defines a stochastic flow of homeomorphisms.

We next consider the case where the support of the Lévy measure is unbounded. Then the Lipschitz constants and linear growth constants $K(z)$ are unbounded. In this case we can construct the solution by the composition formula (3.26). It defines a stochastic flow of homeomorphisms.

Now assume that the first derivatives of $v_j$, $j = 1, \ldots, m$ and $\sum_{jk} \frac{\partial v_j}{\partial x_k} v_j^k$ are bounded and $\delta$-Hölder continuous. It is known that the Jacobian matrix of the map $\phi_z$ is invertible. Then the solution defines a stochastic flow of $C^1$-diffeomorphisms by Theorem 3.11.

Summing up the above discussions, we have

**Theorem 3.16.** *Assume that coefficients $v_j, j = 1, ..., m$ of the canonical SDE are of $C^2$-class and that the first derivatives of $v_1, ..., v_m$ and that of $\sum_{jk} \frac{\partial v_j}{\partial x_k} v_j^k$ are all bounded. Then the solution of the canonical SDE defines a stochastic flow of homeomorphisms.*

*Assume further that the first derivatives of the above are $\delta$-Hölder continuous, then the solution defines a stochastic flow of $C^1$-diffeomorphisms.*

**Remark.** Consider Ito's SDE of separating type, instead of the canonical SDE. We have $g(x, r, z) = \sum_j v_j(x)z^j$, so that $\phi_{r,z}(x) = x + \sum_j v_j(x)z^j$. It is homeomorphic if $|z|$ is sufficiently small. Therefore the solution defines a stochastic flow of homeomorphisms if the size of jumps of the Lévy process is sufficiently small. See Protter [20].

We shall consider the backward SDE for the inverse flow. We need a lemma.

**Lemma 3.17.** *The function $h(x, z) = -\text{Exp}(-\sum_j z^j v_j)(x) + x$ satisfies*

$$\int_{|z|\le c} |g - h|\nu(dz) < \infty, \quad \int_{|z\le c} |h(z)|^2 \nu(dz) < \infty$$

*for any $c > 0$.*

*Proof.* We have by Taylor's theorem,

$$\begin{aligned} g(x, z) - \sum_j z^j v_j(x) &= \text{Exp}(\sum_j z^j v_j)(x) - x - \sum_j z^j v_j(x) \\ &= \frac{1}{2} \sum_{j,k} z^j z^k \nabla v_j \cdot v_k \text{Exp}(\theta \sum_j z^j v_j)(x) \end{aligned}$$

and

$$\begin{aligned} -h(x, z) + \sum_j z^j v_j(x) &= \text{Exp}(-\sum_j z^j v_j)(x) - x + \sum_j z^j v_j(x) \\ &= \frac{1}{2} \sum_{j,k} z^j z^k (\nabla v_j \cdot v_k) \text{Exp}(-\theta' \sum_j z^j v_j)(x), \end{aligned}$$

where $0 < \theta, \theta' \le 1$. Therefore we have

$$|g(x, z) - h(x, z)| \le \frac{1}{2} \sum_{j,k} |z^j z^k| |R_{j,k}(x, z, \theta, \theta')|,$$

where $R_{j,k}(x, z, \theta, \theta')$ are bounded with respect to $z$. Then $g - h$ is integrable with respect to the Lévy measure on the set $\{|z| \le c\}$. The square integrability of $h$ will be obvious.

Now the inverse flow $\xi_{s,t}^{-1}$ satisfies the backward SDE based on $\hat{X}(x, t)$ of (3.36) in view of Remark 2) after Theorem 3.13. It is written as

$$\hat{X}(x,t) = \hat{b}(x)t + \sum_j v_j(x)W^j(t) + \int_{t_0}^t \int_{\mathbf{R}^m-\{0\}} \Big\{ -\operatorname{Exp}(-\sum_j z^j v_j)(x) + x \Big\} \tilde{N}(drdz)$$

where

$$\hat{b}(x) = b(x) - 2c(x) - \int_{\mathbf{R}^m-\{0\}} (g(x,z) - h(x,z))\nu(dz)$$

and $b(x)$ is given by (3.43). Then the drift term is written by

$$\hat{b}(x) = \sum_j b^j v_j(x) - \frac{1}{2}\sum_{j,k} \frac{\partial v_j}{\partial x_k}(x) v_j^k(x) - \int_{\mathbf{R}^m-\{0\}} \Big\{ \operatorname{Exp}(-\sum_j z^j v_j)(x) - x + \sum_j z_j v_j(x) \Big\} \nu(dz).$$

Also we have the equality

$$\begin{aligned}
&-\int_s^t \hat{X}(\xi_{r,t}^{-1}(y), \hat{d}r) \\
&= -\sum_j \int_s^t v_j(\xi_{r,t}^{-1}(y)) \circ \hat{d}Z_c(t) - \sum_j \int_s^t v_j(\xi_{r,t}^{-1}(y))\hat{d}Z_d(t) \\
&\quad - \sum_r \Big\{ -\operatorname{Exp}(-\sum_j \Delta Z^j(r) v_j)(\xi_{r,t}^{-1}(y)) + \xi_{r,t}^{-1}(y) \\
&\quad - \sum_j \Delta Z^j(r) v_j(\xi_{r,t}^{-1}(y)) \Big\} \\
&= \sum_j \int_s^t v_j(\xi_{r,t}^{-1}(y)) \circ \hat{d}(-Z_c)(t) + \sum_j \int_s^t v_j(\xi_{r,t}^{-1}(y))\hat{d}(-Z_d)(t) \\
&\quad + \sum_r \Big\{ \operatorname{Exp}(\sum_j \Delta(-Z^j)(r) v_j)(\xi_{r,t}^{-1}(y)) - \xi_{r,t}^{-1}(y) \\
&\quad - \sum_j \Delta(-Z^j)(r) v_j(\xi_{r,t}^{-1}(y)) \Big\} \\
&= \sum_j \int_s^t v_j(\xi_{r,t}^{-1}(y)) \diamond \hat{d}(-Z^j)(r) \\
&= -\sum_j \int_s^t v_j(\xi_{r,t}^{-1}(y)) \diamond \hat{d}Z^j(r).
\end{aligned}$$

Consequently we have the following.

**Theorem 3.18.** *Assume the same conditions as in Theorem 3.16. Then the inverse flow* $\xi_{s,t}^{-1}(y)$ *satisfies*

$$\xi_{s,t}^{-1}(y) = x - \sum_j \int_s^t v_j(\xi_{r,t}^{-1}(y)) \diamond \hat{d}Z^j(r). \tag{3.44}$$

**Further generalizations.** Let $X(x,t)$, $x \in \mathbf{R}^d$, $t \in [0,T]$, be stochastic processes with values in $\mathbf{R}^d$ with parameter $x$. It is called a semimartingale with spatial parameter, if for each fixed $x$, it is an $\mathbf{R}^d$-valued semimartingale. The $\mathbf{R}^d$-valued processes $X(x,t)$ introduced in (3.3) etc. are examples of such semimartingales with spatial parameter.

Let $\eta_t$ be an adapted cadlag process with values in $\mathbf{R}^d$. Under some regularity conditions on $X(x,t)$, we can define Itô's stochastic integral $\int_0^t X(\eta_{r-}, dr)$ as a semimartingale. Then SDEs based on $X(x,t)$ can be defined. We may also define canonical stochastic integral $\int_0^t X(\eta_{r-}, \diamond dr)$ and canonical SDE. The regularity of the solution with respect to the initial data, the homeomorphic property and the diffeomorphic property have all been studied. We refer to Fujiwara–Kunita [7,8], Carmona–Nualart [3] and Applebaum–Tang [2].

# 4 Appendix. Kolmogorov's criterion for the continuity of random fields and the uniform convergence of random fields

We shall introduce Kolmogorov's criterion for a given random field to have a modification of a continuous random field.

**Theorem 4.1 (Kolmogorov–Totoki).** *Let* $X(x)$, $x \in \mathbf{D}$ *be a random field with values in a normed space* $B$ *where* $\mathbf{D}$ *is a domain in* $\mathbf{R}^d$. *Assume that there exist positive constants* $\gamma$, $C$ *and* $\alpha > d$ *satisfying*

$$E[\|X(x) - X(y)\|^\gamma] \le C|x-y|^\alpha, \quad \forall x, y \in \mathbf{D}. \tag{4.1}$$

*Then* $X(x)$ *has a continuous modification. Further, the continuous modification is uniformly continuous in* $\mathbf{D}$ *and it can be extended continuously to the closure* $\bar{\mathbf{D}}$ *of the domain* $\mathbf{D}$, *i.e., there exists a continuous random field* $\tilde{X}(x)$, $x \in \bar{\mathbf{D}}$ *such that* $\tilde{X}(x) = X(x)$ *holds a.s. for any* $x \in \mathbf{D}$.

In order to prove the theorem, we will introduce a modulus of continuity of a map from $\mathbf{D}$ to $B$. Let $\Pi_n$ be the set of all lattice points in $\mathbf{R}^d$ of the form $(i_1/2^n, ..., i_d/2^n)$, where $i_1, ..., i_d$ are integers. We set $\Pi = \cup_n \Pi_n$. It is a dense subset of $\mathbf{R}^d$. We let

$$\mathbf{D}_n = \{x \in \mathbf{D};\ \exists x_n \in \Pi_n \cap \mathbf{D} \text{ such that } 0 \le x^i - x_n^i < 2^{-n},\ i = 1, ..., d\}.$$

Then $\mathbf{D}_n$ is increasing with $n$ and $\cup_n \mathbf{D}_n = \mathbf{D}$ holds.

Given a map $f : \Pi \cap \mathbf{D} \to B$, we define for each $n$ a modulus of continuity and the modulus of $\beta$-Hölder continuity of $f$ by

$$\Delta_n(f) = \max_{x,y \in \Pi_n \cap \mathbf{D}, |x-y| = 2^{-n}} \|f(x) - f(y)\|,$$
$$\Delta_n^\beta(f) = 2^{-\beta n} \Delta_n(f).$$

**Lemma 4.2.** *The inequality*

$$\|f(x) - f(y)\| \le 2^{d+1} \left( \sum_{n=1}^{\infty} \Delta_n^{\beta}(f) \right) |x - y|^{\beta}, \quad \forall x, y \in \Pi \cap \mathbf{D} \tag{4.2}$$

*holds for any map* $f : \Pi \cap \mathbf{D} \to B$.

*Proof.* Given a function $f$ on $\Pi \cap \mathbf{D}$, we define a sequence of simple functions $g_n : \mathbf{D}_n \to B, n = 1, 2, ....$ by $g_n(x) = f(x_n)$, where $x_n \in \Pi_n \cap \mathbf{D}$ is the point such that $0 \le x^i - x_n^i < 2^{-n}$ holds for any $i = 1, ..., d$. Then it holds

$$|g_{n+1}(x) - g_n(x)| \le \Delta_{n+1}(f), \quad \forall x \in \mathbf{D}_n.$$

For any $x \in \mathbf{D}$, let $k$ be a positive integer such that $x \in \mathbf{D}_k$. Then

$$\sum_{n=k}^{\infty} \|g_{n+1}(x) - g_n(x)\| \le \sum_{n=1}^{\infty} \Delta_{n+1}(f) \le \sum_{n=1}^{\infty} \Delta_{n+1}^{\beta}(f) < \infty.$$

Then the sequence of simple functions $g_n(x)$ converges on $\mathbf{D}$. Let $g(x)$ be the limit function. Then $g(x) = f(x)$ holds valid for $x \in \Pi \cap \mathbf{D}$. In the sequel we prove the lemma for $g$ instead of $f$.

Now let $x, y$ be any points in $\mathbf{D}$. Then $x, y \in \mathbf{D}_n$ for a sufficiently large $n$. Take $k$ such that $2^{-(k+1)} \le |x - y| < 2^{-k}$. Then for $n \ge k$,

$$\begin{aligned} \|g(x) - g_n(x)\| &\le \sum_{m=n+1}^{\infty} \Delta_m(f) \\ &\le 2^{-(n+1)} \sum_{m=n+1}^{\infty} \Delta_m^{\beta}(f) \\ &\le \left( \sum_{m=n+1}^{\infty} \Delta_m^{\beta}(f) \right) |x - y|^{\beta}. \end{aligned}$$

Further since $2^{-(n+1)} \le |x - y|$, we have

$$\|g_{n+1}(x) - g_{n+1}(y)\| \le 2^d \Delta_{n+1}(f) \le 2^d \Delta_{n+1}^{\beta}(f) |x - y|^{\beta}.$$

Therefore

$$\begin{aligned} \|g(x) - g(y)\| &\le \|g(x) - g_{n+1}(x)\| + \|g_{n+1}(x) - g_{n+1}(y)\| \\ &\quad + \|g_{n+1}(y) - g(y)\| \\ &\le 2^{d+1} \left( \sum_{k=1}^{\infty} \Delta_k^{\beta}(f) \right) |x - y|^{\beta}. \end{aligned}$$

The proof is complete.

By the above lemma, the map $f : \Pi \cap \mathbf{D} \to B$ satisfying $\sum_k \Delta_k^\beta(f) < \infty$ for some $\beta > 0$ is uniformly continuous on $\mathbf{D}$ and has a continuous extension $g : \bar{\mathbf{D}} \to B$ ,i.e., there exists a continuous map $g : \bar{\mathbf{D}} \to B$ such that $g(x) = f(x)$ holds for $x \in \Pi \cap \mathbf{D}$. The function $g$ is $\beta$-Hölder continuous.

We shall apply the above lemma to the random field $X(x)$. Observe that for each $\omega$, $X(\cdot, \omega)$ restricting $x$ to $\Pi \cap \mathbf{D}$ can be regarded as a map form $\Pi \cap \mathbf{D}$ to $B$. Then we have

$$\|X(x,\omega) - X(y,\omega)\| \le 2^{d+1}\left(\sum_k \Delta_k^\beta(X(\omega))\right)|x-y|^\beta, \quad \forall x, y \in \Pi \cap \mathbf{D}. \tag{4.3}$$

**Lemma 4.3.** *Let $X(x), x \in \mathbf{D}$ be a random field satisfying the inequality (4.1). Let $\beta$ be a positive number satisfying $\beta\gamma < \alpha - d$. Then,*

$$E[(\sum_k \Delta_k^\beta(X))^\gamma]^{1/\gamma} \le \left(\sum_{k=1}^\infty 2^{-k\{(\alpha-d)/\gamma-\beta\}}\right)\cdot(2^dC)^{1/\gamma} < \infty. \tag{4.4}$$

*Proof.* We will consider the case $\gamma \ge 1$ only. Observe the inequality

$$\begin{aligned}(\Delta_k^\beta(X))^\gamma &\le (\sup_{x,y\in\Pi\cap\mathbf{D}, |x-y|=1/2^k} \|X(x)-X(y)\|2^{k\beta})^\gamma \\ &\le \sum(\|X(x')-X(y')\|2^{k\beta})^\gamma,\end{aligned}$$

where the summations are taken over all $x', y' \in \Pi \cap \mathbf{D}$ such that $|x'-y'| = 1/2^k$. Then the number of summations are at most $2^{(k+1)d}$. Therefore

$$\begin{aligned}E[\Delta_k^\beta(X)^\gamma] &\le 2^{(k+1)d+k\gamma\beta}E[\|X(x')-X(y')\|^\gamma] \\ &\le 2^{k(d+\gamma\beta-\alpha)}2^dC.\end{aligned}$$

In the last inequality, we applied the inequality (4.1). Therefore we get

$$\begin{aligned}E[(\sum_k \Delta_k^\beta(X))^\gamma]^{1/\gamma} &\le \sum_{k=1}^\infty E[\Delta_k^\beta(X)^\gamma]^{1/\gamma} \\ &\le \left(\sum_{k=1}^\infty 2^{-k\{(\alpha-d)/\gamma-\beta\}}\right)\cdot(2^dC)^{1/\gamma} < \infty. \end{aligned} \tag{4.5}$$

The proof is complete.

*Proof of Theorem 4.1.* The random field $X(x)$ restricting $x$ on $\Pi \cap \mathbf{D}$ satisfies the inequality (4.3), where $\sum_k \Delta_k^\beta(X) < \infty$ holds a.s. Therefore $X(x), x \in \Pi \cap \mathbf{D}$ is uniformly $\beta$-Hölder continuous a.s. Then there exists a continuous random field $\tilde{X}(x)$ defined on $\bar{\mathbf{D}}$ such that $X(x) = \tilde{X}(x)$ holds a.s. for any $x \in \Pi \cap \mathbf{D}$. Since $X(x)$ is continuous in probability, the equality $X(x) = \tilde{X}(x)$ holds a.s for any $x \in \mathbf{D}$. Thus we have proved the theorem.

We shall consider the uniform convergence of a sequence of continuous random fields. Let $\gamma$ and $\alpha$ be positive numbers. Let $\{X_n(x)\}$ be a sequence of $\gamma$ integrable random fields. We define a $(\gamma,\alpha)$-Hölder norm for a random field $X(x)$ by

$$\|X\|_{\gamma,\alpha} = \sup_{x\in\mathbf{D}} E[\|X(x)\|^\gamma]^{1/\gamma} + \sup_{x\neq y,x,y\in\mathbf{D}} \frac{E[\|X(x)-X(y)\|^\gamma]^{1/\gamma}}{|x-y|^{\alpha/\gamma}}.$$

**Theorem 4.4.** *Let $\{X_n(x)\}$ be a sequence of continuous random fields such that*

$$\lim_{n\to\infty} \|X_n - X\|_{\gamma,\alpha} = 0$$

*holds for some $\alpha > d$ and $\gamma > 0$. Then we have for any $N$,*

$$\lim_{n\to\infty} E[\sup_{x\in\mathbf{D},|x|\le N} \|X_n(x) - X(x)\|^\gamma] \to 0. \tag{4.6}$$

*Proof.* Take $\beta > 0$ such that $\gamma\beta < \alpha - d$. We want to prove

$$\lim_{n\to\infty} E[(\sum_k \Delta_k^\beta (X_n - X))^\gamma] = 0.$$

Set $C_n = \|X_n - X\|_{\gamma,\alpha}^\gamma$. Then it holds

$$E[\|X_n(x) - X(x)\|^\gamma] \le C_n,$$
$$E[\|X_n(x) - X(x) - (X_n(y) - X(y))\|^\gamma] \le C_n|x-y|^\alpha.$$

We have by Lemma 4.3,

$$E[(\sum_k \Delta^\beta(X_n - X))^\gamma]^{1/\gamma} \le \left(\sum_{k=1}^\infty 2^{-k\{(\alpha-d)/\gamma-\beta\}}\right)\cdot(2^d C_n)^{1/\gamma} = C_n',$$

which converges to 0 as $n\to\infty$. Then we get from (4.3),

$$E[\sup_{|x|\le N} \|X_n(x) - X(x) - X_n(x_0) + X(x_0)\|^\gamma]^{1/\gamma} \le C_n' \sup_{|x|\le N} |x - x_0| \to 0$$

for a fixed $x_0$. Now observe

$$\begin{aligned} &E[\sup_{|x|\le N} \|X_n(x) - X(x)\|^\gamma]^{1/\gamma} \\ &\le E[\|X_n(x_0) - X(x_0)\|^\gamma]^{1/\gamma} \\ &+E[\sup_{|x|\le N} \|X_n(x) - X(x) - X_n(x_0) + X(x_0)\|^\gamma]^{1/\gamma}. \end{aligned}$$

It converges to 0 again. Then we get the assertion of the theorem.

## References


[1] D. Applebaum and H. Kunita, Lévy flows and Lévy processes on Lie groups, *J. Math. Kyoto Univ.* **33** (1993), 1103–1123.

[2] D. Applebaum and F.Tang, The interlacing construction for stochastic flows of diffeomorphisms on Euclidean spaces, *Sankhya, Series A*, **63** (2001), 139–178.

[3] R.A. Carmona and D. Nualart, Nonlinear Stochastic Integrators. Equations and Flows, Stochastic Monographs 6, Gordon and Breach Science Publishers, 1990.

[4] C. Dellacherie and P.A. Meyer, *Probabilities and Potential B — Theory of Martingales*, North Holland, Amsterdam, 1982.

[5] K.D. Elworthy, *Stochastic Differential Equations on Manifolds*, LMS Lecture Note Series, 70, Cambridge University Press, Cambridge, UK, 1982.

[6] T. Fujiwara, Stochastic differential equations of jump type on manifolds and Lévy flows, *J. Math. Kyoto Univ.* **31**(1991), 99–119.

[7] T. Fujiwara and H. Kunita, Stochastic differential equations of jump type and Lévy flows in diffeomorphisms group, *J. Math. Kyoto Univ.* **25** (1985), 71–106.

[8] T. Fujiwara and H. Kunita, Canonical SDEs based on semimartingales with spatial parameters, Part I Stochastic flows of diffeomorphisms, *Kyushu J. Math.* **53** (1999), 265–300.

[9] T. Fujiwara and H. Kunita, Canonical SDEs based on semimartingales with spatial parameters, Part II Inverse flows and backward SDEs, *Kyushu J. Math.* **53** (1999), 301–331.

[10] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North Holland, Amsterdam. 1981.

[11] K. Itô, Spectral type of the shift transformation of differential processes with stationary increments, *Trans. Amer. Math. Soc.* **81**(1956), 253–263.

[12] J. Jacod and A.N. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer, 1987.

[13] H. Kunita, *Stochastic Flows and Stochastic Differential Equations*, Cambridge University Press, 1990.

[14] H. Kunita, Stochastic flows with jumps and stochastic flows of diffeomorphisms. In: "Itô's Stochastic Calculus and Probability Theory." N. Ikeda et al., Eds., Springer, 1996, 197–211.

[15] H. Kunita. Representation of martingales with jumps and applications to mathematical finance, Stochastic Analysis and Related Topics in Kyoto, Advanced Studies in Pure Mathematics, 41 (2004), 209–232.

[16] H. Kunita, S. Watanabe, On square integrable martingales, *Nagoya Math. J.* **30** (1967), 209–245.

[17] R. Léandre, Flot d'une équation differentielle stochastique avec semimartingale directrice discontinue. Séminaire Probab. XIX, Lecture Notes in Math. 1123(1985), 271–275.

[18] S.I. Marcus, Modelling and approximations of stochastic differential equations driven by semimartingales, *Stochastics* **4**(1981), 223–245.

[19] P.A. Meyer, Un cours sur integrales stochastiques, Seminaire Proba. X, Lecture Notes in Math. 511 Springer, 1976, 246–400.

[20] P. Protter, Stochastic Integration and Differential Equations. A New Approach, Applied Math. 21, Springer, 1992.

[21] K. Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge, UK, 1999.

[22] D.W. Stroock, Markov Processes from K. Itô's Perspective, Annals of Mathematical Studies 155, Princeton and Oxford. (2003)