Stocastic Process

Xiaolong Han
12313305@mail.sustech.edu.cn

January 24, 2026

1 Review Probability Theory

Definition 1.1. Let Ω be a sample space,. A **σ-algebra** $\mathcal{F}$ is a collection of subsets of Ω satisfying:
(1)$\Omega \in \mathcal{F}$.
(2)If $A \in \mathcal{F}$, then $A^C \in \mathcal{F}$.
(3)If $A_n \in \mathcal{F}, n \in \mathbb{N}$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Definition 1.2. If a function $\mathbb{P} : \mathcal{F} \to [0,1]$ satisfies:
(1)$\mathbb{P}(\Omega) = 1$.
(2)If $A_n \in \mathcal{F}, n \in \mathbb{N}$ are disjoint, then $\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$, then $\mathbb{P}$ is a **probability measure**. And $(\Omega, \mathcal{F}, \mathbb{P})$ is called a **probability space**.

Definition 1.3. A **ramdom variable** X is a measurable function from $(\Omega, \mathcal{F})$ to $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, i.e. $\forall A \in \mathcal{B}(\mathbb{R}^d)$, $\{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}$.

Definition 1.4. The **distribution** of X is a probability measure μ_X on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ defined as $\mu_X(A) = \mathbb{P}(X \in A) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$ for $\forall A \in \mathcal{B}(\mathbb{R}^d)$.

Remark. $\mu((-\infty, x_1] \times \cdots \times (-\infty, x_d]) = F_X(x_1, \cdots, x_d)$ is the cumulative distribution function of X.

Definition 1.5. We say two ramdom variables X and Y are **identically distributed** if $\mu_X = \mu_Y$.

Definition 1.6. For a ramdom variable X, if $\mathbb{E}[|X|] < \infty$, then the **expectation** of X is defined as $\mathbb{E}[X] = \int_\Omega X(\omega)\mathbb{P}(d\omega)$.

Remark. $\mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x)\mu_X(dx)$ for measurable function f with $\mathbb{E}[|f(X)|] < \infty$.

Definition 1.7. Let $\mathcal{F}_1, \mathcal{F}_2, \cdots, \mathcal{F}_n$ be σ-algebras, if $\forall A_i \in \mathcal{F}_i, i = 1, 2, \cdots, n$, we have $\mathbb{P}(\cap_{i=1}^n A_i) = \prod_{i=1}^n \mathbb{P}(A_i)$, then $\mathcal{F}_1, \mathcal{F}_2, \cdots, \mathcal{F}_n$ are **independent**.

Definition 1.8. Ramdom variables $X_1, X_2, \cdots, X_n$ are **independent** if $\sigma(X_1), \sigma(X_2), \cdots, \sigma(X_n)$ are independent.

Propsition 1.1. Let $X_1, X_2, \cdots, X_n$ be independent ramdom variables, then $\mathbb{E}[\prod_{i=1}^n X_i] = \prod_{i=1}^n \mathbb{E}[X_i]$, if $\mathbb{E}[|X_i|] < \infty$ for $\forall i$.

Definition 1.9. A ramdom variable Y is said to be a **conditional expectation** of X given $\mathcal{G} \subset \mathcal{F}$ if
(1)Y is $\mathcal{G}$ measurable.
(2)$\mathbb{E}[|Y|] < \infty$.
(3)$\forall A \in \mathcal{G}, \mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$.
And we denote $Y = \mathbb{E}[X|\mathcal{G}]$.

Definition 1.10. We call $\mathbb{P}(B|\mathcal{G}) = \mathbb{E}[\mathbb{1}_B|\mathcal{G}]$ the **conditional probability** of B given $\mathcal{G}$ for $B \in \mathcal{F}$.

Remark. $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, then $\mathbb{E}[X|\mathcal{G}]$ exists and is unique a.s..

Remark. $\forall A \in \mathcal{G}, \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A] = \mathbb{E}[X\mathbb{1}_A]$.

Propsition 1.2. Properties of conditional expectation:
(1)$\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}]$.
(2)If $Y$is $\mathcal{G}$ measurable and $\mathbb{E}[|XY|] < \infty$, then $\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}]$.
(3)If $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$, then $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}]$.
(4)If X is independent of $\mathcal{G}$, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.
(5)$X \geq Y$, then $\mathbb{E}[X|\mathcal{G}] \geq \mathbb{E}[Y|\mathcal{G}]$.
(6)Jesen's inequality: If ϕ is a convex function and $\mathbb{E}[|\phi(X)|] < \infty$, then $\mathbb{E}[\phi(X)|\mathcal{G}] \geq \phi(\mathbb{E}[X|\mathcal{G}])$.
(7)Holder's inequality: If $p, q > 1, \frac{1}{p} + \frac{1}{q} = 1$, $\mathbb{E}[|X|^p] < \infty, \mathbb{E}[|Y|^q] < \infty$, then $\mathbb{E}[|XY||\mathcal{G}] \leq (\mathbb{E}[|X|^p|\mathcal{G}])^{\frac{1}{p}}(\mathbb{E}[|Y|^q|\mathcal{G}])^{\frac{1}{q}}$.

Proof:
(1)(2) Obvious from definition.
(3) For $\forall A \in \mathcal{H}$, $\mathbb{E}[\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}]\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A] = \mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]\mathbb{1}_A]$.
(4) For $\forall A \in \mathcal{G}$, $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A] = \mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[X]\mathbb{P}(A) = \mathbb{E}[\mathbb{E}[X]\mathbb{1}_A]$.
(5) Let $Z = \mathbb{E}[X|\mathcal{G}] - \mathbb{E}[Y|\mathcal{G}]$, then for $\forall A \in \mathcal{G}$, $\mathbb{E}[Z\mathbb{1}_A] = \mathbb{E}[(X-Y)\mathbb{1}_A] \geq 0$. Take $A = \{Z < 0\} \in \mathcal{G}$, then $\mathbb{E}[Z\mathbb{1}_{\{Z<0\}}] \geq 0$, thus $\mathbb{P}(Z < 0) = 0$. Hence $\mathbb{E}[X|\mathcal{G}] \geq \mathbb{E}[Y|\mathcal{G}]$.
(6) By definition of convex function, $\phi(x) = sup\{ax + b : \ a, b \in \mathbb{Q}, \phi(y) \geq ay + b, \forall y \in \mathbb{R}\}$. Thus there is a $\mathbb{P}$-null set $N_{a,b}$ s.t. $\phi(X(\omega)) \geq aX(\omega) + b$ for $\forall \omega \in \Omega \setminus N_{a,b}$. Let $N = \cup_{a,b\in\mathbb{Q}} N_{a,b}$, then $\mathbb{P}(N) = 0$ and $\phi(X(\omega)) \geq aX(\omega) + b$ for $\forall \omega \in \Omega \setminus N$. Thus for $\forall \omega \in N$,

$$\mathbb{E}[\phi(X)|\mathcal{G}](\omega) \geq \mathbb{E}[(aX + b)|\mathcal{G}](\omega) = a\mathbb{E}[X|\mathcal{G}](\omega) + b$$

By taking supremum over a, b, we have $\mathbb{E}[\phi(X)|\mathcal{G}] \geq \phi(\mathbb{E}[X|\mathcal{G}])$ a.s..
(7) Case 1: Either $\mathbb{E}[|X|^p|\mathcal{G}] = 0$ or $\mathbb{E}[|Y|^q|\mathcal{G}] = 0$, then the inequality holds trivially.

Case 2: $\mathbb{E}[|X|^p|\mathcal{G}] > 0$ and $\mathbb{E}[|Y|^q|\mathcal{G}] > 0$. Define

$$U = \frac{|X|}{(\mathbb{E}[|X|^p|\mathcal{G}])^{\frac{1}{p}}}, \quad V = \frac{|Y|}{(\mathbb{E}[|Y|^q|\mathcal{G}])^{\frac{1}{q}}}$$

Then by Young's inequality, $UV \leq \frac{U^p}{p} + \frac{V^q}{q}$. Thus

$$\frac{|XY|}{(\mathbb{E}[|X|^p|\mathcal{G}])^{\frac{1}{p}}(\mathbb{E}[|Y|^q|\mathcal{G}])^{\frac{1}{q}}} \leq \frac{|X|^p}{p(\mathbb{E}[|X|^p|\mathcal{G}])} + \frac{|Y|^q}{q(\mathbb{E}[|Y|^q|\mathcal{G}])}$$

Taking conditional expectation on both sides, we have the conclusion. $\square$

2 Discrete Time Stochastic Process

Definition 2.1. On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a sequence of ramdom variables $X = \{X_n : n = 0, 1, 2, \cdots\}$ is called a **discrete-time stochastic process**. If they are independent, then X is called independent process. If they are independently and identically distributed (i.i.d. in short), then X is called a stationary independent process.

Definition 2.2. Let $\{X_n\}$ be a stationary independent process, $S_0 = 0$, $S_n = \sum_{i=1}^n X_i$ for $n \geq 1$, then $S = \{S_n : n = 0, 1, 2, \cdots\}$ is called a **random walk**.

Definition 2.3. Let $S = \{S_n\}$ be a random walk. For $x \in \mathbb{R}$, it is called **recurrent** value of S if $\forall \epsilon > 0$, $\mathbb{P}(|S_n - x| < \epsilon \text{ i.o.}) = 1$. If $\forall \epsilon > 0$, $\exists n$ s.t. $\mathbb{P}(|S_n - x| < \epsilon) > 0$, then x is called a possible value of S.

Definition 2.4. $\{\mathcal{F}_n\}$ is a **filtration** if $\mathcal{F}_n \subset \mathcal{F}_{n+1} \subset \mathcal{F}$ for $\forall n \geq 0$. If X_n is $\mathcal{F}_n$ measurable for $\forall n \geq 0$, then X is adapted to $\{\mathcal{F}_n\}$.

Notation: $\mathcal{F}_\infty = \sigma(\cup_{n=0}^\infty \mathcal{F}_n)$.

Remark. Set $\mathcal{F}_n^X = \sigma(X_0, X_1, \cdots, X_n)$, then $\{\mathcal{F}_n^X\}$ is a filtration and X is adapted to $\{\mathcal{F}_n^X\}$. It is called the natural filtration of X.

Definition 2.5. For $\{\mathcal{F}_n\}$, if r.v. α satisfies (1)$\alpha \in \mathbb{N} \cup \{\infty\}$, (2)$\{\alpha \leq n\} \in \mathcal{F}_n$ for $\forall n \geq 0$, then α is a **stopping time** w.r.t. $\{\mathcal{F}_n\}$.

Example 2.1. $X = \{X_n\}$, $\alpha = min\{n : X_n < c\}$, then α is a stopping time w.r.t. the natural filtration of X.

Definition 2.6. For a filtration $\{\mathcal{F}_n\}$, and a stopping time α, define $\mathcal{F}_\alpha = \{A \in \mathcal{F}_\infty : A \cap \{\alpha \leq n\} \in \mathcal{F}_n, \forall n \geq 0\}$, which is called **pre-$\alpha$ σ-algebra**.

Propsition 2.1. $\mathcal{F}_\alpha$ is a σ-algebra and $\alpha \in \mathcal{F}_\alpha$.

Proof:
(1) $\Omega \in \mathcal{F}_\alpha$ since $\Omega \cap \{\alpha \le n\} = \{\alpha \le n\} \in \mathcal{F}_n$ for $\forall n \ge 0$.
(2) If $A \in \mathcal{F}_\alpha$, then for $\forall n \ge 0$, $A^C \cap \{\alpha \le n\} = (\Omega \cap \{\alpha \le n\}) \setminus (A \cap \{\alpha = n\}) \in \mathcal{F}_n$, thus $A^C \in \mathcal{F}_\alpha$.
(3) If $A_i \in \mathcal{F}_\alpha, i \in \mathbb{N}$, then for $\forall n \ge 0$, $(\cup_{i=1}^\infty A_i) \cap \{\alpha \le n\} = \cup_{i=1}^\infty (A_i \cap \{\alpha \le n\}) \in \mathcal{F}_n$, thus $\cup_{i=1}^\infty A_i \in \mathcal{F}_\alpha$.
Finally, for $\forall n \ge 0$, $\{\alpha \le n\} \cap \{\alpha \le n\} = \{\alpha \le n\} \in \mathcal{F}_n$, thus $\alpha \in \mathcal{F}_\alpha$. $\square$

Remark. For $\forall A \in \mathcal{F}_\infty$, $A = \cup_{n \in \mathbb{N} \cup \{\infty\}} (A \cap \{\alpha = n\})$.

Propsition 2.2. Let α, β be two stopping times w.r.t. $\{\mathcal{F}_n\}$.
(1) If $\alpha \le \beta$ a.s., then $\mathcal{F}_\alpha \subset \mathcal{F}_\beta$.
(2) $\mathcal{F}_{\alpha \wedge \beta} = \mathcal{F}_\alpha \cap \mathcal{F}_\beta$, where $\alpha \wedge \beta = min\{\alpha, \beta\}$.

Proof:
(1) For $\forall A \in \mathcal{F}_\alpha$, $A \cap \{\beta \le n\} \subset A \cap \{\alpha \le n\} \in \mathcal{F}_n$ for $\forall n \in \mathbb{N} \cup \{\infty\}$, thus $A \in \mathcal{F}_\beta$.
(2) By (1), since $\alpha \wedge \beta \le \alpha$ and $\alpha \wedge \beta \le \beta$, we have $\mathcal{F}_{\alpha \wedge \beta} \subset \mathcal{F}_\alpha \cap \mathcal{F}_\beta$.
For $\forall A \in \mathcal{F}_\alpha \cap \mathcal{F}_\beta$, $A \cap \{\alpha \wedge \beta \le n\} = A \cap (\{\alpha \le n\} \cup \{\beta \le n\}) = (A \cap \{\alpha \le n\}) \cup (A \cap \{\beta \le n\}) \in \mathcal{F}_n$ for $\forall n \in \mathbb{N} \cup \{\infty\}$, thus $A \in \mathcal{F}_{\alpha \wedge \beta}$. $\square$

Definition 2.7. Let $X = \{X_n\}$ be adapted to $\{\mathcal{F}_n\}$, $X_n \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ for $\forall n \ge 0$.
(1) X is called a **martingale** if $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$ a.s. for $\forall n \ge 0$.
(2) X is called a **supermartingale** if $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \le X_n$ a.s. for $\forall n \ge 0$.
(3) X is called a **submartingale** if $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \ge X_n$ a.s. for $\forall n \ge 0$.

Remark. (1)If X is a martingale, then $-X$ is also a martingale.
(2)A martingale is both a supermartingale and a submartingale.
(3)$X_n \ge \mathbb{E}[X_{n+1}|\mathcal{F}_n] \iff X_n \ge \mathbb{E}[X_{n+k}|\mathcal{F}_n]$ for $\forall k \ge 1$. Similar for submartingale and martingale.

Example 2.2. Choose $X_n = a_n$. If $\{a_n\}$ is decreasing (increasing, constant), then X is a supermartingale (submartingale, martingale).

Example 2.3. Let $X = \{X_n\}$ be an independent process, $\{\mathcal{F}_n^X\}$ be its natural filtration, and $\mathbb{E}[|X_n|] < \infty$ for $\forall n \ge 0$. Then $S_0 = 0$, $S_n = \sum_{i=1}^n X_i$ for $n \ge 1$ is a martingale w.r.t. $\{\mathcal{F}_n^X\}$.

Proof:
For $\forall n \ge 0$, $\mathbb{E}[S_{n+1}|\mathcal{F}_n^X] = \mathbb{E}[S_n + X_{n+1}|\mathcal{F}_n^X] = S_n + \mathbb{E}[X_{n+1}|\mathcal{F}_n^X] = S_n + \mathbb{E}[X_{n+1}] = S_n$. $\square$

Example 2.4 (Doob's Martingale)**.** Let $Y \in L^1$, $\{\mathcal{F}_n\}$ be a filtration, then $X_n = \mathbb{E}[Y|\mathcal{F}_n]$ is a martingale w.r.t. $\{\mathcal{F}_n\}$.

Definition 2.8. $X = \{X_n\}$, $X_n \in L^1$, $\forall n \in \mathbb{N}$, $X_n \le X_{n+1}$ a.s., $X_1 = 0$ a.s., then X is called an **increasing process**.

Definition 2.9. $X = \{X_n\}$ is called **predictable** w.r.t. $\{\mathcal{F}_n\}$ if X_n is $\mathcal{F}_{n-1}$ measurable for $\forall n \ge 1$.

Theorem 2.1 (Doob's Decomposition Theorem)**.** Let $X = \{X_n\}$ be a submartingale w.r.t. $\{\mathcal{F}_n\}$, $X_n \in L^1$ for $\forall n \ge 0$. Then there exists a unique decomposition $X_n = M_n + A_n$, where $M = \{M_n\}$ is a martingale w.r.t. $\{\mathcal{F}_n\}$, $A = \{A_n\}$ is an increasing predictable process w.r.t. $\{\mathcal{F}_n\}$ with $A_0 = 0$ a.s..

Proof:

Define $A_0 = 0$ a.s., and for $n \ge 1$, $A_n = \sum_{i=1}^{n} \mathbb{E}[X_i - X_{i-1}|\mathcal{F}_{i-1}]$. Then A_n is $\mathcal{F}_{n-1}$ measurable, and $A_n - A_{n-1} = \mathbb{E}[X_n - X_{n-1}|\mathcal{F}_{n-1}] \ge 0$ a.s.. Thus A is an increasing predictable process. Define $M_n = X_n - A_n$, then $\mathbb{E}[M_n|\mathcal{F}_{n-1}] = \mathbb{E}[X_n - A_n|\mathcal{F}_{n-1}] = \mathbb{E}[X_n|\mathcal{F}_{n-1}] - A_{n-1} = X_{n-1} + \mathbb{E}[X_n - X_{n-1}|\mathcal{F}_{n-1}] - A_{n-1} = M_{n-1}$. Thus M is a martingale.

For uniqueness, suppose there is another decomposition $X_n = M_n' + A_n'$, where M' is a martingale, A' is an increasing predictable process with $A_0' = 0$ a.s.. Then $M_n - M_n' = A_n' - A_n$. Since the left side is a martingale and the right side is an increasing process, both sides must be 0 a.s.. Thus the decomposition is unique. □

Remark. If we do not require A to be predictable, then the decomposition may not be unique.

Corollary 2.1. Let $X = \{X_n\}$ be a submartingale w.r.t. $\{\mathcal{F}_n\}$, $X_n \in L^1$ for $\forall n \ge 0$, whose Doob's decomposition is $X_n = M_n + A_n$ where $M = \{M_n\}$ is a martingale, $A = \{A_n\}$ is an increasing predictable process w.r.t. $\{\mathcal{F}_n\}$ with $A_0 = 0$ a.s..

(1) If X is L^1 bounded, i.e. $\sup_{n\ge 0} \mathbb{E}[|X_n|] < \infty$, then so are $M = \{M_n\}$ and $A = \{A_n\}$.

(2) If X is uniformly integrable, i.e. $\lim_{c\to\infty} \sup_{n\ge 0} \mathbb{E}[|X_n| \mathbb{1}_{\{|X_n|>c\}}] = 0$, then so are $M = \{M_n\}$ and $A = \{A_n\}$.

Lemma 2.1. Uniformly integrable $\iff$ L^1 bounded and $\forall \epsilon > 0$, $\exists \delta > 0$ s.t. $\forall A \in \mathcal{F}$ with $\mathbb{P}(A) < \delta$, we have $\sup_{n\ge 0} \mathbb{E}[|X_n| \mathbb{1}_A] < \epsilon$.

Proof:

"$\Rightarrow$":

$$\begin{aligned}
0 \le \mathbb{E}[A_n] = \mathbb{E}[X_n - M_n] &= \mathbb{E}[X_n] - \mathbb{E}[M_n] \\
&\le \sup_{n\in\mathbb{N}} \mathbb{E}[X_n] - \mathbb{E}[M_0] \\
&\le \sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|] - \mathbb{E}[|M_0|] < \infty
\end{aligned}$$

Thus $A_n \in L^1$ for $\forall n \geq 0$. And hence $M_n = X_n - A_n \in L^1$ for $\forall n \geq 0$. Also,

$$\begin{aligned}\mathbb{E}[|M_n|] &\leq \mathbb{E}[|X_n|] + \mathbb{E}[|A_n|] = \mathbb{E}[|X_n|] + \mathbb{E}[A_n] \\ &\leq 2\sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|] - \mathbb{E}[|M_0|] < \infty\end{aligned}$$

Thus $\sup_{n\geq 0} \mathbb{E}[|M_n|] < \infty$ and $\sup_{n\geq 0} \mathbb{E}[|A_n|] < \infty$.
"$\Leftarrow$": Since X is uniformly integrable, by the lemma, X is L^1 bounded. Thus by (1), M and A are also L^1 bounded. Since A_n is increasing and non-negative, we have $A' = lim_{n\to\infty} A_n \geq 0$, which can be infinity. By Fatou theorem, $\mathbb{E}[|A'|] = \mathbb{E}[\liminf_{n\to\infty} |A_n|] \leq \liminf_{n\to\infty} \mathbb{E}[|A_n|] < \infty$, thus $A' \in L^1$. For $\forall n \geq 0$, $\forall a \in \mathbb{R}^+$,

$$0 \leq \mathbb{E}[|A_n| \mathbb{1}_{\{|A_n|>a\}}] \leq \mathbb{E}[|A'| \mathbb{1}_{\{|A'|>a\}}]$$

Taking limit as $a \to \infty$ on both sides, we have $\lim_{a\to\infty} \mathbb{E}[|A_n| \mathbb{1}_{\{|A_n|>a\}}] = 0$ for $\forall n \geq 0$. Thus A is uniformly integrable.
For $\forall E \in \mathcal{F}$ we have $0 \leq \mathbb{E}[|M_n| \mathbb{1}_E] \leq \mathbb{E}[|X_n| \mathbb{1}_E] + \mathbb{E}[|A_n| \mathbb{1}_E]$. For $\forall \epsilon > 0$, since X and A are uniformly integrable, there exists $\delta > 0$ s.t. $\mathbb{P}(E) < \delta$ implies $\mathbb{E}[|X_n| \mathbb{1}_E] < \frac{\epsilon}{2}$ and $\mathbb{E}[|A_n| \mathbb{1}_E] < \frac{\epsilon}{2}$ for $\forall n \geq 0$. Thus $\mathbb{E}[|M_n| \mathbb{1}_E] < \epsilon$ for $\forall n \geq 0$. Hence M is uniformly integrable. □

Propsition 2.3. Let $X = \{X_n\}$ be a process.
(1) If X is a martingale, then $\mathbb{E}[X_n] = \mathbb{E}[X_0]$ for $\forall n \geq 0$.
(2) For $\forall A_n \in \mathcal{F}_n$ we have $\mathbb{E}[X_{n+1} \mathbb{1}_{A_n}] = \mathbb{E}[X_n \mathbb{1}_{A_n}]$ for $\forall n \geq 0$.
And similar for supermartingale and submartingale.

Proof:
$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$, thus we have

$$\mathbb{E}[X_n \mathbb{1}_{A_n}] = \mathbb{E}[\mathbb{E}[X_{n+1}|\mathcal{F}_n] \mathbb{1}_{A_n}] = \mathbb{E}[X_{n+1} \mathbb{1}_{A_n}]$$

for $\forall A_n \in \mathcal{F}_n$. Taking $A_n = \Omega$, we have $\mathbb{E}[X_{n+1}] = \mathbb{E}[X_n]$. By induction, we have $\mathbb{E}[X_n] = \mathbb{E}[X_0]$ for $\forall n \geq 0$. □

Propsition 2.4. X, Y are martingales w.r.t. $\{\mathcal{F}_n\}$, then $aX + bY$ is also a martingale w.r.t. $\{\mathcal{F}_n\}$ for $\forall a, b \in \mathbb{R}$. For supermartingale and submartingale, a, b are required to be non-negative.

Propsition 2.5. X is a submartingale w.r.t. $\{\mathcal{F}_n\}$, and f is an increasing convex function s.t. $\mathbb{E}[|f(X_n)|] < \infty$ for $\forall n \geq 0$, then $f(X) = \{f(X_n)\}$ is also a submartingale w.r.t. $\{\mathcal{F}_n\}$.

Proof:

$$\mathbb{E}[f(X_{n+1})|\mathcal{F}_n] \geq f(\mathbb{E}[X_{n+1}|\mathcal{F}_n]) \geq f(X_n)$$

□

Remark. For a martingale X, f is only required to be convex, and then $f(X)$ is a submartingale.

Theorem 2.2. $Y \in L^1$, $\{\mathcal{F}_n | n \in \mathbb{N} \cup \{\infty\}\}$ is a filtration, α is a stopping time w.r.t. $\{\mathcal{F}_n\}$. Let $X_n = \mathbb{E}[Y|\mathcal{F}_n]$, then $X_\alpha = \mathbb{E}[Y|\mathcal{F}_\alpha]$ a.s..

Proof:

By Jesen's inequality, $|X_n| = |\mathbb{E}[Y|\mathcal{F}_n]| \leq \mathbb{E}[|Y||\mathcal{F}_n]$, thus $\mathbb{E}[|X_n|] < \infty$ for $\forall n \geq 0$.

$$\begin{aligned}
\mathbb{E}[X_\alpha] &= \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathbb{E}[X_n \mathbb{1}_{\{\alpha=n\}}] = \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathbb{E}[\mathbb{E}[Y|\mathcal{F}_n]\mathbb{1}_{\{\alpha=n\}}] \\
&= \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathbb{E}[Y\mathbb{1}_{\{\alpha=n\}}] = \mathbb{E}[Y \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathbb{1}_{\{\alpha=n\}}] = \mathbb{E}[Y]
\end{aligned}$$

And X_α is $\mathcal{F}_\alpha$ measurable. For $\forall A \in \mathcal{F}_\alpha$,

$$\begin{aligned}
\mathbb{E}[X_\alpha \mathbb{1}_A] &= \mathbb{E}[\sum_{n\in\mathbb{N}\cup\{\infty\}} X_n \mathbb{1}_{\{\alpha=n\}}\mathbb{1}_A] = \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathbb{E}[X_n \mathbb{1}_{\{\alpha=n\}}\mathbb{1}_A] \\
&= \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathbb{E}[\mathbb{E}[Y|\mathcal{F}_n]\mathbb{1}_{\{\alpha=n\}}\mathbb{1}_A] = \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathbb{E}[Y\mathbb{1}_{\{\alpha=n\}}\mathbb{1}_A] \\
&= \mathbb{E}[Y \sum_{n\in\mathbb{N}\cup\{\infty\}} \mathbb{1}_{\{\alpha=n\}}\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]
\end{aligned}$$

Thus $X_\alpha = \mathbb{E}[Y|\mathcal{F}_\alpha]$ a.s.. □

Corollary 2.2. $Y \in L^1$, $\{\mathcal{F}_n | n \in \mathbb{N} \cup \{\infty\}\}$ is a filtration, α, β are two stopping times w.r.t. $\{\mathcal{F}_n\}$ with $\alpha \leq \beta$ a.s.. Let $X_n = \mathbb{E}[Y|\mathcal{F}_n]$, then $\{X_\alpha, X_\beta\}$ is a martingale w.r.t. the filtration $\{\mathcal{F}_\alpha, \mathcal{F}_\beta\}$.

Remark. Let $\{\alpha_n\}$ be a sequence of stopping times w.r.t. $\{\mathcal{F}_n\}$ with $\alpha_n \leq \alpha_{n+1}$ a.s. for $\forall n \geq 0$. Let $X_n = \mathbb{E}[Y|\mathcal{F}_n]$, then $\{X_{\alpha_n}\}$ is a martingale w.r.t. the filtration $\{\mathcal{F}_{\alpha_n}\}$.

Theorem 2.3. $X = \{X_n\}$ is a martingale w.r.t. $\{\mathcal{F}_n\}$, and α, β are two bounded stopping times w.r.t. $\{\mathcal{F}_n\}$ with $\alpha \leq \beta$ a.s., and $\mathbb{E}[\|X_\alpha\|] < \infty$, then $\{X_\alpha, X_\beta\}$ is a martingale w.r.t. the filtration $\{\mathcal{F}_\alpha, \mathcal{F}_\beta\}$.

Proof:

First, X_α and X_β are integrable. Since α and β are bounded, there exists $M \geq 0$ s.t. $\alpha \leq M$,

$\beta \le M$ a.s.. Thus

$$\begin{aligned}
\mathbb{E}[|X_\alpha|] &= \sum_{n=0}^{M} \mathbb{E}[|X_\alpha| \mathbb{1}_{\{\alpha=n\}}] \\
&= \sum_{n=0}^{M} \mathbb{E}[|X_n| \mathbb{1}_{\{\alpha=n\}}] \\
&\le \sum_{n=0}^{M} \mathbb{E}[|X_n|] < \infty
\end{aligned}$$

Thus $X_\alpha \in L^1$. Similarly, $X_\beta \in L^1$.
Second, let us start with supermartingale. For $\forall A \in \mathcal{F}_\alpha$, take $k \ge j$, then $A \cap \{\alpha = j\} \in \mathcal{F}_j \subset \mathcal{F}_k$, thus $A \cap \{\alpha = j\} \cap \{\beta \le k\} \in \mathcal{F}_k$. Hence

$$\begin{aligned}
\mathbb{E}[X_{k+1} \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta>k\}}] &= \mathbb{E}[\mathbb{E}[X_{k+1}|\mathcal{F}_k] \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta>k\}}] \\
&\le \mathbb{E}[X_k \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta>k\}}]
\end{aligned}$$

Then we have

$$\begin{aligned}
&\mathbb{E}[X_k \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta\ge k\}}] - \mathbb{E}[X_{k+1} \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta\ge k+1\}}] \\
&= \mathbb{E}[X_k \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta=k\}}] + \mathbb{E}[X_k \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta>k\}}] - \mathbb{E}[X_{k+1} \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta>k\}}] \\
&\ge \mathbb{E}[X_k \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta=k\}}]
\end{aligned}$$

Thus sum k from j to M, we have

$$\begin{aligned}
\mathbb{E}[X_\beta \mathbb{1}_{A\cap\{\alpha=j\}}] &\le \mathbb{E}[X_j \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta\ge j\}}] \\
&= \mathbb{E}[X_j \mathbb{1}_{A\cap\{\alpha=j\}}] \\
&= \mathbb{E}[X_\alpha \mathbb{1}_{A\cap\{\alpha=j\}}]
\end{aligned}$$

Now sum j from 1 to M, we have $\mathbb{E}[X_\beta \mathbb{1}_A] \le \mathbb{E}[X_\alpha \mathbb{1}_A]$, where $\mathbb{E}[X_\beta \mathbb{1}_A] = \mathbb{E}[\mathbb{E}[X_\beta|\mathcal{F}_\alpha]\mathbb{1}_A]$ for $\forall A \in \mathcal{F}_\alpha$. Thus $\mathbb{E}[X_\beta|\mathcal{F}_\alpha] \le X_\alpha$ a.s.. Similarly, for submartingale we have $\mathbb{E}[X_\beta|\mathcal{F}_\alpha] \ge X_\alpha$ a.s.. Hence for martingale we have $\mathbb{E}[X_\beta|\mathcal{F}_\alpha] = X_\alpha$ a.s.. □

Remark. For a sequence of bounded stopping times $\{\alpha_n\}$ w.r.t. $\{\mathcal{F}_n\}$ with $\alpha_n \le \alpha_{n+1}$ a.s. for $\forall n \ge 0$, then $\{X_{\alpha_n}\}$ is a martingale w.r.t. the filtration $\{\mathcal{F}_{\alpha_n}\}$. And similar for supermartingale and submartingale.

Example 2.5. Let α be a stopping time, then $\alpha \wedge n$ is a bounded stopping time for $\forall n \ge 0$. Thus $\{X_{\alpha\wedge n}\}$ is a martingale w.r.t. the filtration $\{\mathcal{F}_{\alpha\wedge n}\}$.

Theorem 2.4 (Discrete Optional Sampling Theorem)**.** Define $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$. Let $X = \{X_n : n \in \bar{\mathbb{N}}\}$ be a martingale (supermartingale, submartingale) w.r.t. $\{\mathcal{F}_n\}$, and α, β are two stopping times w.r.t. $\{\mathcal{F}_n\}$ with $\alpha \leq \beta$ a.s.. Then $\{X_\alpha, X_\beta\}$ is a martingale (supermartingale, submartingale) w.r.t. $\{\mathcal{F}_\alpha, \mathcal{F}_\beta\}$.

Proof:

Step1: Start with supermartingale with $X_\infty = 0$.

On the one hand, for $\forall \epsilon > 0$, $\{X_n < -\epsilon\} \in \mathcal{F}_n$. Take $A = \{X_n < -\epsilon\}$. Since $X_n \geq \mathbb{E}[X_\infty|\mathcal{F}_n]$, we have $\mathbb{E}[X_n \mathbb{1}_A] \geq \mathbb{E}[X_\infty \mathbb{1}_A|\mathcal{F}_n] = \mathbb{E}[X_\infty \mathbb{1}_A] = 0$. On the other hand, $\mathbb{E}[X_n \mathbb{1}_A] \leq -\epsilon \mathbb{P}(A)$. Thus $\mathbb{P}(A) = 0$. Since ϵ is arbitrary, we have $X_n \geq 0$ a.s. for $\forall n \geq 0$.

Notice that $X_\alpha \leq \liminf_{n\to\infty} X_{\alpha\wedge n}$ a.s.. By Fatou theorem, we have

$$0 \leq \mathbb{E}[X_\alpha] \leq \mathbb{E}[\liminf_{n\to\infty} X_{\alpha\wedge n}] \leq \liminf_{n\to\infty} \mathbb{E}[X_{\alpha\wedge n}] \leq \mathbb{E}[X_1] \leq \infty$$

For $\forall A \in \mathcal{F}_\alpha$, we have

$$\mathbb{E}[X_k \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta\geq k\}}] - \mathbb{E}[X_{k+1} \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta\geq k+1\}}] \geq \mathbb{E}[X_k \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta=k\}}]$$

Then sum k from j to M, we have

$$\mathbb{E}[X_j \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta\geq j\}}] - \mathbb{E}[X_{M+1} \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta\geq M+1\}}] \geq \mathbb{E}[X_\beta \mathbb{1}_{A\cap\{\alpha=j\}\cap\{j\leq\beta\leq M\}}]$$

Thus we have

$$\begin{aligned}
\mathbb{E}[X_j \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta\geq j\}}] &= \mathbb{E}[X_\alpha \mathbb{1}_{A\cap\{\alpha=j\}}] \\
&\geq \mathbb{E}[X_\beta \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta\leq M\}}] \\
&\geq \lim_{M\to\infty} \mathbb{E}[X_\beta \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta\leq M\}}] \\
&\geq \mathbb{E}[X_\beta \lim_{M\to\infty} \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta\leq M\}}] = \mathbb{E}[X_\beta \mathbb{1}_{A\cap\{\alpha=j\}\cap\{\beta<\infty\}}]
\end{aligned}$$

Now sum j in $\mathbb{N}$, we get $\mathbb{E}[X_\alpha \mathbb{1}_{A\cap\{\alpha<\infty\}}] \geq \mathbb{E}[X_\beta \mathbb{1}_{A\cap\{\alpha<\infty\}\cap\{\beta<\infty\}}] = \mathbb{E}[X_\beta \mathbb{1}_{A\cap\{\beta<\infty\}}]$.

And we also have $\mathbb{E}[X_\alpha \mathbb{1}_{A\cap\{\alpha=\infty\}}] = 0 = \mathbb{E}[X_\beta \mathbb{1}_{A\cap\{\beta=\infty\}}]$. Thus for $forall A \in \mathcal{F}_\alpha$, we have

$$\begin{aligned}
\mathbb{E}[X_\alpha \mathbb{1}_A] &= \mathbb{E}[X_\alpha \mathbb{1}_{A\cap\{\alpha<\infty\}}] + \mathbb{E}[X_\alpha \mathbb{1}_{A\cap\{\alpha=\infty\}}] \\
&\geq \mathbb{E}[X_\beta \mathbb{1}_{A\cap\{\beta<\infty\}}] + \mathbb{E}[X_\beta \mathbb{1}_{A\cap\{\beta=\infty\}}] \\
&= \mathbb{E}[X_\beta \mathbb{1}_A]
\end{aligned}$$

Take $A = \{\mathbb{E}[X_\beta|\mathcal{F}_\alpha] > X_\alpha\}$, we have $\mathbb{E}[X_\beta \mathbb{1}_A] > \mathbb{E}[X_\alpha \mathbb{1}_A]$, which is a contradiction. Thus $\mathbb{E}[X_\beta|\mathcal{F}_\alpha] \leq X_\alpha$ a.s..

Step2: For supermartingale with general X_∞.

Define $X'_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$, then $X_n - X'_n$ is a supermartingale with $X'_\infty = X_\infty$. By Step1, we have $\mathbb{E}[X_\beta - X'_\beta|\mathcal{F}_\alpha] \leq X_\alpha - X'_\alpha$ a.s.. Thus $\mathbb{E}[X_\beta|\mathcal{F}_\alpha] \leq X_\alpha$ a.s..

Step3: For submartingale and martingale, similar proof. $\square$

Definition 2.10. $Y = \{Y_n\}$ is a martingale (submartingale, supermartingale) w.r.t. $\{\mathcal{F}_n\}$ with $Y_0 = 0$. $Z = \{Z_n\}$ is $\{\mathcal{F}_n\}$-predictable. Set $X_0 = 0$, $X_n = \sum_{k=1}^{n} Z_k(Y_k - Y_{k-1})$. Then X is called a **martingale transformation** of Y through Z.

Propsition 2.6. Y is a martingale, Z is predictable, X is martingale transformation of Y through Z. If X is integrable, then X is a martingale.

Proof:
Clearly, X is adapted. And we have

$$\begin{aligned}\mathbb{E}[X_{n+1}|\mathcal{F}_n] &= X_n + \mathbb{E}[Z_{n+1}(Y_{n+1} - Y_n)|\mathcal{F}_n] \\ &= X_n - Z_{n+1}\mathbb{E}[Y_n - Y_{n+1}|\mathcal{F}_n] \\ &= X_n\end{aligned}$$

$\square$

Remark. Similar conclusion holds for supermartingale and submartingale.

Definition 2.11. $X = \{X_n\}$ are integrable, $\{\mathcal{F}_n\}$ is a decreasing filtration, i.e. $\mathcal{F}_n \supset \mathcal{F}_{n+1}$ for $\forall n \geq 0$.
(1) If $\mathbb{E}[X_n|\mathcal{F}_{n+1}] = X_{n+1}$ for $\forall n \geq 0$, then X is called a **backward martingale** w.r.t. $\{\mathcal{F}_n\}$.
(2) If $\mathbb{E}[X_n|\mathcal{F}_{n+1}] \leq X_{n+1}$ for $\forall n \geq 0$, then X is called a **backward submartingale** w.r.t. $\{\mathcal{F}_n\}$.
(3) If $\mathbb{E}[X_n|\mathcal{F}_{n+1}] \geq X_{n+1}$ for $\forall n \geq 0$, then X is called a **backward supermartingale** w.r.t. $\{\mathcal{F}_n\}$.

Propsition 2.7. Let ϕ be a increasing convex function, and $X = \{X_n\}$ be a backward submartingale w.r.t. $\{\mathcal{F}_n\}$ with $\mathbb{E}[|\phi(X_n)|] < \infty$ for $\forall n \geq 0$. Then $\phi(X) = \{\phi(X_n)\}$ is also a backward submartingale w.r.t. $\{\mathcal{F}_n\}$.

Corollary 2.3. If in addition, $\phi \geq 0$, then $\{\phi(X_n)\}$ is uniformly integrable.

Proof:
By Jesen's inequality, we have

$$\mathbb{E}[\phi(X_n)|\mathcal{F}_{n+1}] \geq \phi(\mathbb{E}[X_n|\mathcal{F}_{n+1}]) \geq \phi(X_{n+1})$$

Thus $\phi(X)$ is a backward submartingale.
For $\forall A \in \mathcal{F}_n$, we have $\mathbb{E}[\phi(X_1)\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[\phi(X_1)|\mathcal{F}_n]\mathbb{1}_A] \geq \mathbb{E}[\phi(X_n)\mathbb{1}_A]$. Let $A = \{\phi(X_n) > M\}$,

we have $\mathbb{E}[\phi(X_n)\mathbb{1}_{\{\phi(X_n)>M\}}] \leq \mathbb{E}[\phi(X_1)\mathbb{1}_{\{\phi(X_n)>M\}}]$. By chebyshev inequality,

$$\mathbb{P}(\phi(X_n) > M) \leq \frac{\mathbb{E}[\phi(X_1)]}{M}$$

Thus $\lim_{M\to\infty} \sup_{n\geq 0} \mathbb{E}[\phi(X_n)\mathbb{1}_{\{\phi(X_n)>M\}}] = 0$. Hence $\{\phi(X_n)\}$ is uniformly integrable. □

Propsition 2.8. Let $X = \{X_n\}$ be a backward martingale w.r.t. $\{\mathcal{F}_n\}$, then X is uniformly integrable.

Definition 2.12. $X = \{X_n\}$ is a stochastic process with **Markov property** if for $\forall n \geq 0$, $\forall A \in \mathcal{B}(\mathbb{R}^d)$,

$$\mathbb{P}(X_{n+1} \in A|\mathcal{F}_n^X) = \mathbb{P}(X_{n+1} \in A|X_n).$$

And X is said to be a **Markov process**.

Propsition 2.9. X has Markov property $\iff$ for $\forall Y \in \sigma(X_{n+1})$, $\mathbb{E}[Y|\mathcal{F}_n^X] = \mathbb{E}[Y|X_n]$.

Proof:
"$\Rightarrow$": For $\forall A \in \mathcal{B}(\mathbb{R}^d)$, let $Y = \mathbb{1}_{\{X_{n+1}\in A\}}$, then we have $\mathbb{E}[Y|\mathcal{F}_n^X] = \mathbb{E}[Y|X_n]$. By the monotone class theorem, we have $\mathbb{E}[Y|\mathcal{F}_n^X] = \mathbb{E}[Y|X_n]$ for $\forall Y \in \sigma(X_{n+1})$.
"$\Leftarrow$": For $\forall A \in \mathcal{B}(\mathbb{R}^d)$, let $Y = \mathbb{1}_{\{X_{n+1}\in A\}}$, then we have $\mathbb{P}(X_{n+1} \in A|\mathcal{F}_n^X) = \mathbb{P}(X_{n+1} \in A|X_n)$. □

Corollary 2.4. For $\forall k \in \mathbb{N}$, $\forall Y \in \sigma(X_{n+k})$, we still have $\mathbb{E}[Y|\mathcal{F}_n^X] = \mathbb{E}[Y|X_n]$.

Proof:
By Markov property we have $\mathbb{E}[Y|\mathcal{F}_{n+k-1}^X] = \mathbb{E}[Y|X_{n+k-1}]$. For $\forall A \in \sigma(X_n, \cdots, X_{n+k-1})$, we have $\mathbb{E}[\mathbb{E}[Y|\mathcal{F}_{n+k-1}^X]\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$. And hence $\mathbb{E}[Y|\mathcal{F}_{n+k-1}^X] = \mathbb{E}[Y|X_{n+k-1}] = \mathbb{E}[Y|X_n, \cdots, X_{n+k-1}]$. Next, use induction. When $k = 1$, it's true. Assume $k - 1$ is true. Then we have

$$\begin{aligned}
\mathbb{E}[Y|X_1, \cdots, X_n] &= \mathbb{E}[\mathbb{E}[Y|X_1, \cdots, X_{n+k-1}]|X_1, \cdots, X_n] \\
&= \mathbb{E}[\mathbb{E}[Y|X_{n+k-1}]|X_1, \cdots, X_n] \\
&= \mathbb{E}[\mathbb{E}[Y|X_{n+k-1}]|X_n] \\
&= \mathbb{E}[\mathbb{E}[Y|X_n, \cdots, X_{n+k-1}]|X_n] \\
&= \mathbb{E}[Y|X_n]
\end{aligned}$$

□

Theorem 2.5. The following are equivalent:
(1) X is a Markov process.

(2) For $\forall n \in \mathbb{N}$, $\forall M \in \sigma(X_{n+1}, X_{n+2}, \cdots)$, it holds $\mathbb{P}(M|X_1, \cdots, X_n) = \mathbb{P}(M|X_n)$.
(3) For $\forall n \in \mathbb{N}$, $\forall M_1 \in \sigma(X_1, \cdots, X_n), \forall M_2 \in \sigma(X_{n+1}, X_{n+2}, \cdots)$, it holds $\mathbb{P}(M_1M_2|X_n) = \mathbb{P}(M_1|X_n)\mathbb{P}(M_2|X_n)$.

Proof:
"(2) $\Rightarrow$ (3)":

$$\begin{aligned}
\mathbb{P}(M_1|X_n)\mathbb{P}(M_2|X_n) &= \mathbb{E}[\mathbb{1}_{M_1}|X_n]\mathbb{E}[\mathbb{1}_{M_2}|X_n] \\
&= \mathbb{E}[\mathbb{1}_{M_1}\mathbb{E}[\mathbb{1}_{M_2}|X_n]|X_n] \\
&= \mathbb{E}[\mathbb{1}_{M_1}\mathbb{E}[\mathbb{1}_{M_2}|X_1, \cdots, X_n]|X_n] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_{M_1}\mathbb{1}_{M_2}|X_1, \cdots, X_n]|X_n] \\
&= \mathbb{E}[\mathbb{1}_{M_1}\mathbb{1}_{M_2}|X_n] \\
&= \mathbb{P}(M_1M_2|X_n)
\end{aligned}$$

"(3) $\Rightarrow$ (2)": For $\forall A \in \sigma(X_n)$, $\forall M^{'} \in \sigma(X_1, \cdots, X_n)$, we have

$$\begin{aligned}
\mathbb{E}[\mathbb{E}[\mathbb{1}_M|X_N]\mathbb{1}_{AM'}] &= \mathbb{E}[\mathbb{E}[\mathbb{1}_M|X_n]|\mathbb{1}_A\mathbb{1}_{M'}] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{E}[\mathbb{1}_M|X_n]\mathbb{1}_{M'}|X_n]\mathbb{1}_A] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_M|X_n]\mathbb{E}[\mathbb{1}_{M'}|X_n]\mathbb{1}_A] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_{MM'}|X_n]\mathbb{1}_A] \\
&= \mathbb{E}[\mathbb{1}_{MM'}\mathbb{1}_A] = \mathbb{E}[\mathbb{1}_M\mathbb{1}_{AM'}] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_M|X_1, \cdots, X_n]\mathbb{1}_{AM'}]
\end{aligned}$$

And hence $\mathbb{E}[\mathbb{1}_M|X_n] = \mathbb{E}[\mathbb{1}_M|X_1, \cdots, X_n]$.
"(2) $\Rightarrow$ (1)": Take $M = \{X_{n+1} \in B\} \in \sigma(X_{n+1})$.
"(1) $\Rightarrow$ (2)": Markov property $\iff \mathbb{E}[Y|X_1, \cdots, X_n] = \mathbb{E}[Y|X_n]$ for $\forall Y \in \sigma(X_{n+1}, \cdots, X_{n+k})$.

Example 2.6. $X = \{X_n\}$ is an independent process. Then it is a Markov process, and so is $S = \{S_n = X_1 + \cdots + X_n\}$.

Proof:
$\forall Y \in \sigma(X_{n+1})$, $\mathbb{E}[T|X_1, \cdots, X_n] = \mathbb{E}[Y] = \mathbb{E}[Y|X_n]$.
$\mathbb{P}(S_{n+1} \in B|S_1, \cdots, S_n) = \mathbb{P}(S_n + X_{n+1}|S_1, \cdots, S_n) = \mathbb{P}(S_n + X_{n+1}|S_n)$.

Definition 2.13. $X = \{X_n\}$ is a process, α is a stopping time and is finity a.s.. Then $\{X_{\alpha+n} : n \in \mathbb{N}\}$ is called **post-α process**. Correspondingly, $\mathcal{F}_\alpha^{'} = \sigma(X_{\alpha+n} : n \in \mathbb{N})$ is **post-α σ-algebra**.

Theorem 2.6. X is a stationary independent process, α is a stopping time and is finity a.s.. Then:
(1) $\mathcal{F}_\alpha^X$ is independent of $\mathcal{F}_\alpha^{X'}$.
(2) $\{X_{\alpha+n} : n \in \mathbb{N}\}$ has the same distribution as $\{X_n : n \in \mathbb{N}\}$ and is a stationary independent process.

Proof:

(1) For $\forall A \in \mathcal{F}_\alpha^X$, $\{X_{\alpha+j} \in B_j\} \in \mathcal{F}_\alpha^{X'}$, then we have

$$\begin{aligned}
\mathbb{P}(A \cap \{X_{\alpha+j} \in B_j, 1 \le j \le k\}) &= \mathbb{P}(\bigcup_{n=1}^{\infty} [A \cap \{\alpha = n\} \cap \{X_{\alpha+j} \in B_j, 1 \le j \le k\}]) \\
&= \mathbb{P}(\bigcup_{n=1}^{\infty} [A \cap \{\alpha = n\} \cap \{X_{n+j} \in B_j, 1 \le j \le k\}]) \\
&= \sum_{n=1}^{\infty} \mathbb{P}(A \cap \{\alpha = n\} \cap \{X_{n+j} \in B_j, 1 \le j \le k\}) \\
&= \sum_{n=1}^{\infty} \mathbb{P}(A \cap \{\alpha = n\}) \mathbb{P}(\{X_{n+j} \in B_j, 1 \le j \le k\}) \\
&= \sum_{n=1}^{\infty} [\mathbb{P}(A \cap \{\alpha = n\}) \prod_{j=1}^{k} \mathbb{P}(\{X_{n+j} \in B_j\})]
\end{aligned}$$

Take $A = \Omega$, $k = 1$, we have

$$\begin{aligned}
\mathbb{P}(\{X_{\alpha+1} \in B_1\}) &= \sum_{n=1}^{\infty} [\mathbb{P}(\{\alpha = n\}) \mathbb{P}(\{X_{n+1} \in B_1\})] \\
&= \sum_{n=1}^{\infty} [\mathbb{P}(\{\alpha = n\}) \mathbb{P}(\{X_1 \in B_1\})] \\
&= \mathbb{P}(\{X_1 \in B_1\})
\end{aligned}$$

Thus X_1 and $X_{\alpha+1}$ have the same distribution.

$$\begin{aligned}
\mathbb{P}(A \cap \{X_{\alpha+j} \in B_j, 1 \le j \le k\}) &= \sum_{n=1}^{\infty} [\mathbb{P}(A \cap \{\alpha = n\}) \prod_{j=1}^{k} \mathbb{P}(\{X_j \in B_j\})] \\
&= \mathbb{P}(A) \prod_{j=1}^{k} \mathbb{P}(\{X_j \in B_j\})
\end{aligned}$$

Take $A = \Omega$, we have $\mathbb{P}(\{X_{\alpha+j} \in B_j, 1 \le j \le k\}) = \prod_{j=1}^{k} \mathbb{P}(\{X_j \in B_j\})$. Thus $\{X_{\alpha+n} : n \in \mathbb{N}\}$ is a stationary independent process with the same distribution as $\{X_n : n \in \mathbb{N}\}$. $\square$

Theorem 2.7. *X is a Markov process, α is a stopping time and is finity a.s.. Then $\mathbb{P}(M|\mathcal{F}_\alpha^X) = \mathbb{P}(M|X_\alpha, \alpha)$ for $\forall M \in \mathcal{F}_\alpha^{X'}$.*

Remark. The sum of two Markov processes is not necessarily a Markov process.

Example 2.7. Let W be a random variable. Set $X_n = W$, $Y_n = (-1)^n W$. Then both X and Y are Markov processes, but $X + Y$ is not a Markov process.

Remark. A Markov process need not to be a martingale. And a martingale need not to be a Markov process.

Example 2.8. X is independent process. Set $S_n = X_1 + \cdots + X_n$. Then S is a Markov process. But it is not a martingale unless $\mathbb{E}[X_n] = 0$.

Example 2.9. Let X be an independent process $\in L^2$ with $\mathbb{E}[X_n] = 0$. Let W be a bounded random variable and is independent of X. Set $Y_n = W(X_{n-1} + \cdots + X_1 + 1)$, $Y_1 = W$, $\mathcal{F}_n = \sigma(W, X_1, \cdots, X_{n-1})$. Then Y is adapted and integrable, $\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = \mathbb{E}[Y_n + WX_n|\mathcal{F}_n] = Y_n$. Thus Y is a martingale. However, $\mathbb{E}[Y_{n+1}^2|Y_1, \cdots, Y_n] = Y_n^2 + W^2\mathbb{E}[X_n^2]$, while $\mathbb{E}[Y_{n+1}^2|Y_n] = Y_n^2 + 2Y_n\mathbb{E}[WX_n|Y_n] + \mathbb{E}[W^2X_n^2|Y_n] = Y_n^2 + \mathbb{E}[W^2|Y_n]\mathbb{E}[X_n^2]$, since $\mathbb{E}[WX_n|Y_n] = \mathbb{E}[\mathbb{E}[WX_n|Y_1, Y_n]|Y_n] = \mathbb{E}[W\mathbb{E}[X_n]|Y_n] = 0$. Thus, Y is not necessarily a Markov process.

3 Continuous Process

Definition 3.1. $X = \{X_t : t \geq 0\}$ is a **continuous-time stochastic process**.

Remark. $X_t(\omega)$:
(1) $\Omega \times [0, \infty) \to \mathbb{R}^d, (\omega, t) \mapsto X_t(\omega)$ as a funtion.
(2) $\Omega \to \mathbb{R}^d, \omega \mapsto X_t(\omega)$ as a random variable for $\forall t \geq 0$.
(3) $[0, \infty) \to \mathbb{R}^d, t \mapsto X_t(\omega)$ as a function for $\forall \omega \in \Omega$. And such function is called a trajectory or sample path of X corresponding to ω.

Definition 3.2. X is a process on $(\Omega, \mathcal{F}, \mathbb{P})$.
(1) We say sample path is **continuous** if for $\forall \omega \in \Omega$, $t \mapsto X_t(\omega)$ is continuous on $[0, \infty)$.
(2) We say sample path is right continuous (RC) and has left limit (LL) if for $\forall \omega \in \Omega$, $t \mapsto X_t(\omega)$ is right continuous and has left limit on $[0, \infty)$.

Definition 3.3. If for $\forall \epsilon > 0, \forall t \geq 0, \lim_{s \to t} \mathbb{P}(|X_t - X_s| \geq \epsilon) = 0$, then X is **stochastic continuous**.

Definition 3.4. If $X_t(\omega) = Y_t(\omega)$ for $\forall t \geq 0, \forall \omega \in \Omega$, then X and Y are the **same**.

Definition 3.5. If $\mathbb{P}(X_t = Y_t, \forall t \geq 0) = 1$, X and Y are **indistinguishable**.

Definition 3.6. If $\mathbb{P}(X_t = Y_t) = 1$ for $\forall t \geq 0$ then X, Y are **modifications** to each other.

Remark. indistinguishable $\implies$ modification, but not vice versa.

Definition 3.7. If $\forall n \geq 1$, $0 \leq t_1 < \cdots < t_n \leq \infty$, $\forall A \in \mathcal{B}(\mathbb{R}^d)$
$\mathbb{P}((X_{t_1}, \cdots, X_{t_n}) \in A) = \mathbb{P}((Y_{t_1}, \cdots, Y_{t_n}) \in A)$, then X, Y have the same **finite dimentional distributions**.

Theorem 3.1 (Kolmogorov Continuity)**.** For $T_0 > 0$, $X = \{X_t : t \in [0, T_0]\}$ if $\exists \alpha, \beta, c > 0$, s.t. $\mathbb{E}[|X_t - X_s|^\alpha] \leq c|t-s|^{1+\beta}$ for $\forall t, s \in [o, T_0]$, then $\exists Y$ satisfying:
(1)Y is modification of X.
(2)Y has continuous trajactories.
(3)$\forall \gamma \in (0, \frac{\beta}{\alpha})\, \exists \delta(\omega) > 0$, $c' > 0$ s.t.

$$p(\sup_{\substack{s,t\in[0,T_0],\\ 0<|t-s|<\delta(\omega)}} \frac{|Y_t(\omega) - Y_s(\omega)|}{|t-s|^\gamma} < c') = 1$$

Proof: First, $\mathbb{P}(|X_t - X_s| \geq \epsilon) \leq \frac{\mathbb{E}[|X_t - X_s|^\alpha]}{\epsilon^\alpha} \leq \frac{c|t-s|^{1+\beta}}{\epsilon^\alpha}$ for $\forall \epsilon > 0$. Thus $lim_{s\to t}\mathbb{P}(|X_t - X_s| \geq \epsilon) = 0$, and hence $X_s \to X_t$ in probability as $s \to t$ for $\forall t \in [0, T_0]$.
Second, set $A_n = \{\max_{1\leq k\leq 2^n T_0} |X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}| \leq c^{\frac{1}{\alpha}} 2^{-\gamma n}\}$,then we have

$$\begin{aligned}
\mathbb{P}(A_n^C) &= \mathbb{P}(\max_{1\leq k\leq 2^n T_0} |X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}| > c^{\frac{1}{\alpha}} 2^{-\gamma n}) \\
&= \mathbb{P}(\bigcup_{k=1}^{2^n T_0} \{|X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}| > c^{\frac{1}{\alpha}} 2^{-\gamma n}\}) \\
&\leq \sum_{k=1}^{2^n T_0} \mathbb{P}(|X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}| > c^{\frac{1}{\alpha}} 2^{-\gamma n}) \\
&\leq \sum_{k=1}^{2^n T_0} \frac{\mathbb{E}[|X_{\frac{k}{2^n}} - X_{\frac{k-1}{2^n}}|^{1+\beta}]}{c 2^{-\gamma\alpha n}} \\
&= T_0 2^{n(\alpha\gamma - \beta)}
\end{aligned} \tag{1}$$

Thus $\sum_{n=1}^\infty \mathbb{P}(A_n^C) < \infty$ for $\gamma < \frac{\beta}{\alpha}$.And hence $\mathbb{P}(\liminf_{n\to\infty} A_n) = 1$. We set $\Omega^* = \liminf_{n\to\infty} A_n$, then $\Omega^* \in \mathcal{F}$ and $\mathbb{P}(\Omega^\star) = 1$. For $\forall \omega \in \Omega^\star$, there exists $n_0(\omega)$ s.t. for $\forall n \geq n_0(\omega)$ we have $\max_{1\leq k\leq 2^n T_0} |X_{\frac{k}{2^n}}(\omega) - X_{\frac{k-1}{2^n}}(\omega)| \leq c^{\frac{1}{\alpha}} 2^{-\gamma n}$. Denote $D = \cup_{n=1}^\infty D_n$, where $D_n = \{\frac{k}{2^n} : k = 0, 1, 2, \cdots 2^n T_0\}$. D is contable and dense in $[0, T_0]$. Fix $\omega \in \Omega^\star$, for $\forall n \geq n_0(\omega)$ and $\forall s, t \in D_m$ with $|t - s| < 2^{-n}$, we claim that

$$|X_t(\omega) - X_s(\omega)| \leq 2c^{\frac{1}{\alpha}} \sum_{j=n+1}^\infty 2^{-\gamma j}$$

By induction, when $m = n + 1$, take $t = \frac{k}{2^{n+1}}$, s=$\frac{k-1}{2^{n+1}}$, then by (1), it is true. Assume it is true for $m = n + 1, \cdots, M - 1$, now consider $m = M$, $s, t \in D_M$. Take $s' = \min\{u \in D_{M-1} : u \geq s\}$,

$t^{'} = \max\{u \in D_{M-1} : u \leq t\}$. Then $|s^{'} - s|, |t - t^{'}| \leq 2^{-M}$ and $s <= s^{'} \leq t^{'} \leq t$. Thus

$$\begin{aligned}|X_t - X_s| &\leq |X_t - X_{t'}| + |X_{t'} - X_{s'}| + |X_{s'} - X_s| \\ &\leq c^{\frac{1}{\alpha}} 2^{-\gamma M} + 2c^{\frac{1}{\alpha}} \sum_{j=n+1}^{M-1} 2^{-\gamma j} + c^{\frac{1}{\alpha}} 2^{-\gamma M} \\ &= 2c^{\frac{1}{\alpha}} \sum_{j=n+1}^{M} 2^{-\gamma j}\end{aligned}$$

Set $\delta(\omega) = 2^{-n_0(\omega)}$, then for $\forall s, t \in D$ with $|t - s| < \delta(\omega)$, we can choose $m, s, t \in D_m$ and choose $n \geq n_0(\omega)$ s.t. $2^{-(n+1)} \leq |t - s| < 2^{-n}$, then

$$|X_t(\omega) - X_s(\omega)| \leq 2c^{\frac{1}{\alpha}} \sum_{j=n+1}^{m} 2^{-\gamma j} \leq \frac{2c^{\frac{1}{\alpha}}}{1 - 2^{-\gamma}} |t - s|^{\gamma}$$

Define

$$Y_t(\omega) = \begin{cases} X_t(\omega), & \omega \in \Omega^\star, t \in D \\ 0, & \omega \notin \Omega^\star, t \in [0, T_0] \\ \lim_{s \in D, s \to t} X_s(\omega), & \omega \in \Omega^\star, t \in [0, T_0] \setminus D \end{cases}$$

The continuity of $X_t(\omega)$ w.r.t t is uniform, hence $Y_t(\omega)$ is well defined and continuous. And Y is what we desire. $\square$

Definition 3.8. On $(\Omega, \mathcal{F}, \mathbb{P})$, $\{\mathcal{F}_t : t \geq 0\}$ is a **filtration** if $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for $\forall 0 \leq s \leq t$. And $\mathcal{F}_\infty = \sigma(\cup_{t \geq 0} \mathcal{F}_t)$, $\mathcal{F}_t^X = \sigma(X_s : 0 \leq s \leq t)$.

Definition 3.9. X is **adapted** to $\{\mathcal{F}_t\}$ if X_t is $\mathcal{F}_t$-measurable for $\forall t \geq 0$.

Remark. Y is modification of X, X is adapted to $\{\mathcal{F}_t\}$, then Y is also adapted to $\{\mathcal{F}_t\}$.
$\forall A$ with $\mathbb{P}(A) = 0$, then $A \in \mathcal{F}_0$.

Notation: $\mathcal{F}_{t-} = \sigma(\cup_{s<t} \mathcal{F}_s)$, $\mathcal{F}_{t+} = \cap_{s>t} \mathcal{F}_s$. $\{\mathcal{F}_t\}$ is said to be right continuous (RC) if $\mathcal{F}_{t+} = \mathcal{F}_t$ for $\forall t \geq 0$.

Definition 3.10. X is **measurable** if $\forall A \in \mathcal{B}(\mathbb{R}^d)$, $\{(\omega, t) \in \Omega \times [0, \infty) : X_t(\omega) \in A\} \in \mathcal{F} \otimes \mathcal{B}([0, \infty))$.

Definition 3.11. X is **progressively measurable** if $\forall t \geq 0$, $\forall A \in \mathcal{B}(\mathbb{R}^d)$, $\{(\omega, t) \in \Omega \times [0, t] : X_t(\omega) \in A\} \in \mathcal{F} \otimes \mathcal{B}([0, t])$.

Propsition 3.1. (1)Progressively measurable $\Longrightarrow$ measurable and adapted.
(2)If a process is measurable and adapted, then it has a modification which is progressively measurable.

Propsition 3.2. If process X is adapted to $\{\mathcal{F}_t\}$ and every sample path of X is right continuous, then X is progressively measurable.

Proof: For $\forall t \geq 0, n \in \mathbb{N}, 0 \leq s \leq t$, define

$$X_s^{(n)}(\omega) = X_{\frac{(k+1)t}{2^n}}(\omega), \text{ for } \frac{kt}{2^n} < s \leq \frac{(k+1)t}{2^n},\ k \in \mathbb{N}$$

Then $(\omega, s) \mapsto X_s^{(n)}(\omega)$ is $\mathcal{F}_t \otimes \mathcal{B}([0,t])$ measurable since X is adapted. And for $\forall \omega \in \Omega$, $X_s^{(n)}(\omega) \to X_s(\omega)$ as $n \to \infty$ for $\forall s \in [0,t]$ by right continuity of X. Thus $(\omega, s) \mapsto X_s(\omega)$ is also $\mathcal{F}_t \otimes \mathcal{B}([0,t])$ measurable. $\square$

Definition 3.12. A **ramdom time** T is a r.v. with values in $[0, \infty]$. Then $\sigma(X_T) = \{\{X_T \in A\}, \{X_T \in A\} \cup \{T = \infty\} : A \in \mathcal{B}(\mathbb{R}^d)\}$

Definition 3.13. If a ramdom time T satisfying $\{T \leq t\} \in \mathcal{F}_t$, $\forall t \geq 0$, it is a **stopping time**. T is an **optional time** if $\{T < t\} \in \mathcal{F}_t$, $\forall t \geq 0$.

Propsition 3.3. (1)A stopping time is always a optional time. (2) If $\{\mathcal{F}_t\}$ is RC, then any optional time is always a stopping time.

Proof:
(1)For $\forall t \geq 0$, $\{T < t\} = \cup_{n=1}^{\infty}\{T \leq t - \frac{1}{n}\} \in \mathcal{F}_t$ since $\{\mathcal{F}_t\}$ is increasing.
(2)For $\forall t \geq 0$, $\{T \leq t\} = \cap_{n=1}^{\infty}\{T < t + \frac{1}{n}\} \in \mathcal{F}_{t+} = \mathcal{F}_t$ since $\{\mathcal{F}_t\}$ is RC. $\square$

Corollary 3.1. Let $\mathcal{G}_t = \mathcal{F}_{t+}$, then T is an optional time of $\{\mathcal{F}_t\}$ $\Longleftrightarrow$ T is a stopping time of $\{\mathcal{G}_t\}$.

Example 3.1. Let X be an adapted process relative to $\{\mathcal{F}_t\}$, $A \in \mathcal{B}(\mathbb{R}^d)$. Set $T(\omega) = \inf\{t \geq 0 : X_t(\omega) \in A\}$ which is known as hitting time.

- If A is open and X has RC sample paths, then T is a optional time.
- If A is closed and X has continuous sample paths, then T is an stopping time.

Propsition 3.4. If T, S are two stopping times, then $T \wedge S$, $T \vee S$, $T + S$ are stopping times.

Proof:
(1) For $\forall t \geq 0$, $\{T \wedge S \leq t\} = \{T \leq t\} \cup \{S \leq t\} \in \mathcal{F}_t$.
(2) For $\forall t \geq 0$, $\{T \vee S \leq t\} = \{T \leq t\} \cap \{S \leq t\} \in \mathcal{F}_t$.
(3) For $\forall t \geq 0$, $\{T + S \leq t\} = \{T = 0, S \leq t\} \cup \{S = 0, T \leq t\} \cup \{0 < T, S < t, T + S \leq t\}$. And

$$\{0 < T, S < t, T + S > t\} = \cup_{q \in \mathbb{Q}, 0 < q < t}\{q < T < t,\ t - q < S < t\} \in \mathcal{F}_t$$

$\{q < T < t\} = \{r < T\} \cap \{T \le s\} = \{T \le q\}^C \cap \{T \le t\} \in \mathcal{F}_t$.
Thus $\{0 < T < t,\, 0 < S < t,\, T + S \le t\} = \{0 < T, S < t\} \cap \{0 < T, S < t,\, T + S > t\}^C \in \mathcal{F}_t$.
Hence $\{T + S \le t\} \in \mathcal{F}_t$. $\square$

Propsition 3.5. If T, S are two optional times, then $T \wedge S$, $T \vee S$ are optional times.

Propsition 3.6. If T, S are two optional times, then so is $T + S$. If additionally either $T, S > 0$ or $T > 0$ is a stopping time, then $T + S$ is a stopping time.

Proof:
(1) If $T, S > 0$, then $\{T + S \le t\} = \{0 < T.S < t,\, T + S \le t\}$

$$\{0 < T, S < t,\, T + S > t\} = \bigcup_{q_1, q_2 \in \mathbb{Q}, 0 < q_2 < q_1 < t} \{q_1 \le T < t,\, t - q_2 \le S < t\} \in \mathcal{F}_t$$

(2) If T is a stopping time and $S > 0$, then $\{T + S \le t\} = \{0 < T \le t,\, 0 \le S < t,\, T + S \le t\}$ and $\{0 < T \le t, 0 \le S < t\} \in \mathcal{F}_t$, hence

$$\{0 < T \le t,\, 0 \le S < t,\, T + S > t\} = \bigcup_{q_1, q_2 \in \mathbb{Q}, 0 < q_2 < q_1 < t} \{q_1 < T < t,\, t - q_2 < S < t\} \in \mathcal{F}_t$$

$\square$

Propsition 3.7. If $\{T_n,\, n \in \mathbb{N}\}$ is a sequence of optional times, then $\sup_{n \in \mathbb{N}}$, $\inf$, $\limsup_{n \to \infty}$, $\liminf$ of T_n are all optional times.

Proof: $\{\sup T_n < t\} = \cap \{T_n < t\} \in \mathcal{F}_t$. $\square$

Definition 3.14. T is a stopping time, then $\mathcal{F}_T = \{A :\, A \cap \{T \le t\} \in \mathcal{F}_t,\, \forall t \ge 0\}$ is called **pre-T σ-algebra**.

Propsition 3.8. S, T are stopping times, then:
(1)$\forall A \in \mathcal{F}_S, A \cap \{S \le T\} \in \mathcal{F}_T$
(2)$\{T < S\}$, $\{T \le S\}$, $\{T = S\} \in \mathcal{F}_T \cap \mathcal{F}_S$
(3)$\mathcal{F}_{T \wedge S} = \mathcal{F}_T \cap \mathcal{F}_S$

Proof:
(1) For $\forall t \ge 0$, $A \cap \{S \le T\} \cap \{T \le t\} = A \cap \{S \le t\} \cap \{T \le t\} \cap \{S \wedge t \le T \wedge t\} \in \mathcal{F}_t$, since both $S \wedge t$, $T \wedge t \in \mathcal{F}_t$.
(3) $T \wedge S \le T, S$, and $T \wedge S$ is a stopping time, thus by (1), $\mathcal{F}_{T \wedge S} \subset \mathcal{F}_T$ and $\mathcal{F}_S$. Conversely, for $\forall A \in \mathcal{F}_T \cap \mathcal{F}_S$, and $\forall t \ge 0$, we have

$$A \cap \{T \wedge S \le t\} = (A \cap \{T \le t\}) \cap (A \cap \{S \le t\}) \in \mathcal{F}_t$$

Thus $A \in \mathcal{F}_{T\wedge S}$. Hence $\mathcal{F}_{T\wedge S} = \mathcal{F}_T \cap \mathcal{F}_S$.
(2) Let $A = \Omega$, then by (1), $\{S \le T\} \in \mathcal{F}_T$. Thus, $\{S > T\} \in \mathcal{F}_T$. $\{S > T\} = \{S > T \wedge S\} \in \mathcal{F}_{T\wedge S} \subset \mathcal{F}_S \cap \mathcal{F}_T$. Then $\{S \le T\} \in \mathcal{F}_{T\wedge S} \subset \mathcal{F}_S \cap \mathcal{F}_T$. Hence the conclusion holds. □

Propsition 3.9. *X is progressively measurablep process, T is a finite stopping time, then X_T is $\mathcal{F}_T$ measurable and $\{X_{T\wedge t} : t \ge 0\}$ is progressively measurable w.r.t. $\{\mathcal{F}_{t\wedge T}\}$.*

Notation:
T is a optional time, define $\mathcal{F}_{T+} = \{A : A \cap \{T \le t\} \in \mathcal{F}_{t+}, \forall t \ge 0\}$ called post-T σ-algebra.

Remark. If T is a stopping time, then $\mathcal{F}_T \subset \mathcal{F}_{T+}$.

Definition 3.15. $\{\mathcal{F}_t\}$ is said to satisfy **usual conditions** if it is RC and $\mathcal{F}_0$ contains all P-null sets in $\mathcal{F}$.

Propsition 3.10. *$\{\mathcal{F}_t\}$ satisfies usual conditions, X is an adapted process with RCLL sample paths, then there exists a sequence $\{T_n : n \in \mathbb{N}\}$ of stopping times s.t. $\{(\omega, t) : X_t(\omega) \ne X_{t-}(\omega)\} = \cup_{n=1}^{\infty}\{(\omega, t) : T_n(\omega) = t\}$. $\{T_n\}$ exhausts the jumps of X.*

Definition 3.16. X is an adapted precess, $X_t \in L^1$, then
(1) if $\forall 0 < s < s < \infty, \mathbb{E}[X_t|\mathcal{F}_s] \ge X_s$, then X is a **submartingale**.
(2) if $\forall 0 < s < s < \infty, \mathbb{E}[X_t|\mathcal{F}_s] \le X_s$, then X is a **supermartingale**.
(3) if $\forall 0 < s < s < \infty, \mathbb{E}[X_t|\mathcal{F}_s] = X_s$, then X is a **martingale**.
If inaddition, $X_\infty \in L^1, X_\infty \in \mathcal{F}_\infty$, and $\forall s \ge 0, \mathbb{E}[X_\infty|\mathcal{F}_s] \ge X_s$, then X is a submartingale extended to infinity. Similar definition holds for supermartingale and martingale extended to infinity.

Remark. If X is a submartingale, and $\{t_n\}$ is decreasing, non-negative, then $\{X_{t_n}\}$ is a backward submartingale w.r.t. $\{\mathcal{F}_{t_n}\}$.

Propsition 3.11. *X is a submartingale, ϕ is an increasing convex function s.t. $\phi(X_t) \in L^1$ for $\forall t$, then $\phi(X_t)$ is also a submartingale. additionally, if X is a martingale, then $\phi(X_t)$ is a submartingale and we do not need ϕ to be increasing.*

Definition 3.17. X is a real valued process. For $a < b$ and $I \subset [0, \infty)$, define $\tau_1(\omega) = \inf\{t \in I : X_t(\omega) \le a\}$, $\sigma_1(\omega) = \inf\{t \in I : t \ge \tau_1(\omega), X_t(\omega) > b\}$, and for $n \ge 1$, define

$$\tau_{n+1}(\omega) = \inf\{t \in I : t \ge \sigma_n(\omega), X_t(\omega) \le a\}$$

$$\sigma_{n+1}(\omega) = \inf\{t \in I : t \ge \tau_{n+1}(\omega), X_t(\omega) > b\}$$

where we set $\inf \emptyset = \infty$. Then $\{\tau_n\}$ and $\{\sigma_n\}$ are called **upcrossing time** of interval [a,b] by X on I. And the number of upcrossings of $[a, b]$ by X on I is defined as

$$U_I([a, b]; X(\omega)) = \sup\{n \ge 1 : \sigma_n < \infty\}$$

Theorem 3.2. X is a submartingale whose sample paths are RC. let $I \subset [0, \infty)$ be a compact interval, $a < b$, $\lambda > 0$, then we have the following properties.
(1) First submartingale inequality:

$$\lambda \mathbb{P}(\sup_{t \in I} X_t \geq \lambda) \leq \mathbb{E}[X^+_{\sup I}]$$

(2) Second submartingale inequality:

$$\lambda \mathbb{P}(\inf_{t \in I} X_t \leq -\lambda) \leq \mathbb{E}[X^-_{\sup I}] - \mathbb{E}[X_{\inf I}]$$

(3) Upcrossing inequality:

$$(b-a)\mathbb{E}[U_I([a,b];X)] \leq \mathbb{E}[X^+_{\sup I}] + |a|$$

(4) Doob's maximal inequality:

$$\forall p > 1,\ \mathbb{E}[\sup_{t \in I} |X_t|^p] \leq (\frac{p}{p-1})^p \mathbb{E}[|X_{\sup I}|^p]$$

if $X_t > 0$ for $\forall t \in I$.
(5) Regularity of sample paths of X: (i) Almost every sample path is bounded on compact intervals and admits left limits almost everywhere on $t \in (0, \infty)$. (ii) If $\mathcal{F}_t$ satisfies usual conditions the jumps of X are exhausted by a sequence of stopping times.

Lemma 3.1. If $X = \{X_n : 1 \leq n \leq N\}$ is a submartingale, then for $\forall \lambda > 0$ we have:
(1) $\lambda \mathbb{P}(\max_{1 \leq k \leq N} X_k \geq \lambda) \leq \mathbb{E}[X^+_N]$
(2) $\lambda \mathbb{P}(\min_{1 \leq k \leq N} X_k \leq -\lambda) \leq \mathbb{E}[X^+_N] - \mathbb{E}[X_1] - \mathbb{E}[X_N \mathbb{1}_{\{\min_{1 \leq k \leq N} X_k > -\lambda\}}] \leq e[X^+_N] - \mathbb{E}[X_1]$

Proof:
(1) Set $\alpha = \min\{X_n \geq \lambda\}$ with $\min \emptyset = N$. Then α is an stopping time since $\{\alpha \leq k\} = \cup_{i=1}^k \{X_i \geq \lambda\} \in \mathcal{F}_k$. Thus X_α, X_N is a $\{\mathcal{F}_\alpha, \mathcal{F}_N\}$-submartingale. Hence, for $forallk < N$, $\{\alpha \leq k\} \subset \{\max X_n \geq \lambda\}$. Therefore, $\{\alpha \leq k\} \cap \{\max X_n \geq \lambda\} = \{\alpha \leq k\} \in \mathcal{F}_\alpha$. Thus, $\{\max X_n \geq \lambda\} \in \mathcal{F}_\alpha$. On $\{\max X_n \geq \lambda\}$, we have $X_\alpha \geq \lambda$. Thus

$$\begin{aligned}
\lambda \mathbb{P}(\max_{1 \leq k \leq N} X_k \geq \lambda) &= \mathbb{E}[\lambda \mathbb{1}_{\{\max X_n \geq \lambda\}}] \\
&\leq \mathbb{E}[X_\alpha \mathbb{1}_{\{\max X_n \geq \lambda\}}] \\
&\leq \mathbb{E}[\mathbb{E}[X_N | \mathcal{F}_\alpha] \mathbb{1}_{\{\max X_n \geq \lambda\}}] \\
&= \mathbb{E}[X_N \mathbb{1}_{\{\max X_n \geq \lambda\}}] \\
&\leq \mathbb{E}[X^+_N]
\end{aligned}$$

(2) Set $\beta = \min\{X_n \le -\lambda\}$ with $\min\emptyset = N$. Then β is a stopping time since $\{\beta \le k\} = \cup_{i=1}^k \{X_i \le -\lambda\} \in \mathcal{F}_k$.

$$\begin{aligned}
\mathbb{E}[X_1] &\le \mathbb{E}[X_\beta] \\
&= \mathbb{E}[X_\beta \mathbb{1}_{\{\beta \le N-1\}}] + \mathbb{E}[X_\beta \mathbb{1}_{\{\beta = N\} \cap \{\min X_n > -\lambda\}}] + \mathbb{E}[X_\beta \mathbb{1}_{\{\beta = N\} \cap \{\min X_n \le -\lambda\}}] \\
&= \mathbb{E}[X_\beta \mathbb{1}_{\{\beta \le N-1\}}] + \mathbb{E}[X_N \mathbb{1}_{\{\beta = N\} \cap \{\min X_n > -\lambda\}}] + \mathbb{E}[X_N \mathbb{1}_{\{\beta = N\} \cap \{\min X_n \le -\lambda\}}] \\
&\le \mathbb{E}[-\lambda \mathbb{1}_{\{\beta \le N-1\}}] + \mathbb{E}[X_N \mathbb{1}_{\{\beta = N\} \cap \{\min X_n > -\lambda\}}] + \mathbb{E}[X_N \mathbb{1}_{\{\min X_n > -\lambda\}}] \\
&\le -\lambda \mathbb{P}(\beta \le n-1) - \lambda \mathbb{P}(\{\beta = N\} \cap \{\min X_n \le \lambda\}) + \mathbb{E}[X_N \mathbb{1}_{\{\min X_n > -\lambda\}}] \\
&= -\lambda \mathbb{P}(\min X_n \le -\lambda) + \mathbb{E}[X_N \mathbb{1}_{\{\min X_n > -\lambda\}}]
\end{aligned}$$

Thus, $\lambda \mathbb{P}(\min X_n \le -\lambda) \le \mathbb{E}[X_N] - \mathbb{E}[X_1] - \mathbb{E}[X_N \mathbb{1}_{\{\min X_n > -\lambda\}}] \le \mathbb{E}[X_N^+] - \mathbb{E}[X_1]$. □

Proof of (1) and (2) in Theorem:

Choose $\{F_N : N \in \mathbb{N}\}$ to be an increasing sequence of finite set s.t. $\inf I, \sup I \in F_N$. $F = \cup_{N=1}^\infty F_N = \{\inf I, \sup I\} \cup (I \cap \mathbb{Q})$. By lemma 3.1, we have:

$$\lambda \mathbb{P}(\max_{t \in F_N} X_t > \lambda) \le \mathbb{E}[X_{\sup I}^+]$$

$$\lambda \mathbb{P}(\min_{t \in F_N} X_t < -\lambda) \le \mathbb{E}[X_{\sup I}^+] - \mathbb{E}[X_{\inf I}]$$

Since $F_N \subset F_{N+1}$, we have $\{\max_{t \in F_N} X_t\} \subset \{\max_{t \in F_{N+1}} X_t\}$. Let $N \to \infty$, we have

$$\lambda \mathbb{P}(\sup_{t \in F} X_t > \lambda) \le \mathbb{E}[X_{\sup I}^+]$$

$$\lambda \mathbb{P}(\inf_{t \in F} X_t < -\lambda) \le \mathbb{E}[X_{\sup I}^+] - \mathbb{E}[X_{\inf I}]$$

F is dense in I, and X is RC, thus $\sup_{t \in F} X_t = \sup_{t \in I} X_t$, $\inf_{t \in F} X_t = \inf_{t \in I} X_t$. Beacuse $F \subset I$, we have $\{sup_{t \in F} X_t > \lambda\} \subset \{\sup_{t \in I} X_t > \lambda\}$, $\exists t_0(\omega)$ s.t. $X_{t_0} > \lambda$. Choose $\{s_n\} \subset F$ s.t. $s_n \to t_0$, and s_n is decreasing to t_0, then $X_{s_n}(\omega) \to X_{t_0}(\omega) > \lambda$ as $n \to \infty$. Thus, for sufficiently large n, $X_{s_n}(\omega) > \lambda$, i.e. $\omega \in \{\sup_{t \in F} X_t > \lambda\}$. Therefore, for $\forall \lambda > 0$, we have

$$\lambda \mathbb{P}(\sup_{t \in I} X_t > \lambda) \le \mathbb{E}[X_{\sup I}^+]$$

$$\lambda \mathbb{P}(\inf_{t \in I} X_t < -\lambda) \le \mathbb{E}[X_{\sup I}^+] - \mathbb{E}[X_{\inf I}]$$

Choose $\lambda = \lambda_n$, to be a increasing sequence to λ,then we have

$$\lambda_n \mathbb{P}(\sup_{t \in I} X_t > \lambda_n) \le \mathbb{E}[X_{\sup I}^+]$$

Since $\cap_{n=1}^{\infty}\{\sup_{t\in I} X_t > \lambda_n\} = \{\sup_{t\in I} X_t \geq \lambda\}$, we have

$$\lambda\mathbb{P}(\sup_{t\in I} X_t \geq \lambda) \leq \mathbb{E}[X^+_{\sup I}]$$

Similarly, we have

$$\lambda\mathbb{P}(\inf_{t\in I} X_t \leq -\lambda) \leq \mathbb{E}[X^+_{\sup I}] - \mathbb{E}[X_{\inf I}]$$

□

Lemma 3.2. If $X = \{X_n\}$ is a submartingale w.r.t. $\{\mathcal{F}_n\}$, then for $a < b$, we have:

$$(b-a)\mathbb{E}[U_n([a,b];X)] \leq \mathbb{E}[(X_N - a)^+] \leq \mathbb{E}[X_n^+] + |a|$$

Proof:

Set τ_k, σ_k be the upcrossing times of $[a,b]$ by X. Set $\min\emptyset = N$, $U_N([a,b];X) = M$.
Then $\{X_1, X_{\tau_1}, X_{\sigma_1}, \cdots, X_N\}$ is a submartingale w.r.t. $\{\mathcal{F}_1, \mathcal{F}_{\tau_1}, \mathcal{F}_{\sigma_1}, \cdots, \mathcal{F}_N\}$, and $X_{\sigma_k} - X_{\tau_k} \geq b-a$ on $\{\sigma_k < \infty\}$. Thus, $(b-a)U_N([a,b];X) \leq \sum_{k=1}^{M}(X_{\sigma_k} - X_{\tau_k})$, take expectation on both sides, we have

$$\begin{aligned}(b-a)\mathbb{E}[U_N([a,b];X)] &\leq \sum_{k=1}^{M}(\mathbb{E}[X_{\sigma_k}] - \mathbb{E}[X_{\tau_k}]) \\ &= \mathbb{E}[X_{\sigma_M}] + \mathbb{E}[\sum_{j=1}^{M-1}(X_{\tau_j} - X_{\sigma_j+1})] - \mathbb{E}[X_{\tau_1}] \\ &\leq \mathbb{E}[X_{\sigma_M}] - \mathbb{E}[X_{\tau_1}]\end{aligned}$$

Since $\{(X_n - a)^+\}$ is also a submartingale, apply the same argument to it, then lemma is proved.
□

Proof of (3) in Theorem:

Choose $\{F_N : N \in \mathbb{N}\}$ to be an increasing sequence of finite set s.t. $\inf I, \sup I \in F_N$. $F = \cup_{N=1}^{\infty} F_N = \{\inf I, \sup I\} \cup (I \cap \mathbb{Q})$. By lemma 3.2, we have:

$$(b-a)\mathbb{E}[U_{F_N}([a,b];X)] \leq \mathbb{E}[X^+_{\sup I}] + |a|$$

Since $F_N \subset F_{N+1}$, we have $U_{F_N}([a,b];X) \leq U_{F_{N+1}}([a,b];X)$. Let $N \to \infty$, we have

$$(b-a)\mathbb{E}[U_F([a,b];X)] \leq \mathbb{E}[X^+_{\sup I}] + |a|$$

F is dense in I, and X is RC, thus $U_F([a,b];X) = U_I([a,b];X)$. Therefore, we have

$$(b-a)\mathbb{E}[U_I([a,b];X)] \leq \mathbb{E}[X^+_{\sup I}] + |a|$$

□

Proof of (5) in Theorem:
From (1) and (2), $\forall I$ compact interval, we have $-\infty < \inf_{t\in I} X_t < \sup_{t\in I} X_t < \infty$ a.s.. Notice that $\{\omega \in \Omega : \exists t \in I \text{ s.t.} \liminf_{s\to t^-} X_s(\omega) < \limsup_{s\to t^-} X_s(\omega)\} \subset \bigcup_{a<b,a,b\in\mathbb{Q}}\{\omega \in \Omega : U_I([a,b]; X(\omega)) = \infty\}$. Then by (3), we have $\mathbb{P}(\exists t \in I \text{ s.t.} \liminf_{s\to t^-} X_s < \limsup_{s\to t^-} X_s) = 0$. Thus, almost every sample path admits left limits almost everywhere on any compact interval. Set $I = [0,n]$, $n \in \mathbb{N}$, Then $\{\forall t \in I_n, \exists \lim_{s\to t^-} X_s\} \subset \{\forall t \in I_{n+1}, \exists \lim_{s\to t^-} X_s\}$, which implies that $\mathbb{P}(\forall t \in [0,\infty), \exists \lim_{s\to t^-} X_s) = 1$. □

Propsition 3.12. Let X be a submartingale w.r.t. $\{\mathcal{F}_t\}$, then:
(1) $\exists \Omega^\star \in \mathcal{F}$ with $\mathbb{P}(\Omega^\star) = 1$, s.t. for $\forall \omega \in \Omega^\star$, the limits $\lim_{s\to t^-} X_s(\omega) := X_{t^-}$ exist for $\forall t \in (0,\infty)$, and $\lim_{s\to t^+} X_s(\omega) := X_{t^+}$ exists for $\forall t \in [0,\infty)$.
(2) $\mathbb{E}[X_{t^+}|\mathcal{F}_t] \geq X_t$, $\forall t \geq 0$; $\mathbb{E}[X_t|\mathcal{F}_{t^-}] \geq X_{t^-}$, $\forall t > 0$.
(3) $\{X_{t^+} : t \geq 0\}$ is a submartingale with $\mathbb{P}$-a.s. RC sample paths.

Theorem 3.3. Let $\{\mathcal{F}_t\}$ be a filtration satisfying usual conditions, and X be a submartingale w.r.t. $\{\mathcal{F}_t\}$. Then $\exists$ a RC modification $\widetilde{X}$ of $X \iff t \mapsto \mathbb{E}[X_t]$ is right continuous. Moreover, the RC modification can be chosen to be RCLL.

Proof:
"$\Rightarrow$": If $\widetilde{X}$ is a RC modification of X, then for any decreasing sequence $\{t_n\}$ converging to t, $\{X_{t_n}\}$ is a backward submartingale w.r.t. $\{\mathcal{F}_{t_n}\}$. Thus, $\mathbb{E}[X_{t_n}]$ is decreasing and $\mathbb{E}[X_{t_n}] \geq \mathbb{E}[x_t]$. So $\lim_{n\to\infty} \mathbb{E}[X_{t_n}]$ exists. Therefore $\{X_{t_n}\}$ is uniformly integrable. Since $\widetilde{X}$ is RC, and by definition of modification, we have $\mathbb{P}(\widetilde{X}_{t_n} = X_{t_n}, \forall n) = 1$, thus $\lim_{n\to\infty} \widetilde{X}_{t_n} = X_t$ a.s.. Then we have $\lim_{n\to\infty} \mathbb{E}[X_{t_n}] = \mathbb{E}[X_t]$. Hence, $t \mapsto \mathbb{E}[X_t]$ is right continuous.
"$\Leftarrow$": If $t \mapsto \mathbb{E}[X_t]$ is right continuous, then by Proposition, $\{X_{t^+} : t \geq 0\}$ is a submartingale w.r.t. $\{\mathcal{F}_{t^+}\}$ with RC sample paths. Since $\{\mathcal{F}_t\}$ satisfies usual conditions, we have $\mathcal{F}_{t^+} = \mathcal{F}_t$. Thus, $\{X_{t^+} : t \geq 0\}$ is a submartingale w.r.t. $\{\mathcal{F}_t\}$. For $\forall t \geq 0$, since $X_t \in L^1$, we have $\mathbb{E}[X_{t^+}] = \lim_{s\to t^+} \mathbb{E}[X_s] = \mathbb{E}[X_t]$. Thus, X_{t^+} is a modification of X. Hence, the conclusion holds.
□

Theorem 3.4 (Submartingale Convergence Theorem)**.** Let X be a submartingale w.r.t. $\{\mathcal{F}_t\}$. If $\sup_{t\geq 0} \mathbb{E}[X_t^+] < \infty$, then $\exists X_\infty \in L^1$, s.t. $X_t \to X_\infty$ a.s. and in L^1 as $t \to \infty$.

Proof:
For $\forall n \in \mathbb{N}$, by upcrossing inequality in theorem 3.2, we have

$$(b-a)\mathbb{E}[U_n([a,b]; X)] \leq \mathbb{E}[X_n^+] + |a| \leq \sup_{t\geq 0} \mathbb{E}[X_t^+] + |a| < \infty$$

Thus, $\mathbb{E}[U_n([a,b];X)] \leq \frac{\sup_{t\geq 0}\mathbb{E}[X_t^+]+|a|}{b-a} < \infty$. Let $n \to \infty$, by Monotone Convergence Theorem, we have

$$\mathbb{E}[U_\infty([a,b];X)] \leq \frac{\sup_{t\geq 0}\mathbb{E}[X_t^+]+|a|}{b-a} < \infty$$

Thus, $U_\infty([a,b];X) < \infty$ a.s.. Since $a < b$ are arbitrary, we have $\mathbb{P}(\liminf_{s\to t^-} X_s = \limsup_{s\to t^-} X_s, \forall t > 0) = 1$. Set $X_\infty(\omega) = \lim_{t\to\infty} X_t(\omega)$ on $\{\omega : \lim_{t\to\infty} X_t(\omega) \text{ exists}\}$, and $X_\infty(\omega) = 0$ otherwise. Then $X_t \to X_\infty$ a.s.. Since X_t is a submartingale, by Fatou's lemma, we have

$$\mathbb{E}[|X_\infty|] \leq \liminf_{t\to\infty} \mathbb{E}[|X_t|] \leq \liminf_{t\to\infty} (\mathbb{E}[X_t^+] + \mathbb{E}[X_t^-]) \leq 2\sup_{t\geq 0}\mathbb{E}[X_t^+] < \infty$$

Thus, $X_\infty \in L^1$. Finally, since $\{X_t\}$ is uniformly integrable, we have $X_t \to X_\infty$ in L^1. □

Remark. $X_\infty \in \mathcal{F}_\infty$.

Corollary 3.2. Let X be a RC non-negative supermartingale w.r.t. $\{\mathcal{F}_t\}$. Then $\exists X_\infty \in L^1$, s.t. $X_t \to X_\infty$ a.s. and in L^1 as $t \to \infty$.

Definition 3.18. X is called **potential** if it is a non-negative supermartingale w.r.t. $\{\mathcal{F}_t\}$, and $\mathbb{E}[X_t] \to 0$ as $t \to \infty$.

Remark. Potential process is a supermartingale with last element 0 a.s..

Propsition 3.13. X is a non-negative submartingale w.r.t. $\{\mathcal{F}_t\}$, then the following are equivalent:
(1) $\{X_t : t \geq 0\}$ is uniformly integrable.
(2) X_t converges in L^1 as $t \to \infty$.
(3) X_t converges a.s. to some $X_\infty \in L^1$ and after adding X_∞ to the process, $\{X_t : t \geq 0\}$ is a submartingale extended to infinity.

Remark. $X_t \geq 0$ can be replaced by $X_t \geq Y \in L^1$ a.s..

Theorem 3.5. Optional Sampling Theorem: Let X be a submartingale w.r.t. $\{\mathcal{F}_t\}$, and S, T be two optional times with $S \leq T$ a.s.. Then

$$\mathbb{E}[X_T|\mathcal{F}_{S+}] \geq X_S \quad a.s..$$

In particular, if S is a stopping time, then

$$\mathbb{E}[X_T|\mathcal{F}_S] \geq X_S \quad a.s..$$

Proof:

Step 1: For $n \in \mathbb{N}$, define $S_n = \begin{cases} k2^{-n}, & \text{if } (k-1)2^{-n} \leq S < k2^{-n}, k = 1,2,\cdots \\ \infty, & \text{if } S = \infty \end{cases}$

and $T_n = \begin{cases} k2^{-n}, & \text{if } (k-1)2^{-n} \leq T < k2^{-n}, k = 1, 2, \cdots \\ \infty, & \text{if } T = \infty \end{cases}$. Then S_n, T_n are optional times with $S_n \leq T_n$ a.s., and $S_n \downarrow S$, $T_n \downarrow T$ as $n \to \infty$.

$\{S_n \leq t\} = \begin{cases} \{S < t\}, & \text{if } t \in \{k2^{-n} : k = 0, 1, 2, \cdots\} \\ \{S < k2^{-n}\}, & \text{if } k2^{-n} < t < (k+1)2^{-n} \end{cases} \in \mathcal{F}_t$. Thus, S_n is an optional time. Similarly, T_n is also an optional time.

Since X is RC, we have $X_{S_n} \to X_S$, $X_{T_n} \to X_T$ a.s. as $n \to \infty$.

Step 2: Consider it as a discrete time process. Then $\{X_{S_n}, X_{T_n}\}$ is a $\{\mathcal{F}_{S_n}, \mathcal{F}_{T_n}\}$-submartingale. For $\forall A \in \mathcal{F}_{S+}$, we have

$$\begin{aligned} \mathbb{E}[X_{T_n} \mathbb{1}_A] &\geq \mathbb{E}[\mathbb{E}[X_{T_n} | \mathcal{F}_{S_n}] \mathbb{1}_A] \\ &\geq \mathbb{E}[X_{S_n} \mathbb{1}_A] \end{aligned}$$

Step 3: $\{X_{S_n}\}$ and $\{X_{T_n}\}$ are backward submartingales w.r.t. $\{\mathcal{F}_{S_n}\}$ and $\{\mathcal{F}_{T_n}\}$ respectively. Since $\mathbb{E}[X_{T_n}], \mathbb{E}[X_{S_n}] \geq \mathbb{E}[X_0]$. Thus, $\{X_{S_n}\}$ and $\{X_{T_n}\}$ are uniformly integrable. Thus, $X_{S_n} \to X_S$ and $X_{T_n} \to X_T$ in L^1 as $n \to \infty$. And hence X_T, X_S are integrable. Moreover, $\{X_{S_n} \mathbb{1}_A\}$ and $\{X_{T_n} \mathbb{1}_A\}$ are also uniformly integrable for $\forall A \in \mathcal{F}_{S^+}$.

$$\begin{aligned} \mathbb{E}[\mathbb{E}[X_T | \mathcal{F}_{S^+}] \mathbb{1}_A] &= \mathbb{E}[\mathbb{E}[X_T \mathbb{1}_A] | \mathcal{F}_{S^+}] \\ &= \mathbb{E}[X_T \mathbb{1}_A] \\ &= \lim_{n \to \infty} \mathbb{E}[X_{T_n} \mathbb{1}_A] \\ &\geq \lim_{n \to \infty} \mathbb{E}[X_{S_n} \mathbb{1}_A] \\ &= \mathbb{E}[X_S \mathbb{1}_A] \end{aligned}$$

Choose $A = \{\mathbb{E}[X_T | \mathcal{F}_{S^+}] < X_S\} \in \mathcal{F}_{S^+}$, we have $\mathbb{E}[(\mathbb{E}[X_T | \mathcal{F}_{S^+}] - X_S) \mathbb{1}_A] \geq 0$. Thus, $\mathbb{E}[X_T | \mathcal{F}_{S^+}] \geq X_S$ a.s..

Step 4: If S is a stopping time, then in Step 2, change $\mathbb{E}[X_{T_n} \mathbb{1}_A] \geq \mathbb{E}[X_{S_n} \mathbb{1}_A]$ for $\forall A \in \mathcal{F}_S$, and the rest of the proof is the same. And in Step 3, change $\{X_{S_n} \mathbb{1}_A\}$ and $\{X_{T_n} \mathbb{1}_A\}$ are uniformly integrable for $\forall A \in \mathcal{F}_S$. Hence, the conclusion holds. □

Definition 3.19. An adapted process A is called **increasing** if A_t is non-decreasing and RC in t a.s., $A_0 = 0$ a.s., and $A_t \in L^1$ for $\forall t \geq 0$. A is called integrable if $\mathbb{E}[\lim_{t \to \infty} A_t] < \infty$.

Theorem 3.6 (Doob-Meyer Decomposition Theorem)**.** Let X be a submartingale w.r.t. $\{\mathcal{F}_t\}$ satisfying usual conditions, and whose sample paths are RC. For $\forall a > 0$, define $\phi_a = \{T : T \text{ is a stopping time with } \mathbb{P}(T \leq a) = 1\}$. Assume that $\{X_T : T \in \phi_a\}$ is uniformly integrable for each $a > 0$. Then $\exists$ a decomposition $X = M + A$, where M is a RC martingale w.r.t. $\{\mathcal{F}_t\}$, and A is an increasing process w.r.t. $\{\mathcal{F}_t\}$. Set $\phi = \{T : T \text{ is a stopping time with } \mathbb{P}(T < \infty) = 1\}$

and $\{X_T : T \in \phi\}$ is uniformly integrable, then M is uniformly integrable and A is integrable.

Definition 3.20. X is called a **Markov process** w.r.t. $\{\mathcal{F}_t\}$ if $\forall s < t$, and $\forall$ bounded measurable function f, we have

$$\mathbb{E}[f(X_t)|\mathcal{F}_s] = \mathbb{E}[f(X_t)|X_s] \quad a.s..$$

The equation above is called the **Markov property**. The distribution of X_0 is called the **initial distribution** of the Markov process.

Remark. Markov property can be expressed as follows: $\forall s < t$, and $\forall B \in \mathcal{B}(\mathbb{R})$, we have

$$\mathbb{P}(X_t \in B|\mathcal{F}_s) = \mathbb{P}(X_t \in B|X_s) \quad a.s..$$

Remark. The initial distribution satisfies $\mathbb{P}(X_0 \in B) = \int_B \mu(dx)$ for $\forall B \in \mathcal{B}(\mathbb{R})$, i.e. $\mathbb{P} \circ X_0^{-1} = \mu$. Given X and μ, we may write $\mathbb{P}^\mu$ for $\mathbb{P}$.

Remark. Choose a discrete time set $I = \{t_0, t_1, t_2, \cdots\}$ with $t_0 = 0 < t_1 < t_2 < \cdots$, then we have $\{X_{t_n}\}$ is a discrete time Markov process w.r.t. $\{\mathcal{F}_{t_n}\}$.

Definition 3.21. X is called a **strong Markov process** w.r.t. $\{\mathcal{F}_t\}$ if $\forall$ optional time T, $\forall t \geq 0$, and $\forall A \in \mathcal{B}(\mathbb{R})$, we have

$$\mathbb{P}(X_{T+t} \in A|\mathcal{F}_{T^+}) = \mathbb{P}(X_{T+t} \in A|X_T) \quad a.s. \text{ on } \{T < \infty\}.$$

Notation: $B_b(\mathbb{R}^d)$ is the set of all bounded Borel-measurable functions on $\mathbb{R}^d$, which is a Banach space under the sup-norm $\|f\| = \sup_{x \in \mathbb{R}^d} |f(x)|$ for $f \in B_b(\mathbb{R}^d)$.

Remark. For $\forall$ Markov process X w.r.t. $\{\mathcal{F}_t\}$, and $\forall s < t$, we associate a family of operators $\{T_{s,t}\}$, where $T_{s,t} : B_b(\mathbb{R}) \to$ sapce of bounded functions on $\mathbb{R}^d$ is defined by

$$T_{s,t}f(x) = \mathbb{E}[f(X_t)|X_s = x], \quad \forall f \in B_b(\mathbb{R}^d).$$

Remark. A Markov Porcess is normal if $\forall 0 \leq s < t, \forall f \in B_b(\mathbb{R}^d)$, it holds that $T_{s,t}f \in B_b(\mathbb{R}^d)$.

Theorem 3.7. If X is a normal Markov process w.r.t. $\{\mathcal{F}_t\}$, then $\{T_{s,t}\}$ satisfies the following properties:

(1) $T_{s,t}$ is a linear operator.

(2) $T_{s,s}$ is the identity operator for $\forall s \geq 0$.

(3) $T_{s,t} = T_{s,r}T_{r,t}$. (4) $\forall f \geq 0$, $T_{s,t}f \geq 0$.

(5)

$$\|T_{s,t}\| = \sup_{\substack{f \in B_b(\mathbb{R}^d) \\ f \neq 0}} \frac{\|T_{s,t}f\|}{\|f\|} \leq 1.$$

(6) If $f(x) = 1$ for $\forall x \in \mathbb{R}^d$, then $T_{s,t}f(x) = 1$ for $\forall x \in \mathbb{R}^d$.

Remark. The family of operators $\{T_{s,t}\}$ satisfying properties (1)-(6) is called a Markov evolution.

Proof:

(1) For $\forall f, g \in B_b(\mathbb{R}^d)$, and $\forall a, b \in \mathbb{R}$, we have

$$\begin{aligned} T_{s,t}(af + bg)(x) &= \mathbb{E}[af(X_t) + bg(X_t)|X_s = x] \\ &= a\mathbb{E}[f(X_t)|X_s = x] + b\mathbb{E}[g(X_t)|X_s = x] \\ &= aT_{s,t}f(x) + bT_{s,t}g(x) \end{aligned}$$

(2) For $\forall f \in B_b(\mathbb{R}^d)$, we have

$$T_{s,s}f(x) = \mathbb{E}[f(X_s)|X_s = x] = f(x)$$

(3) For $\forall f \in B_b(\mathbb{R}^d)$, we have

$$\begin{aligned} T_{s,r}T_{r,t}f(x) &= \mathbb{E}[T_{r,t}f(X_r)|X_s = x] \\ &= \mathbb{E}[\mathbb{E}[f(X_t)|X_r]|X_s = x] \\ &= \mathbb{E}[f(X_t)|X_s = x] \\ &= T_{s,t}f(x) \end{aligned}$$

(4) For $\forall f \in B_b(\mathbb{R}^d)$ with $f \geq 0$, we have

$$T_{s,t}f(x) = \mathbb{E}[f(X_t)|X_s = x] \geq 0$$

(5) For $\forall f \in B_b(\mathbb{R}^d)$ with $f \neq 0$, we have

$$\begin{aligned} \|T_{s,t}f\| &= \sup_{x \in \mathbb{R}^d} |T_{s,t}f(x)| \\ &= \sup_{x \in \mathbb{R}^d} |\mathbb{E}[f(X_t)|X_s = x]| \\ &\leq \sup_{x \in \mathbb{R}^d} \mathbb{E}[|f(X_t)||X_s = x] \\ &\leq \sup_{x \in \mathbb{R}^d} \|f\| = \|f\| \end{aligned}$$

Thus, $\|T_{s,t}\| \leq 1$.

(6) For $f(x) = 1$ for $\forall x \in \mathbb{R}^d$, we have

$$T_{s,t}f(x) = \mathbb{E}[1|X_s = x] = 1$$

□

Definition 3.22. The **transition probability** $\mathbb{P}_{s,t}$ of a Markov process X w.r.t. $\{\mathcal{F}_t\}$ with the Markov evolution $\{T_{s,t}\}$ is defined by

$$\mathbb{P}_{s,t}(x,A) = \mathbb{P}(X_t \in A | X_s = x) = (T_{s,t}(\mathbb{1}_A))(x), \quad \forall x \in \mathbb{R}^d, A \in \mathcal{B}(\mathbb{R}^d).$$

Remark. If $\forall x \in \mathbb{R}^d$, $\mathbb{P}_{s,t}(x,\cdot)$ admits a density function $\rho_{s,t}(x,y)$, i.e.

$$\mathbb{P}_{s,t}(x,A) = \int_A \rho_{s,t}(x,y)dy, \quad \forall A \in \mathcal{B}(\mathbb{R}^d),$$

then $\rho_{s,t}(x,y)$ is called the **transition probability density function**.

Remark.

$$\begin{aligned} T_{s,t}(f)(x) &= \mathbb{E}[f(X_t)|X_s = x] \\ &= \int_{\mathbb{R}^d} f(y)\mathbb{P}_{s,t}(x,dy) \\ &= \int_{\mathbb{R}^d} f(y)\rho_{s,t}(x,y)dy \end{aligned}$$

Theorem 3.8 (Champman-Kolmogorov Equation)**.** Let X be a normal Markov process w.r.t. $\{\mathcal{F}_t\}$ with transition probability $\mathbb{P}_{s,t}$. Then $\forall 0 \le r < s < t$, we have:

$$\mathbb{P}_{r,t}(x,A) = \int_{\mathbb{R}^d} \mathbb{P}_{r,s}(x,dy)\mathbb{P}_{s,t}(y,A), \quad \forall x \in \mathbb{R}^d, A \in \mathcal{B}(\mathbb{R}^d).$$

Proof:

Since X is normal, $\forall A \in \mathcal{B}(\mathbb{R}^d)$, $\mathbb{P}_{s,t}(\cdot,A) \in B_b(\mathbb{R}^d)$. Thus we have

$$\begin{aligned} \mathbb{P}_{r,t}(x,A) &= (T_{r,t}(\mathbb{1}_A))(x) \\ &= (T_{r,s}T_{s,t}(\mathbb{1}_A))(x) \\ &= \int_{\mathbb{R}^d} \mathbb{P}_{r,s}(x,dy)(T_{s,t}(\mathbb{1}_A))(y) \\ &= \int_{\mathbb{R}^d} (\int_{\mathbb{R}^d} \mathbb{1}_A(z)\mathbb{P}_{s,t}(y,dz))\mathbb{P}_{r,s}(x,dy) \\ &= \int_{\mathbb{R}^d} \mathbb{P}_{r,s}(x,dy)\mathbb{P}_{s,t}(y,A) \end{aligned}$$

□

Corollary 3.3. If the transition probability $\mathbb{P}_{s,t}$ admits a density function $\rho_{s,t}(x,y)$, then $\forall 0 \le$

$r < s < t$, we have:

$$\rho_{r,t}(x,z) = \int_{\mathbb{R}^d} \rho_{r,s}(x,y)\rho_{s,t}(y,z)dy, \quad \forall x, z \in \mathbb{R}^d.$$

Theorem 3.9. Let $\{\mathbb{P}_{s,t}\}$ be a family of mappings from $\mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d)$ to $[0,1]$ satisfying the following properties:
(1) For $\forall x \in \mathbb{R}^d$, $\mathbb{P}_{s,t}(x,\cdot)$ is a probability measure on $\mathcal{B}(\mathbb{R}^d)$.
(2) For $\forall A \in \mathcal{B}(\mathbb{R}^d)$, $\mathbb{P}_{s,t}(\cdot, A)$ is Borel-measurable.
(3) **Chapman-Kolmogorov Equation**: $\forall 0 \le r < s < t$, we have:

$$\mathbb{P}_{r,t}(x,A) = \int_{\mathbb{R}^d} \mathbb{P}_{r,s}(x,dy)\mathbb{P}_{s,t}(y,A), \quad \forall x \in \mathbb{R}^d, A \in \mathcal{B}(\mathbb{R}^d).$$

Let μ be a fixed probability measure on $\mathcal{B}(\mathbb{R}^d)$, then there exists a Markov process X w.r.t. some filtration $\{\mathcal{F}_t\}$, having μ as its initial distribution and $\{\mathbb{P}_{s,t}\}$ as its transition probability.

Definition 3.23. X is called a **time-homogeneous Markov process** w.r.t. $\{\mathcal{F}_t\}$ if $\forall s < t$, it holds $T_{s,t} = T_{0,t-s}$.

Remark. Time-homogeneous $\iff$ $\mathbb{P}_{s,t}(x,A) = \mathbb{P}_{0,t-s}(x,A)$ for $\forall s < t$, $x \in \mathbb{R}^d$, $A \in \mathcal{B}(\mathbb{R}^d)$.

Remark. In this case, we may write T_t and $\mathbb{P}_t$ instead of $T_{0,t}$ and $\mathbb{P}_{0,t}$ respectively.

Remark. For normal time-homogeneous Markov process, we have $T_{s,t} = T_{t-s}$, and $\{T_t : t \ge 0\}$ is a semigroup, i.e. T_0 is the identity operator, and $T_{s+t} = T_s T_t$ for $\forall s, t \ge 0$.

Definition 3.24. $C_0(\mathbb{R}^d)$ is the set of all continuous functions $f : \mathbb{R}^d \to \mathbb{R}$ vanishing at infinity, i.e. $\forall \epsilon > 0$, $\exists$ a compact set $K_\epsilon \subset \mathbb{R}^d$, s.t. $|f(x)| < \epsilon$ for $\forall x \notin K_\epsilon$. It is a Banach space under $\|f\| = \sup_{x \in \mathbb{R}^d} |f(x)|$ for $f \in C_0(\mathbb{R}^d)$.

Definition 3.25. X is called a **Feller process** w.r.t. $\{\mathcal{F}_t\}$ if it is a normal time-homogeneous Markov process w.r.t. $\{\mathcal{F}_t\}$, and satisfies:
(1) $T_t(C_0(\mathbb{R}^d)) \subset C_0(\mathbb{R}^d)$
(2) $\lim_{t\to 0^+} \|T_t f - f\| = 0$ for $\forall f \in C_0(\mathbb{R}^d)$.
In this case, $\{T_t\}$ is called a Feller semigroup.

Definition 3.26. X is called a **strong Feller process** w.r.t. $\{\mathcal{F}_t\}$ if it is a normal time-homogeneous Markov process w.r.t. $\{\mathcal{F}_t\}$, and satisfies:
(1) $T_t(B_b(\mathbb{R}^d)) \subset C_b(\mathbb{R}^d)$ for $\forall t > 0$.
(2) $\lim_{t\to 0^+} \|T_t f - f\| = 0$ for $\forall f \in C_b(\mathbb{R}^d)$.
In this case, $\{T_t\}$ is called a strong Feller semigroup.

Definition 3.27. **Infinitesimal generator** $\mathcal{A}$ of a Feller semigroup $\{T_t\}$ is defined by

$$\mathcal{A}f = \lim_{t\to 0^+} \frac{T_t f - f}{t}$$

4 Brownian Motion

Definition 4.1. A stochastic process $B = \{B_t : t \geq 0\}$ is called a **standard d-dimensional Brownian motion** w.r.t. $\{\mathcal{F}_t\}$ if:
(1) $B_0 = 0$ a.s..
(2) For $\forall 0 \leq s < t$, $B_t - B_s \sim \mathcal{N}(0, (t-s)I_d)$, where I_d is the $d \times d$ identity matrix, and is independent of $\mathcal{F}_s$.

Remark. If $B = (B^{(1)}, B^{(2)}, \cdots, B^{(d)})$ is standard d-dimensional Brownian motion $\iff$ $\{B^{(i)} : t \geq 0\}$ are independent standard one-dimensional Brownian motions w.r.t. $\{\mathcal{F}_t\}$.

Definition 4.2. Let $X = \{X_t\}$ be a stochastic process w.r.t. $\{\mathcal{F}_t\}$. We say X has **independent increments** if $\forall 0 \leq t_0 < t_1 < \cdots < t_n$, the increments $X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \cdots, X_{t_n} - X_{t_{n-1}}$ are independent.

Remark. Standard Brownian motion has independent increments.

Proof:
For $\forall 0 \leq t_0 < t_1 < \cdots < t_n$, since $B_{t_i} - B_{t_{i-1}} \sim N(0, (t_i - t_{i-1})I_d)$ is independent of $\mathcal{F}_{t_{i-1}}$, we have $B_{t_i} - B_{t_{i-1}}$ is independent of $\{B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \cdots, B_{t_{i-1}} - B_{t_{i-2}}\}$ for $\forall i = 1, 2, \cdots, n$. Thus, the increments are independent. $\square$

Propsition 4.1. X is a stochastic process with independent increments w.r.t. $\{\mathcal{F}_t\}$, then $\forall 0 \leq s < t$, $X_t - X_s$ is independent of $\mathcal{F}_s^X$.

Proof:
For $\forall A \in \mathcal{F}_s^X$, since $\mathcal{F}_s^X = \sigma(X_r : 0 \leq r \leq s)$, by π-λ theorem, we only need to prove that $X_t - X_s$ is independent of $\sigma(X_{r_1}, X_{r_2}, \cdots, X_{r_n})$ for any finite collection $0 \leq r_1, r_2, \cdots, r_n \leq s$. Without loss of generality, assume $r_1 < r_2 < \cdots < r_n$. Then we have

$$\begin{aligned}\sigma(X_{r_1}, X_{r_2}, \cdots, X_{r_n}) &= \sigma(X_{r_1}, X_{r_2} - X_{r_1}, \cdots, X_{r_n} - X_{r_{n-1}}) \\ &\subset \sigma(X_{r_1}, X_{r_2} - X_{r_1}, \cdots, X_s - X_{r_n})\end{aligned}$$

Thus, $X_t - X_s$ is independent of $\sigma(X_{r_1}, X_{r_2}, \cdots, X_{r_n})$. Hence, the conclusion holds. $\square$

Definition 4.3 (Another Definition)**.** A continuous process $B = \{B_t : t \geq 0\}$ is called a standard 1-dimensional Brownian motion w.r.t. $\{\mathcal{F}_t\}$ if
(1) $\forall n \in \mathbb{N}$, $0 = t_0 < t_1 < t_2 < \cdots < t_n$, the joint distribution of $(B_{t_1}, B_{t_2}, \cdots, B_{t_n})$ is certered Gaussian $\sim \mathcal{N}(0, \Gamma)$.
(2) $\mathbb{E}[B_t B_s] = \min\{s, t\}$ for $\forall s, t \geq 0$.

Propsition 4.2. The two definitions of standard Brownian motion are equivalent.

Proof:
$\Rightarrow$: For $\forall 0 \leq s < t$, since $(B_t - B_s) \sim \mathcal{N}(0, (t-s))$ is independent of $\mathcal{F}_s$, we have $\mathbb{E}[B_t B_s] = \mathbb{E}[(B_t - B_s + B_s)B_s] = \mathbb{E}[B_s^2] = s = \min\{t, s\}$.
For $\forall n \in \mathbb{N}$, $0 = t_0 < t_1 < t_2 < \cdots < t_n$, since the increments $B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}}$ are independent and $B_{t_i} - B_{t_{i-1}} \sim \mathcal{N}(0, (t_i - t_{i-1}))$, we have $(B_{t_1}, B_{t_2}, \cdots, B_{t_n})$ is centered Gaussian $\sim \mathcal{N}(0, \Gamma)$.
$\Leftarrow$: $\mathbb{E}[B_0^2] = 0$, thus $B_0 = 0$ a.s.. For $\forall 0 \leq s < t$, we have $\mathbb{E}[B_t - B_s] = 0$ and $Var(B_t - B_s) = \mathbb{E}[(B_t - B_s)^2] = \mathbb{E}[B_t^2] - 2\mathbb{E}[B_t B_s] + \mathbb{E}[B_s^2] = t - 2s + s = t - s$. Thus, $B_t - B_s \sim \mathcal{N}(0, (t-s))$.
For $\forall A \in \mathcal{F}_s$, by π-λ theorem, we only need to prove that $B_t - B_s$ is independent of $\sigma(B_{r_1}, B_{r_2}, \cdots, B_{r_n})$ for any finite collection $0 \leq r_1, r_2, \cdots, r_n \leq s$. Without loss of generality, assume $r_1 < r_2 < \cdots < r_n$. Then we have

$$\begin{aligned}\sigma(B_{r_1}, B_{r_2}, \cdots, B_{r_n}) &= \sigma(B_{r_1}, B_{r_2} - B_{r_1}, \cdots, B_{r_n} - B_{r_{n-1}}) \\ &\subset \sigma(B_{r_1}, B_{r_2} - B_{r_1}, \cdots, B_s - B_{r_n})\end{aligned}$$

Since $(B_t - B_s)$ is uncorrelated with B_{r_1}, $B_{r_2} - B_{r_1}$, $\cdots$, $B_s - B_{r_n}$, and they are jointly Gaussian, we have $B_t - B_s$ is independent of $\sigma(B_{r_1}, B_{r_2}, \cdots, B_{r_n})$. Hence, the conclusion holds. $\square$

First Construction of Brownian Motion

Step 1: Set $\Omega = \mathbb{R}^{[0,+\infty)}$, which is the set consists of all real-valued function on $[0, +\infty)$. The sets of the form

$$\{\omega \in \Omega : (\omega(t_1), \omega(t_2), \cdots, \omega(t_n)) \in A\}$$

for $n \in \mathbb{N}$, $0 \leq t_1 < t_2 < \cdots < t_n$, and $A \in \mathcal{B}(\mathbb{R}^n)$ are called cylinder sets. Let $\mathcal{F}$ be the σ-algebra generated by all cylinder sets of finite dimension.

Step 2: Let $T_n = \{(t_1, t_2, \cdots, t_n) : t_i \geq 0 \text{ and distinct }\}$ for $n \in \mathbb{N}$. $T = \bigcup_{n=1}^{\infty} T_n$. The collection $\{Q_{t'} : t' \in T\}$ is called a family of finite-dimensional distribution if $Q_{t'}$ is a probability measure on $\mathcal{B}(\mathbb{R}^n)$ for $t' = (t_1, t_2, \cdots, t_n) \in T_n$. The family $\{Q_{t'} : t' \in T\}$ is called consistent if:
(1) For $\forall t' = (t_1, t_2, \cdots, t_n) \in T_n$, and $\forall A_i \in \mathcal{B}(\mathbb{R})$, $\forall \sigma \in S_n$ (the permutation group on n elements), we have

$$Q_{t'}(A_1 \times A_2 \times \cdots \times A_n) = Q_{(t_{\sigma(1)}, t_{\sigma(2)}, \cdots, t_{\sigma(n)})}(A_{\sigma(1)} \times A_{\sigma(2)} \times \cdots \times A_{\sigma(n)})$$

(2) For $\forall t' = (t_1, t_2, \cdots, t_n) \in T_n$, and $\forall A \in \mathcal{B}(\mathbb{R}^{n-1})$, we have

$$Q_{t'}(A \times \mathbb{R}) = Q_{(t_1, t_2, \cdots, t_n, t_{n-1})}(A)$$

Lemma 4.1. For any consistent family $\{Q_{t'} : t' \in T\}$ of finite-dimensional distribution, there exists a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ such that for $\forall t' = (t_1, t_2, \cdots, t_n) \in T_n$, and $\forall A \in \mathcal{B}(\mathbb{R}^n)$, we have

$$\mathbb{P}(\{\omega \in \Omega : (\omega(t_1), \omega(t_2), \cdots, \omega(t_n)) \in A\}) = Q_{t'}(A)$$

For $n \in \mathbb{N}, \forall t' = (t_1, t_2, \cdots, t_n) \in T_n$, choose $Q_{t'}$ as the joint distribution of $(B_{t_1}, B_{t_2}, \cdots, B_{t_n})$ in the following way:
Choose $\sigma \in S_n$ such that $t_{\sigma(1)} < t_{\sigma(2)} < \cdots < t_{\sigma(n)}$. Then define the increments $B_{t_{\sigma(1)}}, B_{t_{\sigma(2)}} - B_{t_{\sigma(1)}}$, $\cdots$, $B_{t_{\sigma(n)}} - B_{t_{\sigma(n-1)}}$ as independent Gaussian random variables with distributions $\mathcal{N}(0, t_{\sigma(1)})$, $\mathcal{N}(0, t_{\sigma(2)} - t_{\sigma(1)})$, $\cdots$, $\mathcal{N}(0, t_{\sigma(n)} - t_{\sigma(n-1)})$ respectively. Then from the distribution of the increments, we can get the joint distribution of $(B_{t_1}, B_{t_2}, \cdots, B_{t_n})$, and hence by choice of $Q_{t'}$, we can get the distribution of $B_{t_1}, B_{t_2}, \cdots, B_{t_n}$. Now choose $B_t(\omega) = \omega(t)$ for $\forall t \geq 0$, and $\mathcal{F}_t = \mathcal{F}_t^B$. Then by Kolmogorov continuity theorem, since $\mathbb{E}[|B_t - B_s|^{2n}] = c_n|t-s|^n$, we have a continuous modification of B on $[0, s]$, which is denoted by W^s. We define $W_t = W_t^s$ for $t \in [0, s]$. Then $W = \{W_t : t \geq 0\}$ is a standard Brownian motion w.r.t. $\{\mathcal{F}_t^B\}$ or $\{\mathcal{F}_t^W\}$.
Second Construction of Brownian Motion

Lemma 4.2. Let B be a Brownian motion w.r.t. $\{\mathcal{F}_t\}$. For $0 \leq s < t$, let $\theta = \frac{t+s}{2}$. Then conditional on $B_s = x, B_t = y$, we have $B_\theta \sim \mathcal{N}(\frac{x+y}{2}, \frac{t-s}{4})$.

Define $I(n) = \{2k-1 : 1 \leq k \leq 2^n\}$ for $n \in \mathbb{N}$. Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and fix a collection $\{\xi_k^{(n)} : k \in I(n)\}$ be a family of independent $\mathcal{N}(0,1)$-distributed random variables on it. Set $B_0^{(0)} = 0, B_1^{(0)} = \xi_1^{(0)}$. For $n \geq 1$, define $B_{k/2^n}^{(n)}$ for $k = 0, 1, 2, \cdots, 2^n$ inductively as follows:

$$B_{k/2^n}^{(n)} = \begin{cases} B_{k/2^{n-1}}^{(n-1)}, & \text{if } k \text{ is even} \\ \frac{1}{2}(B_{(k-1)/2^n}^{(n-1)} + B_{(k+1)/2^n}^{(n-1)}) + \frac{1}{2^{(n+1)/2}}\xi_k^{(n)}, & \text{if } k \text{ is odd} \end{cases}$$

$forall t \in [0,1]$, define $B_t^{(n)}$ by linear interpolation between the points $\{(k/2^n, B_{k/2^n}^{(n)}) : k = 0, 1, 2, \cdots, 2^n\}$. Then $\{B_t^{(n)} : t \in [0,1]\}$ has continuous sample paths.

Lemma 4.3. For almost surely $\omega \in \Omega$, $\{B_t^{(n)}(\omega) : t \in [0,1]\}$ converges uniformly on $[0,1]$ to a limit $\{B_t(\omega) : t \in [0,1]\}$ as $n \to \infty$.

Proof od Lemma:
For $\forall n \in \mathbb{N}$, and $\forall t \in [0,1]$, there exists $k \in \{0, 1, 2, \cdots, 2^n - 1\}$ such that $t \in [\frac{k}{2^n}, \frac{k+1}{2^n}]$. Then we

have

$$\begin{aligned}
|B_t^{(n+1)} - B_t^{(n)}| &\le \max\{|B^{(n+1)}_{(2k+1)/2^{n+1}} - B^{(n)}_{k/2^n}|, |B^{(n+1)}_{(2k+1)/2^{n+1}} - B^{(n)}_{(k+1)/2^n}|\} \\
&= \max\{|\frac{1}{2}(B^{(n)}_{k/2^n} + B^{(n)}_{(k+1)/2^n}) + \frac{1}{2^{(n+2)/2}}\xi^{(n+1)}_{2k+1} - B^{(n)}_{k/2^n}|, \\
&\quad |\frac{1}{2}(B^{(n)}_{k/2^n} + B^{(n)}_{(k+1)/2^n}) + \frac{1}{2^{(n+2)/2}}\xi^{(n+1)}_{2k+1} - B^{(n)}_{(k+1)/2^n}|\} \\
&= \max\{|\frac{1}{2^{(n+2)/2}}\xi^{(n+1)}_{2k+1} - \frac{1}{2}(B^{(n)}_{(k+1)/2^n} - B^{(n)}_{k/2^n})|, \\
&\quad |\frac{1}{2^{(n+2)/2}}\xi^{(n+1)}_{2k+1} + \frac{1}{2}(B^{(n)}_{(k+1)/2^n} - B^{(n)}_{k/2^n})|\} \\
&\le |\frac{1}{2^{(n+2)/2}}\xi^{(n+1)}_{2k+1}| + \frac{1}{2}|B^{(n)}_{(k+1)/2^n} - B^{(n)}_{k/2^n}|
\end{aligned}$$

Thus, we have

$$\|B^{(n+1)} - B^{(n)}\| \le \max_{k\in I(n+1)} |\frac{1}{2^{(n+2)/2}}\xi_k^{(n+1)}| + \frac{1}{2}\|B^{(n)} - B^{(n-1)}\|$$

By iteration, we have

$$\|B^{(n+1)} - B^{(n)}\| \le \sum_{j=1}^{n+1} \max_{k\in I(j)} |\frac{1}{2^{(j+1)/2}}\xi_k^{(j)}|$$

For $\forall \epsilon > 0$, by Chebyshev's inequality, we have

$$\mathbb{P}(\max_{k\in I(j)} |\xi_k^{(j)}| > 2^{(j+1)/4}\epsilon) \le \sum_{k\in I(j)} \mathbb{P}(|\xi_k^{(j)}| > 2^{(j+1)/4}\epsilon) \le 2^j \frac{\mathbb{E}[|\xi_k^{(j)}|^4]}{2^{(j+1)}\epsilon^4} = \frac{3}{2\epsilon^4}2^{-j}$$

Thus, by Borel-Cantelli lemma, we have $\max_{k\in I(j)} |\frac{1}{2^{(j+1)/2}}\xi_k^{(j)}| \to 0$ a.s. as $j \to \infty$. Hence, $\sum_{j=1}^{\infty} \max_{k\in I(j)} |\frac{1}{2^{(j+1)/2}}\xi_k^{(j)}| < \infty$ a.s., which implies that $\|B^{(n+1)} - B^{(n)}\| \to 0$ a.s. as $n \to \infty$. □

Now verify that $B = \{B_t : t \in [0, 1]\}$ is a standard Brownian motion. Obviously, B has continuous sample paths and $B_0 = 0$ a.s.. For $\forall 0 \le s < t \le 1$, we have

$$B_t - B_s = \sum_{n=1}^{\infty}(B_t^{(n)} - B_t^{(n-1)}) - \sum_{n=1}^{\infty}(B_s^{(n)} - B_s^{(n-1)})$$

Since $\{B_t^{(n)} - B_t^{(n-1)} : t \in [0, 1]\}$ are independent Gaussian processes with mean 0 and covariance function

$$Cov(B_t^{(n)} - B_t^{(n-1)}, B_s^{(n)} - B_s^{(n-1)}) = \begin{cases} \frac{1}{2^n}(min\{t, s\} - \frac{\lfloor 2^n t\rfloor}{2^n}), & \text{if } \lfloor 2^n t\rfloor = \lfloor 2^n s\rfloor \\ 0, & \text{otherwise} \end{cases}$$

we have $B_t - B_s \sim \mathcal{N}(0, t - s)$. Similarly, we can verify that B has independent increments. Thus, B is a standard Brownian motion on $[0, 1]$. By repeating the above construction on $[n, n + 1]$ for $n \in \mathbb{N}$, we can get a standard Brownian motion on $[0, +\infty)$.

Propsition 4.3. The sample paths of Brownian motion are locally Holder continuous of any order $\gamma < \frac{1}{2}$ a.s..

Proof:
For $\forall T > 0$, by Kolmogorov continuity theorem, since $\mathbb{E}[|B_t - B_s|^{2n}] = c_n|t-s|^n$, we have a continuous modification of B on $[0, T]$, which is locally Holder continuous of any order $\gamma < \frac{1}{2}$ a.s.. By taking $T \to \infty$, the conclusion holds. $\square$

Propsition 4.4. A standard Brownian motion B w.r.t. $\{\mathcal{F}_t\}$ is a martingale w.r.t. $\{\mathcal{F}_t\}$.

Proof:
For $\forall 0 \leq s < t$, we have $\mathbb{E}[|B_t|] = \sqrt{\frac{2t}{\pi}} < \infty$, and

$$\mathbb{E}[B_t|\mathcal{F}_s] = \mathbb{E}[B_t - B_s + B_s|\mathcal{F}_s] = \mathbb{E}[B_t - B_s] + B_s = B_s$$

Thus, the conclusion holds. $\square$

Propsition 4.5. A standard Brownian motion B w.r.t. $\{\mathcal{F}_t\}$ is always a time-homogeneous strong Feller process w.r.t. $\{\mathcal{F}_t\}$, with transition probability density function

$$\rho_t(x,y) = \frac{1}{(2\pi t)^{d/2}} \exp(-\frac{|y-x|^2}{2t}), \quad \forall x, y \in \mathbb{R}^d, t > 0.$$

Proof:
For $\forall f \in B_b(\mathbb{R}^d)$, we have

$$\begin{aligned} T_t f(x) &= \mathbb{E}[f(B_t)|B_0 = x] \\ &= \mathbb{E}[f(x + B_t)] \\ &= \int_{\mathbb{R}^d} f(x+y) \frac{1}{(2\pi t)^{d/2}} \exp(-\frac{|y|^2}{2t}) dy \\ &= \int_{\mathbb{R}^d} f(y) \frac{1}{(2\pi t)^{d/2}} \exp(-\frac{|y-x|^2}{2t}) dy \end{aligned}$$

Thus, the transition probability density function is as stated. Since $\rho_t(x,y)$ is continuous in x for $\forall y \in \mathbb{R}^d$, we have $T_t(B_b(\mathbb{R}^d)) \subset C_b(\mathbb{R}^d)$ for $\forall t > 0$. Moreover, by dominated convergence theorem, we have $\lim_{t \to 0^+} \|T_t f - f\| = 0$ for $\forall f \in B_b(\mathbb{R}^d)$. Hence, the conclusion holds. $\square$

Remark. Brownian motion is a strong Markov process.

Theorem 4.1. Let S be a finite optional time w.r.t. $\{\mathcal{F}_t\}$. Set $W_t = B_{S+t} - B_S$ for $\forall t \geq 0$. Then $W = \{W_t : t \geq 0\}$ is a standard Brownian motion w.r.t. the filtration $\{\mathcal{F}_W\}$ and independent of $\mathcal{F}_S^+$.

Proof:
Let us prove with characteristic functions. For $\forall n \in \mathbb{N}$, $0 \leq t_1 < t_2 < \cdots < t_n$, and $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{R}^d$, we have

$$\begin{aligned}
&\mathbb{E}[\exp(i\sum_{j=1}^{n} \lambda_j \cdot W_{t_j})|\mathcal{F}_S^+] \\
&= \mathbb{E}[\exp(i\sum_{j=1}^{n} \lambda_j \cdot (B_{S+t_j} - B_S))|\mathcal{F}_S^+] \\
&= \mathbb{E}[\exp(i\sum_{j=1}^{n} \lambda_j \cdot B_{S+t_j})|\mathcal{F}_S^+] \exp(-i\sum_{j=1}^{n} \lambda_j \cdot B_S)
\end{aligned}$$

Since S is an optional time, for $\forall m \in \mathbb{N}$, we have

$$\begin{aligned}
&\mathbb{E}[\exp(i\sum_{j=1}^{n} \lambda_j \cdot B_{S+t_j})\mathbb{1}_{\{S\leq m\}}|\mathcal{F}_m] \\
&= \sum_{k=0}^{\infty} \mathbb{E}[\exp(i\sum_{j=1}^{n} \lambda_j \cdot B_{S+t_j})\mathbb{1}_{\{k<S\leq k+1\}}|\mathcal{F}_m] \\
&= \sum_{k=0}^{m} \mathbb{E}[\exp(i\sum_{j=1}^{n} \lambda_j \cdot B_{S+t_j})\mathbb{1}_{\{k<S\leq k+1\}}|\mathcal{F}_m]
\end{aligned}$$

For $\forall k = 0, 1, \cdots, m$, on the set $\{k < S \leq k+1\}$, since $B_{S+t_j} - B_{k+1}$ are independent of $\mathcal{F}_{k+1}$ for $\forall j = 1, 2, \cdots, n$, we have

$$\begin{aligned}
&\mathbb{E}[\exp(i\sum_{j=1}^{n} \lambda_j \cdot B_{S+t_j})\mathbb{1}_{\{k<S\leq k+1\}}|\mathcal{F}_m] \\
&= \mathbb{E}[\mathbb{E}[\exp(i\sum_{j=1}^{n} \lambda_j \cdot B_{S+t_j})|\mathcal{F}_{k+1}]\mathbb{1}_{\{k<S\leq k+1\}}|\mathcal{F}_m] \\
&= \mathbb{E}[\exp(i\sum_{j=1}^{n} \lambda_j \cdot B_{k+1})\mathbb{E}[\exp(i\sum_{j=1}^{n} \lambda_j \cdot (B_{S+t_j} - B_{k+1}))]\mathbb{1}_{\{k<S\leq k+1\}}|\mathcal{F}_m] \\
&= \mathbb{E}[\exp(i\sum_{j=1}^{n} \lambda_j \cdot B_{k+1}) \exp(-\frac{1}{2}\sum_{j=1}^{n} |\lambda_j|^2 (S + t_j - (k+1)))\mathbb{1}_{\{k<S\leq k+1\}}|\mathcal{F}_m]
\end{aligned}$$

Thus, we have

$$
\begin{aligned}
&\mathbb{E}[\exp(i\sum_{j=1}^{n}\lambda_j\cdot B_{S+t_j})\mathbb{1}_{\{S\le m\}}|\mathcal{F}_m] \\
=&\sum_{k=0}^{m}\mathbb{E}[\exp(i\sum_{j=1}^{n}\lambda_j\cdot B_{k+1})\exp(-\frac{1}{2}\sum_{j=1}^{n}|\lambda_j|^2(S+t_j-(k+1)))\mathbb{1}_{\{k<S\le k+1\}}|\mathcal{F}_m]
\end{aligned}
$$

By taking $m\to\infty$, we have

$$
\begin{aligned}
&\mathbb{E}[\exp(i\sum_{j=1}^{n}\lambda_j\cdot B_{S+t_j})|\mathcal{F}_S^+] \\
=&\sum_{k=0}^{\infty}\mathbb{E}[\exp(i\sum_{j=1}^{n}\lambda_j\cdot B_{k+1})\exp(-\frac{1}{2}\sum_{j=1}^{n}|\lambda_j|^2(S+t_j-(k+1)))\mathbb{1}_{\{k<S\le k+1\}}|\mathcal{F}_S^+]
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
&\mathbb{E}[\exp(i\sum_{j=1}^{n}\lambda_j\cdot W_{t_j})|\mathcal{F}_S^+] \\
=&\sum_{k=0}^{\infty}\mathbb{E}[\exp(i\sum_{j=1}^{n}\lambda_j\cdot B_{k+1})\exp(-\frac{1}{2}\sum_{j=1}^{n}|\lambda_j|^2(S+t_j-(k+1)))\mathbb{1}_{\{k<S\le k+1\}}|\mathcal{F}_S^+]\exp(-i\sum_{j=1}^{n}\lambda_j\cdot B_S) \\
=&\sum_{k=0}^{\infty}\exp(-\frac{1}{2}\sum_{j=1}^{n}|\lambda_j|^2(S+t_j-(k+1)))\mathbb{1}_{\{k<S\le k+1\}}\mathbb{E}[\exp(i\sum_{j=1}^{n}\lambda_j\cdot(B_{k+1}-B_S))|\mathcal{F}_S^+] \\
=&\sum_{k=0}^{\infty}\exp(-\frac{1}{2}\sum_{j=1}^{n}|\lambda_j|^2(S+t_j-(k+1)))\mathbb{1}_{\{k<S\le k+1\}}\exp(-\frac{1}{2}\sum_{j=1}^{n}|\lambda_j|^2(k+1-S)) \\
=&\exp(-\frac{1}{2}\sum_{j=1}^{n}|\lambda_j|^2t_j)
\end{aligned}
$$

Hence, the conclusion holds. □

Propsition 4.6. Let B be a standard Brownian motion w.r.t. $\{\mathcal{F}_t\}$. Then the following holds:
(1) Scaling property: For $\forall c>0$, $\{c^{-1/2}B_{ct}:t\ge 0\}$ is also a standard Brownian motion w.r.t. $\{\mathcal{F}_{ct}\}$.
(2) Time inversion: Define $W_t=tB_{1/t}$ for $\forall t>0$ and $W_0=0$. Then $W=\{W_t:t\ge 0\}$ is also a standard Brownian motion w.r.t. $\{\mathcal{F}_t^W\}$.
(3) Time reversal: For $\forall T>0$, define $W_t=B_T-B_{T-t}$ for $\forall t\in[0,T]$. Then $W=\{W_t:t\in[0,T]\}$ is also a standard Brownian motion w.r.t. $\{\mathcal{F}_t^W\}$.
(4) Symmetry: For $\forall a\in\mathbb{R}^d$, $\{a-B_t:t\ge 0\}$ is also a standard Brownian motion w.r.t. $\{\mathcal{F}_t\}$.

Proof:

(1) For $\forall 0 \leq s < t$, we have $\mathbb{E}[c^{-1/2}B_{ct} - c^{-1/2}B_{cs}] = 0$ and $Var(c^{-1/2}B_{ct} - c^{-1/2}B_{cs}) = c^{-1}(t-s)c = t-s$. Thus, $c^{-1/2}B_{ct} - c^{-1/2}B_{cs} \sim \mathcal{N}(0, t-s)$. Similarly, we can verify that $\{c^{-1/2}B_{ct} : t \geq 0\}$ has independent increments. Hence, the conclusion holds.
(2) For $\forall 0 < s < t$, we have $\mathbb{E}[W_t - W_s] = \mathbb{E}[tB_{1/t} - sB_{1/s}] = 0$ and

$$\begin{aligned}
Var(W_t - W_s) &= \mathbb{E}[(tB_{1/t} - sB_{1/s})^2] \\
&= \mathbb{E}[(tB_{1/t} - sB_{1/t} + sB_{1/t} - sB_{1/s})^2] \\
&= \mathbb{E}[((t-s)B_{1/t} + s(B_{1/t} - B_{1/s}))^2] \\
&= (t-s)^2\mathbb{E}[B_{1/t}^2] + s^2\mathbb{E}[(B_{1/t} - B_{1/s})^2] \\
&= (t-s)^2\frac{1}{t} + s^2(\frac{1}{s} - \frac{1}{t}) = t - s
\end{aligned}$$

Thus, $W_t - W_s \sim \mathcal{N}(0, t-s)$. Similarly, we can verify that W has independent increments. Hence, the conclusion holds.
(3) For $\forall 0 \leq s < t \leq T$, we have $\mathbb{E}[W_t - W_s] = \mathbb{E}[B_T - B_{T-t} - (B_T - B_{T-s})] = 0$ and

$$Var(W_t - W_s) = \mathbb{E}[(B_{T-s} - B_{T-t})^2] = t - s$$

Thus, $W_t - W_s \sim \mathcal{N}(0, t-s)$. Similarly, we can verify that W has independent increments. Hence, the conclusion holds.
(4) For $\forall 0 \leq s < t$, we have $\mathbb{E}[a - B_t - (a - B_s)] = \mathbb{E}[B_s - B_t] = 0$ and

$$Var(a - B_t - (a - B_s)) = \mathbb{E}[(B_t - B_s)^2] = t - s$$

Thus, $a - B_t - (a - B_s) \sim \mathcal{N}(0, t-s)$. Similarly, we can verify that $\{a - B_t : t \geq 0\}$ has independent increments. Hence, the conclusion holds. $\square$

Propsition 4.7. With probability 1, a standard Brownian motion B w.r.t. $\{\mathcal{F}_t\}$ changes its sign infinitely many times in any interval $[0, \epsilon]$ for $\forall \epsilon > 0$.

Proof:
For $\forall \epsilon > 0$, define $A_\epsilon = \{\omega \in \Omega : B_t(\omega)$ changes its sign infinitely many times in $[0, \epsilon]\}$. For $\forall r \in [0, \epsilon]$, since B has independent and stationary increments, we have

$$\begin{aligned}
\mathbb{P}(A_\epsilon) &= \mathbb{P}(\{\omega \in \Omega : B_{r+t}(\omega) - B_r(\omega) \text{ changes its sign infinitely many times in } [0, \epsilon - r]\}) \\
&= \mathbb{P}(A_{\epsilon - r})
\end{aligned}$$

Thus, $\mathbb{P}(A_\epsilon) = \mathbb{P}(A_{\epsilon-r})$ for $\forall r \in [0, \epsilon]$, which implies that $\mathbb{P}(A_\epsilon) = \mathbb{P}(A_0)$. By taking $r \to \epsilon^-$, we have $\mathbb{P}(A_\epsilon) = \mathbb{P}(A_0)$. Since $A_0 \in \mathcal{F}_{0+} = \bigcap_{t>0} \mathcal{F}_t$, by Blumenthal's $0-1$ law, we have $\mathbb{P}(A_0) \in \{0, 1\}$.

Thus, $\mathbb{P}(A_\epsilon) \in \{0, 1\}$.
Next, we prove that $\mathbb{P}(A_\epsilon) = 1$. For $\forall n \in \mathbb{N}$, define $\tau_n = \inf\{t > 0 : |B_t| = \frac{1}{n}\}$. Then τ_n is a stopping time w.r.t. $\{\mathcal{F}_t\}$. Since B is a standard Brownian motion, we have $\mathbb{P}(\tau_n < \epsilon) > 0$ for $\forall n \in \mathbb{N}$. Thus, by strong Markov property, we have

$$\mathbb{P}(A_\epsilon) \geq \mathbb{P}(A_\epsilon | \tau_n < \epsilon)\mathbb{P}(\tau_n < \epsilon) = \mathbb{P}(A_{\epsilon - \tau_n})\mathbb{P}(\tau_n < \epsilon) = \mathbb{P}(A_\epsilon)\mathbb{P}(\tau_n < \epsilon)$$

which implies that $\mathbb{P}(A_\epsilon) = 1$. Hence, the conclusion holds. $\square$

Propsition 4.8. With probability 1, a standard Brownian motion B w.r.t. $\{\mathcal{F}_t\}$ returns to origion infinitely often in any interval $[0, \epsilon]$ for $\forall \epsilon > 0$.

Theorem 4.2. Let B be a 1-dimensional standard Brownian motion w.r.t. $\{\mathcal{F}_t\}$. For $\forall \omega \in \Omega$, define $\mathcal{L}(\omega) = \{t \geq 0 : B_t(\omega) = 0\}$. Then with probability 1, the following holds:
(1) $\mathcal{L}(\omega)$ has Lebesgue measure 0.
(2) $\mathcal{L}(\omega)$ is a closed and unbounded.
(3) $\mathcal{L}(\omega)$ has accumulated point at $t = 0$
(4) $\mathcal{L}(\omega)$ has no isolated points in $[0, +\infty)$.

Proof:
(1) By Fubini's theorem, we have

$$\mathbb{E}[m(\mathcal{L})] = \mathbb{E}[\int_0^\infty \mathbb{1}_{\{B_t = 0\}} dt] = \int_0^\infty \mathbb{P}(B_t = 0) dt = 0$$

which implies that $m(\mathcal{L}) = 0$ a.s..
(2) Since B has continuous sample paths, $\mathcal{L}(\omega) = B_t^{-1}(\omega)(\{0\})$ is closed for $\forall \omega \in \Omega$. Since B returns to origin infinitely often, $\mathcal{L}(\omega)$ is unbounded a.s..
(3) $B_t(\omega)$ changes its sign infinitely many times in any interval $[0, \epsilon]$ for $\forall \epsilon > 0$ a.s., which implies that $\mathcal{L}(\omega)$ has accumulated point at $t = 0$ a.s..
(4) $\{\omega : \mathcal{L}(\omega) \text{ has isolated point }\} = \bigcup_{0 < a < b \in \mathbb{Q}} \{\omega : \exists! s \in (a, b),\, B_s(\omega) = 0\}$. Set $\beta_t(\omega) = \inf\{s > t : B_s(\omega) = 0\}$. Then $\beta_0 = 0$ and $\beta_t < \infty$ for $\forall t \geq 0$ a.s.. $\{\beta_t < r\} = \{\exists s \in (t, r),\, B_s = 0\} \in \mathcal{F}_r$ for $\forall r > t$, which implies that β_t is a optional time w.r.t. $\{\mathcal{F}_t\}$. And $B_{\beta_t} = 0$ a.s.. $\beta_{\beta_t} = \inf\{s > \beta_t : B_s = 0\} = \beta_t + \inf\{s > 0 : B_{s+\beta_t} - B_{\beta_t} = 0\}$. Then $W_s = B_{s+\beta_t} - B_{\beta_t}$ is a standard Brownian motion. $\inf\{s > 0 : B_{s+\beta_t} - B_{\beta_t} = 0\} = \inf\{s > 0 : W_s = 0\} = 0$ a.s.. Thus, we have $\beta_{\beta_t} = \beta_t$ a.s.. Hence, $\mathbb{P}(\exists! s \in (a, b),\, B_s = 0) = \mathbb{P}(\beta_a < b < \beta_{\beta_a}) = 0$. Thus, the conclusion holds. $\square$

Theorem 4.3. B is a 1-dimensional standard Brownian motion w.r.t. $\{\mathcal{F}_t\}$. Then the sample paths is monotone in no interval a.s..

Proof:
$F = \{\omega : B_t(\omega)$ is monotone in some interval$\} = \bigcup_{0<s<t\in\mathbb{Q}}\{\omega : B_t(\omega)$ is monotone in $[s,t]\}$. For $\forall 0 < s < t \in \mathbb{Q}$, define $F_{s,t} = \{\omega : B_t(\omega)$ is monotone in $[s,t]\}$. Then $F = \bigcup_{0<s<t\in\mathbb{Q}} F_{s,t}$. Let $F^+_{s,t} = \{\omega : B_t(\omega)$ is non-decreasing in $[s,t]\}$ and $F^-_{s,t} = \{\omega : B_t(\omega)$ is non-increasing in $[s,t]\}$. Then $F_{s,t} = F^+_{s,t} \cup F^-_{s,t}$. Define $A^n_{s,t} = \cap_{k=0}^{n-1}\{\omega : B_{s+\frac{(t-s)k}{n}}(\omega) \leq B_{s+\frac{(t-s)(k+1)}{n}}(\omega)\}$ for $\forall n \in \mathbb{N}$. Then $F^+_{s,t} \subset \cap_{n=1}^{\infty} A^n_{s,t}$. Since $\{B_{s+\frac{(t-s)k}{n}} - B_{s+\frac{(t-s)(k-1)}{n}} : k = 1, 2, \cdots, n\}$ are independent and identically distributed random variables with mean 0, we have $\mathbb{P}(B_{s+\frac{(t-s)k}{n}} - B_{s+\frac{(t-s)(k-1)}{n}} \geq 0) = \frac{1}{2}$ for $\forall k = 1, 2, \cdots, n$. Thus, we have

$$\mathbb{P}(A^n_{s,t}) = \prod_{k=1}^{n} \mathbb{P}(B_{s+\frac{(t-s)k}{n}} - B_{s+\frac{(t-s)(k-1)}{n}} \geq 0) = \frac{1}{2^n}$$

which implies that $\mathbb{P}(F^+_{s,t}) \leq \mathbb{P}(\cap_{n=1}^{\infty} A^n_{s,t}) = \lim_{n\to\infty} \mathbb{P}(A^n_{s,t}) = 0$. Similarly, we have $\mathbb{P}(F^-_{s,t}) = 0$. Thus, $\mathbb{P}(F_{s,t}) = 0$. By taking union over all $0 < s < t \in \mathbb{Q}$, we have $\mathbb{P}(F) = 0$. Hence, the conclusion holds. $\square$

Theorem 4.4. Sample paths of a standard Brownian motion B w.r.t. $\{\mathcal{F}_t\}$ have no points of increase or decrease a.s..

Propsition 4.9. Let B be a 1-dimensional standard Brownian motion w.r.t. $\{\mathcal{F}_t\}$. For a.s. $\omega \in \Omega$, the local maxima and minima of the sample path $\{B_t(\omega) : t \geq 0\}$ are strict.

Proof:
$\{\omega :$ All local maxima are strict$\} \supset \bigcup_{0\leq t_1<t_2<t_3<t_4\in\mathbb{Q}}\{\max_{t\in[t_1,t_2]} B_t \neq \max_{t\in[t_3,t_4]} B_t\}$.

$$\begin{aligned}\max_{t\in[t_1,t_2]} B_t(\omega) - \max_{t\in[t_3,t_4]} B_t(\omega) &= \max_{t\in[t_1,t_2]}(B_t(\omega) - B_{t_2}(\omega)) - \max_{t\in[t_3,t_4]}(B_t(\omega) - B_{t_3}(\omega)) + B_{t_2}(\omega) - B_{t_3}(\omega)\\ &= Y_1 + Y_2 + Y_3\end{aligned}$$

Since Y_1, Y_2, Y_3 are independent random variables and have continuous distributions, we have $\mathbb{P}(Y_1 + Y_2 + Y_3 = 0) = 0$. Thus, we have

$$\mathbb{P}(\max_{t\in[t_1,t_2]} B_t = \max_{t\in[t_3,t_4]} B_t) = 0$$

By taking union over all $0 \leq t_1 < t_2 < t_3 < t_4 \in \mathbb{Q}$, we have $\mathbb{P}($ All local maxima are strict$) = 1$. Similarly, we can prove that $\mathbb{P}($ All local minima are strict$) = 1$. Hence, the conclusion holds. $\square$

Corollary 4.1. Define I is the set of points of local maxima (or minima) of a 1-dimensional standard Brownian motion B w.r.t. $\{\mathcal{F}_t\}$. Then with probability 1, I is countable and has no accumulation point in $[0, +\infty)$.

Proof:

$B.(\omega)$ is continuous and is monotone in no interval. If $\exists[a,b]\cap I=\emptyset$, then $B.(\omega)$ is monotone in $[a,b]$, which is impossible. Thus, I is dense in $[0,+\infty)$. Let $M_\epsilon=\{t\in\mathbb{Q}\cap\mathbb{R}^+ : B_t\geq B_s, \forall s\in[t-\epsilon,t+\epsilon]\}\subset\mathbb{Q}$ for $\forall\epsilon>0$. Then $I\subset\bigcup_{\epsilon\in\mathbb{Q}}M_\epsilon$. Since M_ϵ is countable for $\forall\epsilon>0$, I is countable. $\square$

Theorem 4.5. B is a 1-dimensional standard Brownian motion w.r.t. $\{\mathcal{F}_t\}$. For a.s. $\omega\in\Omega$, the sample path $\{B_t(\omega): t\geq 0\}$ is nowhere differentiable.

Recall:

Dini derivatives of a function $f:\mathbb{R}\to\mathbb{R}$ at t are defined as follows:

$$D^+f(t)=\limsup_{h\to0^+}\frac{f(t+h)-f(t)}{h},\quad D_+f(t)=\liminf_{h\to0^+}\frac{f(t+h)-f(t)}{h}$$
$$D^-f(t)=\limsup_{h\to0^-}\frac{f(t+h)-f(t)}{h},\quad D_-f(t)=\liminf_{h\to0^-}\frac{f(t+h)-f(t)}{h}$$

Proof:

Claim: $\exists F\in\mathcal{F},\mathbb{P}(F)=1$, s.t. $F\subset\{\omega:\forall t\geq0 \text{ either } D^+B_t(\omega)=+\infty \text{ or } D_+B_t(\omega)=-\infty\}$.

Set $A_{j,k}=\cup_{t\in[0,1]}\cap_{h\in[0,1/k]}\{\omega:|B_{t+h}(\omega)-B_t(\omega)|\leq jh\}$ for $\forall j,k\in\mathbb{N}$. Then $A_{j,k}\in\mathcal{F}$ for $\forall j,k\in\mathbb{N}$. And $\cup_{j=1}^\infty\cup_{k=1}^\infty A_{j,k}=\{\omega:\exists t\in[0,1] \text{ s.t. } D^+B_t(\omega)<+\infty \text{ and } D_+B_t(\omega)>-\infty\}$. Fix $\omega\in A_{j,k}$, the $\exists t\in[0,1]$ s.t. $\forall h\in[0,1/k]$ we have $|B_{t+h}(\omega)-B_t(\omega)|\leq jh$. For $\forall h_1,h_2\in[0,1/k]$ we have $|B_{t+h_1}(\omega)-B_{t+h_2}(\omega)|\leq j|h_1-h_2|$. Now for $\forall n\geq4k$, $\exists i$ s.t. $\frac{i-1}{n}\leq t<\frac{i}{n}$. Then for $l=1,2,3$ we have

$$|B_{\frac{i+l}{n}}(\omega)-B_{\frac{i+l-1}{n}}(\omega)|\leq j(\frac{i+l}{n}-t)+j(\frac{i+l-1}{n}-t)\leq j\frac{2l+1}{n}$$

$A_{j,k}\subset\cup_{i=1}^n\cap_{l=1}^3\{\omega:|B_{\frac{i+l}{n}}-B_{\frac{i+l-1}{n}}|\leq j\frac{2l+1}{n}\}=C_n$. Since $B_{\frac{i+l}{n}}-B_{\frac{i+l-1}{n}}\sim\mathcal{N}(0,\frac{1}{n})$, we have

$$\begin{aligned}\mathbb{P}(\bigcap_{l=1}^3\{|B_{\frac{i+l}{n}}-B_{\frac{i+l-1}{n}}|\leq j\frac{2l+1}{n}\})&=\prod_{l=1}^3\mathbb{P}(|B_{\frac{i+l}{n}}-B_{\frac{i+l-1}{n}}|\leq j\frac{2l+1}{n})\\&=\prod_{l=1}^3\int_{-j\frac{2l+1}{n}}^{j\frac{2l+1}{n}}\sqrt{\frac{n}{2\pi}}\exp(-\frac{nx^2}{2})dx\\&\leq\prod_{l=1}^3 2j\frac{2l+1}{n}\sqrt{\frac{n}{2\pi}}=\frac{48j^3}{\pi^{3/2}n^{3/2}}\end{aligned}$$

Thus $\mathbb{P}(C_n)\leq n\frac{48j^3}{\pi^{3/2}n^{3/2}}=\frac{48j^3}{\pi^{3/2}n^{1/2}}$. By taking $n\to\infty$, we have $\mathbb{P}(A_{j,k})=0$ for $\forall j,k\in\mathbb{N}$. Thus, we have

$$\mathbb{P}(\exists t\in[0,1] \text{ s.t. } D^+B_t<+\infty \text{ and } D_+B_t>-\infty)=0$$

which implies the claim.

Then $\mathbb{P}(F)=1$ and $F\subset\{\omega:\forall t\geq0 \text{ either } D^+B_t(\omega)=+\infty \text{ or } D_+B_t(\omega)=-\infty\}$. Similarly, we can prove that $\exists G \in \mathcal{F}, \mathbb{P}(G) = 1$, s.t. $G \subset \{\omega : \forall t \geq 0 \text{ either } D^-B_t(\omega) = +\infty \text{ or } D_-B_t(\omega) = -\infty\}$. Set $H = F \cap G$. Then $\mathbb{P}(H) = 1$ and

$$H \subset \{\omega : \forall t \geq 0, \text{ either } D^+B_t(\omega) = +\infty \text{ or } D_+B_t(\omega) = -\infty, \text{ either } D^-B_t(\omega) = +\infty \text{ or } D_-B_t(\omega) = -\infty\}$$

For $\forall \omega \in H$ and $\forall t \geq 0$, if $B.(\omega)$ is differentiable at t, then we have $D^+B_t(\omega) = D_+B_t(\omega) = D^-B_t(\omega) = D_-B_t(\omega) \in \mathbb{R}$, which is impossible. Hence, the conclusion holds. $\square$

Theorem 4.6. B is a 1-dimensional standard Brownian motion w.r.t. $\{\mathcal{F}_t\}$. For a.s. $\omega \in \Omega$, the sample path $\{B_t(\omega) : t \geq 0\}$ is nowhere Hölder continuous of order γ for any $\gamma > \frac{1}{2}$.

Proof:
Define $F = \{\omega : B_t \text{ is Hölder continuous of order } \gamma > 1/2 \text{ in some interval}\}$,
$F_{s,t} = \{\omega : B_t \text{ is Hölder continuous of order } \gamma > 1/2 \text{ in } [s,t]\}$ for $\forall 0 < s < t \in \mathbb{Q}$.
Then $F = \bigcup_{0<s<t\in\mathbb{Q}} F_{s,t}$. Set $A^n_{s,t} = \cap_{k=0}^{n-1}\{\omega : |B_{s+\frac{(t-s)(k+1)}{n}}(\omega) - B_{s+\frac{(t-s)k}{n}}(\omega)| \leq c(\frac{|t-s|}{n})^\gamma\}$ for $\forall n \in \mathbb{N}$. Then $F_{s,t} \subset \cap_{n=1}^\infty A^n_{s,t}$. Since the distrubution of $B_{s+\frac{(t-s)(k+1)}{n}} - B_{s+\frac{(t-s)k}{n}}$ is $\mathcal{N}(0, \frac{|t-s|}{n})$, we have

$$\begin{aligned}
\mathbb{P}(A^n_{s,t}) &= \prod_{k=0}^{n-1} \mathbb{P}(|B_{s+\frac{(t-s)(k+1)}{n}} - B_{s+\frac{(t-s)k}{n}}| \leq c(\frac{|t-s|}{n})^\gamma) \\
&= \prod_{k=0}^{n-1} \int_{-c(\frac{|t-s|}{n})^\gamma}^{c(\frac{t-s}{n})^\gamma} \sqrt{\frac{n}{2\pi|t-s|}} \exp(-\frac{nx^2}{2|t-s|})dx \\
&\leq \prod_{k=0}^{n-1} 2c(\frac{|t-s|}{n})^\gamma \sqrt{\frac{n}{2\pi|t-s|}} = (\frac{2c|t-s|^\gamma}{\sqrt{2\pi|t-s|}})^n \frac{1}{n^{n(\gamma-\frac{1}{2})}}
\end{aligned}$$

Since $\gamma > \frac{1}{2}$, by taking $n \to \infty$, we have $\mathbb{P}(F_{s,t}) \leq \mathbb{P}(\cap_{n=1}^\infty A^n_{s,t}) = \lim_{n\to\infty} \mathbb{P}(A^n_{s,t}) = 0$. By taking union over all $0 < s < t \in \mathbb{Q}$, we have $\mathbb{P}(F) = 0$. Hence, the conclusion holds. $\square$

Remark. $\gamma = \frac{1}{2}$ is the threshold for Hölder continuity of Brownian motion sample paths.

Theorem 4.7 (Lévy Modulus). $\mathbb{P}(\lim_{\delta\to 0^+} \frac{1}{\sqrt{2\delta\ln(1/\delta)}} \max_{0\leq s<t\leq 1, t-s\leq\delta} |B_t - B_s| = 1) = 1$.

5 Poisson Process and Lévy Process

Definition 5.1. A r.v. X takes values in $\mathbb{N} = \{0, 1, 2, \cdots\}$ is said to have **Poisson distribution** with parameter $\lambda > 0$ if

$$\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \cdots$$

And we say $X \sim Poisson(\lambda)$.

Remark. Characteristic function of $X \sim Poisson(\lambda)$ is given by

$$\phi_X(u) = \mathbb{E}[\exp(iuX)] = \exp(\lambda(e^{iu} - 1)), \quad u \in \mathbb{R}$$

Definition 5.2. A r.v. X is called **infinitely divisible** if for $\forall n \in \mathbb{N}$, there exist i.i.d. r.v.s $X_{1,n}, X_{2,n}, \cdots, X_{n,n}$ s.t.

$$X \stackrel{d}{=} X_{1,n} + X_{2,n} + \cdots + X_{n,n}$$

Remark. A probability measure μ on $\mathbb{R}^d$ is called infinitely divisible if for $\forall n \in \mathbb{N}$, there exist r.v. X is infinitely divisible with distribution μ.

Definition 5.3. A r.v. Z is called **compound Poisson** r.v. with parameters $\lambda > 0$ and distribution ν on $\mathbb{R}^d$ if

$$Z = \sum_{i=1}^{N} X_i$$

where $N \sim Poisson(\lambda)$, $\{X_i\}$ are i.i.d. r.v.s with distribution ν and are independent of N.

Remark. Characteristic function of compound Poisson r.v. Z is given by

$$\phi_Z(u) = \mathbb{E}[\exp(iu \cdot Z)] = \exp(\lambda(\int_{\mathbb{R}^d} e^{iu \cdot x} \nu(dx) - 1)), \quad u \in \mathbb{R}^d$$

Definition 5.4. ν is called **Lévy measure** on $\mathbb{R}^d - \{0\}$ if $\nu(\{0\}) = 0$ and

$$\int_{\mathbb{R}^d - \{0\}} (|x|^2 \wedge 1) \nu(dx) < +\infty$$

If additionally ν is absolutely continuous w.r.t. Lebesgue measure, then the Radon-Nikodym derivative $\frac{d\nu}{dx}$ is called **Lévy density**.

Remark. $\int_{\mathbb{R}^d - \{0\}} \min(|x|^2, 1) \nu(dx) < +\infty \iff \int_{\mathbb{R}^d - \{0\}} \frac{x^2}{1+x^2} \nu(dx) < +\infty$.

Remark. A finite measure ν on $\mathbb{R}^d - \{0\}$ is a Lévy measure.

Propsition 5.1. Every Lévy measure ν on $\mathbb{R}^d - \{0\}$ is σ-finite.

Proof:
Set $A_n = \{x \in \mathbb{R}^d : \frac{1}{n} \leq |x| < 1\}$ for $\forall n \in \mathbb{N}$. Then $\nu(A_n) \leq \int_{A_n} |x|^2 \nu(dx) \leq \int_{\mathbb{R}^d - \{0\}} (|x|^2 \wedge 1) \nu(dx) < +\infty$. Thus, ν is σ-finite. $\square$

Theorem 5.1 (Lévy-Khintchine Formula)**.** A probability measure μ on $\mathbb{R}^d$ is infinitely divisible if and only if there exist $a \in \mathbb{R}^d$, a symmetric non-negative definite $d \times d$ matrix A and a Lévy measure ν on $\mathbb{R}^d - \{0\}$ s.t. the characteristic function of μ is given by

$$\phi_\mu(u) = \exp(ia^\top u - \frac{1}{2}u^\top Au + \int_{\mathbb{R}^d-\{0\}}(e^{iu^\top x} - 1 - iu^\top x\mathbb{1}_{\{|x|<1\}})\nu(dx)), \quad u \in \mathbb{R}^d$$

Conversely, for $\forall a \in \mathbb{R}^d$, symmetric non-negative definite $d \times d$ matrix A and Lévy measure ν on $\mathbb{R}^d - \{0\}$, there exists a unique probability measure μ on $\mathbb{R}^d$ whose characteristic function is given by the above formula.

Remark. The triplet (a, A, ν) is called the **characteristic** of μ.

Definition 5.5. $\eta(x) = ia^\top u - \frac{1}{2}u^\top Au + \int_{\mathbb{R}^d-\{0\}}(e^{iu^\top x} - 1 - iu^\top x\mathbb{1}_{\{|x|<1\}})\nu(dx)$ is called the **Lévy symbol**.

Remark. Every infinitely divisible probability measure μ has characteristic function $\phi_\mu(u) = \exp(\eta(u))$ where $\eta(u)$ is the Lévy symbol.

Propsition 5.2. $\eta(x)$ is defined as above, then:
(1) $Re(\eta(u)) \leq 0$ for $\forall u \in \mathbb{R}^d$.
(2) η is continuous on $\mathbb{R}^d$, uniformly on every bounded subset of $\mathbb{R}^d$.
(3) $\exists c > 0$ s.t. $|\eta(u)| \leq c(1 + |u|^2)$ for $\forall u \in \mathbb{R}^d$.

Proof:
(1) Since $|e^{iu^\top x}| = 1$ for $\forall x, u \in \mathbb{R}^d$, we have

$$\begin{aligned} Re(\eta(u)) &= -\frac{1}{2}u^\top Au + \int_{\mathbb{R}^d-\{0\}}(cos(u^\top x) - 1)\nu(dx) \\ &\leq 0 \end{aligned}$$

(2) $\forall x_1, x_2 \in \mathbb{R}^d$, define $I = \int_{\mathbb{R}^d-\{0\}} |e^{ix_1^\top y} - e^{ix_2^\top y} - i(x_1^\top y - x_2^\top y)\mathbb{1}_{\{|y|<1\}}|\nu(dy)$. Then

$$\begin{aligned} |\eta(x_1) - \eta(x_2)| &\leq |a^\top(x_1 - x_2)| + \frac{1}{2}|x_1^\top Ax_1 - x_2^\top Ax_2| + I \\ &\leq |a^\top(x_1 - x_2)| + \frac{1}{2}|(x_1 - x_2)^\top Ax_1| + \frac{1}{2}|x_1^\top A(x_1 - x_2)| + (x_2 - x_1)^\top A(x_1 - x_2) + I \end{aligned}$$

As for I, we have

$$I \leq \int_{\mathbb{R}^d-\{0\}} e^{ix_1^\top y}|1 - e^{i(x_2-x_1)^\top y}| + \int_{\mathbb{R}^d-\{0\}} i(x_2 - x_1)^\top y\mathbb{1}_{\{|y|<1\}}\nu(dy)$$

Hence after applying Taylor's expansion in the first part integral, we get $|\eta(x_1) - \eta(x_2)| \leq c_{x_1}|x_1 - x_2| + o(|x_1 - x_2|)$ where c_{x_1} is a constant depending on x_1.

$$c_{x_1} = |a| + 2|Ax_1| + \int_{\mathbb{R}^d-\{0\}} e^{ix_1^\top y}iy + iy\mathbb{1}_{\{|y|<1\}}\nu(dy)$$

Therefore, on bounded set, η is uniformly continuous.
(3) Since $|e^{iu^\top x} - 1 - iu^\top x \mathbb{1}_{\{|x|<1\}}| \le |u^\top x|^2 \mathbb{1}_{\{|x|<1\}} + 2\mathbb{1}_{\{|x|\ge 1\}}$, we have

$$\begin{aligned}|\eta(u)| &\le |a^\top u| + \frac{1}{2}|u^\top Au| + \int_{\mathbb{R}^d - \{0\}} |e^{iu^\top x} - 1 - iu^\top x \mathbb{1}_{\{|x|<1\}}| \nu(dx) \\ &\le |a^\top u| + \frac{1}{2}|u^\top Au| + |u|^2 \int_{\{|x|<1\}} |x|^2 \nu(dx) + 2\nu(\{|x| \ge 1\}) \\ &\le c(1 + |u|^2)\end{aligned}$$

where $c = |a| + \frac{1}{2}\|A\| + \int_{\{|x|<1\}} |x|^2 \nu(dx) + 2\nu(\{|x| \ge 1\})$. Hence, the conclusion holds. □

Theorem 5.2. Any infinitely divisible probability measure μ on $\mathbb{R}^d$ is a weak limit of a sequence of compound Poisson probability measures.

Proof:
Let ϕ be the characteristic function of μ, $X = Y_1^{(n)} + Y_2^{(n)} + \cdots + Y_n^{(n)}$ where $\{Y_i^{(n)}\}$ are i.i.d. r.v.s with characteristic function $\phi_{Y_1^{(n)}}(u) = \phi(u)^{1/n}$. Define Z_n is a compound Poisson r.v. with parameter $\lambda = n$ and distribution of $Y_1^{(n)}/n$. Then the characteristic function of Z_n is given by

$$\phi_{Z_n}(u) = \exp(n(\phi_{Y_1^{(n)}/n}(u) - 1)) = \exp(n(\phi_{Y_1^{(n)}}(u/n) - 1)) = \exp(n(\phi(u/n)^{1/n} - 1))$$

Thus, we have

$$\lim_{n\to\infty} \phi_{Z_n}(u) = \lim_{n\to\infty} \exp(n(\phi(u/n)^{1/n} - 1)) = \lim_{n\to\infty} \exp(\ln \phi(u/n)) = \phi(u)$$

By Lévy's continuity theorem, we have $Z_n \xrightarrow{d} X$. Hence, the conclusion holds. □

Definition 5.6. A stochastic process X is called **Lévy process** if:
(1) $X_0 = 0$ a.s..
(2) X has independent and stationary increments.
(3) X is stochastically continuous, i.e., for $\forall t \ge 0$, $\forall \epsilon > 0$, $\lim_{s\to t} \mathbb{P}(|X_s - X_t| > \epsilon) = 0$.

Example 5.1. A standard Brownian motion B w.r.t. $\{\mathcal{F}_t\}$ is a Lévy process.

Definition 5.7. A **Poisson process** with intensity $\lambda > 0$ is a Lévy process $N = \{N_t : t \ge 0\}$ with values in $\mathbb{N}$ s.t. for $\forall t \ge 0$, $N_t \sim Poisson(\lambda t)$.

Remark. The sample paths are not continuous.

Remark. Set $\tilde{N}_t = N_t - \lambda t$. Then $\tilde{N} = \{\tilde{N}_t : t \ge 0\}$ is called compensated Poisson process.

Definition 5.8. Compound Poisson process with parameters $\lambda > 0$ and distribution ν on $\mathbb{R}^d$ is a Lévy process $Z = \{Z_t : t \geq 0\}$ defined as follows:

$$Z_t = \sum_{i=1}^{N_t} X_i$$

where $N = \{N_t : t \geq 0\}$ is a Poisson process with intensity λ, $\{X_i\}$ are i.i.d. r.v.s with distribution ν and are independent of N.

Remark. For $\forall t \geq 0$ $Z_t \sim \text{Poisson}(\lambda t, \nu)$.

Propsition 5.3. Compound Poisson process Z with parameters $\lambda > 0$ and distribution ν on $\mathbb{R}^d$ is a Lévy process.

Proof:

(1) $Z_0 = 0$ a.s..

(2) For $\forall 0 \leq s < t$, $Z_t - Z_s = \sum_{i=N_s+1}^{N_t} X_i$. Increments are independent since N has independent increments and $\{X_i\}$ are independent of N. $Z_t - Z_s$ depends only on distribution of X_1 and $N_t - N_s$ that has same distribution as N_{t-s} depending only on $t - s$. Explicitly, we have

$$\begin{aligned}
\mathbb{P}(Z_t - Z_s \in A) &= \mathbb{E}[\mathbb{1}_{Z_t - Z_s \in A}] \\
&= \mathbb{E}[\mathbb{1}_{\{\sum_{i=N_s+1}^{N_t} X_i \in A\}}] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{\sum_{i=N_s+1}^{N_t} X_i \in A\}} | N]] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{\sum_{i=1}^{N_t - N_s} X_i \in A\}}]] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{\sum_{i=1}^{N_t - N_s} X_i \in A\}}] | X] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{\sum_{i=1}^{N_{t-s}} X_i \in A\}}] | X] \\
&= \mathbb{P}(Z_{t-s} \in A)
\end{aligned}$$

Thus, increments are stationary.

(3) For $\forall t \geq 0$, $\forall \epsilon > 0$, we have

$$\begin{aligned}
\lim_{s\to t} \mathbb{P}(|Z_s - Z_t| > \epsilon) &= \lim_{s\to t} \mathbb{P}(|\sum_{i=N_s+1}^{N_t} X_i| > \epsilon) \\
&= \lim_{s\to t} \mathbb{E}[\mathbb{1}_{\{|\sum_{i=N_s+1}^{N_t} X_i| > \epsilon\}}] \\
&= \lim_{s\to t} \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{|\sum_{i=N_s+1}^{N_t} X_i| > \epsilon\}} | N]] \\
&= \lim_{s\to t} \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{|\sum_{i=1}^{N_t - N_s} X_i| > \epsilon\}}]] \\
&= \lim_{s\to t} \mathbb{P}(|\sum_{i=1}^{N_{t-s}} X_i| > \epsilon)
\end{aligned}$$

Since $N_{t-s} \xrightarrow{d} 0$ as $s \to t$, we have $\mathbb{P}(N_{t-s} = 0) \to 1$ as $s \to t$. Thus, we have

$$\lim_{s\to t} \mathbb{P}(|\sum_{i=1}^{N_{t-s}} X_i| > \epsilon) \leq \lim_{s\to t} \mathbb{P}(N_{t-s} \geq 1) = 0$$

Hence, the conclusion holds. □

Remark. In general, if X is a Lévy process, then for $\forall t, s \geq 0$, $|X_t - X_s|$ has the same distribution as $X_{|t-s|}$.

Remark. For compound Poisson process, sample paths are not continuous.

Remark. A compound Poisson process Z is Poisson process if and only if the distribution ν is concentrated at point 1.

Propsition 5.4. X is a Lévy process on $\mathbb{R}^d$. Then for $\forall t \geq 0$, X_t has infinitely divisible distribution.

Proof:
For $\forall n \in \mathbb{N}$, we have

$$X_t = X_{\frac{t}{n}} + (X_{\frac{2t}{n}} - X_{\frac{t}{n}}) + \cdots + (X_t - X_{\frac{(n-1)t}{n}})$$

where $\{X_{\frac{(k+1)t}{n}} - X_{\frac{kt}{n}}\}_{k=0}^{n-1}$ are i.i.d. random variables. Hence, the conclusion holds. □

Lemma 5.1. If X is stochastically continuous Lévy process on $\mathbb{R}^d$, then $t \mapsto \phi_{X_t}(u)$ is uniformly continuous for $\forall u \in \mathbb{R}^d$.

Proof:
For $\forall \epsilon > 0$, $\exists \delta' > 0$ s.t. $\forall |y| < \delta'$, we have $|e^{iu^\top y} - 1| < \epsilon$. Since X is stochastically continuous,

$\exists\delta > 0$ s.t. $\forall|t-s| < \delta$, we have $\mathbb{P}(|X_t - X_s| > \delta') < \epsilon$. Thus, for $\forall|t-s| < \delta$, we have

$$\begin{aligned}
|\phi_{X_t}(u) - \phi_{X_s}(u)| &= |\mathbb{E}[e^{iu^\top X_t}] - \mathbb{E}[e^{iu^\top X_s}]| \\
&= |\mathbb{E}[e^{iu^\top X_s}(e^{iu^\top (X_t - X_s)} - 1)]| \\
&\leq \mathbb{E}[|e^{iu^\top (X_t - X_s)} - 1|] \\
&= \mathbb{E}[|e^{iu^\top (X_t - X_s)} - 1|\mathbb{1}_{\{|X_t - X_s| \leq \delta'\}}] + \mathbb{E}[|e^{iu^\top (X_t - X_s)} - 1|\mathbb{1}_{\{|X_t - X_s| > \delta'\}}] \\
&\leq \epsilon + 2\mathbb{P}(|X_t - X_s| > \delta') \\
&\leq 3\epsilon
\end{aligned}$$

Hence, the conclusion holds. $\square$

Theorem 5.3. *X is a Lévy process with Lévy simbol $\eta(u)$ of X_1. Then for $\forall t \geq 0$, the characteristic function of X_t is given by*

$$\phi_{X_t}(u) = \mathbb{E}[\exp(iu^\top X_t)] = \exp(t\eta(u)), \quad u \in \mathbb{R}^d$$

Proof:
For $\forall s, t \geq 0$, we have

$$\phi_{X_{t+s}}(u) = \mathbb{E}[\exp(iu^\top X_{t+s})] = \mathbb{E}[\exp(iu^\top X_t)\exp(iu^\top (X_{t+s} - X_t))] = \phi_{X_t}(u)\phi_{X_s}(u)$$

Thus, $f(t) = \phi_{X_t}(u)$ satisfies Cauchy's functional equation $f(t+s) = f(t)f(s)$ for $\forall t, s \geq 0$. Since X is stochastically continuous, by Lemma, $t \mapsto \phi_{X_t}(u)$ is continuous for $\forall u \in \mathbb{R}^d$. Hence, we have $\phi_{X_t}(u) = e^{c(u)t}$ for some constant $c(u)$ depending on u. By setting $t = 1$, we have $c(u) = \ln\phi_{X_1}(u) = \eta(u)$. Hence, the conclusion holds. $\square$

Corollary 5.1 (Lévy-Khintchine formula for Lévy process)**.** *X is a Lévy process on $\mathbb{R}^d$. Then for $\forall t \geq 0$, the characteristic function of X_t is given by*

$$\phi_{X_t}(u) = \exp(t(ia^\top u - \frac{1}{2}u^\top Au + \int_{\mathbb{R}^d - \{0\}}(e^{iu^\top x} - 1 - iu^\top x\mathbb{1}_{\{|x|<1\}})\nu(dx))), \quad u \in \mathbb{R}^d$$

where (a, A, ν) is the characteristic of X_1.

Remark. Define Lévy symbol and characteristic triplet of Lévy process X as those of X_1.

Propsition 5.5. *Let X be a Lévy process on $\mathbb{R}^d$ with characteristic triplet (a, A, ν). Then $-X = \{-X_t : t \geq 0\}$ is a Lévy process with characteristic triplet $(-a, A, \tilde{\nu})$, where $\tilde{\nu}(B) = \nu(B)$ for $\forall B \in \mathcal{B}(\mathbb{R}^d - \{0\})$.*

Theorem 5.4. Let $X^{(n)}$ be a sequence of Lévy processes, X be a process. If $\forall t > 0$, $X_t^{(n)} \overset{in\ prob}{\rightarrow} X_t$ as $n \to \infty$, then X is a Lévy process.

Proof:
(1)$X_0^{(n)} = 0$ a.s., thus $X_0^{(n)} \to 0$ a.s.. Since $X_0^{(n)} \to X_0$ in probability, we have $X_0 = 0$ a.s..
(2) For $\forall 0 \le t_1 < t_2 < \cdots < t_k$, we have

$$\lim_{n\to\infty} X_{t_{j+1}}^{(n)} - X_{t_j}^{(n)} = X_{t_{j+1}} - X_{t_j} \quad \text{in probability,} \quad j = 0, 1, \cdots, k-1$$

Thus there exists a subsequence $\{n_m\}$ s.t.

$$\lim_{m\to\infty} X_{t_{j+1}}^{(n_m)} - X_{t_j}^{(n_m)} = X_{t_{j+1}} - X_{t_j} \quad \text{a.s.,} \quad j = 0, 1, \cdots, k-1$$

Since $X_{t_{j+1}}^{(n_m)} - X_{t_j}^{(n_m)}$ are independent for $\forall m \in \mathbb{N}$, we have $X_{t_{j+1}} - X_{t_j}$ are independent. Thus the increments of X are independent.
For $\forall 0 \le s < t$, $X_t^{(n)} - X_s^{(n)} \to X_t - X_s$ in probability. There exists a subsequence $\{n_m\}$ s.t. $X_t^{(n_m)} - X_s^{(n_m)} \to X_t - X_s$ a.s.. By dominate convergence theorem (DCT), we have

$$\lim_{m\to\infty} \mathbb{E}[\exp(iu^\top (X_t^{(n_m)} - X_s^{(n_m)}))] = \mathbb{E}[\exp(iu^\top (X_t - X_s))]$$

Since $X_t^{(n)} - X_s^{(n)}$ has same distribution as $X_{t-s}^{(n)}$, we have

$$\lim_{m\to\infty} \mathbb{E}[\exp(iu^\top X_{t-s}^{(n_m)})] = \mathbb{E}[\exp(iu^\top (X_t - X_s))]$$

Thus, $X_t - X_s$ has same distribution as X_{t-s}. Hence, the increments of X are stationary.
(3) For $\forall t \ge 0$, $\forall \epsilon > 0$, we have

$$\lim_{s\to t} \mathbb{P}(|X_s - X_t| > \epsilon) \le \lim_{s\to t} \mathbb{P}(|X_s - X_s^{(n)}| > \epsilon/3) + \lim_{s\to t} \mathbb{P}(|X_s^{(n)} - X_t^{(n)}| > \epsilon/3) + \lim_{s\to t} \mathbb{P}(|X_t^{(n)} - X_t| > \epsilon/3)$$

By taking $n \to \infty$, we have

$$\lim_{s\to t} \mathbb{P}(|X_s - X_t| > \epsilon) = 0$$

Hence, the conclusion holds. □

Propsition 5.6. X is a process s.t. $X_0 = 0$ a.s.. Let P_t be the distribution of X_t for $\forall t \ge 0$. Then X is stochastically continuous $\iff$ P_t weakly converges to δ_0 as $t \to 0^+$, where δ_0 is the Dirac measure concentrated at point 0.

Definition 5.9 (Weak Convergence)**.** $\forall f \in C_b(\mathbb{R}^d)$, $\lim_{n\to\infty} \int_{\mathbb{R}^d} f(x)\mu_n(dx) = \int_{\mathbb{R}^d} f(x)\mu(dx)$. Then we say μ_n weakly converges to μ and denote $\mu_n \Rightarrow \mu$.

Proof:

"$\Rightarrow$": For $\forall f \in C_b(\mathbb{R}^d)$, $\epsilon > 0$, since f is uniformly continuous on bounded set, $\exists \delta > 0$ s.t. $\forall |x| < \delta$, $|f(x) - f(0)| < \epsilon$. Since X is stochastically continuous, $\exists t_0 > 0$ s.t. $\forall t < t_0$, $\mathbb{P}(|X_t| > \delta) < \epsilon$. Thus, for $\forall t < t_0$, we have

$$\begin{aligned}
|\int_{\mathbb{R}^d} f(x)P_t(dx) - f(0)| &= |\mathbb{E}[f(X_t)] - f(0)| \\
&\leq \mathbb{E}[|f(X_t) - f(0)|\mathbb{1}_{\{|X_t| \leq \delta\}}] + \mathbb{E}[|f(X_t) - f(0)|\mathbb{1}_{\{|X_t| > \delta\}}] \\
&\leq \epsilon + 2\|f\|_\infty \mathbb{P}(|X_t| > \delta) \\
&\leq \epsilon + 2\|f\|_\infty \epsilon
\end{aligned}$$

"$\Leftarrow$": For $\forall \epsilon > 0$, choose $g(x)$ to be supported on $B_r(0)$ with $0 \leq g \leq 1$ and $g(0) < 1 - \epsilon$. Since $P_t \Rightarrow \delta_0$ as $t \to 0^+$, we have

$$\begin{aligned}
\mathbb{P}(|X_t| > r) &= \int_{B_r(0)^c} P_t(dx) \\
&= 1 - \int_{B_r(0)} P_t(dx) \\
&\leq 1 - \int_{\mathbb{R}^d} g(x)P_t(dx) \\
&\to 1 - g(0) < \epsilon \quad \text{as } t \to 0^+
\end{aligned}$$

Hence, the conclusion holds. $\square$

Definition 5.10. A family of probability measures $\{P_t : t \geq 0\}$ on $\mathbb{R}^d$ is called **convolution semigroup** if:

(1) $P_0 = \delta_0$.

(2) $P_{t+s} = P_t * P_s$ for $\forall t, s \geq 0$, i.e. $\forall f \in C_b(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} f(x)P_{t+s}(dx) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x+y)P_t(dx)P_s(dy)$$

The semigroup is weakly continuous if $P_t \Rightarrow \delta_0$ as $t \to 0^+$.

Remark. A convolution semigroup $\{P_t : t \geq 0\}$ is weakly continuous $\iff$ $\forall f \in C_b(\mathbb{R}^d)$,

$$\lim_{s \to t^+} \int_{\mathbb{R}^d} f(x)P_s(dx) = \int_{\mathbb{R}^d} f(x)P_t(dx)$$

Corollary 5.2. X is a Lévy process on $\mathbb{R}^d$. Let P_t be the distribution of X_t for $\forall t \geq 0$. Then $\{P_t : t \geq 0\}$ is a weakly continuous convolution semigroup.

Proof:

Since X is Lévy process, then P_t is the distribution of $X_{t+s} - X_s$ for $\forall t, s \geq 0$. Thus, the semigroup is a convolution semigroup. Since X is stochastically continuous, by Proposition, $P_t \Rightarrow \delta_0$ as $t \to 0^+$. Hence, the conclusion holds. $\square$

Theorem 5.5. If $\{P_t\}$ is a weakly continuous convolution semigroup on $\mathbb{R}^d$, then there exists a Lévy process X on $\mathbb{R}^d$ s.t. for $\forall t \geq 0$, the distribution of X_t is P_t.

Proof:

Set $\Omega = \{w : \mathbb{R}_+ \to \mathbb{R}^d \text{ with } w(0) = 0\}$. Consider n-dimensional cylinder set in the form

$$I_{t_1,t_2,\cdots,t_n}^{A_1,A_2,\cdots,A_n} = \{w \in \Omega : w(t_1) \in A_1, w(t_2) \in A_2, \cdots, w(t_n) \in A_n\}$$

where $0 \leq t_1 < t_2 < \cdots < t_n$, $A_i \in \mathcal{B}(\mathbb{R}^d)$ for $i = 1, 2, \cdots, n$. Let $\mathcal{F}$ be the σ-algebra generated by all cylinder sets. Define probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ as follows:

$$\mathbb{P}(I_{t_1,t_2,\cdots,t_n}^{A_1,A_2,\cdots,A_n}) = \int_{A_1}\int_{A_2}\cdots\int_{A_n} P_{t_1}(dx_1)P_{t_2-t_1}(dx_2 - x_1)\cdots P_{t_n-t_{n-1}}(dx_n - x_{n-1})$$

By Carathéodory's extension theorem, $\mathbb{P}$ can be extended to a probability measure on $(\Omega, \mathcal{F})$. Define stochastic process $X = \{X_t : t \geq 0\}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ as $X_t(w) = w(t)$ for $\forall w \in \Omega$. Now it remains to check X is a Lévy process.

(1) By definition, $X_0(w) = w(0) = 0$ a.s..

(2) For $\forall 0 \leq s < t$, we have

$$\begin{aligned}
\mathbb{P}(X_t - X_s \in A) &= \mathbb{P}(\{w \in \Omega : w(t) - w(s) \in A\}) \\
&= \int_{\mathbb{R}^d}\int_{A+x} P_s(dx)P_{t-s}(dy - x) \\
&= \int_A P_{t-s}(dy)
\end{aligned}$$

Thus, increments are independent and stationary.

(3) For $\forall t \geq 0$, $\forall \epsilon > 0$, we have

$$\begin{aligned}
\lim_{s\to t} \mathbb{P}(|X_s - X_t| > \epsilon) &= \lim_{s\to t} \mathbb{P}(\{w \in \Omega : |w(s) - w(t)| > \epsilon\}) \\
&= \lim_{s\to t} \int_{\{|x|>\epsilon\}} P_{|t-s|}(dx)
\end{aligned}$$

Since $\{P_t\}$ is weakly continuous, by Proposition, we have

$$\lim_{s\to t} \int_{\{|x|>\epsilon\}} P_{|t-s|}(dx) = 0$$

Hence, the conclusion holds. □

Remark. Such X is called canonical Lévy process.

Corollary 5.3. If μ is an infinitely divisible probability measure on $\mathbb{R}^d$, then there exists a Lévy process X on $\mathbb{R}^d$ s.t. the distribution of X_1 is μ.

Proof:

Let ϕ be the characteristic function of μ, then $\phi_\mu(u) = \exp(\eta(u))$ where $\eta(u)$ is the Lévy symbol. Let $\phi_{\mu,t}(u) = \exp(t\eta(u))$ for $\forall t \geq 0$ be the characteristic function of an infinitely divisible probability measure. $\phi_{\mu,0} = 1 \Rightarrow P_0 = \delta_0$. For $\forall t, s \geq 0$, we have

$$\phi_{\mu,t+s}(u) = \exp((t+s)\eta(u)) = \exp(t\eta(u))\exp(s\eta(u)) = \phi_{\mu,t}(u)\phi_{\mu,s}(u)$$

Thus, $P_{t+s} = P_t * P_s$. Since $\eta(u)$ is continuous, $P_t \Rightarrow \delta_0$ as $t \to 0^+$. Hence, by Theorem, there exists a Lévy process X on $\mathbb{R}^d$ s.t. for $\forall t \geq 0$, the distribution of X_t is P_t. In particular, the distribution of X_1 is μ. □

Definition 5.11. A **sub-ordinator** is a one-dimensional Lévy process $T = \{T_t : t \geq 0\}$ with non-decreasing sample paths.

Theorem 5.6. If T is a sub-ordinator, then its Lévy symbol $\eta(u)$ has the form:

$$\eta(u) = ibu + \int_0^\infty (e^{iux} - 1)\nu(dx), \quad u \in \mathbb{R}$$

where $b \geq 0$ and ν is a measure on $(0, \infty)$ satisfying:
(1) ν is supported on $(0, \infty)$;
(2) $\int_0^\infty (1 \wedge x)\nu(dx) < \infty$.

Remark. (b, ν) is called the characteristic of sub-ordinator T.

Definition 5.12. **Laplace transformation** of sub-ordinator T is defined as follows:

$$\mathbb{E}[e^{-uT_t}] = \exp(-t\phi(u)), \quad u > 0$$

where

$$\phi(u) = -\eta_T(iu) = bu + \int_0^\infty (1 - e^{-ux})\nu(dx), \quad u > 0$$

Example 5.2. Poisson process ia always a sub-ordinator with characteristic $(\lambda, 0)$ where $\lambda > 0$ is the intensity.

Theorem 5.7. If μ is an infinitely divisible probability measure on $[0,\infty)$, then there exists a sub-ordinator T s.t. the distribution of T_1 is μ.

Proof:

By Lévy-Khintchine theorem, $\exists$ a Lévy process X on $\mathbb{R}$ s.t. the distribution of X_1 is μ. It remains to show X is a sub-ordinator. Since X is supported on $[0,\infty)$, then $X_1 \geq 0$ a.s.. For $\forall n \in \mathbb{N}$, the distribution fo X_1 is the distribution of the sum of n i.i.d. r.v.s $\{Y_i^{(n)}\}_{i=1}^n$. Suppose $\mathbb{P}(Y_1^{(n)} < 0) = p > 0$. Then we have

$$\mathbb{P}(X_1 < 0) = \mathbb{P}(\sum_{i=1}^n Y_i^{(n)} < 0) \geq \mathbb{P}(Y_i^{(n)} < 0, \forall i = 1, 2, \cdots, n) = p^n > 0$$

This contradicts to $X_1 \geq 0$ a.s.. Thus, we have $\mathbb{P}(Y_1^{(n)} < 0) = 0$ for $\forall n \in \mathbb{N}$. Since $Y_1^{(n)} \overset{d}{=} X_{1/n}$ and $X_{r/n} + X_{s/n} \overset{d}{=} X_{(r+s)/n}$ for $\forall r, s, n \in \mathbb{N}$, we have $X_{k/n} \geq 0$ a.s. for $\forall k = 0, 1, 2, \cdots$. And hence $X_t \geq 0$ a.s. for $\forall t \in \mathbb{Q}_+$. For $\forall t \geq 0$, there exists a sequence $\{t_n\} \subset \mathbb{Q}_+$ s.t. $t_n \downarrow t$. Since X is stochastically continuous, there exists a subsequence $\{t_{n_k}\}$ s.t. $X_{t_{n_k}} \to X_t$ a.s.. Since $X_{t_{n_k}} \geq 0$ a.s. for $\forall k \in \mathbb{N}$, we have $X_t \geq 0$ a.s.. Hence, $\mathbb{P}(X_t \geq X_s) = \mathbb{P}(X_t - X_s \geq 0) = \mathbb{P}(X_{t-s} \geq 0) = 1$ for $\forall t \geq s \geq 0$. Thus, X is a sub-ordinator. $\square$

Theorem 5.8 (Time Changing)**.** X is a Lévy process on $\mathbb{R}^d$, T is a sub-ordinator independent of X. Define a new process $Z = \{Z_t : t \geq 0\}$ as $Z_t = X_{T_t}$ for $\forall t \geq 0$. Then Z is a Lévy process on $\mathbb{R}^d$.

Proof:

(1) Since $X_0 = 0$ a.s. and $T_0 = 0$ a.s., we have $Z_0 = X_{T_0} = X_0 = 0$ a.s..

(2) For $\forall 0 \leq s < t$, we have

$$\begin{aligned}
\mathbb{P}(Z_t - Z_s \in A) &= \mathbb{P}(X_{T_t} - X_{T_s} \in A) \\
&= \mathbb{E}[\mathbb{1}_{\{X_{T_t} - X_{T_s} \in A\}}] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{X_{T_t} - X_{T_s} \in A\}} | T_t, T_s]] \\
&= \mathbb{E}[\mathbb{1}_{\{X_{T_t - T_s} \in A\}} | T_t, T_s] \\
&= \mathbb{E}[\mathbb{1}_{\{X_{T_{t-s}} \in A\}}] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{X_{T_{t-s}} \in A\}} | X]] \\
&= \mathbb{E}[\mathbb{1}_{\{X_{T_{t-s}} \in A\}}] \\
&= \mathbb{P}(Z_{t-s} \in A)
\end{aligned}$$

Thus increments are independent and stationary.

(3) For $\forall \epsilon > 0\,, \eta > 0\,, \exists \delta > 0$ s.t. $\forall h \in (0, \delta)$, we have $\mathbb{P}(|X_h| \geq \eta) < \epsilon$ and $\exists \delta' > 0$ s.t. $\forall h \in (0, \delta')$,

we have $\mathbb{P}(|T_h| > \delta) < \epsilon$. Thus, for $\forall h \in (0, \min\{\delta, \delta'\})$, we have

$$\begin{aligned}
\mathbb{P}(|Z_h| > \eta) &= \mathbb{P}(|X_{T_h}| > \eta) \\
&= \mathbb{P}(X_{T_h} > \eta, T_h \geq \delta) + \mathbb{P}(X_{T_h} > \eta, T_h < \delta) \\
&\leq \mathbb{P}(T_h \geq \delta) + \sup_{0 \leq u < \delta} \mathbb{P}(|X_u| > \eta, T_h < \delta) \\
&\leq \epsilon + \sup_{0 \leq u < \delta} \mathbb{P}(|X_u| > \eta) \\
&\leq 2\epsilon
\end{aligned}$$

Hence, the conclusion holds. □

Corollary 5.4. $\eta_Z(u) = \phi_T(\eta_X(u))$ where $\eta_X(u)$ and $\eta_Z(u)$ are Lévy symbols of X and Z respectively, and ϕ_T is the Laplace transformation of sub-ordinator T.

Proof:

$$\begin{aligned}
\exp(\eta_Z(u)t) &= \mathbb{E}[\exp(iu^\top Z_t)] = \mathbb{E}[\exp(iu^\top X_{T_t})] \\
&= \mathbb{E}[\mathbb{E}[\exp(iu^\top X_{T_t})|T_t]] = \mathbb{E}[\exp(\eta_X(u)T_t)] \\
&= \mathbb{E}[\exp(-(-\eta_X(u))T_t)] \\
&= \exp(-t\phi_T(-\eta_X(u)))
\end{aligned}$$

□

Definition 5.13. A Lévy process X is called **recurrent** (at origion) if $\liminf_{t\to\infty} |X_t| = 0$ a.s.. X is called **transient** if $\lim_{t\to\infty} |X_t| = \infty$ a.s..

Theorem 5.9. Lévy process X is recurrent $\iff$
either for some (and hence all) $a > 0$,

$$\lim_{q \to 0^+} \int_{B_a(0)} Re(\frac{1}{q - \eta(u)}) du = \infty$$

or $\forall a > 0$,

$$\int_{B_a(0)} Re(\frac{1}{-\eta(u)}) du = \infty$$

Theorem 5.10. X is a Lévy process with sumbol $\eta(u)$. Then for $\forall u \in \mathbb{R}^d$,

$$M_t^u = \exp(iu^\top X_t - t\eta(u))$$

is a martingale w.r.t. the natural filtration $\mathcal{F}_t^X = \sigma(X_s : s \leq t)$.

Proof:

(1) $\mathbb{E}[|M_t^u|] = \mathbb{E}[\exp(-tRe(\eta(u)))] < \infty$ for $\forall t \geq 0$.

(2) For $\forall s < t$, we have

$$\begin{aligned}
\mathbb{E}[M_t^u|\mathcal{F}_s^X] &= \mathbb{E}[\exp(iu^\top X_t - t\eta(u))|\mathcal{F}_s^X] \\
&= \exp(iu^\top X_s - t\eta(u))\mathbb{E}[\exp(iu^\top (X_t - X_s))|\mathcal{F}_s^X] \\
&= \exp(iu^\top X_s - t\eta(u))\mathbb{E}[\exp(iu^\top (X_t - X_s))] \\
&= \exp(iu^\top X_s - t\eta(u))\exp((t-s)\eta(u)) \\
&= M_s^u
\end{aligned}$$

Hence, the conclusion holds. $\square$

Theorem 5.11. Every Lévy process X on $\mathbb{R}^d$ has a modification Y that is a RCLL Lévy process.

Proof:

For $\forall u \in \mathbb{R}^d$, since M_t^u is a martingale, $M_{t^-}^u = \lim_{s\in\mathbb{Q}\to t^-} M_s^u$ and $M_{t^+}^u = \lim_{s\in\mathbb{Q}\to t^+} M_s^u$ exists a.s.. Define $\Omega_u^{'} = \{\omega :$ at least one limit fails $\}$, then $\mathbb{P}(\Omega_u^{'}) = 0$. Set $\Omega^{'} = \cup_{u\in\mathbb{Q}^d}\Omega_u^{'}$, then $\mathbb{P}(\Omega^{'}) = 0$. For $\forall \omega \in \Omega \setminus \Omega^{'}$, choose increasing sequence $\{s_n\} \subset \mathbb{Q}^+$ s.t. $s_n \to t$. Now consider $\lim_{n\to\infty} X_{s_n}(\omega)$. Assume $\{X_{s_n}(\omega) : n \in \mathbb{N}\}$ has two different limit points $a \neq b$. Then there exists $u \in \mathbb{Q}^d$ s.t. $e^{iu^\top a} \neq e^{iu^\top b}$. By taking subsequences if necessary, we may assume $\lim_{n\to\infty} X_{s_{2n}}(\omega) = a$ and $\lim_{n\to\infty} X_{s_{2n+1}}(\omega) = b$. Thus, we have

$$\lim_{n\to\infty} M_{s_{2n}}^u(\omega) = e^{iu^\top a - t\eta(u)} \neq e^{iu^\top b - t\eta(u)} = \lim_{n\to\infty} M_{s_{2n+1}}^u(\omega)$$

This contradicts to the existence of $M_{t^-}^u(\omega)$. Thus, $\{X_{s_n}(\omega) : n \in \mathbb{N}\}$ has a unique limit point. Similarly, we can show that $\lim_{n\to\infty} X_{r_n}(\omega)$ exists and is equal to the previous limit for any decreasing sequence $\{r_n\} \subset \mathbb{Q}^+$ s.t. $r_n \to t$. Define

$$Y_t(\omega) = \begin{cases} \lim_{s\in\mathbb{Q}\to t} X_s(\omega), & \omega \in \Omega \setminus \Omega^{'} \\ 0, & \omega \in \Omega^{'} \end{cases}$$

Then $Y = \{Y_t : t \geq 0\}$ is a RCLL process and is a modification of X. It remains to show Y is a Lévy process.

(1) Since $X_0 = 0$ a.s., we have $Y_0 = 0$ a.s..

(2) For $\forall 0 \le s < t$, we have

$$\begin{aligned}\mathbb{P}(Y_t - Y_s \in A) &= \mathbb{P}(\lim_{r\in\mathbb{Q}\to t} X_r - \lim_{q\in\mathbb{Q}\to s} X_q \in A)\\ &= \mathbb{P}(X_t - X_s \in A)\\ &= \mathbb{P}(X_{t-s} \in A)\end{aligned}$$

Thus increments are independent and stationary.
(3) Since Y is a modification of X, Y is stochastically continuous. Hence, the conclusion holds. □

Theorem 5.12. On $(\Omega, \mathcal{F}, \mathbb{P})$, X is a Lévy process w.r.t. filtration $\mathcal{F}_t$. T is a bounded stopping time w.r.t. $\mathcal{F}_t$. Set $X_t^{(T)} = X_{T+t} - X_T$ for $\forall t \ge 0$. Then:
(1) $X^{(T)}$ is a Lévy process independent of $\mathcal{F}_T$.
(2) $X_t^{(T)}$ has same distribution as X_t for $\forall t \ge 0$.
(3) $X^{(T)}$ is RCLL and is adapted to the filtration $\mathcal{F}_{T+t}$.

Proof:
$\forall A \in \mathcal{F}_T$, $\forall u_j \in \mathbb{R}^d$, $1 \le j \le n$, $0 \le t_1 < t_2 < \cdots < t_n$, we have

$$\begin{aligned}&\mathbb{E}[\mathbb{1}_A \exp(i\sum_{j=1}^n u_j^\top (X_{T+t_j} - X_{T+t_{j-1}}))]\\ &= \mathbb{E}[\mathbb{1}_A \exp(i\sum_{j=1}^n u_j^\top [(X_{T+t_j} - t_j\eta(u)) - (X_{T+t_{j-1}} - t_{j-1}\eta(u))] + i\sum_{j=1}^n u_j^\top (t_j - t_{j-1}))]\\ &= \mathbb{E}[\mathbb{1}_A \prod_{j=1}^n \frac{M_{T+t_j}^{u_j}}{M_{T+t_{j-1}}^{u_j}}] \prod_{j=1}^n \phi_{t_j - t_{j-1}}(u_j)\end{aligned}$$

where $\phi_t(u) = \mathbb{E}[\exp(iu^\top X_t)]$. Since $M_t^{u_j}$ is a martingale w.r.t. $\mathcal{F}_t$, by optional sampling theorem, we have

$$\begin{aligned}\mathbb{E}[\mathbb{1}_A \frac{M_{T+b}^u}{M_{T+a}^u}] &= \mathbb{E}[\mathbb{E}[\mathbb{1}_A \frac{M_{T+b}^u}{M_{T+a}^u} | \mathcal{F}_{T+a}]]\\ &= \mathbb{E}[\mathbb{1}_A \frac{1}{M_{T+a}^u} \mathbb{E}[M_{T+b}^u | \mathcal{F}_{T+a}]]\\ &= \mathbb{E}[\mathbb{1}_A \frac{M_{T+a}^u}{M_{T+a}^u}] = \mathbb{P}(A)\end{aligned}$$

And hence $\mathbb{E}[\mathbb{1}_A \frac{M_{T+b}^u}{M_{T+a}^u} Z] = \mathbb{E}[\mathbb{1}_A Z]$ for any bounded $\mathcal{F}_{T+a}$-measurable r.v. Z. By applying this iteratively, we have

$$\mathbb{E}[\mathbb{1}_A \exp(i\sum_{j=1}^n u_j^\top (X_{T+t_j} - X_{T+t_{j-1}}))] = \mathbb{P}(A) \prod_{j=1}^n \phi_{t_j - t_{j-1}}(u_j)$$

Thus, $X^{(T)}$ is independent of $\mathcal{F}_T$ and has independent and stationary increments.
(2) Take $A = \Omega$, $n = 1$, for $\forall u \in \mathbb{R}^d$, $t \geq 0$, we have

$$\mathbb{E}[\exp(iu^\top X_t^{(T)})] = \mathbb{E}[\exp(iu^\top X_t)]$$

Thus, $X_t^{(T)}$ has same distribution as X_t.
(1) $X_0^{(T)} = X_T - X_T = 0$ a.s.. Take $A = \Omega$, for $\forall n \in \mathbb{N}$ it holds that $X_{t_{j+1}}^{(T)} - X_{t_j}^{(T)} = X_{T+t_{j+1}} - X_{T+t_j}$, then

$$\mathbb{E}[\exp(i\sum_{j=1}^{n} u_j^\top (X_{t_{j+1}}^{(T)} - X_{t_j}^{(T)}))] = \prod_{j=1}^{n} \mathbb{E}[\exp(iu_j^\top X_{t_{j+1}-t_j})]$$

If $n = 1$, then we know increments are stationary. For $\forall n \in \mathbb{N}$, we know increments are independent. Now set $v_n = u_n$, $v_k = u_k - u_{k-1}$, $1 \leq k \leq n-1$, then

$$\begin{aligned}
\mathbb{E}[\mathbb{1}_A \exp(i\sum_{j=1}^{n} v_j^\top (X_{t_j}^{(T)}))] &= \mathbb{P}(A)\mathbb{E}[\exp(i\sum_{j=1}^{n} v_j^\top X_{t_j}^{(T)})] \\
&= \mathbb{P}(A)\prod_{j=1}^{n} \mathbb{E}[\exp(iv_j^\top (X_{t_j} - X_{t_{j-1}}))]
\end{aligned}$$

Then by definition, $\mathbb{E}[\exp(i\sum_{j=1}^{n} v_j^\top X_{t_j}^{(T)})|\mathcal{F}_T] = \mathbb{E}[\exp(i\sum_{j=1}^{n} v_j^\top X_{t_j}^{(T)})]$. Since n, v_k are arbitrary, $X^{(T)}$ is independent of $\mathcal{F}_T$.
(3) Since X is RCLL, $X^{(T)}$ is also RCLL. For $\forall t \geq 0$, since T is a stopping time, we have $\{T + t \leq s\} = \{T \leq s - t\} \in \mathcal{F}_{s-t} \subset \mathcal{F}_s$ for $\forall s \geq 0$. Thus, X_{T+t} is $\mathcal{F}_{T+t}$-measurable. Since X_T is also $\mathcal{F}_{T+t}$-measurable, $X_t^{(T)}$ is $\mathcal{F}_{T+t}$-measurable. Hence, the conclusion holds. □

Definition 5.14. Jump process of RCLL Lévy process X is defined as

$$\Delta X = \{\Delta X_t = X_t - X_{t^-} : t \geq 0\}$$

Theorem 5.13. Let N be a $\mathbb{N}$-valued Lévy process. If N is increasing a.s. and $\Delta N \in \{0, 1\}$ a.s., then N is a Poisson process.

Proof:
Define stopping times $T_0 = 0$ and $T_n = \inf\{t > T_{n-1} : N_t - N_{T_{n-1}} \neq 0\}$ for $n \in \mathbb{N}$. Then $\{T_n - T_{n-1}\}$ are i.i.d.. Since $T_1 = \inf\{t > 0 : N_t \neq 0\}$, for $\forall s, t \geq 0$, we have

$$\begin{aligned}
\mathbb{P}(T_1 > s + t) &= \mathbb{P}(N_r = 0, \forall r \in (0, s + t]) \\
&= \mathbb{P}(N_r = 0, \forall r \in (0, s])\mathbb{P}(N_{s+r} - N_s = 0, \forall r \in (0, t]) \\
&= \mathbb{P}(T_1 > s)\mathbb{P}(T_1 > t)
\end{aligned}$$

Let $f(t) = \mathbb{P}(T_1 > t)$ for $\forall t \geq 0$. Then $f(s+t) = f(s)f(t)$ and f is right continuous with $f(0) = 1$. Thus, $f(t) = e^{-\lambda t}$ for some $\lambda \geq 0$. $\mathbb{P}(T_1 \leq t) = 1 - e^{-\lambda t}$ implies $T_n = \sum_{i=1}^{n}(T_i - T_{i-1})$ has Gamma distribution with parameters (n, λ). Thus, for $\forall n \in \mathbb{N}$, we have

$$\begin{aligned}
\mathbb{P}(N_t = n+1) &= \mathbb{P}(T_{n+2} > t, T_{n+1} \leq t) \\
&= \mathbb{P}(T_{n+2} > t) - \mathbb{P}(T_{n+2} > t, T_{n+1} > t) \\
&= \mathbb{P}(T_{n+2} > t) - \mathbb{P}(T_{n+1} > t) \\
&= \int_t^\infty \frac{\lambda^{n+2} s^{n+1} e^{-\lambda s}}{(n+1)!} ds - \int_t^\infty \frac{\lambda^{n+1} s^n e^{-\lambda s}}{n!} ds \\
&= \frac{(\lambda t)^{n+1} e^{-\lambda t}}{(n+1)!}
\end{aligned}$$

□

Lemma 5.2. If X is a Lévy process, then fix $t \geq 0$, $\Delta X_t = 0$ a.s..

Proof:

Assume X is RCLL. Let $\{t_n : n \in \mathbb{N}\} \subset \mathbb{R}^+$ and $t_n \downarrow t$. Since X is stochastically continuous, there exists a subsequence $\{t_{n_k}\}$ s.t. $X_{t_{n_k}} \to X_t$ a.s.. Thus, we have

$$\Delta X_t = X_t - X_{t^-} = X_t - \lim_{k\to\infty} X_{t_{n_k}} = 0 \text{ a.s.}$$

□

Definition 5.15. Let X be a RCLL Lévy process. The **Poisson random measure** associated with the jump process ΔX is defined as follows:

$$N(t, A) = \sum_{0<s\leq t} \mathbb{1}_{\{\Delta X_s \in A\}}, \quad t \geq 0, A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$$

Then $N = \{N(t, A) : t \geq 0, A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})\}$ is a Borel measure on $\mathbb{R}^d \setminus \{0\}$ for $\forall t \geq 0$ and is a counting measure on $\mathbb{R}_+$ for $\forall A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$. We call $\mu(\cdot) = \mathbb{E}[N(1, \cdot)]$ the intensity measure of X.

Remark. If A is bounded away from 0, i.e., $\exists \epsilon > 0$ s.t. $A \subset \{x \in \mathbb{R}^d : |x| \geq \epsilon\}$, then $N(t, A) < \infty$ a.s. for $\forall t \geq 0$.

Theorem 5.14. For $\forall A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ bounded away from 0, $N(t, A)$ is a Poisson process with intensity $\mu(A)$.

Proof:
$N(0,A)=0$ a.s.. For $\forall 0\le s<t$, we have

$$\{N(t,A)-N(s,A)\ge n\}=\{\exists s<t_1<t_2<\cdots<t_n\le t:\Delta X_{t_j}\in A,\,1\le j\le n\}\in\sigma(X_u-X_v:s<v<u<t)$$

And hence increments are independent. Since $\Delta X_{t_j}\overset{d}{=}\Delta X_{t_j-s}$ for $\forall j=1,2,\cdots,n$, increments are stationary.

$$\begin{aligned}\mathbb{P}(N(t,A)=0)&=\mathbb{P}(N(\frac{kt}{n},A)-N(\frac{(k-1)t}{n},A)=0,\,1\le k\le n)\\&=\prod_{k=1}^{n}\mathbb{P}(N(\frac{kt}{n},A)-N(\frac{(k-1)t}{n},A)=0)\\&=\left(\mathbb{P}(N(\frac{t}{n},A)=0)\right)^n\end{aligned}$$

Take $\lim_{n\to\infty}$ on both sides, we have

$$\mathbb{P}(N(\frac{t}{n},A)=0)=\lim_{n\to\infty}\left(\mathbb{P}(N(t,A)=0)\right)^n=\begin{cases}1, & \mathbb{P}(N(t,A)=0)>0\\0, & \mathbb{P}(N(t,A)=0)=0\end{cases}$$

Since $N(t,A)\ge 0$ and is increasing in t, then $\lim_{t\to 0^+}\mathbb{P}(N(t,A)=0)$ exists.

If $\mathbb{P}(N(t,A)=0)=0$ for some t, then $\mathbb{P}(N(nt,A)=0)=\mathbb{P}(N(t,A)=0)^n=0$ for $\forall n\in\mathbb{N}$. And hence $\mathbb{P}(N(t,A)=0)=0$ for $\forall t>0$. Thus X is RC around 0 a.s.. This is because if there exists a sequence $\{t_n\}\subset\mathbb{R}^+$ s.t. $t_n\downarrow 0$ and $\Delta X_{t_n}\in A$ for $\forall n\in\mathbb{N}$, then $\exists s_n<t_n$ s.t. $\Delta X_{s_n}\in A$ and $|X_{s_n}-X_{s_n^-}|\ge\epsilon$ where $\epsilon=\inf\{|x|:x\in A\}>0$. Then either $|X_{s_n}|>\epsilon/2$ or $|X_{s_n^-}|>\epsilon/2$. And then $\exists s_n^{'}$ s.t. $|X_{s_n^{'}}|>\epsilon/3$ and $s_n^{'}\downarrow 0$. This contradicts $X_{s_n^{'}}\to X_0=0$ a.s..

When $\mathbb{P}(N(t,A)=0)>0$ for $\forall t>0$, we have $\lim_{t\to 0^+}\mathbb{P}(N(t,A)=0)=1$. Thus, there exists $\lambda_A\ge 0$ s.t. $\mathbb{P}(N(t,A)=0)=e^{-\lambda_A t}$ for $\forall t\ge 0$. Hence, for $\forall n\in\mathbb{N}$, we have

$$\begin{aligned}\mathbb{P}(N(t,A)=n)&=\mathbb{P}(N(\frac{(n+1)t}{n},A)-N(\frac{nt}{n},A)=0,N(\frac{nt}{n},A)-N(\frac{(n-1)t}{n},A)=0,\cdots,N(\frac{t}{n},A)=n)\\&=\mathbb{P}(N(\frac{(n+1)t}{n},A)-N(\frac{nt}{n},A)=0)\mathbb{P}(N(\frac{nt}{n},A)-N(\frac{(n-1)t}{n},A)=0)\cdots\mathbb{P}(N(\frac{t}{n},A)=n)\\&=\frac{(\lambda_A t)^n e^{-\lambda_A t}}{n!}\end{aligned}$$

□

Remark. If $A_1,A_2,\cdots,A_n\in\mathcal{B}(\mathbb{R}^d\setminus\{0\})$ are disjoint and bounded away from 0, $s_1,s_2,\cdots,s_n\ge 0$ are disjoint, then $N(s_1,A_1),N(s_2,A_2),\cdots,N(s_n,A_n)$ are independent.

Definition 5.16. $(\Omega,\mathcal{F},\mathbb{P})$ is a probability space. Let S be a set, $\mathcal{A}$ be an algebra closed under finite union and complement of subsets of S. A **random measure** M on $(S,\mathcal{A})$ is a collection of

r.v.s $\{M(B) : B \in \mathcal{A}\}$ satisfying:
(1) $M(\emptyset) = 0$
(2) $M(A \sqcup B) = M(A) + M(B)$

Remark. A ramdom measure is σ-additive if (2) is replaced by $M(\bigsqcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} M(A_n)$ with $\bigsqcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Definition 5.17. A ramdom measure M on $(S, \mathcal{A})$ is called **independently scattered** if for any disjoint $A_1, A_2, \cdots, A_n \in \mathcal{A}$, $M(A_1), M(A_2), \cdots, M(A_n)$ are independent r.v.s.

Definition 5.18. S is a set. $\mathcal{S}$ is a σ-algebra on S. $\mathcal{A} \subset \mathcal{S}$ is an algebra. A **Poisson random measure** on $(S, \mathcal{S})$ is an independently scattered random measure M on $(S, \mathcal{S})$ satisfying: for $\forall A \in \mathcal{A}$, $M(A) < \infty$ is a Poisson r.v..

Remark. $\lambda(B) = \mathbb{E}[M(B)]$ can be extended to a σ-finite measure on $(S, \mathcal{S})$.

Theorem 5.15. Let S be set, $\mathcal{S}$ be a σ-algebra on S, λ be a σ-finite measure on $(S, \mathcal{S})$. Then there exists a Poisson random measure M on $(S, \mathcal{S})$ s.t. $\mathbb{E}[M(B)] = \lambda(B)$ for $\forall B \in \mathcal{S}$. In this case $\mathcal{A} = \{A \in \mathcal{S} : \lambda(A) < \infty\}$ is an algebra.

Remark. X is a RCLL Lévy process on $\mathbb{R}^d$. $S = \mathbb{R}^d \setminus \{0\}$, $\mathcal{S} = \mathcal{B}(\mathbb{R}^d \setminus \{0\})$, $\mathcal{A} =$ the algebra generated by subsets of S that are bounded away from 0. For $\forall t \geq 0$, $M_t(A) = N(t, A)$ is a Poisson random measure with $\lambda(A) = t\mu(A)$ where $\mu(A) = \mathbb{E}[N(1, A)]$. Set $M([s, t] \times A) = M_t(A) - M_s(A) = N(t, A) - N(s, A)$ for $\forall 0 \leq s < t$, $A \in \mathcal{S}$. Then M can be extended to a σ-additive Poisson random measure on $(\mathbb{R}_+ \times S, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{S})$ with intensity measure $Leb \times \mu$ where Leb is the Lebesgue measure on $\mathbb{R}_+$. And $\lambda(dt, dx) = dt\mu(dx)$.

Remark. $\tilde{N}(t, A) = N(t, A) - t\mu(A)$ is called the compensated Poisson random measure associated with N.

Definition 5.19. X is a RCLL Lévy process on $\mathbb{R}^d$. N is the Poisson random measure associated with the jump process of X. μ is the intensity measure of X. For $\forall A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ bounded away from 0, $\forall f : \mathbb{R}^d \to \mathbb{R}^d$ define the **Poisson integral** as

$$\int_A f(x) N(t, dx) = \sum_{0 < s \leq t} f(\Delta X_s) \mathbb{1}_{\{\Delta X_s \in A\}}, \quad t \geq 0$$

provided the sum converges a.s..

Propsition 5.7. X is a RCLL Lévy process on $\mathbb{R}^d$. N is the Poisson random measure associated with the jump process of X. μ is the intensity measure of X. For $\forall A \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ bounded away from 0, $\forall f : \mathbb{R}^d \to \mathbb{R}^d$ s.t. $\int_A |f(x)| \mu(dx) < \infty$, We define a measure $\mu_{f,A}(B) = \mu(A \cap f^{-1}(B))$ for

$\forall B \in \mathcal{B}(\mathbb{R}^d)$. Then:
(1) $\forall t \geq 0$, if $\int_A |f(x)|\mu(dx) < \infty$, then $\int_A f(x)N(t,dx)$ has compound Poisson distribution with characteristic function

$$\mathbb{E}[\exp(iu^\top \int_A f(x)N(t,dx))] = \exp(t\int_{\mathbb{R}^d}(e^{iu^\top x} - 1)\mu_{f,A}(dx))$$

(2) If $\int_A |f(x)|\mu(dx) < \infty$, then for $\forall t \geq 0$, we have

$$\mathbb{E}[\int_A f(x)N(t,dx)] = t\int_A f(x)\mu(dx)$$

(3) If $\int_A |f(x)|\mu(dx) < \infty$ and $\int_A |f(x)|^2\mu(dx) < \infty$, then for $\forall t \geq 0$, we have

$$Var(\int_A f(x)N(t,dx)) = t\int_A |f(x)|^2\mu(dx)$$

Proof:
$\mathbb{E}[\exp(iu^\top X)] = \phi(u)$, then take derivative w.r.t. u on both sides, then $i\mathbb{E}[X] = \phi^{'}(0)$, and $-\mathbb{E}[X^2] = \phi^{''}(0)$. And hence (2) and (3) are proved.

(1)We first consider $f = \sum_{j=1}^n c_j \mathbb{1}_{A_j}$ where $c_j \in \mathbb{R}^d$, $A_j \in \mathcal{B}(\mathbb{R}^d \setminus \{0\})$ are disjoint and bounded away from 0 for $j = 1, 2, \cdots, n$. Then $N_(t, A_j)$ are independent.

$$\begin{aligned}
\int_A f(x)N(t,dx) &= \sum_{x\in A} f(x)N(t,\{x\}) \\
&= \sum_{j=1}^n \sum_{x\in A_j} c_j N(t,\{x\}) \\
&= \sum_{j=1}^n c_j N(t, \cup_{x\in A_j}\{x\}) \\
&= \sum_{j=1}^n c_j N(t, A_j)
\end{aligned}$$

Thus, we have

$$\begin{aligned}
\mathbb{E}[\exp(iu^\top \int_A f(x)N(t,dx))] &= \mathbb{E}[\exp(iu^\top \sum_{j=1}^n c_j N(t,A_j))] \\
&= \prod_{j=1}^n \mathbb{E}[\exp(iu^\top c_j N(t,A_j))] \\
&= \prod_{j=1}^n \exp(t\mu(A_j)(e^{iu^\top c_j}-1)) \\
&= \exp(t\sum_{j=1}^n \mu(A_j)(e^{iu^\top c_j}-1)) \\
&= \exp(t\int_{\mathbb{R}^d}(e^{iu^\top f(x)}-1)\mu(dx))
\end{aligned}$$

Since on A, $\exp(iu^\top f(x))-1=\sum_{j=1}^n (e^{iu^\top c_j}-1)\mathbb{1}_{A_j}(x)$.

Next, consider $f\geq 0$. Choose $\phi_k\geq 0$ s.t. $\phi_k\uparrow f$ and $\phi_k=\sum_{j=1}^{n_k} c_{j,k}\mathbb{1}_{A_{j,k}}$ where $c_{j,k}\in\mathbb{R}^d$, $A_{j,k}\in\mathcal{B}(\mathbb{R}^d\setminus\{0\})$ are disjoint and bounded away from 0 for $j=1,2,\cdots,n_k$. Then by Monotone Convergence Theorem, we have

$$\begin{aligned}
\mathbb{E}[\exp(iu^\top \int_A f(x)N(t,dx))] &= \lim_{k\to\infty}\mathbb{E}[\exp(iu^\top \int_A \phi_k(x)N(t,dx))] \\
&= \lim_{k\to\infty}\exp(t\int_{\mathbb{R}^d}(e^{iu^\top \phi_k(x)}-1)\mu(dx)) \\
&= \exp(t\int_{\mathbb{R}^d}(e^{iu^\top f(x)}-1)\mu(dx))
\end{aligned}$$

Finally, consider general f. Set $f^+=\max\{f,0\}$ and $f^-=\max\{-f,0\}$. Then $f=f^+-f^-$. By applying the previous result to f^+ and f^- respectively, we have

$$\begin{aligned}
\mathbb{E}[\exp(iu^\top \int_A f(x)N(t,dx))] &= \mathbb{E}[\exp(iu^\top \int_A f^+(x)N(t,dx))]\mathbb{E}[\exp(-iu^\top \int_A f^-(x)N(t,dx))] \\
&= \exp(t\int_{\mathbb{R}^d}(e^{iu^\top f^+(x)}-1)\mu(dx))\exp(t\int_{\mathbb{R}^d}(e^{-iu^\top f^-(x)}-1)\mu(dx)) \\
&= \exp(t\int_{\mathbb{R}^d}(e^{iu^\top f(x)}-1)\mu(dx))
\end{aligned}$$

$\square$

Theorem 5.16 (Lévy-Itô Decomposition)**.** X is a RCLL Lévy process on $\mathbb{R}^d$. N is the Poisson random measure associated with the jump process of X. μ is the intensity measure of X. Then there exists $a\in\mathbb{R}^d$, a positive semi-definite matrix $Q\in\mathbb{R}^{d\times d}$, and a Brownian motion B on $\mathbb{R}^d$

w.r.t. filtration $\mathcal{F}_t$ s.t.

$$X_t = at + Q^{1/2}B_t + \int_{|x|\geq 1} xN(t,dx) + \int_{|x|<1} x\tilde{N}(t,dx), \quad t \geq 0$$

where the four terms on the right-hand side are independent. And $Q^{1/2}$ is the square root of Q, i.e., $Q^{1/2}(Q^{1/2})^\top = Q$.

Definition 5.20. A r.v. X is **stable** if $\forall n \in \mathbb{N}$, $\exists a_n > 0$, $b_n \in \mathbb{R}$ s.t.

$$X_1 + X_2 + \cdots + X_n \stackrel{d}{=} a_n X + b_n$$

where $X_1, X_2, \cdots, X_n$ are i.i.d. copies of X.

Definition 5.21. X is said to be **strictly stable** if in the above definition, $b_n = 0$ for $\forall n \in \mathbb{N}$.

Remark. If $\exists a_n = \sigma n^{1/\alpha}$ for some $\alpha \in (0,2]$ s.t. X is stable, then α is called the index of stability of X.

Example 5.3. Central Limit Theorem: Let $\{X_n : n \in \mathbb{N}\}$ be i.i.d. r.v.s with mean μ and variance $\sigma^2 < \infty$. Set $S_n = \sum_{i=1}^n X_i$. Then

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{d} N(0,1)$$

as $n \to \infty$.In this case, $N(0,1)$ is stable with index of stability $\alpha = 2$.

Propsition 5.8. Stable $\Rightarrow$ infinitely divisible.

Theorem 5.17. X is real-valued stable r.v. with characteristic (a, A, ν). Then one of the following holds:
(1) $\alpha = 2$, $\nu = 0$ so $X \sim \mathcal{N}(a, A)$;
(2) $\alpha \in (0,2)$, $A = 0$ and ν is given by

$$\nu(dx) = \begin{cases} \frac{c_1}{x^{1+\alpha}} dx, & x > 0 \\ \frac{c_2}{|x|^{1+\alpha}} dx, & x < 0 \end{cases}$$

for some $c_1, c_2 \geq 0$ with $c_1 + c_2 > 0$. In this case, we have $\mathbb{P}(|X| > y) \sim y^{-\alpha}$ as $y \to \infty$. Moreover, $\mathbb{E}[|X|^p] < \infty$ for $0 < p < \alpha$ and $\mathbb{E}[|X|^p] = \infty$ for $p \geq \alpha$.

Definition 5.22. A **stable Lévy process** is a Lévy process X s.t. X_t is a stable r.v. for $\forall t > 0$.

Remark. $\eta(u) = -\sigma^\alpha |u|^\alpha$ for ratational invariant stable Lévy process $X_{at} = a^{1/\alpha} X_t$ for $\forall a > 0$.

6 Brownian Sheet

Definition 6.1. A collection $X = \{X_t : t \in \mathbb{R}_+^d\}$ of r.v.s is called a **random field**.

Notation: For $s = (s_1, s_2, \cdots, s_d), t = (t_1, t_2, \cdots, t_d) \in \mathbb{R}_+^d$, we write $s \prec t$ if $s_i \leq t_i$ for $i = 1, 2, \cdots, d$.

Definition 6.2. $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$ is a family of sub-σ-algebras of $\mathcal{F}$ satisfying:
(1) $\mathcal{F}_z \subset \mathcal{F}_{z'}$ if $z \prec z'$;
(2) $\forall A$ with $\mathbb{P}(A) = 0$, $A \in \mathcal{F}_z$ for $\forall z \in \mathbb{R}_+^d$;
(3) $\forall z \in \mathbb{R}_+^2$, $\mathcal{F}_z = \cap_{z \prec z'} \mathcal{F}_{z'}$;
(4) $\forall z = (s,t) \in \mathbb{R}_+^2$, σ-algebra $\mathcal{F}_z^1 = \mathcal{F}_{s\infty} = \sigma(\cup_{t' \geq 0} \mathcal{F}_{(s,t')})$ and $\mathcal{F}_z^2 = \mathcal{F}_{\infty t} = \sigma(\cup_{s' \geq 0} \mathcal{F}_{(s',t)})$ are conditionally independent given $\mathcal{F}_z$;
Then $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$ is called a **filtration**.

Remark. (4') $\forall$ bounded r.v. X, $\forall z \in \mathbb{R}_+^2$, we have

$$\mathbb{E}[\mathbb{E}[X|\mathcal{F}_z^1]|\mathcal{F}_z^2] = \mathbb{E}[\mathbb{E}[X|\mathcal{F}_z^2]|\mathcal{F}_z^1] = \mathbb{E}[X|\mathcal{F}_z]$$

Remark. $\mathcal{F}_z^1 \cap \mathcal{F}_z^2 = \mathcal{F}_z$ for $\forall z \in \mathbb{R}_+^2$.

Example 6.1. $\{f_s^{(1)} : s \geq 0\}$ and $\{f_t^{(2)} : t \geq 0\}$ are two independent filtrations on $(\Omega, \mathcal{F}, \mathbb{P})$. Define $\mathcal{F}_{(s,t)} = \sigma(\mathcal{F}_s^{(1)} \cup \mathcal{F}_t^{(2)} \cup \{A : \mathbb{P}(A) = 0\})$ for $\forall (s,t) \in \mathbb{R}_+^2$. Then $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$ is a two-parameter filtration.

Definition 6.3. A random field $X = \{X_z : z \in \mathbb{R}_+^2\}$ is **adapted** to filtration $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$ if X_z is $\mathcal{F}_z$-measurable for $\forall z \in \mathbb{R}_+^2$. X is measurable if $(z, \omega) \mapsto X_z(\omega)$ is measurable from $(\mathbb{R}_+^2 \times \Omega, \mathcal{B}(\mathbb{R}_+^2) \otimes \mathcal{F})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Definition 6.4. A ramdom field $M = \{M_z : z \in \mathbb{R}_+^2\}$ is called a **martingale** w.r.t. filtration $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$ if:
(1) M is adapted to $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$;
(2) $\mathbb{E}[|M_z|] < \infty$ for $\forall z \in \mathbb{R}_+^2$;
(3) $z, z' \in \mathbb{R}_+^2$, $z \prec z'$, then $\mathbb{E}[M_{z'}|\mathcal{F}_z] = M_z$ a.s..

Notation: For $z, z' \in \mathbb{R}_+^2$, $z \prec z'$, define $(z, z']$ as the rectangle $(s, s'] \times (t, t']$ where $z = (s,t)$ and $z' = (s', t')$. And the increment of random field X on $(z, z']$ is defined as

$$X(z, z') = X_{z'} - X_{(s,t')} - X_{(s',t)} + X_z$$

Definition 6.5. Let $M = \{X_z : z \in \mathbb{R}_+^2\}$ be integrable and adapted.
(1) M is called a **weak martingale** if for $\forall z, z^{'} \in \mathbb{R}_+^2$, $z \prec z^{'}$, we have $\mathbb{E}[M(z, z^{'})|\mathcal{F}_z] = 0$ a.s..
(2) M is called a **strong martingale** if for $\forall z, z^{'} \in \mathbb{R}_+^2$, $z \prec z^{'}$, we have $\mathbb{E}[M(z, z^{'})|\sigma(\mathcal{F}_z^1, \mathcal{F}_z^2)] = 0$. And additionally, $M_z = 0$ if z has at least one coordinate equal to 0.

Definition 6.6. For $i = 1, 2$, M is called i**-martingale** if
(1) M is adapted and integrable;
(2) $\forall z, z^{'} \in \mathbb{R}_+^2$, $z \prec z^{'}$, we have $\mathbb{E}[M(z, z^{'})|\mathcal{F}_z^i] = 0$ a.s..

Propsition 6.1. Strong martingale $\Rightarrow$ martingale $\Rightarrow$ i-martingale for $i = 1, 2$ $\Rightarrow$ weak martingale.

Proof:
Strong martingale $\Rightarrow$ martingale: For $\forall z, z^{'} \in \mathbb{R}_+^2$, $z \prec z^{'}$, by taking expectation on both sides of $\mathbb{E}[M(z, z^{'})|\sigma(\mathcal{F}_z^1, \mathcal{F}_z^2)] = 0$, we have $\mathbb{E}[M(z, z^{'})] = 0$. And hence

$$\mathbb{E}[M_{z'}|\mathcal{F}_z] = \mathbb{E}[M_{(s,t')}|\mathcal{F}_z] + \mathbb{E}[M_{(s',t)}|\mathcal{F}_z] - \mathbb{E}[M_z|\mathcal{F}_z] = M_{(s,t')} + M_{(s',t)} - M_z = M_z$$

Martingale $\Rightarrow$ i-martingale for $i = 1, 2$: For $\forall z, z^{'} \in \mathbb{R}_+^2$, $z \prec z^{'}$, by taking expectation on both sides of $\mathbb{E}[M_{z'}|\mathcal{F}_z] = M_z$ w.r.t. $\mathcal{F}_z^i$, we have

$$\mathbb{E}[M(z, z^{'})|\mathcal{F}_z^i] = \mathbb{E}[M_{z'}|\mathcal{F}_z^i] - \mathbb{E}[M_{(s,t')}|\mathcal{F}_z^i] - \mathbb{E}[M_{(s',t)}|\mathcal{F}_z^i] + \mathbb{E}[M_z|\mathcal{F}_z^i] = M_z - M_{(s,t')} - M_{(s',t)} + M_z = 0$$

i-martingale for $i = 1, 2$ $\Rightarrow$ weak martingale: For $\forall z, z^{'} \in \mathbb{R}_+^2$, $z \prec z^{'}$, we have

$$\mathbb{E}[M(z, z^{'})|\mathcal{F}_z] = \mathbb{E}[\mathbb{E}[M(z, z^{'})|\mathcal{F}_z^i]|\mathcal{F}_z] = \mathbb{E}[0|\mathcal{F}_z] = 0$$

□

Remark. The converses of the above implications do not hold in general.

Definition 6.7. $X = \{X_z : z \in \mathbb{R}_+^2\}$ is called a **Gaussian random field** if for $\forall n \in \mathbb{N}$, $\forall z_1, z_2, \cdots, z_n \in \mathbb{R}_+^2$, the r.v. $(X_{z_1}, X_{z_2}, \cdots, X_{z_n})$ has a multivariate normal distribution.

Definition 6.8. A centered Gaussian random field $X = \{X_z : z \in \mathbb{R}_+^2\}$ is called a **Brownian sheet** if for $\forall z = (s, t), z^{'} = (s^{'}, t^{'}) \in \mathbb{R}_+^2$, we have

$$Cov(X_z, X_{z'}) = (s \wedge s^{'})(t \wedge t^{'})$$

Remark. $B, B^{'}$ are two independent Brownian motions on $\mathbb{R}$. Define $W_{(s,t)} = B_s B_t^{'}$ for $\forall (s, t) \in \mathbb{R}_+^2$. Then W is not a Brownian sheet, since $B_t B_t^{'}$ is not normal for $\forall t > 0$.

Definition 6.9. Let $(S, \mathcal{S})$ be a σ-finite measurable space. A random set function W on $\{A \in \mathcal{S} : \nu(A) < \infty\}$ satisfying:
(1) For $\forall A \in \mathcal{S}$ with $\nu(A) < \infty$, $W(A) \sim \mathcal{N}(0, \nu(A))$;
(2) For disjoint $A_1, A_2, \cdots, A_n \in \mathcal{S}$ with $\nu(A_j) < \infty$ for $j = 1, 2, \cdots, n$, $W(A_1), W(A_2), \cdots, W(A_n)$ are independent;
(3) For disjoint $A_1, A_2, \cdots, A_n \in \mathcal{S}$ with $\nu(A_j) < \infty$ for $j = 1, 2, \cdots, n$, we have $W(\cup_{j=1}^n A_j) = \sum_{j=1}^n W(A_j)$ a.s..
Then W is called a **white noise** based on measure ν.

Remark. If W is a white noise, ν is Lebesgue measure on $\mathbb{R}^d$, set $X_t = W([0, t_1] \times [0, t_2] \times \cdots \times [0, t_d])$ for $\forall t = (t_1, t_2, \cdots, t_d) \in \mathbb{R}_+^d$. Then X is a Brownian sheet on $\mathbb{R}^d$.

Propsition 6.2. B is a Brownian sheet, then:
(1) $B_t = 0$ if t has at least one coordinate equal to 0;
(2) Set $X_{t_1} = (t_2 t_3 \cdots t_d)^{-1/2} B_{(t_1, t_2, \cdots, t_d)}$ for $\forall t_1, t_2, \cdots, t_d \in \mathbb{R}_+$. Then for fixed $t_2, t_3, \cdots, t_d$, $X = \{X_{t_1} : t_1 \geq 0\}$ is a Brownian motion w.r.t. filtration $\{\mathcal{F}_{(t_1, t_2, \cdots, t_d)} : t_1 \geq 0\}$;

Proof:
(1) For $t = (t_1, t_2, \cdots, t_d) \in \mathbb{R}_+^d$, if $t_i = 0$ for some $i \in \{1, 2, \cdots, d\}$, then

$$Var(B_t) = t_1 t_2 \cdots t_d = 0$$

Thus $B_t = 0$ a.s..
(2) For fixed $t_2, t_3, \cdots, t_d \in \mathbb{R}_+$, we have

$$\begin{aligned} Cov(X_{s_1}, X_{t_1}) &= (t_2 t_3 \cdots t_d)^{-1/2} (s_2 s_3 \cdots s_d)^{-1/2} Cov(B_{(s_1, s_2, \cdots, s_d)}, B_{(t_1, t_2, \cdots, t_d)}) \\ &= (t_2 t_3 \cdots t_d)^{-1/2} (s_2 s_3 \cdots s_d)^{-1/2} (s_1 \wedge t_1)(s_2 \wedge t_2) \cdots (s_d \wedge t_d) \\ &= (s_1 \wedge t_1) \end{aligned}$$

for $\forall s_1, t_1 \geq 0$. Thus, X is a Brownian motion.
□

Propsition 6.3. Brownian sheet has a modification which has continuous sample paths.

Theorem 6.1. A Brownian sheet B is a strong martingale w.r.t. the filtration $\{\mathcal{F}_z : z \in \mathbb{R}_+^2\}$ where $\mathcal{F}_{(s,t)} = \sigma(\{B_{(s',t')} : s' \leq s, t' \leq t\} \cup \{A : \mathbb{P}(A) = 0\})$ for $\forall (s, t) \in \mathbb{R}_+^2$.