Stochastic Analysis

Ito Integrals and Stochastic Differential Equations with Jumps

Seiji HIRABA

December 5, 2024

Contents

In this text, we describe Ito integrals (stochastic integrals) and Ito formula as important tools in the theory of stochastic processes, and we describe a part of analysis of Markov processes with jumps by using stochastic differential equations (SDE's). (We assume the readers are well-known about basics of probability theory.)

1 Definition of Stochastic Processes

1.1 Probability spaces and stochastic processes

A stochastic process is random variables with time parameter; $X_t = X_t(\omega)$ where $t \geq 0$ denotes the time and $\omega \in \Omega$ is a parameter of randomness, i.e., X_t is on a probability space $(\Omega, \mathcal{F}, P)$.

If we consider the time as times like 1st time, 2nd time, 3rd time, ... in a coin tossing, then $n = 1, 2, \dots$ denote the time and we denote the random variables as $X_n = X_n(\omega)$. The previous one is called as "continuous time", the later one is as "discrete time".

For ech $\omega \in \Omega$, $X.(\omega) = (X_t(\omega))_{t \geq 0})$ is called a **sample path**. If (X_t) has continuous sample paths, then it is called a **continuous process** or a C**-process**, or if paths are right-continuous and have left-hand-limits (rcll), then it is called a **discontinuous process** or D**-process**.

Let I be an interval of $\mathbf{R}_+ = [0, \infty)$ or a discrete set of $\mathbf{Z}_+$ or $\mathbf{N}$, Let S be a topology space (we mainly consider the d-dimensional Euclid space $\mathbf{R}^d$).

A stochastic process $\{X_t\}_{t \in I}$ on a probability space $(\Omega, \mathcal{F}, P)$ is a set of S-valued random variables $X_t = X_t(\omega)$ parameterized by time $t \in I$, i.e., $(X_t)_{t \in I}$. (We usually omit the variables $\omega \in \Omega$ if it is not necessary.)

If I is an interval, then it called as a **continuous time (stochastic) process**, or if I is a discrete set, then it is called as a **discrete time (stochastic) process**. However, in the later case, it is mainly denoted by $X_n, n = 0, 1, 2, \dots$. And S is called as a **state space**.

In this text we mainly consider the continuous time processes, so we let $I = [0, T]$ or $[0, \infty)$ and $S = \mathbf{R}^1$ or $\mathbf{R}^d$. (Off course, we note them if we change.)

A **filtration** $(\mathcal{F}_t)_{t \geq 0}$ is a family of sub σ-additive classes of $\mathcal{F}$ such that $\mathcal{F}_s \subset \mathcal{F}_t$ if $s < t$.

$\{X_t\}$ is $(\mathcal{F}_t)$**-adapted** $\overset{\text{def}}{\iff} {}^\forall t \geq 0, X_t \in \mathcal{F}_t$, i.e., X_t is $\mathcal{F}_t$-measurable.

$\{X_t\}$ is measurable if it is measurable function of (t, ω), that is, the following is measurable;

$$(t, \omega) \in ([0, \infty) \times \Omega, \mathcal{B}^1([0, \infty)) \otimes \mathcal{F}) \mapsto X_t(\omega) \in (\mathbf{R}^1, \mathcal{B}^1)$$

For a stochastic process $\{X_t\}$, if we let $\mathcal{F}_t^0 := \sigma(X_s; s \leq t) = \bigvee_{s \leq t} X_s^{-1}(\mathcal{B}^1) = \sigma(\bigcup_{s \leq t} X_s^{-1}(\mathcal{B}^1))$, then $\{X_t\}$ is $(\mathcal{F}_t^0)$-adapted.

In usual, to the above filtration we add $\mathcal{N} = \{N \in \mathcal{F}; P(N) = 0\}$, i.e., $\mathcal{F}_t = \mathcal{F}_t^0 \vee \mathcal{N} = \sigma(\mathcal{F}_t^0 \cup \mathcal{N})$. Under this filtration if ${}^\forall t \geq 0, X_t = Y_t$ a.s. then, $\{Y_t\}$ is also $(\mathcal{F}_t)$-adapted.

This filtration is called the **canonical filtration** by $\{X_t\}$.

For two processes $X = \{X_t\}, Y = \{Y_t\}$,

- X and Y are equivalent $\overset{\text{def}}{\iff} {}^\forall t, P(X_t = Y_t) = 1$.
- X and Y are strong equivalent $P({}^\forall t, X_t = Y_t) = 1$.
- X and Y are equivalent in law $\overset{\text{def}}{\iff} {}^\forall t_1, \cdots, t_n \in I, (X_{t_1}, \dots, X_{t_n}) \overset{(\text{d})}{=} (Y_{t_1}, \dots, Y_{t_n})$ (in the sense of distributions), that is, at any finite time points, the finite dimensional distributions are equal.

Clearly, [**strong equivalence** $\Rightarrow$ **equivalence** $\Rightarrow$ **equivalence in law**], however, the inverses are not true in general.

For example, if both have right-continuous sample paths, then the equivalence implies t strong equivalence. Because the probability of that they are equal at every rational time points is one, and their right-continuities implies at every irrational time points.

1.2 Exponential times and Poisson processes

For a constant $\alpha > 0$, a random variable $\tau = \tau(\omega)$ is distributed by an **exponential distribution with a parameter** α if

$$P(\tau > t) = \int_t^\infty \alpha e^{-\alpha s} ds = e^{-\alpha t}.$$

That is, τ has a distribution with a density function $f(s) = \alpha e^{-\alpha s}$.

In this text, we simply call τ as an α**-exponential time** or an **exponential time**.

The mean and the variance can be easily calculated and they are given as

$$E[\tau] = \int_0^\infty \alpha s e^{-\alpha s} ds = \frac{1}{\alpha}, \quad V(\tau) = E[\tau^2] - (E[\tau])^2 = \frac{1}{\alpha^2}.$$

Question 1.1 *Calculate the above variance.*

Proposition 1.1 *If τ is an exponential time, then it has the following* **memoryless property***: for $t, s \geq 0$,*

$$P(\tau > t + s |\ \tau > s) = P(\tau > t).$$

Proof.

$$P(\tau > t + s |\ \tau > s) = \frac{P(\tau > t + s)}{P(\tau > s)} = \frac{e^{-(t+s)}}{e^{-s}} = e^{-t} = P(\tau > t).$$

■

Proposition 1.2 *$\tau_1, \tau_2, \ldots \tau_n$ are independent and $\alpha_1, \alpha_2, \ldots, \alpha_n$-exponential times, then $\min\{\tau_1, \tau_2, \ldots \tau_n\}$ is an $\alpha_1 + \alpha_2 + \cdots + \alpha_n$-exponential time. Furthermore, it holds*

$$P(\min\{\tau_1, \tau_2, \ldots \tau_n\} = \tau_k) = \frac{\alpha_k}{\alpha_1 + \alpha_2 + \cdots + \alpha_n}.$$

Proof. For simplicity, we show in case of $n = 2, k = 1$.

$$P(\tau_1 \wedge \tau_2\} > t) = P(\tau_1 > t, \tau_2 > t) = P(\tau_1 > t)P(\tau_2 > t) = e^{-(\alpha_1+\alpha_2)t}.$$

The joint distribution of τ_1, τ_2 is a product of each distributions of them by independence. Hence,

$$\begin{aligned} P(\min\{\tau_1, \tau_2\} = \tau_1) &= P(\tau_1 < \tau_2) \\ &= \int_0^\infty ds \alpha_1 e^{-\alpha_1 s} P(s < \tau_2) \\ &= \int_0^\infty ds \alpha_1 e^{-\alpha_1 s} e^{-\alpha_2 s} \\ &= \frac{\alpha_1}{\alpha_1 + \alpha_2}. \end{aligned}$$

The general case is a similar. ■

For a given $\lambda > 0$, a stochastic process $(X_t)_{t \geq 0}$ is a **Poisson process with a parameter** λ (simply called a **λ-Poisson process**) if it satisfies the following:

(1) $X_0 = 0$.

(2) If $0 \leq s < t$, then $X_t - X_s$ is distributed by a Poisson distribution with a parameter $\lambda(t - s)$, i.e,

$$P(X_t - X_s = n) = e^{-\lambda(t-s)} \frac{\lambda^n (t-s)^n}{n!} \quad (n = 0, 1, 2, \ldots).$$

(3) X_t has independent increments, i.e., for $0 < t_1 < t_2 < \cdots < t_n$, $X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent.

Theorem 1.1 (Construction of a Poisson process) *Let $\sigma_1, \sigma_2, \ldots$ be identically independent distributed random variables each be a λ-exponential time. Set $\tau_n = \sum_{k=1}^n \sigma_k$, $\tau_0 = 0$ and let*

$$X_t = n \iff \tau_n \leq t < \tau_{n+1}, \ \textit{i.e.,}\ X_t := \sum_{n=0}^\infty n 1_{[\tau_n, \tau_{n+1})}(t) = \max\{n; \tau_n \leq t\}.$$

Then this is a λ-Poisson process.

Note The inverse of the above result holds, that is, if $(X_t)_{t\geq 0}$ is a λ-Poisson process and let $\tau_1, \tau_2, \dots$ be jump times of it, then $\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, \dots$ are i.i.d. and each of them is a λ-exponential time.

In order to show the above result, we use the following result.

Proposition 1.3 *The sum of independent n-number of λ-exponential times σ_k; $\tau = \sum_{k=1}^n \sigma_k$ is distributed by the gamma distiribution $\Gamma(n, \lambda)$, i.e.,*

$$P(\tau < t) = \int_0^t \frac{1}{(n-1)!} \lambda^n s^{n-1} e^{-\lambda s} ds.$$

Proof. By the independence of (σ_n),

$$P(\sigma_1 + \cdots + \sigma_n < t) = \int_{s_1 + \cdots s_n < t} \lambda^n e^{-\lambda(s_1 + \cdots s_n)} ds_1 \cdots ds_n.$$

By the change of variables such that $u_k = s_1 + \cdots s_k$ $(k = 1, \dots, n)$ and $s = u_n$,

$$\begin{aligned}
\int_{s_1+\cdots s_n<t} \lambda^n e^{-\lambda(s_1+\cdots s_n)} ds_1 \cdots ds_n &= \int_0^t du_n \int_0^{u_n} du_{n-1} \cdots \int_0^{u_2} du_1 \lambda^n e^{-\lambda u_n} \\
&= \int_0^t du_n \int_0^{u_n} du_{n-1} \cdots \int_0^{u_3} du_2 u_2 \ \lambda^n e^{-\lambda u_n} \\
&= \int_0^t du_n \frac{1}{(n-1)!} u_n^{n-1} \lambda^n e^{-\lambda u_n} \\
&= \int_0^t ds \frac{1}{(n-1)!} \lambda^n s^{n-1} e^{-\lambda s}
\end{aligned}$$

■

The proof of Theorem 1.1 Since τ_n is independent of σ_{n+1} and distributed by $\Gamma(n, \lambda)$, we have

$$\begin{aligned}
P(X_t = n) &= P(\tau_n \leq t < \tau_{n+1} = \tau_n + \sigma_{n+1}) \\
&= \int_0^t ds \ \frac{1}{(n-1)!} \lambda^n s^{n-1} e^{-\lambda s} P(t < s + \sigma_{n+1}) \\
&= \int_0^t ds \ \frac{1}{(n-1)!} \lambda^n s^{n-1} e^{-\lambda s} e^{-(t-s)\lambda} \\
&= e^{-\lambda t} \frac{\lambda^n}{(n-1)!} \int_0^t s^{n-1} ds = e^{-\lambda t} \frac{\lambda^n t^n}{n!}.
\end{aligned}$$

By a similar way,

$$\begin{aligned}
P(\tau_{n+1} > t+s, X_t = n) &= P(\tau_{n+1} > t+s, \tau_n \leq t < \tau_{n+1}) \\
&= P(\tau_n + \sigma_{n+1} > t+s, \tau_n \leq t) \\
&= \int_0^t du \ \frac{1}{(n-1)!} \lambda^n u^{n-1} e^{-\lambda u} P(u + \sigma_{n+1} > t+s) \\
&= \int_0^t du \ \frac{1}{(n-1)!} \lambda^n u^{n-1} e^{-\lambda u} e^{-\lambda(t+s-u)} = e^{-\lambda(t+s)} \frac{\lambda^n t^n}{n!}.
\end{aligned}$$

Hence,

(1.1) $$P(\tau_{n+1} > t+s | \ X_t = n) = e^{-\lambda s} = P(\sigma_1 = \tau_1 > s).$$

Moreover,
(1.2)
under the condition $X_t = n$, $\tau_{n+1} - t, \sigma_{n+2}, \dots, \sigma_{n+m}$ has the same distribution as $\sigma_1, \sigma_2, \dots, \sigma_m$.

In fact,

$$
\begin{aligned}
&P(\tau_{n+1}-t>s_1,\sigma_{n+2}>s_2,\dots,\sigma_{n+m}>s_m|\ X_t=n)\\
&\quad=P(\tau_n\le t<\tau_{n+1},\tau_{n+1}-t>s_1,\sigma_{n+2}>s_2,\dots,\sigma_{n+m}>s_m)/P(X_t=n)\\
&\quad=P(\tau_n\le t,\tau_{n+1}-t>s_1)P(\sigma_{n+2}>s_2,\dots,\sigma_{n+m}>s_m)/P(X_t=n)\\
&\quad=P(\tau_{n+1}-t>s_1|\ X_t=n)P(\sigma_2>s_2,\dots,\sigma_m>s_m)\\
&\quad=P(\sigma_1>s_1)P(\sigma_2>s_2,\dots,\sigma_m>s_m)\\
&\quad=P(\sigma_1>s_1,\sigma_2>s_2,\dots,\sigma_m>s_m).
\end{aligned}
$$

By this and noting that $\tau_{n+m}-t=(\tau_{n+1}-t)+\sigma_{n+2}+\cdots+\sigma_{n+m}$, we have in general, for $m\ge 1$, we can get

$$P(\tau_{n+m}>t+s|\ X_t=n)=P(\tau_m>s).$$

By subtracting the above from the above with $m+1$ instead of m, we have

$$P(\tau_{n+m}\le t+s<\tau_{n+m+1}|\ X_t=n)=P(\tau_m\le s<\tau_{m+1})=P(X_s=m).$$

By using this, for $n\ge 0,m\ge 1$,

$$
\begin{aligned}
P(X_t=n,X_{t+s}-X_t=m) &= P(X_t=n,X_{t+s}=n+m)\\
&= P(X_t=n)P(X_{t+s}=n+m|\ X_t=n)\\
&= P(X_t=n)P(\tau_{n+m}\le t+s<\tau_{n+m+1}|\ X_t=n)\\
&= P(X_t=n)P(X_s=m).
\end{aligned}
$$

By summing on $n\ge 0$,

$$P(X_{t+s}-X_t=m)=P(X_s=m)=e^{-\lambda}\frac{\lambda^m s^m}{m!}.$$

In case of $m=0$, it can be seen $P(X_{t+s}-X_t=m)=e^{-\lambda s}$, and this is included in the above. In fact, by

$$P(\tau_n>t+s|\ X_t=n)=P(\tau_n>t+s|\ \tau_n\le t<\tau_{n+1})=0,$$

if we subtract this from (1.1), then

$$P(X_{t+s}=n|\ X_t=n)=P(\tau_n\le t+s<\tau_{n+1}|\ X_t=n)=e^{-\lambda s}.$$

Thus,

$$
\begin{aligned}
P(X_t=n,X_{t+s}-X_t=0) &= P(X_t=n,X_{t+s}=n)\\
&= P(X_t=n)P(X_{t+s}=n|\ X_t=n)\\
&= P(X_t=n)e^{-\lambda s}.
\end{aligned}
$$

Hence, by summing on $n\ge 0$, we have $P(X_{t+s}-X_t=0)=e^{-\lambda s}$, and the independence of $X_t,X_{t+s}-X_t$. Moreover, by a similar way and by using (1.2), we have for $0\le t_1<\cdots<t_k$,

$$
\begin{aligned}
&P(X_{t_0}=n_0,X_{t_1}-X_{t_0}=n_1,\dots,X_{t_k}-X_{t_{k-1}}=n_k)\\
&\quad=P(X_{t_0}=n_0,X_{t_1}=n_0+n_1,\dots,X_{t_k}=n_0+\cdots+n_k)\\
&\quad=P(X_{t_0}=n_0)P(X_{t_1-t_0}=n_1,\dots,X_{t_k-t_0}=n_1+\cdots+n_k).
\end{aligned}
$$

Therefore, by repeating this, we have the following independent increments:

$$
\begin{aligned}
&P(X_{t_0}=n_0,X_{t_1}-X_{t_0}=n_1,\dots,X_{t_k}-X_{t_{k-1}}=n_k)\\
&\quad=P(X_{t_0}=n_0)P(X_{t_1-t_0}=n_1)\cdots P(X_{t_k-t_{k-1}}=n_k)\\
&\quad=P(X_{t_0}=n_0)P(X_{t_1}-X_{t_0}=n_1)\cdots P(X_{t_k}-X_{t_{k-1}}=n_k).
\end{aligned}
$$

■

1.3 Brownian motions (Wiener processes)

A real-valued stochastic process $(B_t)_{t\geq 0}$ is a **(1-dimensional) Brownian motion** is that

(1) $B_0 = 0$ a.s.

(2) (B_t) is continuous, i.e., for a.a.ω, the sample path $B_\cdot(\omega)$ is continuous.

(3) For $0 = t_0 < t_1 < \cdots < t_n$, $\{B_{t_k} - B_{t_{k-1}}\}_{k=1}^n$ are independent and $B_{t_k} - B_{t_{k-1}}$ is distributed by the normal distribution $N(0, t_k - t_{k-1})$.

Moreover, if $B_t = (B_t^1, \dots, B_t^d)$ has d numbers of independent one-dimensional Brownian motions as components, then it is called a **d-dimensional Brownian motion**. (It is realized as a product probability space of d-numbers of independent one-dimensional Brownian motions.)

In this case (B_t) satisfies the same conditions as above with the following (3)' instead of (3);

(3)' For $0 = t_0 < t_1 < \cdots < t_n$, $\{B_{t_k} - B_{t_{k-1}}\}_{k=1}^n$ are independent and $B_{t_k} - B_{t_{k-1}}$ is distributed by the d-dimensional normal distribution $N(0, (t_k - t_{k-1})I_d)$.

Let $W = C([0,\infty) \to \mathbf{R}^1)$ and let $\mathcal{W}$ be the σ-additive class determined by the local uniform convergence topology.

Moreover let $w = w(t) \in W_0 \stackrel{\text{def}}{\iff} w \in W; w(0) = 0$. For any finite number of time points $\mathbf{t}_n = (t_1, \dots, t_n); 0 \leq t_1 < t_2 < \cdots < t_n < \infty$ and for any $A_n \in \mathcal{B}^n$, $C(\mathbf{t}_n, A_n) = \{w \in W_0; (w(t_1), \dots, w(t_n)) \in A_n\}$ is called a **cylinder set)**. We denote the σ-additive class generated by all cylinder sets as $\mathcal{W}_0$ (it is known that this is the same σ-additive class determined by the relative topology of W).

Theorem 1.2 (Existence and uniqueness of Wiener measure) *There exists a unique probability measure P_B on $(\Omega, \mathcal{F}) = (W_0, \mathcal{W}_0)$ such that under this measure $B_t(w) = w(t)$ is a Brown motion.*

P_B is called the **Wiener measure**. The Brownian motion is also called the **(1-dimensional) Wiener process**.

We give the proof at the end of this section.

The distribution of d-dimensional Brownian motion $B_t = (B_t^1, \dots, B_t^d)$ is a probability measure on $W_0^d \ni w; w \in C([0,\infty) \to \mathbf{R}^d), w(0) = 0$, and this is called the **d-dimensional Wiener measure**.

The distribution of B_t is given as $P(B_t \in dx) = p_t(x)dx$, where

$$p_t(x) := \frac{1}{\sqrt{2\pi t}^d} e^{-|x|^2/(2t)} \quad (x = (x_1, \dots, x_d) \in \mathbf{R}^d,\ |x| = \sqrt{x_1^2 + \cdots + x_d^2}).$$

$g_t(x)$ is a density function of d-dimensional normal distribution $N_d(0, t)$.

The **characteristic function** of this normal distribution if given as

$$\varphi(z) = \varphi_{B_t}(z) := E[e^{iz\cdot B_t}] = e^{-t|z|^2/2} \quad (z \in \mathbf{R}^d),$$

where $z \cdot B_t = z_1 B_t^1 + \cdots + z_d B_t^d$.

Let

$$p_t(x, y) := p_t(y - x) = \frac{1}{\sqrt{2\pi t}} e^{-|y-x|^2/(2t)}.$$

Then the finite dimensional distribution of the Brownian motion is given by the following: for $0 < t_1 < t_2 < \cdots < t_n$ and $A_k \in \mathcal{B}^1$,

$$P(B_{t_1} \in A_1, \dots, B_{t_n} \in A_n) = \int_{A_1} dy_1 p_{t_1}(0, y_1) \int_{A_2} dy_2 p_{t_2 - t_1}(y_1, y_2) \cdots \int_{A_n} dy_n p_{t_n - t_{n-1}}(y_{n-1}, y_n).$$

In fact, by the independent increments letting $t_0 = 0$, we have

$$P(B_{t_k} - B_{t_{k-1}} \in A_k, k = 1, 2, \dots, n) = \prod_{k=1}^n \int_{A_k} p_{t_k - t_{k-1}}(x_k) dx_k$$

and by the change of variables $x_k = y_k - y_{k-1}$ ($y_0 = 0$) we get the above equation. Here note that $\{B_{t_1} \in A_1, B_{t_2} \in A_2\} = \{B_{t_1} \in A_1, B_{t_2} - B_{t_1} \in A_2 - A_1\}$ where $A_2 - A_1$ is a family of differences of elements, and this is not the difference set; $A_2 \setminus A_1$.

In the following $(\mathcal{F}_t)$ is a standard filtration by the Brownian motion (B_t) (basically, it is a one dimensional BM except (9)),

[Properties of Brownian motions]

(1) $EB_t^{2n} = (2n-1)!!t^n, EB_t^{2n-1} = 0$ $(n \geq 1)$.

(2) For $0 \leq s < t$, $B_t - B_s$ is independent of $\mathcal{F}_s$.

This is equivalent to independent increments. From this (B_t) is martingale (described latter), i.e., $0 \leq s < t \Rightarrow E[B_t - B_s|\ \mathcal{F}_s] = 0$

(3) The covariance $E[B_t B_s] = t \wedge s$ $(s, t > 0)$.

(4) A continuous process (X_t) is a Brownian motion $\iff$ ${}^\forall 0 \leq s < t$, $E[e^{iz(X_t - X_s)}|\ \mathcal{F}_s] = e^{-(t-s)z^2/2}$, where $(\mathcal{F}_t)$ is the canonical filtration by (X_t).

(5) The Brownian motion is invariant under the following transforms ($a > 0$ is a fixed):

$$B_t^a = B_{a+t} - B_a,\ \overline{B_t} = -B_t,\ S^a(B)_t = \sqrt{a} B_{t/a},$$

where $S^a(B)_t$ is called a **scale conversion** or **scaling**.

(6) The total variation of Brownian motion in $[T_1, T_2]$ is infinite a.s., i.e., denote a division as $\Delta = \{t_k\}; T_1 = t_0 < t_1 < \cdot < t_n = T_2$, then

$$V = \sup_{\Delta} \sum_{k=1} |B_{t_k} - B_{t_{k-1}}| = \infty \quad \text{a.s.}$$

(7) ${}^\forall \varepsilon > 0$, (B_t) has $(1/2 - \varepsilon)$-Hödler uniform continuous paths a.s., i.e, for all $\gamma > 0$,

$$\lim_{h \to 0} \sup_{s \neq t; |t-s| \leq h} \frac{|B_t - B_s|}{|t-s|^\gamma} = 0 \text{ or } \infty \text{ a.s. if } \gamma < 1/2 \text{ or } \gamma \geq 1/2.$$

(8) Sample paths of Brownian motion are not differentiable at every time points a.s.

(9) Let (B_t) be a d-dimensional Brownian motion and T be a $d \times d$ orthogonal matrix. Then (TB_t) is also a Brownian motion. Moreover, let $\tau_S := \inf\{t > 0; B_t \in S = S_r^{d-1}\}$ be a hitting time to the sphere 球面 $S = \partial B^d(0, r)$. Then the distribution of $B_{\tau_S} = B_{\tau_S(\omega)}(\omega)$ is the uniform measure on S.

Furthermore, the Brownian motion (B_t) has the following properties: (We omit the proofs.)

- $X_t = tB_{1/t}$ is also a Brownian motion with $X_0 = 0$.
- $$\limsup_{t \downarrow 0} \frac{B_t}{\sqrt{2t \log\log(1/t)}} = 1 \quad \text{a.s.}$$

  Moreover, by symmetry, $\liminf_{t \downarrow 0}$ is -1, and by scaling,

  $$\limsup_{t \uparrow \infty} \frac{B_t}{\sqrt{2t \log\log t}} = 1 \quad \text{a.s.}$$

- ${}^\forall \varepsilon > 0$, (B_t) has $(1/2 - \varepsilon)$-Hödler uniform continuity a.s. as mentioned, more precisely, it satisfies the following:

  $$\lim_{h \to 0} \sup_{s \neq t; |t-s| \leq h} \frac{|B_t - B_s|}{\sqrt{2|t-s| \log(1/|t-s|)}} = 1.$$

[The proofs of properties of Brownian motions] (1) It can be calculated directly, or by differentiating the both sides of the characteristic function $E[e^{izB_t}] = e^{-tz^2/2}$.

(2) For $0 \le s_1 < \cdots < s_n \le s < t$ and for bounded Borel functions $f(x), g(x_1, \ldots, x_n)$, it can be seen $E[f(B_t - B_s)g(B_{s_1}, \ldots, B_{s_n})] = E[f(B_t - B_s)]E[g(B_{s_1}, \ldots, B_{s_n})]$. In fact, if $n = 2$, then

$$\begin{aligned} & E[f(B_t - B_s)g(B_{s_1}, B_{s_n})] \\ = & \int_{\mathbf{R}^4} f(x_4 - x_3)g(x_1, x_4)p_{s_1}(x_1)p_{s_2-s_1}(x_1, x_2)\cdots p_{s-s_2}(x_2, x_3)p_{t-s}(x_3, x_4)dx_1\cdots dx_4 \\ = & \int_{\mathbf{R}^4} f(y_2)g(x_1, x_2)p_{s_1}(x_1)p_{s_2-s_1}(x_1, x_2)p_{s-s_2}(y_1)p_{t-s}(y_2)dx_1dx_2dy_1dy_2 \\ = & \int_{\mathbf{R}} f(y_2)p_{t-s}(y_2)dy_2 \int_{\mathbf{R}^2} g(x_1, x_2)p_{s_1}(x_1)p_{s_2-s_1}(x_1, x_2)dx_1dx_2 \\ = & E[f(B_t - B_s)]E[g(B_{s_1}, B_{s_2})], \end{aligned}$$

where we use the change of variables of $x_4 - x_3 = y_2, x_3 - x_2 = y_1$ and $\int p_{s-s_2}(y_1)dy_1 = 1$. Thus, $B_t - B_s$ is independent of $(B_{s_1}, \ldots, B_{s_n})$, i.e., $\mathcal{F}_s^0$, and hence, $\mathcal{F}_s$.

Moreover, it can be taken as $f(x) = x$, and hence, $E[B_t - B_s | \mathcal{F}_s] = E[B_t - B_s] = 0$

(3) If $0 \le s < t$, then $E[B_tB_s] = E[(B_t - B_s)B_s + B_s^2] = EB_s^2 = s$.

(4) $(\Rightarrow)$ is obvious from the above. $(\Leftarrow)$ By $E[e^{iz(X_t-X_s)} | \mathcal{F}_s] = e^{-(t-s)z^2/2}$, for $0 \le s_1 < \cdots < s_n < s < t$ and for a bounded Borel function $f(x_1, \ldots, x_n)$, we have

$$\begin{aligned} E[e^{iz(X_t-X_s)}f(X_{s_1}, \ldots, X_{s_n})] &= E[E[e^{iz(X_t-X_s)} | \mathcal{F}_s]f(X_{s_1}, \ldots, X_{s_n})] \\ &= e^{-(t-s)z^2/2}E[f(X_{s_1}, \ldots, X_{s_n})]. \end{aligned}$$

Hence, $X_t - X_s$ is independent of $\mathcal{F}_s$ and it is distributed by a normal distribution $N(0, t)$.

(5) If we denote each transform as X_t, then it is enough to show $E[e^{iz(X_t-X_s)} | \mathcal{F}_s] = e^{-(t-s)z^2/2}$ and to be a continuous process. However,they are almost evident.

(6) We may let $[T_1, T_2] = [0, 1]$ by (5). B_t is uniform continuous at $t \in [0, 1]$ a.s., and hence,

$$\delta_n = \max_{1\le k\le n} |B_{k/n} - B_{(k-1)/n}| \to 0 \ (n \to \infty) \quad \text{a.s.}$$

Let $X_n = \sum_{k=1}^n (B_{k/n} - B_{(k-1)/n})^2$ and $Z_k = (B_{k/n} - B_{(k-1)/n})^2 - 1/n$. Then $EZ_k^2 = 3/n^2 - 2/n^2 + 1/n^2 = 2/n^2$, and thus,

$$E(X_n - 1)^2 = \sum_{k=1}^n EZ_k^2 = \frac{2}{n} \to 0.$$

Therefore, ${}^\exists\{n_k\}; X_{n_k} \to 1$ a.s. From these, we have

$$V = \sup_{\Delta} \sum_{k=1}^n |B_{t_k} - B_{t_{k-1}}| \ge \frac{X_{n_k}}{\delta_{n_k}} \to \infty \quad \text{a.s.}$$

(7) Since $E|B_t - B_s|^{2n} = c_n|t - s|^n$ ($c_n = (2n-1)!!$), and by the **Kolmogorov's continuity theorem** (described at the end of this section), for ${}^\forall\gamma < (n-1)/(2n) \to 1/2$, (B_t) is γ-Hölder uniform continuous.

Moreover, in order to show that it is ∞ if $\gamma = 1/2$, we fix ${}^\forall L \ge 1$ and let $A_n = \{|B_{k/n} - B_{(k-1)/n}| \le L/\sqrt{n}, k = 1, 2, \ldots, n\}$. By the scaling property we have $P(|B_{k/n} - B_{(k-1)/n}| \le L/\sqrt{n}) = P(|B_k - B_{k-1}| \le L) = P(|B_1| \le L) =: p_L \in (0, 1)$. The independence implies $P(A_n) = p_L^n$. Hence, $\sum_{n\ge1} P(A_n) < \infty$. By Borel-Cantelli's lemma, $P(\limsup A_n) = 0$. Thus, it holds that with probability one, ${}^\exists N = N(\omega) \ge 1; {}^\forall n \ge N, {}^\exists k \le n, |B_{k/n} - B_{(k-1)/n}| > L/\sqrt{n}$. By the arbitrariness of $L \ge 1$, we have the desired result.

(8) We may set the time interval as $[0, 1)$. The proof is done by the following steps:

$$P({}^\exists s \in [0, 1); {}^\exists B_s') \le P\left(\bigcup_{m\ge1} \bigcup_{N\ge1} A_{m,N}\right) = 0,$$

where

$$A_{m,N} = \bigcap_{n\geq N} \bigcup_{i=1}^{n+1} \bigcap_{j=i+1}^{i+3} \left\{ |B_{j/n} - B_{(j-1)/n}| \leq \frac{8m}{n} \right\}.$$

First, if ${}^{\exists}s_0 \in [0,1]; {}^{\exists}B'_{s_0}$, then

$${}^{\exists}m \geq 1, {}^{\exists}t_0 > s_0; |B_t - B_{s_0}| \leq m(t - s_0), \quad t \in [s_0, t_0].$$

Moreover, if $s_k = ([ns_0] + k)/n, 1 \leq k \leq 4$, then $s_0 < s_1 \leq \cdots \leq s_4, s_4 - s_0 \leq 4/n$ and ${}^{\exists}N \geq 1; {}^{\forall}n \geq N, s_k \in [s_0, t_0]$. Hence from the above, we have for each $k = 2, 3, 4$,

$$|B_{s_k} - B_{s_{k-1}}| \leq |B_{s_k} - B_{s_0}| + |B_{s_0} - B_{s_{k-1}}| \leq 2m(s_4 - s_0) \leq \frac{8m}{n}.$$

That is, ${}^{\exists}m, N \geq 1; {}^{\forall}n \geq N$, if we set $i = [ns_0] + 1$, then $1 \leq i \leq n+1$ and $|B_{j/n} - B_{(j-1)/n}| \leq 8m/n$ for $j = i+1, i+2, i+3$. Therefore, we have the first inequality. Next, it is enough to show ${}^{\forall}m, N \geq 1, P(A_{m,N}) = 0$. By a simple calculation,

$$P\left(\left| B_{j/n} - B_{(j-1)/n} \right| \leq \frac{8m}{n} \right) = \frac{2}{\sqrt{2\pi/n}} \int_0^{8m/n} e^{-x^2/(2/n)} dx = \frac{2}{\sqrt{2\pi}} \int_0^{8m/\sqrt{n}} e^{-x^2/2} dx \leq \frac{C}{\sqrt{n}}.$$

Thus, we have

$$\begin{aligned} P(A_{m,N}) &\leq \inf_{n\geq N} P\left(\bigcup_{i=1}^{n+1} \bigcap_{j=i+1}^{j+3} \left\{ |B_{j/n} - B_{(j-1)/n}| \leq \frac{8m}{n} \right\}. \right) \\ &\leq \liminf_{n\to\infty} \sum_{i=1}^{n+1} \prod_{j=i+1}^{i+3} P\left(|B_{j/n} - B_{(j-1)/n}| \leq \frac{8m}{n} \right) \leq \lim_{n\to\infty} (n+1) \left(\frac{C}{\sqrt{n}} \right)^3 = 0. \end{aligned}$$

■

(9) The first of half can be immediately obtained by calculating the characteristic function of the increments of finite numbers of time points. In fact, for $z, x \in \mathbf{R}^d$, by $\langle z, Tx \rangle = \langle {}^tTz, x \rangle$, we have, for $0 \leq t_0 < t_1 < \cdots < t_n, z_k \in \mathbf{R}^d (k = 1, 2, \ldots, n)$,

$$E\left[\exp\left\{ i \sum_{k=1}^n \langle z_k, T(B_{t_k} - B_{t_{k-1}}) \rangle \right\} \right] = e^{-\sum_{k=1}^n (t_k - t_{k-1}) |{}^tTz_k|^2} = e^{-\sum_{k=1}^n (t_k - t_{k-1}) |z_k|^2}.$$

Thus, if we let every component of z_k be 0 except z_j, then the j component of (TB_t) is a one-dimensional Brownian motion, and the above last expression is equal to the product of the each forms. Hence, the independence of each components holds.

On the later half, for any orthogonal matrix T, let τ_S^T be the hitting time to S of TB_t, then $\tau_S^T = \tau_S$. By the above result and the uniqueness of the distribution of Brownian motion we have that for ${}^{\forall}A \in \mathcal{B}(S)$,

$$P(B_{\tau_S} \in A) = P(T(B)_{\tau_S^T} \in A) = P(T(B)_{\tau_S} \in A) = P(T(B_{\tau_S}) \in A) = P(B_{\tau_S} \in T^{-1}A).$$

This implies $\mu_S(d\xi) := P(B_{\tau_S} \in d\xi)$ is a rotation invariant measure on S. Moreover, if $t \to \infty$, then $P(B_t \in B(0,r)) = \int_{B(0,r)} p_t(x) dx \to 0$. Thus, we see $P(\tau_S < \infty) = 1$ and we get $\mu_S(S) = 1$. (If $P(\tau_S = \infty) > 0$, then $0 < P({}^{\forall}t > 0, B_t \in B(0,r)) \leq \limsup_{t\to\infty} P(B_t \in B(0,r)) = 0$. This is inconsistent. ■

[Construction of Brownian motions] It is well-known that there are 3 ways, however, we give the simplest way.

It is enough to show the case of $t \in [0,1]$. Because the case of $[0,T]$ is the same, and by the uniqueness it is possible to extend to $[0,\infty)$. Let $D = \bigcup_{n\geq 1} \{k/2^n; k = 0, 1, \ldots, 2^n\}$ be the family of all binary rational numbers in $[0,1]$.

First, by using **Kolmogorov's extension theorem** to the probability space on $\mathbf{R}^\infty$, a probability P_0 can be constructed on $\mathbf{R}^D$ ($\ni w = w(t) : D \to \mathbf{R}$ is a function) such that the every finite dimensional distribution of $X_t(w) = w(t)$ is the same as the Brownian motion.

Furthermore, it is possible to show that (X_t) satisfies the conditions of the following **Kolmogorov's continuity theorem**. Hence, (X_t) is uniform continuous on D a.s., and $\widetilde{X_t} = \lim_{r\downarrow t; r\in D} X_r$ is continuous. Thus, $B_t = \widetilde{X_t}$ is the desired one.

Theorem 1.3 (Kolmogorov's continuity theorem)

(1) Let $D = \bigcup_{n\geq 1}\{k/2^n; k = 0, 1, \dots, 2^n\}$ be the family of all binary rational numbers in $[0,1]$. In general, if a stochastic process $\{X_t\}_{t\in D}$ on a Banach space $(B, \|\cdot\|)$ satisfies

$$^\exists C, \alpha, \beta > 0; E\|X_t - X_s\|^\alpha \leq C|t-s|^{1+\beta},$$

then X_t is uniform continuous on D a.s.

(2) If $\{X_t\}_{t\in[0,1]}$ satisfies the above inequation for $^\forall s, t \in [0,1]$, then there exists a continuous modification $\{\widetilde{X_t}\}_{t\in[0,1]}$ uniquely, and it is γ-Hölder uniform continuous a.s. for $^\forall\gamma < \beta/\alpha$; i.e.,

$$\lim_{h\to 0} \sup_{s\neq t; |t-s|\leq h} \frac{\|X_t - X_s\|}{|t-s|^\gamma} = 0 \quad a.s.$$

Proof. For simplicity, we denote $\|\cdot\| = |\cdot|$. Fix $0 < \gamma < {}^\forall\delta < \beta/\alpha$ and set $\Delta_n = 1/2^n$. Note $\beta - \alpha\delta > 0$.

(1) By a simple calculation,

$$E\left[\sum_{n\geq 1}\sum_{k=1}^{2^n}\left(\frac{|X_{k\Delta_n} - X_{(k-1)\Delta_n}|}{\Delta_n^\delta}\right)^\alpha\right] \leq \sum_{n\geq 1}\sum_{k=1}^{2^n} C\Delta_n^{1+(\beta-\alpha\delta)} = C\sum_{n\geq 1}\Delta_n^{\beta-\alpha\delta} < \infty.$$

Thus, the inside of the above expectation is finite a.s. and hence,

$$\lim_{n\to\infty}\sum_{k=1}^{2^n}\left(\frac{|X_{k\Delta_n} - X_{(k-1)\Delta_n}|}{\Delta_n^\delta}\right)^\alpha = 0 \quad \text{a.s.}$$

Therefore,

$$P\left(^\exists n_0; {}^\forall n \geq n_0, {}^\forall k = 1, 2, \dots, 2^n, |X_{k\Delta_n} - X_{(k-1)\Delta_n}| < \Delta_n^\delta\right) = 1.$$

We denote this event as Ω_0; $P(\Omega_0) = 1$. On Ω_0, we can see that

$$^\exists n_0; {}^\forall r, r' \in D; 0 < r - r' < \Delta_{n_0}, |X_r - X_{r'}| \leq C'|r-r'|^\delta \tag{1.3}$$

with $C' = 2/(1-2^{-\delta})$, and hence, (X_t) is uniform continuous on D. In fact, $^\exists n \geq n_0; \Delta_{n+1} < r - r' \leq \Delta_n$ and r, r' are in the same interval $[k\Delta_n, (k+1)\Delta_n]$, or each is in adjacent intervals $((k-1)\Delta_n, k\Delta_n), (k\Delta_n, (k+1)\Delta_n)$. In case of $r, r' \in [k\Delta_n, (k+1)\Delta_n]$, we estimate $|X_r - X_{r'}| \leq |X_r - X_{k\Delta_n}| + |X_{k\Delta_n} - X_{r'}|$. At first, by using binary notation we denote $r - k\Delta_n = \varepsilon_1\Delta_{n+1} + \cdots + \varepsilon_p\Delta_{n+p}$ ($^\exists p \geq 1, \varepsilon_i = 0$ or 1). Let $r_0 = k\Delta_n, r_j = k\Delta_n + \varepsilon_1\Delta_{n+1} + \cdots + \varepsilon_j\Delta_{n+j}$ ($j = 1, 2, \dots, p$). By $r_j - r_{j-1} = \varepsilon_j\Delta_{n+j} \leq \Delta_{n+j}$, we have, on Ω_0

$$|X_r - X_{k\Delta_n}| \leq \sum_{j=1}^p |X_{r_j} - X_{r_{j-1}}| < \sum_{j=1}^\infty \Delta_{n+j}^\delta = \Delta_{n+1}^\delta/(1-2^{-\delta}).$$

$|X_{k\Delta_n} - X_{r'}|$ satisfies the same estimate, and hence, (noting $r - r' > \Delta_{n+1}$)

$$|X_r - X_{r'}| \leq 2\Delta_{n+1}^\delta/(1-2^{-\delta}) \leq 2(1-2^{-\delta})^{-1}|r-r'|^\delta.$$

In case of $r \in ((k-1)\Delta_n, k\Delta_n), r' \in (k\Delta_n, (k+1)\Delta_n)$, by $k\Delta_n - r, r' - k\Delta_n < \Delta_n$ and by using binary notation, we can get the same estimation.

Therefore, (1.3) holds.

(2) By (1), X_t is uniform continuous. Define $\widetilde{X_t}$ as X_t for $t \in D$, and $\widetilde{X_t} = \lim_{r \in D; r \downarrow t}$ for $t \notin D$. Then it is a continuous process and

$$^{\exists}n_0; {}^{\forall}s, t \in [0,1]; 0 < t - s < \frac{1}{2^{n_0}}, |\widetilde{X_t} - \widetilde{X_s}| \le C'|t-s|^{\delta}$$

holds. By $\gamma < \delta < \beta/\alpha$, it is obvious that it has a γ-Hölder uniform continuity. Moreover, for $t \notin D$, if we take $r_n \in D; r_n \downarrow t$, then

$$E|\widetilde{X_t} - X_t|^{\alpha} \le \liminf_{n\to\infty} E|\widetilde{X_{r_n}} - X_t|^{\alpha} \le C \liminf_{n\to\infty} |r_n - t|^{1+\beta} = 0.$$

Thus, ${}^{\forall}t \in [0,1], P(X_t = \widetilde{X_t}) = 1$, that is, (X_t) and $(\widetilde{X_t})$ are equivalent, and this continuous modification $(\widetilde{X_t})$ is unique in the sense of strong equivalence. ■

1.4 Markov processes and martingales

(X_t) is a **Markov process (MP)** $\overset{\text{def}}{\Longleftrightarrow}$ For all $0 \le s < t$ and for all bounded Borel functions f, $E[f(X_t)|\ \mathcal{F}_s] = E[f(X_t)|\ X_s]$ a.s. Moreover, if (the above)$= E[f(X_{t-s})|\ X_0 = x]|_{x=X_s}$ a.s., then it is called a **time-homogeneous MP)** More precisely, let (X_t, P_x) be a Markov process staring from x; $P_x = P(\cdot|\ X_0 = x)$, Then, it is time-homogeneous if and only if for all x, $E_x[f(X_t)|\ \mathcal{F}_s] = E_{X_s}[f(X_{t-s})]$ a.s.

When $\mathcal{F}_t = \mathcal{F}_t^0 := \sigma(X_s; s \le t)$, the above definition is that for $0 \le s_1 < \cdots < s_n = s < t$, $a_1, \dots, a_n \in \mathbf{R}$,

$$E[f(X_t)|\ X_{s_1} \le a_1, \dots, X_{s_n} \le a_n] = E[f(X_t)|\ X_s \le a_n].$$

Moreover, it is time-homogeneous if and only if (the above)$= E[E[f(X_{t-s})|\ X_0 = x]|_{x=X_s}|\ X_s \le a_n]$. That is, for all x,

$$E_x[f(X_t)|\ X_{s_1} \le a_1, \dots, X_{s_n} \le a_n] = E_x[E_{X_s}[f(X_{t-s})]|\ X_s \le a_n].$$

If a Markov process is on a discrete space, then it is also called a Markov chain.

A countable set S valued stochastic process $(X_t)_{t \ge 0}$ is a **continuous-time Markov chain** if ti has the following Markov property; For all $s, t \ge 0$, $i, j, k_\ell \in S, 0 \le u_\ell < s$ $(\ell \le \ell_0)$,

$$P(X_{t+s} = j|\ X_s = i, X_{u_\ell} = k_\ell\ (\ell \le \ell_0)) = P(X_{t+s} = j|\ X_s = i).$$

Moreover, the following is a time-homogeneous property;

$$P(X_{t+s} = j|\ X_s = i) = P(X_t = j|\ X_0 = i).$$

We denote this as a **transition probability** $q_t(i,j) = P(X_t = j|\ X_0 = i)$.

Theorem 1.4 *A Poisson process is a continuous-time Markov chain.*

It is clear by the following question.

Question 1.2 *In general, for a countable linear space valued continuous-time stochastic process, if it has independent increments, then it is a Markov chain.*

Answer. Let X_t be a stochastic process satisfying the assumption. For $0 \le t_1 < t_2 < \cdots < t_n < t_{n+1}$, the independence of $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_{n+1}} - X_{t_n}$ implies $X_{t_{n+1}} - X_{t_n}$, $(X_{t_1}, \dots, X_{t_n})$ are independent, and $X_{t_{n+1}} - X_{t_n}$, X_{t_n} are independent. Hence, these imply the Markov property;

$$\begin{aligned} P(X_{t_{n+1}} = j_{n+1}|\ X_{t_k} = j_k, 1 \le k \le n) &= P(X_{t_{n+1}} - X_{t_n} = j_{n+1} - j_n|\ X_{t_k} = j_k, 1 \le k \le n) \\ &= P(X_{t_{n+1}} - X_{t_n} = j_{n+1} - j_n) \\ &= P(X_{t_{n+1}} - X_{t_n} = j_{n+1} - j_n|\ X_{t_n} = j_n) \\ &= P(X_{t_{n+1}} = j_{n+1}|\ X_{t_n} = j_n). \end{aligned}$$

■

[Markov property of Brownian motions]

For a Brownian motion (B_t, P) starting from 0, $(x + B_t)$ is a Brownian motion starting from x. We denote the distribution as P_x on $(W = W^d, \mathcal{W})$. Then $(x + B_t)$ is changed to (B_t). Under P_x, this is the same as $B_t(w) = w(t), w \in W$ and $P_x(B_0 = x) = 1$ holds. That is, $(B_t, P_x) \overset{(d)}{=} (x + B_t, P_0)$ and $P = P_0$ hold.

For the filtration, let $\mathcal{N} = \{N \in \mathcal{W}; {}^\forall x \in \mathbf{R}^d, P_x(N) = 0\}$ and the canonical filtration is given as $\mathcal{F}_t = \mathcal{F}_t^0 \vee \mathcal{N}$ with $\mathcal{F}_t^0 = \sigma(B_s; s \le t)$. The right-continuous filtration is given as $\mathcal{F}_t^* \equiv \mathcal{F}_{t+} := \cap_{\varepsilon > 0} \mathcal{F}_{t+\varepsilon}$. Then, it holds that $\mathcal{F}_t^* = \mathcal{F}_t$ (it is shown later). .

For each $s \ge 0$, a **shift operator** θ_s on W is defined by $\theta_s w(t) := w(t + s)$.

Theorem 1.5 (Markov property for $(\mathcal{F}_t)$) *Let Y be a bounded $\mathcal{W}$-measurable function. ${}^\forall x \in \mathbf{R}^d, {}^\forall s \ge 0$, it holds*

$$E_x[Y \circ \theta_s |\, \mathcal{F}_s] = E_{B_s}[Y] \ \ a.s.$$

For a bounded Borel function f and $0 \le s < t$, if we let $Y = f(B_{t-s})$, then $Y \circ \theta_s = f(B_t)$ and the above equation is $E_x[f(B_t)|\, \mathcal{F}_s] = E_{B_s}[f(B_{t-s})] = E[f(B_{t-s})|\, B_0 = x]|\,_{x=B_s}$. This implies (B_t) is a time-homogeneous MP

Proof. Let f, f_j be bounded Borel functions on $\mathbf{R}^1$ and for each $Y = f(B_t), \prod_{j \le n} f_j(B_{t_j})$ $(0 \le t_1 < \cdots < t_n)$ if we show the result, then for a general bounded $\mathcal{W}$-measurable Y, it can be obtained as a limit of linear combinations of the second forms, and hence, we can get the desired result.

If $Y = f(B_t)$, then it is enough to show the case of $f(x) = e^{izx}$ (${}^\forall z \in \mathbf{R}$). That is, $Y = e^{izB_t}$ and $Y \circ \theta_s = e^{izB_{t+s}}$. Thus, we have

$$E_x[e^{izB_{t+s}}|\, \mathcal{F}_s] = E_x[e^{izB_s} e^{iz(B_{t+s} - B_s)}|\, \mathcal{F}_s] = e^{izB_s} E_0[e^{izB_t}] = E_0\ [e^{iz(x+B_t)}]\ \Big|_{x=B_s} = E_{B_s}[e^{izB_t}].$$

If $Y = \prod_{j \le n} f_j(B_{t_j})$ $(0 \le t_1 < \cdots < t_n)$, then the result can be shown by the induction on n. We assume it holds for n, and show the case of $n + 1$.

$$\begin{aligned} E_x[Y \circ \theta_s |\, \mathcal{F}_s] &= E_x\left[E_x\left[\prod_{j \le n+1} f_j(B_{t_j + s}) \middle|\, \mathcal{F}_{t_1+s} \right] \middle|\, \mathcal{F}_s \right] \\ &= E_x\left[f_1(B_{t_1+s}) E_{B_{t_1+s}}\left[\prod_{j=2}^{n+1} f_j(B_{t_j - t_1}) \right] \middle|\, \mathcal{F}_s \right] \\ &= E_{B_s}\left[f_1(B_{t_1}) E_{B_{t_1}}\left[\prod_{j=2}^{n+1} f_j(B_{t_j - t_1}) \right] \right] \\ &= E_{B_s}\left[f_1(B_{t_1}) \prod_{j=2}^{n+1} f_j(B_{t_j}) \right] = E_{B_s}[Y]. \end{aligned}$$

Here we use the assumption of the induction in the second line and use the above result in the third line noting that the following is a bounded Borel function;

$$f(x) = E_x\left[\prod_{j=2}^{n+1} f_j(B_{t_j - t_1}) \right] = E_0\left[\prod_{j=2}^{n+1} f_j(x + B_{t_j - t_1}) \right].$$

In the fourth line, we also use the assumption of the induction under the conditional expectation of $\mathcal{F}_{t_1}$.
■

The above Markov property also holds for $(\mathcal{F}_t^*)$.

Theorem 1.6 (Markov property for $(\mathcal{F}_t^*)$) *For a bounded $\mathcal{W}$-measurable function Y and for $^\forall x \in \mathbf{R}^d, {}^\forall s \geq 0$,*

$$E_x[Y \circ \theta_s | \ \mathcal{F}_s^*] = E_{B_s}[Y] \ a.s.$$

Moreover, this implies $\mathcal{F}_t^* = \mathcal{F}_t$.

Proof. We first show $\mathcal{F}_t^* = \mathcal{F}_t$ by using the Markov property. Let $^\forall A, B \in \mathcal{F}_s^*$ and set $Y = 1_B$, $\widetilde{Y} = E_{B_s}[1_B]$. It holds $E_x[1_B; A] = E_x[\widetilde{Y}; A]$, i.e., $E_x[(1_B - \widetilde{Y}); A] = 0$. Since $\widetilde{Y}$ is $\mathcal{F}_s$-measurable and $A \in \mathcal{F}_s^*$ is arbitrary, we have $1_B = \widetilde{Y}$ a.s. Hence, $^\exists \widetilde{B} \in \mathcal{F}_s; N \in \mathcal{N}; B = \widetilde{B} \cup N \in\in \mathcal{F}_s$. (In fact, noting that $N = \{1_B \neq \widetilde{Y}\} \in \mathcal{N}$, $\{1_B = \widetilde{Y}\} = N^c \in \mathcal{F}_s$, $1_B = \widetilde{Y} 1_{N^c} + 1_{B\cap N}$ and the 1st term is $\mathcal{F}_s$-measurable and take 0 or 1, we can take $\widetilde{B} = \{\widetilde{Y} 1_{N^c} = 1\}$.) Thus, $\mathcal{F}_s^a st \subset \mathcal{F}_s$.

Next we show the Markov property. Set $Y = f(B_t)$ with a bounded Borel function f. $^\forall \varepsilon > 0, A \in \mathcal{F}_{s+\varepsilon}$ implies $E_x[f(B_{t+s+\varepsilon})| \ \mathcal{F}_{s+\varepsilon}] = E_{B_{s+\varepsilon}}[f(B_t)]$ a.s., that is, $E_x[f(B_{t+s+\varepsilon})1_A] = E_x[E_{B_{s+\varepsilon}}[f(B_t)1_A]$. If $\varepsilon \downarrow 0$, then by the continuity of BM, the boundedness of f and by using convergence theorem, we have $E_x[f(B_{t+s})1_A] = E_x[E_{B_s}[f(B_t)1_A]$. Therefore, we can extend to bounded Borel functions Y as above. ■

Theorem 1.7 (Blumenthal's zero-one law) *For all $A \in \mathcal{F}_0 = \mathcal{F}_0^*$, $P(A) = 0$ or 1.*

Proof. If $A \in \mathcal{F}_0$, then

$$P_x(A) = E_x[1_A] = E_x[E_x[1_A \circ \theta_0]; A] = E_x[E_{B_0}[1_A]; A] = E_x[P_x(A); A] = P_x(A)^2.$$

Hence, $P_x(A) = 0$ or 1. ■

From this result, the one-dimensional Brownian motion (B_t) staring from 0 moves immediately to positive (and hence, to negative), that is, for $\tau_{(0,\infty)} := \inf\{t > 0; B_t > 0\}$, $P(\tau_{(0,\infty)} = 0) = 1$ holds.

(X_t) is **martingale** $\overset{\text{def}}{\iff}$ For all $0 \leq s \leq t$, $X_t \in L^1$ and $E[X_t| \ \mathcal{F}_s] = X_s$ a.s.

We often use (M_t) as martingale, so the above conditions are $M_t \in L^1$, $E[M_t| \ \mathcal{F}_s] = M_s$ a.s.

Then, the means are constant; $EM_t = EM_0$.

If $E[X_t| \ \mathcal{F}_s] \geq X_s$ a.s., then it is called as **sub-martingale**. Then, the means are increasing; $EX_0 \leq EX_s \leq EX_t$. Moreover, if the inverse inequality holds, then it is called as **super-martingale)**.

Theorem 1.8 (Doob-Meyer decomposition) *If (X_t) is continuous sub-martingale and if it is in class (DL), i.e., $^\forall a > 0$, $\{X_{\tau\wedge a}\}_\tau$ is uniform integrable (where τ is a stopping time $\overset{\text{def}}{\iff}$ $^\forall t \geq 0, \{\tau \leq t\} \in \mathcal{F}_t$), then $X_t = A_t + M_t$ such that (M_t) is continuous martingale, and (A_t) is continuous increasing process; $A_0 = 0$. This decomposition is unique.*

For a sequence of stopping timess $\tau_n; \uparrow\uparrow \infty$ a.s., if $(X_{t\wedge\tau_n})$ is martingale, then (X_t) is called **a local martinga;e**.

If $\mathcal{M}$ is a family of all martingales, then the family of all local martingales is denoted as $\mathcal{M}_{\text{loc}}$.

Martingale is an important notion and stochastic integrals are continuous or rcll martingales. In that, L^2-martingale plays a key role, however, in the jump type, we also treat L^1-martingale.

Therefore, we assume at least that martingale is rcll and in L^1. The results of discrete-time martingale can be extended to the case of continuous-time martingale by using right-continuity.

2 C-spaces and D-spaces

A **Polish space** is a complete separable metrizable topology space, that is, it is isomorphic to a complete metric space which has a countable dense subset.

d-dimensional Euclid space $\mathbf{R}^d$ and open interval $(0,1)$ are isomorphic to $\mathbf{R}^1$, and hence they are Polish.

Let I be an interval of $\mathbf{R}^1$ and denote the all mappings $f : I \to \mathbf{R}$ as $\mathbf{R}^I$.

The following function spaces are Polish, each is called C-space or D-space.

$$C = C(I) = \{f \in \mathbf{R}^I : \text{continuous}\}, \quad D = D(I) = \{f \in \mathbf{R}^I : \text{1st kind of discontinuous}\},$$

where $f : I \to \mathbf{R}^1$ is 1st kind of discontinuous means that f is right-continuous at each points of I except the right-point and has left-hand limits at each points of I except the left-point.

2.1 C-spaces and uniform convergence topology

The C-space $C = C(I)$ which is a family of all continuous functions on I is complete separable under the following metric if I is compact, i.e., $I = [a, b]$ $(-\infty < a < b < \infty)$. its topology is called the **uniform convergence topology**.

$$d_u(f, g) = \sup_{t \in I} |f(t) - g(t)|.$$

If I is not compact, then $^\exists I_n = [a_n, b_n]; I = \bigcup I_n$ and it is separable and complete by the following metric: Its topology is called **local uniform convergence topology**.

$$d_u(f, g) = \sum_{n \geq 1} 2^{-n}(1 \wedge \sup_{t \in I_n} |f(t) - g(t)|).$$

The completeness is well-known and it is easy to see. On the separability, by Weierstrass's polynomial approximation theorem, we see that the family of rational polynomials is dense, and hence, C is separable.

Question 2.1 *Show the above results.*

2.2 D spaces and Skorohod topology

The D-space $D = D(I)$ which is a family of first kind of discontinuous functions on a interval I is complete, however not separable under the uniform convergence topology. (You may consider $f_\alpha = 1_{[0,\alpha] \cap I}$ $(\alpha > 0)$.)

However, it is Polish under the Skorohod topology determined by the Billingsley's metric.

When I is compact, i.e., $I = [a, b]$ $(-\infty < a < b < \infty)$, $D = D(I)$ is complete and separable under the following Billingsley's metric d_B.

Let Φ be a family of all isomorphic mappings such that preserving the order. For $\varphi \in \Phi$, set

$$\lambda(\varphi) = \sup_{s \neq t} \left| \log \frac{\varphi(t) - \varphi(s)}{t - s} \right|$$

and $\varphi \in \Psi \overset{\text{def}}{\Longleftrightarrow} \lambda(\varphi) < \infty$. We define

$$d_B(f, g) = \inf_{\varphi \in \Psi} \{\|f \circ \varphi - g\|_\infty + \lambda(\varphi)\}.$$

Note that the following metric d_S determines the same topology (which is called the **Skorohod topology**. Under d_S, D is separable, however not complete.

$$d_S(f, g) = \inf_{\varphi \in \Phi} \{\|f \circ \varphi - g\|_\infty + \|\varphi - i\|_\infty\}.$$

If I is not compact, then D can be complete and separable by a similar way to C.

2.3 Continuous processes and discontinuous processes

A stochastic process (X_t) is called **continuous-type** if it has continuous sample paths X. a.s. If it has discontinuous sample paths a.s., then it called **discontinuous-type**, however, we omit the explosion paths and right-discontinuous paths. The discontinuous means right-continuous and having left-hand limits, i.e., **rcll**=right-conti. has left-limit, or **cádlág** (French). It is called **first kind of discontinuous**.

If it has explosion at a time $T > 0$, then it is enough to consider for $t \in [0, T)$. If it has right-hand limit and it is right-discontinuous, and left-continuous at a time point, then it is possible to be right-continuous and to have left-hand limit, hence, it is not essential. Moreover, if it is left-continuous at every point, then the future is determined by the previous value. So it is not so interest.

The Brownian motion is a continuous Markov process and it is a basic and central process in continuous-type.

The Poisson process is the simplest jump process, however, it's just one example of jump-type, and the following Poisson random measure is an important tool.

2.4 Poisson random measures

Let $(Z, \mathcal{Z})$ be a measurable space and $\lambda(dz)$ be a σ-finite measure on it.

$N(dz) = N(\omega; dz)$ is a **Poisson random measure** on Z with a mean measure λ $\overset{\text{def}}{\Longleftrightarrow}$

(1) For a.a. $\omega \in \Omega$, $N(\omega; dz)$ is a measure on $(Z, \mathcal{Z})$.

(2) $^\forall A \in \mathcal{Z}$, if $\lambda(A) < \infty$, then $N(A)$ is a $\lambda(A)$-Poisson random variable, i.e., it is distributed by a Poisson distribution with a parameter $\lambda(A)$. If $\lambda(A) = \infty$, then $N(A) = \infty$ a.s.

(3) If $A_n \in \mathcal{Z}$ are disjoint, then $N(A_n)$ are independent.

By (2), $\widehat{N}(dz) := E[N(dz)] = \lambda(dz)$.

In this text, we set Z be a time-space, i.e., $Z = [0, \infty) \times \mathbf{R}^m \ni (t, z)$, $\mathcal{Z} = \mathcal{B}^1([0, \infty) \times \mathbf{R}^m)$ and let $\nu(dz)$ be a measure on $\mathbf{R}^m$ such that $\nu(\{0\}) = 0$ and that $^\forall n \geq 1, \nu(|z| \geq 1/n) < \infty$ (hence, ν is σ-finite). Then ν is called a **Lévy measure**.

In this case, $N(dtdz)$ is a $dt\nu(dz)$-Poisson random measure.

We have the following result.

Proposition 2.1 *Let (τ_k, ξ_k) be the point masses of $N(dt, dz)$, i.e.,*

$$N(dt, dz) = \sum \delta_{(\tau_k, \xi_k)}(dt, dz).$$

Then, $^\forall k, j \geq 1$, $P(\tau_k \neq \tau_j) = 1$. That is, $N(dt, dz)$ has only one point mass at most at the same time-point. $P(^\forall t \geq 0, N(\{t\} \times \mathbf{R}^m) = 0 \ \text{ or } 1) = 1$.

This comes from the continuity of the time-part of the mean measure.

For simplicity, let $\mathbf{R}^m = \mathbf{R}^1$.

We first give the **construction of $dt\nu(dz)$-Poisson random measure**.

For any fixed $T > 0$ let $t \in [0, T]$ and set $Z_0 = \{|z| \geq 1\}$, $Z_n = \{1/(n+1) \leq |z| < 1/n\}$ $(n \geq 1)$. $^\forall n \geq 0$, let $\nu_n = \nu|_{Z_n}$. and

$$\overline{\lambda_n}(dt, dz) \equiv \frac{\lambda(dt, dz)}{\lambda([0, T] \times Z_n)} := \frac{dt\nu_n(dz)}{T\nu(Z_n)} \quad (\lambda(dt, dz) := dt\nu(dz)) \quad \text{on } [0, T] \times Z_n.$$

Let $\{Y_k^n = (\tau_k^n, \xi_k^n)\}_{k \geq 1}$ be independent random variables distributed by this, i.e., $P(Y_k^n \in dtdz) = \overline{\lambda_n}(dt, dz)$. Moreover, let K_n be a $T\nu(Z_n)$-Poisson variable, and $\{Y_k^n, K_m; n \geq 0, k \geq 1, m \geq 0\}$ be independent. We define

$$N_n(dt, dz) = \sum_{k=1}^{K_n} \delta_{Y_k^n}(dt, dz) = \sum_{k=1}^{K_n} 1_{dtdz}(Y_k^n), \quad N = \sum_{n \geq 0} N_n$$

This N is a desired random measure.

Question 2.2 *Show the above result.*

Note that $N_n(A) = k$, $K_n = m \geq k$ means k numbers of Y_j^n are in A and the rest $m-k$ numbers are in $([0,T] \times Z_n) \setminus A$. We can see that $N_n(A)$ is a $\lambda(A)$-Poisson variable. The sum of independent Poisson variables is again a Poisson variable, and the other properties are obvious.

In fact, we have

$$\begin{aligned} P(N_n(A) = k) &= \sum_{m \geq k} P(K_n = m, N_n(A) = k, N_n(A_n^c) = m-k) \\ &= \sum_{m \geq k} e^{-T\nu(Z_n)} \frac{(T\nu(Z_n))^m}{m!} \binom{m}{k} \overline{\lambda_n}(A)^k \overline{\lambda_n}(A_n^c)^{m-k} \\ &= e^{-\lambda_n(A)} \frac{(\lambda_n(A))^k}{k!}. \end{aligned}$$

(Note that $\lambda_n(A) + \lambda_n(A_n^c) = \lambda([0,T] \times Z_n) = T\nu(Z_n)$.)

On the other hand, the sum $N_1 + N_2$ of independent λ_i-Poisson variables N_i $(i = 1, 2)$ is again $(\lambda_1 + \lambda_2)$-Poisson variable by due to

$$P(N_1 + N_2 = n) = \sum_{k=0}^n P(N_1 = k, N_2 = n-k) = \sum_{k=0}^n e^{-\lambda_1} \frac{\lambda_1^k}{k!} e^{-\lambda_2} \frac{\lambda_2^{n-k}}{(n-k)!} = e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1 + \lambda_2)^n}{n!}.$$

(Proof of the above Proposition)

The point mass (τ_i, ξ_i) of N is the same as ${}^\exists n \geq 0, {}^\exists k \geq 1; (\tau_k^n, \xi_k^n)$. Hence, it is enough to show that for any $n, m \geq 0, k, j \geq 1; (n,k) \neq (m,j)$, $P(\tau_k^n \neq \tau_j^m) = 1$,i.e., $P(\tau_k^n = \tau_j^m) = 0$. We fix ${}^\forall M \geq 1$ and divide $[0,T]$ into M equal parts, i.e., set $Z_{n,\ell} = [(\ell-1)T/M, \ell T/M) \times Z_n$ $(1 \leq \ell < M)$ and $Z_{n,M} = [(M-1)T/M, T] \times Z_n$. Then, $\overline{\lambda_n}(Z_{n,\ell}) = 1/M$. Therefore, if $\tau_k^n = \tau_j^m$, then ${}^\exists \ell; \tau_k^n \in [(\ell-1)T/M, \ell T/M)$. Thus, $Y_k^n = (\tau_k^n, \xi_k^n) \in Z_{n,\ell}, Y_j^m = (\tau_j^m, \xi_j^m) \in Z_{m,\ell}$. By the independence of Y_k^n, Y_j^m, we have

$$P(\tau_k^n = \tau_j^m) \leq P\left(\bigcup_{\ell=1}^M \{Y_k^n \in Z_{n,\ell}, Y_j^m \in Z_{m,\ell}\}\right) \leq \sum_{\ell=1}^M \overline{\lambda_n}(Z_{n,\ell}) \overline{\lambda_m}(Z_{m,\ell}) = \frac{1}{M} \to 0.$$

■

3 Stochastic Integrals

3.1 Ito integrals by Wiener processes

In the following, *let*$(\Omega, \mathcal{F}, P)$ be a complete probability space and $(\mathcal{F}_t)_{t\geq 0}$ be a right-continuous filtration containing all null sets. Let $(B_t)_{t\geq 0}$ be an $(\mathcal{F}_t)$-Brownian motion (simply, we denote a BM); $B_0 = 0$ a.s. and $t \in [0, T]$ denote the time (we extend to $t \in [0, \infty)$ at the end).

We define $f = f(t) = f(t, \omega) \in \mathcal{L}^2_T \overset{\text{def}}{\iff} (f(t))_{t\geq 0}$ is $(\mathcal{F}_t)$-adapted on $[0, T] \times \Omega$ and L^2-integrable under $dtP(d\omega)$. We shall define a **Ito integral (continuous-type stochastic integral)** as

$$\int_0^t f(r)dB_r = \int_0^t f(r, \omega)dB_r(\omega).$$

Moreover, f is finally extended to the following:

$$f = f(t) = f(t, \omega) \in \mathcal{L}^2_{\text{loc}} \overset{\text{def}}{\iff} f : [0, \infty) \times \Omega \to \mathbf{R}^1 \text{ is measurable and } (f(t))_{t\geq 0} \text{ is } (\mathcal{F}_t)\text{-adapted},$$
$$\int_0^t f(r)^2 dr < \infty \text{ a.s. for } {}^\forall t > 0.$$

Stochastic integral can not be defined like Riemannian integral for each paths (i.e., for each ω), because a BM does not have a bounded variation. Hence, it can be defined under the measure $dtP(d\omega)$, for from right-continuous step functions to $f \in \mathcal{L}^2_T$ by using L^2-approximation.

Now we define the following:

· A **defining process** $f(t, \omega) = f_a(\omega)1_{(a,b]}(t); a < b; a, b \in [0, T]$, f_a is bounded and $\mathcal{F}_a$-measurable.

· A **step process** $f \in \mathcal{S}$ is a finite sum of defining processes with disjoint time intervals, i.e.,

$$f(t, \omega) = \sum_{k=1}^n f_{t_{k-1}}(\omega)1_{(t_{k-1},t_k]}(t),$$

we simply denote

$$f(t) = \sum_{k=1}^n f_{t_{k-1}}1_{(t_{k-1},t_k]}(t),$$

where $0 = t_0 \leq t_1 < \cdots < t_n = T$, $f_{t_{k-1}}$ is bounded and $\mathcal{F}_{t_{k-1}}$-measurable.

The norm $\|\cdot\|_T$ is defined as

$$\|f\|^2_T := \int_0^T Ef(t)^2 dt.$$

Proposition 3.1 *$\mathcal{S}$: dense in $\mathcal{L}^2_T$ under $\|\cdot\|_T$, i.e., ${}^\forall f \in \mathcal{L}^2_T, {}^\exists f_n \in \mathcal{S}; \|f - f_n\|_T \to 0$.*

In the proof of this, we use the result such that if a measurable stochastic process $(f(t))_{t\geq 0}$ is $(\mathcal{F}_t)$-adapted, then it has a progressively measurable version, where $(f(t))_{t\geq 0}$ is **progressively measurable** means for ${}^\forall t > 0$, $(s, \omega) \in ([0, t] \times \Omega, \mathcal{B}[0, t] \otimes \mathcal{F}_t) \mapsto f(s, \omega) \in (\mathbf{R}, \mathcal{B}^1)$ is measurable.

If we would not like to use this result, we assume ł'progressively measurable" instead of "$(\mathcal{F}_t)$-adapted" in the definition of $\mathcal{L}^2_T$.

Proof. For $f \in \mathcal{L}^2_T$, by $f1_{\{|f|\leq n\}}$ we may assume f is bounded. Furthermore, for ${}^\forall\varepsilon > 0$, by considering the following we may assume f is continuous; Let

$$f_\varepsilon(t, \omega) = \frac{1}{\varepsilon}\int_{((t-\varepsilon)\vee 0,t]} f(r, \omega)dr \quad \text{Then } f_\varepsilon \to f \text{ in } \mathcal{L}^2_T.$$

(Using the result that in general $g_t \in L^2([0, T])$ is L^2-continuous, i.e, $\int_{[0,T]} |g(t + \varepsilon) - g(t)|^2 dt \to 0$ as $\varepsilon \to 0$. This holds because $C_c \subset L^2$ is dense.) Here, note that f_ε is $(\mathcal{F}_t)$-adapted by the progressively measurability. Therefore, for bounded continuous $f \in \mathcal{L}^2_T$, set

$$f_n(t, \omega) = f(0, \omega)1_{\{0\}}(t) + \sum_{k=1}^n f(t_{k-1}, \omega)1_{(t_{k-1},t_k]}(t), \quad t_k = \frac{k}{n}T.$$

Then $f_n \in \mathcal{S}$ and $\|f - f_n\|_T \to 0$ hold. ∎

(Note) A right (or left) continuous $(\mathcal{F}_t)$-adapted process is progressively measurable. In fact, for a fixed $t > 0$, $f_n(r,\omega) = f(0,\omega)1_{\{0\}}(r) + \sum_{k=1}^n f(t_k,\omega)1_{(t_{k-1},t_k]}(r)$ ($t_k = \frac{k}{n}t$) are $\mathcal{B}^1([0,t]) \otimes \mathcal{F}_t$-measurable and converge to f a.s. by the right-continuity, thus the limit f is so too.

For a step process $f(t) = \sum_{k=1}^n f_{t_{k-1}} 1_{(t_{k-1},t_k]}(t)$, define

$$M_t(f) \equiv \int_0^t f(r)dB_r := \sum_{k=1}^n f_{t_{k-1}}(B_{t_k\wedge t} - B_{t_{k-1}\wedge t}).$$

Remark 3.1 *In the first edition of Funaki [2], the same definition is given for a right-continuous step process $f(t) = \sum_{k=1}^n f_{t_{k-1}} 1_{[t_{k-1},t_k)}(t)$. This give the same definition of stochastic integrals, by a BM (or continuous martingales). However, for the definition of stochastic integrals by Poisson random measures, we need left-continuous step processes. Moreover, we extend f to $(\mathcal{F}_t)$-predictable processes. Note that if $f(t,\omega)$ is measurable and $(\mathcal{F}_t)$-adapted, then it has $(\mathcal{F}_t)$-predictable version, so we may restrict to it here. In fact, $\limsup_{\varepsilon\downarrow 0} \int_{t-\varepsilon}^t f(r,\omega)dr$ is a predictable version.*

We have the following:

Proposition 3.2 *$\{M_t(f)\}$ is continuous and $EM_t(f) = 0, EM_t(f)^2 = \int_0^t Ef(r)^2dr$. If $s < t$, then $E[M_t(f)|\ \mathcal{F}_s] = M_s(f)$ a.s., i.e.,*

$$E\int_0^t f(r)dB_r = 0, E\left(\int_0^t f(r)dB_r\right)^2 = \int_0^t Ef(r)^2dr, E\left[\int_0^t f(r)dB_r \middle|\ \mathcal{F}_s\right] = \int_0^s f(r)dB_r \ \text{a.s.}$$

That is, let $\mathcal{M}_{c,0}^2$ be a family of all continuous L^2-(integrable) martingales, then $\{M_t(f)\} \in \mathcal{M}_{c,0}^2$ and

$$\langle M(f), M(g)\rangle_t = \int_0^t f(r)g(r)dr.$$

About $\langle M(f), M(g)\rangle_t$, if $(M_t), (N_t)$ are continuous L^2-martingale, then M_t^2 is continuous sub-martingale and is in class (DL). Hence, by Doob-Meyer decomposition, $^\exists A_t$ is a continuous increasing process; $A_0 = 0$, $M_t^2 - A_t$ is martingale. It is denoted as $A_t =: \langle M\rangle_t$ and called by a **quadratic variation process of** (M_t). moreover, let

$$\langle M, N\rangle_t := \frac{1}{4}(\langle M+N\rangle_t - \langle M-N\rangle_t) = \frac{1}{2}(\langle M+N\rangle_t - \langle M\rangle_t - \langle N\rangle_t)$$

This has bounded variation and $M_tN_t - \langle M, N\rangle_t$ is a continuous martingale. It called by **quadratic variation of** (M_t) **and** (N_t). It holds that for $\Delta; 0\ t_0 < t_1 < \cdots < t_n = t$, 次が成り立つ.

$$\langle M\rangle_t = \lim_{|\Delta|\to 0} \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})^2 \quad \text{(in prob.)}$$

For a BM, $\langle B\rangle_t = t$.

Proof. If f is a defining process $f(t) = f_a(\omega)1_{[a,b)}(t)$, then $M(f)_t = f_a(B_{t\wedge b} - B_{t\wedge a})$ is continuous and $M(f)_0 = 0$.

We show the martingale property; for $0 \le s < t$, $E[M(f)_t|\ \mathcal{F}_s] = M(f)_s$ a.s. Since

$$E[M(f)_t - M(f)_s|\ \mathcal{F}_s] = E[f_a(B_{t\wedge b} - B_{t\wedge a} - B_{s\wedge b} + B_{s\wedge a})|\ \mathcal{F}_s],$$

if $s \le a$, then $E[E[f_a(B_{t\wedge b} - B_{t\wedge a})|\ \mathcal{F}_a]|\ \mathcal{F}_s] = E[f_aE[(B_{t\wedge b} - B_{t\wedge a})|\ \mathcal{F}_a]|\ \mathcal{F}_s] = 0$, if $s > a$, then $E[f_a(B_{t\wedge b} - B_{s\wedge b})|\ \mathcal{F}_s] = f_aE[B_{t\wedge b} - B_{s\wedge b}|\ \mathcal{F}_s] = 0$. Hence, $E[M(f)_t - M(f)_s|\ \mathcal{F}_s] = 0$ a.s. On a quadratic variation, it is enough to show

$$E\left[M(f)_tM(g)_t - M(f)_sM(g)_s - \int_s^t f(r)g(r)dr \middle|\ \mathcal{F}_s\right] = 0$$

In fact, from this, we see $\langle M(f)\rangle_t = \int_0^t f(r)^2 dr, \langle M(f+g)\rangle_t = \int_0^t (f+g)^2(r)dr$, and thus, we have the result. Let $g(t) = g_c 1_{[c,d)}(t)$ and we may set $a \le c$. Then the above result can be shown by the same way as above according to $s \le c, s > c$. ■

If $f \in \mathcal{L}^2_T$, then by approximating of step processes f_n, $(M(f_n)_t)$ is a Cauchy sequence in $L^2(dP)$. By completeness of $L^2(dP)$ there exists a limit $M(f)_t$ and it is a continuous process $(M(f)_t)$ (see the end of this subsection). A stochastic integral $\int_0^t f(s)dB_s$ is defined by this $M(f)_t$.

It holds that the following:

Theorem 3.1 *For $f, g \in \mathcal{L}^2_T$,*

$$\left(M(f)_t = \int_0^t f(r)dB_r\right)_t \in \mathcal{M}^2_{0,c}, \quad \langle M(f), M(g)\rangle_t = \int_0^t f(r)g(r)dr.$$

Moreover, if $f \in \mathcal{L}^2_{\text{loc}}$, then set

$$\sigma_n = \inf\left\{t > 0; \int_0^t f(r)^2 dr > n\right\}$$

and $f_n(t) := f(t \wedge \sigma_n) \in \mathcal{L}^2_T$, we can define $M(f)_t = \lim M(f(\cdot \wedge \sigma_n))_t$. (Noting that ${}^\forall\omega, {}^\exists N = N(\omega) \ge 1; {}^\forall n \ge N, \sigma_n(\omega) = T$.) Then, it holds that $M(f)_{t\wedge\sigma_n} = M(f(\cdot \wedge \sigma_n))_t$. Moreover, $(M(f)_t) \in \mathcal{M}^{2,\text{loc}}_{0,c}$ is a continuous local L^2-martingale starting from 0.

Proof of "$(M(f)_t)$ is a continuous process":

In general, for a rcll martingale (M_t), let $|M|^*_T := \sup_{0\le t\le T}|M_t|$, then the **continuous time martingale ineqq.** $\||M|^*_T\|_p \le p/(p-1)\|M_T\|_p$ $(p > 1)$ holds. (In fact, by using discrete-time sub-martingale inequality (see Th.3.6, Cor. 3.2 in [3]), we have $aP(\sup_{t\in[0,T]\cap\mathbf{Q}}|M_t| \ge a) \le E|M_T|$, and by right-continuity it holds $aP(\sup_{t\le T}|M_t| \ge a) \le E|M_T|$. Moreover by using this and by the same way as in the case of discrete-time, we have the desired inequality.)

Furthermore, by using Borel-Cantelli's lemma, we can show that ${}^\exists(M_t)$: a continuous process, ${}^\exists\{m_k\}; |M(f_{m_k}) - M|^*_T \to 0$ a.s. $(k \to \infty)$. In fact, by $\||M(f_n) - M(f_m)|^*_T\|_2^2 == E[(|M(f_n) - M(f_m)|^*_T)^2] \to 0$ $(m, n \to \infty)$ ${}^\exists\{m_k\}; P(|M(f_{m_{k+1}}) - M(f_{m_k})|^*_T| \ge 1/2^k) \le 1/2^k$ and bu B-C's Lem., we have ${}^\forall\ell > k \ge 1, |M(f_{m_\ell}) - M(f_{m_k})|^*_T| \le \sum_{j=k}^{\ell-1} 1/2^j \le 1/2^{k-1} \to 0$ $(k \to \infty)$ a.s. That is, $M(f_{m_k})$ is a Cauchy seq. in $C([0,T])$ a.s. and by the completeness of this sp. ${}^\exists M = M(f)$: a continuous process; $|M - M(f_{m_k})|^*_T| \to 0$ a.s.

Moreover we can show

$$E[(|M(f_n) - M(f)|^*_T)^2] \le 4 \varliminf_{k\to\infty} \|M(f_n)_T - M(f_{m_k})_T\|_2^2 = 4 \varliminf_{k\to\infty} \|f_n - f_{m_k}\|_T^2 = 4\|f_n - f\|_T^2 \to 0$$

In fact, for an fixed ${}^\forall n \ge 1$ and the above subseq. $\{m_k\}$, $|M(f_n) - M(f)|^*_T \le \varliminf_{k\to\infty}|M(f_n) - M(f_{m_k})|^*_T$ Furthermore, by Fatou's Lem., $\|\varliminf_{k\to\infty}|M(f_n) - M(f_{m_k})|^*_T\|_2^2 \le \varliminf_{k\to\infty}\||M(f_n) - M(f_{m_k})|^*_T\|_2^2 \le 4\|M(f_n)_T - M(f_{m_k})_T\|_2^2 = \lim_{k\to\infty} 4\|f_n - f_{m_k}\|_T^2 = 4\|f_n - f\|_T^2$. Therefore we have $\|(|M(f_n) - M(f)|^*_T)\|_2^2 \le 4\|f_n - f\|_T^2 \to 0$ $(n \to \infty)$.

[Another Proof (unusing Borel-Cantelli's lemma)]

For each fixed $t \in [0,T]$, $\{M(f_n)_t\}$ is a Cauchy seq. in $L^2(dP)$. Hence there exists a L^2-limit $M_t = M(f)_t$. Let $D := [0,T] \cap \mathbf{Q}$. We can take a suitable subsequence $\{m_k\}$ by diagonal method such that "${}^\forall r \in D, M(f_{m_k})_r \to M_r$" a.s. In fact, let $D = \{r_j\}$ and we take $\{m_{1,k}\}; M_{r_1}^{m_{1,k}} \to M_{r_1}$ a.s. Next we take $\{m_{2,k}\} \subset \{m_{1,k}\}; M_{r_2}^{m_{2,k}} \to M_{r_2}$ a.s. and take $\{m_{j,k}\}$. Let $m_k = m_{k,k}$, then] $M(f_{m_k})$ converges for all $r = r_j$ with probability 1.

Moreover, let $M^n = M(f_n)$ and $|M|^*_D := \sup_{r\in D}|M_r|$. By Fatou's Lem.,

$$\||M^n - M|^*_D\|_2^2 \le \|\varliminf_{k\to\infty}|M^n - M^{m_k}|^*_D\|_2^2 \le \varliminf_{k\to\infty}\||M^n - M^{m_k}|^*_D\|_2^2 \le 4\|f_n - f\|_T^2 \to 0\ (n \to \infty).$$

Thus, ${}^\exists\{n_k\}$ such that $|M(f_{n_k}) - M|^*_D \to 0$ a.s. and $\{M_r\}_{r\in D}$ is continuous in D.

Hence for each $t \in D^c := [0,T] \cap \mathbf{Q}^c$, set $M_t := M_{t+}$, then $|M(f_{n_k}) - M|_T^* = |M(f_{n_k}) - M|_D^* \to 0$ a.s. (by this, $\{M_t = M(f)_t\}$ is continuous in $[0,T]$), and we have $\||M(f_n) - M|_T^*\|_2 = \||M(f_n) - M|_D^*\|_2 \to 0$ $(n \to \infty)$. ■

By the above proof for $M = (M_t) \in \mathcal{M}_{c,0}^2$ restricted in $t \in [0,T]$, if we define a norm $\|M\| \equiv \|M\|_T^* := \||M|_T^*\|_2 = (E[\sup_{t \le T} |M_t|^2])^{1/2}$, then $\mathcal{M}_{c,0}^2$ is complete.

In the next subsection, in case of stochastic integrals by a compensated Poisson random measure, the limit is rcll because an approximate sequence is so.

Question In the above proof, show the existence of M_{t+} for $t \in D^c := [0,T] \cap \mathbf{Q}^c$. (We have the same result in the case that $M^n = M(f_n)$ is rcll.)

[Ans.] It is enough to show the case that M^n is rcll and $|M - M^n|_D^* \to 0$ $(n \to \infty)$ holds. Set $D_n^+(t) := \{r \in D; t < r \le t + 1/2^n\}$ and let

$$\underline{M_{t+}} := \lim_{n\to\infty} \inf_{r \in D_n^+(t)} M_r, \quad \overline{M_{t+}} := lim_{n\to\infty} \sup_{r \in D_n^+(t)} M_r.$$

We may show $\underline{M_{t+}} = \overline{M_{t+}}$. However, this is easily obtained by using $M^n \rightrightarrows M$ on D and M_n is r-c.

3.2 Stochastic integrals by Poisson random measures

Let $\nu(dz)$ be a Lévy measure on $\mathbf{R}^m$, i.e.,

$$\nu(\{0\}) = 0, {}^\forall n \ge 1, \nu(|z| \ge 1/n) < \infty$$

Let $N(dt, dz)$ be the $dtd\nu$-Poisson random measure, that is, it is a Poisson random measure on $[0,\infty) \times \mathbf{R}^m$ with a mean measure $\widehat{N}(dt, dz) := E[N(dt, dz)] = dt\nu(dz)$ For a given filtration $(\mathcal{F}_t)$ and for any $0 \le s < t, U \in \mathcal{B}^m$, $N((s,t] \times U)$ is independent of $\mathcal{F}_s$. Moreover, let $\widetilde{N} := N - \widehat{N}$ be the **compensated Poisson random measure)**.

Let $\{(\tau_k, \xi_k)\}$ be the point masses of $N(dt, dz)$, then

$$N(dt, dz) = \sum \delta_{(\tau_k, \xi_k)}(dt, dz).$$

$\{\tau_k\}$ take a.s. different values, and the number of k such that $|\xi_k| \ge 1/n$ is finite.

Let $f(t,z) = f(t,z,\omega)$, $g(t,z) = g(t,z,\omega)$ be $\mathbf{R}^d$-valued and $(\mathcal{F}_t)$-predictable in $(t,z,\omega) \in [0,\infty) \times \mathbf{R}^m \times \Omega$, i.e., let $\mathcal{P}$ be the smallest σ-field such that the following functions $h(t,z,\omega)$ satisfying (1), (2) are all measurable. Let f, g be $\mathcal{P}$-measurable. Then f, g are called **$(\mathcal{F}_t)$-predictable**.

(1) ${}^\forall(z,\omega)$, $t \mapsto h(t,z,\omega)$ is left continuous. (2) ${}^\forall t > 0$, $(z,\omega) \mapsto h(t,z,\omega)$ is $\mathcal{B}^m \otimes \mathcal{F}_t$-measurable.

(Remark) The stochastic integral by a Poisson random measure can be defined by using the independence of $N((s,t] \times U)$ and $\mathcal{F}_s$, hence left-continuous step functions are basic. Therefore, they are extended to the predicatable functions.

Moreover, let f, g satisfy the following integral conditions;

$${}^\forall t > 0, \int_0^t dr \int_{\mathbf{R}^m} |f(r,z)| \nu(dz) < \infty \quad \text{a.s.},$$

$${}^\forall t > 0, \int_0^t dr \int_{\mathbf{R}^m} |g(r,z)|^2 \nu(dz) < \infty \quad \text{a.s.}$$

where $\int_0^t = \int_{0+}^{t+} = \int_{(0,t]}$. By a similar way to the continuous case, let

$$\eta_n := \inf\left\{t > 0; \int_0^t dr \int_{\mathbf{R}^m} |f(r,z)| \nu(dz) > n\right\}, \ \sigma_n := \inf\left\{t > 0; \int_0^t dr \int_{\mathbf{R}^m} |g(r,z)|^2 \nu(dz) > n\right\}$$

and for any fixed $t > 0$, by changing $0 < r \le t$ to $r \wedge \eta_n$, $r \wedge \sigma_n$ we may assume the followings:

$$^\forall t > 0, \quad \int_0^t dr \int_{\mathbf{R}^m} E|f(r,z)|\nu(dz) < \infty.$$

$$^\forall t > 0, \quad \int_0^t dr \int_{\mathbf{R}^m} E|g(r,z)|^2 \nu(dz) < \infty.$$

First, define

$$X_t = \int_0^t \int_{\mathbf{R}^m} f(r,z)N(dr,dz) := \sum_{k;\tau_k \le t} f(\tau_k, \xi_k).$$

Then, for $0 \le s < t$,

$$E[X_t |\ \mathcal{F}_s] = X_s + \int_s^t \int_{\mathbf{R}^m} E[f(r,z)|\ \mathcal{F}_s] dr\nu(dz),$$

and

$$E|X_t| \le \int_0^t dr \int_{\mathbf{R}^m} E|f(r,z)|\nu(dz) < \infty.$$

If g satisfies

$$^\forall t > 0, \int_0^t dr \int_{\mathbf{R}^m} E|g(r,z)|^2 \nu(dz) < \infty,$$

then define

$$Y_t = \int_0^t \int_{\mathbf{R}^m} g(r,z)\widetilde{N}(dr,dz) := L^2\text{-} \lim_{n\to\infty} \int_0^t \int_{|z|\ge 1/n} g(r,z)\widetilde{N}(dr,dz).$$

In this case, Y_t is rcll L^2-martingale with mean 0, i.e.,

$$E\left[\int_0^t \int_{\mathbf{R}^m} g(r,z)\widetilde{N}(dr,dz) \,\middle|\, \mathcal{F}_s \right] = \int_0^s \int_{\mathbf{R}^m} g(r,z)\widetilde{N}(dr,dz),$$

and

$$E\left[\left(\int_0^t \int_{\mathbf{R}^m} g(r,z)\widetilde{N}(dr,dz) \right)^2 \right] = \int_0^t dr \int_{\mathbf{R}^m} Eg(r,z)^2 \nu(dz).$$

(Note that this is rcll by the same way as in the case of Ito integrals by Brownian motions because of an approximate sequence is so.)

For general g, define $Y_t = \int_0^t \int_{\mathbf{R}^m} g(r,z)\widetilde{N}(dr,dz) = \lim_{n\to\infty} \int_0^t \int_{\mathbf{R}^m} g(r\wedge\sigma_n, z)\widetilde{N}(dr,dz)$.. Then this is rcll local L^2-martingale, i.e., $Y_{t\wedge\sigma_n}$ is rcll L^2-martingale and satisfies the above properties.

For simplicity of the proof, let $m = 1$, and for $T > 0$, by restricting the time interval to $[0,T]$, it is enough to consider the following left-continuous step function $f(r,z)$. Hence the first half is easy.

$$f(r,z) = \sum_{k=1}^{2^n} f(r_{k-1}^n) 1_{(r_{k-1}^n, r_k^n]}(r) 1_U(z)$$

($r_k^n = kT/2^n$, $f(t) = f(\omega;t)$ is $(\mathcal{F}_t)$-adapted and $U \in \mathcal{B}^1; \nu(U) < \infty$.)

Proposition 3.3 *Let* $\mathbf{F}$ *be a linear sub-space of a space of all bounded measurable real-valued functions* $f(t,z,\omega)$*. If it satisfies the following two conditions, then it contains all bounded predictable functions.*

(1) $\mathbf{F}$ *contains all bounded functions* $f(t,z,\omega)$ *such that left-continuous in* $t \ge 0$*, and* $\mathcal{B}^m \otimes \mathcal{F}_t$*-measurable in* (z,ω)*.*

(2) $f_n \in \mathbf{F}; \uparrow f \implies f \in \mathbf{F}$

(We give the proof at the end of this section.)

It is the same for $g(r,z)$. We show the martingale property. It is enough to show

$$E\left[\int_s^t \int_{\mathbf{R}} g(r,z)\widetilde{N}(dr,dz)\middle|\ \mathcal{F}_s\right] = 0 \text{ a.s.}$$

We divide $(s,t]$ into 2^n parts and denote the points as $\{r_k^n\}$, and approximate g by step functions. In the way, we restrict to $Z_n = \{|z| \geq 1/n\}$ and let $U \in \mathcal{Z}_n = \mathcal{B}^m \cap Z_n$. The above is easily obtained by the following holds. Note that $s \leq r_{k-1}^n < r_k^n \leq t$.

$$\begin{aligned}
&E\left[\int_s^t \int_{\mathbf{R}} g(r_{k-1}^n)1_{(r_{k-1}^n,r_k^n]}(r)1_U(z)\widetilde{N}(dr,dz)\ \middle|\mathcal{F}_s\right]\\
&\quad = E\left[E\left[g(r_{k-1}^n)\widetilde{N}((r_{k-1}^n,r_k^n]\times U)\ \middle|\mathcal{F}_{r_{k-1}}\right]\ \middle|\mathcal{F}_s\right]\\
&\quad = E\left[g(r_{k-1}^n)E\left[\widetilde{N}((r_{k-1}^n,r_k^n]\times U)\right]\ \middle|\mathcal{F}_s\right] = 0 \text{ a.s.}
\end{aligned}$$

the last two equals comes from the following:

$$E\left[\widetilde{N}((r_{k-1}^n,r_k^n]\times U)\ \middle|\mathcal{F}_{r_{k-1}}\right] = E\left[\widetilde{N}((r_{k-1}^n,r_k^n]\times U)\right] = 0$$

Let $g_n(r,z) = g(r,z)1_{|z|\geq 1/n}$ and by using L^2-approximating, it is easy to show the above desired equation. In fact, it holds that

$$E\left[\int_s^t \int_{\mathbf{R}} g_n(r,z)\widetilde{N}(dr,dz)\ \middle|\mathcal{F}_s\right] = 0 \text{ a.s.}$$

and in order to show this for g, it is enough to show that for $^\forall A \in \mathcal{F}_s$,

$$E\left[\int_s^t \int_{\mathbf{R}} g(r,z)\widetilde{N}(dr,dz); A\right] = 0.$$

However, this is obtained by the following:

$$\begin{aligned}
&\left(E\left[\int_s^t \int_{\mathbf{R}} g(r,z)\widetilde{N}(dr,dz); A\right] - E\left[\int_s^t \int_{\mathbf{R}} g_n(r,z)\widetilde{N}(dr,dz); A\right]\right)^2\\
&\qquad = \int_s^t dr \int_{\mathbf{R}} E[(g(r,z)-g_n(r,z))^2 1_A]\nu(dz) \to 0.
\end{aligned}$$

Moreover, the square moment is obtained by the following: For the step function, each terms are for $j \leq k$,

$$g(r_{k-1}^n)\widetilde{N}((r_{k-1}^n,r_k^n]\times U)g(r_{j-1}^n)\widetilde{N}((r_{j-1}^n,r_j^n]\times U),$$

and if $j < k$, i.e., $j \leq k-1$, then $g(r_{j-1}^n)\widetilde{N}((r_{j-1}^n,r_j^n]\times U)g(r_{k-1}^n)$ is independent of $\widetilde{N}((r_{k-1}^n,r_k^n]\times U)$, and the expectation of the last is 0. If $j = k$, then

$$E[g(r_{k-1}^n)^2\widetilde{N}((r_{k-1}^n,r_k^n]\times U)^2] = E[g(r_{k-1}^n)^2]E[\widetilde{N}((r_{k-1}^n,r_k^n]\times U)^2]$$

and

$$E[\widetilde{N}((r_{k-1}^n,r_k^n]\times U)^2] = EN((r_{k-1}^n,r_k^n]\times U) = (r_k^n - r_{k-1}^n)\nu(U).$$

Hence, the desired equation is obtained by using L^2 approximation. ■

[Proof of Proposition 3.3] For simplicity, we omit $z \in \mathbf{R}^m$. (e.g. in $\mathbf{F}$, we change $\mathcal{B}^m \otimes \mathcal{F}_t$ to $\mathcal{F}_t$.)

$\mathcal{D} \subset 2^{[0,\infty)\times\Omega}$ is **Dynkin family (d-system)** $\overset{\text{def}}{\Longleftrightarrow}$

(i) $[0,\infty)\times\Omega \in \mathcal{D}$.
(ii) $A, B \in \mathcal{D}; A \subset B \Rightarrow B \setminus A \in \mathcal{D}$.
(iii) $A_n \in \mathcal{D}\uparrow \Rightarrow \bigcup A_n \in \mathcal{D}$.

For any $\mathcal{C} \subset 2^{[0,\infty)\times\Omega}$, there exists the smallest d-system $d(\mathcal{C})$ containing $\mathcal{C}$.

Then, the following holds.

Lemma 3.1 *If $\mathcal{C} \subset 2^{[0,\infty)\times\Omega}$ is closed under intersection of finite number of elements, then $d(\mathcal{C}) = \sigma(\mathcal{C})$.*

The proof is easy (the next question).

We show Proposition 3.3 by using this. A bounded non-negative predictable function is approximated by increasing sequence of non-negative predictable simple functions. Hence, it is enough to show $^\forall A \in \mathcal{P}, 1_A \in \mathbf{F}$. Thus, define $A \in \mathcal{P}' \overset{\text{def}}{\iff} 1_A \in \mathbf{F}$, then this is a Dynkin family. For $k \le n$, let (Y_t^k) be a left continuous $(\mathcal{F}_t)$-adapted process. Let $\mathcal{C}$ be the family of all $\bigcap_{k\le n}\{Y_t^k \in B_k\}$ $(B_k \in \mathcal{B}^1)$. Then, it can be seen $\mathcal{C} \subset \mathcal{P}'$, and hence, $d(\mathcal{C}) \subset \mathcal{P}'$. By the above lemma, $d(\mathcal{C}) = \sigma(\mathcal{C}) = \mathcal{P}$, and we have $\mathcal{P} \subset \mathcal{P}'$. On $\mathcal{C} \subset \mathcal{P}'$, it is enough to show the case of $n = 1$, i.e., $A_t = \{Y_t \in B\} \in \mathcal{P}'$. since $^\exists\varphi_n$ is non-negative continuous on $\mathbf{R}^1$ such that $0 \le \varphi_n \uparrow 1_B$, we have $1_{A_t} = 1_B(Y_t) = \lim \varphi_n(Y_t) \in \mathbf{F}$. ∎

Question 3.1 *Show the above lemma. The proof is a similar to the case of Monotone Class Theorem.*

Since a σ-additive class is a Dynkin family, $d(\mathcal{C}) \subset \sigma(\mathcal{C})$ is clear. It is enough to show that $d(\mathcal{C})$ is a σ-additive class. Moreover, it reduce to show that $A, B \in d(\mathcal{C}) \Rightarrow A \cap B \in d(\mathcal{C})$. For a fixed $A \in d(\mathcal{C})$, set

$$\mathcal{D}_A = \{B \subset [0,\infty) \times \Omega; A \cap B \in d(\mathcal{C})\}.$$

Then it can be easily seen that if $A \in d(\mathcal{C})$, then $\mathcal{D}_A$ is a Dynkin family. Hence, $d(\mathcal{C}) \subset \mathcal{D}_A$. This implies the above result. In fact, if $B_1, B_2 \in \mathcal{D}; B_1 \subset B_2$, then $A \cap (B_2 \setminus B_1) = (A \cap B_2) \setminus (A \cap B_1) \in d(\mathcal{C})$. If $B_n \in \mathcal{D}_A; \uparrow$, then $A \cap (\bigcup B_n) = \bigcup (A \cap B_n) \in d(\mathcal{C})$. Therefore, by the assumption of$\mathcal{C}$ and the above results, if $A \in \mathcal{C}$, then $\mathcal{D}_A \supset d(\mathcal{C})$, that is, if $B \in d(\mathcal{C})$, then $B \in \mathcal{D}_A$, i.e., $A \cap B \in d(\mathcal{C})$. and hence, by exchanging A and B, we have if $A \in d(\mathcal{C})$, then $\mathcal{D}_A \supset d(\mathcal{C})$, ∎

3.3 Ito formula (continuous type)

Let (X_t) be an $\mathbf{R}^d$-valued stochastic process defined by the following stochastic integrals;

$$X_t(\omega) = x + \int_0^t a(r,\omega)dr + \int_0^t b(r,\omega)dB_r(\omega).$$

if we denote the components, $X_t = (X_t^i) = (X_t^1, \ldots, X_t^d)$, then

$$X_t^i = x^i + \int_0^t a^i(r)dr + \int_0^t \sum_{j\le N} b_k^i(r)dB_r^k,$$

where $B_t = (B_t^k)_{k\le N}$: an N-dimensional Brownian motion.

- $a(t) = a(t,\omega) = (a^i(t,\omega))_{i\le d}$: $(\mathcal{F}_t)$-adapted; $^\forall T > 0, \displaystyle\int_0^T |a(t)|dt < \infty$ a.s.
- $b(t) = b(t,\omega) = (b_k^i(t,\omega))_{i\le d, k\le N}$: $(\mathcal{F}_t)$-adapted; $^\forall T > 0, \displaystyle\int_0^T \|b(t)\|^2 dt < \infty$ a.s., i.e, $\{b(t)\}_{t\ge 0} \in \mathcal{L}^2_{0,\mathrm{loc}}$. Note that

$$b(t)dB_t = \sum_{k\le N} b_k^i(t)dB_t^k, \quad \|b(t)\|^2 = \sum_{i,k} (b_k^i)^2(t).$$

This is simply denoted as follows: (it is formal, however, it is convenient for calculus,)

$$dX_t = a(t)dt + b(t)dB_t.$$

Then, **Ito formula** is given as follows: for $\varphi(x) \in C^2(\mathbf{R}^d)$,

$$d\varphi(X_t) = a(t) \cdot D\varphi(X_t)dt + b(t) \cdot D\varphi(X_t)dB_t + \frac{1}{2}b^2(t) \cdot D^2\varphi(X_t)dt.$$

where $a(t) \cdot D = a^i(t)\partial_i$, $b^2(t) \cdot D^2 = \sum_{k\le m} b_k^i(t)b_k^j \partial_{ij}^2$ (moreover, we sum on the same character of superscripts and subscripts). $\partial_i = \partial/\partial x_i, \partial_{ij}^2 = \partial^2/\partial x_i \partial x_j$.

Let $d = 1, N = 1$.

$$d\varphi(X_t) = \varphi'(X_t)a(t)dt + \varphi'(X_t)b(t)dB_t + \frac{1}{2}\varphi''(X_t)b(t)^2dt.$$

If φ has a Taylor's expansion, then by using the following:

$$(dB_t)^2 = dt, (dt)^2 = dtdB_t = 0, \text{ i.e., } (dX_t)^2 = b(t)^2dt \text{ and } {}^\forall n \geq 3, (dX_t)^n = 0$$

formally, we can get the result as follows;

$$\begin{aligned} d\varphi(X_t) &= \varphi'(X_t)dX_t + \frac{1}{2}\varphi''(X_t)(dX_t)^2 \\ &= \varphi'(X_t)\{a(t)dt + b(t)dB_t\} + \frac{1}{2}\varphi''(X_t)b(t)^2dt. \end{aligned}$$

Furthermore, if $\varphi(t,x) \in C_b^{1,2}([0,\infty) \times \mathbf{R}^d)$, then we have $\varphi(0, X_0) = \varphi(0, x)$,

$$d\varphi(t, X_t) = \partial_t\varphi(t, X_t)dt + a(t) \cdot D\varphi(t, X_t)dt + b(t) \cdot D\varphi(t, X_t)dB_t + \frac{1}{2}b^2(t) \cdot D^2\varphi(t, X_t)dt.$$

Proof. We only show the case of $d = N = 1$. It is enough to show the case $a(r), b(r)$ are step processes. If we consider an infinitesimal interval $[s,t]$ which is contained in a small time interval of $a(r), b(r)$, then we may assume both are constants $a(s), b(s) \in \mathcal{F}_s$ for the time $r \in [s,t]$ and bounded a.s. Here, we divide $[s,t]$ into equal n parts, however, it is essential the same as $s = 0$, so we may let $a_0, b_0 \in \mathcal{F}_0$ and bounded a.s., and divide $[0,t]$ into equal n parts; $t_k = t_k^n = tk/n, k = 0, 1, \ldots, n$. By Taylor's Theorem, we have

$$\varphi(X_{t_k}) - \varphi(X_{t_{k-1}}) = \varphi'(X_{t_{k-1}})(X_{t_k} - X_{t_{k-1}}) + \frac{1}{2}\varphi''(Y_k)(X_{t_k} - X_{t_{k-1}})^2,$$

where ${}^\exists\theta = \theta(\omega) \in (0,1)$; $Y_k = X_{t_{k-1}} + \theta(X_{t_k} - X_{t_{k-1}})$.

$$X_{t_k} - X_{t_{k-1}} = a_0(t_k - t_{k-1}) + b_0(B_{t_k} - B_{t_{k-1}}).$$

Hence, it is enough to show the following:

$$\begin{aligned} &X_{t_k} - X_{t_{k-1}} \to a(t)dt + b(t)dB_t \\ &(X_{t_k} - X_{t_{k-1}})^2 \to (a(t)dt + b(t)dB_t)^2 = b(t)^2dt \end{aligned}$$

as $n \to \infty$ in the sense of a.s. or in L^2. In order to it, we divide

$$\varphi(X_t) - \varphi(X_0) = \sum_{k=1}^n (\varphi(X_{t_k}) - \varphi(X_{t_{k-1}})) = \sum_{j=1}^5 I_n^j$$

and assume $\varphi \in C_b^2$ (i.e., φ'' is bounded). Then we can get the following as $n \to \infty$.

$$\begin{aligned} I_n^1 &= \sum_{k=1}^n \varphi'(X_{t_{k-1}})a_0(t_k - t_{k-1}) \to \int_0^t \varphi'(X_r)a(r)dr \quad \text{a.s.}, \\ I_n^2 &= \sum_{k=1}^n \varphi'(X_{t_{k-1}})b_0(B_{t_k} - B_{t_{k-1}}) \to \int_0^t \varphi'(X_r)b(r)dB_r \quad \text{in } L^2, \\ I_n^3 &= \sum_{k=1}^n \frac{1}{2}\varphi''(Y_k)b_0^2(B_{t_k} - B_{t_{k-1}})^2 \to \frac{1}{2}\int_0^t \varphi''(X_r)b^2(r)dr \quad \text{in } L^2, \\ I_n^4 &= \sum_{k=1}^n \frac{1}{2}\varphi''(Y_k)a_0b_0(t_k - t_{k-1})(B_{t_k} - B_{t_{k-1}}) \to 0 \quad \text{a.s.}, \\ I_n^5 &= \sum_{k=1}^n \frac{1}{2}\varphi''(Y_k)a_0^2(t_k - t_{k-1})^2 \to 0 \quad \text{a.s.} \end{aligned}$$

Therefore, by taking a suitable sub-sequence, all convergence holds in the sense of a.s. Hence, we get the desired result.

In fact, for I_n^4, by the continuity of B_r, $\max|B_{t_k} - B_{t_{k-1}}| \to 0$ a.s. On $I_n^5 \to 0$, by $t_k - t_{k-1} = 1/n \to 0$, it is clear. We consider on I_n^1, I_n^2, I_n^3. On I_n^1, since $\varphi'(X_r)$ is continuous in r, we have

$$I_n^1 \to a_0 \int_0^t \varphi'(X_r)dr = \int_0^t \varphi'(X_r)a(r)dr \quad \text{p.w.}$$

On I_n^2, by a step process $f_n(r) = \sum_{k=1}^n \varphi'(X_{t_{k-1}})1_{(t_{k-1},t_k]}(r)$, $I_n^2 = b_0 \int_0^t f_n(r)dB_r$, and since φ' is bounded and $\varphi'(X_r)$ is continuous, we have $\|f_n(r) - \varphi'(X_r)1_{(0,t]}(r)\|_T \to 0$. Thus,

$$I_n^2 = b_0 \int_0^t f_n(r)dB_r \to b_0 \int_0^t \varphi'(X_r)dB_r = \int_0^t \varphi'(X_r)b(r)dB_r \quad \text{in } L^2.$$

On I_n^3, let $\psi = \varphi'' \in C_b(\mathbf{R})$, and it is enough to show

$$\sum_{k=1}^n \psi(Y_k)(B_{t_k} - B_{t_{k-1}})^2 \to \int_0^t \psi(X_r)dr \quad \text{in } L^2.$$

By the continuities of X_r and ψ, we have $\max_{1\le k\le n} E|\psi(Y_k) - \psi(X_{t_{k-1}})|^2 \to 0$ and by the boundedness of ψ and by the following it is obvious.

$$E\left(\sum_{k=1}^n \{(B_{t_k} - B_{t_{k-1}})^2 - (t_k - t_{k-1})\}\right)^2 = \sum_{k=1}^n E\{(B_{t_k} - B_{t_{k-1}})^2 - (t_k - t_{k-1})\}^2 = \frac{t^2}{n} \to 0$$

At last, if $\varphi \in C^2$, then by letting $\sigma_n = \inf\{t > 0; |X_t| \ge n\}$ and by considering $X_{t\wedge\sigma_n}$, since φ', φ'' are both bonded on $\{|x| \le n\}$, Ito formula holds, and hence, by letting $n \to \infty$, we have the result. ■

3.4 Ito formula 2 (jump type)

Let $\nu(dz)$ be the Lévy measure on $\mathbf{R}^m$, i.e.,

$$\nu(\{0\}) = 0, {}^\forall n \ge 1, \nu(|z| \ge 1/n) < \infty$$

and let $N(dt, dz)$ be the $dtd\nu$-Poisson random measure, i.e., the Poisson random measure on $[0,\infty) \times \mathbf{R}^m$ with the mean measure $\widehat{N}(dt, dz) := E[N(dt, dz)] = dt\nu(dz)$. Moreover, set $\widetilde{N} := N - \widehat{N}$ as the compensated Poisson random measure.

Let $f(t, z) = f(t, z, \omega)$, $g(t, z) = g(t, z, \omega)$ be $\mathbf{R}^d$-valued $(\mathcal{F}_t)$-predictable functions satisfying the following integral conditions;

$${}^\forall t > 0, \quad \int_0^t dr \int_{|z|\ge 1} |f(r, z)|\nu(dz) < \infty \quad \text{a.s.},$$

$${}^\forall t > 0, \int_0^t dr \int_{|z|<1} |g(r, z)|^2\nu(dz) < \infty \quad \text{a.s.}$$

Let (X_t) be the stochastic process starting at x defined as follows;

$$dX_t(\omega) = a(t, \omega)dt + b(t, \omega)dB_t(\omega) + \int_{|z|\ge 1} f(t, z, \omega)N(\omega; dt, dz) + \int_{|z|<1} g(t, z, \omega)\widetilde{N}(\omega; dt, dz),$$

where a, b are the same as in the previous section.

Then **Ito formula (jump-type)** is given as the following: For $\varphi(x) \in C^2(\mathbf{R}^d)$,

$$\begin{aligned} d\varphi(X_t) &= a(t)\cdot D\varphi(X_t)dt + b(t)\cdot D\varphi(X_t)dB_t + \frac{1}{2}b^2(t)\cdot D^2\varphi(X_t)dt \\ &+ \int_{|z|\geq 1}[\varphi(X_{t-}+f(t,z)) - \varphi(X_{t-})]N(dt,dz) \\ &+ \int_{|z|<1}[\varphi(X_{t-}+g(t,z)) - \varphi(X_{t-})]\widetilde{N}(dt,dz) \\ &+ \int_{|z|<1}[\varphi(X_{t-}+g(t,z)) - \varphi(X_{t-}) - g(t,z)\cdot D\varphi(X_{t-})]\nu(dz)dt \end{aligned}$$

Remark: In Ikeda-Watanabe [5], they assume $fg = 0$ and f, g are integrated on $z \in \mathbf{R}$.

If $\varphi(t,x) \in C_b^{1,2}([0,\infty)\times\mathbf{R}^d)$, then add the term $\partial_t\varphi(t,X_t)dt$.

Especially, in the pure-jump type, i.e., $a = 0, b = 0$, we have

$$\begin{aligned} dX_t &= \int_{|z|\geq 1} f(t,z)N(dt,dz) + \int_{|z|<1} g(t,z)\widetilde{N}(dt,dz), \\ d\varphi(X_t) &= \int_{|z|\geq 1}[\varphi(X_{t-}+f(t,z)) - \varphi(X_{t-})]N(dt,dz) \\ &+ \int_{|z|<1}[\varphi(X_{t-}+g(t,z)) - \varphi(X_{t-})]\widetilde{N}(dt,dz) \\ &+ \int_{|z|<1}[\varphi(X_{t-}+g(t,z)) - \varphi(X_{t-}) - g(t,z)\cdot D\varphi'(X_{t-})]\nu(dz)dt. \end{aligned}$$

For simplicity of the proof, let $d = m = 1$ and $a = 0, b = 0$. Moreover, it is enough to show the case that coefficients satisfies the following.

$$\int_0^t dr\int_{|z|\geq 1} E|f(r,z)|\nu(dz) < \infty, \quad \int_0^t dr\int_{|z|<1} E|g(r,z)|^2\nu(dz) < \infty.$$

First, if $g = 0$, then X_t is a pure jump process and $\varphi(X_t)$ is changed by only jumps of $f(r,z)$, and hence,

$$\begin{aligned} \varphi(X_t) - \varphi(X_0) &= \sum_{r\leq t;\Delta X_r\neq 0}(\varphi(X_r) - \varphi(X_{r-})) = \sum_{r\leq t;\Delta X_r\neq 0}(\varphi(X_{r-}+f(r,z)) - \varphi(X_{r-})) \\ &= \int_0^t\int_{|z|\geq 1}(\varphi(X_{r-}+f(r,z)) - \varphi(X_{r-}))N(dr,dz). \end{aligned}$$

In case of $g \neq 0$, for simplicity, let $f = 0$, (however if $f \neq 0$, then it is enough to add the above term). Let

$$dX_t^n := \int_{1/n\leq|z|<1} g(t,z)\widetilde{N}(dt,dz).$$

Then it is also expressed as

$$dX_t^n := \int_{1/n\leq|z|<1} g(t,z)N(dt,dz) - \int_{1/n\leq|z|<1} g(t,z)dt\nu(dz)$$

We denote as $dX_t^n = d(X_t^n)^d + d(X_t^n)^c = \Delta X_t^n + d(X_t^n)^c$ with $\Delta X_t^n = X_t^n - X_{t-}^n$, i.e, $X_t^n = X_{t-}^n + \Delta X_t^n$. If $N(\{(t,z)\}) = 1$, then $\Delta X_t^n = g(t,z)$ and $\Delta\varphi(X_t^n) = \varphi(X_t^n) - \varphi(X_{t-}^n) = \varphi(X_{t-}^n + g(t,z)) - \varphi(X_{t-}^n)$.

We set $Z_n = \{1/n \le |z| < 1\}$.

$$\begin{aligned}
\varphi(X_t^n) - \varphi(X_s^n) &= \int_s^t \varphi'(X_r^n) d(X_r^n)^c + \sum_{r;\Delta X_r^n \neq 0} \Delta\varphi(X_r^n) \\
&= -\int_s^t \int_{Z_n} \varphi'(X_r^n) g(r,z) dr \nu(dz) + \int_s^t \int_{Z_n} [\varphi(X_{r-}^n + g(r,z)) - \varphi(X_{r-}^n)] N(dt,dz) \\
&= -\int_s^t \int_{Z_n} \varphi'(X_r^n) g(r,z) dr \nu(dz) + \int_s^t \int_{Z_n} [\varphi(X_{r-}^n + g(r,z)) - \varphi(X_{r-}^n)] \widetilde{N}(dr,dz) \\
&\quad + \int_s^t \int_{Z_n} [\varphi(X_{r-}^n + g(r,z)) - \varphi(X_{r-}^n)] dr \nu(dz) \\
&= \int_s^t \int_{Z_n} [\varphi(X_{r-}^n + g(r,z)) - \varphi(X_{r-}^n)] \widetilde{N}(dr,dz) \\
&+ \int_s^t \int_{Z_n} [\varphi(X_{r-}^n + g(r,z)) - \varphi(X_{r-}^n) - \varphi'(X_r^n) g(r,z)] dr \nu(dz).
\end{aligned}$$

Thus, if $n \to \infty$, then we can get the desired equation as a.s.-limits of a suitable subsequence. In fact, since X^n is a martingale, by using martingale inequality, we have $\sup_{t \le T} |X_t - X_t^n| \to 0$ in L^2. Hence, for a suitable subsequence, it convereges a.s. The above 1st term convereges in l^2 and the 2nd term convereges a.s. Therefore, by taking a sutable subesequence, both converge a.s. ■

4 Stochastic Differential Equations

For simplicity, we consider the one-dimensional case; $d = m = 1$.

Let $a(t,x), b(t,x), f(t,x,z), g(t,x,z); g(t,x,0) = 0$ be $\mathbf{R}^1$-valued functions of (time, state-space, jump-parameter) $(t, x.z) \in [0.T] \times \mathbf{R}^1 \times \mathbf{R}^1$

A stochastic differential equation is given by the following: the solution X_t is a $\mathbf{R}^1$-valued stochastic process such that $X_0 = x$.

$$dX_t = a(t, X_t)dt + b(t, X_t)dB_t + \int_{|z|\geq 1} f(t, X_{t-}, z)N(dt, dz) + \int_{|z|<1} g(t, X_{t-}, z)\widetilde{N}(dt, dz).$$

$X_0 = x$ is called an **initial condition** and a, b, f, g are called **coefficients** (by adding the Lévy measure ν, a, b, f, g, ν are also called coefficients).

Each of terms shows instantaneous hourly variation rate, and a is rate according to time, b is to a Brownian motion, f is to large jumps and g is to small jumps (note that the number of large jumps in a finite time interval is finite a.s.)

If a SDE without f, g, i.e, $f = g = 0$, then it is called a **continuous type)**, otherwise it is called a **jump type**. Moreover, if $a = b = 0$, then it is called **pure jump type**.

The SDE is actually defined by stochastic integrals.

$$\begin{aligned} X_t = x + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dB_s + \int_0^t \int_{|z|\geq 1} f(s, X_{s-}, z)N(ds, dz) \\ + \int_0^t \int_{|z|<1} g(s, X_{s-}, z)\widetilde{N}(ds, dz). \end{aligned}$$

4.1 Continuous-type SDE

Let $d \geq 1, N \geq 1$. Let $B_t = (B_t^i)_{i\leq N}$ be an N-dimensional Brownian motion; $B_0 = 0$.

Let a, b be Borel functions of $(t, x) \in [0, T] \times \mathbf{R}^d$ such that $a = a(t, x) = (a^i(t, x))_{i\leq d}$ is d-dimensional vector valued, and $b = b(t, x) = (b_k^i(t, x))_{i\leq d, k\leq N}$ is $d \times N$-matrix valued. Set $X_0 = x \in \mathbf{R}^d$.

$$dX_t = a(t, X_t)dt + b(t, X_t)dB_t.$$

Each components is

$$dX_t^i = a^i(t, X_t)dt + \sum_{k=1}^N b_k^i(t, X_t)dB_t^k, \quad 1 \leq i \leq d.$$

a is called a drift coefficient and b is called a diffusion coefficient. The integral form is

$$X_t = x + \int_0^t a(r, X_r)dr + \int_0^t b(r, X_r)dB_r.$$

Each components is

$$X_t^i = x^i + \int_0^t a^i(r, X_r)dr + \sum_{k=1}^N \int_0^t b_k^i(r, X_r)dB_r^k, \quad 1 \leq i \leq d.$$

Set $|a|^2 = (a^1)^2 + \cdots + (a^d)^2$ and $\|b\|^2 = \sum_{i\leq d,,k\leq N}(b_k^i)^2$.

Theorem 4.1 *For the continuous type SDE, if coefficients are linear increasing and Lipschitz continuous, i.e.,* $^\forall T > 0, ^\exists K = K_T > 0; ^\forall t \in [0, T], x \in \mathbf{R}^d$,

$$|a(t, x)| + \|b(t, x)\| \leq K(1 + |x|), \qquad |a(t, x) - a(t, y)| + \|b(t, x) - b(t, y)\| \leq K|x - y|,$$

then the solution $(X_t)_{t\geq 0}$ *exists uniquely and it is continuous such that* $(X_t) \in \mathcal{L}^2$, *i.e.,* $^\forall T > 0, \int_0^T E|X_t|^2 dt < \infty$.

The uniqueness means if $(\widetilde{X_t})$ is also a solution, then it is strong equivalent to (X_t), i.e., $P(X_t = \widetilde{X_t}, {}^\forall t \geq 0) = 1$.

Proof. We use Picard iteration. The approximating sequence $X^n = (X_t^n)$ is defined as follows: Let $X_t^1 = x$ and if X^n is defined, then set

$$dX_t^{n+1} = a(t, X_t^n)dt + b(t, X_t^n)dB_t, \quad X_0^{n+1} = x.$$

It also holds that $X^n = (X_t^n)$ は $(\mathcal{F}_t)$-adapted and

$$E\left[\sup_{t\leq T} |X_t^n|^2\right] < \infty.$$

From these, X^{n+1} is well-defined and the following holds ${}^\forall T > 0, {}^\forall n \geq 1$,

$$E\left[\sup_{t\leq T} |X_t^{n+1} - X_t^n|^2\right] \leq C_2 \frac{(C_1 T)^{n-1}}{(n-1)!}.$$

Therefore in $L^2(\Omega \to C([0,T]))$ under the norm $\|X\|^* := \||X|_T^*\|_2 = (E[\sup_{t\leq T} |X_t|^2])^{1/2}$, the infinite series $\sum_{n\geq 1} \|X^{n+1} - X^n\|^*$ converges. Hence $\{X^n\}$ is a Cauchy seq. and by completeness of $L^2(\Omega \to C([0,T]))$ under $\|\cdot\|^*$, there exists a limit $X = (X_t)_{t\leq T}$ a.s. That is, $\{X^n\}$ converges to X uniformly a.s. and it also converges under $\|\cdot\|_T$. Hence, it is a solution of the SDE.

At last, for the uniqueness, if $X, \widetilde{X}$ are the solutions and let $\tau_L = \inf\{t \geq 0; |X_t| \vee |\widetilde{X}_t| \geq L\}$, then by the same way as above we have

$$E|X_{t\wedge\tau_L} - \widetilde{X}_{t\wedge\tau_L}|^2 \leq C_1 \int_0^t E|X_{r\wedge\tau_L} - \widetilde{X}_{r\wedge\tau_L}|^2 dr.$$

Hence, by the Gronwall's inequality (see the next proposition), we get

$$E|X_{t\wedge\tau_L} - \widetilde{X}_{t\wedge\tau_L}|^2 = 0, \quad 0 \leq t \leq T$$

and by the continuities of $X, \widetilde{X}$, we have $P({}^\forall t \leq T, X_{t\wedge\tau_L} = \widetilde{X}_{t\wedge\tau_L}) = 1$ and $\tau_L \to \infty$ a.s. Thus, this implies the string equivalence. ■

Proposition 4.1 (Gronwall's inequality) *For a continuous function $g(t)$ on $[0,T]$,*
$${}^\exists C_1, C_2 \geq 0;\ 0 \leq g(t) \leq C_1 + C_2 \int_0^t g(r)dr \Rightarrow g(t) \leq C_1 e^{C_2 t}$$

Proof. If we set $h(t) = e^{-C_2 t} \int_0^t C_2 g(r)dr$, then the assumption implies $h'(t) \leq C_1 C_2 e^{-C_2 t}$. By integrating on $[0,t]$ and $h(0) = 0$, we have $h(t) \leq C_1(1 - e^{-C_2 t})$. Again, by the assumption, the desired inequality is obtained. ■

4.2 Jump-type SDE

The jump-type is an adding jump terms to a continuous-type. The coefficients f, g of jump terms are Borel functions of $(t, x, z) \in [0,\infty) \times \mathbf{R}^d \times \mathbf{R}^m$ and $f = f(t,x,z) = (f^i(t,x,z))_{i\leq d}; f = 0$ on $|z| < 1$, $g = g(t,x,z) = (g^i(t,x,z))_{i\leq d}; g(t,x,0) = 0$ かつ $g = 0$ on $|z| \geq 1$. The jump-type SDE is

$$dX_t = a(t, X_t)dt + b(t, X_t)dB_t + \int_{|z|\geq 1} f(t, X_{t-}, z)N(dt, dz). + \int_{|z|<1} g(t, X_{t-}, z)\widetilde{N}(dt, dz),$$

where $\nu(dz)$ is a measure on $\mathbf{R}^m$ such that ${}^\forall n \geq 1, \nu(|z| \geq 1/n) < \infty$. $N(dt, dz)$ is a $dt\nu(dz)$-Poisson r.m., $\widehat{N}(dt, dz) = dt\nu(dz)$ and $\widetilde{N} = N - \widehat{N}$. Let assume the following on f:

$${}^\forall t > 0, x \in \mathbf{R}^d,\ \int_0^t dr \int_{|z|\geq 1} |f(r, x, z)|\nu(dz) < \infty.$$

Theorem 4.2 *For the jump-type SDE, by adding the following conditions to the continuous-type:* $^\forall T > 0, ^\exists K = K_T > 0; ^\forall t \in [0,T], x \in \mathbf{R}^d$,

$$\int_{|z|<1} |g(t,x,z)|^2 \nu(dz) \le K(1+|x|^2),$$

and

$$\int_{|z|<1} |g(t,x,z) - g(t,y,z)|^2 \nu(dz) \le K|x-y|^2,$$

the solution $(X_t)_{t\ge 0}$ *exists uniquely and it is first kind of discontinuous process. Moreover, if* $f=0$*, then* $(X_t) \in \mathcal{L}^2$ *holds.*

Proof. It can be shown by a similar way to the continuous-type. However, in order to treat large jumps, let $(\tau_k, \xi_k); 0 < \tau_k \uparrow\uparrow \infty$ be a family of all jump masses, i.e., sets of jump times and jumps. (It is possible because for any $T>0$, $N([0,T]\times\{|z|\ge 1\}) < \infty$ a.s.)

First if $f=0$, then the existence of the solution Y_t^1 can be the same way as in case of the continuous-type. Thus, let $X_t^1 = Y_t^1$ for $t \in [0,\tau_1)$, and $X_t^1 = Y_{\tau_1}^1 + f(\tau_1, Y(\tau_1-), \xi_1)$ if $t=\tau_1$). Next, let Y_t^2 be the solution for $t \in [0, \tau_2 - \tau_1]$ starting at $Y_0^2 = X_{\tau_1}^1$ corresponding to $B_t^{\tau_1} := B_{t+\tau_1} - B_{\tau_1}$ and $N^{\tau_1}(dt,dz) := N(dt+\tau_1, dz)$, determined by the same way as Y_t^1 (note that we have to change the time variable r to $r+\tau_1$). Let $X_t^2 = X_t^1$ if $t \in [0.\tau_1]$ and $X_t^2 = Y_{t-\tau_1}^2$ if $t \in [\tau_1, \tau_2]$, then this is a solution for $t \in [0,\tau_2]$. Here, more precisely, the definition of Y_t^2 is that we first change τ_1 to non-random time s and consider the starting point as $^\forall x \in \mathbf{R}^d$, and for the solution $Y_t^2 = Y_t^2(\omega; x, s)$, set $s = \tau_1(\omega), x = X_{\tau_1(\omega)}^1(\omega)$.

By continuing this operation the solution X_t is defined as $X_t = X_t^k$ for $t \in [0,\tau_k]$, and this is the unique solution. ■

[**Adding**] The proof of that X_t^2 is a solution for $t \in [0,\tau_2]$, especially for $t \in (\tau_1, \tau_2]$.
Let $s = \tau_1$ and if $t \in [0, \tau_2 - s)$, then

$$\begin{aligned} Y_t^2 - X_s^1 &= \int_0^t a(r+s, Y_r^2)dr + \int_0^t b(r+s, Y_r^2)dB_r^s + \int_0^t \int_{|z|<1} g(r+s, Y_{r-}^2, z)\widetilde{N}^s(dr,dz) \\ &= \int_s^{t+s} a(r, Y_{r-s}^2)dr + \int_s^{t+s} b(r, Y_{r-s}^2)dB_r + \int_s^{t+s}\int_{|z|<1} g(r, Y_{(r-s)-}^2, z)\widetilde{N}(dr,dz), \end{aligned}$$

where we use for a fixed $s \ge 0$, $(B_r^s = B_{r+s} - B_s) \overset{\text{(d)}}{=} (B_r)$ and

$$\int_0^t b(r+s, Y_r^2)dB_r^s = \int_0^t b(r+s, Y_r^2)d(B_{r+s} - B_s) = \int_s^{t+s} b(r, Y_{r-s}^2)dB_r.$$

Therefore, for $t \in [s, \tau_2)$, $X_t^2 = Y_{t-s}^2$ is the same as above Y_t^2, and it is a solution to the original equation. Of course,

$$X_{\tau_2}^2 = Y_{\tau_2 - s}^2 + f(\tau_2, Y_{(\tau_2-s)-}, \xi_2).$$

■

5 Transition Probabilities and Generators

Let (X_t, P_x) be a time-homogeneous Markov process on $\mathbf{R}^d$ starting from x. For a bounded Borel function φ, set

$$P_t(x, dy) := P_x(X_t \in dy), \quad P_t\varphi(x) := E_x[\varphi(X_t)] = \int_{\mathbf{R}^d} \varphi(y) P_t(x, dy).$$

$P_t(x, dy)$ is called a **transition probability**. $(P_t)_{t\geq 0}$ is a **transition semi-group** on a family of bounded continuous functions $C_b \equiv C_b(\mathbf{R}^d)$, i.e.,

$$P_0 = I, P_s P_t = P_{s+t} \ (s, t \geq 0)$$

Moreover, if (X_t) is a Markov process with sample paths in $D \equiv D(\mathbf{R}^d)$, then $\lim_{t\downarrow 0} P_t\varphi(x) = \varphi(x)$, i.e., $\lim_{t\downarrow 0} P_t \to I$ on C_b. Furthermore, it is right-continuous: $P_{t+h} \to P_t$ $(h \downarrow 0)$ on C_b for $^\forall t \geq 0$.

On the other hand, if (P_t) is right-continuous on C_b, then (X_t) is right-continuous in probability, i.e., $^\forall \varepsilon > 0$, $P(|X_{t+h} - X_t| < \varepsilon) \to 1$ $(h \downarrow 0)$ for $^\forall t \geq 0$. In fact, if $\varphi \in C_b$ satisfies $\varphi(0) = 0$ and $\varphi(x) = 1$ for $|x| \geq \varepsilon$, then

$$P(|X_{t+h} - X_t| \geq \varepsilon) \leq E[\varphi(X_{t+h} - X_t)] = E[\int \varphi(y - x) P_h(x, dy)|_{\ x=X_t}] \to E[\varphi(0)] = 0.$$

5.1 Generators

Let (P_t) be a right-continuous semi-group on C_b.

A **generator** L on $\mathcal{D}(L) \subset C_b$ is defined as

$$L := \lim_{h\downarrow 0} \frac{1}{h}(P_h - I), \quad \text{i.e.,} \quad L\varphi(x) := \lim_{h\downarrow 0} \frac{1}{h}(P_h\varphi(x) - \varphi(x)),$$

where $\varphi \in \mathcal{D}(L)$ is a family of all $\varphi \in C_b$ such that the above limit exists.

In this case. we get $P_t = e^{tL}$ formally.

Moreover, if the transition probabilities depend on the time as $P_{s,t}(x, dy) = P(X_t \in dy|\ X_s = x)$ for $0 \leq s \leq t$, then $P_{s,s} = I$, $P_{s,t}P_{t,u} = P_{s,u}$. We also assume the right-continuity of $P_{s,t}$); i.e., $P_{s+,t} = P_{s,t}, P_{s,t+} = P_{s,t}$. In this case, the generator also depends the time and it is defined by $L_t = \lim_{h\downarrow 0}(P_{t,t+h} - I)/h$ on $\mathcal{D}(L_t)$. Formally, it is given as $P_{s,t} = \exp[\int_s^t L_r dr]$.

we consider the following function spaces:

· $\varphi \in C_b^2 \equiv C_b^2(\mathbf{R}^d) \overset{\text{def}}{\iff} {}^\exists \partial_{ij}^2\varphi \in C_b \ (i, j \leq d)$, where $\partial_{ij}^2 = \partial^2/\partial x_i \partial x_j$.

· $\varphi \in C_c^\infty \overset{\text{def}}{\iff} \varphi \in C^\infty$, supp φ is compact.

Theorem 5.1 *Let (X_t) be the solution of the following jump type SDE starting from $X_0 = x$, where, coefficients (a, b, f, g) satsfy the same condition as in Theorem 4.2.*

$$dX_t = a(t, X_t)dt + b(t, X_t)dB_t + \int_{|z|\geq 1} f(t, X_{t-}, z)N(dt, dz). + \int_{|z|<1} g(t, X_{t-}, z)\widetilde{N}(dt, dz).$$

Then, the generator is given as follows: for $\varphi \in C_b^2$,

$$\begin{aligned} L_t\varphi(x) &= a^i(t, x)\partial_i\varphi(x) + \frac{1}{2}\sum_{k=1}^N b_k^i b_k^j(t, x)\partial_{ij}^2\varphi(x) \\ &\quad + \int_{|z|\geq 1} [\varphi(x + f(t, x, z)) - \varphi(x)]\nu(dz) \\ &\quad + \int_{|z|<1} [\varphi(x + g(t, x, z)) - \varphi(x) - \partial_i\varphi(x)g^i(t, x, z)]\nu(dz), \end{aligned}$$

where we sum on the same characters of superscripts and subscripts, e.g. $a^i x_i = \sum_{i\leq d} a^i x_i$.

Proof. For simplicity, let $d = 1$. By applying Ito formula to $\varphi(X_{t+h})$,

$$\begin{aligned}\varphi(X_{t+h}) - \varphi(X_t) &= \int_t^{t+h} \varphi'(X_s)d(X_s)^c + \frac{1}{2}\int_t^{t+h} \varphi''(X_s)d(X_s^2)^c \\ &+ \int_t^{t+h}\int_{|z|\geq 1} [\varphi(X_{s-} + f(s, X_{s-}, z)) - \varphi(X_{s-})]\, N(ds, dz) \\ &+ \int_t^{t+h}\int_{|z|<1} [\varphi(X_{s-} + g(s, X_{s-}, z)) - \varphi(X_{s-})]\, \widetilde{N}(ds, dz). \\ &+ \int_t^{t+h}\int_{|z|<1} [\varphi(X_{s-} + g(s, X_{s-}, z)) - \varphi(X_{s-}) - \varphi'(X_{s-})g(s, X_{s-}, z)]\, ds\nu(dz),\end{aligned}$$

where $d(X_s)^c = a(s, X_s)ds + b(s, X_s)dB_s$, $d(X_s^2)^c = (dB_s)^2 = ds$. Noting that $P_{t,t+h}\varphi(x) = E[\varphi(X_{t+h})|\ X_t = x]$, that is,

$$P_{t,t+h}\varphi(x) - \varphi(x) = E[\varphi(X_{t+h}) - \varphi(X_t)|\ X_t = x],$$

since the expectations of stochastic integrals by dB_s and $d\widetilde{N}$ are 0, we get the desired result. ■

The simplest example is for a Brownian motion. Let $X_t = x + B_t$. This is time-homogeneous and $L_t \equiv L = \triangle_x/2$.

If all coefficients does not depend on time and space, then the solution is a time-space inhomogeneous Markov process, and this is a **Lévy process**. Especially, if $f = g = z$ $(m = d)$, then ν satisfies that

$$\nu(\{0\}) = 0, \qquad \int_{\mathbf{R}^d} (1 \wedge |z|^2)\nu(dz) < \infty.$$

The generator is given as

$$L\varphi = a \cdot D\varphi + \frac{1}{2}b^2 \cdot D^2\varphi + \int_{\mathbf{R}^d} [\varphi(\cdot + z) - \varphi(\cdot) - z \cdot D\varphi(\cdot)1_{|z|<1}]\nu(dz),$$

where $a \cdot D = a^i\partial_i$, $b^2 \cdot D^2 = \sum_k b_k^i b_k^j \partial_{ij}^2$.

The characteristic function is given as

$$E[e^{i\xi\cdot X_t}] = e^{t\Psi(\xi)}; \quad \Psi(\xi) = ia \cdot \xi - \frac{1}{2}b^2 \cdot \xi^2 + \int \left[e^{iz\cdot\xi} - 1 - iz \cdot \xi 1_{|z|<1}\right] \nu(dz),$$

where $\xi \in \mathbf{R}^d$, $b^2 \cdot \xi^2 = \sum_k b_k^i b_k^j \xi_i\xi_j$.

5.2 Martingale problems

The solution of the previous SDE is called a **string solution**.

On the other hand, for the given coefficients (a, b, f, g, ν), if there exist an appropriate probability space and the solution of SDE on it, then it called a **weak solution**.

In general, for a given operator L_t, it is difficult that whether if the corresponding transition semi-group $P_{s,t}$ exists.

Let L_t be the same as in (5.1) and denote the domain as $D(L_t)$. (we sometime use this as $\bigcap_{t>0} D(L_t)$ if necessary.)

On $\Omega = D([0, \infty))$, set $\mathcal{F} = \sigma(\mathcal{C})$; $\mathcal{C}$ is a family of all cylinder sets, i.e.,

$$C \in \mathcal{C} \overset{\text{def}}{\iff} C = \{(\omega(t_1), \dots, \omega(t_n)) \in A_{t_1} \times \cdots \times A_{t_n}, \quad (0 \leq t_1 < \cdots < t_n, A_{t_j} \in \mathcal{B}^1, n \geq 1).$$

Also let $\mathcal{F}_t^0$ be the σ-additive class generated by a family of all cylinder sets until the time $t \geq 0$. and set $\mathcal{F}_t := \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}^0$.

For any fixed $x \in \mathbf{R}^d$, if exists a probability measure $P = P_x$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t))$ such that for $X_t(\omega) = \omega(t)$, it satisfies the following, then $\{P_x\}_x$ is called a **solution of a martingale problem**: with an appropriate function space D_0 (e.g., C_c^∞, C_b^2, etc),

(1) $P(X_0) = x$

(2) $^\forall \varphi \in D_0$, $M(\varphi) := \varphi(X_t) - \varphi(x) - \int_0^t L_r\varphi(X_r)dr$ is $(\mathcal{F}_t)$-martingale.

Of course, this solution is a Markov process with the generator L_t.

Clearly, a weal solution is a solution of the martingale problem. The inverse is also ture, and hence, these are equivalent. In this case $D_0 \subset D(L_t)$ is called a **core** of (L_t).

Let P_x be a solution of the martingale problem. For $0 \le s < t$, by $P_{s,t}\varphi(y) = E_x[\varphi(X_t)|\ X_s = y]$,

$$P_{s,t}\varphi(y) - \varphi(y) = E_[\varphi(X_t) - \varphi(X_s)|\ X_s = y] = \int_s^t E_x[L_r\varphi(X_r)]|\ X_s = y]dr = \int_s^t P_{s,r}L_r\varphi(y)dr.$$

This implies the generator is L_t.

Moreover, if we differentiate in t, then

$$\partial_t P_{s,t}\varphi(x) = P_{s,t}L_t\varphi(x) = \int L_t\varphi(y)P_{s,t}(x, dy).$$

For simplicity, let $d = 1$ and we consider the time-homogeneous case. Let $P_t(x, dy) = p_t(x, y)dy$ and $\partial_y p, \partial_y^2 p$, $\partial_t p$ are bounded in y. Furthermore, if L is a continuous type ($f = g = 0$), and if a, b is in C_b^2, then for $\varphi \in C_c^2$, we have (by integration-by-parts)

$$\int \varphi(y)\partial_t p_t(x, y)dy = \int \varphi(y)L^* p_t(x, y)dy.$$

That is, $\partial_t p_t(x, y) = L^* p_t(x, y)$ a.e. holds, where L^* is an adjoint operator of L and it given as

$$L^* p(y) = -\partial_i \left(a^i(y)p(y)\right) + \frac{1}{2}\partial_{ij}^2 \left(\sum_{k \le N} b_k^i b_k^j(y)p(y)\right).$$

If we don't fix a starting point of X_t, that is, if we let X_0 be a random variable, then for $u(t, y) = E[p_t(X_0, y)]$, under the same conditions, $\partial_t u = L^* u$ a.e. holds. 同じ計算で導かれる.)

References

[1] ITO, Kiyoshi "Probability Theory" (in Japanese) , Iwanami books (1953).

[2] FUNAKI, Naohisa, "Stochastic Differential Equations" (in Japanese) , Iwanami books (1997, 2nd ed. 2000).

[3] HIRABA, Seiji, "Basics of Stochastic Processes", ProbBE.pdf (2023). http://hiraba.starfree.jp/Math/eindex.html

[4] KARAZAS, SHREAVE, "Brownian Motions and Stochastic Integrals" Springer

[5] N. IKEDA, S. WATANABE, "Stochastic Differential Equations and Diffusion Processes", 2nd Ed., Holland-Kodansha (1988, 2nd.ed. 1991).