

Lecture Notes on Brownian Motion, Continuous Martingale and Stochastic Analysis (Itô's Calculus)

This lecture notes mainly follows Chapter 11, 15, 16 of the book *Foundations of Modern Probability* by Olav Kallenberg.

Recall a normal distribution $N(\mu, \sigma^2)$, $\sigma > 0$, is a probability measure on \mathbb{R} with a density function:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

If $X \sim N(\mu, \sigma^2)$, then $\mathbb{E}[X] = \mu$ and $\text{var}(X) = \mathbb{E}[(X - \mathbb{E}X)^2] = \sigma^2$. This distribution has a characteristic function:

$$\phi(t) = \mathbb{E}[e^{itX}] = e^{i\mu t - \frac{\sigma^2 t^2}{2}}.$$

We also think of the point mass δ_μ as a normal distribution $N(\mu, 0)$. It has no density function, but has characteristic function $e^{i\mu t}$. A random variable with normal distribution is called a Gaussian random variable.

Recall that for a \mathbb{R}^d -valued random vector $X = (X_1, \dots, X_d)$ with law ν , the characteristic function for X or ν is defined by

$$\phi(\underline{t}) = \mathbb{E}[e^{i \sum_{j=1}^d t_j X_j}], \quad \underline{t} = (t_1, \dots, t_d) \in \mathbb{R}^d.$$

It is known that ϕ determines μ . This implies Gramér and Wold Theorem (Corollary 4.5): for two random vectors X and Y in \mathbb{R}^d , if for any $\underline{c} \in \mathbb{R}^d$, $\sum c_j X_j \stackrel{d}{=} \sum c_j Y_j$, then $X \stackrel{d}{=} Y$. This holds because the assumption implies that X and Y have the same characteristic function.

A probability measure on \mathbb{R}^d is called a normal distribution $N(\mu, A)$, where $\mu \in \mathbb{R}^d$ and $A \in \mathbb{R}^{d^2}$ if its characteristic function is

$$\phi(\underline{t}) = e^{i \sum_{j=1}^d \mu_j t_j - \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d A_{j,k} t_j t_k}.$$

The density function exists whenever A is strictly positive. If A is singular, the measure is supported by a lower dimensional space. A random vector with normal distribution is called a Gaussian random vector.

1 Gaussian Processes and Brownian Motion

A family of random variables $(X_t)_{t \in I}$ is said to be jointly Gaussian if for any $t_1, \dots, t_n \in I$ and any $c_1, \dots, c_n \in \mathbb{R}$, $\sum_{j=1}^n c_j X_{t_j}$ is Gaussian, and is said to be centered if $\mathbb{E}[X_t] = 0$ for all t . This property holds, for example, if X_t are independent Gaussian random variables. In fact, if $X_t \sim N(a_t, \sigma_t^2)$ and are independent, then $\sum_{j=1}^n c_j X_{t_j} \sim N(\sum_{j=1}^n c_j a_{t_j}, \sum_{j=1}^n c_j^2 \sigma_{t_j}^2)$. A random process (X_t) is called a Gaussian process if the random variables are jointly Gaussian.

Lemma 11.1. *(Covariance function)*

- (i) Suppose $(X_t)_{t \in T}$ and $(Y_t)_{t \in T}$ are both jointly Gaussian. If for any $t \in T$, $\mathbb{E}[X_t] = \mathbb{E}[Y_t]$; and for any $s, t \in T$, $\text{cov}(X_s, X_t) = \text{cov}(Y_s, Y_t)$. Then $(X_t)_{t \in T}$ and $(Y_t)_{t \in T}$ have the same distribution.
- (ii) If the family $(X_t)_{t \in T}$ is jointly Gaussian, then the family is independent if and only if $\text{cov}(X_s, X_t) = 0$ for any $s \neq t \in T$.
- (iii) If $X = (X_1, \dots, X_d)$ are jointly Gaussian, then the law of X is $N(\mu, A)$, where $\mu_j = \mathbb{E}[X_j]$ and $A_{j,k} = \text{cov}(X_j, X_k)$, $1 \leq j, k \leq d$.

Proof. (i) For any $t_1, \dots, t_n \in T$ and c_1, \dots, c_n , we have

$$\mathbb{E}\left[\sum_{j=1}^n c_j X_{t_j}\right] = \sum_{j=1}^n c_j \mathbb{E}[X_{t_j}];$$

$$\text{var}\left[\sum_{j=1}^n c_j X_{t_j}\right] = \mathbb{E}\left[\left(\sum_{j=1}^n c_j (X_{t_j} - \mathbb{E}[X_{t_j}])\right)^2\right] = \sum_{j=1}^n \sum_{k=1}^n c_j c_k \text{cov}(X_{t_j}, X_{t_k}).$$

The same computation holds for $\sum_{j=1}^n c_j Y_{t_j}$. From the assumption we know that $\sum_{j=1}^n c_j X_{t_j}$ and $\sum_{j=1}^n c_j Y_{t_j}$ have the same expectation and variance. Since they are both Gaussian random variables, they must have the same distribution. By Gramér and Wold Theorem (Corollary 4.5), $(X_{t_1}, \dots, X_{t_n}) \stackrel{d}{=} (Y_{t_1}, \dots, Y_{t_n})$ for any $t_1, \dots, t_n \in T$. So by Proposition 2.2, $X \stackrel{d}{=} Y$.

(ii) The only if part is trivial. The if part follows from (i) since if we set $Y_t, t \in T$, to be independent Gaussian random variables such that $Y_t \stackrel{d}{=} X_t$ for every $t \in T$, then by (i), $X \stackrel{d}{=} Y$. So $X_t, t \in T$, must also be independent.

(iii) For any $t_1, \dots, t_d \in \mathbb{R}$, $\sum t_j X_j$ is a Gaussian variable with $\mathbb{E}[\sum t_j X_j] = \sum t_j \mathbb{E}[X_j]$ and $\text{var}(\sum t_j X_j) = \sum_j \sum_k t_j t_k \text{cov}(X_j, X_k)$. Thus,

$$\mathbb{E}\left[e^{i \sum t_j X_j}\right] = e^{i \sum t_j \mathbb{E}[X_j] - \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \text{cov}(X_j, X_k) t_j t_k}.$$

This implies that the characteristic function of X agrees with that of $N(\mu, A)$ with $\mu_j = \mathbb{E}[X_j]$ and $A_{j,k} = \text{cov}(X_j, X_k)$. So X has the law of $N(\mu, A)$. \square

Suppose $(X_t)_{t \in T}$ is jointly centered Gaussian. Let H^0 be the linear space spanned by $(X_t)_{t \in T}$. By definition, every element in H^0 is a Gaussian random variable centered at 0. Assign the L^2 norm to H^0 : $\|X\| = \mathbb{E}[X^2]^{1/2}$. Let H be the completion of H^0 . Then H is a Hilbert space. Every element of H is a Gaussian random variable. To see this, note that if $X_n \in H^0$ tends to $X \in H$, then $\phi_{X_n}(t) \rightarrow \phi_X(t)$, which implies that $\phi_X(t)$ has the form of $e^{-\frac{\sigma^2 t^2}{2}}$ for some $\sigma \geq 0$. We call H a Gaussian Hilbert space. From the lemma we know that, for $X, Y \in H$, X and Y are independent iff $\text{cov}(X, Y) = 0$. We may construct an infinite dimensional separable Gaussian Hilbert space H as follows. Let $\zeta_n, n \in \mathbb{N}$, be i.i.d. $N(0, 1)$ random variables, and let H be the L^2 -closed linear space spanned by ζ_n 's.

Proposition 11.2. *Let ζ_1, \dots, ζ_d be i.i.d. random variables with $d \geq 2$. Then the distribution of $(\zeta_1, \dots, \zeta_d)$ is spherically symmetric iff they are jointly centered Gaussian.*

Proof. If $\zeta_1, \dots, \zeta_d \stackrel{d}{=} N(0, \sigma^2)$ with $\sigma > 0$, then each ζ_j has a density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}}.$$

So the random vector $(\zeta_1, \dots, \zeta_d)$ has a density function

$$f(\underline{x}) = \prod_{j=1}^d f(x_j) = \frac{1}{(\sqrt{2\pi\sigma})^d} e^{-\frac{|\underline{x}|^2}{2\sigma^2}},$$

which is spherically symmetric. So the distribution of $(\zeta_1, \dots, \zeta_d)$ is spherically symmetric. If $\sigma = 0$, then $(\zeta_1, \dots, \zeta_d)$ is a point mass at $(0, \dots, 0)$, which is still spherically symmetric.

Now suppose the distribution of $(\zeta_1, \dots, \zeta_d)$ is spherically symmetric. Let ϕ be the characteristic function of any ζ_j , i.e., $\phi(t) = \mathbb{E}[e^{it\zeta_j}]$. Then by symmetry, $\phi(t) = \phi(-t) = \overline{\phi(t)}$, $t \in \mathbb{R}$. So ϕ takes real values on \mathbb{R} . For any $s, t \in \mathbb{R}$, by spherical symmetry, $s\zeta_1 + t\zeta_2 \stackrel{d}{=} \sqrt{s^2 + t^2}\zeta_1$. Thus,

$$\phi(s)\phi(t) = \mathbb{E}[e^{is\zeta_1 + it\zeta_2}] = \mathbb{E}[e^{i\sqrt{s^2 + t^2}\zeta_1}] = \phi(\sqrt{s^2 + t^2}).$$

Then we get $\phi(\sqrt{2}t) = \phi(t)^2$. This shows that ϕ is non-negative. By induction, $\phi(\sqrt{n}t) = \phi(t)^n$. Thus, for any $q = n/m \in \mathbb{Q}_+$, $\phi(\sqrt{q}t) = \phi(t)^q$. By continuity, we get $\phi(\sqrt{x}) = \phi(1)^x$ for any $x \geq 0$, i.e., $\phi(x) = \phi(1)^{x^2}$ for $x > 0$. Since $\phi(-t) = \phi(t)$, we have $\phi(x) = \phi(1)^{x^2}$ for any $x \in \mathbb{R}$. Since $|\phi| \leq 1$, $\phi(1) = e^{-\frac{\sigma^2}{2}}$ for some $\sigma \geq 0$. So $\phi(t) = e^{-\frac{\sigma^2 t^2}{2}}$, which means that each ζ_j is a centered Gaussian random variable with variance σ^2 . \square

We omit Theorems 11.3 and 11.4.

Theorem 11.5, part 1 (Existence of Brownian motion). *There exists a Gaussian process B with independent increments such that $B_0 = 0$ and for any $t \geq s$, $B_t - B_s \sim N(0, t - s)$.*

Proof. There are several ways to prove the existence. One way is to show that the transition kernels

$$\mu_{s,t}(x, \cdot) = N(x, t - s), \quad t \geq s \geq 0,$$

satisfy the Chapman-Kolmogorov relation, and apply Kolmogorov extension theorem to construct a Markov process with this transition kernel.

We now give another proof. First, we construct an infinite dimensional separable Gaussian Hilbert space H . Such H is isomorphic to $L^2(\mathbb{R}_+, \lambda)$, where λ is the Lebesgue measure on \mathbb{R} . Let $f : L^2(\mathbb{R}_+, \lambda) \rightarrow H$ be an isomorphism. Define

$$B_t = f(\mathbf{1}_{[0,t]}) \in H.$$

Then B is a centered Gaussian process, and from $f(0) = 0$ we get a.s. $B_0 = 0$. For $t > s$, $\text{var}(B_t - B_s) = \|\mathbf{1}_{[0,t]} - \mathbf{1}_{[0,s]}\|^2 = t - s$, so $B_t - B_s \sim N(0, t - s)$. Finally, if $t_0 \leq t_1 \leq \dots \leq t_n$, then $B_{t_1} - B_{t_0}, \dots, B_{t_n} - B_{t_{n-1}}$ are f -images of $\mathbf{1}_{(t_0,t_1]}, \dots, \mathbf{1}_{(t_{n-1},t_n]}$, which are orthogonal in $L^2(\mathbb{R}_+, \lambda)$. So $B_{t_1} - B_{t_0}, \dots, B_{t_n} - B_{t_{n-1}}$ are orthogonal in H . Thus, they are independent. \square

Theorem 11.5, part 2 (Continuity of Brownian motion). *The B in the previous theorem has a continuous version, which is locally Hölder continuous of any exponent $a \in (0, 1/2)$.*

Recall that B' is called a version of B if for every $t \geq 0$, a.s. $B'_t = B_t$. It is weaker than the condition that, a.s. for every $t \geq 0$, $B'_t = B_t$, in which case we say that B' and B are indistinguishable. If B has two continuous versions B' and B'' , then B' and B'' are indistinguishable because first, for every $t \in \mathbb{Q}_+$, a.s. $B'_t = B''_t$; second a.s. for every $t \in \mathbb{Q}_+$, $B'_t = B''_t$ because \mathbb{Q}_+ is countable; and finally a.s. $B' \equiv B''$ by continuity of B' and B'' , and the denseness of \mathbb{Q}_+ in \mathbb{R}_+ .

A function f is said to be Hölder continuous of exponent a if there is a constant $C > 0$ such that for any x, y in the domain of f , $|f(x) - f(y)| \leq C|x - y|^a$. If called locally Hölder continuous of exponent a if for any x in the domain of f , there is a neighborhood U of x such that $f|_U$ is Hölder continuous of exponent a . To prove the theorem, we use the following theorem.

Theorem 2.23. *Let $X_t, t \geq 0$, be a real valued process. Assume that there are $a, b, C > 0$ such that for any $s, t \in \mathbb{R}_+$,*

$$\mathbb{E}[|X_t - X_s|^a] \leq C|t - s|^{1+b}.$$

Then X has a continuous version, which is a.s. locally Hölder continuous with exponent c for any $c \in (0, b/a)$.

Proof. It suffices to consider $X|_{[0,1]}$. Fix $c \in (0, b/a)$. For $n \in \mathbb{N}$, let D_n be the set of binary points of level n in $[0, 1]$, i.e., $\frac{k}{2^n}, 0 \leq k \leq 2^n$. We first consider the Hölder property of $X|_{D_n}$. Let

$$\zeta_n = \max\{|X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}| : 1 \leq k \leq 2^n\}.$$

Then we have

$$\begin{aligned} \mathbb{P}[\zeta_n > (\frac{1}{2^n})^c] &\leq \sum_{k=1}^{2^n} \mathbb{P}[|X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}|^a > (\frac{1}{2^n})^{ca}] \\ &\leq 2^{can} \sum_{k=1}^{2^n} \mathbb{E}[|X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}|^a] \leq 2^{can} 2^n C 2^{-n(1+b)} = C 2^{n(ca-b)}. \end{aligned}$$

Since $ca < b$,

$$\sum_{n=1}^{\infty} \mathbb{P}[\zeta_n > (\frac{1}{2^n})^c] \leq \sum_{n=1}^{\infty} C 2^{n(ca-b)} < \infty.$$

By Borel-Cantelli lemma, a.s. there is a random N such that for any $n > N$, $\zeta_n \leq (\frac{1}{2^n})^c$. So there a.s. exists a random $C > 0$ such that for any $n \in \mathbb{N}$, and $t, s \in D_n$ with $|t - s| = \frac{1}{2^n}$, $|X_t - X_s| \leq C|t - s|^c$. Let E denote the event that the inequality holds, which has probability 1. Let $D = \bigcup_n D_n$. Then D is dense in $[0, 1]$.

We now show that $X|_D$ is Hölder continuous with exponent c on the event E . Suppose E occurs. Let $x < y \in D$. Let n_0 be the smallest $n \in \mathbb{N}$ such that $D_n \cap [x, y] \neq \emptyset$. Let $z_0 \in D_{n_0} \cap [x, y]$. Then $|x - z_0|, |y - z_0| \leq \frac{1}{2^{n_0}}$. We first estimate $|X_x - X_{z_0}|$. If $x = z_0$, it is trivial. Suppose $x < z_0$. Then $x \in D_m$ for some $m > n_0$. Then we can find $n_0 < n_1 < \dots < n_k = m$ and $x = x_k < x_{k-1} < \dots < x_0 = z_0$ such that $x_k \in D_{n_k}$, $0 \leq k \leq n$ and $|x_k - x_{k-1}| = 2^{-n_k}$. Then we get

$$|X_x - X_{z_0}| \leq \sum_{j=1}^k |X_{x_j} - X_{x_{j-1}}| \leq C \sum_{j=1}^k C(2^{-n_j})^c \leq C 2^{-cn_1} \sum_{s=0}^{\infty} 2^{-sc} = \frac{C 2^{-cn_1}}{1 - 2^{-c}}.$$

On the other hand, $|x - z_0| \geq |x_1 - x_0| = 2^{-n_1}$. So $|X_x - X_{z_0}| \leq \frac{C}{1 - 2^{-c}} |x - z_0|^c \leq \frac{C}{1 - 2^{-c}} |x - y|^c$. Symmetrically, $|X_y - X_{z_0}| \leq \frac{C}{1 - 2^{-c}} |x - y|^c$. Thus,

$$|X_x - X_y| \leq |X_x - X_{z_0}| + |X_y - X_{z_0}| \leq \frac{2C}{1 - 2^{-c}} |x - y|^c.$$

Since $X|_D$ is Hölder continuous on the event E , it extends to a continuous process Y on $[0, 1]$ by the denseness of D in $[0, 1]$. We now show that Y is a version of X , i.e., for any $t \in [0, 1]$, a.s. $X_t = Y_t$. Fix $t \in [0, 1]$. Let $t_n \in D_n$, $n \in \mathbb{N}$, be such that $t_n \rightarrow t$. Then a.s. $X_{t_n} \rightarrow Y_t$. Since $\mathbb{E}[|X_{t_n} - X_t|^a] \leq C|t_n - t|^{1+b}$, we have $X_{t_n} \rightarrow X_t$ in L^a , and so $X_{t_n} \xrightarrow{P} X_t$. So we must have a.s. $X_t = Y_t$. \square

Proof of Theorem 11.5, part 2. For any $t > s$, since $X_t - X_s \sim N(0, t - s)$, $(X_t - X_s)/\sqrt{t - s} \sim N(0, 1)$. For $a > 2$, we have

$$\mathbb{E}[|X_t - X_s|^a] = |t - s|^{a/2} \int_{\mathbb{R}} |x|^a \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = C|t - s|^{a/2}.$$

for some constant C depending on a . Let $b = a/2 - 1$. By Theorem 2.23, X has a continuous version, which is Hölder continuous with any exponent in $(0, \frac{a/2-1}{a})$. By letting $a \rightarrow \infty$, we find that the continuous version of X is Hölder continuous with any exponent in $(0, 1/2)$. \square

The continuous centered Gaussian process given by Theorem 11.5 is called a (standard) Brownian motion (BM for short) or a Wiener process. For $x \in \mathbb{R}$, $x + B$ is called a Brownian motion started from x . A Brownian motion in \mathbb{R}^d is a process $B_t = (B_t^1, \dots, B_t^d)$ such that B_t^j , $1 \leq j \leq d$, are independent 1d Brownian motions. For $\underline{x} \in \mathbb{R}^d$, $\underline{x} + B$ is called a Brownian motion in \mathbb{R}^d started from \underline{x} .

Recall that the distribution of a centered Gaussian process X is determined by its covariance function, i.e., $\text{cov}(X_t, X_s)$ for $t, s \in T$. Let $t, s \geq 0$. If $t \geq s$, then

$$\text{cov}(B_t, B_s) = \text{cov}(B_t - B_s, B_s) + \text{var}(B_s) = 0 + s = s = t \wedge s.$$

Symmetrically, if $s \geq t$, we also have $\text{cov}(B_t, B_s) = t \wedge s$. For $c \in \mathbb{R}$,

$$\text{cov}(cB_t, cB_s) = c^2 \text{cov}(B_t, B_s) = c^2(t \wedge s) = (c^2 t) \wedge (c^2 s) = \text{cov}(B_{c^2 t}, B_{c^2 s}), \quad t, s \geq 0.$$

So $cB \stackrel{d}{=} B_{c^2 t}$, i.e., scaling the space by a factor c is equivalent to scaling the time by a factor c^2 . This is called the Brownian scaling. We call $x + B_{c^2 t}$ a Brownian motion started from x with speed c^2 . Taking $c = -1$, we find the symmetry of BM: $-B \stackrel{d}{=} B$. For any fixed t_0 , consider the process $B_{t_0+t} - B_{t_0}$, $t \geq 0$. Since for $t, s \geq 0$,

$$\begin{aligned} \text{cov}(B_{t_0+t} - B_{t_0}, B_{t_0+s} - B_{t_0}) &= \text{cov}(B_{t_0+t}, B_{t_0+s}) - \text{cov}(B_{t_0}, B_{t_0+s}) - \text{cov}(B_{t_0+t}, B_{t_0}) + \text{var}(B_{t_0}) \\ &= (t_0 + t) \wedge (t_0 + s) - t_0 - t_0 + t_0 = t \wedge s = \text{cov}(B_t, B_s), \end{aligned}$$

$$(B_{t_0+t} - B_{t_0})_{t \geq 0} \stackrel{d}{=} B.$$

We may use a similar idea to construct Brownian sheet, which is a Gaussian process with two time variables: $X(t_1, t_2)$, $t_1, t_2 \geq 0$. For the construction, consider a isomorphism $f : L^2(\mathbb{R}_+^2, \lambda^2) \rightarrow H$, where H is a Gaussian Hilbert space. The Brownian sheet is defined by $X(t_1, t_2) = f(\mathbf{1}_{[0, t_1] \times [0, t_2]})$.

Exercise. Find the covariance function for Brownian sheet.

A Brownian motion is defined on \mathbb{R}_+ . We may extend it to \mathbb{R} as follows. Let B^+ and B^- be two independent BM on \mathbb{R}_+ . Define $B_t^{\mathbb{R}} = B_t^+$ if $t \geq 0$, $B_t^{\mathbb{R}} = B_{-t}^-$ if $t \leq 0$. This process is not stationary because $B_0^{\mathbb{R}}$ is constant 0 but other $B_t^{\mathbb{R}}$ is not. But it has the following nice properties.

Exercise. Prove that for any fixed $t_0 \in \mathbb{R}$, $B_{t_0+t}^{\mathbb{R}} - B_{t_0}^{\mathbb{R}}$, $t \in \mathbb{R}$, has the same distribution as $B^{\mathbb{R}}$.

Let \mathcal{F}^0 be the natural filtration generated by B . This means that for every $t \geq 0$, \mathcal{F}_t^0 is the σ -algebra generated by B_s , $0 \leq s \leq t$. Let \mathcal{N} denote the family of subsets N of Ω such that

there exists some $A \in \mathcal{F}$ with $\mathbb{P}[A] = 0$ and $N \subset A$. We let \mathcal{F} denote the completion of \mathcal{F}^0 , i.e., for every $t \geq 0$, \mathcal{F}_t is the σ -algebra generated by \mathcal{F}_t^0 and \mathcal{N} .

Fix $t_0 \geq 0$. Since for every $t \geq 0$ and $s \in [0, t_0]$, $B_{t_0+t} - B_{t_0}$ is independent of B_s , we see that the process $B_{t_0+t} - B_{t_0}$, $t \geq 0$, is independent of $\mathcal{F}_{t_0}^0$, and so is also independent of \mathcal{F}_{t_0} .

Proposition . \mathcal{F} is right-continuous.

Proof. Fix $t_0 \geq 0$ and let $t_1 > t_2 > \dots \rightarrow t_0$. For $n \in \mathbb{N}$, let \mathcal{G}_n be the σ -algebra generated by $B_t - B_{t_n}$, $t \in [t_{n+1}, t_n]$. By independent increment property of B , \mathcal{G}_n , $n \in \mathbb{N}$, are independent. We also note that

$$\mathcal{F}_{t_n}^0 = \sigma(\mathcal{F}_{t_0}^0, \mathcal{G}_n, \mathcal{G}_{n+1}, \dots), \quad n \in \mathbb{N},$$

since if $t \in (t_0, t_n)$, $B_t - B_{t_0}$ can be expressed as an infinite sum of random variables, each of which is \mathcal{G}_k -measurable for some $k \geq n$. By Corollary 6.25, we then get a.s. $\bigcap_n \mathcal{F}_{t_n}^0 = \mathcal{F}_{t_0}^0$, which implies that $\bigcap_n \mathcal{F}_{t_n} = \mathcal{F}_{t_0}$. \square

Proposition . B is a martingale and time-homogeneous and space-homogeneous Markov process with transition kernel $\mu_{s,t}(x, A) = N(x, t-s)(A)$ w.r.t. \mathcal{F}^0 or \mathcal{F} .

Proof. We first work on \mathcal{F}^0 . To check the martingale property, fix $t_0, t \geq 0$. Since $N(0, t) \sim B_{t_0+t} - B_{t_0} \perp\!\!\!\perp \mathcal{F}_{t_0}^0$, we have $\mathbb{E}[B_{t_0+t} - B_{t_0} | \mathcal{F}_{t_0}^0] = \mathbb{E}[B_{t_0+t} - B_{t_0}] = 0$. So B is an \mathcal{F}^0 -martingale. The space-homogeneous Markov property of B follows from Proposition 7.5 and the fact that B has \mathcal{F}^0 -independent increments, i.e., $B_{t_0+t} - B_{t_0} \perp\!\!\!\perp \mathcal{F}_{t_0}^0$ for any $t_0, t \geq 0$. The proposition also tells us that the transition kernels are given by $\mu_{s,t}(x, A) = \mathbb{P}[X_t - X_s \in A - x] = N(0, t-s)(A - x) = N(x, t-s)(A)$. Since $\mu_{s,t}$ depends only on $t-s$, B is time-homogeneous. The above argument also works for \mathcal{F} . \square

We may use the Brownian motion B to define other continuous Gaussian processes. Let

$$X_t = B_t - tB_1, \quad 0 \leq t \leq 1.$$

Then $X_0 = X_1 = 0$. Such X is called a Brownian bridge (from 0 to 0 with time span 1). It is a Gaussian process with covariance function

$$\begin{aligned} \text{cov}(X_t, X_s) &= \text{cov}(B_t - tB_1, B_s - sB_1) \\ &= \text{cov}(B_t, B_s) - t \text{cov}(B_1, B_s) - s \text{cov}(B_1, B_t) + st \text{cov}(B_1, B_1) \\ &= t \wedge s - ts - ts + ts = t \wedge s - ts = t(1-s) \wedge s(1-t). \end{aligned}$$

We then immediately see that $(X_{1-t}) \stackrel{d}{=} X$ because

$$\text{cov}(X_{1-t}, X_{1-s}) = ((1-t)s) \wedge ((1-s)t) = \text{cov}(X_t, X_s).$$

Lemma 11.6. *If B is a Brownian motion, then $tB_{1/t}$ is also a Brownian motion, and $(1-t)B_{t/(1-t)}$ and $tB_{(1-t)/t}$ are Brownian bridges. If X is a Brownian bridge, then $(1+t)X_{t/(1+t)}$ and $(1+t)X_{1/(1+t)}$ are Brownian motions.*

Proof. All processes are centered Gaussian processes, whose distributions are determined by their covariance functions. To prove that the process $tB_{1/t}$ has the same distribution as B , we note that

$$\text{cov}(tB_{1/t}, sB_{1/s}) = ts \text{cov}(B_{1/t}, B_{1/s}) = ts((1/t) \wedge (1/s)) = s \wedge t.$$

□

Exercise. Prove other statements in Lemma 11.6.

Lemma . *Let H be a Gaussian Hilbert space. Let $\zeta \in H$. Let \mathcal{F}_0 be the σ -algebra generated by a set $S \subset H$, and let \mathcal{F} be its completion. Let H_0 be the closed linear space spanned by S . Let ζ_0 be the orthogonal projection of ζ onto H_0 , and let $\sigma = \text{dist}(\zeta, H_0) = \|\zeta - \zeta_0\|_{L^2}$. Then the conditional law of ζ given \mathcal{F} , i.e., $\text{Law}(\zeta|\mathcal{F}) = N(\zeta_0, \sigma^2)$.*

Proof. Since ζ_0 belongs to the closed linear space spanned by S , which generates \mathcal{F}_0 , we have $\zeta_0 \in \mathcal{F}$. In fact, there is a sequence ζ_n , each of which belongs to the linear space spanned by S , which converges a.s. to ζ_0 . Then each $\zeta_n \in \mathcal{F}^0$, so $\zeta_0 \in \mathcal{F}$. Since $\zeta - \zeta_0 \perp H_0$, we have $\zeta - \zeta_0 \perp \mathcal{F}$. Since $\zeta - \zeta_0 \in H$ and $\sigma = \|\zeta - \zeta_0\|_{L^2}$, $\text{Law}(\zeta - \zeta_0) = N(0, \sigma^2)$. Since $\zeta - \zeta_0 \perp \mathcal{F}$, the conditional law $\text{Law}(\zeta - \zeta_0|\mathcal{F})$ is the same as the unconditional law, i.e., $N(0, \sigma^2)$. Since $\zeta_0 \in \mathcal{F}$, $\text{Law}(\zeta|\mathcal{F}) = \zeta_0 + N(0, \sigma^2) = N(\zeta_0, \sigma^2)$. □

Proposition 11.7 (Gaussian Markov Processes). *Let $X_t, t \in T$, be a centered Gaussian process. Define $r_{s,t} = \text{cov}(X_s, X_t)$, $s, t \in T$. Then X is Markov iff for any $s \leq t \leq u$, if $r_{t,t} \neq 0$, then $r_{s,u}r_{t,t} = r_{s,t}r_{t,u}$; if $r_{t,t} = 0$, then $r_{s,u} = 0$.*

Proof. Fix $t \leq u \in T$. We need to show that $\text{Law}(X_u|\mathcal{F}_t) = \text{Law}(X_u|X_t)$. Let H_t be the closed linear space spanned by $X_s, s \leq t$. Let L_t be the linear space spanned by X_t . By the previous lemma, it suffices to show that the orthogonal projection of X_u to H_t agrees with the orthogonal projection of X_u to L_t . If $r_{t,t} = 0$, then $L_t = \{0\}$, and we need that $X_u \perp H_t$, i.e., $r_{s,u} = 0$ for any $s \leq t$. If $r_{t,t} \neq 0$, then the projection of X_u to L_t can be described as aX_t , where $a = \frac{\text{cov}(X_t, X_u)}{\text{cov}(X_t, X_t)} = \frac{r_{t,u}}{r_{t,t}}$. So we have $X'_u := X_u - aX_t \perp X_s$ for any $s \leq t$, which gives

$$0 = \text{cov}(X'_u, X_s) = \text{cov}(X_u - aX_t, X_s) = r_{s,u} - ar_{s,t} = r_{s,u} - r_{t,u}r_{s,t}/r_{t,t},$$

if $r_{t,t} \neq 0$. The above argument can be reversed. So we get the equivalence. □

We note that Brownian Bridge is Markov because for $s \leq t \leq u$, $r_{s,u} = s - su$, $r_{t,t} = t(1-t)$, $r_{s,t} = s(1-t)$, $r_{t,u} = t(1-u)$.

For a Brownian motion B , we define a centered Gaussian process

$$Y_t = e^{-t}B\left(\frac{1}{2}e^{2t}\right), \quad t \in \mathbb{R}.$$

It is called a stationary Ornstein-Uhlenbeck process. It has covariance

$$\text{cov}(Y_t, Y_s) = e^{-t}e^{-s}\frac{1}{2}e^{2t \wedge 2s} = \frac{1}{2}e^{-|t-s|}.$$

Since for $s \leq t \leq u$, $\frac{1}{2}e^{-|t-s|} \cdot \frac{1}{2}e^{-|u-t|} = \frac{1}{2}e^{-|u-s|} \cdot \frac{1}{2}e^{-|t-t|}$, it is a Gaussian Markov process. From the covariance function, we find that Y is stationary and time-reversible, i.e., for any $t_0 \in \mathbb{R}$, $(Y_{t_0+t}) \stackrel{d}{=} Y$ and $(Y_{-t}) \stackrel{d}{=} Y$.

Exercise. Find the Markov transition kernels of a Brownian bridge and a stationary Ornstein-Uhlenbeck process.

There is another important Gaussian process, called fractional Brownian motion (fBM for short). A fBM with Hurst index $H \in (0, 1)$ is a centered Gaussian process B^H indexed by \mathbb{R} with the covariance function

$$\text{cov}(B_t^H, B_s^H) = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}), \quad t, s \in \mathbb{R}.$$

It may be constructed using an isomorphism from $L^2(\mathbb{R}, \lambda)$ to a Gaussian Hilbert space.

Exercise. Prove the following. (i) $B^{1/2}$ agrees with BM on \mathbb{R} . (ii) For any $a > 0$, $(B_{at}^H) \stackrel{d}{=} |a|^H B^H$. (iii) For any fixed $t_0 \in \mathbb{R}$, $(B_{t_0+t}^H - B_{t_0}^H)_{t \in \mathbb{R}} \stackrel{d}{=} B^H$. (iv) B^H has a continuous version, which is locally Hölder continuous with any exponent less than H .

We omit Lemma 11.8.

We will study sample path properties of Brownian motion.

Recall that for a function f defined on $[s, t]$, the (total) variation of f is defined to be

$$\sup \sum_{k=1}^n |f(t_k) - f(t_{k-1})|,$$

where the supremum is over all partitions $\{s = t_0 < t_1 < \dots < t_n = t\}$ of $[s, t]$. For Brownian motion, we are not interested in its variation, but its quadratic variation.

Theorem 11.9 (quadratic variation). *Let B be a Brownian motion. Fix $t > 0$. Let $\Delta = \{0 = t_0 < \dots < t_n = t\}$ be a partition of $[0, t]$. Let $h_\Delta := \max\{t_k - t_{k-1} : 1 \leq k \leq n\}$ be the mesh size of Δ . Let*

$$\zeta_\Delta = \sum_{k=1}^n (B_{t_k} - B_{t_{k-1}})^2.$$

Then as $h_\Delta \rightarrow 0$, $\zeta_\Delta \rightarrow t$ in L^2 . Moreover, if $(\Delta_n)_{n \in \mathbb{N}}$ is a nested sequence of partitions of $[0, t]$, i.e., $\Delta_1 \subset \Delta_2 \subset \dots$, and $h_{\Delta_n} \rightarrow 0$, then a.s. $\zeta_{\Delta_n} \rightarrow t$.

Proof. Since $B_{t_k} - B_{t_{k-1}}$, $1 \leq k \leq n$, are independent, and $B_{t_k} - B_{t_{k-1}} \sim N(0, t_k - t_{k-1}) \sim (t_k - t_{k-1})^{1/2} B_1$, we have

$$\mathbb{E}[\zeta_\Delta] = \sum_{k=1}^n \mathbb{E}[(B_{t_k} - B_{t_{k-1}})^2] = \sum_{k=1}^n (t_k - t_{k-1}) = t_n - t_0 = t;$$

$$\begin{aligned}\|\zeta_\Delta - t\|_{L^2}^2 &= \mathbb{E}[(\zeta_\Delta - t)^2] = \text{var}(\zeta_\Delta) = \sum_{k=1}^n \text{var}((B_{t_k} - B_{t_{k-1}})^2) \\ &= \sum_{k=1}^n (t_k - t_{k-1})^2 \text{var}(B_1^2) \leq t \text{var}(B_1^2) h_\Delta.\end{aligned}$$

Thus, as $h_\Delta \rightarrow 0$, $\zeta_\Delta \rightarrow t$ in L^2 .

Now suppose $\Delta_1 \subset \Delta_2 \subset \dots$, and $h_{\Delta_n} \rightarrow 0$. By inserting more partitions, we may assume that Δ_{n+1} contains exactly one more point than Δ_n . To prove that a.s. $\zeta_{\Delta_n} \rightarrow t$, we will use the martingale convergence theorem. We will show that (ζ_{Δ_n}) is a reverse martingale w.r.t. the filtration $(\mathcal{F}_{-n})_{n \in \mathbb{N}}$, where \mathcal{F}_{-n} is the σ -algebra generated by ζ_{Δ_k} , $k \geq n$. This means that, for any $n \in \mathbb{N}$,

$$\mathbb{E}[\zeta_{\Delta_{n-1}} - \zeta_{\Delta_n} | \zeta_{\Delta_n}, \zeta_{\Delta_{n+1}}, \dots] = 0.$$

Once this is proved, by martingale convergence theorem (Theorem 6.18), we know that (ζ_{Δ_n}) a.s. converges. Since we already know that it converges to t in L^2 , the a.s. limit is also t .

Suppose $\Delta_n \setminus \Delta_{n-1} = \{b\}$, and the partition interval of Δ_{n-1} that contains b is $[a, c]$. Then

$$\zeta_{\Delta_{n-1}} - \zeta_{\Delta_n} = (B_c - B_a)^2 - ((B_c - B_b)^2 + (B_b - B_a)^2) = 2(B_c - B_b)(B_b - B_a).$$

We now introduce another probability space $\Sigma = \{1, -1\}$ with $\mathbb{P}\{1\} = \mathbb{P}\{-1\} = 1/2$. Consider the product space $\Omega' = \Sigma \times \Omega$. On the product space Ω' , we have a random variable $\theta(\sigma, \omega) = \sigma$, which is independent of Ω , and has distribution $\mathbb{P}[\theta = \pm 1] = \frac{1}{2}$. Define on Ω' a new process $B'_s = B_s$, if $s \leq b$; $B'_s = B_b + \theta(B_s - B_b)$ if $s \geq b$. Since $B_{b+s} - B_b$, $s \geq 0$, is a BM independent of B_s , $s \leq b$, by the symmetry of BM and the independence between θ and B , $\theta \cdot (B_{b+s} - B_b)$ is also a BM independent of B_s , $s \leq b$. Since B (resp. B') can be recovered from B_s , $s \leq b$ and $B_{b+s} - B_b$, $s \geq 0$ (resp. $\theta \cdot (B_{b+s} - B_b)$, $s \geq 0$) in the same way, we see that B' has the same distribution as B , and so is also a BM. We define ζ'_{Δ_k} for B' in the same way as ζ_{Δ_k} for B . Then the process (ζ'_{Δ_k}) has the same distribution as (ζ_{Δ_k}) . Now it suffices to show that

$$\mathbb{E}[\zeta'_{\Delta_{n-1}} - \zeta'_{\Delta_n} | \zeta'_{\Delta_n}, \zeta'_{\Delta_{n+1}}, \dots] = 0.$$

We observe that $\zeta'_{\Delta_k} = \zeta_{\Delta_k}$ for $k \geq n$ because each partition interval of Δ_k , $k \geq n$, lies either in $[0, b]$ or in $[b, t]$. But $\zeta'_{\Delta_n} - \zeta'_{\Delta_{n-1}} = 2(B'_c - B'_b)(B'_b - B'_a) = \theta \cdot (\zeta_{\Delta_n} - \zeta_{\Delta_{n-1}})$. So the equality becomes

$$\mathbb{E}[\theta \cdot (\zeta_{\Delta_n} - \zeta_{\Delta_{n-1}}) | \zeta_{\Delta_n}, \zeta_{\Delta_{n+1}}, \dots] = 0,$$

which follows from the independence of θ and ζ_k 's. This shows that (ζ'_{Δ_k}) is a reversed martingale, and so is (ζ_{Δ_k}) . The proof is now complete. \square

If f is continuous and has bounded variation V on $[s, t]$, and Δ_n , $n \in \mathbb{N}$, is a nested sequence of partitions of $[s, t]$ with $h_{\Delta_n} \rightarrow 0$, then for any partition of $[s, t]$,

$$\sum_{k=1}^n |f(t_k) - f(t_{k-1})|^2 \leq \sup_{x, y \in [s, t], |x-y| \leq h_{\Delta_n}} |f(x) - f(y)| V \rightarrow 0, \quad n \rightarrow \infty.$$

If f is Brownian motion, this convergence a.s. does not hold. So we have the corollary:

Corollary 11.10. *Brownian motion a.s. has unbounded variation on every interval $[s, t]$ with $s < t$.*

Suppose we have a filtration $(\mathcal{F}_t)_{t \geq 0}$. We say that $B_t, t \geq 0$, is an \mathcal{F} -Brownian motion if

- (i) B is \mathcal{F} -adapted;
- (ii) for any $s \geq 0$, $B_{s+t} - B_s, t \geq 0$, is a Brownian motion, and is independent of \mathcal{F}_s .

For example, if B is a Brownian motion, and \mathcal{F}^0 is the natural filtration generated by B , and \mathcal{F} is the completion of \mathcal{F}^0 , then B is a Brownian motion w.r.t. \mathcal{F}^0 or \mathcal{F} . Here we use the property that B has time-homogeneous independent increments. If B is an \mathcal{F} -Brownian motion, and \mathcal{F}' is the completion of \mathcal{F} , then B is also an \mathcal{F}' -Brownian motion. This holds because the independence property does not care about null sets. If \mathcal{F}^+ is the right-continuous augmentation of \mathcal{F} , then B is also an \mathcal{F}^+ -Brownian motion. To see this, fix $s \geq 0$. Then for any $\delta > 0$, $B_{s+\delta+t} - B_{s+\delta}$ is independent of $\mathcal{F}_{s+\delta}$, which contains \mathcal{F}_s^+ . So $B_{s+\delta+t} - B_{s+\delta}, t \geq 0$, is independent of \mathcal{F}_s^+ . Letting $\delta \downarrow 0$, we see that $B_{s+t} - B_s, t \geq 0$, is independent of \mathcal{F}_s^+ .

If B is an \mathcal{F} -Brownian motion, then it is a time-homogeneous \mathcal{F} -Markov process because it has \mathcal{F} -independent increments. By Proposition 7.9 (strong Markov property), if τ is a finite stopping time taking countably many values, then conditional on \mathcal{F}_τ , $B_{\tau+t}, t \geq 0$, is a Brownian motion started from B_τ , which implies that $B_{\tau+t} - B_\tau, t \geq 0$, is a Brownian motion independent of \mathcal{F}_τ . We now improve this result without assuming that τ takes finitely many values.

Theorem 11.11 (Strong Markov Property). *Let B be an \mathcal{F} -Brownian motion. Let τ be a finite weak \mathcal{F} -stopping time. Then $B_{\tau+t} - B_\tau, t \geq 0$, is a Brownian motion independent of \mathcal{F}_τ^+ . Thus, conditionally on \mathcal{F}_τ^+ , $B_{\tau+t}$ is a Brownian motion started from B_τ .*

Proof. By Lemma 6.4, we may take a decreasing sequence of finite \mathcal{F} -stopping times τ_n such that each τ_n takes values in $\frac{\mathbb{Z}^+}{2^n}$ and $\tau_n \downarrow \tau$. Then by Lemma 6.3, $\mathcal{F}_\tau^+ = \bigcap_n \mathcal{F}_{\tau_n}$.

Since B is time-homogeneous and space-homogeneous Markov, and τ_n takes countably many values, by Proposition 7.9 (Strong Markov Property), $B_{\tau_n+t} - B_{\tau_n}, t \geq 0$, is a Brownian motion independent of \mathcal{F}_{τ_n} , and so is also independent of \mathcal{F}_τ^+ . Letting $\tau_n \rightarrow \tau$, we conclude that $B_{\tau+t} - B_\tau, t \geq 0$, is a Brownian motion independent of \mathcal{F}_τ^+ . \square

Omit Corollary 11.12.

Proposition 11.13. *Let B be a Brownian motion, and define $M_t = \sup_{s \leq t} B_s, t \geq 0$. Then for any $t \geq 0$,*

$$M_t \stackrel{d}{=} M_t - B_t \stackrel{d}{=} |B_t|.$$

Here we remark that as processes, M and $|B|$ do not have the same distribution because M is increasing but $|B|$ is not.

Lemma 11.14 (reflection principle). *For any stopping time τ , a Brownian motion B has the same distribution as the process*

$$\tilde{B}_t = B_{t \wedge \tau} - (B_t - B_{t \wedge \tau}), \quad t \geq 0.$$

Note that $\tilde{B}_t = B_t$ if $t \leq \tau$; $\tilde{B}_{\tau+t} - \tilde{B}_\tau = -(B_{\tau+t} - B_\tau)$ for $\tau < \infty$ and $t \geq 0$. This means that the increment of \tilde{B} after τ is a reflection of the increment of B after τ .

Proof. We first assume that τ is a bounded stopping time. By strong Markov property of BM, $B_t^\tau := B_{\tau+t} - B_\tau$ is a Brownian motion independent of \mathcal{F}_τ , and so is independent of the process B_t , $0 \leq t \leq \tau$. By symmetry of BM, $-B_t^\tau$ is also a BM independent of B_t , $0 \leq t \leq \tau$. Thus, $(B_t^\tau, B_{t \wedge \tau})$ and $(-B_t^\tau, B_{t \wedge \tau})$ have the same law. Since we may obtain B from B_t , $0 \leq t \leq \tau$, and B_t^τ , and obtain \tilde{B} from B_t , $0 \leq t \leq \tau$, and $-B_t^\tau$ in the same way, we see that B has the same law as \tilde{B} .

Now we do not assume that τ is bounded. It suffices to show that, for any $a \geq 0$, \tilde{B}_t , $0 \leq t \leq a$, has the same law as B_t , $0 \leq t \leq a$. Fix $a \geq 0$. If we define \hat{B}_t in the same way as \tilde{B}_t but with $\tau \wedge a$ in place of τ , then $\hat{B}_t = \tilde{B}_t$, $0 \leq t \leq a$. Since $\tau \wedge a$ is a bounded stopping time, by the last paragraph, \hat{B} has the same law as B . Thus, \tilde{B}_t , $0 \leq t \leq a$, has the same law as B_t , $0 \leq t \leq a$. This completes the proof. \square

Proof of Proposition 11.13. Since $M_0 = B_0 = 0$, the statement is trivial if $t = 0$. Suppose $t > 0$. We are going to find the joint distribution of (M_t, B_t) . Note that (M_t, B_t) takes values in $\{(x, y) \in \mathbb{R}^2 : x \geq y \vee 0\}$. It suffices to know $\mathbb{P}[M_t \geq x, B_t \leq y]$ for any pair (x, y) such that $x \geq y \vee 0$.

Fix $x, y \in \mathbb{R}$ such that $x \geq y \vee 0$. Let $\tau = \inf\{s : B_s = x\}$. Then τ is a stopping time, and $M_t \geq x$ iff $\tau \leq t$. Define \tilde{B} as in Lemma 11.14. If $M_t \geq x$ and $B_t \leq y$, then $\tilde{B}_t = 2x - B_t \geq 2x - y$. On the other hand, if $\tilde{B}_t \geq 2x - y \geq x$, then we must have $\tau \leq t$ because if $\tau > t$, then $M_t \geq B_t = \tilde{B}_t \geq x$, which contradicts that $\tau > t$. Thus, $B_t = 2x - \tilde{B}_t \leq y$. So we have $\mathbb{P}[M_t \geq x, B_t \leq y] = \mathbb{P}[\tilde{B}_t \geq 2x - y] = \int_{2x-y}^{\infty} \phi_t(s) ds$, where $\phi_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}$ is the density function of $\tilde{B}_t \sim N(0, t)$. We may rewrite $\mathbb{P}[M_t \geq x, B_t \leq y]$ as

$$\int_{2x-y}^{\infty} \phi_t(s) ds = \int_{-\infty}^y \phi_t(2x - b) db = \int_{-\infty}^y \int_x^{\infty} -2\phi_t'(2a - b) da db.$$

So (M_t, B_t) has a joint density in \mathbb{R}^2 , which is $\mathbf{1}_{\{x \geq y \vee 0\}}(-2\phi_t'(2x - y))$. By changing variables, we know that $(M_t, M_t - B_t)$ has density $\mathbf{1}_{\{x, y \geq 0\}}(-2\phi_t'(x + y))$. Thus, both M_t and $M_t - B_t$ have the same density, which is 0 on $(-\infty, 0)$, and equals $\int_0^{\infty} -2\phi_t'(x + y) dy = 2\phi_t(x)$ at $x > 0$. This is the density function for $|B_t|$. \square

Omit Lemma 11.15, Theorem 11.16, Theorem 11.17.

Theorem 11.18 (laws of iterated logarithm). *For a Brownian motion B , we have a.s.*

$$\limsup_{t \rightarrow 0^+} \frac{B_t}{\sqrt{2t \log \log(1/t)}} = \limsup_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \log \log(t)}} = 1.$$

Here $\log \log(t) = \log(\log(t))$, which is positive when $t > e$.

Lemma . For any $x > 0$,

$$\frac{1}{x}e^{-\frac{x^2}{2}} > \int_x^\infty e^{-\frac{u^2}{2}} du > \frac{x}{x^2 + 1}e^{-\frac{x^2}{2}}.$$

Proof. For the first inequality, note that

$$\int_x^\infty e^{-\frac{u^2}{2}} du < \int_x^\infty \frac{u}{x}e^{-\frac{u^2}{2}} du = \frac{1}{x}e^{-\frac{x^2}{2}}.$$

For the second inequality, define

$$f(x) = xe^{-x^2/2} - (x^2 + 1) \int_x^\infty e^{-u^2/2} du.$$

Observe that $\lim_{x \rightarrow \infty} f(x) = 0$ and for $x > 0$,

$$\begin{aligned} f'(x) &= e^{-x^2/2} - x^2e^{-x^2/2} - 2x \int_x^\infty e^{-u^2/2} du + (x^2 + 1)e^{-x^2/2} \\ &= -2x \left(\int_x^\infty e^{-u^2/2} du - \frac{e^{-x^2/2}}{x} \right) > 0, \end{aligned}$$

where we used the first inequality at the last step. So $f < 0$ on $[0, \infty)$, and we get the second inequality. \square

Corollary . Almost surely, $\limsup_{t \rightarrow \infty} B_t = \infty$, $\liminf_{t \rightarrow \infty} B_t = -\infty$, and for any $x \in \mathbb{R}$, the level set $\{t \geq 0 : B_t = x\}$ is unbounded. So we say that B is recurrent.

Proof. That a.s. $\limsup_{t \rightarrow \infty} B_t = \infty$ follows from a.s. $\limsup_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \log \log(t)}} = 1$. Since $-B \sim B$, we have a.s. $\limsup_{t \rightarrow \infty} (-B_t) = \infty$, which implies that $\liminf_{t \rightarrow \infty} B_t = -\infty$. Then almost surely both equalities hold, and when this happens, for any $x \in \mathbb{R}$, $\{t \geq 0 : B_t = x\}$ is unbounded by intermediate value theorem for continuous functions. \square

For example, if we define τ to be the first time that B_t visits some $x \in \mathbb{R}$. Then a.s. τ is finite. If $I = [a, b]$ with $a < 0 < b$, and τ is the first time that B exits I , then τ is a finite stopping time. We have $B_\tau = a$ if B visits a before b , and $B_\tau = b$ if otherwise. Since B is a martingale, $B_t^\tau := B_{\tau \wedge t}$ is a bounded martingale. By Optional Stopping Theorem (Theorem 6.29), $0 = B_0 = \mathbb{E}[B_\tau^\tau | \mathcal{F}_0] = \mathbb{E}[B_\tau] = a\mathbb{P}[B_\tau = a] + b\mathbb{P}[B_\tau = b]$. Since $1 = \mathbb{P}[B_\tau = a] + \mathbb{P}[B_\tau = b]$, we calculate $\mathbb{P}[B_\tau = a] = \frac{b}{b-a}$ and $\mathbb{P}[B_\tau = b] = \frac{-a}{b-a}$.

Corollary . A Brownian motion B is a.s. NOT locally Hölder continuous of exponent $1/2$.

Proof. If B is Hölder continuous of exponent $1/2$ in a neighborhood of 0, then $\limsup_{t \rightarrow 0^+} |B_t - B_0|/\sqrt{t} < \infty$, which implies that $\limsup_{t \rightarrow 0^+} |B_t - B_0|/\sqrt{2t \log \log(t)} = 0$, which almost surely contradicts Theorem 11.18. \square

Corollary . For any fixed $t_0 \geq 0$, B is a.s. not differentiable at t_0 .

Proof. We use the fact that $B_{t_0+t} - B_{t_0}$, $t \geq 0$, has the same distribution as B , and B is a.s. not differentiable at 0 by the law of iterated logarithm. \square

Theorem (Nowhere differentiability). *Almost surely B is not differentiable at any $t \geq 0$.*

Proof. We may consider $B|_{[0,1]}$. In fact, if we have proved that a.s. B is nowhere differentiable on $[0, 1]$, then by scaling, we can conclude that for any $N \in \mathbb{N}$, B is a.s. nowhere differentiable on $[0, N]$. Since there are countably such N , a.s. B is nowhere differentiable on any $[0, N]$, $N \in \mathbb{N}$, and so is nowhere differentiable on \mathbb{R}_+ . If B is differentiable at $t_0 \in [0, 1)$, then

$$\limsup_{h \downarrow 0} \frac{|B_{t_0+h} - B_{t_0}|}{h} < \infty.$$

This means that there are constants $\delta \in (0, 1)$ and $C_1 > 0$ such that if $0 \leq h < \delta$, then $|B_{t_0+h} - B_{t_0}| \leq C_1 h$. Let $R > 0$ be an upper bound of B on $[0, 2]$. Let $M \in \mathbb{N}$ be such that $M \geq C_1 \vee (2R/\delta)$. Then for $h \in [\delta, 1]$, $|B_{t_0+h} - B_{t_0}| \leq 2R \leq M\delta \leq Mh$. So for any $0 \leq h \leq 1$, $|B_{t_0+h} - B_{t_0}| \leq Mh$. Suppose for a fixed $M \in \mathbb{N}$, there exists $t_0 \in [0, 1]$ such that $|B_{t_0+h} - B_{t_0}| \leq Mh$ for all $0 \leq h \leq 1$. Let $n \in \mathbb{N}$ and $n \geq 2$. There is $k \in \{1, 2, \dots, 2^n\}$ such that $t_0 \in [\frac{k-1}{2^n}, \frac{k}{2^n}]$. Then for $1 \leq j \leq 3$,

$$|B_{\frac{k+j}{2^n}} - B_{\frac{k+j-1}{2^n}}| \leq |B_{\frac{k+j}{2^n}} - B_{t_0}| + |B_{\frac{k+j-1}{2^n}} - B_{t_0}| \leq \frac{M(2j+1)}{2^n}.$$

Let E denote the event that B is differentiable at some $t_0 \in [0, 1]$, and define

$$E_{n,k}^M = \{|B_{\frac{k+j}{2^n}} - B_{\frac{k+j-1}{2^n}}| \leq \frac{M(2j+1)}{2^n}, \quad 1 \leq j \leq 3\}.$$

Then

$$E \subset \bigcup_{M=1}^{\infty} \bigcap_{n=2}^{\infty} \bigcup_{k=1}^{2^n} E_{n,k}^M.$$

We have

$$\mathbb{P}[E_{n,k}^M] = \prod_{j=1}^3 \mathbb{P}[|B_{\frac{j}{2^n}}| \leq \frac{M(2j+1)}{2^n}] \leq \mathbb{P}[|B_1| \leq \frac{7M}{2^{n/2}}]^3 \leq C(\frac{7M}{2^{n/2}})^3.$$

Here we used the fact that the density function of B_1 is bounded by $\frac{1}{\sqrt{2\pi}}$. Hence

$$\mathbb{P}[\bigcup_{k=1}^{2^n} E_{n,k}^M] \leq 2^n C (\frac{7M}{2^{n/2}})^3 = 7^3 C M^3 2^{-n/2},$$

which then implies that $\mathbb{P}[\bigcap_{n=2}^{\infty} \bigcup_{k=1}^{2^n} E_{n,k}^M] = 0$. So we get $\mathbb{P}[E] = 0$. \square

Proof of Theorem 11.8. Since $tB_{1/t}$ has the same law as B , the two equalities are equivalent to each other. So it suffices to prove the case $\limsup_{t \rightarrow \infty}$. Let $M_t = \sup_{s \leq t} B_s$ be the maximum process. By the above lemma, for any $x > 0$,

$$\frac{1}{x} \frac{e^{-x^2/2}}{\sqrt{2\pi}} > \mathbb{P}[B_1 > x] > \frac{x}{x^2 + 1} \frac{e^{-x^2/2}}{\sqrt{2\pi}}. \quad (1.1)$$

Since $B_t \stackrel{d}{=} \sqrt{t}B_1$, and $M_t \stackrel{d}{=} |B_t|$, we have

$$\mathbb{P}[M_t > ut^{1/2}] = 2\mathbb{P}[B_t > ut^{1/2}] \leq \frac{2e^{-u^2/2}}{\sqrt{2\pi u}}.$$

Let $h(t) = \sqrt{2t \log \log(t)}$. Fix $R, \rho > 1$. Then $h(\rho^n) = \sqrt{2\rho^n \log(n \log(\rho))}$. Applying the above formula to $t = \rho^n$ and $u = R\sqrt{2 \log(\log(\rho^n))}$ (so $e^{-u^2/2} = (n \log \rho)^{-R^2}$), we get

$$\mathbb{P}[M_{\rho^n} > Rh(\rho^n)] = \mathbb{P}[M_{\rho^n} > \sqrt{\rho^n} R \sqrt{2 \log(\log(\rho^n))}] \leq \frac{2(n \log \rho)^{-R^2}}{\sqrt{2\pi} R \sqrt{2 \log(n \log \rho)}}, \quad n \geq 2.$$

Since $R > 1$, $\sum_{n=2}^{\infty} n^{-R^2} < \infty$. By Borel-Cantelli Lemma, the probability that there exist infinitely many n such that $M_{\rho^n} > Rh(\rho^n)$ is 0, i.e., a.s. there exist N such that for $n > N$, $M_{\rho^n} \leq Rh(\rho^n)$. So a.s. $\limsup M_{\rho^n}/h(\rho^n) \leq R$. Since this holds for any $R > 1$, we get a.s. $\limsup M_{\rho^n}/h(\rho^n) \leq 1$. Since $h(\rho^n)/h(\rho^{n-1}) \rightarrow \sqrt{\rho}$, we get a.s. $\limsup M_{\rho^n}/h(\rho^{n-1}) \leq \sqrt{\rho}$. Since for any $t > 1$ there exist n such that $\rho^{n-1} \leq t \leq \rho^n$, we get $M_t/h(t) \leq M_{\rho^n}/h(\rho^{n-1})$. So a.s. $\limsup_{t \rightarrow \infty} B_t/h(t) \leq \limsup_{t \rightarrow \infty} M_t/h(t) \leq \rho$. Since this holds for any $\rho > 1$, we have a.s. $\limsup_{t \rightarrow \infty} B_t/h(t) \leq 1$.

To prove the reverse inequality, let $R > 1$ and $c = \sqrt{(R-1)/R} < 1$. Since $B_{R^n} - B_{R^{n-1}} \stackrel{d}{=} \sqrt{R^{n-1}(R-1)}B_1$, by (1.1) we have

$$\begin{aligned} \mathbb{P}[B_{R^n} - B_{R^{n-1}} \geq ch(R^n)] &= \mathbb{P}[\sqrt{R^{n-1}(R-1)}B_1 \geq \sqrt{(R-1)/R} \sqrt{2R^n \log \log(R^n)}] \\ &= \mathbb{P}[B_1 \geq \sqrt{2 \log \log(R^n)}] \gtrsim \frac{1/\sqrt{2\pi}}{\log(R^n) \sqrt{\log \log(R^n)}}. \end{aligned}$$

Since the sum of the RHS over n is infinity, and $B_{R^n} - B_{R^{n-1}}$, $n \geq 1$, are independent, By Borel-Cantelli Lemma, the event that $B_{R^n} - B_{R^{n-1}} \geq ch(R^n)$ will happen infinitely often. This implies that a.s.

$$\sqrt{(R-1)/R} = c \leq \limsup (B_t - B_{t/R})/h(t) \leq \limsup B_t/h(t) + \limsup (-B_{t/R})/h(t).$$

By the upper bound, $\limsup (-B_{t/R})/h(t) \leq R^{-1/2}$. Thus, $\limsup B_t/h(t) \geq \sqrt{(R-1)/R} - R^{-1/2}$. Letting $R \rightarrow \infty$, we then get $\limsup B_t/h(t) \geq 1$. \square

In contrast to the law of iterated logarithm, we have the following result.

Theorem (Lévy's modulus of continuity). *Almost surely, we have*

$$\limsup_{h \downarrow 0} \sup_{0 \leq t \leq 1-h} \frac{|B_{t+h} - B_t|}{\sqrt{2h \log(1/h)}} = 1.$$

Proof. This is Theorem 1.14 of *Brownian motion* by Peter Möters and Yuval Peres. □

Omit the part of Chapter 11 after Theorem 11.18.

2 Stochastic Integrals and Quadratic Variation

We first introduce a new object: local martingale. Fix a filtration $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$, which is right-continuous and complete. Recall that for an \mathcal{F} -adapted process X and an \mathcal{F} -stopping time τ , we may define a new process X^τ by $X_t^\tau = X_{t \wedge \tau}$. This is X stopped at τ .

Definition . An \mathcal{F} -adapted process X is called a local martingale if there is a sequence of \mathcal{F} -stopping times τ_n with $\tau_n \uparrow \infty$ such that for each n , $X^{\tau_n} - X_0$ is an \mathcal{F} -martingale. The sequence (τ_n) is called a localizing sequence. If in addition, X is sample-wise continuous, then it is called a continuous local martingale.

Remark . We have the following facts. Recall that by Optional Stopping Theorem (Theorem 6.29), if M is a continuous martingale, and τ is a stopping time, then the stopped process M^τ is also a martingale.

1. If X_0 is integrable, $X^{\tau_n} - X_0$ is an \mathcal{F} -martingale iff X^{τ_n} is an \mathcal{F} -martingale.
2. A martingale is a local martingale. We may simply take all $\tau_n = \infty$.
3. A uniformly bounded local martingale is a martingale. This follows from Dominated Convergence Theorem. To see this, suppose X is a uniformly bounded local martingale with localizing sequence (τ_n) , then for any $t \geq s \geq 0$ and $n \in \mathbb{N}$, $\mathbb{E}[X_{t \wedge \tau_n} | \mathcal{F}_s] = X_{s \wedge \tau_n}$. Letting $n \rightarrow \infty$, we get $\mathbb{E}[X_t | \mathcal{F}_s] = X_s$.
4. A local martingale may not be a martingale. A concrete example will be given later. Sometimes we call a martingale a true martingale to emphasize.
5. For a continuous process X , the condition that $\tau_n \uparrow \infty$ can be slightly weakened to a.s. $\tau_n \rightarrow \infty$. First, a null event does not affect the martingale property. Second, if τ_n is a sequence tending to ∞ , then we may define an *increasing* sequence of stopping times tending to ∞ by $\tau'_n = \inf_{m \geq n} \tau_m$. Since \mathcal{F} is right-continuous, each τ'_n is a stopping time. Whenever X^{τ_n} is a martingale, $X^{\tau'_n} = (X^{\tau_n})^{\tau'_n}$ is also a martingale. Here we use the optional stopping theorem.
6. A continuous local martingale stopped at any stopping time is still a local martingale. In fact, if X is a continuous local martingale with localizing sequence (τ_n) , and if τ is any stopping time, then for any n , $(X^\tau)^{\tau_n} - X_0^\tau = (X^{\tau_n})^\tau - X_0$ is a martingale. So (τ_n) is localizing sequence for X^τ .

7. If X is an \mathcal{F} -adapted continuous process, and if there exists an increasing sequence of \mathcal{F} -stopping times τ_n with $\tau_n \uparrow \infty$ such that for each n , X^{τ_n} is a local martingale, then X is a local martingale. To see this, suppose $X_0 = 0$. For each n , let $(\sigma_k^n)_{k \geq 1}$ be the localizing sequence of X^{τ_n} . For each n , since $\sigma_k^n \rightarrow \infty$, we may choose k_n such that $\mathbb{P}[\sigma_{k_n}^n < n] < \frac{1}{2^n}$. By Borel-Cantellie lemma, a.s. $\sigma_{k_n}^n \rightarrow \infty$. Let $\tau'_n = \tau_n \wedge \sigma_{k_n}^n$. Then a.s. $\tau'_n \rightarrow \infty$ and for each n , $X^{\tau'_n} = (X^{\tau_n})^{\sigma_{k_n}^n}$ is a true martingale.
8. The sum of two continuous local martingales is a local martingale. If X and Y are both continuous local martingales with localizing sequences (τ_n) and (σ_n) , then $\tau_n \wedge \sigma_n \uparrow \infty$, and each $\tau_n \wedge \sigma_n$ localizes both X and Y , and so also localizes $X + Y$. Thus, $X + Y$ is also a local martingale.
9. Suppose X is a continuous local martingale with $X_0 = 0$. Then we have a natural choice of localizing sequence: $\tau_n := \inf\{t \geq 0 : |X_t| \geq n\}$. By convention, we set $\inf \emptyset = \infty$. Then $\tau_n \uparrow \infty$, and for each n , $|X^{\tau_n}|$ is bounded by n , and so X^{τ_n} is a true martingale.

Proposition 15.2. *If X is a continuous local martingale with locally finite variation, then a.s. $X_t = X_0$ for all $t \geq 0$.*

Proof. By considering $X - X_0$, we may assume $X_0 = 0$. We first prove the proposition assuming that X has finite total variation on $[0, \infty)$, which is uniformly bounded by V . Then X itself is bounded in absolute value by V . So X is a true martingale. Fix $t > 0$ and a partition $\Delta = \{0 = t_0 < \dots < t_n = t\}$ of $[0, t]$. Then

$$X_t^2 = \sum_{k=1}^n (X_{t_k} - X_{t_{k-1}})^2 + 2 \sum_{1 \leq j < k \leq n} (X_{t_j} - X_{t_{j-1}})(X_{t_k} - X_{t_{k-1}}).$$

By the martingale property of X , for each $1 \leq j < k \leq n$,

$$\mathbb{E}[(X_{t_j} - X_{t_{j-1}})(X_{t_k} - X_{t_{k-1}}) | \mathcal{F}_{t_{k-1}}] = 0$$

because $\mathbb{E}[X_{t_k} - X_{t_{k-1}} | \mathcal{F}_{t_{k-1}}] = 0$ and $X_{t_j} - X_{t_{j-1}} \in \mathcal{F}_{t_{k-1}}$. So we get $\mathbb{E}[X_t^2] = \mathbb{E}[Q_\Delta]$, where

$$Q_\Delta := \sum_{k=1}^n (X_{t_k} - X_{t_{k-1}})^2 \leq V \cdot \sup_{s_1, s_2 \in [0, t] : |s_1 - s_2| \leq |\Delta|} |X_{s_1} - X_{s_2}|,$$

where $|\Delta| = \max_{1 \leq k \leq n} |t_k - t_{k-1}|$ is the mesh size of Δ . Since X is continuous, if the mesh size of Δ tends to 0, then the RHS of the above formula tends to 0. By Dominated Convergence Theorem, $\mathbb{E}[Q_\Delta] \rightarrow 0$ as $|\Delta| \rightarrow 0$. Since $\mathbb{E}[X_t^2] = \mathbb{E}[Q_\Delta]$ holds for any partition Δ on $[0, t]$, we get a.s. $X_t = 0$. Thus, a.s. for every $t \in \mathbb{Q}_+$, $X_t = 0$. By continuity of X , we get a.s. $X \equiv 0$.

For the general case, let V_t be the total variation of X on $[0, t]$. Then V is continuous, nondecreasing, and adapted. For each $n \in \mathbb{N}$, let τ_n be the first time that $V_t \geq n$. If such time does not exist, then $\tau_n = \infty$. This is a stopping time. Then X^{τ_n} has total variation on $[0, \infty)$ bounded by n , and so itself is also bounded by n , and so is a true martingale. From the last paragraph, we know that a.s. $X_n^\tau \equiv 0$. Since $\tau_n \rightarrow \infty$, we then get a.s. $X \equiv 0$. \square

Definition . A predictable step process has the form

$$V_t = \sum_{j=0}^{\infty} \zeta_j \mathbf{1}_{(\tau_j, \infty)}(t) = \sum_{k=0}^{\infty} \eta_k \mathbf{1}_{(\tau_k, \tau_{k+1}]}(t), \quad t \geq 0, \quad (2.1)$$

where (τ_k) is a sequence of stopping times with $\tau_0 = 0$ and $\tau_k \uparrow \infty$, and ζ_k, η_k is \mathcal{F}_{τ_k} -measurable for each k . Note that $\eta_k = \sum_{j=0}^k \zeta_j$. Also note that V is adapted. To see this, fix any $t \geq 0$. For every k , the event $\{\tau_k < t \leq \tau_{k+1}\}$ is \mathcal{F}_t -measurable. It suffices to show that V_t restricted to this event is \mathcal{F}_t -measurable, which follows from the facts that $\eta_k \in \mathcal{F}_{\tau_k}$ and $\mathcal{F}_{\tau_k} \cap \{\tau_k < t\} \subset \mathcal{F}_t$.

Definition . Given a predictable step process with the form (2.1) and an adapted process X , we define the elementary integral process $V \cdot X$ as in Chapter 6 by

$$(V \cdot X)_t \equiv \int_0^t V dX = \sum_{k=0}^{\infty} \eta_k (X_{t \wedge \tau_{k+1}} - X_{t \wedge \tau_k}) = \sum_{j=0}^{\infty} \zeta_j (X_t - X_{t \wedge \tau_j}).$$

Note that $(V \cdot X)_0 = 0$. To see that the last equality holds, and this is a finite sum, note that if $t \in (\tau_n, \tau_{n+1}]$, then

$$\begin{aligned} \sum_{k=0}^{\infty} \eta_k (X_{t \wedge \tau_{k+1}} - X_{t \wedge \tau_k}) &= \sum_{k=0}^{n-1} \eta_k (X_{\tau_{k+1}} - X_{\tau_k}) + \eta_n (X_t - X_{\tau_n}); \\ \sum_{j=0}^{\infty} \zeta_j (X_t - X_{t \wedge \tau_j}) &= \sum_{j=1}^n \zeta_j (X_t - X_{\tau_j}) = \eta_n X_t - \sum_{j=1}^n (\eta_j - \eta_{j-1}) X_{\tau_j} \\ &= \eta_n (X_t - X_{\tau_n}) + \sum_{j=0}^{n-1} \eta_j (X_{\tau_{j+1}} - X_{\tau_j}). \end{aligned}$$

Moreover, $V \cdot X$ is adapted because the RHS of the above formulas restricted to the event $\tau_n < t$ is \mathcal{F}_t -measurable. Also note that if X is continuous, then so is $V \cdot X$ because by definition and the continuity of X , $V \cdot X$ is continuous on $[\tau_n, \tau_{n+1}]$ for each n .

Lemma . If M is a martingale, and (τ_n) is an increasing sequence of stopping times with $\tau_0 = 0$ and $\tau_n \uparrow \infty$, then for for any $t \geq 0$,

$$\mathbb{E}[M_t^2] = \sum_{k=0}^{\infty} \mathbb{E}[(M_{t \wedge \tau_{k+1}} - M_{t \wedge \tau_k})^2].$$

Proof. We may write

$$M_{t \wedge \tau_n} = \sum_{k=0}^{n-1} (M_{t \wedge \tau_{k+1}} - M_{t \wedge \tau_k}).$$

So

$$M_{t \wedge \tau_n}^2 = \sum_{k=0}^{n-1} (M_{t \wedge \tau_{k+1}} - M_{t \wedge \tau_k})^2 + 2 \sum_{0 \leq j < k \leq n-1} (M_{t \wedge \tau_{j+1}} - M_{t \wedge \tau_j})(M_{t \wedge \tau_{k+1}} - M_{t \wedge \tau_k}).$$

For each $j < k$,

$$\mathbb{E}[(M_{t \wedge \tau_{j+1}} - M_{t \wedge \tau_j})(M_{t \wedge \tau_{k+1}} - M_{t \wedge \tau_k}) | \mathcal{F}_{\tau_k}] = 0,$$

which follows from Optional Stopping Theorem and the fact that $(M_{t \wedge \tau_{j+1}} - M_{t \wedge \tau_j}) \in \mathcal{F}_{\tau_k}$. Here we used the fact that $M_{t \wedge \tau_{j+1}} - M_{t \wedge \tau_j}$ and $M_{t \wedge \tau_{k+1}} - M_{t \wedge \tau_k}$ are in L^2 . So we get

$$\mathbb{E}M_{t \wedge \tau_n}^2 = \sum_{k=0}^{n-1} \mathbb{E}[(M_{t \wedge \tau_{k+1}} - M_{t \wedge \tau_k})^2].$$

Since as $n \rightarrow \infty$, $M_{t \wedge \tau_n} \rightarrow M_t$, by Corollary 6.22 and Proposition 3.12,

$$\mathbb{E}M_t^2 = \lim_{n \rightarrow \infty} \mathbb{E}M_{t \wedge \tau_n}^2 = \sum_{k=0}^{\infty} \mathbb{E}[(M_{t \wedge \tau_{k+1}} - M_{t \wedge \tau_k})^2].$$

□

Lemma 15.3. *Let V be a predictable process with $|V| \leq C$. Let M be a continuous L^2 -martingale (i.e., $\mathbb{E}[M_t^2] < \infty$ for every $t \geq 0$) with $M_0 = 0$. Then $V \cdot M$ is still a continuous L^2 martingale; and for any $t \geq 0$, $\mathbb{E}[(V \cdot M)_t^2] \leq C^2 \mathbb{E}[M_t^2]$.*

Proof. The continuity of $V \cdot M$ follows from the definition and the continuity of M . Recall Corollary 6.14: if M is a martingale, τ is a stopping time that takes countably many values, and η is a bounded \mathcal{F}_τ -measurable random variable, then the process $N_t := \eta(M_t - M_{t \wedge \tau})$ is again a martingale. For a continuous martingale M , we can remove the assumption that τ takes countably many values because we may find a sequence of stopping times $\tau^n \downarrow \tau$ such that each τ^n takes countably many values. Then $\mathbb{E}[\eta(M_t - M_{t \wedge \tau^n}) | \mathcal{F}_s] = \eta(M_s - M_{s \wedge \tau^n})$ for any $t \geq s \geq 0$. Letting $n \rightarrow \infty$ we get $\mathbb{E}[\eta(M_t - M_{t \wedge \tau})] = \eta(M_s - M_{s \wedge \tau})$ using the continuity of M and the uniform integrability of $\{\eta M_{t \wedge \tau^n} : n \in \mathbb{N}\}$. Here we use the fact that $M_{t \wedge \tau^n} = \mathbb{E}[M_t | \mathcal{F}_{t \wedge \tau^n}]$. Suppose V has the form of (2.1). Suppose there are only finitely many k such that $\zeta_k \neq 0$, then $V \cdot M$ is a martingale since it is a finite sum of $\zeta_j(M_t - M_{t \wedge \tau_j})$, each of which is a martingale. Fix any $t \geq 0$, by the above lemma, we have

$$\begin{aligned} \mathbb{E}[(V \cdot M)_t^2] &= \sum_{j=0}^{\infty} \mathbb{E}[(V \cdot M)_{t \wedge \tau_{j+1}} - (V \cdot M)_{t \wedge \tau_j}]^2 \\ &= \sum_{j=0}^{\infty} \mathbb{E}[\eta_j^2 (M_{t \wedge \tau_{j+1}} - M_{t \wedge \tau_j})^2] \leq C^2 \sum_{j=0}^{\infty} \mathbb{E}[(M_{t \wedge \tau_{j+1}} - M_{t \wedge \tau_j})^2] = C^2 \mathbb{E}[M_t^2]. \end{aligned}$$

Thus, $V \cdot M$ is also an L^2 -martingale. Finally, we consider the general case, i.e., there may be infinitely many k such that $\zeta_k \neq 0$. We write $V_t^n = \sum_{k=0}^n \zeta_k \mathbf{1}_{(\tau_k, \infty)}(t)$. Then each V^n is a predictable process with finitely many nonzero ζ_k . So $\mathbb{E}[(V^n \cdot M)_t^2] \leq \mathbb{E}[M_t^2]$. Since $(V^n \cdot M)_t \rightarrow (V \cdot M)_t$, by Fatou's lemma, $\mathbb{E}[(V \cdot M)_t^2] \leq C^2 \mathbb{E}[M_t^2]$. Fix $t \geq s \geq 0$. For each $n \in \mathbb{N}$, $V^n \cdot M$ is a martingale, so $\mathbb{E}[(V^n \cdot M)_t | \mathcal{F}_s] = (V^n \cdot M)_s$. Since the sequence $(V^n \cdot M)_t$ is L^2 -bounded, it is uniformly integrable. So the a.s. convergence implies L^1 -convergence. Letting $n \rightarrow \infty$, we get $\mathbb{E}[(V \cdot M)_t | \mathcal{F}_s] = (V \cdot M)_s$. So $V \cdot M$ is a martingale. \square

Recall the following facts from Theorem 6.18, Theorem 6.21, Corollary 6.22, and Theorem 6.29. If M is a uniformly integrable martingale (i.e., $(M_t)_{t \geq 0}$ is uniformly integrable), then $M_\infty := \lim_{t \rightarrow \infty} M_t$ converges a.s. and in L^1 , and for any stopping time τ , $\mathbb{E}[M_\infty | \mathcal{F}_\tau] = M_\tau$. In particular, we have $\|M_\tau\|_{L^p} \leq \|M_\infty\|_{L^p}$ for any $p \geq 1$, and $\mathbb{E}[M_\infty] = \mathbb{E}[M_0]$. Also recall that for any $p > 1$, if M is L^p -bounded, i.e., $\sup_{t \geq 0} \|M_t\|_{L^p} < \infty$, then M is uniformly integrable. In that case, we have $M_t \rightarrow M_\infty$ in L^p as $t \rightarrow \infty$. In particular, if M is uniformly bounded, it is uniformly integrable, and $M_t \rightarrow M_\infty$ in L^p for every $p \geq 1$.

Let \mathcal{M}^2 denote the space of all L^2 -bounded continuous martingale M with $M_0 = 0$. Equip \mathcal{M}^2 with the norm $\|M\|_{\mathcal{M}^2} = \|M_\infty\|_{L^2} = \|\lim_{t \rightarrow \infty} M_t\|_{L^2}$. Then \mathcal{M}^2 is clearly an inner product space with $\langle M, N \rangle = \mathbb{E}[M_\infty N_\infty]$. Recall that by Proposition 6.16, for any $t \geq 0$, $\|\sup_{0 \leq s \leq t} |M_t|\|_{L^2} \leq 2\|M_t\|_{L^2}$. Since $\|M_t\|_{L^2} \uparrow \|M_\infty\|_{L^2} = \|M\|$, we get $\|M_\infty^*\|_{L^2} \leq 2\|M\|_{\mathcal{M}^2}$, where $M_\infty^* := \sup_{t \geq 0} |M_t|$.

Lemma 15.4. \mathcal{M}^2 is a Hilbert space.

Proof. Let (M^n) be a Cauchy sequence in \mathcal{M}^2 . Since $\|(M^n - M^m)_\infty^*\|_{L^2} \leq 2\|M_\infty^n - M_\infty^m\|$, by choosing a subsequence, we may assume that

$$\mathbb{E}[\sup_{t \geq 0} |M_t^n - M_t^{n+1}|^2] < 2^{-3n},$$

which implies that

$$\mathbb{P}[\sup_{t \geq 0} |M_t^n - M_t^{n+1}| > 2^{-n}] < 2^{-n}.$$

By Borel-Cantelli lemma, a.s. there is (a random) N such that for $n > N$, $\sup_{t \geq 0} |M_t^n - M_t^{n+1}| \leq 2^{-n}$. Let E denote the event that such N exists. Then $\mathbb{P}[E] = 1$, and (M^n) is uniformly Cauchy on \mathbb{R}_+ on the event E . Let $M = \lim_{n \rightarrow \infty} M^n$ on E , and $M = 0$ on E^c . Then M is continuous on \mathbb{R}_+ . Since a.s. $M_t^n \rightarrow M_t$ for all $t \geq 0$, each M^n is adapted, and \mathcal{F} is complete, M is also adapted. The martingale property of M follows from the martingale property of each M^n , and the uniform integrability of $(M_t^n)_{n \in \mathbb{N}}$ for every fixed t . In fact, for any fixed $t \geq s \geq 0$, $\mathbb{E}[M_t^n | \mathcal{F}_s] = M_s^n$. The sequence M_t^n is uniformly integrable because their L^2 norms are uniformly bounded. Letting $n \rightarrow \infty$, we then get $\mathbb{E}[M_t | \mathcal{F}_s] = M_s$. The L^2 -boundedness of M follows from Fatou's lemma: $\mathbb{E}[M_t^2] \leq \liminf \mathbb{E}[(M_t^n)^2] \leq \liminf \mathbb{E}[(M_\infty^n)^2]$ for all $t \geq 0$. Finally, in the formula $\lim_{n, m \rightarrow \infty} \mathbb{E}[\sup_{t \geq 0} |M_t^n - M_t^m|^2] \rightarrow 0$, if we fix n , let $m \rightarrow \infty$, and use Fatou's lemma, we then get $\lim_{n \rightarrow \infty} \mathbb{E}[\sup_{t \geq 0} |M_t^n - M_t|^2] \rightarrow 0$, and so $M^n \rightarrow M$. \square

Theorem 15.5. *For any continuous local martingales M and N , there a.s. exists a unique continuous process $[M, N]$ of locally finite variation and with $[M, N]_0 = 0$ such that $NM - [N, M]$ is a local martingale. The form $[M, N]$ is a.s. symmetric and bilinear with $[M - M_0, N - N_0] = [M, N]$. Furthermore, for any stopping time τ , a.s. $[M, N]^\tau = [M^\tau, N^\tau] = [M^\tau, N]$. The process $[M] := [M, M]$ is a.s. nondecreasing; and if M is bounded with $M_0 = 0$, then $M^2 - [M] \in \mathcal{M}^2$.*

We call $[M]$ the quadratic variation of M , and $[M, N]$ the quadratic covariation of M and N . The name comes from the following fact (Proposition 15.18): Fix $t > 0$. For any partition $\Delta = \{0 = t_0 < t_1 < \dots < t_n = t\}$ of $[0, t]$, define

$$T^\Delta(M, N) = \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})(N_{t_k} - N_{t_{k-1}}).$$

Then for any sequence of partitions (Δ_n) of $[0, t]$ with mesh size $|\Delta_n| \rightarrow 0$, we have $T^\Delta(M, N) \xrightarrow{P} [M, N]_t$. For example, if M is the Brownian motion B , then $[B]_t = t$ for any $t \geq 0$.

Lemma . *For a continuous local martingale M , a stopping time τ , and an \mathcal{F}_τ -measurable random variable ζ , $\zeta(M - M^\tau)$ is a continuous local martingale.*

Proof. The continuity is obvious. We know that $M - M^\tau$ is a local martingale. We construct localizing sequence σ_n for $\zeta(M - M^\tau)$ as follows. If $|\zeta| > n$, then $\sigma_n = \tau$; if $|\zeta| \leq n$, then $\sigma_n = \inf\{t \geq \tau : |M_t - M_\tau| > n\}$. Then $\sigma_n \uparrow \infty$, and each σ_n is a stopping time because for any $t \geq 0$,

$$\{\sigma_n < t\} = (\{\tau < t\} \cap \{|\zeta| > n\}) \cup \bigcup_{q \in [0, t) \cap \mathbb{Q}} (\{\tau \leq q\} \cap \{|M_q - M_\tau| > n\}) \in \mathcal{F}_t.$$

Since $(M - M^\tau)^{\sigma_n}$ is a local martingale bounded by n , it is a true martingale, which vanishes on $[0, \tau]$. If $|\zeta| > n$, $\sigma_n = \tau$, and so $(\zeta(M - M^\tau))^{\sigma_n} = \zeta(M^\tau - M^\tau) = 0$. Thus,

$$(\zeta(M - M^\tau))^{\sigma_n} = (\mathbf{1}_{|\zeta| \leq n} \zeta)(M - M^\tau)^{\sigma_n}$$

is a martingale by Corollary 6.14 because $\mathbf{1}_{|\zeta| \leq n} \zeta$ is bounded and \mathcal{F}_τ -measurable. \square

Proof of Theorem 15.5. If there are two continuous processes V^1 and V^2 with locally finite variation such that $V_0^j = 0$ and $NM - V^j$ is a local martingale. Then $V^1 - V^2$ is a continuous local martingale with locally finite variation and $(V^1 - V^2)_0 = 0$. By Proposition 15.2, a.s. $V^1 - V^2 \equiv 0$. So we get the uniqueness.

If $[M, N]$ exists, then for any stopping time τ ,

$$(MN - [M, N])^\tau = M^\tau N^\tau - [M, N]^\tau \tag{2.2}$$

is a local martingale. By the uniqueness, we get $[M^\tau, N^\tau] = [M, N]^\tau$. By the lemma, $M^\tau N - M^\tau N^\tau = M^\tau(N - N^\tau) = M_\tau(N - N^\tau)$ is a local martingale. Combining it with (2.2), we see that $M^\tau N - [M, N]^\tau$ is a local martingale, which implies that $[M^\tau, N] = [M, N]^\tau$.

We may assume that $M_0 = N_0 = 0$. In fact, if we know that $[M, N]$ exists whenever $M_0 = N_0 = 0$, then for general M, N , $[M - M_0, N - N_0]$ exists, i.e., $(M - M_0)(N - N_0) - [M - M_0, N - N_0]$ is a local martingale. By the lemma applied to $\tau = 0$, we see that $M_0(N - N_0)$ and $N_0(M - M_0)$ are local martingales, which then implies that $MN - M_0N_0 - [M - M_0, N - N_0]$ is a local martingale. Thus, $[M, N]$ exists and a.s. equals $[M - M_0, N - N_0]$.

To prove the existence and properties of $[M, N]$, by polarization, it suffices to prove the existence, continuity, and monotonicity of $[M]$. In fact, if $(M + N)^2 - [M + N]$ and $(M - N)^2 - [M - N]$ are local martingales, then $4MN - ([M + N] - [M - N])$ is a local martingale, and so $[M, N] = ([M + N] - [M - N])/4$. Since $[M + N]$ and $[M - N]$ are continuous and nondecreasing, $[M, N]$ is continuous and has locally finite total variation. The symmetry and bilinear property of $[M, N]$ is obvious.

We first assume that $|M|$ is bounded by $C < \infty$. Then M is a true martingale. Fix $n \in \mathbb{N}$. Define a sequence of stopping times $(\tau_k^n)_{k \geq 0}$ such that $\tau_0^n = 0$ and for each $k \in \mathbb{N}$,

$$\tau_k^n = \inf\{t \geq \tau_{k-1}^n : |M_t - M_{\tau_k^n}| \geq 2^{-n}\}.$$

As usual, if such time does not exist, then $\tau_k^n = \infty$, and so are all τ_j^n , $j \geq k$. We have $\tau_k^n \uparrow \infty$. Introduce two processes

$$V_t^n = \sum_{k=0}^{\infty} M_{\tau_k^n} \mathbf{1}_{(\tau_k^n, \tau_{k+1}^n]}(t), \quad Q_t^n = \sum_{k=0}^{\infty} (M_{t \wedge \tau_{k+1}^n} - M_{t \wedge \tau_k^n})^2. \quad (2.3)$$

This means that if $t \in (\tau_k^n, \tau_{k+1}^n]$, then $V_t^n = M_{\tau_k^n}$ and $Q_t^n = \sum_{j=0}^{k-1} (M_{\tau_{j+1}^n} - M_{\tau_j^n})^2 + (M_t - M_{\tau_k^n})^2$. Note that V^n is a bounded predictable step process. We have

$$M_t^2 = 2(V^n \cdot M)_t + Q_t^n, \quad t \geq 0.$$

To see this, note that if $t \in (\tau_k^n, \tau_{k+1}^n]$,

$$M_t^2 - M_{\tau_k^n}^2 = 2M_{\tau_k^n}(M_t - M_{\tau_k^n}) + (M_t - M_{\tau_k^n})^2;$$

$$M_{\tau_{j+1}^n}^2 - M_{\tau_j^n}^2 = 2M_{\tau_j^n}(M_{\tau_{j+1}^n} - M_{\tau_j^n}) + (M_{\tau_{j+1}^n} - M_{\tau_j^n})^2, \quad 0 \leq j \leq k-1.$$

By Lemma 15.3, each $V^n \cdot M$ is a continuous L^2 -martingale. Since $|V^n - M| \leq 2^{-n}$ by the definition of (τ_k^n) , we have for $n \geq m$,

$$\|V^n \cdot M - V^m \cdot M\|_{\mathcal{M}^2} = \|(V^n - V^m) \cdot M\|_{\mathcal{M}^2} \leq 2^{1-m} \|M\|_{\mathcal{M}^2} \rightarrow 0, \quad \text{as } n, m \rightarrow \infty.$$

By the completeness of \mathcal{M}^2 , there is $N \in \mathcal{M}^2$ such that $V^m \cdot M \rightarrow N$. So we get

$$\|(V^n \cdot M - N)_\infty^*\|_{L^2} \leq 2\|V^n \cdot M - N\|_{\mathcal{M}^2} \rightarrow 0.$$

The process $[M] := M^2 - 2N$ is continuous, $M^2 - [M] = 2N \in \mathcal{M}^2$, and by L^2 -convergence we have

$$(Q^n - [M])_\infty^* = 2(N - V^n \cdot M)_\infty^* \xrightarrow{\text{P}} 0.$$

Thus, there is a subsequence (Q^{n_k}) of (Q^n) , which a.s. converges uniformly to $[M]$. We claim that $[M]$ is a.s. nondecreasing. Since $[M]$ is continuous, it suffices to show that for any $p < q \in \mathbb{Q}_+$, a.s. $[M]_p < [M]_q$. There are two cases: Case (i) M is constant on $[p, q]$. In this case, for any $n \geq 1$, $[p, q]$ lies in one interval $[\tau_k^n, \tau_{k+1}^n]$, and Q^n is constant on $[p, q]$. By the a.s. uniform convergence of (Q^{n_k}) to $[M]$, we conclude that a.s. $[M]$ is constant on $[p, q]$. Case (ii) M is not constant on $[p, q]$. When n is big enough, we have $2^{-n} < \max M([p, q]) - \min M([p, q])$. Then there exists k such that $\tau_k^n \in [p, q]$. From the definition of Q^n , we see that $Q_p^n \leq Q_{\tau_k^n}^n \leq Q_q^n$. By the a.s. uniform convergence of (Q^{n_k}) to $[M]$, we conclude that a.s. $[M]_p \leq [M]_q$ in Case (ii). So we get a.s. $[M]_p \leq [M]_q$ as desired. So $[M]$ is a.s. nondecreasing.

Now we do not assume that M is bounded. Recall that $M_0 = 0$. Let τ_n be the first time such that $|M_t| \geq n$. Then $\tau_n \uparrow \infty$, and for each n , M^{τ_n} is a bounded martingale. So there exists a continuous nondecreasing process $[M^{\tau_n}]$ starting from 0 such that $(M^{\tau_n})^2 - [M^{\tau_n}]$ is a martingale. For $n \leq m$, since $M^{\tau_n} = (M^{\tau_m})^{\tau_n}$, we have a.s. $[M^{\tau_n}] = [M^{\tau_m}]^{\tau_n}$. So we may define a process $[M]$ such that for any n , a.s. $[M]^{\tau_n} = [M^{\tau_n}]$. In fact, we may define $[M]_t = [M^{\tau_n}]_t$ if $t \leq \tau_n$ on the event that $[M^{\tau_n}] = [M^{\tau_m}]^{\tau_n}$ for any $n \leq m$, which has probability 1. Then for any n , $(M^2 - [M])^{\tau_n} = (M^{\tau_n})^2 - [M^{\tau_n}]$ is a martingale. So $M^2 - [M]$ is a local martingale. \square

The quadratic covariation will be used to define the stochastic integral $V \cdot M$, where M is a continuous local martingale, and V is an adapted process satisfying some properties. The $V \cdot M$ is a continuous local martingale that satisfies $[V \cdot M, N] = V \cdot [M, N]$, where N is any other continuous local martingale, and $V \cdot [M, N]$ is the Stieltjes integral of V against the measure generated by $[M, N]$, i.e., $(V \cdot [M, N])_t = \int_0^t V_s d[M, N]_s$, $t \geq 0$.

One equality we will often use is: if $\sigma \leq \tau$ are two stopping times, then

$$[M^\tau - M^\sigma] = [M^\tau] + [M^\sigma] - 2[M^\sigma, M^\tau] = [M]^\tau + [M]^\sigma - [M, M]^{\sigma \wedge \tau} = [M]^\tau - [M]^\sigma.$$

Since $[M]$ is nondecreasing, we define $[M]_\infty = \lim_{t \rightarrow \infty} [M]_t$. From Theorem 15.5, we know that if M is uniformly bounded with $M_0 = 0$, then $M^2 - [M] \in \mathcal{M}^2$. So we get $\mathbb{E}[M_\infty^2 - [M]_\infty] = \mathbb{E}[M_0^2 - [M]_0] = 0$, and

$$\|M\|_{\mathcal{M}^2}^2 = \mathbb{E}[M_\infty^2] = \mathbb{E}[M]_\infty = \|[M]_\infty\|_{L^1}. \quad (2.4)$$

We now extend this result to more general case.

Lemma . *For a continuous local martingale M with $M_0 = 0$, we have $M \in \mathcal{M}^2$ iff $[M]_\infty \in L^1$, and then $\|M\|_{\mathcal{M}^2}^2 = \|[M]_\infty\|_{L^1}$.*

Proof. For $n \in \mathbb{N}$, let τ_n be the first time such that $|M_t| \geq n$. Then $\tau_n \uparrow \infty$, and M^{τ_n} is bounded by n . From (2.4), we get

$$\mathbb{E}[M_{\tau_n}^2] = \mathbb{E}[(M^{\tau_n})_\infty^2] = \mathbb{E}[M^{\tau_n}]_\infty = \mathbb{E}[M]_{\tau_n}. \quad (2.5)$$

First, suppose $M \in \mathcal{M}^2$. Then $\mathbb{E}[M_{\tau_n}^2] \leq \mathbb{E}[M_\infty^2] = \|M\|_{\mathcal{M}^2}^2$. So $\mathbb{E}[M]_{\tau_n} \leq \|M\|_{\mathcal{M}^2}^2$ for any $n \in \mathbb{N}$. Since $[M]_{\tau_n} \uparrow [M]_\infty$, we get $[M]_\infty \in L^1$, and $\|[M]_\infty\|_{L^1} \leq \|M\|_{\mathcal{M}^2}^2$.

Second, suppose $[M]_\infty \in L^1$. Let $C = \|[M]_\infty\|_{L^1}$. Then for any n , by (2.5), $\|M^{\tau_n}\|_{\mathcal{M}^2}^2 \leq C$, which implies that for any $t \geq 0$, $\mathbb{E}[M_{t \wedge \tau_n}^2] \leq C$. Letting $n \rightarrow \infty$ and using Fatou's lemma, we get $\mathbb{E}[M_t^2] \leq C$ for any $t \geq 0$. To prove the martingale property of M , we fix $t \geq s \geq 0$. For any $n \in \mathbb{N}$, since M^{τ_n} is a true martingale, $\mathbb{E}[M_{t \wedge \tau_n} | \mathcal{F}_s] = M_{s \wedge \tau_n}$. Note that as $n \rightarrow \infty$, $M_{t \wedge \tau_n} \rightarrow M_t$. Since the family $(M_{t \wedge \tau_n})$ is L^2 -bounded, by letting $n \rightarrow \infty$, we get a.s. $\mathbb{E}[M_t | \mathcal{F}_s] = M_s$. From $\mathbb{E}[M_t^2] \leq C$ for any $t \geq 0$ we get $M \in \mathcal{M}^2$ and $\|M\|_{\mathcal{M}^2}^2 = \lim_{t \rightarrow \infty} \mathbb{E}[M_t^2] \leq C = \|[M]_\infty\|_{L^1}$. \square

Corollary . *For a continuous local martingale M with $M_0 = 0$, M is a.s. constant 0 iff $[M]$ is a.s. constant 0.*

Proposition (Interval of Constancy). *For a continuous local martingale M , a.s. M and $[M]$ have the same interval of constancy. This means that, for almost every $\omega \in \Omega$, for any open interval $I \subset \mathbb{R}_+$ such that M is constant, $[M]$ is also constant, and vice versa.*

Proof. Since $[M - M_0] = [M]$, we may assume that $M_0 = 0$. It suffices to show that, for any fixed $p < q \in \mathbb{Q}_+$, on the event that M is constant on $[p, q]$, a.s. $[M]$ is constant on $[p, q]$, and vice versa.

We first show that, on the event E_0 that M is constant 0, $[M]$ is a.s. also constant 0, and vice versa. Define the stopping time $\tau = \inf\{t \geq 0 : M_t \neq 0\}$. Then M^τ is constant 0. By the last corollary, $[M^\tau] = [M]^\tau$ is a.s. constant 0. So $[M]$ is a.s. constant on $[0, \tau)$. Since $\tau = \infty$ on the event E_0 , $[M]$ is a.s. 0 on \mathbb{R}_+ on the event E_0 . The reverse direction can be proved similarly by defining the stopping time $\sigma = \inf\{t \geq 0 : [M]_t \neq 0\}$.

Fixe $p < q \in \mathbb{Q}_+$. Let $E_{p,q}$ denote the event that M is constant on (p, q) . On this event the local martingale $M^q - M^p$ is constant 0. From the last paragraph, we know that $[M^q - M^p]$ is a.s. constant 0 on the event $E_{p,q}$. Since $[M^q - M^p] = [M]^q - [M]^p$, by the last paragraph $[M]^q - [M]^p$ is a.s. constant 0 on $E_{p,q}$, and so $[M]$ is a.s. constant on $[p, q]$ on $E_{p,q}$. On the other hand, let $F_{p,q}$ be the event that $[M]$ is constant on $[p, q]$. Then $[M^q - M^p] = [M]^q - [M]^p$ is constant 0 on $F_{p,q}$, which implies that $M^q - M^p$ is a.s. constant 0 on $F_{p,q}$, and so M is a.s. constant on $[p, q]$ on the event $F_{p,q}$. \square

Exercise. For a continuous local martingale M , prove that on the event that $[M]_\infty < \infty$, a.s. $\lim_{t \rightarrow \infty} M_t$ converges.

Proposition 15.6. *For any sequence of continuous local martingales (M^n) with $M_0^n = 0$, $(M^n)_\infty \xrightarrow{\mathbb{P}} 0$ iff $[M^n]_\infty \xrightarrow{\mathbb{P}} 0$, and for any $t \geq 0$, $(M^n)_t^* \xrightarrow{\mathbb{P}} 0$ iff $[M^n]_t \xrightarrow{\mathbb{P}} 0$.*

Proof. First, let $(M^n)_\infty \xrightarrow{\mathbb{P}} 0$. Fix $\varepsilon > 0$. Let τ^n be the first time that $|M_t^n| \geq \varepsilon$. Then $(M^n)^{\tau^n}$ is a bounded martingale. So $((M^n)^{\tau^n})^2 - [(M^n)^{\tau^n}]$ is an L^2 -martingale. This implies that $\mathbb{E}[(M_{\tau^n}^n)^2] = \mathbb{E}[[M^n]_{\tau^n}]$. Here if $\tau^n = \infty$, $M_{\tau^n}^n$ is understood as $\lim_{t \rightarrow \infty} (M^n)_t^{\tau^n} = \lim_{t \rightarrow \infty} M_t^n$, which a.s. converges. In particular, we have $\mathbb{E}[[M^n]_{\tau^n}] \leq \varepsilon^2$ since $|M_{\tau^n}^n| \leq \varepsilon$. So we get

$$\mathbb{P}[[M^n]_\infty > \varepsilon] \leq \mathbb{P}[\tau^n < \infty] + \mathbb{P}[[M^n]_{\tau^n} > \varepsilon]$$

$$\leq \mathbb{P}[\tau^n < \infty] + \varepsilon^{-1} \mathbb{E}[[M^n]_{\tau^n}] \leq \mathbb{P}[\tau^n < \infty] + \varepsilon,$$

where the second inequality follows from Chebyshev's inequality. Since $(M^n)_\infty \xrightarrow{\mathbb{P}} 0$, $\mathbb{P}[\tau^n < \infty] \rightarrow 0$ as $n \rightarrow \infty$. So for n big enough, $\mathbb{P}[[M^n]_\infty > \varepsilon] < 2\varepsilon$. Then we get $[M^n]_\infty \xrightarrow{\mathbb{P}} 0$.

Second, let $[M^n]_\infty \xrightarrow{\mathbb{P}} 0$. Now we define τ^n to be the first time that $[M^n]_t \geq \varepsilon^3$. Then $[(M^n)^{\tau^n}] = [M^n]^{\tau^n}$ is bounded by ε^3 .

By the previous lemma, $(M^n)^{\tau^n} \in \mathcal{M}$ with $\|(M^n)^{\tau^n}\|_{\mathcal{M}^2}^2 = \mathbb{E}[M^n]_\infty^{\tau^n} \leq \varepsilon^3$, which implies that $\|((M^n)^{\tau^n})_\infty^*\|_{L^2}^2 \leq 4\varepsilon^3$. By Chebyshev's inequality,

$$\begin{aligned} \mathbb{P}[|(M^n)_\infty^*| > \varepsilon] &\leq \mathbb{P}[\tau^n < \infty] + \mathbb{P}[|((M^n)^{\tau^n})_\infty^*|^2 > \varepsilon^2] \\ &\leq \mathbb{P}[\tau^n < \infty] + \varepsilon^{-2} \|((M^n)^{\tau^n})_\infty^*\|_{L^2}^2 \leq \mathbb{P}[\tau^n < \infty] + 4\varepsilon. \end{aligned}$$

Here we use that $((M^n)^{\tau^n})_\infty^* = (M^n)_{\tau^n}^*$. Since $[M^n]_\infty \xrightarrow{\mathbb{P}} 0$, $\mathbb{P}[\tau^n < \infty] \rightarrow 0$ as $n \rightarrow \infty$. So for n big enough, $\mathbb{P}[|(M^n)_\infty^*| > \varepsilon] < 5\varepsilon$. Then we get $(M^n)_\infty^* \xrightarrow{\mathbb{P}} 0$.

The statement about $(M^n)_t^*$ and $[M^n]_t$ follows by considering the stopped process $(M^n)^t$. \square

Skip Proposition 15.7 (BDG inequalities) for now. Will come back later. Also skip Lemma 15.8 and Corollary 15.9.

A continuous adapted process A with locally finite total variation and $A_0 = 0$ will be simply called a finite variation process. For example, $[M, N]$ is a finite variation process. Such A determines a continuous, adapted, and non-decreasing process V such that V_t is the total variation of A over $[0, t]$. Then $\frac{1}{2}(V + A)$ and $\frac{1}{2}(V - A)$ are continuous non-decreasing, and determine two (positive) measures μ_+ and μ_- on \mathbb{R}_+ , which are locally finite and have no point mass. Since A is the difference of the two processes, dA determines a signed measure $\mu = \mu_+ - \mu_-$. The μ_+ and μ_- are singular to each other. The Stieltjes integral against dA and $|dA|$ are the Lebesgue integral against the signed measure μ and the (positive) measure $|\mu| := \mu_+ + \mu_-$, respectively. From $\frac{1}{2}(V \pm A) \geq 0$ we easily see that $|A_t| \leq V_t$ for all $t \geq 0$.

Proposition 15.10. *For any continuous local martingales M, N , we have a.s.*

$$|[M, N]_t| \leq \int_0^t |d[M, N]| \leq [M]_t^{1/2} [N]_t^{1/2}, \quad t \geq 0, \quad (2.6)$$

and for any measurable processes U, V , a.s.

$$\int_0^t |UV| |d[M, N]| \leq \left(\int_0^t U^2 d[M] \right)^{1/2} \left(\int_0^t V^2 d[N] \right)^{1/2}, \quad t \geq 0.$$

Proof. For every $a, b \in \mathbb{R}$, a.s.

$$0 \leq [aM + bN]_t = a^2[M]_t + b^2[N]_t + 2ab[M, N]_t, \quad t \geq 0.$$

Thus, a.s. for every $a, b \in \mathbb{Q}$, the above formula holds, which then implies $[M, N]_t^2 \leq [M]_t [N]_t$, $t \geq 0$.

Fix $t > s \geq 0$. Since $M - M^s$ and $N - N^s$ are continuous local martingales, we have

$$[M - M^s, N - N^s]_t^2 \leq [M - M^s]_t [N - N^s]_t.$$

Recall that $[M^s, N^s]_t = [M, N^s]_t = [M^s, N]_t = [M, N]_t^s = [M, N]_s$. The LHS equals $([M, N]_t - [M, N]_s)^2$. Similarly, the RHS equals $([M]_t - [M]_s)([N]_t - [N]_s)$. Thus, a.s.

$$([M, N]_t - [M, N]_s)^2 \leq ([M]_t - [M]_s)([N]_t - [N]_s).$$

Then a.s. the above inequality holds for any $t > s \in \mathbb{Q}_+$. By continuity of $[M, N], [M], [N]$, the above inequality a.s. holds for any $t > s \geq 0$. Thus, a.s. for any $t > 0$ and any partition $0 = t_0 < \dots < t_n = t$, we have

$$\begin{aligned} \sum_{k=1}^n |[M, N]_{t_k} - [M, N]_{t_{k-1}}| &\leq \sum_{k=1}^n ([M]_{t_k} - [M]_{t_{k-1}})^{1/2} ([N]_{t_k} - [N]_{t_{k-1}})^{1/2} \\ &\leq \left(\sum_{k=1}^n ([M]_{t_k} - [M]_{t_{k-1}}) \right)^{1/2} \left(\sum_{k=1}^n ([N]_{t_k} - [N]_{t_{k-1}}) \right)^{1/2} = [M]_t^{1/2} [N]_t^{1/2}. \end{aligned}$$

Taking supremum over all partitions of $[0, t]$, we conclude that a.s. (2.6) holds. A same argument shows that, for any $s < t \in \mathbb{Q}_+$, a.s.,

$$\int_s^t |d[M, N]| \leq ([M]_t - [M]_s)^{1/2} ([N]_t - [N]_s)^{1/2}.$$

Next, write $d\mu = d[M]$, $d\nu = d[N]$, and $d\rho = |d[M, N]|$. The last displayed formula implies that, a.s. for any interval $I = (s, t)$, where $s < t \in \mathbb{Q}_+$, we have $(\rho I)^2 \leq (\mu I)(\nu I)$. By continuity, a.s. the inequality holds for any interval $I \subset \mathbb{R}_+$. Outside the exceptional event, we have the following. Suppose $G \subset \mathbb{R}_+$ is open. Then G is a disjoint union of open intervals I_k . So by Cauchy inequality

$$\rho G = \sum_k \rho I_k \leq \sum_k (\mu I_k)^{1/2} (\nu I_k)^{1/2} \leq \left(\sum_k \mu I_k \right)^{1/2} \left(\sum_k \nu I_k \right)^{1/2} = (\mu G)^{1/2} (\nu G)^{1/2}.$$

Since every Borel set B can be approximated by open sets, we get $\rho B \leq (\mu B)^{1/2} (\nu B)^{1/2}$.

Suppose $U = \sum_k a_k \mathbf{1}_{B_k}$ and $V = \sum_k b_k \mathbf{1}_{B_k}$ are simple functions such that B_k 's are mutually disjoint. Then

$$\begin{aligned} \rho|UV| &= \sum_k |a_k b_k| \rho B_k \leq \sum_k |a_k b_k| (\mu B_k)^{1/2} (\nu B_k)^{1/2} \\ &\leq \left(\sum_k |a_k|^2 \mu B_k \right)^{1/2} \left(\sum_k |b_k|^2 \nu B_k \right)^{1/2} = (\mu U^2)^{1/2} (\nu V^2)^{1/2}. \end{aligned}$$

The inequality then extends to any measurable functions U, V on \mathbb{R}_+ . \square

Corollary . *If M and N are two local martingales that agree (throughout \mathbb{R}_+) on an event E , then a.s. $[M]$ and $[N]$ also agree on the event E .*

Proof. Since $M - N$ is constant 0 on the event E , by the proposition on intervals of constancy, $[M - N]$ is a.s. constant 0 on E . Since $[M] - [N] = [M + N, M - N]$, by Proposition 15.10, $[M] - [N]$ is a.s. constant 0 on the event E . \square

For every finite variation process A , let $L(A)$ denote the space of progressive processes V such that the Stieltjes integrals $\int_0^t |V| |dA|$ is finite for all $t \geq 0$. Recall that a process V is called progressive, if for any fixed $t \geq 0$, the map $[0, t] \times \Omega \ni (s, \omega) \mapsto V_s(\omega)$ is $\mathcal{B}[0, t] \times \mathcal{F}_t$ -measurable. A progressive process is adapted, but the converse is not true. An adapted right or left-continuous process is progressive. Thus, $L(A)$ contains all right-continuous or left-continuous adapted processes, which are locally bounded. For every $V \in L(A)$, we may define the process $V \cdot A$ by $(V \cdot A)_t = \int_0^t V dA$. Such $V \cdot A$ is also a finite variation process. We use the assumption that V is progressive to guarantee that $V \cdot A$ is adapted. Moreover, $d(V \cdot A)$ is absolutely continuous w.r.t. dA with Radon-Nikodym derivative being V , $U \in L(V \cdot A)$ iff $UV \in L(A)$, and $U \cdot (V \cdot A) = (UV) \cdot A$.

For $V \in \mathcal{E}$, $V \cdot A$ defined in this way agrees with the elementary integral. Then $V \cdot A$ is also a finite variation process, and $d(V \cdot A)$ is absolutely continuous w.r.t. dA with Radon-Nikodym derivative being V . Thus, $U \in L(V \cdot A)$ iff $UV \in L(A)$, and $U \cdot (V \cdot A) = (UV) \cdot A$.

Let \mathcal{E} denote the space of uniformly bounded predictable step processes, which may change values only at finitely many fixed deterministic times. This means that each $U \in \mathcal{E}$ can be expressed as

$$U_t = \sum_{j=0}^{n-1} \zeta_j \mathbf{1}_{(t_j, \infty)}(t) = \sum_{k=0}^{n-1} \eta_k \mathbf{1}_{(t_k, t_{k+1}]}(t), \quad (2.7)$$

where $n \in \mathbb{N}$, $0 = t_0 < t_1 < \dots < t_n$, and for each k , ζ_k and η_k are bounded \mathcal{F}_{t_k} -measurable random variables. For any finite variation process A , we have $\mathcal{E} \subset L(A)$, and the Stieltjes integral $V \cdot A$ agrees with the elementary integral defined earlier.

Lemma 15.11. *For any continuous local martingales M, N , and any $U \in \mathcal{E}$, the elementary integral $U \cdot M$ is also a continuous local martingale, and a.s.*

$$[U \cdot M, N] = U \cdot [M, N].$$

Here the dot on the left is an elementary integral, and the dot on the right is a Stieltjes integral.

Proof. We may take $M_0 = N_0 = 0$. The fact that $U \cdot M$ is a local martingale follows from localization: we use a sequence of stopping times $\tau_n \uparrow \infty$ to make M^{τ_n} bounded martingales. Then Lemma 15.3 implies that $(U \cdot M)^{\tau_n} = U \cdot M^{\tau_n}$ is a martingale. So $U \cdot M$ is a local martingale. To prove $[U \cdot M, N] = U \cdot [M, N]$, we may assume by localization that M, N , and $[M, N]$ are uniformly bounded. Then $M, N, MN - [M, N]$ are all bounded martingales.

To prove that a.s. $[U \cdot M, N] = U \cdot [M, N]$, it suffices to prove that $(U \cdot M)N - U \cdot [M, N]$ is a martingale. Suppose U has the form of (2.7). Then

$$(U \cdot M)_t N_t = \sum_{j=0}^{n-1} \zeta_j (M_t - M_{t \wedge t_j}) N_t;$$

$$(U \cdot [M, N])_t = \sum_{j=0}^{n-1} \zeta_j ([M, N]_t - [M, N]_{t \wedge t_j}).$$

Then we have

$$(U \cdot M)_t N_t - (U \cdot [M, N])_t = \sum_{j=0}^{n-1} \zeta_j ((M_t N_t - [M, N]_t) - (M_t^{t_j} N_t - [M^{t_j}, N]_t))$$

For each j , $MN - [M, N]$ and $M^{t_j}N - [M, N]^{t_j}$ are martingales, and so is their difference. Since the two processes agree on $[0, t_j]$, their difference vanishes on $[0, t_j]$. By Corollary 6.14, $\zeta_j((M_t N_t - [M, N]_t) - (M_t^{t_j} N_t - [M^{t_j}, N]_t))$ is a martingale. So $(U \cdot M)N - (U \cdot [M, N])$ is a martingale. \square

Given a continuous local martingale M , let $L(M)$ denote the class of all progressive processes V with $(V^2 \cdot [M])_t < \infty$ for all $t \geq 0$, i.e., $V^2 \in L([M])$.

Theorem 15.12. *For every continuous local martingale M and $V \in L(M)$, there exists an a.s. unique continuous local martingale M' with $M'_0 = 0$ such that for any continuous local martingale N , a.s. $[M', N] = V \cdot [M, N]$.*

Remark: If $V \in L(M)$, then $V \in L([M, N])$ by Proposition 15.10. So $V \cdot [M, N]$ is well defined.

We will use $V \cdot M$ or $\int V dM$ to denote the process M' , and call it the stochastic integral of V against dM . The equality then reads

$$[V \cdot M, N] = V \cdot [M, N].$$

By Lemma 15.11, this stochastic integral extends the elementary integral in the case that $V \in \mathcal{E}$. By iteration, we find that for two continuous local martingales M and N , and $U \in L(M)$, $V \in L(N)$, a.s.

$$[U \cdot M, V \cdot N] = U \cdot [M, V \cdot N] = U \cdot (V \cdot [M, N]) = UV \cdot [M, N].$$

In particular, $[U \cdot M] = U^2 \cdot [M]$.

Proof. To prove the uniqueness, suppose M' and M'' are two continuous local martingales started from 0 such that for every continuous local martingale N , a.s. $[M', N] = [M'', N] = V \cdot [M, N]$. Taking $N = M' - M''$, we get a.s. $[M' - M''] = 0$, which implies that a.s. $M' = M''$.

For the existence, first assume that $\|V\|_M^2 := \mathbb{E}(V^2 \cdot [M])_\infty < \infty$. Consider the linear function on \mathcal{M}^2 : $N \mapsto \mathbb{E}(V \cdot [M, N])_\infty$. By Proposition 15.10, $(V \cdot [M, N])$ is a process with finite total variation on $[0, \infty)$. So $(V \cdot [M, N])_\infty := \lim_{t \rightarrow \infty} (V \cdot [M, N])_t$ exists and is finite. Also by Proposition 15.10,

$$|\mathbb{E}(V \cdot [M, N])_\infty| \leq \mathbb{E}[(V^2 \cdot [M])_\infty^{1/2} [N]_\infty^{1/2}] \leq \|V\|_M \mathbb{E}[[N]_\infty]^{1/2} = \|V\|_M \|N\|_{\mathcal{M}^2}.$$

Thus, the linear function is bounded. Since \mathcal{M}^2 is a Hilbert space, by Riesz representation theorem, there is $M' \in \mathcal{M}^2$ such that for any $N \in \mathcal{M}^2$,

$$\mathbb{E}(V \cdot [M, N])_\infty = \langle M', N \rangle = [M', N]_\infty = \mathbb{E}M'_\infty N_\infty.$$

By replacing N by N^τ for some stopping time τ , we get

$$\mathbb{E}(V \cdot [M, N])_\tau = \mathbb{E}(V \cdot [M, N]^\tau)_\infty = \mathbb{E}(V \cdot [M, N^\tau])_\infty = \mathbb{E}M'_\infty N_\infty^\tau = \mathbb{E}M'_\infty N_\tau = \mathbb{E}M'_\tau N_\tau.$$

Since V is progressive, $V \cdot [M, N]$ is adapted. Now $\mathbb{E}[(V \cdot [M, N] - M'N)_\tau] = 0$ for any stopping time τ . By Lemma 6.13, $V \cdot [M, N] - M'N$ is a martingale. So we get a.s. $[M', N] = V \cdot [M, N]$ for any $N \in \mathcal{M}^2$. By localization, this extends to any continuous local martingale N .

We may remove the assumption that $\|V\|_M^2 < \infty$ by localization. More specifically, define $\tau_n = \inf\{t \geq 0 : (V^2 \cdot [M])_t \geq n\}$, $n \in \mathbb{N}$. Then for every n , $\|V\|_{M^{\tau_n}} \leq n$. By the previous argument, for each n , there exists a continuous local martingale $M^{(n)}$ such that for any continuous local martingale N , a.s.

$$[M^{(n)}, N] = V \cdot [M^{\tau_n}, N]. \quad (2.8)$$

For $m < n$, we have that for any continuous local martingale N , a.s.

$$[(M^{(n)})^{\tau_m}, N] = [M^{(n)}, N]^{\tau_m} = V \cdot [M^{\tau_n}, N]^{\tau_m} = V \cdot [M^{\tau_m}, N] = [M^{(m)}, N].$$

Taking $N = (M^{(n)})^{\tau_m} - M^{(m)}$, we get a.s. $(M^{(n)})^{\tau_m} = \widetilde{M}^{(m)}$. So there exists a continuous process M' such that $(M')^{\tau_n} = M^{(n)}$ for any $n \in \mathbb{N}$. Since $\tau_n \uparrow \infty$, and $(M')^{\tau_n}$ is a local martingale for every n , M' is a continuous local martingale. Finally, for any continuous local martingale N , and any $n \in \mathbb{N}$, a.s.

$$[M', N]^{\tau_n} = [M^{(n)}, N] = V \cdot [M^{\tau_n}, N] = (V \cdot [M, N])^{\tau_n}.$$

Since $\tau_n \uparrow \infty$, we get a.s. $[\widetilde{M}, N] = V \cdot [M, N]$. □

Definition . A process X is called a continuous semimartingale if it can be written as $X = M + A$, where M is a continuous local martingale, and A is a finite variation process with $A_0 = 0$. An \mathbb{R}^d -valued process $X = (X^1, \dots, X^d)$ is called a continuous vector semimartingale if every X^j is a continuous semimartingale.

The local martingales and semimartingales considered in this course are all continuous. So we will omit the words “continuous”.

We call $M + A$ a canonical decomposition of X . If a semimartingale X has two canonical decompositions $X = M + A = M' + A'$, then $M - M' = A' - A$ is a local martingale with locally finite variation starting from 0, which is a.s. 0 by Proposition 15.2. Thus, the canonical decomposition is a.s. unique.

If X is a semimartingale with canonical decomposition $M + A$, then for any stopping time τ , X^τ is a semimartingale with canonical decomposition $M^\tau + A^\tau$. On the other hand, if there is a sequence of stopping times $\tau_n \uparrow \infty$ such that for every n , X^{τ_n} is a semimartingale, then X is a semimartingale. In fact, if $M^n + A^n$ is the canonical decomposition of X^{τ_n} , then for $n < m$, $(M^m)^{\tau_n} + (A^m)^{\tau_n}$ is also a canonical decomposition of X^{τ_n} . By the uniqueness, we have a.s. $(M^m)^{\tau_n} = M^n$. Then we may define processes M and A such that on an event E with probability 1, for any n , $M^{\tau_n} = M^n$, $A^{\tau_n} = A^n$; and on E^c , $M = X$ and $A = 0$. Then A is a finite variation process, and for every n , a.s. $M^{\tau_n} = M^n$, and so M is a local martingale. Since $X = M + A$, X is a semimartingale.

For two semimartingales X and X' with canonical decompositions $X = M + A$ and $X' = M' + A'$, we define

$$[X] = [M], \quad [X, X'] = [M, M'].$$

For a semimartingale X with canonical decomposition $M + A$, let $L(X) = L(M) \cap L(A)$. For $V \in L(X)$, we define

$$V \cdot X = \int V dX = V \cdot M + V \cdot A.$$

So $V \cdot X$ is a semimartingale with canonical decomposition $V \cdot M + V \cdot A$.

Exercise. Let X and Y be two continuous adapted process, and $E \in \mathcal{F}_0$. Suppose $X = Y$ on E , and Y is constant on E^c . Show that if X is a martingale, local martingale, or semimartingale, then Y is respectively a martingale, local martingale, or semimartingale.

Exercise. Let X and Y be semimartingales. Prove that (i) For any stopping time τ , $[X^\tau, Y^\tau] = [X^\tau, Y] = [X, Y]^\tau$. (ii) On any event such that $X = Y$, a.s. X and Y have the same decomposition, and so $[X] = [Y]$. (iii) Almost surely on any interval such that X is constant, $[X]$ is also constant. To prove (ii), we need to improve Proposition 15.2: If M is a local martingale, then on any event E such that M has locally finite variation, M is a.s. constant.

Corollary 15.14 (Stochastic Dominated Convergence). *Fix a semimartingale X . Let $U, V, V^1, V^2, \dots \in L(X)$ satisfy a.s. $|V_t^n| \leq U_t$ for all $n \in \mathbb{N}$ and $t \geq 0$, and a.s. $V_t^n \rightarrow V_t$ for all $t \geq 0$ with at most countably many possible exceptions. Then for all $t \geq 0$,*

$$(V^n \cdot X - V \cdot X)_t^* = \sup_{0 \leq s \leq t} |(V^n \cdot X)_s - (V \cdot X)_s| \xrightarrow{\mathbb{P}} 0.$$

Proof. Assume that $X = M + A$. Since $U \in L(X)$, $U \in L(A)$ and $U^2 \in L([M])$. By DCT for Stieltjes integral and the fact that dA and $d[M]$ have no point mass, we have a.s. for all $t \geq 0$,

$\int_0^t |V_s^n - V_s| |dA| \rightarrow 0$ and $((V^n - V)^2 \cdot [M])_t \rightarrow 0$, where the former convergence further implies that $(V^n \cdot A - V \cdot A)_t^* \xrightarrow{P} 0$. Since a.s. $((V^n - V)^2 \cdot [M])_t = [(V^n - V) \cdot M]_t$, by Proposition 15.6, $((V^n - V) \cdot M)_t^* \xrightarrow{P} 0$ for all $t \geq 0$. Thus, $((V^n - V) \cdot X)_t^* \xrightarrow{P} 0$ for all $t \geq 0$. \square

Exercise. Prove that if two semimartingales X and Y agree on an event E , then for any continuous adapted process V , a.s. $V \cdot X = V \cdot Y$ on E . Hint: Use predictable step processes to approximate V , and apply stochastic dominated convergence theorem.

Proposition 15.15 (Chain Rule). *Let X be a semimartingale. Let U and V be two progressive processes such that $V \in L(X)$. Then $U \in L(V \cdot X)$ iff $UV \in L(X)$, in which case, a.s. $U \cdot (V \cdot X) = (UV) \cdot X$.*

Proof. Assume $X = M + A$ is the canonical decomposition of X . Then $V \cdot X = V \cdot M + V \cdot A$, and $[V \cdot M] = V^2 \cdot [M]$. Now $U \in L(V \cdot X)$ iff $U \in L(V \cdot A)$ and $U^2 \in L([V \cdot M]) = L(V^2 \cdot [M])$, which is further equivalent to $UV \in L(A)$ and $U^2 V^2 \in L([M])$, which is equivalent to $UV \in L(X)$. We know that $U \cdot (V \cdot A) = (UV) \cdot A$. It remains to prove that a.s. $U \cdot (V \cdot M) = (UV) \cdot M$. To see this, note that for any local martingale, a.s. $[U \cdot (V \cdot M), N] = U \cdot [V \cdot M, N] = U \cdot (V \cdot [M, N]) = (UV) \cdot [M, N] = [(UV) \cdot M, N]$. \square

Proposition 15.16 (Optional Stopping). *Let X be a semimartingale. Let $V \in L(X)$. Let τ be a stopping time. Then a.s.*

$$(V \cdot X)^\tau = V \cdot X^\tau = (\mathbf{1}_{[0, \tau]} \cdot V) \cdot X.$$

Proof. The statement is obvious if X is a finite variation process because in that case $dX^\tau = \mathbf{1}_{[0, \tau]} dX$. It suffices to prove the statement in the case that $X = M$ is a local martingale. Now for any local martingale N ,

$$\begin{aligned} [(V \cdot M)^\tau, N] &= [V \cdot M, N^\tau] = V \cdot [M, N^\tau] = V \cdot [M^\tau, N] = [V \cdot M^\tau, N] \\ &= V \cdot [M, N]^\tau = (\mathbf{1}_{[0, \tau]} V) \cdot [M, N] = [(\mathbf{1}_{[0, \tau]} V) \cdot M, N]. \end{aligned}$$

This then implies that $(V \cdot M)^\tau = V \cdot M^\tau = (\mathbf{1}_{[0, \tau]} V) \cdot M$. \square

We briefly recall some basic definition and results.

1. Local martingales are natural extensions of continuous martingales.
2. For local martingales M and N , $[M, N]$ is the finite variation process such that $MN - [M, N]$ is a local martingale.
3. For a finite variation process A , we define $U \cdot A$ using Stieltjes integral for $U \in L(A)$.
4. For a local martingale M and $U \in L(M)$, i.e., $U^2 \in L([M])$, $U \cdot M$ is the local martingale such that for any local martingale N , $[U \cdot M, N] = U \cdot [M, N]$.

5. A semimartingale X is the sum of a local martingale M with a finite variation process A . We define $[X] = [M]$ and $U \cdot X = U \cdot M + U \cdot A$ for $U \in L(X) = L(M) \cap L(A)$.
6. We use localization to gain boundedness assumptions on local martingales.

We are going to prove the celebrated Itô's formula, which says, if $X = (X^1, \dots, X^d)$ is a vector semimartingale, and if $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is C^2 differentiable, then $Y := f(X)$ is also a semimartingale, and a.s.

$$Y = Y_0 + \sum_{j=1}^d \partial_j f(X) \cdot X^j + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_j \partial_k f(X) \cdot [X^j, X^k].$$

We will first prove a simple case: $f(x, y) = xy$. Note that $\partial_x f = y$, $\partial_y f = x$, $\partial_x^2 f = \partial_y^2 f = 0$, and $\partial_x \partial_y f = \partial_y \partial_x f = 1$. Later we will use the special case to prove the general result.

Theorem 15.17 (Product Formula). *For any two semi-martingales X, Y , their product XY is also a semi-martingale, and a.s.*

$$XY = X_0 Y_0 + X \cdot Y + Y \cdot X + [X, Y].$$

Proof. By polarization, it suffices to consider the case that $X = Y$. First, suppose $X = M \in \mathcal{M}^2$ and $M_0 = 0$. Fix $n \in \mathbb{N}$. Define V^n and Q^n as in the proof of Theorem 15.5: We first define a sequence of stopping times $\tau_k^n \uparrow 0$ by $\tau_0^n = 0$ and $\tau_k^n = \inf\{t \geq \tau_{k-1}^n : |M_t - M_{\tau_{k-1}^n}| = 2^{-n}\}$, $k \in \mathbb{N}$. Then define V^n and Q^n by (2.3), i.e.,

$$V_t^n = \sum_{k=0}^{\infty} M_{\tau_k^n} \mathbf{1}_{(\tau_k^n, \tau_{k+1}^n]}(t), \quad Q_t^n = \sum_{k=0}^{\infty} (M_{t \wedge \tau_{k+1}^n} - M_{t \wedge \tau_k^n})^2.$$

Recall that we have

$$M^2 = 2(V^n \cdot M) + Q^n.$$

Since $|V^n|_t \leq M_t^*$ for all $n \in \mathbb{N}$ and $t \geq 0$, and $V^n \rightarrow M$, by stochastic dominated convergence, $(V^n \cdot M)_t \xrightarrow{P} (M \cdot M)_t$ for every $t \geq 0$. From the proof of Theorem 15.5, for every $t \geq 0$, $Q_t^n \xrightarrow{P} [M]_t$. So for any $t \geq 0$, a.s. $M_t^2 = 2(M \cdot M)_t + [M]_t$. By continuity, we get a.s. $M^2 = 2(M \cdot M) + [M]$.

Next, assume that $X = M$ is a local martingale and $M_0 = 0$. Let $\tau_n = \inf\{t \geq 0 : |M_t| \geq n\}$, $n \in \mathbb{N}$. Then $M^{\tau_n} \in \mathcal{M}^2$. So a.s. for any $n \in \mathbb{N}$,

$$(M^2)^{\tau_n} = (M^{\tau_n})^2 = 2(M^{\tau_n} \cdot M^{\tau_n}) + [M^{\tau_n}] = (2(M \cdot M) + [M])^{\tau_n},$$

which implies that a.s. $M^2 = 2(M \cdot M) + [M]$.

If $X = A$ is a finite variation process, then we need to show that $A^2 = 2A \cdot A$. To see this, suppose $\mu = dA$. Then by Fubini Theorem,

$$A_t^2 = \iint \mathbf{1}_{[0,t]^2} \mu^2(ds_1 \otimes ds_2) = \iint \mathbf{1}_{\Delta_1} \mu^2(ds_1 \otimes ds_2) + \iint \mathbf{1}_{\Delta_2} \mu^2(ds_1 \otimes ds_2)$$

$$= \int_0^t A_{s_2} \mu(ds_2) + \int_0^t A_{s_1} \mu(ds_1) = 2(A \cdot A)_t,$$

where $\Delta_1 = \{(s_1, s_2) \in \mathbb{R}^2 : 0 \leq s_1 \leq s_2 \leq t\}$, and $\Delta_2 = \{(s_1, s_2) \in \mathbb{R}^2 : 0 \leq s_2 < s_1 \leq t\}$.

Suppose X is a semi-martingale with canonical decomposition $M + A$, and satisfies $X_0 = M_0 = 0$. We already have a.s. $M^2 = 2M \cdot M + [M]$ and $A^2 = 2A \cdot A$. To prove that a.s. $X^2 = 2X \cdot X + [X]$, it suffices to show that a.s. $AM = A \cdot M + M \cdot A$. Fix $t \geq 0$ and $n \in \mathbb{N}$. Let $t_k = \frac{k}{n}t$. Define processes A^n and M^n by

$$A^n = \sum_{k \in \mathbb{N}} A_{t_{k-1}} \mathbf{1}_{(t_{k-1}, t_k]}, \quad M^n = \sum_{k \in \mathbb{N}} M_{t_k} \mathbf{1}_{(t_{k-1}, t_k]}.$$

Note that A^n is adapted, but M^n is not. We have

$$A_t M_t = (A^n \cdot M)_t + (M^n \cdot A)_t, \quad n \in \mathbb{N}.$$

To see this, note that

$$(A^n \cdot M)_t = \sum_{k=1}^n A_{t_{k-1}} (M_{t_k} - M_{t_{k-1}}), \quad (M^n \cdot A)_t = \sum_{k=1}^n M_{t_k} (A_{t_k} - A_{t_{k-1}}).$$

Since $|A_s^n| \leq (A_s)^*$ and $A_s^n \rightarrow A_s$, we get a.s. $(A^n \cdot M)_t \rightarrow (A \cdot M)_t$ by stochastic dominated convergence. Since $|M_s^n| \leq M_s^*$ for $0 \leq s \leq t$, we get $(M^n \cdot A)_t \rightarrow (M \cdot A)_t$ by ordinary dominated convergence. Thus, for any fixed $t \geq 0$, a.s. $(MA)_t = (A \cdot M)_t + (M \cdot A)_t$. By continuity, we then get a.s. $(MA)_t = (A \cdot M)_t + (M \cdot A)_t$ for any $t \geq 0$.

Finally, we may remove the assumption that $X_0 = 0$ because from above we have a.s.

$$(X - X_0)^2 = 2(X - X_0) \cdot (X - X_0) + [X - X_0] = 2X \cdot X - 2X_0(X - X_0) + [X],$$

which implies that $X^2 = X_0^2 + 2X \cdot X + [X]$. \square

Remark . The theorem implies that semimartingales are closed under multiplication. This is not true for local martingales are not. We may rewrite the formula in the above theorem as $d(XY) = XdY + YdX + d[X, Y]$. If X and Y are positive, then we may write the formula as

$$\frac{d(XY)}{XY} = \frac{dX}{X} + \frac{dY}{Y} + \frac{d[X, Y]}{XY}.$$

By induction, this extends to the product of finitely many positive semimartingales:

$$\frac{d \prod_{j=1}^n X^j}{\prod_{j=1}^n X^j} = \sum_{j=1}^n \frac{dX^j}{X^j} + \sum_{1 \leq j < k \leq n} \frac{d[X^j, X^k]}{X^j X^k}.$$

If we move the denominator on the LHS to the right, then the formula holds without assuming that the X^j are positive.

Proposition 15.18. Let X, Y be two semimartingales. Let $t > 0$. Let Δ^n be a sequence of partitions of $[0, t]$. For each $\Delta^n = \{0 = t_0^n < \dots < t_{k_n}^n\}$, we define

$$T^{\Delta^n} = \zeta^{\Delta^n}(X, Y) = \sum_{k=1}^{k_n} (X_{t_k^n} - X_{t_{k-1}^n})(Y_{t_k^n} - Y_{t_{k-1}^n}). \quad (2.9)$$

Suppose the mesh size $|\Delta^n| \rightarrow 0$. Then $T^{\Delta^n} \xrightarrow{\mathbb{P}} [X, Y]_t$.

Remark . This proposition shows that the quadratic variation a.s. does not depend on the underlying filtration. If \mathcal{F}' is another filtration w.r.t. which X is also a semimartingale, then a.s. $[X]^{\mathcal{F}'} = [X]^{\mathcal{F}}$. This proposition also implies that if under a new measure $\mathbb{P}' \ll \mathbb{P}$, X is also a semimartingale, then the quadratic variation of X a.s. does not change. However, the canonical decomposition of X may be different. See Girsanov Theorem later.

Proof of Proposition 15.18. We may assume that $X_0 = Y_0 = 0$. Introduce the predictable step processes

$$X^n = \sum_{k=1}^n X_{t_{k-1}^n} \mathbf{1}_{(t_{k-1}^n, t_k^n]}, \quad Y^n = \sum_{k=1}^n Y_{t_{k-1}^n} \mathbf{1}_{(t_{k-1}^n, t_k^n]}.$$

Then we have

$$X_t Y_t = (X^n \cdot Y)_t + (Y^n \cdot X)_t + T^{\Delta^n}, \quad n \in \mathbb{N}.$$

Since $X^n \rightarrow X$, $Y^n \rightarrow Y$, and $|X_s^n| \leq X_s^*$, $|Y_s^n| \leq Y_s^*$, $0 \leq s \leq t$, and $X^* \in L(Y)$ and $Y^* \in L(X)$, by the stochastic dominated convergence, we get $(X^n \cdot Y)_t \xrightarrow{\mathbb{P}} (X \cdot Y)_t$ and $(Y^n \cdot X)_t \xrightarrow{\mathbb{P}} (Y \cdot X)_t$. Since by Theorem 15.17, a.s. $X_t Y_t = (X \cdot Y)_t + (Y \cdot X)_t + [X, Y]_t$, we get $T^{\Delta^n} \xrightarrow{\mathbb{P}} [X, Y]_t$. \square

Example . If B is a Brownian motion, then $[B]_t = t$ for all $t \geq 0$. This follows from Proposition 15.18 and the quadratic variation of Brownian motion. One may also prove this directly by showing that $B_t^2 - t$ is a martingale. If B' is another Brownian motion independent of B , then $[B, B'] = 0$. To see this, by computing covariance function, we find that $(B + B')/\sqrt{2}$ is a Brownian motion, which implies that $[B + B']_t = 2t$ for all $t \geq 0$. So $[B, B'] = ([B + B'] - [B] - [B'])/2 = 0$. Below is a more general statement.

Exercise. Suppose that X and Y are two \mathcal{F} -continuous local martingale, which are independent of each other. Prove that a.s. $[X, Y] = 0$, and so XY is also an \mathcal{F} -local martingale.

Hint: Since $[X, Y]$ does not depend on the filtration, it suffices to show that a.s. $[X, Y] = 0$ if the filtration is the natural filtration $\mathcal{F}^{(X, Y)}$ generated by (X, Y) . By localization, we may assume that X and Y are bounded martingales. Then a.s. $[X, Y] = 0$ is equivalent to that XY is a martingale w.r.t. $\mathcal{F}^{(X, Y)}$, which further follows from a monotone class argument.

Theorem 15.19 (Itô's Formula). *Let $X = (X^1, \dots, X^d)$ be a vector semimartingale. Let $f \in C^2(\mathbb{R}^d, \mathbb{R})$. Then $f(X)$ is a semimartingale, and a.s.*

$$f(X) = f(X_0) + \sum_{j=1}^d \partial_j f(X) \cdot X^j + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_j \partial_k f(X) \cdot [X^j, X^k]. \quad (2.10)$$

We will often express the Itô's formula as a differential form:

$$df(X) = \sum_{j=1}^d \partial_j f(X) dX^j + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_j \partial_k f(X) d[X^j, X^k].$$

In the case that $d = 1$, the formula becomes

$$df(X) = f'(X) dX + \frac{1}{2} f''(X) d[X].$$

Proof. Let Σ denote the set of all $f \in C^2(\mathbb{R}^d \rightarrow \mathbb{R})$ such that a.s. (2.10) holds. Then Σ is a vector space, and contains all constant functions and coordinate functions: $f_j(x) = x_j$, $1 \leq j \leq d$. Suppose $f, g \in \Sigma$. By Theorem 15.17, a.s.

$$\begin{aligned} f(X)g(X) &= f(X_0)g(X_0) + f(X) \cdot g(X) + g(X) \cdot f(X) + [f(X), g(X)] \\ &= (fg)(X_0) + \sum_{j=1}^d f(X) \partial_j g(X) \cdot X^j + \sum_{j=1}^d g(X) \partial_j f(X) \cdot X^j + \\ &+ \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d f(X) \partial_j \partial_k g(X) d[X_j, X_k] + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d g(X) \partial_j \partial_k f(X) d[X_j, X_k] \\ &+ \left[\sum_{j=1}^d \partial_j f(X) dX^j, \sum_{k=1}^d \partial_k g(X) dX^k \right] \\ &= (fg)(X_0) + \sum_{j=1}^d (f(X) \partial_j g(X) + g(X) \partial_j f(X)) \cdot X^j \\ &+ \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d (f(X) \partial_j \partial_k g(X) + g(X) \partial_j \partial_k f(X) + \partial_j f(X) \partial_k g(X) + \partial_k f(X) \partial_j g(X)) d[X_j, X_k] \\ &= (fg)(X_0) + \sum_{j=1}^d \partial_j (fg)(X) \cdot X^j + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_j \partial_k (fg)(X) d[X_j, X_k]. \end{aligned}$$

Thus, $fg \in \Sigma$. This means that Σ is closed under multiplication. So Σ contains all polynomials in x_1, \dots, x_d . We now prove that if X is uniformly bounded, i.e., there is a constant $R < \infty$

such that $|X_t| \leq R$ for all $t \geq 0$, then for any $f \in C^2(\mathbb{R}^d, \mathbb{R})$, (2.10) holds. Recall the (higher-order) Weierstrass Approximation Theorem: There is a sequence of polynomials q_n such that q_n and its up to second order partial derivatives converge to f and its corresponding partial derivatives, respectively, uniformly on $\{x \in \mathbb{R}^d : |x| \leq R\}$. For each $n \in \mathbb{N}$, since $q_n \in \Sigma$, we have

$$q_n(X) = q_n(X_0) + \sum_{j=1}^d \partial_j q_n(X) \cdot X^j + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_j \partial_k q_n(X) [X^j, X^k].$$

Letting $n \rightarrow \infty$ and using the ordinary and stochastic dominated convergence theorems, we then conclude that (2.10) holds.

We intend to use localization to remove the boundedness assumption on X , and define for each $n \in \mathbb{N}$, $\tau_n = \inf\{t \geq 0 : |X_t| \geq n\}$. However, since we did not assume that $|X_0| \leq n$, we do not have $|X^{\tau_n}| \leq n$. In fact, if $|X_0| > n$, then $\tau_n = 0$, and X^{τ_n} is a constant value outside $\{x : |x| \leq n\}$. We observe that X^{τ_n} is uniformly bounded on the event $\{|X_0| \leq n\} \in \mathcal{F}_0$, and is constant on $\{|X_0| > n\}$. This motivates us to slightly relax the boundedness assumption on X . We claim that, if there is an \mathcal{F}_0 -measurable event E , such that X is uniformly bounded on the event E , and is constant on the event E^c , then (2.10) holds for any $f \in C^2(\mathbb{R}^d, \mathbb{R})$. To see this, we may define another process \tilde{X} such that $\tilde{X} = X$ on E and $\tilde{X} = 0$ on E^c . By an exercise, \tilde{X} is also a semimartingale. Let $f \in C^2(\mathbb{R}^d, \mathbb{R})$. Since \tilde{X} is uniformly bounded, from the last paragraph, we see that $f(\tilde{X})$ is a semimartingale, and (2.10) holds for \tilde{X} . Now $f(X) = f(\tilde{X})$ on E , and $f(X)$ is constant on E^c . By the same exercise again $f(X)$ is also a semimartingale. On the event E , since $X = \tilde{X}$, by another exercise the RHS of (2.10) for \tilde{X} agrees with the RHS of (2.10) for X . So (2.10) holds for X a.s. on E . Since X is constant on E^c , (2.10) also holds for X on E^c . Thus, (2.10) holds a.s. for X .

Now we start the proof of the general case. For every $n \in \mathbb{N}$, let $\tau_n = \inf\{t \geq 0 : |X_t| \geq n\}$. Then X^{τ_n} is uniformly bounded on the event $E_n = \{|X_0| < n\} \in \mathcal{F}_0$, and is constant on E_n^c . From the last paragraph, we know that (2.10) holds for X^{τ_n} . So (2.10) holds for X up to τ_n . Since $\tau_n \uparrow \infty$, we get (2.10) throughout \mathbb{R}_+ . \square

Remark . 1. Itô's formula is a very powerful tool. The rest of this course can be viewed as a non-ending series of applications of Itô's formula.

2. The differentiability assumption of f can be relaxed. If some component X^j of X is of locally finite total variation, then we only need that f is C^1 in the j -th coordinate. The proof goes through just the same.
3. Itô's formula shows that the class of semimartingales is invariant under composition with C^2 -functions, which gives a reason for the introduction of semimartingales. If M is a local martingale, or even a martingale, $f(M)$ is usually not a local martingale, but only a semimartingale.

Example . We use (t) to denote the process which equals t at time t . Let B be a Brownian motion. Then $[B] = (t)$. Let σ and μ be two continuous adapted processes. Let $A = \mu \cdot (t)$, i.e.,

$A_t = \int_0^t \mu_s ds$. Then A is an adapted C^1 process with $A'_t = \mu_t$. Then we write $dA = \mu dt$. Let $M = \sigma \cdot B$. Then M is a local martingale with $[M] = \sigma^2 \cdot [B] = \sigma^2 \cdot (t)$. We write $dM = \sigma dB$. If a semimartingale $X = X_0 + \sigma \cdot B + \mu \cdot (t)$, then we write $dX = \sigma dB + \mu dt$. It satisfies $[X] = [\sigma \cdot B] = \sigma^2 \cdot (t)$. So we write $d[X] = \sigma^2 dt$. If $f \in C^2(\mathbb{R}, \mathbb{R})$, then $Y = f(X)$ is a semimartingale that satisfies

$$\begin{aligned} Y - Y_0 &= f'(X) \cdot X + \frac{1}{2} f''(X) \cdot [X] = f'(X) \cdot (\sigma \cdot B + \mu \cdot (t)) + \frac{1}{2} f''(X) \cdot (\sigma^2 \cdot (t)) \\ &= (f'(X)\sigma) \cdot B + (f'(X)\mu) \cdot (t) + \frac{1}{2} f''(X)\sigma^2 \cdot (t) = (f'(X)\sigma) \cdot B + (f'(X)\mu + \frac{1}{2} f''(X)\sigma^2) \cdot (t). \end{aligned}$$

So $dY = f'(X)\mu dB + (f'(X)\mu + \frac{1}{2} f''(X)\sigma^2)dt$, which means that Y has the same form as X .

Example . If $X = B = (B^1, \dots, B^d)$ is a d -dimensional Brownian motion, and $f \in C^2(\mathbb{R}^d, \mathbb{R})$, then since $[B^j, B^k]_t = \delta_{j,k}t$,

$$df(B) = \sum_{j=1}^d \partial_j f(B) dB_j + \frac{1}{2} \Delta f(B) dt.$$

Thus, if f is harmonic, then $f(B)$ is a local martingale.

Example (Complex Itô's Formula). A complex semimartingale is of the form $Z = X + iY$, where X and Y are real semimartingales. Its quadratic variation is $[Z] = [X + iY, X + iY] = [X] - [Y] + 2i[X, Y]$. Suppose $f : \mathbb{C} \rightarrow \mathbb{C}$ is analytic. Then $f(Z)$ is a complex semimartingale, and satisfies

$$df(Z) = f'(Z)dZ + \frac{1}{2} f''(Z)d[Z].$$

Here f' and f'' are first and second order complex derivatives of f . The formula also extends to complex vector semimartingales. The details are left as an exercise.

Example . We define the stochastic exponential of a semimartingale M as

$$\mathcal{E}(M) = \exp(M - \frac{1}{2}[M]).$$

Then $\mathcal{E}(M)$ satisfies $\mathcal{E}(M) = e^{M_0} + \mathcal{E}(M) \cdot M$, which may be also written as

$$\frac{d\mathcal{E}(M)}{\mathcal{E}(M)} = dM.$$

If M is a local martingale, then $\mathcal{E}(M)$ is also a local martingale. To derive the formula, let $f = e^x$ and $X = M - \frac{1}{2}[M]$, then by Itô's formula and that $f'' = f' = f$,

$$d\mathcal{E}(M) = f'(X)dX + \frac{1}{2} f''(X)d[X] = f(X)(dM - \frac{1}{2}d[X]) + \frac{1}{2} f(X)d[X] = f(X)dM = \mathcal{E}(M)dM.$$

For $\lambda \in \mathbb{C}$, we write $\mathcal{E}^\lambda(M)$ for the stochastic exponential of λM , then

$$\mathcal{E}^\lambda(M) = \exp(\lambda M - \frac{\lambda^2}{2}[M]), \quad \mathcal{E}^\lambda(M) = e^{\lambda M_0} + \lambda \mathcal{E}^\lambda(M) \cdot M.$$

So far, all local martingales or semimartingales are defined on \mathbb{R}_+ . We will relax this assumption and study continuous processes with (possibly finite) random lifetime. Let

$$\Sigma = \bigcup_{T \in (0, \infty]} C([0, T], \mathbb{R}).$$

For every $f \in \Sigma$, let $T(f) \in [0, \infty]$ be the lifetime of f , i.e., the domain of f is $[0, T(f))$. A continuous stochastic process with a random lifetime is a map $X : \Omega \rightarrow \Sigma$. We may also write X as X_t , $0 \leq t < T(X)$. Note that for each $t \geq 0$, X_t is only defined on the event $\{T(X) > t\}$. We say that X is \mathcal{F} -adapted if for every $t \geq 0$, $\{T(X) > t\} \in \mathcal{F}_t$, and X_t restricted to $\{T(X) > t\}$ is \mathcal{F}_t -measurable. In this case, $T(X)$ is a stopping time.

Definition . A random map $\tau : \Omega \rightarrow [0, \infty]$ is called a predictable time if there exists a sequence of stopping times $\tau_n \uparrow \tau$ such that $\tau_n < \tau$ when $\tau > 0$. The stopping times (τ_n) are said to announce τ .

As an increasing limit of stopping times, a predictable time must be a stopping time. If τ and σ are predictable times, then so are $\tau \wedge \sigma$ and $\tau \vee \sigma$ because if (τ_n) announce τ and (σ_n) announce σ , then $(\tau_n \wedge \sigma_n)$ and $(\tau_n \vee \sigma_n)$ announce $\tau \wedge \sigma$ and $\tau \vee \sigma$, respectively.

Example . Let X be a continuous process from \mathbb{R}_+ to \mathbb{R}^d , and F be a closed subset of \mathbb{R}^d . Then $\tau_F := \inf\{t \geq 0 : X_t \in F\}$ is a predictable time because there is a sequence of open sets $G_n \downarrow F$, and so the stopping times $\tau_{G_n} := \inf\{t \geq 0 : X_t \in G_n\}$ announce τ_F .

Definition . Suppose τ is a positive predictable time. Let X be a continuous stochastic process with lifetime τ . We say that X is a local martingale (resp. semimartingale) with lifetime τ , if there is a sequence of stopping times (τ_n) announcing τ such that for every n , $X_t^{\tau_n} = X_{\tau_n \wedge t}$, $0 \leq t < \infty$, is a local martingale (resp. semimartingale) defined on $[0, \infty)$. A vector semimartingale with lifetime τ is defined similarly.

A continuous process X with lifetime τ is a local martingale (resp. semimartingale) iff for any stopping time $\sigma < \tau$, X^σ is a local martingale (resp. semimartingale). The “if” part is obvious. For the “only if” part, let σ be a stopping time with $\sigma < \tau$. Suppose (τ_n) announces τ , and for every n , X^{τ_n} is a local martingale (resp. semimartingale). Define $\hat{\tau}_n = \tau_n$ if $\tau_n < \sigma$, and $\hat{\tau}_n = \infty$ if $\tau_n \geq \sigma$. Then for any $t \geq 0$, $\{\hat{\tau}_n < t\} = \{\tau_n < \sigma\} \cap \{\tau_n < t\} = \{\tau_n < \sigma \wedge t\} \in \mathcal{F}_{\sigma \wedge t} \subset \mathcal{F}_t$. Thus, $\hat{\tau}_n$ is a stopping time. Since $\tau_n \geq \sigma$ for n big enough, we have $\hat{\tau} \uparrow \infty$. Since for every n , $(X^\sigma)^{\hat{\tau}_n} = X^{\sigma \wedge \hat{\tau}_n} = X^{\sigma \wedge \tau_n} = (X_n^\tau)^\sigma$ is a local martingale (resp. semimartingale), X^σ is a local martingale (resp. semimartingale).

Thus, the definition of local martingale or semimartingale with lifetime τ does not depend on the choice of the sequence (τ_n) that announce τ ; and the set of local martingales (resp. semimartingales) with the same lifetime τ form a linear space.

Example . Let τ be a positive predictable time. Let X be a local martingale (resp. semimartingale) on \mathbb{R}_+ . Then $X|_{[0, \tau)}$ is a local martingale (resp. semimartingale) with lifetime τ because for any stopping time $\sigma < \tau$, $(X|_{[0, \tau)})^\sigma = X^\sigma$ is a local martingale (resp. semimartingale).

A lifetime- τ -local martingale M with locally finite total variation must be a.s. constant because for any stopping time $\sigma < \tau$, M^σ is a.s. constant by Proposition 15.2.

An adapted stochastic process A with lifetime τ will be simply called a finite variation process if $A_0 = 0$ and for each $t \in [0, \tau)$, A has finite total variation on $[0, t]$.

If X is a semimartingale with lifetime τ , then it has an a.s. unique canonical decomposition $X = M + A$ such that M is a local martingale with lifetime τ , and A is a finite variation process with lifetime τ . For the existence, suppose (τ_n) announces τ . Then for every n , X^{τ_n} is a semimartingale on \mathbb{R}_+ with canonical decomposition $M^n + A^n$. Let $n < m \in \mathbb{N}$. By the a.s. uniqueness of the decomposition of semimartingales on \mathbb{R}_+ , we have a.s. $M^n = (M^m)^{\tau_n}$ and $A^n = (A^m)^{\tau_n}$. Thus, we may define M and A with lifetime τ such that on an event E with probability 1, for each n , $M^{\tau_n} = M^n$, $A^{\tau_n} = A^n$; and on E^c , $M = X$ and $A = 0$. Then $X = M + A$, where M is a local martingale with lifetime τ , and A is a finite variation process with lifetime τ .

If M is a local martingale with lifetime τ , then M^2 is a semimartingale because for each stopping time $\sigma < \tau$, $(M^2)^\sigma = (M^\sigma)^2$ is a semimartingale. Let $N + A$ be a canonical decomposition of M^2 . Then we call A the quadratic variation of M , and denote it by $[M]$. For any stopping time $\sigma < \tau$, $N^\sigma + [M]^\sigma$ is a canonical decomposition of $(M^\sigma)^2$. So we have a.s. $[M]^\sigma = [M^\sigma]$, which is nondecreasing. Thus, $[M]$ is nondecreasing on $[0, \tau)$. For two local martingales M and N with lifetime τ , we define $[M, N] = ([M + N] - [M - N])/4$. If X and Y are lifetime- τ -semimartingales with canonical decomposition $M + A$ and $N + B$, then we define $[X] = [M]$ and $[X, Y] = [M, N]$.

For a finite variation process A with lifetime τ , we define $L(A)$ to be the set of progressive stochastic processes U with lifetime τ , such that for any $t \in [0, \tau)$, $\int_0^t |U_s| |dA_s| < \infty$. Here by saying that U is progressive, we mean that for any fixed $t_0 \geq 0$, the set $D_{t_0} := \{(\omega, t) \in \Omega \times [0, t_0] : t < T(U(\omega))\} \in \mathcal{F}_{t_0} \times \mathcal{B}_{[0, t_0]}$, and $U|_{D_{t_0}}$ is $\mathcal{F}_{t_0} \times \mathcal{B}_{[0, t_0]}$ -measurable. This is the case if U is adapted and continuous. For every $U \in L(A)$, we may define $(U \cdot A)_t = \int_0^t U_s dA_s$, $0 \leq t < \tau$, which is a finite variation process with lifetime τ . Note that for every stopping time $\sigma < \tau$, $U^\sigma \in L(A^\sigma)$, and $U^\sigma \cdot A^\sigma = (U \cdot A)^\sigma$.

For a local martingale M with lifetime τ , we define $L(M)$ to be the set of progressive processes U with lifetime τ such that $U^2 \in L([M])$. If $U \in L(M)$, then for every stopping time $\sigma < \tau$, $(U^\sigma)^2 = (U^2)^\sigma \in L([M]^\sigma)$. So $U^\sigma \in L(M^\sigma)$, and we may define $U^\sigma \cdot M^\sigma$. Suppose (τ_n) announce τ . Then the family $U^{\tau_n} \cdot M^{\tau_n}$, $n \in \mathbb{N}$, is consistent, i.e., for any $n < m$, a.s. $U^{\tau_n} \cdot M^{\tau_n} = (U^{\tau_m} \cdot M^{\tau_m})^{\tau_n}$. So we may define the local martingale $U \cdot M$ with lifetime τ such that for each n , a.s. $(U \cdot M)^{\tau_n} = U^{\tau_n} \cdot M^{\tau_n}$. If $\sigma < \tau$ is a stopping time, then a.s. for any n , $((U \cdot M)^\sigma)^{\tau_n} = (U^{\tau_n} \cdot M^{\tau_n})^\sigma = (U^\sigma \cdot M^\sigma)^{\tau_n}$. Since $\tau = \lim \tau_n > \sigma$, we get a.s. $(U \cdot M)^\sigma = (U^\sigma) \cdot M^\sigma$. If N is another local martingale with lifetime τ , a similar argument shows that a.s. $[U \cdot M, N] = U \cdot [M, N]$.

For a lifetime- τ -semimartingale X with the canonical decomposition $M + A$, let $L(X) = L(M) \cap L(A)$. For $U \in L(X)$, we define $U \cdot X = U \cdot M + U \cdot A$. Then $U \cdot X$ is a lifetime- τ -semimartingale with canonical decomposition $U \cdot M + U \cdot A$. It is the a.s. unique process with lifetime τ that satisfies the property that for any stopping time $\sigma < \tau$, $(U \cdot X)^\sigma = U^\sigma \cdot X^\sigma$.

Corollary 15.20 (Local Itô's Formula). *Fix an open set $D \subset \mathbb{R}^d$. Let $f \in C^2(D, \mathbb{R})$. Let X be a vector semimartingale with random lifetime τ such that $X_t \in D$ for all $t \in [0, \tau)$. Then $f(X)$ is a semimartingale with lifetime τ , and Itô's formula holds up to τ , and a.s.*

$$f(X) = f(X_0) + \sum_{j=1}^d \partial_j f(X) \cdot X^j + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_j \partial_k f(X) \cdot [X^j, X^k], \quad \text{on } [0, \tau). \quad (2.11)$$

Here the meaning of “local” is twofold: first, X is defined not on $[0, \infty)$, but on $[0, \tau)$; second, f is not defined on \mathbb{R}^d , but on $D \subset \mathbb{R}^d$.

Proof. The main idea of the proof is “localization”. Suppose τ_n is a sequence of stopping times announcing τ . Let G_n be an increasing sequence of bounded open subsets of D such that $D = \bigcup G_n$, and $\overline{G_n} \subset D$ for each $n \in \mathbb{N}$. Let $\tau_{G_n} = \inf(\{\tau\} \cup \{t \geq 0 : X_t \notin G_n\})$, $n \in \mathbb{N}$. Let $\sigma_n = \tau_n \wedge \tau_{G_n}$, $n \in \mathbb{N}$. Then (σ_n) also announces τ . Fix $n \in \mathbb{N}$. Then X^{σ_n} is contained in $\overline{G_n}$ on the event $E_n := \{X_0 \in G_n\} \in \mathcal{F}_0$, and is constant on E_n^c . Pick $x_0 \in G_0$. Define Y such that $Y = X^{\sigma_n}$ on E_n , and $Y \equiv x_0$ on E_n^c . Then Y is contained in $\overline{G_n}$. By an exercise, Y^n is also a vector semimartingale. We may define a function $f_n \in C^2(\mathbb{R}^d, \mathbb{R})$ such that f_n agrees with f in G_{n+1} . Applying the global Itô's formula to Y and f_n and using the fact that $f, \partial_j f, \partial_j \partial_k f$ respectively agree with $f_n, \partial_j f_n, \partial_j \partial_k f_n$ on the range of Y^n , we find that $f(Y) = f_n(Y)$ is a semimartingale, and a.s.

$$f(Y) = f(Y_0) + \sum_{j=1}^d \partial_j f(Y) \cdot Y^j + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_j \partial_k f(Y) \cdot [Y^j, Y^k]. \quad (2.12)$$

Since $f(Y)$ agrees with $f(X^{\sigma_n})$ on E_n , and $f(X^{\sigma_n})$ is constant on E_n^c . By the exercise again, $f(X^{\sigma_n})$ is also a semimartingale. Since X^{σ_n} agrees with Y on the event E_n , by another exercise, the RHS of (2.12) agrees a.s. on E_n with the same formula with X^{σ_n} in place of Y . Thus, a.s. on E_n ,

$$f(X)^{\sigma_n} = f(X_0) + \sum_{j=1}^d (\partial_j f(X) \cdot X^j)^{\sigma_n} + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_j \partial_k f(X) \cdot [X^j, X^k]^{\sigma_n}. \quad (2.13)$$

Since X^{σ_n} is constant on E_n^c , (2.11) also holds on E_n^c . Thus, for any n , (2.13) a.s. holds. This means that (2.11) a.s. holds on $[0, \sigma_n)$. Since σ_n announce τ , we see that (2.11) a.s. holds throughout $[0, \tau)$. \square

From now on, a local martingale or semimartingale may have infinite or finite (and random) lifetime.

Example . The quotient of two semimartingales X/Y is a semimartingale (assuming that Y does not take value 0). To see this, write $X/Y = f(X, Y)$, where $f(x, y) := x/y$ is C^2 on $\mathbb{R} \times (\mathbb{R} \setminus \{0\})$.

Example . Suppose X is a positive semimartingale, and $\alpha \in \mathbb{R}$. Then X^α is a semimartingale. To see this, we apply the local Itô's formula to $f(x) = x^\alpha$ and $D = (0, \infty)$. Since $f'(x) = \frac{\alpha}{x}f(x)$ and $f''(x) = \frac{\alpha(\alpha-1)}{x^2}f(x)$, we get

$$dX^\alpha = f'(X)dX + \frac{1}{2}f''(X)d[X] = \frac{\alpha}{X}X^\alpha dX + \frac{1}{2}\frac{\alpha(\alpha-1)}{X^2}X^\alpha d[X].$$

We may rewrite this formula as

$$\frac{dX^\alpha}{X^\alpha} = \alpha \frac{dX}{X} + \frac{1}{2}\alpha(\alpha-1)\frac{d[X]}{X^2}.$$

Example . Let D be a domain in \mathbb{R}^d . Let B be a Brownian motion in \mathbb{R}^d (started from $x \in \mathbb{R}^d$). Let τ be the first time that B exits D . Let $f : D \rightarrow \mathbb{R}$ be harmonic. Then $f(B)|_{[0,\tau]}$ is a local martingale with lifetime τ . In fact, by Itô's formula, and that $[B^j, B^k]_t = \delta_{j,k}t$, we have

$$df(B) = \sum_{j=1}^d \partial_j f(B)dB^j + \frac{1}{2}\Delta f(B)dt = \sum_{j=1}^d \partial_j f(B)dB^j.$$

We now study time-changes. Let T be a positive predictable time. Suppose u is an adapted, continuous, and strictly increasing process with lifetime T and $u_0 = 0$. This is the case, for example, if $u_t = \int_0^t \rho_s ds$, where ρ_t , $0 \leq t < T$, is a positive continuous adapted process. Let $S = \sup\{u_t : t \in [0, T)\}$. Then u is a homeomorphism from $[0, T)$ onto $[0, S)$. Let $v : [0, S) \rightarrow [0, T)$ be the inverse of u . We further extend v to $[0, \infty)$ such that $v_s = \infty$ for $s \in [S, \infty)$. Then for every $s \geq 0$, v_s is an \mathcal{F} -stopping time because

$$v_s = \inf\{t \in [0, T) : u_t > s\}, \quad s \geq 0. \quad (2.14)$$

, and for $t > 0$, $\{v_s < t\} = \{T > t\} \cap \{u_t > s\} \in \mathcal{F}_t$. We call the v a time-change.

For each $s \geq 0$, since v_s is a stopping time, we have a σ -algebra \mathcal{F}_{v_s} . Since v_s is increasing in s , we get a new filtration $\widehat{\mathcal{F}} := (\mathcal{F}_{v_s})_{s \geq 0}$, called the filtration induced by the time-change v . Since \mathcal{F} is right-continuous, and $v_{s_n} \downarrow v_{s_0}$ if $s_n \downarrow s$, $\widehat{\mathcal{F}}$ is also right-continuous. If τ is an \mathcal{F} -stopping time with $\tau < T$, then u_τ is an $\widehat{\mathcal{F}}$ -stopping time. To see this, note that for any $a, b \geq 0$,

$$\{u_\tau < a\} \cap \{v_a < b\} = \{u_\tau < a\} \cap \{\tau < b\} \cap \{a < u_b\} \cap \{b < T\} \in \mathcal{F}_b,$$

which implies that $\{u_\tau < a\} \in \mathcal{F}_{v_a} = \widehat{\mathcal{F}}_a$ for any $a \geq 0$. If τ_n is a sequence of \mathcal{F} -stopping times announcing T , then u_{τ_n} is a sequence of $\widehat{\mathcal{F}}$ -stopping times announcing S . So S is a positive $\widehat{\mathcal{F}}$ -predictable time.

Suppose X is a left- or right- continuous \mathcal{F} -adapted process with lifetime T . Then X is progressive. By Lemma 6.5, for any $s \geq 0$, X_{v_s} restricted to $\{s < S\}$ is \mathcal{F}_{v_s} -measurable. So $X \circ v$ is $\widehat{\mathcal{F}}$ -adapted. We then define the $\widehat{\mathcal{F}}$ -adapted process \widehat{X} with lifetime S by $\widehat{X}_s = X_{v_s}$, $0 \leq s < S$, and call it the time-change of X via v . Since v is continuous, $\widehat{X} = X \circ v$ is also a

left- or right-continuous process. On the other hand, now v_s , $0 \leq s < S$, is a continuous and strictly increasing $\widehat{\mathcal{F}}$ -adapted process with $v(0) = 0$, and u is the inverse of v . The time-change of \widehat{X} via u is just the original X . Moreover, we have $\widehat{\mathcal{F}}_{u(t)} = \mathcal{F}_t$, $t \geq 0$. So the status of the two triples (X, T, \mathcal{F}) and $(\widehat{X}, S, \widehat{\mathcal{F}})$ are symmetric.

Exercise. Check that the two filtrations $\widehat{F}_{u(\cdot)}$ and \mathcal{F} agree.

Theorem 15.25 (A modified version). *Let X be an \mathcal{F} -semimartingale with lifetime T with canonical decomposition $M + A$. Let $\widehat{X}, \widehat{M}, \widehat{A}, [\widehat{X}]$ be the time-changes of $X, M, A, [X]$, respectively, via v . Then \widehat{X} is an $\widehat{\mathcal{F}}$ -semimartingale with lifetime S with canonical decomposition $\widehat{M} + \widehat{A}$, and the quadratic variation of \widehat{X} is $[\widehat{X}]$. Moreover, if U is a right- or left- continuous \mathcal{F} -adapted process with lifetime T and $U \in L(X)$, then the time-change of U via v , i.e., $\widehat{U} := U \circ v$ belongs to $L(\widehat{X})$, and a.s. $\widehat{U} \cdot \widehat{X}$ is the time-change of $U \cdot X$ through v .*

Proof. Since A has finite variation, it can be expressed as $A_+ - A_-$, where A_+, A_- are increasing. Then $\widehat{A} = \widehat{A}_+ - \widehat{A}_-$, where $\widehat{A}_\pm := A_\pm \circ v$ are both increasing. So \widehat{A} also has finite variation. Since A is continuous and \mathcal{F} -adapted, \widehat{A} is $\widehat{\mathcal{F}}$ -adapted. Thus, \widehat{A} is an $\widehat{\mathcal{F}}$ -finite variation process. To prove that \widehat{M} is an $\widehat{\mathcal{F}}$ -local martingale, we may assume $M_0 = \widehat{M}_0 = 0$. Then we can find a sequence τ_n announcing T such that for each n , M^{τ_n} is uniformly bounded. By Optional Stopping Theorem (for uniformly integrable martingale), for any $s_2 \geq s_1 \geq 0$,

$$\mathbb{E}[M_{v_{s_2}}^{\tau_n} | \widehat{F}_{s_1}] = \mathbb{E}[M_{v_{s_2}}^{\tau_n} | \widehat{F}_{v_{s_1}}] = M_{v_{s_1}}^{\tau_n}.$$

Thus,

$$(\widehat{M})_s^{u_{\tau_n}} = \widehat{M}_{u_{\tau_n} \wedge s} = M_{v(u_{\tau_n} \wedge s)} = M_{\tau_n \wedge v_s} = M_{v_s}^{\tau_n}$$

is an $\widehat{\mathcal{F}}$ -martingale. Since u_{τ_n} announce S , \widehat{M} is an $\widehat{\mathcal{F}}$ -local martingale with lifetime S . Since $\widehat{X} = \widehat{M} + \widehat{A}$, \widehat{X} is an $\widehat{\mathcal{F}}$ -semimartingale with canonical decomposition $\widehat{M} + \widehat{A}$. Since $[\widehat{X}]$ is the finite variation component of the canonical decomposition of \widehat{M}^2 , we see that a.s. $[\widehat{X}]$ is the time-change of $[M]$ via v . Then we immediately get that if X and Y are semimartingales with lifetime T , and $\widehat{X} = X \circ v$ and $\widehat{Y} = Y \circ v$, then $[\widehat{X}, \widehat{Y}] = [X, Y] \circ v$.

Let U be a right- or left- continuous \mathcal{F} -adapted process with lifetime T . Then \widehat{U} is a right- or left- continuous $\widehat{\mathcal{F}}$ -adapted process. For any $0 \leq s < S$, $\int_0^s |\widehat{U}| |d\widehat{A}| = \int_0^{v_s} |U| |dA|$ since the signed measure $\widehat{\mu}$ on $[0, S)$ determined by \widehat{A} is related to the signed measure μ on $[0, T)$ determined by A by $\widehat{\mu} = \mu \circ v$. Thus, if $U \in L(A)$, then $\widehat{U} \in L(\widehat{A})$ and $\widehat{U} \cdot \widehat{A} = (U \cdot A) \circ v$.

If $U \in L(M)$, then $U^2 \in L([M])$, and so $\widehat{U}^2 \in L([\widehat{M}]) = L([M])$. So $\widehat{U} \in L(\widehat{M})$. Moreover, if \widehat{N} is an $\widehat{\mathcal{F}}$ -local martingale with lifetime S , then $N := \widehat{N} \circ v$ is an \mathcal{F} -local martingale with lifetime T , and

$$[(U \cdot M) \circ v, \widehat{N}] = [U \cdot M, N] \circ v = (U \cdot [M, N]) \circ v = \widehat{U} \cdot ([M, N] \circ v) = \widehat{U} \cdot ([\widehat{M}, \widehat{N}]).$$

Thus, a.s. $\widehat{U} \cdot \widehat{M} = (U \cdot M) \circ v$. So if $U \in L(X) = L(A) \cap L(M)$, then $\widehat{U} \in L(\widehat{A}) \cap L(\widehat{M}) = L(\widehat{X})$, and a.s. $\widehat{U} \cdot \widehat{X} = (U \cdot X) \circ v$. \square

3 Continuous Martingales and Brownian Motion

Theorem 16.3 (Lévy's Characterization of Brownian Motion). *For a continuous d -dimensional adapted process $X = (X^1, \dots, X^d)$ with $X_0 = 0$, the following are equivalent.*

- (i) X is a d -dimensional \mathcal{F} -Brownian motion.
- (ii) X is a local martingale, and $[X^j, X^k]_t = \delta_{j,k}t$, $t \geq 0$, for every $1 \leq j, k \leq d$.

Proof. We already know that (i) implies (ii). Suppose (ii) is true. Fix $v = (v_1, \dots, v_d) \in \mathbb{R}^d$ and $t_1 \geq 0$. Let $(v, X) = \sum_{j=1}^d v_j X^j$. Then (v, X) is a local martingale, and $[(v, X)]_t = \sum_{j=1}^d v_j^2 [X^j, X^j]_t = |v|^2 t$. Let

$$M_t = \mathcal{E}^i((v, X)_t) = \exp\left(i(v, X)_t - \frac{i^2}{2}[(v, X)]_t\right) = \exp\left(i(v, X)_t + \frac{1}{2}|v|^2 t\right), \quad t \geq 0.$$

Then M is the stochastic exponential of $i(v, X)$, and so is also a local martingale. We have $|M_t| = e^{\frac{1}{2}|v|^2 t}$. So for any $t_1 \geq 0$, M^{t_1} is uniformly bounded on $[0, t_1]$, and so is a true martingale. Let $t_0 \in [0, t_1]$. Then $\mathbb{E}[M_{t_1} | \mathcal{F}_{t_0}] = M_{t_0}$. So for any $A \in \mathcal{F}_{t_0}$,

$$\mathbb{E}[\mathbf{1}_A \exp(-i(v, X_{t_0}) \exp(i(v, X_{t_1}) + \frac{1}{2}|v|^2 t_1))] = \mathbb{E}[\mathbf{1}_A \exp(-i(v, X_{t_0}) \exp(i(v, X_{t_0}) + \frac{1}{2}|v|^2 t_0))],$$

which reduces to

$$\mathbb{E}[\mathbf{1}_A \exp(i(v, X_{t_1} - X_{t_0}))] = \exp\left(-\frac{1}{2}|v|^2(t_1 - t_0)\right) \mathbb{P}[A].$$

Since this holds for any $A \in \mathcal{F}_{t_0}$ and $v \in \mathbb{R}^d$, we conclude that $X_{t_1} - X_{t_0}$ is independent of \mathcal{F}_{t_0} , and has characteristic function $v \mapsto \exp\left(-\frac{1}{2}|v|^2(t_1 - t_0)\right)$. So $X_{t_1} - X_{t_0}$ is a Gaussian vector with covariation matrix $(t_1 - t_0)I_d$ independent of \mathcal{F}_{t_0} . Since this holds for any $t_1 \geq t_0 \geq 0$, X is a d -dimensional \mathcal{F} -Brownian motion. \square

Remark . If we do not assume that $X_0 = 0$, then (ii) is equivalent to that X is an \mathcal{F} -Brownian motion started from some point.

If B is a d -dimensional Brownian motion, and T is a positive predictable time, then $X := B|_{[0, T)}$ is a vector local martingale with lifetime T , and for any $1 \leq j, k \leq d$, $[X^j, X^k]_t = \delta_{j,k}t$, $0 \leq t < T$. On the other hand, if $X = (X^1, \dots, X^d)$ is a d -dimensional local martingale with lifetime T such that for any $1 \leq j, k \leq d$, $[X^j, X^k]_t = \delta_{j,k}t$, $0 \leq t < T$, then a natural question to ask is whether X a.s. extends to a d -dimensional Brownian motion. In general, the answer is no. It is simply because the probability space may not be big enough to support a Brownian motion with full range. For example, we start with a Brownian motion B_t , $0 \leq t < \infty$, defined on $(\Omega, \mathcal{A}, \mathbb{P})$, which generates filtration \mathcal{F} . Let $X = B|_{[0, 1)}$. Let $\mathcal{B} = \mathcal{F}_1$, and restrict our attention to the new probability space $(\Omega, \mathcal{B}, \mathbb{P})$. Define a filtration $\widehat{\mathcal{F}}$ on (Ω, \mathcal{B}) by $\widehat{\mathcal{F}}_t = \mathcal{F}_{t \wedge 1}$,

$t \geq 0$. Now X is an $\widehat{\mathcal{F}}$ -adapted continuous local martingale with lifetime 1, and $[X]_t = t$ for $0 \leq t < 1$. But we can not extend X to a Brownian motion on $(\Omega, \mathcal{B}, \mathbb{P})$ because if such extension \widehat{B} exists, then on the one hand, $\widehat{B}_2 - \widehat{B}_1$ is \mathcal{B} -measurable, and on the other hand, it is independent of X_t , $0 \leq t < 1$, which generates \mathcal{B} . Then $\widehat{B}_2 - \widehat{B}_1$ must be a.s. constant, which is not possible.

We may overcome the issue by expanding the probability space and filtration. Recall that we have been working on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ throughout, where $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ is a filtration. By an enlargement of $(\Omega, \mathcal{F}, \mathbb{P})$, we mean another filtered probability space $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ together with a map $\pi : \widehat{\Omega} \rightarrow \Omega$ such that $\mathbb{P} = \widehat{\mathbb{P}} \circ \pi^{-1}$, and for any t , $\pi^{-1}(\mathcal{F}_t) \subset \widehat{\mathcal{F}}_t$. A process X defined on Ω may be viewed as defined on $\widehat{\Omega}$ by setting $\widehat{X}(\widehat{\omega}) = X(\pi(\omega))$. Since π is $(\widehat{\mathcal{F}}_t, \mathcal{F}_t)$ -measurable, if X is \mathcal{F} -adapted, then \widehat{X} is $\widehat{\mathcal{F}}$ -adapted. For any event E ,

$$\widehat{\mathbb{P}} \circ \widehat{X}^{-1} = \widehat{\mathbb{P}} \circ (X \circ \pi)^{-1}(E) = \widehat{\mathbb{P}} \circ \pi^{-1} \circ X^{-1}(E) = \mathbb{P} \circ X^{-1}(E).$$

So \widehat{X} (with underlying measure $\widehat{\mathbb{P}}$) has the same law as X (with underlying measure \mathbb{P}).

Theorem . *For the above vector local martingale X with $[X^j, X^k]_t = \delta_{j,k}t$, $0 \leq t < T$, there is an enlargement $(\widehat{\Omega}, \widehat{\mathcal{F}}, \widehat{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$, and a d -dimensional $\widehat{\mathcal{F}}$ -Brownian motion B defined on $\widehat{\Omega}$ such that a.s. $X_t = X_0 + B_t$, $0 \leq t < T$.*

Proof. We may assume that $X_0 = 0$. First, we show that X a.s. extends to a continuous martingale Y with lifetime ∞ without enlarging the probability space. We will show that on the event $T < \infty$, a.s. $\lim_{t \uparrow T} X_t$ converges, and if we define $Y_t = X_t$ on $[0, T)$, and $Y_t = \lim_{t \uparrow T} X_t$ on $[T, \infty)$, then Y is a continuous vector martingale.

There is a sequence of stopping times (τ_n) that announces T such that for each n , $|X^{\tau_n}|$ is bounded by n . Fix $j \in \{1, \dots, d\}$ and $m \in \mathbb{N}$. For any $n_1 < n_2 \in \mathbb{N}$,

$$[(X^j)^{m \wedge \tau_{n_2}} - (X^j)^{m \wedge \tau_{n_1}}]_t = [(X^j)^{m \wedge \tau_{n_2}}]_t - [(X^j)^{m \wedge \tau_{n_1}}]_t = m \wedge \tau_{n_2} \wedge t - m \wedge \tau_{n_1} \wedge t.$$

Thus, $[(X^j)^{m \wedge \tau_{n_2}} - (X^j)^{m \wedge \tau_{n_1}}]_\infty \leq m \wedge \tau_{n_2} - m \wedge \tau_{n_1}$. As $n_1, n_2 \rightarrow \infty$, $\tau_{n_1}, \tau_{n_2} \rightarrow T$. By Dominated Convergence Theorem, $\mathbb{E}[(X^j)^{m \wedge \tau_{n_2}} - (X^j)^{m \wedge \tau_{n_1}}]_\infty \rightarrow 0$ as $n_1, n_2 \rightarrow \infty$. Since for $M \in \mathcal{M}^2$, $\|M\|_{\mathcal{M}^2}^2 = \mathbb{E}[M_\infty^2] = \mathbb{E}[M]_\infty$, $(X^j)^{m \wedge \tau_n}$, $n \in \mathbb{N}$, form a Cauchy sequence in \mathcal{M}^2 .

Let $Y^{j,m} \in \mathcal{M}^2$ be its limit. Then for each $t \geq 0$, $(X^j)_t^{m \wedge \tau_n} \xrightarrow{P} Y_t^{j,m}$ as $n \rightarrow \infty$. For every $t \in \mathbb{Q}_+ \cap [0, m)$, on the event $\{t < T\}$, since $\tau_n \uparrow T$, we have $(X^j)_t^{m \wedge \tau_n} = X_t^j$ for n big enough, which then implies that $Y_t^{j,m} = X_t^j$ a.s. on the event $\{t < T\}$. Since both $Y^{j,m}$ and X^j are continuous, we see that a.s. $X^j = Y^{j,m}$ on $[0, T \wedge m)$. Fix $t_1 < t_2 \in \mathbb{Q}_+$. For each n , since $\tau_n < T$, $(X^j)_{t_1}^{m \wedge \tau_n} = (X^j)_{t_1}^{m \wedge \tau_n}$ on $\{T \leq t_1\}$. Thus, $Y_{t_1}^{j,m} = Y_{t_2}^{j,m}$ a.s. on $\{T \leq t_1\}$. Then a.s. for any $t_1 < t_2 \in \mathbb{Q}_+$, $Y_{t_1}^{j,m} = Y_{t_2}^{j,m}$ on $\{T \leq t_1\}$. Since $Y^{j,m}$ is continuous, we see that it is a.s. constant on $[T, \infty)$.

Now we still fix j , but let m vary. For $m_1 \leq m_2$, from $(X^j)^{m_1 \wedge \tau_n} = ((X^j)^{m_2 \wedge \tau_n})^{m_1}$ we get a.s. $Y^{j,m_1} = (Y^{j,m_2})^{m_1}$. This means that the family $Y^{j,m}$, $m \in \mathbb{N}$, are consistent, and we may define a process Y^j on \mathbb{R}_+ such that for any $m \in \mathbb{N}$, $Y^j|_{[0,m]}$ a.s. agrees with $Y^{j,m}|_{[0,m]}$. Since each $Y^{j,m}$ is a continuous martingale on \mathbb{R}_+ , Y^j is also a continuous martingale on \mathbb{R}_+ . Since

each $Y^{j,m}$ a.s. takes constant value on $[T, \infty)$, so does Y^j . Since for every m , $Y^{j,m}$ a.s. agrees with X^j on $[0, T \wedge m)$, we see that Y^j a.s. agrees with X^j on $[0, T)$. Then $Y = (Y^1, \dots, Y^d)$ is the extension that we want. Since a.s. Y agrees with X on $[0, T)$, and is constant on $[T, \infty)$, we have

$$[Y^j, Y^k]_t = \delta_{j,k} t \wedge T, \quad 1 \leq j, k \leq d, \quad t \geq 0. \quad (3.1)$$

Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ be a filtered probability space that supports a d -dimensional Brownian motion $\tilde{\beta} = (\tilde{\beta}^1, \dots, \tilde{\beta}^d)$. Consider the product space

$$\hat{\Omega} = \Omega \times \tilde{\Omega}, \quad \hat{\mathcal{F}}_t = \mathcal{F}_t \times \tilde{\mathcal{F}}_t, \quad \hat{\mathbb{P}} = \mathbb{P} \times \tilde{\mathbb{P}}.$$

Then $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ with the map $\pi(\omega, \tilde{\omega}) = \omega$ is the enlargement of $(\Omega, \mathcal{F}, \mathbb{P})$ that we need. The process Y on Ω is then viewed as a process defined on $\hat{\Omega}$ by $Y_t(\omega, \tilde{\omega}) = Y_t(\omega)$. By the independence of \mathcal{F} and $\tilde{\mathcal{F}}$, we see that the new Y is still a martingale. By Proposition 15.18, the new Y also satisfies (3.1). We also understand $\tilde{\beta}$ as a process on $\hat{\Omega}$ by $\tilde{\beta}(\omega, \tilde{\omega}) = \tilde{\beta}(\tilde{\omega})$. Then $\tilde{\beta}$ is an $\hat{\mathcal{F}}$ -Brownian motion on $\hat{\Omega}$ independent of Y . So for any $1 \leq j, k \leq d$,

$$[\tilde{\beta}^j, \tilde{\beta}^k]_t = \delta_{j,k} t, \quad [\tilde{\beta}^j, Y^k]_t = 0, \quad t \geq 0.$$

Since T is an \mathcal{F} -stopping time, it is also an $\hat{\mathcal{F}}$ -stopping time. So for $1 \leq j, k \leq d$ and $t \geq 0$, $(\tilde{\beta}^k)^T$ is a local martingale, and

$$[\tilde{\beta}^j, (\tilde{\beta}^k)^T] = [(\tilde{\beta}^j)^T, (\tilde{\beta}^k)^T] = \delta_{j,k} t \wedge T, \quad [(\tilde{\beta}^j)^T, Y^k]_t = [\tilde{\beta}^j, Y^k]_t^T = 0,$$

which implies that

$$[\tilde{\beta}^j - (\tilde{\beta}^j)^T, \tilde{\beta}^k - (\tilde{\beta}^k)^T]_t = \delta_{j,k} (t - t \wedge T), \quad [\tilde{\beta}^j - (\tilde{\beta}^j)^T, Y^k] = 0.$$

Define $B = Y + (\tilde{\beta} - \tilde{\beta}^T)$. Then B is an $\hat{\mathcal{F}}$ -vector local martingale with lifetime ∞ , and for $1 \leq j, k \leq d$,

$$[B^j, B^k]_t = [Y^j, Y^k]_t + [\tilde{\beta}^j - (\tilde{\beta}^j)^T, \tilde{\beta}^k - (\tilde{\beta}^k)^T]_t = \delta_{j,k} t, \quad t \geq 0.$$

By Lévy's Characterization of Brownian Motion, B is an $\hat{\mathcal{F}}$ -Brownian motion. Since $\tilde{\beta} - \tilde{\beta}^T = 0$ on $[0, T)$, B agrees with Y on $[0, T)$, and so a.s. agrees with X on $[0, T)$. \square

Because of this theorem, we may call a vector local martingale X_t , $0 \leq t < T$, that satisfies $[X^j, X^k]_t = \delta_{j,k} t$, $0 \leq t < T$, for any $1 \leq j, k \leq d$, a stopped d -dimensional Brownian motion.

Suppose now M is a (one-dimensional) local martingale with lifetime T such that $[M]$ is strictly increasing, i.e., M does not take constant value on any time interval. Let $u_t = [M]_t$. Then u is a continuous and strictly increasing adapted process. Let v be the inverse of u , and let \widehat{M} be the time-change of M via v . Then \widehat{M} is a local martingale with lifetime $S := \sup_{0 \leq t < T} [M]_t$, and $[\widehat{M}]_s = [M]_{v_s} = u_{v_s} = s$, $0 \leq s < S$. By the above theorem, we know that in an enlarged probability space, \widehat{M} extends to a Brownian motion. So M is a time-change of a stopped Brownian motion. Since a Brownian motion B is continuous on $[0, \infty)$ and satisfies

that a.s. $\limsup_{t \rightarrow \infty} \widehat{B}_t = \infty$ and $\liminf_{t \rightarrow \infty} \widehat{B}_t = -\infty$, we can then conclude that a.s. on the event that $\sup_{0 \leq t < T} [M]_t = \infty$, $\limsup_{t \uparrow T} M_t = \infty$ and $\liminf_{t \uparrow T} M_t = -\infty$; and on the event that $\sup_{0 \leq t < T} [M]_t < \infty$, $\lim_{t \uparrow T} M(t)$ converges to a finite number. Thus, on the event that M is bounded from either above or below, a.s. $\lim_{t \uparrow T} M(t)$ converges.

The above result also holds even if M does take constant values on some intervals. In that case, we still define $u_t = [M]_t$, and define v_s by $v_s = \inf\{t : u_t > s\}$. Now v may not be continuous. However, since every discontinuity of v corresponds to an interval of constancy of $u = [M]$, on which M is constant. So $\widehat{M}_s := M_{v_s}$ is still a continuous local martingale. We then still have $[\widehat{M}]_s = s$, $0 \leq s < S$, and so \widehat{M} is a stopped Brownian motion. Then the original M is a time-change of a stopped Brownian motion with possible pauses. So the result at the end of the previous paragraph still holds. The above argument works only for one-dimensional local martingale. We can not always transform a vector local martingale to a stopped (multi-dimensional) Brownian motion using a time-change.

Suppose that M is a local martingale such that $[M]$ is C^1 with positive derivative. Besides time-change, there is another way to obtain a (stopped) Brownian motion from M , which is a stochastic integral against M . Let $\sigma_t = \sqrt{d[M]_t/dt} > 0$, and define $B_t = \int_0^t \sigma_s^{-1} dM_s$. Then we have $d[B]_t = \sigma_t^{-2} d[M]_t = 1$ (up to its lifetime). So B is a stopped Brownian motion, and $M = M_0 + \sigma \cdot B$. We may then write $dM = \sigma dB$.

We are interested in the kind of semimartingales X (with lifetime T) with canonical decomposition $M + A$ such that there are continuous adapted processes σ and μ such that $M = M_0 + \sigma \cdot B$ and $A = \mu \cdot (t)$. Then we may express X in the following form:

$$dX_t = \sigma_t dB_t + \mu_t dt.$$

If X takes values in an open set $U \subset \mathbb{R}$, and $f \in C^2(U, \mathbb{R})$, then by Itô's formula, $Y_t := f(X_t)$ satisfies the equation $dY_t = f'(X_t) \sigma_t dB_t + (f'(X_t) \mu_t + \frac{1}{2} f''(X_t) \sigma_t^2) dt$.

Let ρ_t , $0 \leq t < T$, be a positive continuous adapted process. Let $u_t = \int_0^t \rho_s ds$, $0 \leq t < T$. Let v be the inverse of u . Then $v'_s = 1/\widehat{\rho}_s$, where $\widehat{\rho}_s = \rho_{v_s}$. Let $\widehat{X}, \widehat{M}, \widehat{A}$ be the time-changes of X, M, A via v . Then from $\widehat{A} = A \circ v$, we find that $d\widehat{A}_s = \widehat{\mu}_s v'_s ds$, where $\widehat{\mu}_s = \mu_{v_s}$. From $\widehat{M} = M \circ v$, we find that $[\widehat{M}]_s = [M] \circ v$, and so $d[\widehat{M}]_s = d[M]_{v_s} v'_s ds = \widehat{\sigma}_s^2 v'_s ds$, where $\widehat{\sigma}_s = \sigma_{v_s}$. So there is **another** Brownian motion \widehat{B} such that $\widehat{X}_s = X_{v_s}$ satisfies the SDE:

$$d\widehat{X}_s = \frac{\widehat{\sigma}_s}{\sqrt{\widehat{\rho}_s}} d\widehat{B}_s + \frac{\widehat{\mu}_s}{\widehat{\rho}_s} ds.$$

Note the two additional factors $1/\sqrt{\widehat{\rho}_s}$ and $1/\widehat{\rho}_s$.

A two-dimensional Brownian motion $B = (B^1, B^2)$ may also be viewed as a complex Brownian motion: $B_t = B_t^1 + iB_t^2$.

Theorem (Conformal Invariance of Complex Brownian Motion). *Let $U \subset \mathbb{C}$ be open, and let $f : U \rightarrow \mathbb{C}$ be analytic and injective. Let B be a complex Brownian motion started from some point in $z_0 \in U$. Let τ_U be the first time that B exits U . Then $f(B_t)$, $0 \leq t < \tau_U$, is a time-change of a complex Brownian motion started from $f(z_0)$ up to the time $\tau_{f(U)}$.*

Proof. Write $f = a + ib$ and $X = f(B) = a(B^1, B^2) + ib(B^1, B^2)$. By Itô's formula, $a(B)$ and $b(B)$ satisfy the SDEs: (We write a_x for $\partial_x a(B)$)

$$da(B_t) = (a_x dB_t^1 + a_y dB_t^2) + \frac{1}{2}a_{xx}d[B^1]_t + \frac{1}{2}a_{yy}d[B^2]_t + a_{xy}d[B^1, B^2]_t = a_x dB_t^1 + a_y dB_t^2;$$

$$db(B_t) = (b_x dB_t^1 + b_y dB_t^2) + \frac{1}{2}b_{xx}d[B^1]_t + \frac{1}{2}b_{yy}d[B^2]_t + b_{xy}d[B^1, B^2]_t = b_x dB_t^1 + b_y dB_t^2.$$

Here we used the fact that $[B^1] = [B^2]$, $[B^1, B^2] = 0$, and $a_{xx} + a_{yy} = b_{xx} + b_{yy} = 0$. So we have

$$d[a(B_t)] = a_x(B)^2 dt + a_y(B)^2 dt = |f'(B)|^2 dt, \quad d[b(B_t)] = b_x(B)^2 dt + b_y(B)^2 dt = |f'(B)|^2 dt,$$

$$d[a(B_t), b(B_t)] = (a_x(B)b_x(B) + a_y(B)b_y(B))dt = 0.$$

Here we use that $[B^1]_t = [B^2]_t = t$, $[B^1, B^2] = 0$, $a_x^2 + a_y^2 = b_x^2 + b_y^2 = |f'|^2$, and $a_x b_x + a_y b_y = 0$. If we define $u_t = \int_0^t |f'(B_s)|^2 ds$, $0 \leq t < \tau_U$, and let $v = u^{-1}$, then the time-change of $X = f(B) = a(B) + ib(B)$ via v , i.e., $f(B_{v_s}) = a(B_{v_s}) + ib(B_{v_s})$ satisfies $[a(B_{v_s})]_t = [b(B_{v_s})]_t = dt$, and $[a(B_{v_s}), b(B_{v_s})] = 0$. So $f(B_{v_s})$, $0 \leq s < \tau_U$, is a stopped complex Brownian motion. Since $B_t \rightarrow \partial D$ as $t \rightarrow \tau_U$, we have $f(B_{v_s}) \rightarrow \partial f(D)$ as $s \rightarrow S = u(\tau_U)$. So $v(\tau_U)$ corresponds to the time that the complex Brownian motion in $f(D)$ exits $f(D)$. Then we get the conclusion. \square

Corollary . For an open set $U \subset \mathbb{C}$ and $z \in U$, we use $S_{D,z}$ to denote the image of a complex Brownian motion B^z started from z up to the time that it exits U , i.e., $S_{D,z} = \{B_t^z : 0 \leq t < \tau_U\}$. Such $S_{D,z}$ is a random set, whose law satisfies conformal invariance in the following sense. If f maps U conformally onto \tilde{U} , and $f(z) = \tilde{z}$, then $f(S_{D,z})$ has the same law as $S_{\tilde{D},\tilde{z}}$.

Example . Brownian bridge is a semimartingale. This is not obvious from the original definition: $X_t = B_t - tB_1$, $0 \leq t \leq 1$, where B is a Brownian motion. However, we know that X_t , $0 \leq t < 1$, has the same distribution as the process $X_t := (1-t)B_{t/(1-t)}$, $0 \leq t < 1$, for a Brownian motion B . We know that $B_{t/(1-t)}$ is a local martingale since it is a time-change of Brownian motion. Then its product with $1-t$ is a semimartingale. We write $Y_t = B_{t/(1-t)}$. Then $[Y]_t = t/(1-t)$, $0 \leq t < 1$, and by Itô's formula,

$$dX_t = (1-t)dY_t - Y_t dt = d\tilde{B}_t - \frac{X_t}{1-t} dt,$$

where $\tilde{B} := (1-t) \cdot Y$ is a local martingale with $[\tilde{B}]_t = \int_0^t (1-s)^2 [Y]_s' ds = t$, $0 \leq t < 1$. So \tilde{B} is a stopped Brownian motion

Example (Bessel Processes). Let $B = (B^1, \dots, B^\delta)$ be a Brownian motion in \mathbb{R}^δ started from some $x_0 \neq 0$. Let τ be the first time that B reaches 0. Then $|B| = \sqrt{(B^1)^2 + \dots + (B^\delta)^2}$ restricted to $[0, \tau)$ is a semimartingale.

We now do the calculation. Write $f(x_1, \dots, x_\delta) = \sqrt{x_1^2 + \dots + x_\delta^2}$. Then

$$\partial_j f = \frac{x_j}{(x_1^2 + \dots + x_\delta^2)^{1/2}}, \quad \partial_j^2 f = \frac{1}{(x_1^2 + \dots + x_\delta^2)^{1/2}} - \frac{x_j^2}{(x_1^2 + \dots + x_\delta^2)^{3/2}}, \quad \Delta f = \frac{\delta - 1}{|x|}.$$

So

$$d|B| = \sum_{j=1}^{\delta} \partial_j f(B) dB^j + \frac{1}{2} \Delta f(B) dt = \sum_{j=1}^{\delta} \frac{B^j dB^j}{|B|} + \frac{1}{2} \frac{\delta - 1}{|B|} dt.$$

Since $[\sum_{j=1}^{\delta} \frac{B^j \cdot B^j}{|B|}]_t = \sum_{j=1}^{\delta} (\frac{B^j}{|B|})^2 [B^j]_t = t$, by Lévy's characterization of Brownian motion, $\sum_{j=1}^{\delta} \frac{B^j \cdot B^j}{|B|}$ is a one-dimensional Brownian motion. Denote it by \tilde{B} . Then $X = |B|$ satisfies the stochastic differential equation:

$$dX_t = d\tilde{B}_t + \frac{(\delta - 1)/2}{X_t} dt, \quad X_0 = |x_0|.$$

The lifetime of X is $\tau = \inf\{t \geq 0 : X_t = 0\}$.

Definition . Let B be a (one-dimensional) Brownian motion, $x_0 > 0$, and $\delta \in \mathbb{R}$. The semimartingale X that satisfies the stochastic differential equation (SDE)

$$dX_t = dB_t + \frac{(\delta - 1)/2}{X_t} dt, \quad X_0 = x_0, \quad (3.2)$$

is called a Bessel process of dimension δ started from x_0 .

We first prove the existence of Bessel processes. Let $Y_t = X_t - B_t$. Then Y_t satisfies the ODE

$$Y_t' = \frac{(\delta - 1)/2}{Y_t + B_t}, \quad Y_0 = x_0.$$

For a given Brownian motion B , by the ODE theory, the solution Y_t , $0 \leq t < T$, exists and is unique, where T is some blowup time, which could be finite or infinite. Let \mathcal{F} be the complete and right-continuous filtration generated by B . By the ODE theory again, we also know that Y depends continuously on B . So Y is \mathcal{F} -adapted. Let $X = Y + B$. Then X is an \mathcal{F} -adapted process with lifetime T , and $X_t - B_t = X_0 + \int_0^t \frac{(\delta-1)/2}{X_s} ds$, $0 \leq t < T$. So X is a semimartingale satisfying (3.2), and so is a Bessel process of dimension δ .

We now derive some deterministic properties of the Bessel process. Fix a Brownian motion B . By discarding a null event, we may assume that $\limsup_{t \rightarrow \infty} B_t = \infty$ and $\liminf_{t \rightarrow \infty} B_t = -\infty$. Since $Y_0 = x_0 > 0$, and $Y_t + B_t \neq 0$ on $[0, T)$, we have $X_t = Y_t + B_t > 0$ on $[0, T)$. By the formula for Y_t' , Y is monotone increasing or decreasing depending on the sign of $\delta - 1$.

First, we claim that, if $T < \infty$, it can not happen that $Y_t \rightarrow \infty$ or $Y_t \rightarrow -\infty$ as $t \uparrow T$. If that is the case, since $\lim_{t \uparrow T} B(t)$ converges to $B(T)$, we see that for some $\varepsilon > 0$, $1/(Y_t + B_t)$ is bounded on $(T - \varepsilon, T)$, which implies that Y is Lipschitz continuous on $(T - \varepsilon, T)$, and so Y is bounded on $(T - \varepsilon, T)$. This contradicts that $Y_t \rightarrow \infty$ or $Y_t \rightarrow -\infty$. Thus, if $T < \infty$, then the only way that the solution blows up at the time T is that $Y_t + B_t \rightarrow 0$ as $t \uparrow T$. So we find that, if $T < \infty$, then $\lim_{t \uparrow T} X_t = 0$.

Second, we claim that, if $T = \infty$, then $\delta \geq 1$, and X is unbounded on $[0, \infty)$. If $\delta \leq 1$, Y is decreasing. So $0 < X_t = Y_t + B_t \leq Y_0 + B_t = x_0 + B_t$, which implies that $B_t > -x_0$ on $[0, T)$.

Since $\liminf_{t \rightarrow \infty} B_t = -\infty$, we can not have $T = \infty$. So $\delta \geq 1$, and Y is increasing. Then we get $X_t \geq Y_0 + B_t = x_0 + B_t$. From $\limsup_{t \rightarrow \infty} B_t = \infty$ we then see that X is unbounded.

From the strong Markov property of Brownian motion and the property of solutions of ODE, we see that a Bessel process also satisfies strong Markov property: if X is a Bessel process of dimension δ , and τ is a stopping time less than the lifetime T of X , then conditionally on \mathcal{F}_τ , the law of $X_{\tau+}$ is a Bessel process started from X_τ . To see this, note that $X_{\tau+}$ satisfies

$$X_{\tau+t} = X_\tau + (B_{\tau+t} - B_\tau) + \int_0^t \frac{(\delta - 1)/2}{X_{\tau+s}} ds, \quad 0 \leq t < T - \tau,$$

and recall that the conditional law of $\tilde{B}_{\tau+} - \tilde{B}_\tau$ given \mathcal{F}_τ is that of a Brownian motion.

Theorem (Transience and Recurrence of Bessel Processes). *Let X be a Bessel process of dimension δ started from $x_0 > 0$. Let T be its lifetime. Then*

- (i) *If $\delta \geq 2$, then a.s. $T = \infty$.*
- (ii) *If $\delta < 2$, then a.s. $T < \infty$ and $\lim_{t \rightarrow T} X_t = 0$.*
- (iii) *If $\delta > 2$, then a.s. $\lim_{t \rightarrow \infty} X_t = \infty$, and for any $r \in (0, x_0)$, $\mathbb{P}[X_t > r : 0 \leq t < T] > 0$.*
- (iv) *If $\delta = 2$, then a.s. $\limsup_{t \rightarrow \infty} X_t = \infty$ and $\liminf_{t \rightarrow \infty} X_t = 0$.*

Remark . The strategy of the proof will be used many times later. First, we construct a local martingale using Itô's formula. Second, we obtain a true martingale from the local martingale using stopping and boundedness. Third, we use the martingale to derive some equality or inequality about the probability of some event.

Proof. Using Itô's formula, we may construct a function f defined on $(0, \infty)$ such that $f(X_t)$ is a local martingale. Since

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)dt = f'(X_t)dB_t + (f'(X_t)\frac{(\delta - 1)/2}{X_t} + \frac{1}{2}f''(X_t))dt,$$

we need f to make the drift term vanish, i.e., $f'(x)\frac{(\delta-1)/2}{x} + \frac{1}{2}f''(x) = 0$. Solving this ODE, we may choose $f'(x) = x^{1-\delta}$. For $\delta = 2$, we let $f(x) = \log(x)$; for $\delta \neq 2$, let $f(x) = x^{2-\delta}$.

Suppose X starts from $x_0 \in (0, \infty)$. Let $\tau_y = \inf\{t \geq 0 : X_t = y\}$, $y > 0$. Fix $R > x_0 > r > 0$. Since we know that X either tends to 0 (when $T < \infty$) or is unbounded (when $T = \infty$), it must exit the interval (r, R) . So $\tau_r \wedge \tau_R < T$

Since $Y_t := f(X_t)$ is a local martingale, and is bounded before $\tau_r \wedge \tau_R$, $Y_t^{\tau_r \wedge \tau_R}$ is a bounded martingale. So we have $\mathbb{E}[f(X_{\tau_r \wedge \tau_R})] = \mathbb{E}[f(X_0)] = f(x_0)$. We have either $\tau_r < \tau_R$ or $\tau_R < \tau_r$. In the former case $X_{\tau_r \wedge \tau_R} = r$, and in the latter case $X_{\tau_r \wedge \tau_R} = R$. Thus,

$$f(r)\mathbb{P}[\tau_r < \tau_R] + f(R)\mathbb{P}[\tau_R < \tau_r] = f(x_0).$$

Combining this formula with $\mathbb{P}[\tau_r < \tau_R] + \mathbb{P}[\tau_R < \tau_r] = 1$, we get

$$\mathbb{P}[\tau_r < \tau_R] = \frac{f(R) - f(x_0)}{f(R) - f(r)}, \quad \mathbb{P}[\tau_R < \tau_r] = \frac{f(x_0) - f(r)}{f(R) - f(r)}. \quad (3.3)$$

If $\delta \geq 2$, $|f(r)| = r^{2-\delta}$ or $|\log(r)|$, which tends to ∞ as $r \rightarrow 0^+$. So by (3.3),

$$\mathbb{P}\left[\bigcap_{r \in (0, x)} \{\tau_r < \tau_R\}\right] = \lim_{r \downarrow 0} \mathbb{P}[\tau_r < \tau_R] = \lim_{r \downarrow 0} \frac{f(R) - f(x_0)}{f(R) - f(r)} = 0.$$

If $T < \infty$, then $\lim_{t \uparrow T} X_t = 0$, and $X([0, T])$ is bounded. So there is $R \in \mathbb{N}$ with $R > x_0$ such that $\tau_r < \tau_R$ for any $r \in (0, x_0)$. This means that

$$\{T < \infty\} \subset \bigcup_{R \in \mathbb{N}: R > x_0} \bigcap_{r \in (0, x_0)} \{\tau_r < \tau_R\}.$$

Thus, $\mathbb{P}[T < \infty] \leq \mathbb{P}[\bigcup_{R \in \mathbb{N}: R > x_0} \bigcap_{r \in (0, x_0)} \{\tau_r < \tau_R\}] = 0$. So we get (i).

If $\delta < 2$, then $f(r) = r^{2-\delta} \rightarrow 0$ as $r \rightarrow 0^+$ and $f(R) \rightarrow \infty$ as $R \rightarrow \infty$. So

$$\mathbb{P}\left[\bigcap_{r \in (0, x_0)} \{\tau_r < \tau_R\}\right] = \lim_{r \downarrow 0} \mathbb{P}[\tau_r < \tau_R] = \frac{f(R) - f(x_0)}{f(R)},$$

and

$$\mathbb{P}\left[\bigcup_{R > x_0} \bigcap_{r \in (0, x_0)} \{\tau_r < \tau_R\}\right] = \lim_{R \rightarrow \infty} \frac{f(R) - f(x_0)}{f(R)} = 1.$$

This means that a.s. there is $R > x_0$ such that for any $r \in (0, x_0)$, $\tau_r < \tau_R$. This means that a.s. X is bounded, which then implies that $T < \infty$ and $\lim_{t \rightarrow \infty} X_t = 0$. Then we get (ii).

(iii) If $\delta > 2$, then we already know that a.s. $T = \infty$ and X is unbounded on $[0, \infty)$, i.e., $\limsup_{t \rightarrow \infty} X_t = \infty$. Since $Y = X^{2-\delta}$ is a positive local martingale, we have a.s. $\lim_{t \rightarrow \infty} Y_t$ converges. From $2 - \delta < 0$ and that a.s. $\limsup_{t \rightarrow \infty} X_t = \infty$ we get a.s. $\liminf_{t \rightarrow \infty} Y_t = 0$. So a.s. $\lim_{t \rightarrow \infty} Y_t = 0$, which implies that $\lim_{t \rightarrow \infty} X_t = \infty$. For $r \in (0, x_0)$, we have

$$\mathbb{P}[X_t > r : 0 \leq t < \infty] = \mathbb{P}[\tau_r = \infty] = \lim_{n \rightarrow \infty} \mathbb{P}[\tau_n < \tau_r] = \lim_{n \rightarrow \infty} \frac{x_0^{2-\delta} - r^{2-\delta}}{n^{2-\delta} - r^{2-\delta}} = 1 - \left(\frac{x_0}{r}\right)^{2-\delta} > 0.$$

(iv) Let $\delta = 2$. We know that $Y = \log(X)$ is a local martingale. If it does not hold that $\limsup_{t \rightarrow \infty} X_t = \infty$ and $\liminf_{t \rightarrow \infty} X_t = 0$, then it does not hold that $\limsup_{t \rightarrow \infty} Y_t = \infty$ and $\liminf_{t \rightarrow \infty} Y_t = -\infty$. Since Y is a time-change of a Brownian motion up to some stopping time, we then have a.s. $\lim_{t \rightarrow \infty} Y_t$ converges. So $\lim_{t \rightarrow \infty} X_t$ converges, which contradicts that X is unbounded because $T = \infty$. \square

Theorem (Transience and Recurrence of Multidimensional Brownian Motion). *Let $B = (B^1, \dots, B^d)$ be a Brownian motion in \mathbb{R}^d , with $d \geq 2$, started from some x_0 . Let $x_1 \in \mathbb{R}^d \setminus \{x_0\}$ and $r \in (0, |x_1 - x_0|)$, let $\tau_{x_1} = \inf\{t \geq 0 : B_t = x_1\}$ and $\tau_{x_1, r} = \inf\{t \geq 0 : |B_t - x_1| \leq r\}$. Then*

(i) Almost surely $\tau_{x_1} = \infty$.

(ii) If $d \geq 3$, a.s. $|B_t| \rightarrow \infty$; and for any $r \in (0, |x_1 - x_0|)$, $\mathbb{P}[\tau_{x_1, r} = \infty] > 0$.

(iii) If $d = 2$, a.s. $\liminf_{t \rightarrow \infty} |B_t - x_1| = 0$.

Proof. This follows from the previous theorem and the fact that $|B - x_1|$ is a Bessel process of dimension d started from $|x_1 - x_0|$. Note that $\{\tau_{x_1} < \infty\}$ is the event that $|B - x_1|$ reaches 0, $\{\tau_{x_1, r} < \infty\}$ is the event that $|B - x_1|$ reaches r , and $|B_t| \rightarrow \infty$ iff $|B_t - x_1| \rightarrow \infty$. \square

Because of this theorem, we say that Brownian motions of dimensions ≥ 3 are transient, and the Brownian motion of dimension 2 is neighborhood recurrent (but not point recurrent). Note that (i) states that, when $d \geq 2$, for any fixed $x_1 \in \mathbb{R}^d \setminus \{x_0\}$, a.s. B does not visit x . It does not imply that B a.s. does not visit any point other than x_0 . By Fubini Theorem, we see that for $d \geq 2$, a.s. the range of a Brownian motion of dimension d has Lebesgue measure zero.

Example . We may now construct a positive local martingale X , which is not a true martingale, and for any $t \geq 0$, X_t is integrable. Let B be a Brownian motion in \mathbb{R}^3 started from 0. Let $v = (1, 0, 0)$. We know that B a.s. does not pass through v . Since $|B - v|$ is a Bessel process of dimension 3, $X := |B - v|^{-1}$ is a local martingale. We may show that $\mathbb{E}[X_t] < 1 = X_0$ by calculation using the density of B_t . Note that the density of B_t is spherically symmetric. We may use the following two facts to do the computation. First, the average of $1/|x - v|$ over the surface $|x| = R$ equals $1/|v| = 1$ if $R < |v| = 1$. This follows from the fact that $f(x) := 1/|x - v|$ is harmonic on $\mathbb{R}^3 \setminus \{v\}$. Second, if $R > |v| = 1$, the average of $1/|x - v|$ over the surface $|x| = R$ equals $1/R$, which is strictly less than 1. This follows from the fact that, when $|x| = R$, $1/|x - v| = R/|x - \hat{v}|$, where $\hat{v} = (R^2, 0, 0)$.

Exercise Provide details of the above example.

Exercise Let X_t , $0 \leq t < \infty$, be a nonnegative local martingale with $\mathbb{E}[X_0] < \infty$. Prove the following. (i) X is a supermartingale. Hint: Use Fatou's lemma. (ii) If $\mathbb{E}[X_\infty] = \mathbb{E}[X_0]$, then X is a uniformly integrable martingale. (iii) If for some stopping time T , $\mathbb{E}[X_T] = \mathbb{E}[X_0]$, then X^T is a uniformly integrable martingale.

Example (Brownian motion with a linear drift). Let B be a one-dimensional Brownian motion. Let $a > 0$ and $X_t = B_t - at$. By the law of iterated logarithm, we have $\lim_{t \rightarrow \infty} X_t/t = -a < 0$. So $\lim_{t \rightarrow \infty} X_t = -\infty$. Since $X_0 = 0$, $\zeta := \sup_{t \geq 0} X_t$ is a positive finite random number. We wish to find the distribution of ζ . For $b > 0$, let τ_b denote the first time that $X_t \geq b$. Then $\zeta \geq b$ iff $\tau_b < \infty$. We now compute $\mathbb{P}[\tau_b < \infty]$. The X satisfies the SDE:

$$dX_t = dB_t - a dt.$$

We may find f such that $f(X)$ is a local martingale. Since

$$df(X_t) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)dt = f'(X_t)dB_t + \left(\frac{1}{2}f''(X_t) - af'(X_t)\right)dt,$$

we need f that satisfies $\frac{1}{2}f''(x) - af'(x) = 0$. Then $f(x) = e^{2ax}$ is a solution. Thus, $Y := e^{2aX}$ is a local martingale. For $t \leq \tau_b$, we have $X_t \leq b$, which implies that $0 < Y_t \leq e^{2ab}$. Thus, Y^{τ_b} is uniformly integrable, and we have $1 = Y_0 = \mathbb{E}[Y_{\tau_b}]$. Note that on the event $\tau_b < \infty$, $Y_{\tau_b} = e^{2ab}$, and on the event $\tau_b = \infty$, from $\lim_{t \rightarrow \infty} X_t = -\infty$, we get $Y_\infty = 0$. Thus, $1 = e^{2ab}\mathbb{P}[\tau_b < \infty] + 0 \cdot \mathbb{P}[\tau_b = \infty]$, which implies that $\mathbb{P}[\tau_b < \infty] = e^{-2ab}$. Then we get $\mathbb{P}[\zeta \geq b] = e^{-2ab}$. So ζ has an exponential distribution with rate $2a$.

Actually, we proved the following probability: for any $a, b > 0$,

$$\mathbb{P}[B_t < at + b, \forall t \geq 0] = 1 - e^{-2ab}.$$

4 Girsanov's Theorem and Applications

At the beginning of the lecture, we fixed a space Ω , a right-continuous and complete filtration \mathcal{F} , and a probability measure \mathbb{P} . We also assumed that \mathcal{F} is \mathbb{P} -complete. When we did a time-change, we changed the filtration. We now fix Ω and \mathcal{F} , but change the probability measure.

Consider two different probability measures P and Q on (Ω, \mathcal{F}) . Suppose for each $t \geq 0$, $Q \ll P$ on \mathcal{F}_t , i.e., $A \in \mathcal{F}_t$ and $P[A] = 0$ implies $Q[A] = 0$. We assume that \mathcal{F} is P -complete. Then it is also Q -complete. By Radon-Nikodym Theorem, there is a positive \mathcal{F}_t -measurable random variable Z_t such that $Q = Z_t \cdot P$, i.e., $Q[A] = \int_A Z_t dP$ for any $A \in \mathcal{F}_t$. Such Z_t is P -a.s. unique, and we call it the RN process from P to Q .

The next lemma describes the relation between P -martingales and Q -martingales.

Lemma 16.15. *Suppose $Q = Z_t \cdot P$ on \mathcal{F}_t for each $t \geq 0$. Then Z is a P -martingale. It is further P -uniformly integrable iff $Q \ll P$ on $\mathcal{F}_\infty := \bigvee_{t \geq 0} \mathcal{F}_t$. Any adapted process X is a Q -martingale iff XZ is a P -martingale.*

Proof. For any adapted process X , X is a Q -martingale iff for any $t > s \geq 0$ and $A \in \mathcal{F}_s$, $\int_A X_t dQ = \int_A X_s dQ$. Since $A \in \mathcal{F}_s \subset \mathcal{F}_t$, this equality becomes

$$\int_A X_t Z_t dP = \int_A X_s Z_s dP,$$

which is equivalent to that XZ is a P -martingale. Taking $X \equiv 1$, we see that Z is a P -martingale.

If $Q \ll P$ on \mathcal{F}_∞ , let $Z_\infty \in \mathcal{F}_\infty$ be the RN derivative. For any $t \geq 0$ and $A \in \mathcal{F}_t \subset \mathcal{F}_\infty$, from $Q = Z_t \cdot P$ on \mathcal{F}_t and $Q = Z_\infty \cdot P$ on \mathcal{F}_∞ , we get $Q(A) = \int_A Z_t dP = \int_A Z_\infty dP$. So we have $Z_t = \mathbb{E}_P[Z_\infty | \mathcal{F}_t]$, $t \geq 0$, which implies that Z is uniformly integrable.

Now suppose Z is uniformly integrable. Then $Z_t \rightarrow Z_\infty$ as $t \rightarrow \infty$ in L^1 , and $Z_t = \mathbb{E}_P[Z_\infty | \mathcal{F}_t]$ for all $t \geq 0$. Especially, $\mathbb{E}_P[Z_\infty] = \mathbb{E}_P[Z_t] = 1$. So we may define a probability measure \tilde{Q} on \mathcal{F}_∞ by $\tilde{Q} = Z_\infty \cdot P$. For any $t \geq 0$ and $A \in \mathcal{F}_t$, since $\mathbb{E}_P[Z_\infty] = Z_t$, $\tilde{Q}[A] = \int_A Z_\infty dP = \int_A Z_t dP = Q[A]$. So \tilde{Q} agrees with Q on \mathcal{F}_t for each $t \geq 0$. By a monotone class argument, we conclude that $\tilde{Q} = Q$ on \mathcal{F}_∞ . So $Q \ll P$ on \mathcal{F}_∞ . \square

By Theorem 6.27, every martingale on \mathcal{F} has an rcll version. We now assume that Z is rcll.

Lemma 16.16. *Suppose $Q = Z_t \cdot P$ on \mathcal{F}_t for each $t \geq 0$, and Z is rcll. Then for any stopping time τ , $Q = Z_\tau \cdot P$ on $\mathcal{F}_\tau \cap \{\tau < \infty\}$, i.e., if $A \in \mathcal{F}_\tau$ and $A \subset \{\tau < \infty\}$, then $Q[A] = \int_A Z_\tau dP$. Furthermore, an rcll process X is a Q -local martingale iff XZ is a P -local martingale.*

Proof. By Optional Stopping Theorem (6.29), $\mathbb{E}_P[Z_t | \mathcal{F}_{\tau \wedge t}] = Z_{\tau \wedge t}$ for any $t \geq 0$. Fix $t \geq 0$ and $A \in \mathcal{F}_{\tau \wedge t} \subset \mathcal{F}_t$. We have

$$Q[A] = \int_A Z_t dP = \int_A Z_{\tau \wedge t} dP.$$

Let $A \in \mathcal{F}_\tau$. For $t \geq 0$, let $A_t = A \cap \{\tau \leq t\} \in \mathcal{F}_{\tau \wedge t}$. Here we used that $\mathcal{F}_\tau \cap \{\tau \leq t\} = \mathcal{F}_\tau \cap \{\tau \leq \tau \wedge t\} \subset \mathcal{F}_{\tau \wedge t}$. So we have

$$Q[A_t] = \int_{A_t} Z_{\tau \wedge t} dP = \int_{A_t} Z_\tau dP,$$

where in the last equality we used that $\tau = \tau \wedge t$ on $A_t \subset \{\tau \leq t\}$. If $A \subset \{\tau < \infty\}$, then $A = \bigcup_{t \geq 0} A_t$. By Monotone Convergence Theorem, we get $Q[A] = \int_A Z_\tau dP$.

To prove the last assertion, it suffices to show that for any stopping time τ , X^τ is a (Q, \mathcal{F}) -martingale iff $(XZ)^\tau$ is a (P, \mathcal{F}) -martingale. Here we use (Q, \mathcal{F}) and (P, \mathcal{F}) to emphasize the dependence on \mathcal{F} . By the last paragraph, for any $t \geq 0$, $Q = Z_{\tau \wedge t} \cdot P$ on $\mathcal{F}_t^\tau := \mathcal{F}_{\tau \wedge t}$. By Lemma 16.15, X^τ is a (Q, \mathcal{F}^τ) -martingale iff $X^\tau Z^\tau$ is a (P, \mathcal{F}^τ) -martingale. It remains to show that X^τ is a (Q, \mathcal{F}) -martingale iff X^τ is a (Q, \mathcal{F}) -martingale, and $X^\tau Z^\tau$ is a (P, \mathcal{F}) -martingale iff $X^\tau Z^\tau$ is a (P, \mathcal{F}^τ) -martingale.

The two statements are similar. It suffices to prove the first one. Note that since X is rcll and \mathcal{F} -adapted, X^τ is both \mathcal{F}^τ and \mathcal{F} -adapted. Now suppose X^τ is an (P, \mathcal{F}) -martingale. Let $t \geq s \geq 0$ and $A \in \mathcal{F}_{\tau \wedge s}$. Then $A \in \mathcal{F}_s$, and so $\int_A X_t^\tau dP = \int_A X_s^\tau dP$. Thus, X^τ is also an (P, \mathcal{F}^τ) -martingale. Next, suppose X^τ is an (P, \mathcal{F}^τ) -martingale. Let $t \geq s \geq 0$ and $A \in \mathcal{F}_s$. Then $A \cap \{s \leq \tau\} = A \cap \{s \leq \tau \wedge s\} \in \mathcal{F}_{\tau \wedge s}$. So

$$\int_{A \cap \{s \leq \tau\}} X_t^\tau dP = \int_{A \cap \{s \leq \tau\}} X_s^\tau dP.$$

Since $X_t^\tau = X_s^\tau = X_\tau$ on the event $\{\tau < s\}$, we have

$$\int_{A \cap \{\tau < s\}} X_t^\tau dP = \int_{A \cap \{\tau < s\}} X_s^\tau dP.$$

Combining the last two displayed formulas, we get $\int_A X_t^\tau dP = \int_A X_s^\tau dP$. Thus, X^τ is also an (P, \mathcal{F}) -martingale. \square

Lemma 16.17. *For every $t \geq 0$, Q -a.s. $\inf_{s \in [0, t]} Z_s > 0$.*

Proof. For any $t \geq 0$, $Q[\{Z_t = 0\}] = \int_{\{Z_t = 0\}} Z_t dP = 0$. So Q -a.s. $Z_t > 0$. Let $\tau_n = \inf\{t \geq 0 : Z_t < 1/n\}$, $n \in \mathbb{N}$. Then each τ_n is a stopping time. By right-continuity of Z , $Z_{\tau_n} \leq 1/n$ on the

event $\tau_n < \infty$. Moreover, we have $\inf_{s \in [0, t]} Z_s = 0$ iff $\tau_n \leq t$ for each n . By Optional Stopping Theorem,

$$Q[\tau_n \leq t] = \mathbb{E}_Q[\mathbf{1}_{\{\tau_n \leq t\}}] = \mathbb{E}_P[\mathbf{1}_{\{\tau_n \leq t\}} Z_t] = \mathbb{E}_P[\mathbf{1}_{\{\tau_n \leq t\}} Z_{\tau_n}] \leq 1/n.$$

So we get $Q[\{\inf_{s \in [0, t]} Z_s = 0\}] = \lim_{n \rightarrow \infty} Q[\tau_n \leq t] = 0$. \square

Theorem 16.19 (Girsanov Theorem). *Suppose $Q = Z_t \cdot P$ on \mathcal{F}_t for each $t \geq 0$, and Z is continuous. Then for any continuous local P -martingale M , the process $\widetilde{M} = M - Z^{-1} \cdot [M, Z]^P$ is a continuous Q -local martingale, where we use the superscript P to indicate the dependence on P .*

Proof. First assume that $Z > 0$ and Z^{-1} is bounded. Then \widetilde{M} is a continuous P -semimartingale, and by Proposition 15.15 (Chain Rule) and Theorem 15.17 (Product Formula),

$$\begin{aligned} \widetilde{M}Z - \widetilde{M}_0Z_0 &= \widetilde{M} \cdot Z + Z \cdot \widetilde{M} + [\widetilde{M}, Z]^P \\ &= \widetilde{M} \cdot Z + Z \cdot M - [M, Z]^P + [\widetilde{M}, Z]^P = \widetilde{M} \cdot Z + Z \cdot M. \end{aligned}$$

So $\widetilde{M}Z$ is a continuous P -local martingale. By Lemma 16.16, \widetilde{M} is a continuous local Q -martingale.

For general case, define $\tau_n = \inf\{t \geq 0 : Z_t \leq 1/n\}$. Then $Z^{\tau_n} > 0$ and $(Z^{\tau_n})^{-1}$ is bounded. Since $\widetilde{M}^{\tau_n} = M^{\tau_n} - (Z^{\tau_n})^{-1} \cdot [M^{\tau_n}, Z^{\tau_n}]^P$, by the last paragraph, \widetilde{M}^{τ_n} is a Q -local martingale. By Lemma 16.17, Q -a.s. $\tau_n \uparrow \infty$. So by Lemma 15.1, \widetilde{M} is a Q -local martingale. \square

From now on, we assume that Z is continuous. We add superscript P or Q to $[X]$, $L(X)$ and $V \cdot X$ to indicate the dependence on P or Q .

Proposition 16.20. *Any P -semimartingale X is also a Q -semimartingale, and for any semimartingales X and Y , Q -a.s. $[X]^P = [X]^Q$ and $[X, Y]^P = [X, Y]^Q$. Furthermore, for any $V \in L^P(X)$, Q -a.s. $V \in L^Q(X)$, and $(V \cdot X)^P = (V \cdot X)^Q$. Finally, if M is a P -local martingale, then Q -a.s. $(\widetilde{V \cdot M}) = V \cdot \widetilde{M}$, where $\widetilde{M} = M - Z^{-1} \cdot [M, Z]^P$ is as in Theorem 16.19.*

Proof. Let $M + A$ be the P -decomposition of X . Since $M = \widetilde{M} + Z^{-1} \cdot [M, Z]^P$, \widetilde{M} is a Q -local martingale, and $Z^{-1} \cdot [M, Z]^P$ is a finite variation process, we see that $X = \widetilde{M} + Z^{-1} \cdot [M, Z]^P + A$ is a Q -semimartingale. From Proposition 15.18, we see that Q -a.s., $[X]^P = [X]^Q$. In fact, for any fixed $t_0 > 0$ and a sequence $\Delta_n = \{0 = t_0^n < \dots < t_{k_n}^n = t_0\}$ of partitions of $[0, t_0]$ with mesh size tending to 0, we have $T_X^{\Delta_n} := \sum_{j=1}^{k_n} |X_{t_j^n} - X_{t_{j-1}^n}|^2$ tends to $[X]_{t_0}^P$ in probability w.r.t. P . Since $Q \ll P$ on \mathcal{F}_{t_0} , the convergence also holds w.r.t. Q . By polarization, we then have Q -a.s. $[X, Y]^P = [X, Y]^Q$. From now on, we drop the superscript P or Q after $[\cdot]$.

Let $V \in L^P(X)$. Then $V \in L^P(A)$ and $V^2 \in L^P([M]) = L^P([\widetilde{M}])$. For a finite variation process A , $L(A)$ does not depend on the underlying probability. So $L^P(A) = L^Q(A)$ and $L^P([\widetilde{M}]) = L^Q([\widetilde{M}])$. Thus, $V \in L^Q(A) \cap L^Q(\widetilde{M})$. Since $X = \widetilde{M} + Z^{-1} \cdot [M, Z]^P + A$, to prove

that $V \in L^Q(X)$, it remains to show that Q -a.s. $V \in L(Z^{-1} \cdot [M, Z])$. This is true because by Proposition 15.10, Q -a.s.

$$\int_0^t |V_s Z_s^{-1}| |d[M, Z]_s| \leq \left(\int_0^t V_s^2 d[M]_s \right) \cdot \left(\int_0^t Z_s^{-2} d[Z]_s \right) < \infty, \quad \forall t \geq 0.$$

Here we used the fact that $V^2 \in L([\widetilde{M}]) = L([M])$ and Q -a.s. Z^{-1} is continuous.

We now prove that Q -a.s., $(V \cdot X)^P = (V \cdot X)^Q$. The equality is trivial if X is a finite variation process. It remains to prove the case that X is the local martingale M . To prove $(V \cdot M)^Q = (V \cdot M)^P$, it suffices to show that $(V \cdot \widetilde{M})^Q = (V \cdot M)^P - V \cdot (Z^{-1} \cdot [M, Z])$. Recall that (under P or Q) if M is a local martingale, then $V \cdot M$ is the local martingale such that for any local martingale N , $[V \cdot M, N] = V \cdot [M, N]$. The statement is also true if N is any semimartingale since the finite variation part of N does not contribute to the quadratic covariation. Now $\widetilde{M} = M - Z^{-1} \cdot [M, Z]$ is a Q -local martingale. To prove that $(V \cdot \widetilde{M})^Q = (V \cdot M)^P - V \cdot (Z^{-1} \cdot [M, Z])$, we need to show

- (i) $(V \cdot M)^P - V \cdot (Z^{-1} \cdot [M, Z])$ is a Q -local martingale;
- (ii) for any Q -semimartingale N , $[(V \cdot M)^P - V \cdot (Z^{-1} \cdot [M, Z]), N] = V \cdot [\widetilde{M}, N]$.

For (i), by Lemma 16.16 we need to show that $Z((V \cdot M)^P - V \cdot (Z^{-1} \cdot [M, Z]))$ is a P -local martingale. This follows from a straightforward Itô's calculation: the finite variation part comes from the first term is $[Z, V \cdot M] = V \cdot [Z, M]$, while the finite variation part comes from the second term is $-Z \cdot (V \cdot (Z^{-1} \cdot [M, Z])) = -V \cdot [M, Z]$. So (i) is proved. Part (ii) also follows from a straightforward calculation:

$$[(V \cdot M)^P - V \cdot (Z^{-1} \cdot [M, Z]), N] = [(V \cdot M)^P, N] = V \cdot [M, N] = V \cdot [\widetilde{M}, N].$$

From now on, we drop superscripts P and Q after stochastic integral.

Suppose M is a P -local martingale. Then \widetilde{M} is a Q -local martingale, and so is $V \cdot \widetilde{M}$. We know that $V \cdot M - V \cdot \widetilde{M} = V Z^{-1} \cdot [M, Z]$ is a finite variation process. We also know that $V \cdot M - \widetilde{V \cdot M}$ is a finite variation process. So the two Q -local martingales $V \cdot \widetilde{M}$ and $\widetilde{V \cdot M}$ differ by a finite variation process. Thus, Q -a.s. $(V \cdot \widetilde{M}) = \widetilde{V \cdot M}$. \square

Remark . Thanks to Proposition 16.20, we do not need to distinguish $[\cdot]^P$ from $[\cdot]^Q$, and $(V \cdot X)^P$ from $(V \cdot X)^Q$, and so may remove the superscript.

We now explain how to remove the drift term of a semimartingale using Girsanov Theorem.

Theorem . *Suppose B is a Brownian motion under P . Let f be progressive. Suppose there is a positive uniformly integrable continuous martingale Z , which satisfies $dZ/Z = f dB$. We now define another probability measure Q by $Q = Z_\infty / Z_0 \cdot P$. Then under the new measure Q , $\widetilde{B} = B - f \cdot (t)$ is a Brownian motion, and so $dB_t = d\widetilde{B}_t + f_t dt$. Here we use (t) to denote the process $X_t = t$ for all $t \geq 0$.*

Proof. Since $Z = (fZ) \cdot B$, we have $Z^{-1} \cdot [B, Z] = Z^{-1} \cdot (fZ \cdot (t)) = f \cdot (t)$. Since $Z_t = \mathbb{E}[Z_\infty | \mathcal{F}_t]$, we have $Q = Z_t \cdot P$ on \mathcal{F}_t for each $t \geq 0$. So by Theorem 16.19, $\tilde{B} = B - f \cdot (t) = B - Z^{-1} \cdot [B, Z]$ is a local martingale under Q . By Theorem 16.20, $[\tilde{B}]^Q = [B]^Q = [B]^P = (t)$. By Levy's characterization of Brownian motion, \tilde{B} is a Brownian motion under Q . \square

Remark . The above theorem also holds for n -dimensional, in which case $B = (B^1, \dots, B^n)$ is a n -dimensional Brownian motion, and $f = (f^1, \dots, f^n)$ is progressive. We use $f \cdot (t)$ to denote the process (h_t^1, \dots, h_t^n) such that $h_t^j = \int_0^t f_s^j ds$. If a positive uniformly integrable continuous martingale Z satisfies $dZ_t/Z_t = \sum_{j=1}^n f_t^j dB_t^j$, and we define a new measure Q by $Q = Z_\infty/Z_0 \cdot P$. Then $\tilde{B} = B - f \cdot (t)$ is an n -dimensional Brownian motion under Q .

Remark . We may remove the assumption that Z is uniformly integrable by localization. Suppose there is a stopping time τ such that Z^τ is uniformly bounded. Then we apply the theorem to Z^τ instead. We note that $dZ_t^\tau = \mathbf{1}_{[0, \tau]} f_t Z_t^\tau dB_t$. If we define a new measure Q^τ by $Q^\tau = Z_\tau/Z_0 \cdot P$, then under Q^τ , $\tilde{B}_t^\tau = B_t - \int_0^{t \wedge \tau} f_s ds$, $t \geq 0$, is a Brownian motion.

Example . The following example concerns Brownian motion and Brownian bridge. Recall that if B is a Brownian motion (started from 0), and $X_t = B_t - tB_1$, $0 \leq t \leq 1$, then X is a Brownian bridge. It has lifetime 1, and starts and ends at 0. It is a semimartingale: there is another Brownian motion \tilde{B} such that X satisfies the SDE:

$$dX_t = d\tilde{B}_t - \frac{X_t}{1-t} dt, \quad 0 \leq t < 1.$$

For $x \in \mathbb{R}$, define $X_t^x = X_t + tx$, $0 \leq t \leq 1$. Then $X_0^x = 0$ and $X_1^x = x$. We call X^x a Brownian bridge from 0 to x (in time 1). Then X^x satisfies the SDE:

$$dX_t^x = dX_t + xdt = d\tilde{B}_t - \frac{X_t - (1-t)x}{1-t} dt = d\tilde{B}_t - \frac{X_t^x - x}{1-t} dt.$$

We now start with a Brownian motion B . Recall that B has transition density:

$$f(t, x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}, \quad t > 0, \quad x \in \mathbb{R}.$$

We compute

$$\begin{aligned} \partial_x f &= \frac{-x}{\sqrt{2\pi t} \cdot t} e^{-\frac{x^2}{2t}} \\ \partial_x^2 f &= \frac{-1}{\sqrt{2\pi t} \cdot t} e^{-\frac{x^2}{2t}} + \frac{x^2}{\sqrt{2\pi t} \cdot t^2} e^{-\frac{x^2}{2t}} \\ \partial_t f &= \frac{-1/2}{\sqrt{2\pi t} \cdot t} e^{-\frac{x^2}{2t}} + \frac{x^2}{\sqrt{2\pi t} \cdot 2 \cdot t^2} e^{-\frac{x^2}{2t}}. \end{aligned}$$

So we have $\partial_t f = \frac{1}{2} \partial_x^2 f$. For each $x \in \mathbb{R}$, define

$$Z_t^x = f(1-t, B_t - x) = \frac{1}{\sqrt{2\pi(1-t)}} e^{-\frac{|B_t - x|^2}{2(1-t)}}, \quad 0 \leq t < 1.$$

Then

$$\begin{aligned} dZ_t^x &= \partial_x f(1-t, B_t-x)dB_t - \partial_t f(1-t, B_t-x)dt + \frac{1}{2}\partial_x^2 f(1-t, B_t-x)dt \\ &= \frac{-(B_t-x)e^{-\frac{(B_t-x)^2}{2(1-t)}}}{\sqrt{2\pi(1-t)}(1-t)}dB_t = Z_t^x \cdot \frac{-(B_t-x)}{1-t} dt. \end{aligned}$$

So Z_t^x , $0 \leq t < 1$, is a local martingale. In fact, Z_t^x is a true martingale (on $[0, 1)$) because for any $t_0 \in (0, 1)$, Z^x is uniformly bounded on $[0, t_0]$ by $\frac{1}{\sqrt{2\pi(1-t_0)}}$. Moreover, we have

$$\frac{dZ_t^x}{Z_t^x} = -\frac{B_t-x}{1-t}dB_t.$$

Fix $t_0 \in (0, 1)$ and define $\tilde{\mathbb{P}}_{t_0}^x$ such that $\tilde{\mathbb{P}}_{t_0}^x = Z_{t_0}^x/Z_0^x \cdot \mathbb{P}$. Then by Girsanov Theorem, under the new measure $\tilde{\mathbb{P}}_{t_0}^x$, the process

$$\tilde{B}_t^x = B_t + \int_0^{t \wedge t_0} \frac{B_s-x}{1-s} ds$$

is a Brownian motion. Since $dB_t = d\tilde{B}_t^x - \frac{B_t-x}{1-t}dt$ for $0 \leq t \leq t_0$, we see that the part of B up to t_0 under the new measure has the law of a Brownian bridge from 0 to x (with lifetime 1).

Now we change the point of view. Let μ_1 denote the law of B_t , $0 \leq t \leq 1$. For each $x \in \mathbb{R}$, let ν_1^x denote the law of a Brownian bridge from 0 to x with lifetime 1. They are all probability measures on $C([0, 1])$, which has a coordinate process X_t ($X_t(f) = f(t)$ for $f \in C([0, 1])$), $0 \leq t \leq 1$, and a natural filtration $(\mathcal{F}_t)_{0 \leq t \leq 1}$ generated by X . The above argument shows that, for any $x \in \mathbb{R}$ and $0 < t_0 < 1$,

$$\nu_1^x = f(1-t_0, X_{t_0}-x)/f(1, -x) \cdot \mu_1, \quad \text{on } \mathcal{F}_{t_0}.$$

Since the density function is measurable in x , we see that $x \mapsto \nu_1^x$ is a kernel on \mathcal{F}_{t_0} . Moreover, we may integrate and get

$$\int_{\mathbb{R}} \nu_1^x f(1, x) dx = \int_{\mathbb{R}} f(1-t_0, X_{t_0}-x) dx \cdot \mu_1 = \int_{\mathbb{R}} f(1-t_0, x) dx \cdot \mu_1 = \mu_1, \quad \text{on } \mathcal{F}_{t_0}.$$

Here in the last step we used the fact that f is a transition density. Since $\mathcal{F}_1 = \bigvee_{0 < t_0 < 1} \mathcal{F}_{t_0}$, by a monotone class argument, we can show that $x \mapsto \nu_1^x$ is a kernel on \mathcal{F}_1 , and

$$\mu_1 = \int_{\mathbb{R}} \nu_1^x f(1, x) dx = \mu_1, \quad \text{on } \mathcal{F}_1.$$

This formula can be viewed as a decomposition of a Brownian motion B stopped at time 1 according to its value at 1, and we may say that the conditional law of B_t , $0 \leq t \leq 1$, given the value of B_1 , is that of a Brownian bridge from 0 to B_1 in time 1.

Let $t_0 > 0$ and $x \in \mathbb{R}$. If B is a Brownian motion, then $X_t = B_t - \frac{t}{t_0}B_{t_0} + \frac{t}{t_0}x$, $0 \leq t \leq t_0$, is called a Brownian bridge with any lifetime t_0 from 0 to x . Let its law be denoted by $\nu_{t_0}^x$. Let μ_{t_0} denote the law of B_t , $0 \leq t \leq t_0$. Then a similar argument shows that

$$\mu_{t_0} = \int_{\mathbb{R}} \nu_{t_0}^x f(t_0, x) dx.$$

Exercise Extend the above result to n -dimensional Brownian motion and Brownian bridge.

Exercise Let B be a Brownian motion in \mathbb{C} started from 0. Let $\mathbb{D} = \{|z| < 1\}$. Let $\tau_{\mathbb{D}}$ be the first time that B exits \mathbb{D} . For each $w \in \partial\mathbb{D} = \{|z| = 1\}$, let $P_w(z) = \operatorname{Re} \frac{w-z}{w+z}$. Note that P_w is positive and harmonic in \mathbb{D} , and equals 0 on $\overline{\mathbb{D}} \setminus \{w\}$. It is called the Poisson kernel in D with pole w . An h -process X^w is a two-dimensional process satisfying the SDE:

$$dX_t^w = \tilde{d}B_t + \frac{\nabla P_w}{P_w}(X_t^w) dt, \quad X_0^w = 0,$$

where \tilde{B} is another Brownian motion in \mathbb{C} started from 0. The process is well defined up to the time that X^w exits \mathbb{D} . One can actually show that X^w exits \mathbb{D} through w . Let μ denote the law of B_t , $0 \leq t < \tau_{\mathbb{D}}$, and let ν^w denote the law of X^w . Prove that $\mu = \frac{1}{2\pi} \int_0^{2\pi} \nu^{e^{i\theta}} d\theta$. Hint: Consider the local martingales $Z_t^w := P_w(B_t)$, $0 \leq t < \tau_{\mathbb{D}}$, for $w \in \{|z| = 1\}$. This formula may be viewed as a decomposition of B up to $\tau_{\mathbb{D}}$ according to its value at $\tau_{\mathbb{D}}$.

Example . Let $\delta > 2$ and $x_0 > 0$. Let X be Bessel processes of dimensions δ started from x_0 . Recall that X has infinite lifetime and $\lim_{t \rightarrow \infty} X_t = \infty$. Moreover, $Z = (X/x_0)^{2-\delta}$ is a positive local martingale. Let τ_n be the first $t \geq 0$ such that $X_t \leq x_0/n$, $n \in \mathbb{N}$. Then Z^{τ_n} is a bounded true martingale. So $\mathbb{E}[Z_{\tau_n}] = Z_0 = 1$. Let $\tilde{\mathbb{P}}$ denote the underlying probability measure. Then we may define a new probability measure $\tilde{\mathbb{P}}_n$ by $d\tilde{\mathbb{P}}_n = Z_{\tau_n} d\tilde{\mathbb{P}}$. Then for any $t \geq 0$, $d\tilde{\mathbb{P}}_n = Z_{\tau_n \wedge t} d\tilde{\mathbb{P}}$ on \mathcal{F}_t . By Itô's formula, Z satisfies the SDE

$$\frac{dZ_t}{Z_t} = \frac{2-\delta}{X_t} dX_t + \frac{1}{2} \frac{(2-\delta)(1-\delta)}{X_t^2} dt = \frac{2-\delta}{X_t} dB_t.$$

So by Girsanov Theorem, under the new measure $\tilde{\mathbb{P}}_n$,

$$\tilde{B}_t^n := B_t - \int_0^{\tau_n \wedge t} \frac{2-\delta}{X_s} ds$$

is a Brownian motion. We find that

$$d\tilde{B}_t^n = dB_t - \mathbf{1}_{[0, \tau_n]} \frac{2-\delta}{X_t} dt.$$

Since $dX_t = dB_t + \frac{\delta/2-1/2}{X_t} dt$, we get

$$dX_t = d\tilde{B}_t^n + \frac{3/2-\delta/2}{X_t} dt = d\tilde{B}_t^n + \frac{\tilde{\delta}/2-1/2}{X_t} dt$$

up to τ_n , where we define $\tilde{\delta} = 4 - \delta < 2$. So under $\tilde{\mathbb{P}}_n$, X is a Bessel process of dimension $\tilde{\delta}$ up to τ_n .

Now we change the point of view. We use μ and $\tilde{\mu}$ to denote the laws of a Bessel process of dimension δ and $\tilde{\delta}$, respectively, started from x_0 . They are probability measures on $\Sigma := \bigcup_{T \in (0, \infty]} C([0, T], \mathbb{R})$. Each $f \in \Sigma$ has a finite or infinite lifetime, which is denoted by T_f . Then μ -a.s. $T_f = \infty$ and $\tilde{\mu}$ -a.s. $T_f < \infty$ because a Bessel process of dimension > 2 and < 2 have infinite and finite lifetime, respectively. Let X be the coordinate process on Σ , i.e., $X_t(f) = f(t)$, $0 \leq t < T_f$. There is a natural filtration \mathcal{F}^Σ on Σ , which is generated by X . Let τ_n be the first $t \geq 0$ such that $X_t \leq 1/n$. Let $Z = (X/x_0)^{\delta-2}$. The above argument shows that, under the measure μ , Z is a local martingale and $Z_{\tau_n}^\tau$ is a true martingale. Moreover,

$$\tilde{\mu} = Z_{\tau_n} \cdot \mu \quad \text{on } \mathcal{F}_{\tau_n}^\Sigma.$$

We claim that for every $t \geq 0$, $\tilde{\mu} = (X_t/x_0) \cdot \mu$ on $\mathcal{F}_t^\Sigma \cap \{f : T_f > t\}$, i.e., if $A \in \mathcal{F}_t^\Sigma$ and $A \subset \{T > t\}$, then $\tilde{\mu}[A] = \int_A Z_t d\mu$. To see this, fix $A \in \mathcal{F}_t^\Sigma$ and let $A_n = A \cap \{t \leq \tau_n\} \in \mathcal{F}_{\tau_n}$. Since $\tilde{\mu} = Z_{\tau_n} \cdot \mu$ on \mathcal{F}_{τ_n} , by Optional Stopping Theorem and the martingale property of Z_{τ_n} , we get

$$\tilde{\mu}[A_n] = \int_{A_n} Z_{\tau_n} d\mu = \int_{A_n} Z_t d\mu.$$

From $A \subset \{T > t\}$ we get $A_n \uparrow A$. Sending $n \rightarrow \infty$ in the above displayed formula and using monotone convergence theorem, we get $\tilde{\mu}[A] = \int_A Z_t d\mu$, as desired.

We observe that the Z in the above example is a local martingale but not a true martingale. In fact, since under μ , a.s. $T = \infty$, we have,

$$\mathbb{E}_\mu[Z_t] = \int_\Omega Z_t d\mu = \int_{\{T > t\}} Z_t d\mu = \tilde{\mu}[T > t] \rightarrow 0, \quad \text{as } t \rightarrow \infty,$$

where the limit holds because $\tilde{\mu}$ -a.s. $T < \infty$. So we can not have $\mathbb{E}_\mu[Z_t] = 1 = Z_0$ for all $t \geq 0$.

Example . In the previous example, we changed a Bessel process of dimension $\delta > 2$ to a Bessel process of dimension $\tilde{\delta} = 4 - \delta < 2$ using the local martingale $(X_t/x_0)^{2-\delta}$. For general $\delta, \tilde{\delta} \in \mathbb{R}$, we can still find a local martingale for the same purpose. The local martingale may be expressed as

$$Z_t = (X_t/x_0)^{(\tilde{\delta}-\delta)/2} \times \text{some } C^1 \text{ function.}$$

We first calculate

$$\begin{aligned} \frac{dX_t^{(\tilde{\delta}-\delta)/2}}{X_t^{(\tilde{\delta}-\delta)/2}} &= \frac{(\tilde{\delta}-\delta)/2}{X_t} dX_t + \frac{1}{2} \frac{(\tilde{\delta}-\delta)/2((\tilde{\delta}-\delta)/2-1)}{X_t^2} dt \\ &= \frac{(\tilde{\delta}-\delta)/2}{X_t} dB_t + \frac{(\tilde{\delta}-\delta)(\tilde{\delta}+\delta-4)}{8X_t^2} dt. \end{aligned}$$

So the correct formula for Z is

$$Z_t = (X_t/x_0)^{(\tilde{\delta}-\delta)/2} \times \exp\left(\int_0^t -\frac{(\tilde{\delta}-\delta)(\tilde{\delta}+\delta-4)}{8X_s^2} ds\right)$$

because

$$\frac{dZ_t^{\delta,\tilde{\delta}}}{Z_t^{\delta,\tilde{\delta}}} = \frac{dX_t^{(\tilde{\delta}-\delta)/2}}{X_t^{(\tilde{\delta}-\delta)/2}} + \frac{(\delta-\tilde{\delta})(\tilde{\delta}+\delta-4)}{8X_t^2} dt = \frac{(\tilde{\delta}-\delta)/2}{X_t} dB_t.$$

If we let μ_δ and $\mu_{\tilde{\delta}}$ be respectively the laws of a Bessel process of dimension δ and $\tilde{\delta}$ started from x_0 , then for any $t \geq 0$, $\mu_{\tilde{\delta}} = Z_t^{\delta,\tilde{\delta}} \cdot \mu_\delta$ on $\mathcal{F}_t^\Sigma \cap \{T > t\}$.

In the above examples, we find explicit formulas for the RN process Z . For a general progressive process f , if $f \in L(B)$, i.e., $\int_0^t f_s^2 ds < \infty$ for all $t \geq 0$, then a solution of $dZ = fZdB$ may be expressed as a stochastic exponential:

$$Z = \mathcal{E}(f \cdot B) = \exp(f \cdot B - \frac{1}{2}f^2 \cdot (t)).$$

If such Z is uniformly integrable, then we may define a new measure Q by $Q = Z_\infty/Z_0 \cdot P$. Then under Q , $\tilde{B} := B - f \cdot (t)$ is a Brownian motion. Let P_B and $P_{B-f \cdot (t)}$ denote the law of B and $B - f \cdot (t)$, respectively. Then under P and Q , \tilde{B} respectively has the law $P_{B-f \cdot (t)}$ and P_B . From $Q \ll P$, we find $P_B \ll P_{B-f \cdot (t)}$. If we further now that $Z_\infty > 0$. Then $P \ll Q$ as well, and so $P_{B-f \cdot (t)} \ll P_B$. In that case, we can say that $B - f \cdot (t)$ satisfies every almost sure property of a Brownian motion.

If we do not have the condition that Z is uniformly integrable, then we may find a sequence of stopping times (τ_n) , which increases to ∞ , such that Z^{τ_n} is uniformly bounded for each n . We may then apply the theorem to Z^{τ_n} . In this case we can conclude that $B - f \cdot (t)$ satisfies every almost sure **local** property of a Brownian motion, such as Hölder continuity and the law of iterated logarithm at 0.

Finally, we state a criterion to check when $\mathcal{E}(f \cdot B)$ is a uniformly integrable martingale so that we do not need to do localization.

Theorem 16.23 (Nivikov condition). *Let M be a local martingale with $M_0 = 0$ and*

$$\mathbb{E}[e^{[M]_\infty/2}] < \infty.$$

Then $\mathcal{E}(M) = \exp(M - \frac{1}{2}[M])$ is a uniformly integrable martingale.

In particular, if $M = f \cdot B$, then $[M]_\infty = \int_0^\infty f_s^2 ds$, and the Nivikov condition becomes $\mathbb{E}[e^{\frac{1}{2} \int_0^\infty f_s^2 ds}] < \infty$.

Lemma . *Let B be a Brownian motion. Then $\mathcal{E}(B)_t = e^{B_t-t/2}$ is a martingale.*

Proof. Fix $t > s \geq 0$. Since $B_t - B_s \perp\!\!\!\perp \mathcal{F}_s$, we have $e^{B_t - B_s} \perp\!\!\!\perp \mathcal{F}_s$. Recall that $B_t - B_s$ has density $f(t - s, x) = \frac{1}{\sqrt{2\pi(t-s)}} e^{-\frac{x^2}{2(t-s)}}$. So

$$\begin{aligned} \mathbb{E}[e^{B_t - B_s} | \mathcal{F}_s] &= \mathbb{E}[e^{B_t - B_s}] = \int_{\mathbb{R}} e^x f(t - s, x) dx = \int_{\mathbb{R}} \frac{e^{-\frac{x^2}{2(t-s)} + x}}{\sqrt{2\pi(t-s)}} dx \\ &= e^{\frac{t-s}{2}} \int_{\mathbb{R}} \frac{e^{-\frac{(x-(t-s))^2}{2(t-s)}}}{\sqrt{2\pi(t-s)}} dx = e^{(t-s)/2}, \end{aligned}$$

which implies that $\mathbb{E}[e^{B_t - t/2} | \mathcal{F}_s] = e^{B_s - s/2}$. \square

Lemma 16.22. *Suppose $\Omega = C(\mathbb{R}_+, \mathbb{R})$, and \mathbb{P} is the law of Brownian motion. This means that the coordinate process under \mathbb{P} is a Brownian motion. Let B denote this coordinate process. Let \mathcal{F} be the right-continuous and P -complete filtration generated by B . Let \mathbb{P}' denote the law of a Brownian motion plus the process (t) . Then for any $t \geq 0$, $\mathbb{P}' = \mathcal{E}(B)_t \cdot \mathbb{P}$ on \mathcal{F}_t .*

Proof. Let $Z = \mathcal{E}(B)$. Then Z is a martingale by the lemma. Fix $t_0 \geq 0$. Then $Z^{t_0} = \mathcal{E}(\mathbf{1}_{[0, t_0]} B)$ is a uniformly integrable martingale. If we define a new probability measure Q_{t_0} on Ω by $dQ_{t_0}/d\mathbb{P} = Z_{t_0}^{t_0} = Z_{t_0}$, then under the new measure Q_{t_0} , $\tilde{B}_t^{(t_0)} = B_t - \int_0^t \mathbf{1}_{[0, t_0]}(s) ds = B_t - t \wedge t_0$ is a Brownian motion. In other words, under Q_{t_0} , up to the time t_0 , B is the sum of a Brownian motion \tilde{B} and the function t , i.e., the law of B under Q_{t_0} is \mathbb{P}' on \mathcal{F}_{t_0} . Since $Q_{t_0} = Z_{t_0} \cdot \mathbb{P}$, and Z_{t_0} is \mathcal{F}_{t_0} -measurable, we get the conclusion. \square

Lemma 16.24. *Let B be a Brownian motion. Let τ be a stopping time such that $\mathbb{E}[e^{\tau/2}] < \infty$. Then $\mathbb{E}[\mathcal{E}(B)_\tau] = \mathbb{E}[e^{B_\tau - \frac{\tau}{2}}] = 1$.*

Proof. Let $Z = \mathcal{E}(B)$. Although Z is a martingale with $Z_0 = 1$, we can not apply Optional Stopping Theorem directly because Z is not uniformly integrable and τ may not be bounded.

We may suppose that $\Omega = C(\mathbb{R}_+, \mathbb{R})$, \mathbb{P} is the law of a Brownian motion, and B is the coordinate process (so is a Brownian motion under \mathbb{P}). Let \mathbb{P}' denote the law of a Brownian motion plus the process (t) . By Lemma 16.22, for every $t_0 \geq 0$, $\mathbb{P}' = Z_{t_0} \cdot \mathbb{P}$ on \mathcal{F}_{t_0} . By Lemma 16.16, for any stopping time T , $\mathbb{P}' \ll \mathbb{P}$ on $\mathcal{F}_T \cap \{T < \infty\}$, and the Radon-Nikodym derivative is Z_T . Thus, T is such that $\mathbb{P}'[T < \infty] = 1$, we then have $\mathbb{E}[Z_T] = 1$.

Fix $b > 0$ and let τ_b be the first time that $B_t - t = -b$; and $\tau_b = \infty$ when such time does not exist. Then $\mathbb{P}'[\tau_b < \infty] = 1$ since under \mathbb{P}' , $B_t - t$ has the law of a Brownian motion, which is a.s. recurrent. So we get $\mathbb{E}[Z_{\tau_b}] = 1 = Z_0$. Since $Z > 0$, from $\mathbb{E}[Z_{\tau_b}] = 1$ we know that Z^{τ_b} is a uniformly integrable martingale (exercise). Thus, by Optional Stopping Theorem,

$$1 = Z_0 = \mathbb{E}[Z_{\tau_b \wedge \tau}] = \mathbb{E}[\mathbf{1}_{\{\tau \leq \tau_b\}} e^{B_\tau - \tau/2}] + \mathbb{E}[\mathbf{1}_{\{\tau_b < \tau\}} e^{B_{\tau_b} - \tau_b/2}]. \quad (4.1)$$

When $\tau_b < \infty$, $B_{\tau_b} - \tau_b = -b$. So $B_{\tau_b} - \tau_b/2 = \tau_b/2 - b$. From $\mathbb{E}[e^{\tau/2}] < \infty$ we get

$$\mathbb{E}[\mathbf{1}_{\{\tau_b < \tau\}} e^{B_{\tau_b} - \tau_b/2}] = e^{-b} \mathbb{E}[\mathbf{1}_{\{\tau_b < \tau\}} e^{\tau_b/2}] \leq e^{-b} \mathbb{E}[e^{\tau/2}] \rightarrow 0, \quad \text{as } b \rightarrow \infty.$$

Letting $b \rightarrow \infty$ in (4.1) we get $\mathbb{E}[e^{B_\tau - \tau/2}] = 1$. \square

Proof of Theorem 16.23. Let $u_t = [M]_t$, $S = \sup\{u_t : t \geq 0\} = [M]_\infty$, and $v_s = \inf\{t \geq 0 : u_t > s\}$. Then in an enlarged probability space M_{v_s} , $0 \leq s < S$, extends to a Brownian motion B , and S is a stopping time for B . By the assumption, we have $\mathbb{E}[e^{S/2}] < \infty$. By Lemma 16.24, $\mathbb{E}[e^{M_\infty - [M]_\infty/2}] = \mathbb{E}[e^{B_S - S/2}] = 1$. Thus, $e^{M - [M]/2}$ is a uniformly integrable martingale. \square

5 Stochastic Differential Equations

A stochastic differential equation (SDE) has the following general form.

$$dX_t^i = \sum_{j=1}^{\delta} \sigma_j^i(t, X) dB_t^j + b^i(t, X) dt, \quad 1 \leq i \leq d.$$

or equivalently,

$$X_t^i - X_0^i = \sum_j \int_0^t \sigma_j^i(s, X) dB_s^j + \int_0^t b^i(s, X) ds, \quad 1 \leq i \leq d.$$

Here $X = (X^1, \dots, X^d)$ is a continuous vector semimartingale and $B = (B^1, \dots, B^\delta)$ is a δ -dimensional Brownian motion. The σ_j^i and b^i , $1 \leq i \leq d$, $1 \leq j \leq \delta$, are real valued functions defined on

$$\mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^d),$$

which are progressive in the sense that for any fixed $t_0 > 0$, σ_j^i and b^i restricted to the subspace

$$[0, t_0] \times C(\mathbb{R}_+, \mathbb{R}^d)$$

are measurable w.r.t. $\mathcal{B}([0, t_0]) \times \mathcal{F}_{t_0}^c$, where $\mathcal{F}_{t_0}^c$ is the σ -algebra on $C(\mathbb{R}_+, \mathbb{R})$ generated by the coordinate process up to t_0 . This means that, for a fixed t , the values of $\sigma_j^i(t, w)$ and $b^i(t, w)$, $w \in C(\mathbb{R}_+, \mathbb{R})$, depend only on the values of $w(s)$, $0 \leq s \leq t$. Note that the second variable of σ_j^i and b^i is not a vector in \mathbb{R}^d but an \mathbb{R}^d -valued continuous function, and there are no randomness in σ_j^i and b^i . We write $\sigma = (\sigma_j^i)$ and $b = (b^i)$, which are $\mathbb{R}^{d \times \delta}$ -valued and \mathbb{R}^d -valued, respectively. We often write the SDE in the vector form:

$$dX_t = \sigma(t, X) \circ dB_t + b(t, X) dt, \tag{5.1}$$

or

$$X_t - X_0 = \int_0^t \sigma(s, X) \circ dB_s + \int_0^t b(s, X) ds. \tag{5.2}$$

A simple case of σ and b is that there exist measurable functions $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times \delta}$ and $b : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $\sigma(t, w) = \sigma(t, w_t)$ and $b(t, w) = b(t, w_t)$. Then the SDE becomes

$$dX_t = \sigma(t, X_t) \circ dB_t + b(t, X_t) dt.$$

If σ and b further do not depend on t , then the equation becomes

$$dX_t = \sigma(X_t) \circ dB_t + b(X_t)dt,$$

and is said to be time-homogeneous. We emphasize that there are also other possibilities of σ and b . For example, $b(t, w) = \int_0^t w_s ds$.

We call (5.1) or (5.2) SDE(σ, b). If we further require that $X_0 = x$ for some $x \in \mathbb{R}^d$, then the equation is called SDE $_x(\sigma, b)$. The x may also be replaced by a random variable ζ , which must be independent of B , or a probability measure μ on \mathbb{R}^d .

There are two kinds of solutions of SDE(σ, b): *strong solution* and *weak solution*.

For the strong solution, besides σ and b , we are also given a d -dim Brownian motion B on some probability space and an initial value, which is a deterministic point $x \in \mathbb{R}^d$ or a random vector ζ in \mathbb{R}^d independent of B . We require that the solution X to be adapted to \mathcal{F}^{B, X_0} , which is the complete and right-continuous filtration generated by B and X_0 . The SDE is required to hold w.r.t. \mathcal{F}^{B, X_0} .

For the weak solution, besides σ, b , we are only given the initial value $x \in \mathbb{R}^d$ or an initial distribution μ on \mathbb{R}^d . We are not given the Brownian motion or the probability space. Instead, we have the freedom to choose a Brownian motion. A weak solution of the SDE $_\mu(\sigma, b)$ is a package:

$$(\Omega, \mathcal{F}, \mathbb{P}), \quad B, \quad X,$$

where $(\Omega, \mathcal{F}, \mathbb{P})$ is a filtered space such that \mathcal{F} is complete and right-continuous, B is a d -dim \mathcal{F} -Brownian motion, X is a d -dim \mathcal{F} -semimartingale with $X_0 = x$ or $X_0 \sim \mu$, and (5.2) holds w.r.t. the filtration \mathcal{F} . Note that the filtered space and the Brownian motion are also parts of the solution.

The relations between the two kinds of solutions are (i) if X is a strong solution for the given B , then $(\Omega, \mathcal{F}^{B, X_0}, B, X)$ together is a weak solution. (ii) if (\mathcal{F}, B, X) is a weak solution, and X is \mathcal{F}^{B, X_0} -adapted, then X is a strong solution for such B . Condition (ii) does not hold in general.

We will discuss the uniqueness of the solution. If two strong solutions X and X' of SDE(σ, b) with the same initial value both have infinite lifetime, the uniqueness means that a.s. $X = X'$. We may allow that a solution to have a finite lifetime. In that case, the uniqueness means that if $X_t, 0 \leq t < T$, and $X'_t, 0 \leq t < T'$, are strong solutions, then a.s. $X = X'$ on $[0, T \wedge T')$. If the uniqueness holds, then there exists an a.s. unique solution with the maximal interval so that all other solutions are its restrictions.

There are two kinds of uniqueness for weak solutions. We say that pathwise uniqueness holds if for any two weak solutions (B, X) and (B', X') of SDE(σ, b) defined on the same filtered space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $B = B'$ and $X_0 = X'_0$, we have a.s. X and X' agree on the common time interval. We say that uniqueness in law holds if any two weak solutions (B, X) and (B', X') of SDE $_\mu(\sigma, b)$ with the maximal interval have the same joint finite dimensional distribution.

Example (A trivial example). Let $d = \delta$; $\sigma_j^i = 1$ if $i = j$, $= 0$ if $i \neq j$; and $b^i \equiv 0$. The SDE(σ, b) becomes

$$dX_t = dB_t.$$

For ζ independent of B , $X = \zeta + B$ is a strong solution of the SDE with initial value ζ .

Example . Let $d = \delta = 1$, $\sigma(t, w) = w_t$, and $b \equiv 0$. The SDE(σ, b) becomes

$$dX_t = X_t dB_t.$$

The stochastic exponential $\mathcal{E}(B)_t = e^{B_t - t/2}$ is a strong solution of the SDE with initial value 1. For ζ independent of B , $X = \zeta \mathcal{E}(B)$ is a strong solution of the SDE with initial value ζ .

Example . Bessel process of dimension δ is the solution of the SDE:

$$dX_t = dB_t + \frac{\delta - 1}{2X_t} dt.$$

Recall that we get the existence and uniqueness of the strong solution by transforming it into an ODE.

Example (Brownian motion on the unit circle). Let B be a 1-dimensional Brownian motion. Let $X = \cos(B)$ and $Y = \sin(B)$. Then (X, Y) takes values in $\{(x, y) : x^2 + y^2 = 1\}$, and the argument runs as a Brownian motion. It is called a Brownian motion on the unit circle. Itô's formula shows that they satisfy the SDE:

$$\begin{aligned} dX_t &= -Y_t dB_t - \frac{1}{2} X_t dt; \\ dY_t &= X_t dB_t - \frac{1}{2} Y_t dt. \end{aligned}$$

On the other hand, if (X, Y) solves this SDE, then (X, Y) stays on a circle because

$$d(X_t^2 + Y_t^2) = 2X_t dX_t + 2Y_t dY_t + Y_t^2 dt + X_t^2 dt = 0.$$

Example (A weak solution not a strong solution). Let $d = \delta = 1$. Let W be a 1-dimensional Brownian motion. Let

$$B_t = \int_0^t \text{sign}(W_s) dW_s, \quad t \geq 0,$$

where $\text{sign}(x) \in \{1, 0, -1\}$ depending on the sign of x . Then B is a local martingale with $[B]_t = \int_0^t \text{sign}(W_s)^2 ds = t$, $t \geq 0$. Here we used the fact that the set $B^{-1}(0)$ has Lebesgue measure zero. By Levy's characterization of Brownian motion, B is also a Brownian motion. Since $W_t = \int_0^t \text{sign}(W_s) dB_s$, $t \geq 0$, (W, B) is a weak solution of

$$dW = \text{sign}(W) dB, \quad W(0) = 0. \tag{5.3}$$

Here the underlying filtration is \mathcal{F}^W generated by W . This solution does not satisfy the pathwise uniqueness. In fact, $(-W, B)$ is also a weak solution.

We claim that W is not a strong solution of (5.3). We show that W is not adapted to the filtration \mathcal{F}^B generated by B . For the proof, take a sequence of convex even C^2 functions F_n

such that $F_n(x) = |x|$ if $|x| \geq 1/n$, and for each $x \in \mathbb{R}$, $F_n(x) \downarrow |x|$. Let $f_n = F'_n$. Then $|f_n| \leq 1$ and $f_n(x) \rightarrow \text{sign}(x)$ on \mathbb{R} . By Itô's formula, for any n ,

$$F_n(W_t) - F_n(W_0) = \int_0^t f_n(W_s) dW_s + \frac{1}{2} \int_0^t F''_n(W_s) ds, \quad t \geq 0.$$

Let $n \rightarrow \infty$. Then $F_n(W_t) - F_n(W_0) \rightarrow |W_t|$. By dominated convergence theorem for stochastic integral, $\int_0^t f_n(W_s) dW_s \rightarrow \int_0^t \text{sign}(W_s) dW_s = B_t$. Since F''_n is even, we get

$$B_t = |W_t| - \lim_{n \rightarrow \infty} \frac{1}{2} \int_0^t F''_n(|W_s|) ds, \quad t \geq 0.$$

This means that B is adapted to the filtration $\mathcal{F}^{|W|}$ generated by $|W|$. If W is adapted to \mathcal{F}^B , then $\mathcal{F}^W \subset \mathcal{F}^B \subset \mathcal{F}^{|W|} \subset \mathcal{F}^W$. We should have $\mathcal{F}^W = \mathcal{F}^{|W|}$. However, for a Brownian motion W , $\mathcal{F}^W \neq \mathcal{F}^{|W|}$ because for any fixed $t_0 > 0$, the event $\{W(t_0) > 0\}$ belongs to \mathcal{F}^W but not to $\mathcal{F}^{|W|}$. This is a contradiction.

Let $L_t = \lim_{n \rightarrow \infty} L_t^n = \lim_{n \rightarrow \infty} \frac{1}{2} \int_0^t F''_n(|W_s|) ds$, $t \geq 0$. Then

$$|W| = B + L.$$

Since for each n , $F''_n \geq 0$, each L^n is increasing, and so is L . Thus, $|W|$ is a semimartingale with

$$d|W_t| = dB_t + dL_t = \text{sign}(W_t) dW_t + dL_t.$$

The equation is called Itô-Tanaka equation, and L is called the local time of the Brownian motion W . Since $F''_n(x) = 0$ for $|x| \geq 1/n$, L^n stays constant on each interval on which $|W| \geq 1/n$, which implies that L stays constant on each open interval on which $W \neq 0$. So it increases only at the time when $W = 0$.

Example . Consider the d -dimensional SDE

$$dX_t = dB_t + b(t, X_t) dt, \quad (5.4)$$

where B is a d -dimensional Brownian motion and $b : [0, \infty) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$. Under certain conditions, we may use Girsanov Theorem to construct a weak solution. First let B be a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. Define a local martingale $Z = \mathcal{E}(\sum_j b^j(s, B) \cdot B^j)$. Suppose the Nivikov condition is satisfied, i.e.,

$$\mathbb{E} \left[\exp \left(\frac{1}{2} \int_0^\infty \|b(s, B_s)\|^2 ds \right) \right] < \infty.$$

This is the case, for example, if there is a function $b \in L^2(\mathbb{R}_+, \lambda)$ such that $\|b(t, x)\| \leq |b(t)|$ for each $t \geq 0$ and $x \in \mathbb{R}^d$. Then Z is a uniformly integrable martingale. We define another probability measure $\tilde{\mathbb{P}}$ on (Ω, \mathcal{F}) such that $\tilde{\mathbb{P}} = Z_\infty \cdot \mathbb{P}$. Then $\tilde{\mathbb{P}} = Z_t \cdot \mathbb{P}$ on \mathcal{F}_t for each $t \geq 0$. By Girsanov Theorem, under the new measure $\tilde{\mathbb{P}}$, $\tilde{B}_t := B_t - \int_0^t b(s, B_s) ds$, $t \geq 0$, is a d -dimensional Brownian motion. Now B and \tilde{B} satisfy

$$dB_t = d\tilde{B}_t + b(t, B_t) dt.$$

Thus, (B, \tilde{B}) on $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$ is a weak solution of (5.4). Now B plays the role of X , while \tilde{B} plays the role of B in the SDE.

5.1 Strong solution: existence and uniqueness

We now consider the existence and uniqueness of the strong solution of the following SDE

$$X_t - X_0 = \int_0^t \sigma(s, X_s) \cdot dB_s + \int_0^t b(s, X_s) ds, \quad (5.5)$$

where $\sigma = (\sigma_j^i) \in C(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R}^{d \times \delta})$ and $b = (b^i) \in C(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R}^d)$. The theory parallels the main existence/uniqueness result for ordinary differential equations.

We use L^2 to denote the space of random vectors ζ in \mathbb{R}^d such that $\|\zeta\|_2 = \mathbb{E}[\|\zeta\|^2]^{1/2} < \infty$.

Theorem 18.3 (A simplified version). *Fix a d -dimensional Brownian motion B . Suppose that there is $K > 0$ such that for any $t \in \mathbb{R}_+$ and $x, y \in \mathbb{R}^d$,*

$$\|\sigma(t, x)\| + \|b(t, x)\| \leq K(1 + \|x\|); \quad (5.6)$$

$$\|\sigma(t, x) - \sigma(t, y)\| + \|b(t, x) - b(t, y)\| \leq K\|x - y\|. \quad (5.7)$$

Then we have the following.

- (i) For any $\zeta \in L^2$ independent of B , a strong solution X^ζ of $SDE_\zeta(\sigma, b)$ (with initial value ζ) exists, and $X_t^\zeta \in L^2$ for any $t \geq 0$.
- (ii) The strong solution in (i) is unique, and any weak solution of $SDE_\zeta(\sigma, b)$ with initial value $\zeta \in L^2$ is a strong solution.
- (iii) We may choose a version of the solution X^x with initial value x for each $x \in \mathbb{R}^d$ such that $\mathbb{R}^d \times \mathbb{R}_+ \ni (x, t) \mapsto X_t^x$ is continuous. This means that the solution depends continuously on the initial value.
- (iv) Let X_t^x , $x \in \mathbb{R}^d$, $t \in \mathbb{R}_+$, be as in (iii). Then for any $\zeta \in L^2$, $X^x|_{x=\zeta}$ is a strong solution of $SDE_\zeta(\sigma, b)$. In other words, the solution X^ζ in (i) a.s. equals $X^x|_{x=\zeta}$.

Here the norm $\|\sigma(t, x)\|$ is defined by $\|\sigma(t, x)\|^2 = \sum_i \sum_j |\sigma_j^i(t, x)|^2$, and the norm $\|b(t, x)\|$ is defined by $\|b(t, x)\|^2 = \sum_i |b^i(t, x)|^2$.

Lemma 18.4 (Gronwall). *Let $a, b > 0$. Let $f \in C(\mathbb{R}_+, \mathbb{R}_+)$ such that*

$$f(t) \leq a + b \int_0^t f(s) ds, \quad t \geq 0.$$

Then $f(t) \leq ae^{bt}$ for all $t \geq 0$.

Proof. Let $F(t) = \int_0^t f(s) ds$, $t \geq 0$. Then $F \in C^1(\mathbb{R}_+, \mathbb{R})$, and $F' = f \leq a + bF$. Let $G(t) = e^{-bt}F(t)$. Then $G'(t) = e^{-bt}(F'(t) - bF(t)) \leq ae^{-bt}$ and $G(0) = 0$. So $G(t) \leq \int_0^t ae^{-bs} ds = a(1 - e^{-bt})/b$. Thus, $F(t) = e^{bt}G(t) \leq a(e^{bt} - 1)/b$, and $f(t) \leq a + bF(t) \leq ae^{bt}$. \square

Let \mathcal{F} be a filtration, and B be an \mathcal{F} -Brownian motion. For any continuous \mathcal{F} -adapted process X in \mathbb{R}^d , we define an \mathcal{F} -vector semimartingale $S(X)$ in \mathbb{R}^d by

$$S(X)_t = \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) \circ dB_s, \quad t \geq 0.$$

If $X = X_0 + S(X)$, then (X, B) is a weak solution of SDE(σ, b). If we further have $\mathcal{F} = \mathcal{F}^{B, X_0}$, then X is a strong solution.

Lemma 18.5. *Let X and Y be adapted continuous processes in \mathbb{R}^d such that $X_t, Y_t \in L^2$ for all $t \geq 0$. (i) If σ and b satisfy (5.6), then*

$$\mathbb{E}[\sup_{0 \leq s \leq t} \|S(X)_s\|^2] \leq 4K^2(t+4)(t + \int_0^t \mathbb{E}[\|X_s\|^2] ds). \quad (5.8)$$

(ii) If σ and b satisfy (5.7), then

$$\mathbb{E}[\sup_{0 \leq s \leq t} \|S(X)_s - S(Y)_s\|^2] \leq 2K^2(t+4) \int_0^t \mathbb{E}[\|X_s - Y_s\|^2] ds. \quad (5.9)$$

Proof. (i) Using that $(x+y)^2 \leq 2x^2 + 2y^2$, we get

$$\mathbb{E}[\sup_{0 \leq s \leq t} \|S(X)_s\|^2] \leq 2\mathbb{E}[\sup_{0 \leq s \leq t} \|\int_0^s b(u, X_u) du\|^2] + 2\mathbb{E}[\sup_{0 \leq s \leq t} \|\int_0^s \sigma(u, X_u) \circ dB_u\|^2].$$

For the first term on the RHS, since $\|b(t, x)\| \leq K(1 + \|x\|)$, by Cauchy-Schwarz inequality,

$$\begin{aligned} & \mathbb{E}[\sup_{0 \leq s \leq t} \|\int_0^s b(u, X_u) ds\|^2] \leq \mathbb{E}[(\int_0^t \|b(s, X_s)\| ds)^2] \leq \mathbb{E}[t \int_0^t \|b(s, X_s)\|^2 ds] \\ & = t \int_0^t \mathbb{E}[\|b(s, X_s)\|^2] ds \leq t \int_0^t \mathbb{E}[K^2(1 + \|X_s\|)^2] ds \leq 2K^2 t \int_0^t (1 + \mathbb{E}[\|X_s\|^2]) ds. \end{aligned} \quad (5.10)$$

For the second term on the RHS, since $\|\sigma(t, x)\| \leq K(1 + \|x\|)$, using Doob's martingale inequality, we get

$$\begin{aligned} & \mathbb{E}[\sup_{0 \leq s \leq t} \|\int_0^s \sigma(u, X_u) \circ dB_u\|^2] \leq 4\mathbb{E}[\|\int_0^t \sigma(s, X_s) \circ dB_s\|^2] \\ & = 4\mathbb{E}[\sum_{i=1}^d (\int_0^t \sum_{j=1}^{\delta} \sigma_j^i(s, X_s) dB_s^j)^2] = 4 \sum_{i=1}^d \mathbb{E}[\int_0^t \sum_{j=1}^{\delta} \sigma_j^i(s, X_s)^2 ds] \\ & = 4 \int_0^t \mathbb{E}[\|\sigma(s, X_s)\|^2] ds \leq 4 \int_0^t \mathbb{E}[(K(1 + \|X_s\|))^2] ds \leq 8K^2 \int_0^t (1 + \mathbb{E}[\|X_s\|^2]) ds. \end{aligned} \quad (5.11)$$

Combining (5.10) with (5.11) we get (5.8).

(ii) We have

$$\mathbb{E}[\|S(X)_t - S(Y)_t\|^2] \leq 2\mathbb{E}[\|\int_0^t b(s, X_s) - b(s, Y_s) ds\|^2] + 2\mathbb{E}[\|\int_0^t (\sigma(s, X_s) - \sigma(s, Y_s)) \circ dB_s\|^2].$$

For the first term, using a similar argument as in (5.10) and that $\|b(t, x) - b(t, y)\| \leq K\|x - y\|$, we get

$$\mathbb{E}[\|\int_0^t b(s, X_s) - b(s, Y_s) ds\|^2] \leq K^2 t \int_0^t \mathbb{E}[\|X_s - Y_s\|^2] ds.$$

For the second term, using a similar argument as in (5.11) and that $\|\sigma(t, x) - \sigma(t, y)\| \leq K\|x - y\|$, we get

$$\mathbb{E}[\|\int_0^t (\sigma(s, X_s) - \sigma(s, Y_s)) \circ dB_s\|^2] \leq 4K^2 \int_0^t \mathbb{E}[\|X_s - Y_s\|^2] ds.$$

Combining the two displayed formulas, we get (5.9). \square

Corollary . *Suppose σ and b satisfy (5.6,5.7). If (X, B) is a weak solution of (5.5) such that $X_0 \in L^2$, then for any $t \geq 0$,*

$$\mathbb{E}[\|X_t\|^2] \leq (2\mathbb{E}[\|X_0\|^2] + 8K^2(t+4)t)e^{8K^2(t+4)t} < \infty, \quad (5.12)$$

and so $X_t \in L^2$. If (Y, B) is another weak solution of (5.5) with $Y_0 \in L^2$, then for any $t \geq 0$,

$$\mathbb{E}[\sup_{0 \leq s \leq t} \|X_s - Y_s\|^2] \leq 2\mathbb{E}[\|X_0 - Y_0\|^2]e^{4K^2(t+4)t}. \quad (5.13)$$

In particular, the weak solution of (5.5) with L^2 -initial value satisfies pathwise uniqueness.

Proof. Fix $R > 0$. Let τ_R be the first $t \geq 0$ such that $\|X_t\| \geq R$. Let $\sigma_R(t, x) = \mathbf{1}_{\|x\| < R}\sigma(t, x)$ and $b_R(t, x) = \mathbf{1}_{\|x\| < R}b(t, x)$. Then X^{τ_R} is a solution of $\text{SDE}_{X_0}(\sigma_R, b_R)$, which means that $X_{t \wedge \tau_R} = X_0 + S_R(X_{\cdot \wedge \tau_R})$, where

$$S_R(X)_t := \int_0^t b_R(s, X_s) ds + \int_0^t \sigma_R(s, X_s) \circ dB_s, \quad t \geq 0.$$

Note that $X_{t \wedge \tau_R} \in L^2$ for all $t \geq 0$ because $\|X_{t \wedge \tau_R}\| \leq R \vee \|X_0\|$. Since σ_R and b_R satisfy (5.6), by Lemma 18.5 (i), (5.8) holds for S_R . Since $X_{t \wedge \tau_R} = X_0 + S_R(X_{\cdot \wedge \tau_R})_t$, we get

$$\mathbb{E}[\|X_{t \wedge \tau_R}\|^2] \leq 2\mathbb{E}[\|X_0\|^2] + 2\mathbb{E}[\|S_R(X_{\cdot \wedge \tau_R})_t\|^2] \leq 2\|X_0\|^2 + 8K^2(t+4)(t + \int_0^t \mathbb{E}[\|X_{s \wedge \tau_R}\|^2] ds).$$

Let $f_R(t) = \mathbb{E}[\|X_{t \wedge \tau_R}\|^2]$. Then f_R is continuous by Dominated convergence theorem. By the displayed formula, we have for any $T > 0$,

$$f_R(t) \leq 2\mathbb{E}[\|X_0\|^2] + 8K^2(T+4)T + 8K^2(T+4) \int_0^t f(s) ds, \quad 0 \leq t \leq T.$$

By Gronwall's lemma, $f_R(T) \leq (2\mathbb{E}[\|X_0\|^2] + 8K^2(T+4)T)e^{8K^2(T+4)T}$, $T \geq 0$. Since $X_{t \wedge \tau_R} \rightarrow X_t$ as $R \rightarrow \infty$, by Fatou's lemma, we get (5.12), so $X_t \in L^2$ for all $t \geq 0$.

Since $X_t = X_0 + S(X)_t$ and $Y_t = Y_0 + S(Y)_t$, we get

$$\begin{aligned} \mathbb{E}[\sup_{0 \leq s \leq t} \|X_s - Y_s\|^2] &\leq 2\mathbb{E}[\|X_0 - Y_0\|^2] + 2\mathbb{E}[\sup_{0 \leq s \leq t} \|S(X)_s - S(Y)_s\|^2] \\ &\leq 2\mathbb{E}[\|X_0 - Y_0\|^2] + 4K^2(T+4) \int_0^t \mathbb{E}[\|X_s - Y_s\|^2] ds, \quad 0 \leq t \leq T \end{aligned}$$

If we let $f(t) = \mathbb{E}[\sup_{0 \leq s \leq t} \|X_s - Y_s\|^2]$, then $f(t) \leq 2\mathbb{E}[\|X_0 - Y_0\|^2] + 4K^2(T+4) \int_0^t f(s) ds$. Then we get (5.13) using Gronwall's lemma.

If $X_0 = Y_0$, then (5.13) implies that a.s. $X = Y$. So we get the pathwise uniqueness. \square

Proof of Theorem 18.3. (i) For the existence, we use the well-known Picard iteration for ODE. Let the δ -dimensional Brownian motion B and a random vector $\zeta \in L^2$ be fixed such that $\zeta \perp B$. A solution X of SDE $_{\zeta}(\sigma, b)$ satisfies that $X = \zeta + S(X)$. Define a sequence of $\mathcal{F}^{B, \zeta}$ -semimartingales (X^n) such that $X^0 \equiv \zeta$, and for any $n \geq 0$, $X^{n+1} = \zeta + S(X^n)$. If the sequence converges to some process X in some sense, then we expect that $X = \zeta + S(X)$.

Fix $T > 0$. Let $C_1 = 4K^2(T+4)T(1 + \mathbb{E}[\|\zeta\|^2])$ and $C_2 = 2K^2(T+4)$. We will prove by induction that for any $n \geq 0$,

$$\mathbb{E}[\sup_{0 \leq s \leq t} \|X_s^{n+1} - X_s^n\|^2] \leq C_1 \frac{(C_2 t)^n}{n!}, \quad \forall 0 \leq t \leq T. \quad (5.14)$$

For $n = 0$, we have $X^1 - X^0 = S(X^0) = S(\zeta)$. By (5.8) we get (5.14) for $n = 0$. Assume that (5.14) holds for $n - 1$. Since $X^{n+1} - X^n = S(X^n) - S(X^{n-1})$, by (5.9),

$$\mathbb{E}[\sup_{0 \leq s \leq t} \|X_s^{n+1} - X_s^n\|^2] \leq C_2 \int_0^t \mathbb{E}[\|X_s^n - X_s^{n-1}\|^2] ds \leq C_2 \int_0^t C_1 \frac{(C_2 s)^{n-1}}{(n-1)!} ds = C_1 \frac{(C_2 t)^n}{n!}.$$

Thus, (5.14) holds for all $n \geq 0$.

By Chebyshev's inequality,

$$\mathbb{P}[\sup_{0 \leq s \leq T} \|X_s^{n+1} - X_s^n\|^2 > 4^{-n}] \leq C_1 \frac{(4C_2 T)^n}{n!}, \quad n \geq 0.$$

Since $\sum_{n=0}^{\infty} C_1 \frac{(4C_2 T)^n}{n!} < \infty$, by Borel-Cantelli Lemma,

$$\mathbb{P}[\text{for infinitely many } n, \sup_{0 \leq s \leq T} \|X_s^{n+1} - X_s^n\| > 2^{-n}] = 0.$$

Thus, a.s. there is a random number N such that for $n \geq N$, $\sup_{0 \leq s \leq T} \|X_s^{n+1} - X_s^n\| \leq 2^{-n}$. Since $\sum_{n=0}^{\infty} 2^{-n} < \infty$, this implies that a.s. $X^n|_{[0, T]}$ converges as $n \rightarrow \infty$ uniformly on $[0, T]$. The limit is a continuous process on $[0, T]$. Then we can conclude that a.s. for every $N \in \mathbb{N}$, $\lim_{n \rightarrow \infty} X^n|_{[0, N]}$ converges uniformly on $[0, N]$ to some $X^{(N)}$. If $N < N'$, then we have a.s.

$X^{(N)} = X^{(N')}|_{[0,N]}$. Thus, outside a null event, we have $X^{(N)} = X^{(N')}|_{[0,N]}$ for any $N < N' \in \mathbb{N}$. So there is a continuous process X on $[0, \infty)$ such that a.s. $X^{(n)}$ converges to X uniformly on $[0, T]$ for any $T > 0$. Since each X^n is $\mathcal{F}^{B, \zeta}$ -adapted, so is X . By (5.14), for every $t \geq 0$, X_t^n converges in L^2 . So $X_t \in L^2$ is the L^2 -limit of (X_t^n) . By (5.14) and the fact that $\rho(X, Y) := \mathbb{E}[\sup_{0 \leq s \leq T} \|X_s - Y_s\|^2]^{1/2}$ is a metric, we have

$$\mathbb{E}[\sup_{0 \leq s \leq T} \|X_s - X_s^n\|^2]^{1/2} \leq \sum_{m=n}^{\infty} \left(C_1 \frac{(C_2 t)^m}{m!} \right)^{1/2}. \quad (5.15)$$

We now prove that a.s. $\zeta + S(X) = X$. First, we have

$$\mathbb{E}[\sup_{0 \leq s \leq T} \|X_s - X_s^{n+1}\|^2] \leq \left(\sum_{m=n+1}^{\infty} \left(C_1 \frac{(C_2 t)^m}{m!} \right)^{1/2} \right)^2$$

which tends to 0 as $n \rightarrow \infty$. Since $X^{n+1} = \zeta + S(X^n)$, by (5.9) and (5.15),

$$\begin{aligned} \mathbb{E}[\sup_{0 \leq s \leq T} \|\zeta + S(X)_s - X_s^{n+1}\|^2] &= \mathbb{E}[\sup_{0 \leq s \leq T} \|S(X)_s - S(X^n)_s\|^2] \\ &\leq 2K^2(T+4) \int_0^T \mathbb{E}[\|X_s - X_s^n\|^2] ds \leq 2K^2(T+4) \int_0^T \left(\sum_{m=n}^{\infty} \left(C_1 \frac{(C_2 s)^m}{m!} \right)^{1/2} \right)^2 ds, \end{aligned}$$

which also tends to 0 as $n \rightarrow \infty$. Thus, by sending $n \rightarrow \infty$, we get $\mathbb{E}[\sup_{0 \leq s \leq T} \|\zeta + S(X)_s - X_s\|^2] = 0$, i.e., a.s. $\zeta + S(X) = X$ on $[0, T]$. Since this holds for any $T > 0$, we get a.s. $\zeta + S(X) = X$. So X is a solution of the SDE $_{\zeta}(\sigma, b)$. Since X is $\mathcal{F}^{B, \zeta}$ -adapted, it is a strong solution.

(ii) The uniqueness of the strong solution with initial value in L^2 follows from the pathwise uniqueness of the weak solution since every strong solution is also a weak solution. By the pathwise uniqueness of the weak solution and the existence of the strong solution, a weak solution with initial value in L^2 should also be a strong solution.

(iii) We will apply Theorem 2.23. It is about the existence of a continuous version of a process with given finite dimensional distribution. We used it to prove the existence of continuous Brownian motion. It says that if $X_t, t \in \mathbb{R}^d$, is a process taking values in a complete metric space (S, ρ) , and there exists $a, b, C > 0$ such that for any $s, t \in \mathbb{R}^d$,

$$\mathbb{E}[\rho(X_s, X_t)^a] \leq C \|s - t\|^{d+b},$$

then X has a continuous version. We apply Theorem 2.23 here to X_t^x with index set $(x, t) \in \mathbb{R}^d \times \mathbb{R}_+$. It suffices to prove the estimate for a bounded region, i.e., for any $R, T > 0$, there exists $a, b, C > 0$ such that

$$\mathbb{E}[\|X_s^x - X_t^y\|^a] \leq C(\|x - y\| + |s - t|)^{d+b}, \quad \|x\|, \|y\| \leq R, \quad s, t \in [0, T]. \quad (5.16)$$

When this holds, we get the existence of a continuous version of X_t^x on $\{x \in \mathbb{R}^d : \|x\| \leq R\} \times [0, T]$. Letting $R, T \rightarrow \infty$, we then get the continuous version of X_t^x on $\mathbb{R}^d \times \mathbb{R}_+$. The

proof of (5.16) requires BDG inequality, which we have skipped. So we now skip the technical part of the proof of (iii).

(iv) Suppose now X_t^x , $x \in \mathbb{R}^d$, $t \in \mathbb{R}_+$, is jointly continuous in x and t , and for every x , X^x is the strong solution of $\text{SDE}_x(\sigma, b)$. Let $\zeta \in L^2$ be independent of B . We want to show that the solution X^ζ of $\text{SDE}_\zeta(\sigma, b)$ is a.s. equal to $X^x|_{x=\zeta}$. First, suppose ζ takes values in a countable set Σ . Recall that when we constructed the solution X^ζ by successively defined a sequence of processes X^n , which a.s. converges to X^ζ locally uniformly. We now write them as $X^{\zeta, n}$ to emphasize the dependence on ζ . Similarly, for any $x \in \mathbb{R}^d$, we have a sequence $X^{x, n}$, which a.s. converges to the solution X^x with initial value x . We can prove by induction that for any n and any $x \in \Sigma$, on the event $\{\zeta = x\}$, a.s. $X^{\zeta, n} = X^{x, n}$. Thus, as their limits, we have a.s. $X^\zeta = X^x$ on the event $\{\zeta = x\}$. Since ζ takes values in Σ , and Σ is countable, we then get a.s. $X^\zeta = X^x|_{x=\zeta}$. For a general ζ , we may define a sequence (ζ_m) , $m \in \mathbb{N}$, such that when ζ lies in the cube $\prod_{j=1}^d [\frac{k_j}{2^m}, \frac{k_j+1}{2^m})$ for some $(k_1, \dots, k_d) \in \mathbb{Z}^d$, ζ_m takes values $(\frac{k_1}{2^m}, \dots, \frac{k_d}{2^m})$. Then each ζ_m takes countably many values, belongs to L^2 and is independent of B , and as $m \rightarrow \infty$, $\zeta_m \rightarrow \zeta$ pointwise and in L^2 . Now for each m , we have a.s. $X^{\zeta_m} = X^x|_{x=\zeta_m}$. Fix $t \geq 0$. By the continuity of X_t^x in x , we know that $X_t^x|_{x=\zeta_m}$ converges to $X_t^x|_{x=\zeta}$ as $m \rightarrow \infty$. By a corollary we know that $X_t^{\zeta_m} \rightarrow X_t^\zeta$ in L^2 , which implies that $(X_t^{\zeta_m})$ has a subsequence, which converges a.s. to X_t^ζ . Thus, we get a.s. $X_t^\zeta = X_t^x|_{x=\zeta}$. Since both sides are continuous in t , we then get a.s. $X^\zeta = X^x|_{x=\zeta}$. \square

Remark . The condition of Theorem 18.3 is somewhat too strong because σ and b have to be defined on $\mathbb{R}_+ \times \mathbb{R}^d$, and the Lipschitz condition should hold throughout. However, we may use localization and Lipschitz extension to weaken the assumptions. We may assume that σ and b are defined on $\mathbb{R}_+ \times U$, where U is a domain in \mathbb{R}^d containing the initial value x_0 , and σ and b are locally bounded and locally Lipschitz continuous in x . The weaker assumption is satisfied if σ and b are continuous and continuously differentiable in x . In that case, we may not have a solution defined on $[0, \infty)$. But for any initial value $x_0 \in U$, the solution X^{x_0} exists up to some positive random lifetime T . If Y_t , $0 \leq t < S$, also solves the initial value problem, then a.s. Y is a restriction of X^{x_0} . Moreover, if $T < \infty$, then as $t \uparrow T$, either $\|X_t^{x_0}\| \rightarrow \infty$ or $X_t^{x_0} \rightarrow \partial U$.

5.2 Weak solution and martingale problem

Suppose (X, B) defined on the filtered space $(\Omega, \mathcal{F}, \mathbb{P})$ is a weak solution of

$$dX_t = \sigma(t, X_t) \circ dB_t + b(t, X_t)dt$$

for two measurable functions $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times \delta}$ and $b : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$. We assume that σ and b are locally bounded, i.e., for any $R > 0$ and $T > 0$, $\|\sigma\|$ and $\|b\|$ are bounded on $[0, T] \times \{x \in \mathbb{R}^d : \|x\| \leq R\}$. Let $C_0^2(\mathbb{R}^d, \mathbb{R})$ denote the space of C^2 functions on \mathbb{R}^d with compact supports. For $f \in C_0^2(\mathbb{R}^d, \mathbb{R})$, by Itô's formula, we have

$$df(X_t) = \sum_{i=1}^d \partial_i f(X_t) dX_t^i + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_j \partial_k f(X_t) d[X^j, X^k]_t$$

$$\begin{aligned}
&= \sum_{i=1}^d \sum_{j=1}^{\delta} \partial_i f(X_t) \sigma_j^i(t, X_t) dB_t^j + \sum_{i=1}^d \partial_i f(X_t) b^i(t, X_t) dt \\
&\quad + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \partial_j \partial_k f(X_t) \sum_{i=1}^d \sigma_i^j(t, X_t) \sigma_i^k(t, X_t) dt.
\end{aligned}$$

Let $a^{j,k} = \sum_{i=1}^{\delta} \sigma_i^j \sigma_i^k$, i.e., as matrixes, $a = \sigma \sigma'$. Define

$$M_t^f := f(X_t) - f(X_0) - \int_0^t A_s f(X_s) ds, \quad t \geq 0, \quad (5.17)$$

where

$$A_t f(x) = \sum_{i=1}^d b^i(t, x) \partial_i f(x) + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d a^{j,k}(t, x) \partial_j \partial_k f(x).$$

Then M^f is an \mathcal{F} -local martingale. We will see that M^f is in fact a true martingale.

Since $f \in C_0^2(\mathbb{R}^d, \mathbb{R})$, we know that f , $\partial_i f$, and $\partial_j \partial_k f$ are all bounded, and there are $R > 0$ such that $\partial_i f$ and $\partial_j \partial_k f$ vanish outside $\{x \in \mathbb{R}^d : |x| < R\}$. Since b and $a = \sigma \sigma'$ are locally bounded, we can then conclude that for any $T > 0$, $t \mapsto \sup_{x \in \mathbb{R}^d} |A_t f(x)|$ is bounded on $[0, T]$. Thus, M_t^f is bounded on $[0, T]$ for any $T \in \mathbb{R}_+$. So M^f is a true \mathcal{F} -martingale. In fact, we do not need to mention \mathcal{F} because M^f is a martingale w.r.t. the natural filtration generated by itself.

We say that X solves the martingale problem (a, b) if for any $f \in C_0^2(\mathbb{R}^d, \mathbb{R})$ with a compact support, the M^f defined by (5.17) is a martingale (w.r.t. the natural filtration generated by itself). Thus, if (X, B) is a weak solution of SDE (σ, b) , then X solves the martingale problem (a, b) with $a = \sigma \sigma'$.

Theorem 18.7. *The SDE (σ, b) has a weak solution (X, B) if and only if X solves the martingale problem for (a, b) with $a = \sigma \sigma'$.*

Proof. We have proved the “only if” part. For the “if” part, suppose X is such that for any $f \in C^2(\mathbb{R}, \mathbb{R})$ with compact support, M^f is a martingale.

We use a localization argument. Fix $i \in \{1, \dots, d\}$. For any $R > 0$, there is a function $f_R \in C^2(\mathbb{R}, \mathbb{R})$ with compact support such that $f_R^i(x) = x_i$ for $|x| \leq R$. Note that $f_R''(x) = 0$ for $|x| \leq R$. Using $f = f_R^i$, we find that

$$M_t := f_R^i(X_t) - f_R^i(X_0) - \sum_j \int_0^t \partial_j f_R^i(X_s) b^j(s, X_s^j) ds - \frac{1}{2} \sum_{j,k} \int_0^t a^{j,k} \partial_j \partial_k f_R^i(X_s) [X_s^j, X_s^k] ds$$

is a martingale. Stopping this process at the time τ_R , which is the first time that X exits $\{|x| \leq R\}$, we still get a martingale. Since $f_R^i(x) = x_i$ for $|x| \leq R$, $\partial_j x_i = \delta_{i,j}$, and $\partial_j \partial_k x_k = 0$, the stopped process M^{τ_R} can be expressed as

$$M_t^{\tau_R} = X_{t \wedge \tau_R}^i - X_0^i - \int_0^{t \wedge \tau_R} b^i(s, X_s) ds.$$

Letting $n \rightarrow \infty$, we conclude that

$$M_t^i := X_t^i - X_0^i - \int_0^t b^i(s, X_s) ds$$

is a local martingale. A similar localization argument (with $f_R^{j,k}$ in place of f_R^i which satisfies $f_R^{j,k}(x) = x_j x_k$ for $|x| \leq R$) shows that for any j, k ,

$$M_t^{j,k} := X_t^j X_t^k - X_0^j X_0^k - \int_0^t b^j(s, X_s) X_s^k ds - \int_0^t b^k(s, X_s) X_s^j ds - \int_0^t a^{j,k}(s, X_s) ds$$

is a local martingale. By the definition of M^i and $M^{j,k}$, we have, for any $1 \leq i, j, k \leq d$,

$$dX_t^i = dM_t^i + b^i(t, X_t) dt;$$

$$d(X_t^j X_t^k) = dM_t^{j,k} + b^j(t, X_t) X_t^k dt + b^k(t, X_t) X_t^j dt + a^{j,k}(t, X_t) dt.$$

Applying the product formula to the first group of SDE, we get

$$d(X_t^j X_t^k) = X_t^j dM_t^k + X_t^k dM_t^j + b^j(t, X_t) X_t^k dt + b^k(t, X_t) X_t^j dt + d[M^j, M^k]_t.$$

Compared it with the second equation, we get

$$d[M^j, M^k]_t = a^{j,k}(t, X_t) dt = \sum_{i=1}^d \sigma_i^j(t, X_t) \sigma_i^k(t, X_t) dt.$$

Then we use Theorem 16.12, which can be viewed as a higher dimensional counterpart of Levy's characterization theorem. We are not going to give a proof of this theorem.

Theorem 16.12. *Let $M = (M^1, \dots, M^d)$ be a d -dimensional vector martingale with $M_0 = 0$ such that $d[M^j, M^k]_t = \sum_i \sigma_i^j \sigma_i^k dt$ for some progressive processes σ_i^j , $1 \leq j \leq d$, $1 \leq i \leq \delta$. Then in an extended filtered space, there exists an δ -dimensional Brownian motion $B = (B^1, \dots, B^\delta)$ such that $dM_t^j = \sum_{i=1}^\delta \sigma_i^j dB_t^i$, $1 \leq j \leq d$.*

Applying this theorem we get a δ -dimensional Brownian motion $B = (B^1, \dots, B^\delta)$ (in an enlarged filtered space) such that $M_t^j = \sum_{i=1}^\delta \sigma_i^j(t, X_t) dB_t^i$ for all $1 \leq j \leq d$. Then we get

$$dX_t^j = dM_t^j + b^j(t, X_t) dt = \sum_{i=1}^\delta \sigma_i^j(t, X_t) dB_t^i + b^j(t, X_t) dt, \quad 1 \leq j \leq d.$$

So X is a weak solution of the SDE $dX_t = \sigma(t, X_t) \circ dB_t + b(t, X_t) dt$. □

5.3 Diffusion and Markov property

We will focus on the SDE of the kind

$$dX_t = \sigma(X_t) \circ dB_t + b(X_t)dt, \quad X_0 = x, \quad (5.18)$$

where $\sigma \in C(\mathbb{R}^d, \mathbb{R}^{d \times \delta})$ and $b \in C(\mathbb{R}^d, \mathbb{R}^d)$. Note that σ and b do not depend on t . Suppose that they satisfy the conditions in Theorem 18.3, i.e., there is a constant $K > 0$ such that for any $x, y \in \mathbb{R}^d$,

$$\begin{aligned} \|\sigma(x)\| + \|b(x)\| &\leq K(1 + \|x\|); \\ \|\sigma(x) - \sigma(y)\| + \|b(x) - b(y)\| &\leq K\|x - y\|. \end{aligned}$$

The (unique) strong solution X^x of the SDE with initial value x is called a diffusion process. Let P^x denote the law of X^x . Since we can find a version of X^x for each $x \in \mathbb{R}^d$ such that X_t^x is jointly continuous in t and x , we see that $x \rightarrow P^x$ is a probability kernel.

Theorem . *The strong solution X of (5.18) with any initial value $\zeta \in L^2$ is a time-homogeneous Markov process w.r.t. the complete filtration \mathcal{F}^B generated by B , and for any fixed $t_0 \geq 0$, $\text{Law}[X_{t_0+} | \mathcal{F}_{t_0}^B] = P^{X_{t_0}}$.*

Proof. Fix $t_0 \geq 0$. By Theorem 18.3, $X_{t_0} \in L^2$. Define $B_t^{t_0+} = B_{t_0+t} - B_{t_0}$ and $\mathcal{F}_t^{B, t_0+} = \mathcal{F}_{t_0+t}^B$, $t \geq 0$. Then B^{t_0+} is an \mathcal{F}^{B, t_0+} -Brownian motion independent of $\mathcal{F}_{t_0}^B$. Let $X_t^{t_0+} = X_{t_0+t}$, $t \geq 0$. Since X satisfies (5.18), we have for any $t \geq 0$,

$$X_t^{t_0+} - X_0^{t_0+} = \int_{t_0}^{t_0+t} \sigma(X_s) \circ dB_s + \int_{t_0}^{t_0+t} b(X_s)ds = \int_0^t \sigma(X_s^{t_0+}) \circ dB_s^{t_0+} + \int_0^t b(X_s^{t_0+})ds.$$

Thus, (X^{t_0+}, B^{t_0+}) w.r.t. the filtration \mathcal{F}^{B, t_0+} is a weak solution of (5.18) with initial value X_{t_0} , which is in L^2 . By Theorem 18.3, this weak solution is a strong solution. Moreover, by the same theorem there is a family of \mathbb{R}^d -valued random vectors $Z(x, t)$, $x \in \mathbb{R}^d$, $t \in \mathbb{R}_+$, such that

- (i) For each $x \in \mathbb{R}^d$, $Z(x, t)$, $t \geq 0$, is a strong solution of (5.18) with initial value x and with B^{t_0+} in place of B .
- (ii) The map $\mathbb{R}^d \times \mathbb{R}_+ \ni (x, t) \mapsto Z_t^x$ is continuous.
- (iii) Almost surely $X^{t_0+} = Z(X_{t_0}, \cdot)$.

For each $x \in \mathbb{R}^d$, we have $\text{Law}[Z(x, \cdot) | \mathcal{F}_{t_0}^B] = \text{Law}[Z(x, \cdot)] = \mathbb{P}^x$ since $Z(x, \cdot)$ is measurable w.r.t. B^{t_0+} , and $B^{t_0+} \perp \mathcal{F}_{t_0}^B$ (and so $Z(x, \cdot) \perp \mathcal{F}_{t_0}^B$). From the facts that a.s. $X^{t_0+} = Z(X_{t_0}, \cdot)$ and that X_{t_0} is $\mathcal{F}_{t_0}^B$ -measurable, we then know that $\text{Law}[X_{t_0+} | \mathcal{F}_{t_0}^B] = P^{X_{t_0}}$. \square

6 Connections with Partial Differential Equations

In this section we study diffusion processes from the perspective of their Markov properties, and connect them with the partial differential equations. The material of this section is chosen from multiple references.

6.1 Transition kernel and operator

We now briefly review the notation of probability kernels and Markov process needed here. A probability kernel from \mathbb{R}^d to \mathbb{R}^d is map μ defined on the product space $\mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d)$ such that for any $x \in \mathbb{R}^d$, $\mu(x, \cdot)$ is a probability measure on \mathbb{R}^d , and for any Borel measurable set $A \subset \mathbb{R}^d$, $\mu(\cdot, A)$ is a Borel measurable function on \mathbb{R}^d . From now on we simply call such μ a kernel in \mathbb{R}^d . If μ and ν are two kernels in \mathbb{R}^d , then we may define a new kernel $\mu\nu$ in \mathbb{R}^d by

$$\mu\nu(x, A) = \int \mu(x, dy)\nu(y, A), \quad x \in \mathbb{R}^d, \quad A \in \mathcal{B}(\mathbb{R}^d).$$

The formula means that we integrate the measurable function $\nu(\cdot, A)$ on \mathbb{R}^d against the measure $\mu(x, \cdot)$. We have the associative law: $(\mu\nu)\lambda = \mu(\nu\lambda)$, but not the commutative law.

An \mathcal{F} -adapted \mathbb{R}^d -valued process X_t , $t \in \mathbb{R}_+$, is called \mathcal{F} -Markov if there is a family of probability kernels $\mu_{s,t}$, $0 \leq s \leq t < \infty$, from \mathbb{R}^d to \mathbb{R}^d , such that for any $0 \leq s \leq t$, a.s.

$$\text{Law}[X_t | \mathcal{F}_s] = \text{Law}[X_t | X_s] = \mu_{s,t}(X_s, \cdot).$$

This means that if $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is bounded and measurable, then

$$\mathbb{E}[f(X_t) | \mathcal{F}_s] = \int f(y)\mu_{s,t}(X_s, dy).$$

The family $\mu_{s,t}$ must satisfy the Chapman-Kolmogorov relation:

$$\mu_{r,s}\mu_{s,t} = \mu_{r,t} \quad 0 \leq r \leq s \leq t.$$

If $\mu_{s,t}$ depends only on $t-s$, i.e., there is a one parameter family μ_t , $t \in \mathbb{R}_+$, such that $\mu_{s,t} = \mu_{t-s}$, then X is called time-homogeneous, and the Chapman-Kolmogorov relation becomes

$$\mu_s\mu_t = \mu_{s+t}, \quad s, t \geq 0.$$

Recall that the diffusion processes are time-homogeneous Markov processes.

For every kernel μ in \mathbb{R}^d , we associate it with an operator T_μ defined by

$$T_\mu f(x) = \int f(y)\mu(x, dy),$$

where f is a bounded measurable function from \mathbb{R}^d to \mathbb{R} . Then $|T_\mu f|$ is also bounded. If $f = \mathbf{1}_A$ for some $A \in \mathcal{B}(\mathbb{R}^d)$, then $T_\mu f = \mu(\cdot, A)$ is measurable. Thus, for any measurable simple function f , $T_\mu f$ is measurable. By approximation, $T_\mu f$ is measurable for any bounded measurable function f .

If μ and ν are two kernels in \mathbb{R}^d , then we have the following equality

$$T_\mu \circ T_\nu = T_{\mu\nu}$$

because for any f ,

$$T_\mu \circ T_\nu f(x) = \int T_\nu f(y) \mu(x, dy) = \int \int f(z) \nu(y, dz) \mu(x, dy) = \int f(z) \mu \nu(x, dz).$$

Suppose (μ_t) is a family of probability kernels in \mathbb{R}^d associated with a time-homogeneous Markov process X in \mathbb{R}^d . Let $T_t = T_{\mu_t}$. Then from Chapman-Kolmogorov relation, we get

$$T_s \circ T_t = T_{s+t}, \quad s, t \geq 0.$$

Thus, $\{T_t : t \geq 0\}$ forms a semigroup of operators on the space of bounded measurable functions on \mathbb{R}^d . If X^x is the Markov process starting from x , then for any $t \geq 0$, the law of X_t^x is $\mu_t(x, \cdot)$. Then the integral $\int f(y) \mu_t(x, dy)$ is exactly the expectation of $f(X_t)$. So we get an expression of T_t in terms of X^x :

$$T_t f(x) = \mathbb{E}[f(X_t^x)] = \mathbb{E}^x[f(X_t)],$$

where \mathbb{E}^x means the expectation w.r.t the law of the Markov process started from x .

6.2 Boundary value problems

Let D be a bounded domain in \mathbb{R}^d . Let f be a continuous function on ∂D . Consider the following Laplace's equation

$$\begin{cases} \frac{1}{2} \Delta u(x) = 0, & x \in D; \\ u(x) = f(x), & x \in \partial D. \end{cases}$$

Here u is continuous on \bar{D} and C^2 on D , and $\Delta u = \sum_{j=1}^d \partial_j^2 u$. The factor $\frac{1}{2}$ is to make the connection with Itô's formula.

Suppose u is a solution of the above PDE. Let $x \in D$. Let B^x be a Brownian motion in \mathbb{R}^d started from x . Let τ_D be the first time that B^x exits D . By Itô's formula,

$$X_t := u(B_{t \wedge \tau_D}^x), \quad t \geq 0,$$

is a bounded martingale. Then we get

$$u(x) = X_0 = \mathbb{E}[X_{\tau_D}] = \mathbb{E}[u(B_{\tau_D}^x)] = \mathbb{E}^x[f(B_{\tau_D})].$$

So we get a formula of u in terms of f . On the other hand, if we define u by the above formula, then it can be shown that it is the solution of the Laplace's equation. But it takes some work.

Next, we consider a modified PDE problem:

$$\begin{cases} \frac{1}{2} \Delta u(x) = g(x), & x \in D; \\ u(x) = f(x), & x \in \partial D. \end{cases}$$

The f is the same as before. The g is a continuous and bounded function on D . If u is a solution, by Itô's formula, the

$$X_t := u(B_{t \wedge \tau_D}^x) - \int_0^{t \wedge \tau_D} g(B_s^x) ds, \quad t \geq 0,$$

is a local martingale. One can actually show that it is a uniformly integrable true martingale. So we get

$$u(x) = X_0 = \mathbb{E}[X_{\tau_D}] = \mathbb{E}[f(B_{\tau_D}^x) - \int_0^{\tau_D} g(B_s^x) ds].$$

On the other hand, if we define u using the above formula, then it turns out that it is the solution of the PDE.

Thirdly, we consider the heat equation.

$$\begin{cases} \partial_t u(t, x) = \frac{1}{2} \Delta u(t, x), & t \in (0, \infty), x \in D; \\ u(t, x) = f(x), & t \in [0, \infty), x \in \partial D; \\ u(0, x) = u_0(x), & x \in \bar{D}. \end{cases}$$

Here u_0 is continuous on \bar{D} , f is continuous on ∂D , and $u_0|_{\partial D} = f$. The solution $u(t, x)$ is continuous on $[0, \infty) \times \bar{D}$, and is $C^{1,2}$ on $(0, \infty) \times D$.

Suppose u is a solution of the heat equation. Fix $t_0 > 0$ and $x_0 \in D$. Then the process

$$X_t := u(t_0 - t \wedge \tau_D, B_{t \wedge \tau_D}^{x_0}), \quad 0 \leq t \leq t_0,$$

is a true martingale. So we get

$$u(t_0, x_0) = X_0 = \mathbb{E}[X_{t_0}] = \mathbb{E}^{x_0}[\mathbf{1}_{\{\tau_D \leq t_0\}} f(B_{\tau_D})] + \mathbb{E}^{x_0}[\mathbf{1}_{\{\tau_D > t_0\}} u_0(B_{t_0})].$$

This formula does give the solution of the PDE.

A more general heat equation is

$$\begin{cases} \partial_t u(t, x) = \frac{1}{2} \Delta u(t, x) + g(x), & t \in (0, \infty), x \in D; \\ u(t, x) = f(x), & t \in [0, \infty), x \in \partial D; \\ u(0, x) = u_0(x), & x \in \bar{D}. \end{cases}$$

If u is a solution, we construct a martingale

$$X_t = u(t_0 - t \wedge \tau_D, B_{t \wedge \tau_D}^{x_0}) - \int_0^{t \wedge \tau_D} g(B_s^{x_0}) ds, \quad 0 \leq t \leq t_0.$$

Using $u(t_0, x_0) = X_0 = \mathbb{E}[X_{t_0}]$, we get a formula of the solution u .

6.3 Infinitesimal generator and parabolic PDE

Suppose that X is a diffusion process as strong solutions of (5.18), where σ and b satisfy the conditions in Theorem 18.3. For a fixed bounded continuous function f , we define

$$u(t, x) = T_t f(x) = \mathbb{E}[f(X_t^x)] = \mathbb{E}^x[f(X_t)].$$

Note that $u(0, x) = f(x)$. Since f is bounded, u is bounded on $\mathbb{R}_+ \times \mathbb{R}^d$. Since X_t^x is jointly continuous in x and t , by dominated convergence theorem, $u(t, x)$ is jointly continuous in t and x .

The infinitesimal generator A associated to the SDE is

$$Af(x) = \lim_{t \downarrow 0} \frac{(T_t f)(x) - f(x)}{t} = \lim_{t \downarrow 0} \frac{1}{t} (u(t, x) - u(0, x)), \quad x \in \mathbb{R}^d.$$

Let \mathcal{D}_A denote the set of all bounded continuous functions f on \mathbb{R}^d such that the limit exists for every $x \in \mathbb{R}^d$. For $f \in \mathcal{D}_A$, Af is a well-defined measurable function on \mathbb{R}^d .

Let $a = \sigma\sigma'$. Define the second order differential operator \mathcal{A} by

$$\mathcal{A}f(x) = \frac{1}{2} \sum_{j,k=1}^d a^{j,k}(x) \partial_j \partial_k f(x) + \sum_{i=1}^d b^i(x) \partial_i f(x)$$

Example . Suppose $b \equiv 0$ and $\sigma_j^i = \delta_{i,j}$, i.e., σ is the identity matrix, then the same is true for a . Then $\mathcal{A}f = \frac{1}{2} \Delta f$. We will also deal with more complicated cases.

Theorem . *The space $C_0^2(\mathbb{R}^d)$ (C^2 functions with compact support) belongs to \mathcal{D}_A , and for any $f \in C_0^2(\mathbb{R}^d)$,*

$$Af(x) = \mathcal{A}f(x). \tag{6.1}$$

Proof. Let $f \in C_0^2(\mathbb{R}^d)$. Since X^x solves the martingale problem (a, b) ,

$$M_t^x := f(X_t^x) - f(x) - \int_0^t \mathcal{A}f(X_s^x) ds$$

is a martingale. From $\mathbb{E}[M_t^x] = \mathbb{E}[M_0^x] = 0$, we get

$$\frac{1}{t} (\mathbb{E}[f(X_t^x)] - f(x)) = \mathbb{E} \left[\frac{1}{t} \int_0^t \mathcal{A}f(X_s^x) ds \right].$$

The quantity inside the square bracket on the RHS is uniformly bounded because $\partial_i f$ and $\partial_j \partial_k f$ have compact supports, and b^i and $a^{j,k}$ are locally bounded. Letting $t \downarrow 0$, by dominated convergence theorem, the RHS tends to $\mathcal{A}f(x)$. So $f \in \mathcal{D}_A$, and (6.1) holds. \square

From the expression of M_t^x and its martingale property ($0 = M_0 = \mathbb{E}[M_\tau]$), we get the Dynkin's formula

Theorem (Dynkin's formula). *For any bounded stopping time τ , if $f \in C_0^2(\mathbb{R}^d)$, then*

$$\mathbb{E}[f(X_\tau^x)] = f(x) + \mathbb{E} \left[\int_0^\tau \mathcal{A}f(X_s^x) ds \right].$$

Remark . If τ is the less than τ_U , which is the first time that X exits a bounded domain U , then Dynkin's formula holds for any $f \in C^2(\mathbb{R}^d)$ with Af replaced by $\mathcal{A}f$. This is because we may find another function $f_0 \in C_0^2(\mathbb{R}^d)$ such that $f = f_0$ on U .

Recall that $u(t, x) = T_t f(x) = \mathbb{E}[f(X_t^x)]$. When τ in the theorem is a deterministic time t , Dynkin's formula becomes

$$u(t, x) = f(x) + \int_0^t \mathbb{E}[Af(X_s^x)]ds = u(0, x) + \int_0^t u^A(s, x)ds,$$

where $u^A(t, x)$ is defined in the same way as $u(t, x)$ except with Af in place of f . Since $Af = \mathcal{A}f \in C_0(\mathbb{R}^d)$, $u^A(t, x)$ is continuous in t and x . So u is differentiable in t , and

$$\partial_t u(t, x) = \partial_t T_t f(x) = u^A(t, x) = T_t \circ Af(x), \quad f \in C_0^2(\mathbb{R}^d). \quad (6.2)$$

Theorem (Kolmogorov's backward equation). *For $f \in C_0^2(\mathbb{R}^d)$, $u(t, \cdot) \in \mathcal{D}_A$ for each $t > 0$; and u satisfies the following equation with initial value*

$$\begin{cases} \partial_t u(t, x) = Au(t, x), & t > 0, x \in \mathbb{R}^d; \\ u(0, x) = f(x), & x \in \mathbb{R}^d. \end{cases} \quad (6.3)$$

Here the operator A acts on the second variable x .

Proof. Clearly, $u(0, x) = f(x)$, $x \in \mathbb{R}^d$. Fix $t > 0$. Let $g(x) = u(t, x)$. Since u is differentiable in t , using the Markov property of X we get for small r ,

$$\begin{aligned} \frac{1}{r}(\mathbb{E}^x[g(X_r)] - g(x)) &= \frac{1}{r}(\mathbb{E}^x[\mathbb{E}^{X_r}[f(X_t)]] - \mathbb{E}^x[f(X_t)]) \\ &= \frac{1}{r}(\mathbb{E}^x[\mathbb{E}^x[f(X_{r+t})|\mathcal{F}_r]] - \mathbb{E}^x[f(X_t)]) \\ &= \frac{1}{r}\mathbb{E}^x[f(X_{r+t}) - f(X_t)] = \frac{1}{r}(u(r+t, x) - u(t, x)) \rightarrow \partial_t u, \quad r \downarrow 0. \end{aligned}$$

Hence

$$Au = \lim_{r \downarrow 0} \frac{1}{r}(\mathbb{E}^x[g(X_r)] - g(x)), \quad \text{exists and equals } \partial_t u.$$

□

We have proved that $A = \mathcal{A}$ on $C_0^2(\mathbb{R}^d)$. Here we can not say that $Au = \mathcal{A}u$ because u in general does not belong to C_0^2 . If we know that $Au = \mathcal{A}u$. Then Equation (6.3) becomes the following second-order parabolic PDE

$$\begin{cases} \partial_t u(t, x) = \mathcal{A}u(t, x), & t > 0, x \in \mathbb{R}^d; \\ u(0, x) = f(x), & x \in \mathbb{R}^d. \end{cases} \quad (6.4)$$

The PDE is parabolic since the coefficient matrix $(a^{j,k})_{1 \leq j, k \leq d}$ is semi-positive definite (because $a = \sigma\sigma'$). At this moment we can not immediately say that $u(t, x) := \mathbb{E}^x[f(X_t)]$ gives the solution of (6.4) because we have not shown that $Au = \mathcal{A}u$. However, we have the following theorem, which says that such u is the only bounded candidate of the solution. We let $C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ denote the space of bounded functions defined on $\mathbb{R}_+ \times \mathbb{R}^d$, which are once continuously differentiable in the first variable, and twice continuously differentiable in all other variables.

Theorem . If $u \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ is a solution of the PDE (6.4), then $u(t, x) = \mathbb{E}^x[f(X_t)]$.

Proof. Fix $t_0 > 0$ and $x \in \mathbb{R}^d$. We want to show that $u(t_0, x) = \mathbb{E}^x[f(X_{t_0})]$. Let X be a strong solution of the SDE(σ, b) started from x . Define

$$M_t = u(t_0 - t, X_t^x), \quad 0 \leq t \leq t_0.$$

Then M is a semimartingale, and by Itô's formula, we calculate

$$dM_t = -\partial_t u(t_0 - t, X_t) + \sum_{i=1}^d \partial_i u(t_0 - t, X_t) dX_t^i + \frac{1}{2} \sum_{j,k=1}^d \partial_j \partial_k u(t_0 - t, X_t) d[X^j, X^k]_t.$$

Since X is a strong solution of SDE(σ, b), we have

$$dX_t^i = \sum_{j=1}^{\delta} \sigma_j^i(X_t) dB_t^j + b^i(X_t) dt,$$

and

$$d[X^j, X^k]_t = \sum_{i=1}^{\delta} \sigma_i^j(X_t) \sigma_i^k(X_t) dt = a^{j,k}(X_t) dt.$$

Thus,

$$dM_t = \sum_{i=1}^d \partial_i u(t_0 - t, X_t) \sum_{j=1}^{\delta} \sigma_j^i(X_t) dB_t^j - \partial_t u(t_0 - t, X_t) dt + \mathcal{A}u(t_0 - t, X_t) dt.$$

The drift term disappears because $\partial_t u = \mathcal{A}u$. So M is a local martingale. Since u is bounded, M is a true martingale. So $\mathbb{E}[M_{t_0}] = M_0$. From the definition of M , we get $\mathbb{E}[f(X_{t_0})] = \mathbb{E}[u(0, X_{t_0})] = \mathbb{E}[M_{t_0}] = M_0 = u(t_0, x)$. \square

For the operator \mathcal{A} defined by

$$\mathcal{A} = \frac{1}{2} \sum_{j,k=1}^d a^{j,k}(x) \partial_j \partial_k + \sum_{i=1}^d b^i(x) \partial_i,$$

we define its adjoint \mathcal{A}^* by

$$\mathcal{A}^* g(x) = \frac{1}{2} \sum_{j,k=1}^d \partial_j \partial_k (a^{j,k}(x) g(x)) - \sum_{i=1}^d \partial_i (b_i(x) g(x)), \quad g \in C^2(\mathbb{R}^d).$$

It is called the adjoint of \mathcal{A} because

$$\int_{\mathbb{R}^d} \mathcal{A}f(x) g(x) dx = \int_{\mathbb{R}^d} f(x) \mathcal{A}^* g(x) dx, \quad f \in C_0^2(\mathbb{R}^d), g \in C^2(\mathbb{R}^d).$$

This fact follows from integration by parts (and that f vanishes outside a bounded set).

Theorem (Kolmogorov's forward equation). *Assume that the diffusion process X that satisfies the SDE*

$$dX_t = \sigma(X_t) \circ dB_t + b(X_t)dt$$

has a transition density $p_t(x, y)$, i.e.,

$$T_t f(x) = \mathbb{E}^x[f(X_t)] = \int_{\mathbb{R}^d} f(y)p_t(x, y)dy, \quad f \in C_0^2(\mathbb{R}^d),$$

and $p_t(x, y)$ is C^2 in y for every $t > 0$ and $x \in \mathbb{R}^d$. Then $p_t(x, y)$ satisfies the Kolmogorov's forward equation:

$$\partial_t p_t(x, \cdot) = \mathcal{A}_y^* p_t(x, \cdot), \quad t > 0,$$

where \mathcal{A}_y^ acts on the second variable of $p_t(x, y)$.*

Proof. Let $f \in C_0^2(\mathbb{R}^d)$. By Dynkin's formula, $\mathbb{E}^x[f(X_t)] = f(x) + \int_0^t \mathbb{E}^x[\mathcal{A}f(X_s)]ds$. By the definition of density function, for $t_2 > t_1 > 0$,

$$\int_{\mathbb{R}^d} f(y)(p_{t_2}(x, y) - p_{t_1}(x, y))dy = \int_{t_1}^{t_2} \int_{\mathbb{R}^d} \mathcal{A}f(y)p_s(x, y)dyds = \int_{\mathbb{R}^d} f(y) \int_{t_1}^{t_2} \mathcal{A}_y^* p_s(x, y)dsdy.$$

Since the equality holds for any $f \in C_0^2(\mathbb{R}^d)$, we get

$$p_{t_2}(x, y) - p_{t_1}(x, y) = \int_{t_1}^{t_2} \mathcal{A}_y^* p_s(x, y)ds, \quad t_2 > t_1 > 0.$$

Letting $t_2 - t_1 \rightarrow 0$, we get the conclusion. \square

6.4 The Feynman-Kac formula

Let the diffusion processes X_t^x , the infinitesimal generator A , and differential operator \mathcal{A} be as before. Suppose $f, q \in C_b(\mathbb{R}^d)$. Consider the following equation:

$$\begin{cases} \partial_t u = \mathcal{A}u + qu, & t > 0, x \in \mathbb{R}^d; \\ u(0, x) = f(x), & x \in \mathbb{R}^d. \end{cases} \quad (6.5)$$

Theorem . *If u is a solution of (6.5) such that for any t_0 , u is bounded on $[0, t_0] \times \mathbb{R}^d$, then*

$$u(t, x) = \mathbb{E}^x \left[f(X_t) \exp \left(\int_0^t q(X_s)ds \right) \right], \quad t \geq 0, x \in \mathbb{R}^d. \quad (6.6)$$

Proof. Fix $t_0 > 0$. Define $N(t) = u(t_0 - t, X_t)$, $0 \leq t \leq t_0$. By Itô's formula,

$$dN_t = -\partial_t u(t_0 - t, X_t)dt + \sum_{i=1}^d \partial_i u(t_0 - t, X_t)dX_t^i + \frac{1}{2} \sum_{j,k=1}^d \partial_j \partial_k u(t_0 - t, X_t)d[X^j, X^k]_t$$

$$\begin{aligned}
&= \sum_{i=1}^d \partial_i u(t_0 - t, X_t) \sum_{j=1}^{\delta} \sigma_j^i(X_t) dB_t^j - \partial_t u(t_0 - t, X_t) dt + \mathcal{A}u(t_0 - t, X_t) dt \\
&= \sum_{i=1}^d \partial_i u(t_0 - t, X_t) \sum_{j=1}^{\delta} \sigma_j^i(X_t) dB_t^j - q(X_t) u(t_0 - t, X_t) dt.
\end{aligned}$$

Let

$$M_t = N_t \exp\left(\int_0^t q(X_s) ds\right).$$

By product formula,

$$\begin{aligned}
dM_t &= \exp\left(\int_0^t q(X_s) ds\right) dN_t + M_t q(X_t) dt \\
&= \exp\left(\int_0^t q(X_s) ds\right) \sum_{i=1}^d \partial_i u(t_0 - t, X_t) \sum_{j=1}^{\delta} \sigma_j^i(X_t) dB_t^j.
\end{aligned}$$

Thus, M is a local martingale. For any $t_0 \geq 0$, since q is bounded on \mathbb{R}^d and u is bounded on $[0, t_0] \times \mathbb{R}^d$, we see that M is bounded on $[0, t_0]$. Thus, M is a true martingale. We then get $u(t_0, x) = M_0 = \mathbb{E}[M_{t_0}] = \mathbb{E}[f(X_{t_0}) \exp(\int_0^{t_0} q(X_s) ds)]$. \square

We call (6.6) the Feynman-Kac formula. If we define u by (6.6), we can not immediately say that u solves (6.5). Instead we have the following theorem.

Theorem . *Let $f \in C_0^2(\mathbb{R}^d)$ and $q \in C_b(\mathbb{R}^d)$. Define u by (6.6). Then u satisfies the equation*

$$\begin{cases} \partial_t u = \mathcal{A}u + qu, & t > 0, x \in \mathbb{R}^d; \\ u(0, x) = f(x), & x \in \mathbb{R}^d. \end{cases} \quad (6.7)$$

Proof. Let $Y_t = f(X_t)$ and $Z_t = \exp(\int_0^t q(X_s) ds)$. Then Z is an adapted C^1 process, and $dZ_t = Z_t q(X_t) dt$. Since q is bounded, say by R , we have $0 < Z_t \leq e^{Rt}$ for any $t \geq 0$. Recall that for some local martingale L_t , Y_t satisfies the SDE:

$$dY_t = dL_t + \mathcal{A}f(X_t) dt.$$

Since Z_t is a C^1 adapted process,

$$d(Y_t Z_t) = Y_t dZ_t + Z_t dY_t = Z_t dL_t + Z_t (\mathcal{A}f(X_t) + q(X_t) Y_t) dt.$$

Thus, the M defined by

$$M_t = Y_t Z_t - Y_0 Z_0 - \int_0^t Z_s (\mathcal{A}f(X_s) + q(X_s) f(X_s)) ds$$

is a local martingale. Fixed $t_0 > 0$. Since Z is uniformly bounded on $[0, t_0]$, and $Y = f(X)$, $\mathcal{A}f(X)$, and $q(X)f(X)$ are all uniformly bounded, M is also uniformly bounded on $[0, t_0]$. So M is a true martingale. Thus, $\mathbb{E}[M_t] = \mathbb{E}[M_0] = 0$, which implies that

$$u(t, x) = \mathbb{E}^x[Y_t Z_t] = f(x) + \int_0^t \mathbb{E}^x[Z_s(\mathcal{A}f(X_s) + q(X_s)f(X_s))]ds$$

is differentiable in t .

Let $r \geq 0$. We have $Z_{r+t}/Z_r = \exp(\int_0^t q(X_{r+s})ds)$. By the Markov property of X , we have

$$\mathbb{E}^x[Y_{r+t}Z_{r+t}|\mathcal{F}_r] = Z_r \mathbb{E}^x[f(X_{r+t}) \exp(\int_0^t q(X_{r+s})ds)|\mathcal{F}_r] = Z_r \mathbb{E}^{X_r}[Y_t Z_t] = Z_r u(t, X_r).$$

Thus,

$$u(r+t, x) = \mathbb{E}^x[Y_{r+t}Z_{r+t}] = \mathbb{E}^x[\mathbb{E}^x[Y_{r+t}Z_{r+t}|\mathcal{F}_r]] = \mathbb{E}^x[Z_r u(t, X_r)].$$

Recall that $Au(t, x) = \lim_{r \downarrow 0} \frac{1}{r}(T_r u(t, x) - u(t, x))$. Since $T_r u(t, x) = \mathbb{E}^x[u(t, X_r)]$, we get

$$\begin{aligned} \frac{1}{r}(T_r u(t, x) - u(t, x)) &= \frac{1}{r}(\mathbb{E}^x[u(t, X_r)] - u(r+t, x)) + \frac{1}{r}(u(r+t, x) - u(t, x)) \\ &= \mathbb{E}^x[u(t, X_r)] \frac{1}{r}(1 - Z_r) + \frac{1}{r}(u(r+t, x) - u(t, x)). \end{aligned}$$

As $r \downarrow 0$, the second term on the RHS tends to $\partial_t u(t, x)$. By dominated convergence theorem, the first term tends to $-q(x)u(t, x)$. So we get $u \in \mathcal{D}_A$ and $Au = \partial_t u - qu$. \square

6.5 Feller semigroup

We will learn some abstract theory in this subsection. It has a flavor of operator theory in Functional Analysis. Let $(T_t)_{t \geq 0}$ be a semigroup of transition operators associated with a time-homogeneous Markov process X , i.e.,

$$T_t f(x) = \mathbb{E}^x[f(X_t)].$$

Recall that T_t is a map from $C_b(\mathbb{R}^d)$ to $C_b(\mathbb{R}^d)$. Now we let $C_0 = C_0(\mathbb{R}^d)$ denote the space of continuous functions f on \mathbb{R}^d , which satisfy that $f(x) \rightarrow 0$ as $\|x\| \rightarrow \infty$. The space C_0 is a Banach space equipped with the uniform norm $\|f\| = \sup |f(x)|$.

Definition . We call $(T_t)_{t \geq 0}$ a Feller semigroup, and X a Feller process if

(F1) $T_t C_0 \subset C_0$, $t \geq 0$,

(F2) For any $f \in C_0$, $(t, x) \mapsto T_t f(x)$ is continuous on $[0, \infty) \times \mathbb{R}^d$.

By the definition of T_t , it is clear that each T_t is a contraction map, i.e., $\|T_t f\| \leq \|f\|$.

Example . Let $\sigma \in C(\mathbb{R}^d, \mathbb{R}^{d \times \delta})$ and $b \in C(\mathbb{R}^d, \mathbb{R}^d)$ be Lipschitz continuous. Let B be a δ -dimensional Brownian motion. Consider the SDE

$$dX_t^x = \sigma(X_t^x) \circ dB_t + b(X_t^x)dt, \quad X_0^x = x.$$

We have known that the solution exists uniquely, and we may choose a version of X^x for every $x \in \mathbb{R}^d$ such that X_t^x is jointly continuous in x and t . Now we further assume that σ and b are bounded. Then X is a Feller process. In fact, for $f \in C_0$, we already know that $T_t f \in C_b$. It remains to show that for any $t > 0$, $\mathbb{E}[f(X_t^x)] \rightarrow 0$ as $\|x\| \rightarrow \infty$. Fix $t > 0$. By dominated convergence theorem and the fact that $f \in C_0$, it suffices to show that $\|X_t^x\| \xrightarrow{P} \infty$ as $\|x\| \rightarrow \infty$. This holds because by the boundedness assumption on σ and b , $\|X_t^x - x\| = \|\int_0^t \sigma(X_s^x) \circ dB_s + \int_0^t b(X_s^x)ds\|$ is bounded in L^2 by a constant independent of x . Since $f(X_t^x)$ is jointly continuous in x and t , we get (F2) by dominated convergence theorem.

Recall that we say that (T_t) is a semigroup because it satisfies $T_t \circ T_s = T_{t+s}$. A semigroup of operators on a Banach space B may be constructed from bounded linear operators. For a bounded linear operator A on B , its exponential is the bounded linear operator e^A defined by

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!},$$

where A^0 is the identity I , $A^1 = A$, and $A^{n+1} = A \circ A^n$. The series converges in the operator norm because $\|A^n\| \leq \|A\|^n$. It is easy to see that $e^{tA} \circ e^{sA} = e^{(t+s)A}$. So $(e^{tA})_{t \geq 0}$ form a semigroup. We say that A is the infinitesimal generator of this semigroup, and may recover A from $(e^{tA})_{t \geq 0}$ because

$$\frac{1}{t}(e^{tA} - I) = \frac{1}{t} \sum_{n=1}^{\infty} \frac{t^n A^n}{n!} = A + t \sum_{n=2}^{\infty} \frac{t^{n-2} A^n}{n!} \rightarrow A, \quad \text{as } t \downarrow 0.$$

So we make the following definition for the Feller semigroup.

Definition . For a Feller semigroup (T_t) , its infinitesimal generator A is defined by

$$Af = \lim_{t \downarrow 0} \frac{1}{t}(T_t f - f), \quad f \in \mathcal{D}_A,$$

where \mathcal{D}_A is the set of all $f \in C_0$ such that the above limit converges in the norm topology.

It turns out that A is not defined on the whole space C_0 , and is not a bounded linear operator. We can not thus express T_t as e^{tA} . But we can still do something on it. The following theory is inspired by, but does not make reference to diffusion processes.

Proposition . If $f \in \mathcal{D}_A$, then (i) $T_t f \in \mathcal{D}_A$ for every $t \geq 0$; (ii) the function $t \mapsto T_t f$ is strongly differentiable in C_0 , and $\partial_t T_t f = T_t A f = A T_t f$; (iii) $T_t f - f = \int_0^t T_s A f ds = \int_0^t A T_s f ds$.

Proof. For fixed $t \geq 0$, using the semigroup property, we get

$$\lim_{r \downarrow 0} \frac{1}{r} [T_r T_t f - T_t f] = \lim_{r \downarrow 0} T_t \left[\frac{1}{r} (T_r f - f) \right] = T_t A f.$$

So we get $T_t f \in \mathcal{D}_A$, which is (i), and $A T_t f = T_t A f$. Since $T_r T_t = T_{r+t}$, the above shows that $t \mapsto T_t f$ has right-hand derivative $T_t A f = A T_t f$. Since $A f \in C_0$, $T_t A f$ is continuous in t . So the right-hand derivative is actually the two-sided derivative, and (ii) and (iii) both hold. \square

An operator A with domain \mathcal{D}_A on a Banach space B is called closed if its graph $\{(f, A f) : x \in \mathcal{D}_A\}$ is a closed subset of B^2 . This is a natural extension of bounded linear operators. An operator A is closed if and only if for any sequence (f_n) in \mathcal{D}_A , the two conditions “ $f_n \rightarrow f$ in B ” and “ $A f_n \rightarrow g$ in B ” together imply that $f \in \mathcal{D}_A$ and $A f = g$.

Proposition . *The domain \mathcal{D}_A of the infinitesimal generator A of a Feller semigroup is dense in C_0 , and A is a closed operator.*

Proof. Set $A_h f = \frac{1}{h} (T_h f - f)$ and $B_s f = \frac{1}{s} \int_0^s T_t f dt$. Then A_h and B_s are bounded operators on C_0 , and

$$A_h B_s = B_s A_h = A_s B_h = B_h A_s.$$

Here the first and the third “=” follows easily from the fact that T_h commutes with T_t . The second “=” follows from

$$\begin{aligned} B_s A_h f &= \frac{1}{hs} \int_0^s T_t (T_h f - f) dt = \frac{1}{hs} \int_0^s (T_{t+h} f - T_t f) dt = \frac{1}{hs} \left(\int_h^{h+s} T_t f dt - \int_0^s T_t f dt \right) \\ &= \frac{1}{hs} \left(\int_s^{h+s} T_t f dt - \int_0^h T_t f dt \right) = A_s B_h f. \end{aligned}$$

For every $f \in C_0$, we have $B_h f \rightarrow f$ as $h \downarrow 0$. Thus, for $s > 0$,

$$A_h B_s f = A_s B_h f \rightarrow A_s f, \quad h \downarrow 0.$$

So $B_s f \in \mathcal{D}_A$. Since $B_s f \rightarrow f$ as $s \downarrow 0$, \mathcal{D}_A is dense in C_0 .

Let (f_n) be a sequence in \mathcal{D}_A converging to f , and suppose $A f_n \rightarrow g$. Then

$$\begin{aligned} B_s g &= \lim_{n \rightarrow \infty} B_s A f_n = \lim_{n \rightarrow \infty} B_s \lim_{h \downarrow 0} A_h f_n = \lim_{n \rightarrow \infty} \lim_{h \downarrow 0} B_s A_h f_n \\ &= \lim_{n \rightarrow \infty} \lim_{h \downarrow 0} A_s B_h f_n = \lim_{n \rightarrow \infty} A_s f_n = A_s f. \end{aligned}$$

It follows that $f \in \mathcal{D}_A$, and $A f = \lim_{s \downarrow 0} A_s f = \lim_{s \downarrow 0} B_s g = g$. So A is closed. \square

Remark . For the Feller semigroup obtained from the SDE $dX_t = \sigma(X_t) \circ dB_t + b(X_t) dt$, we have known that the infinitesimal generator A agrees with the second order differential operator $\mathcal{A} = \frac{1}{2} \sum_{j,k} a^{j,k} \partial_j \partial_k + \sum_i b^i \partial_i$ on C_0^2 . From the above proposition, we know that, for any $f \in C_0$, if there is a sequence (f_n) in C_0^2 such that $f_n \rightarrow f$ and $\mathcal{A} f_n \rightarrow g \in C_0$, then $f \in \mathcal{D}_A$ and $A f = g$.

Let $\lambda > 0$, we define

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} T_t f(x) dt.$$

Since $T_t f \in C_0$ and $\|T_t f\| \leq \|f\|$ for all $t \geq 0$, by dominated convergence theorem, $R_\lambda f \in C_0$. By triangle inequality, $\|R_\lambda f\| \leq \int_0^\infty e^{-\lambda t} \|f\| dt = \frac{1}{\lambda} \|f\|$, i.e., $\|R_\lambda\| \leq \frac{1}{\lambda}$, and λR_λ is a contraction operator on C_0 . Each R_λ is called a resolvent of the Feller semigroup.

Proposition . *For any $\lambda > 0$, the operator $\lambda I - A$ from \mathcal{D}_A to C_0 is one-to-one and onto, and its inverse is R_λ .*

Proof. For $f \in \mathcal{D}_A$,

$$\begin{aligned} R_\lambda(\lambda I - A)f &= \int_0^\infty e^{-\lambda t} T_t(\lambda f - Af) dt \\ &= \lambda \int_0^\infty e^{-\lambda t} T_t f dt - \int_0^\infty e^{-\lambda t} \partial_t(T_t f) dt \\ &= - \int_0^\infty \partial_t(e^{-\lambda t}) T_t f dt - \int_0^\infty e^{-\lambda t} \partial_t(T_t f) dt \\ &= - \int_0^\infty \partial_t(e^{-\lambda t} T_t f) dt = -e^{-\lambda t} T_t f|_0^\infty = T_0 f = f. \end{aligned}$$

Conversely, if $f \in C_0$, then

$$\begin{aligned} A_h R_\lambda f &= R_\lambda A_h f = \int_0^\infty e^{-\lambda t} T_t \frac{1}{h} (T_h f - f) dt = \frac{1}{h} \int_0^\infty e^{-\lambda t} (T_{h+t} f - T_t f) dt \\ &= \frac{1}{h} \int_0^\infty e^{-\lambda t} T_{h+t} f dt - \frac{1}{h} \int_h^\infty e^{-\lambda t} T_t f dt - \frac{1}{h} \int_0^h e^{-\lambda t} T_t f dt \\ &= \frac{1}{h} \int_0^\infty (e^{-\lambda t} - e^{-\lambda(t+h)}) T_{h+t} f dt - \frac{1}{h} \int_0^h e^{-\lambda t} T_t f dt. \end{aligned}$$

As $h \downarrow 0$, the LHS tends to $A R_\lambda f$, while the RHS tends to $\int_0^\infty \lambda e^{-\lambda t} T_t f dt - f = \lambda R_\lambda f - f$. So $R_\lambda f \in \mathcal{D}_A$ and $A R_\lambda f = \lambda R_\lambda f - f$, which implies that $(\lambda I - A) R_\lambda f = f$. \square

From this proposition we see that the infinitesimal generator A determines all resolvents R_λ . Since $\lambda \mapsto R_\lambda f$ is the Laplace transform of $t \mapsto T_t f$, the resolvents in turn determine the Feller semigroup (T_t) .

Proposition (Resolvent identity). *(i) For any $\lambda, \mu > 0$,*

$$R_\lambda - R_\mu = (\mu - \lambda) R_\lambda R_\mu = (\mu - \lambda) R_\mu R_\lambda.$$

(ii) For any $f \in C_0$, $\|\lambda R_\lambda f - f\| \rightarrow 0$ as $\lambda \rightarrow \infty$.

Proof. (i) From $R_\lambda = (\lambda I - A)^{-1}$ we get

$$R_\mu^{-1} - R_\lambda^{-1} = \mu I - \lambda I.$$

Composing R_λ on the left of both sides, and composing R_μ from the right of both sides, we get $R_\lambda - R_\mu = (\mu - \lambda)R_\lambda R_\mu$. Switching R_λ and R_μ , we get $R_\lambda - R_\mu = (\mu - \lambda)R_\mu R_\lambda$.

(ii) First suppose $f \in \mathcal{D}_A$. Let $g = (I - A)f \in C_0$. Then $f = R_1 g$, and

$$\lambda R_\lambda f - f = \lambda R_\lambda R_1 g - R_1 g = R_\lambda R_1 g - R_\lambda g = R_\lambda(R_1 g - g) \rightarrow 0$$

in the norm topology because $R_1 g - g \in C_0$, and $\|R_\lambda\| \leq 1/\lambda$. For a general $f \in C_0$, since \mathcal{D}_A is dense in C_0 , we may find a sequence (f_n) in \mathcal{D}_A such that $f_n \rightarrow f$. For each n , $\lambda R_\lambda f_n \rightarrow f_n$ as $\lambda \rightarrow \infty$. Since λR_λ is a contraction for each $\lambda > 0$, $\lambda R_\lambda f_n \rightarrow \lambda R_\lambda f$ as $n \rightarrow \infty$ uniformly in λ . So we have $\lambda R_\lambda f \rightarrow f$ as $\lambda \rightarrow \infty$. \square

6.6 One-dimensional diffusion process

At the end, we focus on the solution of the following SDE

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt, \quad (6.8)$$

where $\sigma, b \in C^1(I, \mathbb{R})$ is positive, and I is an open interval on \mathbb{R} . Since σ and b are locally Lipschitz continuous, the strong solution X^x with any initial value $x \in I$ exists uniquely, which may blow up at some random finite time. When that happens, X_t^x tends to one end point of I (which could be $+\infty$ or $-\infty$) at that time.

We may simplify the SDE as follows. Suppose $Y = f(X)$ for some $f \in C^2(I)$. By Itô's formula, Y satisfies the SDE:

$$dY_t = f'(X_t)\sigma(X_t)dB_t + f'(X_t)b(X_t)dt + \frac{1}{2}f''(X_t)\sigma(X_t)^2dt.$$

If f is strictly increasing and satisfies the equation

$$f'(x)b(x) + \frac{1}{2}f''(x)\sigma(x)^2 = 0, \quad (6.9)$$

then the SDE for Y simplifies to

$$dY_t = (\sigma f') \circ f^{-1}(Y_t)dB_t, \quad (6.10)$$

and so Y is a local martingale.

Solving (6.9), we get $f''(x)/f'(x) = 2b(x)/\sigma(x)^2$. Since $f''/f' = \frac{d}{dx}(\log f')$, we get $\log f'(x) = C + \int_{x_0}^x 2b(s)/\sigma(s)^2 ds$, where x_0 is any point in I , and $C \in \mathbb{R}$ is a constant. So $f'(x) = \exp(\int_{x_0}^x 2b(s)/\sigma(s)^2 ds)$ is one solution. Integrating f' , we get f . Since f' is C^1 and positive, f is C^2 and strictly increasing. Since $\sigma \in C^1$, we see that $(\sigma f') \circ f^{-1} \in C^1$. So we can reduce (6.8) to the equation

$$dX_t = \sigma(X_t)dB_t, \quad (6.11)$$

where $\sigma \in C^1(I, \mathbb{R})$ is positive.

Lemma . *Let $a < b \in I$. The process X that satisfies (6.8) does not stay in $[a, b]$ during its life period $[0, T)$.*

Proof. By applying a function $f \in C^2(I)$, we may assume that X satisfies (6.11). Then X is a local martingale, and so is a time-change of a full or stopped Brownian motion. This means that there is a Brownian motion W defined on an extended probability space such that $X_t = W_{[X]_t}$, $0 \leq t < T$. Then almost surely either $[X]_\infty = \infty$ and $\limsup_{t \uparrow T} X_t = +\infty$ and $\liminf_{t \uparrow T} X_t = -\infty$, or $[X]_T < \infty$ and $\lim_{t \uparrow T} X_t$ converges to $W_{[X]_T}$. Suppose $X_t \in [a, b]$ for $0 \leq t < T$. Then the first case can not happen, and we have $[X]_T < \infty$ and $\lim_{t \uparrow T} X_t$ converges to some $x_0 \in [a, b]$. If $T < \infty$, then X_t approaches some end point of I . So we must have $T = \infty$. But now $[X]_T = \int_0^\infty \sigma(X_t)^2 = \infty$ because $\lim_{t \uparrow \infty} \sigma(X_t) = \sigma(x_0) > 0$, which is a contradiction. \square

Proposition . *For the solution X of (6.11), if $a < x < b \in I$, then*

$$\mathbb{P}^x[\tau_a < \tau_b] = \frac{b-x}{b-a}, \quad \mathbb{P}^x[\tau_b < \tau_a] = \frac{x-a}{b-a}. \quad (6.12)$$

where τ_a and τ_b are the first time that X reaches a and b , respectively.

Proof. We know that it does not happen that $X([0, T)) \subset [a, b]$. So $\tau_a \wedge \tau_b < T$. Since X is a local martingale, and X is bounded on $[0, \tau_a \wedge \tau_b]$, we get

$$x = X_0^x = \mathbb{E}[X_{\tau_a \wedge \tau_b}^x] = a\mathbb{P}^x[\tau_a < \tau_b] + b\mathbb{P}^x[\tau_b < \tau_a].$$

Since $\mathbb{P}^x[\tau_a < \tau_b] + \mathbb{P}^x[\tau_b < \tau_a] = 1$, we get (6.12). \square

Corollary . *For the solution X of (6.8) and an injective C^2 function f on I such that $f(X)$ is a local martingale, if $a < x < b \in I$, then*

$$\mathbb{P}^x[\tau_a < \tau_b] = \frac{f(b) - f(x)}{f(b) - f(a)}, \quad \mathbb{P}^x[\tau_b < \tau_a] = \frac{f(x) - f(a)}{f(b) - f(a)}. \quad (6.13)$$

Proof. If X starts from x , then $Y := f(X)$ satisfies (6.10) and starts from $f(x)$. The time that X reaches a or b is the time that Y reaches $f(a)$ or $f(b)$. Then we apply the previous proposition to Y . \square

The increasing function f with the property of (6.13) is called a scale function for X . Such scale function is unique up to an affine map. If $f(x) = x$ is a scale function, then X is a local martingale, and we say that X is on a natural scale.

For X on a natural scale, the behavior of X at its terminal time, i.e., the limit of X_t as $t \uparrow T$, is determined by the interval I . This is the statement of the following proposition.

Proposition . *Suppose X solves (6.11) on an interval I , and starts from $x \in I$.*

(i) *If $I = \mathbb{R}$, then a.s. $T = \infty$, $\limsup_{t \uparrow T} X_t = +\infty$ and $\liminf_{t \uparrow T} X_t = -\infty$.*

(ii) If $I = (a, \infty)$ or $(-\infty, a)$ for some $a \in \mathbb{R}$, then a.s. $\lim_{t \uparrow T} X_t = a$.

(iii) If $I = (a, b)$ for some $a < b \in \mathbb{R}$, then

$$\mathbb{P}^x[\lim_{t \uparrow T} X_t = a] = \frac{b - x}{b - a}, \quad \mathbb{P}^x[\lim_{t \uparrow T} X_t = b] = \frac{x - a}{b - a}.$$

Proof. Since X is a local martingale, on an extended probability space there is a Brownian motion W such that $X_t = W_{[X]_t}$, $0 \leq t < T$. So a.s. either Case 1: $[X]_T < \infty$ and $\lim_{t \uparrow T} X_t$ converges, or Case 2: $[X]_T = \infty$, $\limsup_{t \uparrow T} X_t = +\infty$ and $\liminf_{t \uparrow T} X_t = -\infty$. When $I = \mathbb{R}$, Case 1 can not happen because if $\lim_{t \uparrow T} X_t = x_0 \in \mathbb{R}$, then $[X]_T = \int_0^T \sigma(X_s)^2 ds < \infty$, which implies that $T < \infty$. But $T < \infty$ implies that $X_t \rightarrow +\infty$ or $-\infty$ as $t \uparrow T$, a contradiction. So Case 2 must happen, i.e., $[X]_T = \infty$, $\limsup_{t \uparrow T} X_t = +\infty$ and $\liminf_{t \uparrow T} X_t = -\infty$. We then get a.s. $T = \infty$ because if $T < \infty$, then X_t tends to ∞ or $-\infty$ as $t \uparrow T$. So we get (i). For (ii) and (iii), Case 2 can not happen because $I \neq \mathbb{R}$, and so a.s. $\lim_{t \uparrow T} X_t$ converges. The limit must be a boundary point because if it is $x_0 \in I$, then from $[X]_T = \int_0^T \sigma(X_s)^2 ds < \infty$ we get $T < \infty$, which implies that X_t tends to an end point of I , a contradiction. So we get (ii). For (iii), we have $\mathbb{P}^x[\lim_{t \uparrow T} X_t = a] + \mathbb{P}^x[\lim_{t \uparrow T} X_t = b] = 1$. Then we may compute $\mathbb{P}[\lim_{t \uparrow T} X_t = a]$ using the martingale property of X . \square

Corollary . Suppose X solves (6.8) on an interval I with a scale function f , and starts from $x \in I$.

(i) If $f(I) = \mathbb{R}$, then a.s. $T = \infty$, $\limsup_{t \uparrow T} X_t = \sup I$ and $\liminf_{t \uparrow T} X_t = \inf I$.

(ii) If $f(I) = (a, \infty)$ for some $a \in \mathbb{R}$, then a.s. $\lim_{t \uparrow T} X_t = \inf I$; if $f(I) = (-\infty, a)$ for some $a \in \mathbb{R}$, then a.s. $\lim_{t \uparrow T} X_t = \sup I$.

(iii) If $f(I) = (a, b)$ for some $a < b \in \mathbb{R}$, then

$$\mathbb{P}^x[\lim_{t \uparrow T} X_t = \inf I] = \frac{b - f(x)}{b - a}, \quad \mathbb{P}^x[\lim_{t \uparrow T} X_t = \sup I] = \frac{f(x) - a}{b - a}.$$

Suppose X is on a natural scale, i.e., solves (6.11). We are specially interested in the case that $I = \mathbb{R}$. In this case, X visits every point infinitely many times, and is recurrent. We call $\nu = \sigma(x)^{-2} dx$ the speed measure for X . The name comes from the following ergodicity theorem. We will not work on its proof.

Theorem . Suppose X is a diffusion on \mathbb{R} on a natural scale with speed measure ν . Then for any two nonnegative measurable functions f and g on I with $\nu f < \infty$ and $\nu g > 0$, and any $x \in \mathbb{R}$, we have \mathbb{P}^x -a.s.

$$\lim_{t \rightarrow \infty} \frac{\int_0^t f(X_s) ds}{\int_0^t g(X_s) ds} = \frac{\nu f}{\nu g}.$$

In particular, if ν is a finite measure, and we normalize it to a probability measure $\mathbb{P}_0 = \nu/|\nu|$, then we get by choosing $g \equiv 1$,

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f(X_s) ds = \int f(x) \mathbb{P}_0(dx).$$

Moreover, if the process X starts with initial distribution \mathbb{P}_0 , then it is a stationary process.