Wiener - Ito Integrals

MS Project, Academic Year 2022-23

Mrittika Nandi

2021SMCS007

School of Mathematical and Computational Sciences
Indian Association for the Cultivation of Science
Jadavpur, Kolkata

Project Supervisor : Prof. Alok Goswami

25 May, 2023

Contents

INTRODUCTION

For a long time during the first half of the last century, a group of mathematicians, physicists and economists felt a strong need for developing a way of integrating functions (deterministic as well as random) with respect to Brownian motion. This seemed like a fundamental need in order to properly understand various phenomenon in kinetic dynamics, in p.d.e. theory and in mathematics finance. However, since the paths of Brownian motion, except for being continuous, are fairly erratic and, in particular, of unbounded variation on every interval. This meant that none of the existing integration theories would work. It was therefore a major breakthrough, when a complete new integration theory involving Brownian motion was discovered, first by Norbert Wiener (for deterministic integrands) and then by Kyosi Ito (for stochastic integrands) in the 1930s and 40s. This came to be known as Ito Integral (or Wiener-Ito Integral or more generally, Stochastic integral) Thus a completely new integration theory was defined for integrating with respect to Brownian motion, called Ito integral (more generally, Stochastic integral). The aim of this project is to study this theory and then study one or two of the numerous applications of this theory. I start with an attempt first to get an understanding of Brownian motion first before being able to get to the calculus of Brownian motion.

Chapter 1

Preliminaries

1.1 Normal/Gaussian Distribution

1.1.1 Normal Density in One Dimension (centred at zero)

If X is a one-dimensional random variable, then the probability density function of X on $\mathbb{R}$ is given by

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

Consequently,

Expectation of X , $E(X) = \int_{\mathbb{R}} x\, f(x)\, dx \ = 0$

Variance of X , $Var(X) = E(X^2) - (E(X))^2 = \sigma^2$

Thus, X is said to follow a Normal Distribution with mean 0 and variance σ^2 , $X \sim N(0, \sigma^2)$.

1.1.2 Normal Density in Higher Dimension (centred at origin)

If $\mathbf{X}$=$(X_1, X_2, \cdots, X_n)$ is a n-dimensional random variable, then the probability density function of X on $\mathbb{R}^n$ is given by

$$f(\mathbf{x}) = (\mathbf{2}\pi)^{-\frac{\mathbf{n}}{\mathbf{2}}} (\mathbf{det\Sigma})^{-\frac{1}{2}} \mathbf{e}^{-\mathbf{q(x)}}, \ \textbf{where } \mathbf{q(x)} = \frac{\mathbf{1}}{\mathbf{2}} \mathbf{x'\Sigma^{-1}x}$$

Here, Σ is a symmetric positive definite matrix , denoted by, $((\sigma_{ij}))$, say.

Expectation, $E(X_i) = \int_{\mathbb{R}^n} x_i\, f(\mathbf{x})\, \mathbf{dx} \ = \mathbf{0}$, for $1 \le i \le n$

$E(X_i X_j) = \int_{\mathbb{R}^n} x_i\, x_j\, f(\mathbf{x})\, \mathbf{dx} \ = \sigma_{\mathbf{ij}}$, for $1 \le i \le n, 1 \le j \le n$

Variance, $Var(X_i) = E({X_i}^2) - (E(X_i))^2 = \sigma_{ii}$, for $1 \le i \le n$

Covariance, $Cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j) = \sigma_{ij}$, for $1 \le i \le n, 1 \le j \le n$

$\mathbf{X}$ is said to follow n-dimensional Normal Distribution with mean vector $\mathbf{0}$ and dispersion matrix $\mathbf{\Sigma}$, $\mathbf{X \sim N_n(0, \Sigma)}$

Important facts:

1. If $\mathbf{\Sigma} = \mathbf{diag}(\sigma_{\mathbf{1}}{}^{\mathbf{2}}, \sigma_{\mathbf{2}}{}^{\mathbf{2}}, \cdots, \sigma_{\mathbf{n}}{}^{\mathbf{2}})$, $X_i's$ are independent with $X_i \sim N(0, {\sigma_i}^2)$.
2. If $\mathbf{X \sim N_n(0, \Sigma)}$ and $\mathbf{A}$ is non-singular, then $\mathbf{AX \sim N(0, A^T \Sigma A)}$.
3. If $\mathbf{X \sim N_n(0, \Sigma)}$, then all the marginals of X have normal densities.
4. For $\mathbf{X \sim N_n(0, \Sigma)}$, $\exists$ a non-singular $\mathbf{A}$ s.t. $\mathbf{AX}$=$\mathbf{Y}$=$(Y_1, \cdots, Y_n)$ with each $Y_i \sim N(0,1)$ iid .
5. If $\mathbf{X \sim N_n(0, \Sigma)}$ and $\mathbf{Y{=}X{+}m}$ (m $\in \mathbb{R}^n$), then $\mathbf{Y \sim N_n(m, \Sigma)}$,i.e., $\mathbf{X = Y - m \sim N_n(0, \Sigma)}$.

1.2 Degenerate Normal Distribution

In one dimension,
$X \sim N(0, \sigma^2)$ and $Y = aX + b, a \neq 0$, then $Y \sim N(b, a^2)$.
Putting b= μ and $|a| = \sigma > 0$, $Y \sim N(\mu, \sigma^2)$
Here, if σ= 0, the Y is said to follow degenerate normal distribution and $Y = \mu$.

In higher dimension,
$\mathbf{X} \sim \mathbf{N_n}(\mathbf{0}, \mathbf{\Sigma})$, non-degenerate (i.e.,Σ symmetric positive definite) , then for any singular $\mathbf{A}$, $\mathbf{Z} = \mathbf{AY} \sim \mathbf{N}(\mathbf{0}, \mathbf{A^T \Sigma A})$ is degenerate.
Here, degeneracy of $\mathbf{Z}$ means
(i) $\mathbf{Z}$ does **NOT** have a density but
(ii) for any $\alpha = (\alpha_1, \cdots, \alpha_n)$, $\alpha\mathbf{Z} = \alpha_\mathbf{1}\mathbf{Z_1} + \cdots \alpha_\mathbf{n}\mathbf{Z_n}$ is either degenerate at $\mathbf{0}$ or has a one-dimensional $N(0, \sigma^2)$ density.

Chapter 2

Brownian Motion

Definition 2.1. *A Standard Brownian Motion (SBM) is a continuous-time stochastic process* $(B_t, t \geq 0)$ *on* $(\Omega, \mathcal{A}, \mathbf{P})$ *satisfying the properties:*

(a) $B_0(\omega) = 0 \, \forall \, \omega \in \Omega$

(b) (Independent increments) For $0 < t_1 < t_2 < \cdots < t_n$, $B_{t_1}, B_{t_2} - B_{t_1} \cdot \cdots, B_{t_n} - B_{t_{n-1}}$ *are independent random variables.*

(c) (Stationary Gaussian increments) $B_{s+t} - B_s \sim N(0, t)$ *, for any* $s \geq 0$ *and t>0.*

(d) The map $t \to B_t(\omega)$ *is continuous on [0,∞)* $\forall \, \omega \in \Omega$ *.*

For each ω*, the map* $t \to B_t(\omega)$ *on [0,∞) is called the* ω*-trajectory or* ω*-path of the Brownian motion.*

Definition 2.2. *(Gaussian Process) Let* $T \subseteq \mathbb{R}$ *be an interval. A stochastic process* $\{X_t, t \in T\}$ *is said to be a Gaussian if , for every* $n \geq 1$ *and every* $0 < t_1 < t_2 < \cdots < t_n$ *in* T *, the random vector* $(X_{t_1}, X_{t_2}, \cdots, X_{t_n})$ *has a Gaussian distribution (possibly degenerate).*

Remark 2.1. $\mathbf{m(t) = E(X_t), t \in T}$ *is called the* ***mean function*** *and* $\mathbf{C(s,t) = Cov(X_s, X_t), s, t \in T}$ *is called the* ***covariance kernel*** *of the Gaussian Process* $\{X_t, t \in T\}$.
m(·) is a real function on T, C(·,·) is a real symmetric non-negative definite kernel on T.

Definition 2.3. *(2nd definition of SBM) A Standard Brownian Motion is a Gaussian process* $\{B_t, t \geq 0\}$ *with mean function* $m \equiv 0$*, covariance kernel* $C(s,t) = s \wedge t$ *and continuous trajectories (almost surely).*

Remark 2.2. *1. For any* $n \geq 1$ *and any* $0 < t_1 < t_2 < \cdots < t_n$, $(B_0, B_{t_1}, B_{t_2}, \cdots, B_{t_n})$ *has a degenerate normal distribution .*
2. For any $n \geq 1$ *and any* $0 < t_1 < t_2 < \cdots < t_n$, $(B_{t_1}, B_{t_2}, \cdots, B_{t_n}) \sim N(\mathbf{0}, \mathbf{\Sigma})$*, where* $\mathbf{\Sigma} = ((\mathbf{t_i} \wedge \mathbf{t_j}))$

2.1 Properties of SBM

Let $\{B_t, t \geq 0\}$ be a SBM. Then,

1. (Scaling) $\{X_t = a^{-1} B_{a^2 t}, \in [0, \infty)\}$ for any $a \neq 0$, is a SBM.
2. (Markov) For $s \geq 0$, $\{X_t = B_{s+t} - B_s, t \in [0, \infty)\}$ is a SBM , which is independent of $\{B_u, u \leq s\}$.
3. (Time Reversal) $\{X_t = B_{T-t} - B_T, t \in [0, T]\}$ is a SBM on [0,T], for any $T > 0$.
4. (Time Inversion) $\{X_t, t \in [0, \infty)\}$ defined as $X_0 \equiv 0$ and $X_t = t B_{1/t}$ for $t > 0$ is a SBM.

2.2 Martingale Properties of SBM

Definition 2.4. *(Continuous-time version) Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space . Let each $\mathcal{A}_t$ is a subfield of $\mathcal{A}$ for all t and $\mathcal{A}_s \subseteq \mathcal{A}_t$ for $s \leq t$,i.e., $\{\mathcal{A}_t, t \geq 0\}$ is a filtration. Then a family $\{X_t, t \geq 0\}$ of integrable real random variables on $(\Omega, \mathcal{A}, \mathbf{P})$ is said to be a **martingale** with respect to $\{\mathcal{A}_t, t \geq 0\}$ if X_t is $\mathcal{A}_t$-measurable for all $t \geq 0$,i.e., $\{X_t\}$ is adapted to $\{\mathcal{A}_t, t \geq 0\}$ and if*

$$E(X_t/\mathcal{A}_s) = X_s \text{ for all } 0 \leq s \leq t.$$

If $E(X_t/\mathcal{A}_s) \geq X_s$ for all $0 \leq s \leq t$, then $\{X_t, t \geq 0\}$ is called a submartingle.
If $E(X_t/\mathcal{A}_s) \leq X_s$ for all $0 \leq s \leq t$, then $\{X_t, t \geq 0\}$ is called a supermartingle.

Remark 2.3. *For a filtration $\{\mathcal{A}_t, t \in [0, \infty)\}$,*

1. $\mathcal{A}_\infty = \vee_t \mathcal{A}_t$
2. $\mathcal{A}_{t+} = \cap_{s>t} \mathcal{A}_s$ *for each* $t \geq 0$
3. $\mathcal{A}_t \subseteq \mathcal{A}_{t+}$ *for each* $t \geq 0$
4. $\{\mathcal{A}_{t+}, t \geq 0\}$ *is also a filtration.*

Theorem 2.1. *Let $\{X_t\}$ is a martingle w.r.t. $(\mathcal{A}_t)$ and has continuous trajectories. Then,*

1. $(\Phi(X_t))$ *, where Φ is a convex function s.t. $\Phi(X_t) \in L_1 \; \forall \, t$, is a submartingle w.r.t. $\mathcal{A}_t$.*
2. $E(X_t) = E(X_0)$, $\forall \, t$
3. *(Maximal Inequality) For any $t > 0$ any $\lambda > 0$,*

$$P(sup_{0 \leq s \leq t}|X_s| > \lambda) \leq \tfrac{1}{\lambda} E|X_t|$$

4. *(L_p inequality) Suppose $(X_t) \in L_p$ for all t , where $p > 1$. Then, for all $t > 0$,*

$$|| \quad sup_{0 \leq s \leq t}|X_s| \quad ||_p \leq q||X_t||_p \text{ , where } p \text{ and } q \text{ are conjugates.}$$

2.3 Properties of Brownian Paths

Let $\{B_t, t \in [0, \infty\}$ be a SBM on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$. Assuming that all its trajectories are continuous (by discarding a P-null set, if necessary), we look at some properties of Brownian paths.

Path Property 1. For a P-almost every ω, the map $t \to B(t, \omega)$ is not monotone on any interval in $[0, \infty)$.

Proof. Every non-degenerate interval in $[0, \infty)$ contains an interval $[a, b]$ with rational $0 \leq a < b$. Since there are countably many intervals $[a, b]$ with rational $0 \leq a < b$, it is enough to fix a pair of points $0 \leq a < b$ and prove that,

$$\text{for } \mathbf{P}\text{-a.e.}\,\omega\text{,the map } t \mapsto B(t, \omega) \text{ is not monotone on } [a, b] \; \cdots \; (*)$$

Now, from the scaling and markov property , we have $\{X(t) = (b-a)^{-1/2}[B(a+(b–a)t)–B(a)],\ 0 \leq t \leq 1\}$ is a SBM on $[0, 1]$. Clearly, $t \mapsto B(t, \omega)$ is not monotone on $[a, b]$ if and only if $t \mapsto X(t, \omega)$ is not monotone on $[0, 1]$. Thus it is enough to prove $(*)$ just for $[a, b] = [0, 1]$. Since $\{-B(t)\}$ is also a SBM , by scaling, it is enough to prove that,

$$\mathbf{P}(\,\omega : t \mapsto B(t, \omega) \text{ is non-decreasing on } [0, 1]) = 0$$

For each $n \geq 1$, define

$$A_n = \{\,\omega : B(\tfrac{i}{n}, \omega) – B(\tfrac{i-1}{n}, \omega) \geq 0 \text{ for all } i = 1, \cdots, n\}.$$

Since the increments $B(\frac{i}{n}) - B(\frac{i-1}{n})$, $i = 1, 2, \cdots, n$ are i.i.d. $N(0, \frac{1}{n})$ random variables, by definition, we get $\mathbf{P}(A_n) = 2^{-n}$, for each $n \geq 1$.
Clearly $\mathbf{P}(\,\omega : t \mapsto B(t, \omega)$ is non-decreasing on $[0, 1])$ is bounded above by $\mathbf{P}(A_n) = 2^{-n}$, for each $n \geq 1$, hence, $\mathbf{P}(\,\omega : t \mapsto B(t, \omega)$ is non-decreasing on $[0, 1]) = 0$. This complete the proof. $\square$

Path Property 2. For each $t > 0$, denote $m_t(\omega) = min\{B_s(\omega) : 0 \leq s \leq t\}$ and $M_t(\omega) = max\{B_s(\omega) : 0 \leq s \leq t\}$. Then,

$$P(m_t < 0 < M_t, \forall\, t > 0) = 1$$

Proof. By scaling, $\{-B_t\}$ is an SBM, so it is enough to prove,

$$\mathbf{P}(M_t > 0 \text{ for all } t > 0) = 1.$$

Fix any sequence $t_1 > t_2 > \cdots > t_n > \cdots \downarrow 0$ and let $A_n = \{B(t_n) > 0\}$, $n \geq 1$. Since $B(t_n) \sim \mathrm{N}(0, t_n)$, we have $\mathbf{P}(A_n) = \frac{1}{2}$ for all n. By Fatou's Lemma,

$$\mathbf{P}(\limsup_n A_n) = \int \limsup_n \mathbf{I}_{A_n}\, d\mathbf{P} \geq \limsup_n \mathbf{P}(A_n) = \tfrac{1}{2}.$$

Let $X_n = B(t_n) - B(t_{n+1})$. Thus we get a sequence $\{X_n, n \geq 1\}$ of independent random variables such that $B(t_n) = \sum_{k \geq n} X_k$. Now, $\limsup_n A_n = \{B(t_n) > 0 \text{ for infinitely many } n\} = \{B(t_m) > 0 \text{ for infinitely many } m \geq n\}$, for all n . This means that,

$$\limsup_n A_n \in \mathcal{T}_n = \sigma\big(\{X_k, k \geq 1\}\big), \text{ for all } n, \text{ i.e., } \limsup_n A_n \in \mathcal{T} = \bigcap_n \mathcal{T}_n$$

Since $\{X_n\}$ is a sequence of independent random variables, Kolmogorov's Zero-One Law implies that

$$\mathbf{P}(\limsup_n A_n) = \text{either } 0 \text{ or } 1$$

But, we already know that $\mathbf{P}(\limsup_n A_n) \geq \frac{1}{2}$ and so we conclude that $\mathbf{P}(\limsup_n A_n) = 1$.
From the definitions of A_n and $\limsup_n A_n$, it is clear that

$$\limsup_n A_n = \{M_t > 0 \text{ for all } t > 0\}$$

and that completes the proof. □

Path Property 3. For each $t > 0$, denote $\tilde{m_t} = min\{B_s : s \geq t\} and \tilde{M_t} = max\{B_s : s \geq t\}$. Then,

$$P(\tilde{m_t} < 0 < \tilde{M_t}\, \forall\, t > 0) = 1$$

Proof. By the time inversion property, we can deduce information on behaviour of Brownian path near infinity from known information on behaviour near zero, and, vice versa. Thus , path property 2 and the time inversion property proves this property. □

Lemma 2.1. *(Borel-Cantelli lemma) For any sequence $\{A_n\}$ of events in a probability space $(\Omega, \mathcal{A}, \mathbf{P})$, if $\mathbf{P}(A_n) < \infty$, then $\mathbf{P}(limsup_n A_n) = 0$, where $limsup_n A_n = \cap_n \cup_{k \geq n} A_k$.*

Lemma 2.2. *(Borel-Cantelli Lemma 2) If $\{A_n\}$ is a sequence of* **independent** *events such that $\sum_n \mathbf{P}(A_n) = +\infty$, then $\mathbf{P}\big(\limsup_n A_n\big) = 1$.*

Proof. For any $n \geq 1$, $\mathbf{P}(\cap_{k=n}^{n+m} A_k^c) = \prod_{k=n}^{n+m} \mathbf{P}(A_k^c)$, by the independence of the sequence $\{A_n\}$.
Using the inequality $1 - x \leq e^{-x}$, one gets,

$$\mathbf{P}(\cap_{k=n}^{n+m} A_k^c) \leq e^{-\sum\limits_{k=n}^{n+m} \mathbf{P}(A_k)}$$

Let $m \to \infty$, then by the hypothesis $\sum_n \mathbf{P}(A_n) = +\infty$, we have, $\mathbf{P}(\cap_{k \geq n} A_k^c) \leq e^{-\sum\limits_{k\geq n} \mathbf{P}(A_k)} = 0, \forall\, n \geq 1$.
Then, $\mathbf{P}(\limsup_n A_n)^c) = \mathbf{P}(\cup_n \cap_{k \geq n} A_k^c) = 0$, which proves the lemma. □

Path Property 4. For a SBM $\{B_t, t \in [0, \infty)\}$,

$$P(\lim_{n\to\infty} sup B_n = +\infty, \lim_{n\to\infty} inf B_n = -\infty) = 1$$

Proof. It is enough to prove that $\mathbf{P}(\limsup_n B_n = +\infty) = 1$, since the infimum case follows, as $\{-B_t\}$ is also a SBM, by scaling. Fix an integer $m \geq 9$ and, for each $n \geq 1$, consider the events

$$A_n = \{B(m^n) < -3\sqrt{m^n \log n}\}, C_n = \{B(m^{n+1}) - B(m^n) \geq \sqrt{2(m-1)m^n \log n}\}$$

Now, $B(m^n) \sim \mathrm{N}(0, m^n), i.e., \frac{1}{\sqrt{m^n}} B(m^n) \sim \mathrm{N}(0,1)$ and , $\frac{1}{\sqrt{(m-1)m^n}}(B(m^{n+1}) - B(m^n)) \sim \mathrm{N}(0,1)$.
Using these, it follows that, for each $n \geq 1$,

$$\mathbf{P}(A_n) = \mathbf{P}(Z < -3\sqrt{\log n}) \text{ and } \mathbf{P}(C_n) = \mathbf{P}(Z \geq \sqrt{2 \log n}),$$

where Z denotes a $\mathrm{N}(0,1)$ random variable.
By the inequality, $\frac{a}{1+a^2} e^{-a^2/2} < \int_0^\infty e^- x^2/2 dx < \frac{1}{a} e^{-a^2/2}$, we have $\sum_n \mathbf{P}(A_n) < +\infty, \sum_n \mathbf{P}(C_n) = +\infty$
Now, by using Borel-Cantelli Lemma on the sequence $\{A_n\}$ and Borel-Cantelli Lemma 2 on the independent sequence $\{C_n\}$, one gets that, $B(m^n) \geq -3\sqrt{m^n \log n}$ for all but finitely many n, and, $B(m^{n+1}) - B(m^n) \geq \sqrt{2(m-1)m^n \log n}$ for infinitely many n, both with probability 1.
Thus, $B(m^{n+1}) \geq [\sqrt{2(m-1)} - 3]\sqrt{m^n \log n}$ for infinitely many n, with probability 1.

Since we have taken $m \geq 9$, we have,

$$\limsup_n B_n = +\infty, \text{ with probability one}$$

This completes the proof. $\square$

Definition 2.5. *(Zero Set) For a SBM $\{B_t, t \geq 0\}$, the set $Z(\omega) = \{t \in [0, \infty) : B(t, \omega) = 0\}$ is called the zero set.*

Path Property 5. For a SBM $\{B_t, t \geq 0\}$, the set $Z(\omega)$ is, for almost every ω,closed, unbounded set of zero lebesgue measure,which is dense in itself , for almost every ω.

Proof. The trajectories , $t \to B_t(\omega)$, are a.s. continuous, so by definition, $\mathbf{Z}(\omega)$ is a **closed** set for a.e. ω, since $\{0\}$ is closed. Again, the trajectories have a.s. infinitely many crossings between positive and negative values, near $+\infty$ and near 0, so the set $\mathbf{Z}(\omega)$, for a.e. ω, is **unbounded** and has 0 as a limit point. Now, by using Fubini for non-negative functions, we have ,

$$\int_\Omega \lambda(\{t : B(t,\omega) = 0\})\, d\mathbf{P}(\omega) = \int_\Omega \int_0^\infty 1_{\{B(t,\omega)=0\}}\, dt\, d\mathbf{P} = \int_0^\infty \int_\Omega 1_{\{B(t,\omega)=0\}}\, d\mathbf{P}\, dt$$

Again, $\int_0^\infty \int_\Omega 1_{\{B(t,\omega)=0\}}\, d\mathbf{P}\, dt = \int_0^\infty \mathbf{P}(\omega : B(t,\omega) = 0)\, dt = 0$, which implies that $\mathbf{Z}(\omega)$ has **zero Lebesgue measure**, for a.e. ω.

Finally, we prove that, $\mathbf{Z}(\omega)$ in itself for a.e. ω,i.e., no point of the Zero-set $\mathbf{Z}(\omega)$ is an isolated point in $\mathbf{Z}(\omega)$. This can be proved by showing that, for a.e. ω, any point of $\mathbf{Z}(\omega)$, which is isolated in $\mathbf{Z}(\omega)$ from the left, has to be a limit point of $\mathbf{Z}(\omega)$ from the right. We prove this using the idea that if $t_0 \in (0, \infty)$ is a point in $\mathbf{Z}(\omega)$, which is isolated in $\mathbf{Z}(\omega)$ from the left, then there must be a rational number $r \in (0, \infty)$, such that $t_0 = \min\{t \geq r : B(t, \omega) = 0\}$.

By Path Property 4, $\{t \geq r : B(t, \omega) = 0\} \neq \emptyset$, $\mathbf{P}$-a.e. ω, so let us define, for each rational $r \in (0, \infty)$, $\tau_r(\omega) = \min\{t \geq r : B(t, \omega) = 0\}$ and, $A_r = \{\omega : \tau_r(\omega)$ is a right limit point of $\mathbf{Z}(\omega)\}$. Now, $\mathbf{P}(A_r) = 1$ for each rational $r \in (0, \infty)$, will imply $\mathbf{P}(\cap_r A_r) = 1$, which in turn proves the claim.

Thus it is enough to show that

$$\mathbf{P}(A_r) = 1 \text{ for each rational } r \in (0, \infty).$$

Now, by the strong markov property (discussed later), for any random time τ, the process, $\{X_t = B_{\tau+t} - B_\tau,\ t \in [0, \infty)\}$ is a SBM. In this case, $\tau = \tau_r$ and since $B_{\tau_r} \equiv 0$, by definition, we get that the process $\{X_t = B_{\tau_r+t},\ t \in [0, \infty)\}$ is a SBM. But, then by the Path Property 3, we would have

$$\mathbf{P}(\omega : \exists\ t_1 > t_2 > \cdots \downarrow 0 : X_{t_n}(\omega) = 0) = 1$$

Since $X_t = B_{\tau_r+t}$, $t \geq 0$, we have $\mathbf{P}(A_r) = 1$, thus proving $\mathbf{P}(A_r) = 1$ for each rational $r \in (0, \infty)$. $\square$

Path Property 6. If $M_I = max\{B_t : t \in T\}$ for any closed bounded interval $I \subseteq [0, \infty)$, then $P(M_I \neq M_J) = 1$, for disjoint I and J.

Proof. For a SBM $\{B_t,\ t \in [0, \infty)\}$, and for any closed bounded interval $I \subset [0, \infty)$, let

$$M_I = \max\{B_t : t \in I\}.$$

Let I and J be two disjoint closed bounded subintervals of $[0, \infty)$, say, $I = [a, b]$, $J = [c, d]$, where $a < b < c < d$. Let us put $X = \max\{B_{c+t} - B_c : 0 \leq t \leq d - c\}$, $Y = B_c - B_b$ and $Z = M_I - B_b$.

Now, $\{\mathcal{A}_t,\ t \in [0, \infty)\}$ being the natural filtration of the SBM, the Markov property asserts that X is independent of $\mathcal{A}_c$. Now, Y and Z are both measurable with respect to $\mathcal{A}_c$ and, therefore, X is independent of the vector (Y, Z). Again, using Markov property, we can say that Y is independent of $\mathcal{A}_b$. Thus the random variables X, Y, Z are mutually independent. The definition of X implies that $X = M_J - Y - B_b$, so that $M_J = X + Y + B_b$. Thus,

$$M_I = M_J \iff M_I = X + Y + B_b \iff Z - X = Y, i.e., \mathbf{P}(M_I = M_J) = \mathbf{P}(Z - X = Y)$$

Now, using the result , if U and V are independent random variables and V has continuous distribution, then $\mathbf{P}(U = V) = 0$, we have $\mathbf{P}(M_I = M_J) = \mathbf{P}(Z - X = Y) = 0$, since $Y = B_c - B_b \sim \mathrm{N}(0, c - b)$, thus having a continuous distribution and Y is independent of $Z - X$. This completes the proof. $\square$

Definition 2.6. *Let f is a real valued function on a closed bounded interval $[a, b]$, then for a finite partition $\pi : a = t_0 < t_1 < \cdots < t_k = b$ of $[a, b]$, the* ***variation of f over the partition*** *π is defined as*

$$V(f, \pi, [a, b]) = \sum_{i=1}^k |f(t_i) - f(t_{i-1})|.$$

*f is said to be **of bounded / finite variation** on $[a,b]$ if*

$$V(f,[a,b]) = \sup_\pi V(f,\pi,[a,b]) < \infty,$$

where the sup is over all finite partitions π of $[a,b]$.

Path Property 7. For a SBM $\{B_t, t \geq 0\}$, the paths $B(\cdot,\omega)$ are, for P-a.e. ω , of unbounded variation on every interval $\subseteq [0,\infty)$.

Proof. Let $\{B_t,\, t \in [0,\infty)\}$ be a SBM and fix an interval $[a,b] \subset [0,\infty)$. Denote $\delta = b-a > 0$.
Now, consider the finite partition $\pi_n : a = t_0 < t_1 < \cdots < t_{2^n} = b$, where, for $0 \leq i \leq 2^n$, $t_i = a + \frac{i}{2^n}\delta$.
For each n, define

$$Y_n(\omega) = \sum_{i=1}^{2^n} \xi_{i,n}(\omega)$$

where $\xi_{i,n}(\omega) = \big|B(t_i,\omega) - B(t_{i-1},\omega)\big|$, for $i = 1,\cdots,2^n$. Clearly, $Y_n(\omega) = V(B(\cdot,\omega),\pi_n,[a,b])$.
For each i, $B(t_i) - B(t_{i-1}) \sim \mathrm{N}(0,\delta/2^n)$ and so, $\xi_{i,n} \overset{d}{=} \sqrt{\delta/2^n}\,\big|Z\big|$, where $Z \sim \mathrm{N}(0,1)$.
Let $\mathbf{E}(|Z|) = \alpha$, thus we get $\mathbf{E}(Y_n) = 2^n\sqrt{\delta/2^n}\,\alpha = \sqrt{2^n\delta}\,\alpha$. Since $\xi_{i,n}$, $i = 1,\cdots,2^n$ are independent, $\mathrm{Var}(Y_n) = 2^n(\mathrm{Var}(\sqrt{\delta/2^n}\,|Z|\,)) = \delta\,\mathrm{Var}(|Z|) = \delta(1-\alpha^2)$. Thus we have, $\mathbf{E}(Y_n) = \sqrt{2^n\delta}\,\alpha \uparrow \infty$, as $n \to \infty$, while $\mathrm{Var}(Y_n) = \delta(1-\alpha^2) > 0$, for all n. Now, for any constant $C > 0$, we have

$$\mathbf{P}\big(Y_n \geq C\big) \geq \mathbf{P}\big(|Y_n - \mathbf{E}(Y_n)| \leq \mathbf{E}(Y_n) - C\big) \geq 1 - \tfrac{\delta(1-\alpha^2)}{(\mathbf{E}(Y_n)-C)^2}, \text{ by Chebyshev's inequality} \cdots \ (*)$$

This implies that $\mathbf{P}(Y_n \geq C) \to 1$, as $n \to \infty$. Since $Y_n(\omega) = V(B(\cdot,\omega),\pi_n,[a,b])$ is increasing in n, we may conclude from (*) that $\mathbf{P}(\lim_n Y_n \geq C) = 1$. Since (*) holds for every $C > 0$, we conclude that $Y_n(\omega) \uparrow \infty$, for $\mathbf{P}$-a.e. ω. ,which implies,

$$V(B(\cdot,\omega),[a,b]) = +\infty, \text{ for } \mathbf{P}\text{-a.e. } \omega\ .$$

Since this holds for every $[a,b] \subset [0,\infty)$ with rational end points $a < b$, the proof is done. □

Definition 2.7. *Let f be a real valued function on a closed bounded interval $[a,b]$. For any finite partition $\pi : a = t_0 < t_1 < \cdots < t_k = b$, the sum*

$$\sum_i (f(t_i) - f(t_{i-1}))^2$$

*is called the **Quadratic Variation** of f w.r.t. the partition π and is denoted $Q(f,\pi,[a,b])$. For a sequence $\{\pi_n\}$ of finite partitions with $\|\pi_n\| \downarrow 0$, if the limit $\lim_n Q(f,\pi_n,[a,b])$ exists, then the limit is called the **Quadratic variation of** f **along** $\{\pi_n\}$, and is denoted $Q(f,\{\pi_n\},[a,b])$.*

Path Property 8. For a SBM $\{B_t, t \geq 0\}$, let $I \subseteq [0,\infty)$. Then, for any sequence $\{\pi_n\}$ of finite partitions of I with $||\pi_n|| \downarrow 0$,

$$Q(B(\cdot,\omega),\pi_n,I) \xrightarrow{L_2} \lambda(I) \text{ as } n \to \infty$$

Proof. Let $\{B_t,\, t \in [0,\infty)\}$ be a SBM and fix a closed bounded interval $[a,b] \subset [0,\infty)$. For any finite partition $\pi : a = t_0 < t_1 < \cdots < t_k = b$, we have

$$Q(B(\cdot,\omega),\pi,[a.b]) = Y_1^2 + \cdot + Y_k^2,$$

where Y_i, $1 \leq i \leq k$ are independent and $Y_i \sim \mathrm{N}(0, t_i - t_{i-1})$. Hence it follows that, $\mathbf{E}(Y_i^2) = t_i - t_{i-1}$ and $\mathrm{Var}(Y_i^2) = \mathbf{E}(Y_i^4) - (\mathbf{E}(Y_i^2))^2 = 2(t_i - t_{i-1})^2$. Thus $\mathbf{E}(Q(B(\cdot,\omega),\pi,[a.b])) = b-a$ and $\mathbf{E}\,|Q(B(\cdot,\omega),\pi,[a.b]) - (b-a)|^2 = \mathrm{Var}(Q(B(\cdot,\omega),\pi,[a.b]) = 2\sum_i (t_i - t_{i-1})^2 \leq 2(b-a)\|\pi\|$.
Thus, if $\{\pi_n\}$ is any sequence of partitions with $\|\pi_n\| \downarrow 0$, then

$$\mathbf{E}\,|Q(B(\cdot,\omega),\pi_n,[a.b]) - (b-a)|^2 \to 0,$$

$$\text{that is,}\quad Q(B(\cdot,\omega),\pi_n,[a.b]) \xrightarrow{L_2} (b-a), \text{ as } n \to \infty.$$

This completes the proof. □

Definition 2.8. *(Dini Derivatives) For $f : [0,\infty) \to \mathbb{R}$, the **right upper** and **lower Dini Derivatives** of f at $t \in [0,\infty)$ are defined as*

$$D^* f(t+) = \limsup_{h\downarrow 0} \tfrac{f(t+h)-f(t)}{h} = \lim_{h\downarrow 0} \sup_{0<s<h} \big\{\tfrac{f(t+s)-f(t)}{s}\big\}$$

$$D_* f(t+) = \liminf_{h\downarrow 0} \tfrac{f(t+h)-f(t)}{h} = \lim_{h\downarrow 0} \inf_{0<s<h} \big\{\tfrac{f(t+s)-f(t)}{s}\big\}$$

*Similarly, the **left Dini Derivatives** at $t \in (0, \infty)$ are*

$$D^* f(t-) = \limsup_{h \uparrow 0} \frac{f(t+h)-f(t)}{h} = \lim_{h \uparrow 0} \sup_{h<s<0} \left\{ \frac{f(t+s)-f(t)}{s} \right\}$$
$$D_* f(t-) = \liminf_{h \uparrow 0} \frac{f(t+h)-f(t)}{h} = \lim_{h \uparrow 0} \inf_{h<s<0} \left\{ \frac{f(t+s)-f(t)}{s} \right\}$$

Remark 2.4. *Clearly, $-\infty \le D_* f(t+) \le D^* f(t+) \le +\infty$, for $t \in [0, \infty)$ $-\infty \le D_* f(t-) \le D^* f(t-) \le +\infty$, for $t \in (0, \infty)$.*
Further, right-derivative $Df(t+)$ at $t \in [0, \infty)$ exists if and only if $-\infty < D_* f(t+) = D^* f(t+) < +\infty$ and, in that case,

$$Df(t+) = D_* f(t+) = D^* f(t+)$$

Similarly, left-derivative $Df(t-)$ at $t \in (0, \infty)$ exists if and only if $-\infty < D_* f(t-) = D^* f(t-) < +\infty$ and, in that case,

$$Df(t-) = D_* f(t-) = D^* f(t-)$$

Finally, f is differentiable at $t \in (0, \infty)$ if and only if all four Dini Derivatives are equal and finite and, in that case, the common value gives the derivative $Df(t)$.

Path Property 9. (Everywhere Non-differentiability) Let $\{B_t, t \ge 0\}$ be a SBM on $(\Omega, \mathcal{A}, \mathbf{P})$. Then $\exists$ a $\mathbf{P}$ -null set $N \in \mathcal{A}$ s.t. $\forall\, \omega \notin N$, either $D^* B(t+, \omega) = +\infty$ or $D_* B(t+, \omega) = -\infty$ for every $t \in [0, \infty)$ and also either $D^* B(t-, \omega) = +\infty$ or $D_* B(t-, \omega) = -\infty$ for every $t \in (0, \infty)$.

Proof. Since SBM has time reversal property, it is enough to prove the statement on the right Dini Derivatives only. Now, the Markov property would imply that it is enough to prove the statement on the right Dini Derivatives only for $t \in [0, 1)$. We are going to show that there is a $\mathbf{P}$-null set N such that, if $\omega \notin N$, then for each $t \in [0, 1)$,

$$\text{either } D^* B(t+, \omega) = +\infty \text{ or } D_* B(t+, \omega) = -\infty.$$

Let $A = \{\omega : -\infty < D_* B(t+, \omega) \le D^* B(t+, \omega) < +\infty, \text{ for some } t \in [0, 1)\}$. We cannot show that A itself is a $\mathbf{P}$-null set, because A may not even belong $\mathcal{A}$. What we instead do is that we get a set $N \in \mathcal{A}$ with $A \subset N$ and show that $\mathbf{P}(N) = 0$. Towards this, note first that $\omega \in A$ implies that there exists a $t \in [0, 1)$ such that $\frac{B(t+h,\omega)-B(t,\omega)}{h}$ remains bounded for all sufficiently small $h > 0$.
In other words, $\omega \in A$ implies that there exist $t \in [0, 1)$ and integers $j \ge 1, k \ge 1$
such that $\left| \frac{B(t+h,\omega)-B(t,\omega)}{h} \right| \le j$ for all $0 < h \le \frac{1}{k}$, or equivalently,

$$\big| B(t+h, \omega) - B(t, \omega) \big| \le jh \ \text{ for } 0 \le h \le \tfrac{1}{k} \ \cdots \ (*)$$

Since the $t \in [0, 1)$ and the integers $j \ge 1$, $k \ge 1$ all depend on $\omega \in A$, for any $\omega \in A$, there exist such t, j and k, for which $(*)$ holds.Without loss of generality, we may assume that the integer k above also satisfies $t + \frac{1}{k} < 1$ (by going for a larger k, if necessary). Now, for any integer $n > 4k$, we will have an i (again depending on ω) with $1 \le i \le n$, such that $t \in \left[\frac{i-1}{n}, \frac{i}{n}\right)$.
Using $t + \frac{1}{k} < 1$ and $n > 4k$, we get $\frac{i+3}{n} = \frac{i-1}{n} + \frac{4}{n} < t + \frac{1}{k} < 1$, implying that $i \le n - 4$,i.e., for $m = 0, 1, 2, 3$, the inequality $(*)$ would hold with $t + h = \frac{i+m}{n}$, that is,

$$\left| B\big(\tfrac{i+m}{n}, \omega\big) - B\big(t, \omega\big) \right| \le \left(\tfrac{i+m}{n} - t\right) j < \tfrac{m+1}{n} j$$

That is $\left| B\big(\frac{i+1}{n}, \omega\big) - B\big(\frac{i}{n}, \omega\big) \right| \le \left| B\big(\frac{i+1}{n}, \omega\big) - B\big(t, \omega\big) \right| + \left| B\big(\frac{i}{n}, \omega\big) - B\big(t, \omega\big) \right| \le \frac{2}{n} j + \frac{1}{n} j = \frac{3}{n} j$. Similarly, $\left| B\big(\frac{i+2}{n}, \omega\big) - B\big(\frac{i+1}{n}, \omega\big) \right| \le \frac{3}{n} j + \frac{2}{n} j = \frac{5}{n} j$, $\left| B\big(\frac{i+3}{n}, \omega\big) - B\big(\frac{i+2}{n}, \omega\big) \right| \le \frac{4}{n} j + \frac{3}{n} j = \frac{7}{n} j$
For positive integers i, j, n, let us define

$$A_{i,j,n} = \left\{ \omega : \begin{array}{rcl} \left| B\big(\frac{i+1}{n}, \omega\big) - B\big(\frac{i}{n}, \omega\big) \right| & \le & \frac{3}{n} j \\ \left| B\big(\frac{i+2}{n}, \omega\big) - B\big(\frac{i+1}{n}, \omega\big) \right| & \le & \frac{5}{n} j \\ \left| B\big(\frac{i+3}{n}, \omega\big) - B\big(\frac{i+2}{n}, \omega\big) \right| & \le & \frac{7}{n} j \end{array} \right\}$$

Clearly, $A_{i,j,n} \in \mathcal{A}$, for every i, j, n and we have shown that, for every $\omega \in A$, there exist integers $j \ge 1$, $k \ge 1$, such that, for all $n > 4k$,

$$\omega \in A_{i,j,n} \text{ for some } 1 \le i \le n - 4$$

In other words, we have just proved that

$$A \subset \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{n>4k} \bigcup_{i=1}^{n-4} A_{i,j,n},$$

where the set on the right clearly belongs to $\mathcal{A}$. Denoting the set on the right as N, we now show that N is a $\mathbf{P}$-null set and that will complete the proof.
Since the union over j and the union over k are both countable unions, it is enough to show that

$$\mathbf{P}\big(\bigcap_{n>4k}\bigcup_{i=1}^{n-4} A_{i,j,n}\big) = 0, \quad \text{for each } j \text{ and } k.$$

From the properties (b) and (c) in the definition of SBM, the three increments appearing in the definition of the event $A_{i,j,n}$ are each $\sim \mathrm{N}(0, \frac{1}{n}) \overset{d}{=} \frac{1}{\sqrt{n}}\, Z$, where Z has a $\mathrm{N}(0,1)$ distrbution, and they are independent. From this, it immediately follows that, for each i, j, n ,

$$\mathbf{P}\big(A_{i,j,n}\big) = \mathbf{P}\big(|Z| \le \tfrac{3}{\sqrt{n}} j\big) \times \mathbf{P}\big(|Z| \le \tfrac{5}{\sqrt{n}} j\big) \times \mathbf{P}\big(|Z| \le \tfrac{7}{\sqrt{n}} j\big)$$

Using the crude bound, $\mathbf{P}(|Z| \le \alpha) = 2\,\frac{1}{\sqrt{2\pi}}\int_0^\alpha e^{-x^2/2}\,dx \le \alpha$, one gets,

$$\mathbf{P}\big(A_{i,j,n}\big) \le 105\, j^3/n^{3/2}, \quad \text{for each } i, j, n$$

implying that

$$\mathbf{P}\big(\bigcup_{i=1}^{n-4} A_{i,j,n}\big) \le (n-4)\mathbf{P}\big(A_{i,j,n}\big) \le 105\, j^3/\sqrt{n}$$

This finally gives, $\mathbf{P}\big(\bigcap_{n>4k}\bigcup_{i=1}^{n-4} A_{i,j,n}\big) \le \inf_{n>4k}\big(\frac{105\, j^3}{\sqrt{n}}\big) = 0$. This completes the proof. □

Definition 2.9. *(Hölder Continuity) A real function f on an interval I is said to be* ***locally Hölder continuous*** *of exponent $\gamma > 0$ at $t \in I$ if there exists $\delta > 0$ and $c < \infty$,s.t.,*

$$|f(s) - f(t)| \le c|s-t|^\gamma \;\forall\; s \in [t-\delta, t+\delta] \cap I$$

Path Property 10. (Non-Hölder Continuity) Almost every trajectory of a SBM is not locally Hölder continuous anywhere of any exponent $\gamma > 1/2$.

Proof. This can be proved using similar ideas as in the proof of nowhere differentiability of a.e. trajectory of SBM. □

If a function is locally Hölder continuous of an exponent γ, then it is also locally Hölder continuous of exponent γ' for any $\gamma' < \gamma$.

These properties give testimony of the erratic nature of a typical Brownian motion.

2.4 Markov property of Brownian Motion

If $\{B_t, t \ge 0\}$ is a SBM, then for any $s \ge 0$, the conditional distribution of the process $\{B_{s+t}, t \ge 0\}$, given $\{B_u, u \le s\}$, is that of a Brownian motion starting at B_s.
Here, $\{B_{s+t}, t \ge 0\}$ represented the motion after time s, $s \ge 0$,i.e., the $future$ motion beyond time s,i.e., a Brownian motion *starting* at B_s and $\{B_u, u \le s\}$ represents *past* history upto and including time s.

Definition 2.10. *(Regular Conditional Distribution) Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space and let $\mathcal{G} \subseteq \mathcal{A}$ be a sub σ-field. Then, for a measurable $X : (\Omega, \mathcal{A}, \mathbf{P}) \to (E, \mathcal{E})$, a Regular Conditional Distribution of X, given $\mathcal{G}$, means a function $Q : \Omega \times E \to [0,1]$, satisfying,*

1. *for each $\omega \in \Omega, Q(\omega, \cdot)$ is a probability measure on (E).*
2. *for each $B \in \mathcal{E}, Q(\cdot, B)$ is $\mathcal{G}$ - measurable.*
3. *for each $B \in \mathcal{E}$ and $G \in \mathcal{G}, P(G \cap X^{-1}(B)) = \int_G Q(\omega, B)dP(\omega)$*

Regular Conditional Distributions need not always exist !!

2.4.1 The Space $C[0,\infty)$ and Law of Brownian Motion

Consider the space $C[0,\infty)$ consisting of all real-valued continuous functions on $[0,\infty)$ and let $C[0,\infty) = C$. Further we denote by $\mathcal{C}$, the Borel σ-field on C containing all the open sets.
Let $\{B(t), t \in [0,\infty\}$ be a SBM on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ and assume that all trajectories are continuous $\mathbf{P}$-a.s..
Since $B(t)$ is a real random variable for each t, the map,

$$B(\cdot) : (\Omega, \mathcal{A}, \mathbf{P}) \to (C, \mathcal{C}),$$

defined as $\omega \to B(\cdot, \omega)$ is a measurable map.

Definition 2.11. *(Wiener Measure) Denote $\mathbf{P}_0$ to be the probability on $(C,\mathcal{C})$ induced by the measurable map $B(\cdot) : (\Omega, \mathcal{A}), \mathbf{P}) \to (C, \mathcal{C})$, that is,*

$$\mathbf{P}_0(E) = P\{\omega : B(\cdot,\omega) \in E\}, for E \in \mathcal{C}$$

The probability P_0 is commonly known as the ***Wiener measure*** *on $(C,\mathcal{C})$*

The SBM $\{B(t), t \in [0,\infty\}$ defined on $(\Omega, \mathcal{A}, \mathbf{P})$ can be viewed as a C-valued random variable of the SBM or what is often referred as the **law of SBM**. Clearly, the coordinate process on $(C,\mathcal{C})$ under the probability $\mathbf{P}_0$, will be a SBM.

The law of a Brownian Motion starting at x : For $x \in \mathbb{R}$, if $\{B^x(t), t \in [0,\infty)\}$ defined on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ is a Brownian motion starting from x and having all trajectories continuous , then, $B^x(\cdot) : \omega \to B^x(\cdot,\omega)$ is a C-valued random variable. The probability $\mathbf{P}_x$ on $\mathcal{C}$ induced by the map $B^x(\cdot)$ describes the law of a Brownian motion starting at x. and it is clear that the coordinate process on $(C,\mathcal{C})$, under $\mathbf{P}_x$ will be a Brownian motion starting at x.

Note:

1. For a Brownian motion starting at x, for any $x \in \mathbb{R}$, $B_0 \equiv x$.
2. $\{B^x(t), t \in [0,\infty)\}$ is a Brownian Motion starting at x if and only if $B^x(\cdot) = x + B(\cdot)$, where $\{B(t), t \in [0,\infty)\}$ is a SBM.
3. $\mathbf{P}_x(E) = \mathbf{P}_0(\{f \in C : x + f(\cdot) \in E\}), E \in \mathcal{C}, x \in \mathbb{R}$
4. For any $E \in \mathcal{C}$, the map $x \to \mathbf{P}_x(E)$ is a measurable map on $(\mathbb{R}, \mathcal{B})$

Lemma 2.3. *Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space and let $\mathcal{G} \subset \mathcal{A}$ be a sub-σ-field. Consider a C-valued random variable Y on $(\Omega, \mathcal{A}, \mathbf{P})$ with $\mathbf{P}Y^{-1} = \mu_0$, which is independent of $\mathcal{G}$. If X is a $\mathcal{G}$-measurable real random variable on $(\Omega, \mathcal{A}, \mathbf{P})$ and Z is the C-valued random variable defined as $Z = X + Y$ (or , $Z(\cdot,\omega) = X(\omega) + Y(\cdot,\omega)$, $\omega \in \Omega$, to be precise) then a regular conditional distribution of Z, given $\mathcal{G}$, is given by*

$$Q(\omega, E) = \mu_0\big(\{f \in C : X(\omega) + f(\cdot) \in E\}\big),\ \omega \in \Omega, E \in \mathcal{C}$$

Proof. That Z is a C-valued random variable is clear, since the map $\omega \mapsto Z(t,\omega) = X(\omega) + Y(t,\omega)$, for each $t \in [0,\infty)$, is a real random variable on $(\Omega, \mathcal{A}, \mathbf{P})$. The hypothesis that Y is independent of $\mathcal{G}$ means that

$$\mathbf{P}\big(G \bigcap Y^{-1}(E)\big) = \mathbf{P}(G) \cdot \mathbf{P}\big(Y^{-1}(E)\big) = \mathbf{P}(G) \cdot \mu_0(E), \text{ for all } G \in \mathcal{G},\ E \in \mathcal{C}$$

It is easy to see that, for each $\omega \in \Omega$, $Q(\omega,\cdot)$, as defined in the statement, is a probability measure on $(C,\mathcal{C})$. Now, the map $x \mapsto \mu_0\big(\{f \in C : x + f \in E\}\big)$ is a borel measurable map on $\mathbb{R}$, for each $E \in \mathcal{C}$. Since X is a $\mathcal{G}$-measurable real random variable, it follows that, for each $E \in \mathcal{C}$, the map $\omega \mapsto Q(\omega, E)$ is $\mathcal{G}$-measurable. So, to complete the proof, it is enough to show that, for all $G \in \mathcal{G}$ and all $E \in \mathcal{C}$,

$$\mathbf{P}\big(G \bigcap \{\omega : Z(\cdot\,,\omega) \in E\}\big) = \textstyle\int_G Q(\omega, E)\, d\mathbf{P}(\omega)$$

This is proved through the following steps.

Step 1 : The map $\varphi : (\Omega, \mathcal{A}, \mathbf{P}) \longrightarrow (\Omega \times C, \mathcal{G} \otimes \mathcal{C})$, given by $\omega \mapsto (\omega, Y(\cdot\,,\omega))$ is a measurable map and

$$\mathbf{P}\varphi^{-1} = \mathbf{P} \otimes \mu_0.$$

For any $G \in \mathcal{G}$, $E \in \mathcal{C}$, we have $\varphi^{-1}(G \times E) = G \bigcap \{\omega : Y(\cdot,\omega) \in E\} \in \mathcal{A}$, and, $\mathbf{P}\varphi^{-1}(G \times E) = \mathbf{P}\big(G \bigcap \{\omega : Y(\cdot,\omega) \in E\}\big) = \mathbf{P}(G) \cdot \mathbf{P}\big(\{\omega : \{\omega : Y(\cdot,\omega) \in E\}\big) = \mathbf{P}(G) \cdot \mu_0(E)$, where the last two equalities are consequences of the hypotheses that Y is independent of $\mathcal{G}$ and $\mathbf{P}Y^{-1} = \mu_0$.
Since the class $\big\{G \times E : G \in \mathcal{G},\, E \in \mathcal{C}\big\}$ of measurable rectangles form a semi-field that generates the product σ-field $\mathcal{G} \otimes \mathcal{C}$, the above proves Step 1.

Step 2 : Fix $E \in \mathcal{C}$. If $E^* = \{(\omega, f) \in \Omega \times C : X(\omega) + f(\cdot) \in E\}$, then $E^* \in \mathcal{G} \otimes \mathcal{C}$. For this, it si enough to show that the map $(\omega, f) \mapsto X(\omega) + f(\cdot)$ is a measurable map on $(\Omega \times C, \mathcal{G} \otimes \mathcal{C})$ into $(C, \mathcal{C})$,that is , to show that, for each $t \in [0,\infty)$, the map

$$(\omega, f) \mapsto X(\omega) + f(t) \text{ on } (\Omega \times C, \mathcal{G} \otimes \mathcal{C}) \text{ into } (\mathbb{R}, \mathcal{B}) \text{ is a measurable map.}$$

Now, X is a real-valued $\mathcal{G}$-measurable map on Ω and $f \mapsto f(t)$, for each t, is a real-valued $\mathcal{C}$-measurable map. It easily follows that, for each t, $(\omega, f) \mapsto \big(X(\omega), f(t)\big)$ is a measurable map on $(\Omega \times C, \mathcal{G} \otimes \mathcal{C})$ into $(\mathbb{R}^2, \mathcal{B}^2)$, making $(\omega, f) \mapsto X(\omega) + f(t)$ a $\mathcal{G} \otimes \mathcal{C}$-measurable map. Thus the assertion of Step 2 is proved.

Final Step : Fix $G \in \mathcal{G}$ and let $H = \{(\omega, f) \in \Omega \times C : \omega \in G, X(\omega) + f(\cdot) \in E\}$.
Clearly, $H = (G \times C) \cap E^* \in \mathcal{G} \otimes \mathcal{C}$ and, for $\omega \in \Omega$,

$$H_\omega = \begin{cases} \emptyset & \text{if } \omega \notin G \\ E^*_\omega = \{f \in C : X(\omega) + f(\cdot) \in E\} & \text{if } \omega \in G \end{cases},$$

and so $\mu_0(H_\omega) = \mathbf{I}_G\, \mu_0\big(\{f : X(\omega) + f(\cdot) \in E\}\big) = \mathbf{I}_G\, Q(\omega, E)$. Now, from the map φ defined in Step 1, we get $G \bigcap \{\omega : Z(\cdot,\omega) \in E\} = \{\omega : \omega \in G, X(\omega) + Y(\cdot,\omega) \in E\} = \{\omega : (\omega, Y(\omega)) \in H\} = \{\omega : \varphi(\omega) \in H\}$, which implies, that $\mathbf{P}\big(G \bigcap \{\omega : Z(\cdot,\omega) \in E\}\big) = \mathbf{P}\varphi^{-1}(H)$. Hence, by $\mathbf{P}\varphi^{-1} = \mathbf{P} \otimes \mu_0$, as proved in Step 1 and $\mu_0(H_\omega) = \mathbf{I}_G\, \mu_0\big(\{f : X(\omega) + f(\cdot) \in E\}\big) = \mathbf{I}_G\, Q(\omega, E)$, we finally get

$$\mathbf{P}\big(G \bigcap \{\omega : Z(\cdot,\omega) \in E\}\big) = \mathbf{P} \otimes \mu_0(H) = \int_G Q(\omega, E)\, d\mathbf{P}(\omega)$$

This completes the proof. □

Theorem 2.2. *(Markov property of Brownian motion) Let $\{B^x(t), t \in [0,\infty)\}$ defined on $(\Omega, \mathcal{A}, \mathbf{P})$ be a Brownian motion starting from $x \in \mathbb{R}$. Then, for any $s \geq 0$,a regular conditional distribution of $\{B^x(s+t), t \in [0,\infty)\}$, given $\{B^x(u), u \leq s\}$, is given by the kernel*

$$Q(\omega, E) = \mathbf{P}_{B^x_s(\omega)}(E), \omega \in \Omega, E \in \mathcal{C}$$

Here, $\mathbf{P}_{B^x_s(\omega)}(E)$ is the law of a Brownian motion starting at $B^x_s(\omega)$.

Proof. Let us define , $\mathcal{A}_s = \sigma\big(\{B^x(u),\, u \leq s\}\big)$. Since, for each $\omega \in \Omega$, $Q(\omega, E) = \mathbf{P}_{B^x_s(\omega)}(E)$, $E \in \mathcal{C}$ is a probability on $\mathcal{C}$, the property (1) of the definition 2.10 of regular condtional distribution holds. Next, we know that $x \mapsto \mathbf{P}_x(E)$ is, for each $E \in \mathcal{C}$, a measurable map on $(\mathbb{R}, \mathcal{B})$ and the map $\omega \mapsto X_s(\omega)$ into $(\mathbb{R}, \mathcal{B})$ is $\mathcal{A}_s$-measurable.
By composition of these two maps, we get that, for every $E \in \mathcal{C}$, the map

$$\omega \mapsto Q(\omega, E) = \mathbf{P}_{B^x_s(\omega)}(E) \text{ is measurable with respect to } \mathcal{A}_s$$

Thus property (2) of the definition 2.10 is satisfied too. So, to complete the proof of the theorem, we only need to proof property (3), that is,

$$\mathbf{P}\big(G \bigcap \{\omega : B^x(s+\cdot,\omega) \in E\}\big) = \int_G Q(\omega, E)\, d\mathbf{P}(\omega) \;\; \forall\, G \in \mathcal{A}_s, E \in \mathcal{C} \;\cdots\; (*)$$

where $Q(\omega, E) = \mathbf{P}_{B^x_s(\omega)}(E)$, $\omega \in \Omega$, $E \in \mathcal{C}$.
Let $\mathcal{G} = \mathcal{A}_s$, $Y(\cdot,\omega) = B^x(s+\cdot,\omega) - B^x(s,\omega)$ and $X = B^x_s$, we have, from the Markov property of a Brownian motion starting at x, (i) Y is a C-valued random variable, independent of $\mathcal{G}$ and with law $\mathbf{P}\, Y^{-1} = \mathbf{P}_0$, and , (ii) X is a $\mathcal{G}$-measurable real random variable.
Now, $B^x(s+\cdot,\omega) = X(\omega) + Y(\cdot,\omega)$ and we know, for $y \in \mathbb{R}$, $\mathbf{P}_y(E) = \mathbf{P}_0(\{f : y + f(\cdot) \in E\})$, $E \in \mathcal{C}$, thus (*) follows immediately as a consequence of the lemma 2.3. This completes the proof. □

The idea of the theorem is that, for the $\{B^x_t, t \in [0,\infty)\}$, given the history upto and including time $s \geq 0$, the future motion after time s, i.e., $\{B^x(s+t), t \geq 0\}$ is distributed conditionally like a Brownian motion starting at the present state B^x_s thus depicting the general idea of Markov property.

2.4.2 Strong Markov Property of Brownian Motion

This property asserts that the Markov property of Brownian Motion holds not just for fixed times $s \geq 0$, but also for a certain class of **random** times τ.
If $\{B_t^x, t \in [0,\infty)\}$ is a Brownian motion starting at x, then for certain non-negative , finite random times τ, the conditional evolution of the post-τ process $\{B^x(\tau + \cdot)\}$, given the "history of $\{B^x(\cdot)\}$ upto and including time τ", is the same as that of a Brownian motion starting at the "present state B_τ^x".

Definition 2.12. *(Stopping time) $\tau : \Omega \to [0,\infty]$ is called a* ***stopping time*** *w.r.t. $\{\mathcal{A}_t, t \in [0,\infty)\}$ if $\{\tau \leq t\} \in \mathcal{A}_t \,\forall\, t > 0$.*

Definition 2.13. *(Optional time) $\tau : \Omega \to [0,\infty]$ is called an* ***optional time*** *w.r.t. $\{\mathcal{A}_t, t \in [0,\infty)\}$ if $\{\tau < t\} \in \mathcal{A}_t \,\forall\, t > 0$.*

Definition 2.14. *(Hitting time) If $\{B_t, t \geq 0\}$ is a SBM on $(\Omega, \mathcal{A}, \mathbf{P})$, $\{\mathcal{A}_t\}$ ia a natural filtration . Then $\tau_A = inf\{t \geq 0 : B_t \in A\}$ is the* ***hitting times*** *of subsets $A \subseteq \mathbb{R}$ by the SBM.*

Theorem 2.3. *(Strong Markov Property of Brownian Motion) Let $\{B^x(t), t \in [0,\infty)\}$ defined on $(\Omega, \mathcal{A}, \mathbf{P})$ be a Brownian motion starting from $x \in \mathbb{R}$ and let $\{\mathcal{A}_t, t \in [0,\infty)\}$ be its natural filtration. Then, for any finite optional time τ with respect to $\{\mathcal{A}_t\}$, a regular conditional distribution of $\{B_{\tau+t}^x, t \in [0,\infty)\}$, given $\mathcal{A}_{\tau+}$, is given by the kernel,*

$$Q(\omega, E) = \mathbf{P}_{B_\tau^x(\omega)}(E), \omega \in \Omega, E \in \mathcal{C}$$

Here, $\{B_{\tau+t}^x(\cdot) = B^x(\tau(\cdot)+t, \cdot), t \in [0,\infty)\}$ represents the post-τ process, while $B_\tau^x(\omega) = B^x(\tau(\omega), \omega), \omega \in \Omega$ represents the state at the the time τ.

Let us recall that, for any $s \geq 0$, the constant random variable $\tau \equiv s$ is an optional time and, in this case, $\mathcal{A}_{\tau+} = \mathcal{A}_{s+}$.

Now, using the Strong Markov Property for $\tau \equiv s$ gives us the following slightly improved version of the Markov property.

Theorem 2.4. *(Markov Property(improved version)) Let $\{B^x(t), t \in [0,\infty)\}$ defined on $(\Omega, \mathcal{A}, \mathbf{P})$ be a Brownian motion starting from $x \in \mathbb{R}$ and let $\{\mathcal{A}_t, t \in [0,\infty)\}$ be its natural filtration. Then, for any $s \geq 0$, a regular conditional distribution of $\{B^x(s+t), t \in [0,\infty)\}$, given $\mathcal{A}_{s+}$, is given by the kernel,*

$$Q(\omega, E) = \mathbf{P}_{B_s^x(\omega)}(E), \omega \in \Omega, E \in \mathcal{C}$$

2.5 A Canonical formulation of Brownian Motion as Markov Process

Such canonical representation will also provide a convenient setup for discussing the general theory of Markov processes. In this formulation, the space of trajectories ,that is, the path space , is taken as the underlying space and the process is always the coordinate process.
That is, we take $\Omega = C$ and consider the coordinate process $\{X_t, t \geq 0\}$ on Ω defined as,

$$X_t(\omega) = \omega(t), \omega \in \Omega, t \geq 0.$$

The σ-field $\mathcal{A}$ on Ω is taken to be the one generated by $\{X_t, t \geq 0\}$ and, as we know, $\mathcal{A}$ is just the σ-field $\mathcal{C}$.
On $(\Omega, \mathcal{A}) = (C, \mathcal{C})$, we consider the natural filtration $\{\mathcal{A}_t, t \geq 0\}$ associated to the coordinate process $\{X_t\}$. Thus, we have a family of probabilities $(\mathbf{P}_x, x \in \mathbb{R})$ on $(\Omega, \mathcal{A}) = (C, \mathcal{C})$, with $\mathbf{P}_x$ representing the law of the process $\{B_t^x, t \geq 0\}$.However,what this also means is that, under the probability $\mathbf{P}_x$ on $(\Omega, \mathcal{A}) = (C, \mathcal{C})$, the process $\{X_t, t \geq 0\}$ is itself a Brownian motion starting from x.
The main issue here is that instead of having different processes $\{B_t^x, t \geq 0\}$, one for each $x \in \mathbb{R}$, we have the same process, namely the coordinate process $\{X_t, t \geq 0\}$ and it is the different probabilities $\mathbf{P}_x, x \in \mathbb{R}$ that determine the law of the process.
Hence, the markov property can be differently stated as, For each $x \in \mathbb{R}$ and $s \geq 0$,

$$P_x(G \cap \theta_s^{-1}(A)) = \textstyle\int_G P_{X_s(\omega)}(A) d\mathbf{P}_x(\omega) \,\forall\, G \in \mathcal{G}, A \in \mathcal{A}$$

where for each $t \geq 0$, $\theta_t : \Omega \to \Omega$ is the shift operator defined as $(\theta_t(\omega))(u) = \omega(t+u), u \geq 0, \omega \in \Omega$,that is, $\theta_s^{-1}(A) = \{\omega : X(s+, \cdot, \omega) \in A\} = \{\omega : (\theta_s\omega)(\cdot) \in A\}$ for any $A \in \mathcal{A}$.
Hence, $Q(\omega, A) = P_{X_s(\omega)}(A)$ is a RCD of θ_s,given $\mathcal{A}_s$ under $\mathbf{P}_x$ for all $x \in \mathbb{R}$.

Chapter 3

Markov Process

In this section, we are going to study what is called the semigroup theory of continuous time Markov processes. As was seen with Brownian motion, a convenient formulation is when the underlying space is taken as the path space with the process being the canonical coordinate process on it.
The **state space** of the Markov process can be taken as a **second countable, locally compact** and **Hausdorff space** S, equipped with its Borel σ-field $\mathcal{S}$, or , a **separable metric space** S with its Borel σ-field $\mathcal{S}$. In either case, $\mathcal{S}$ equals the smallest σ-field that makes all bounded real continuous functions on S measurable.
Since we study Markov processes with **continuous trajectories** only, we take the state space and underlying space of the Markov process as,

$$(S, \mathcal{S}) = \text{ a second countable, locally compact Hausdorff space } S, \text{ with its Borel } \sigma\text{-field } \mathcal{S}$$
$$\text{and, } \Omega = \text{ space of all continuous functions on } [0, \infty) \text{ into } S$$

And, the coordinate process on Ω is the process $\{X_t,\, t \geq 0\}$ defined on Ω as

$$X_t(\omega) = \omega(t),\, \omega \in \Omega, t \geq 0.$$

which, by definition, has continuous trajectories.
Let $\mathcal{A}$ be defined as the smallest σ-field on Ω such that $X_t : (\Omega, \mathcal{A}) \to (S, \mathcal{S})$ is measurable for all $t \geq 0$. It can be seen that the map $(t, \omega) \mapsto X_t(\omega)$ defined on the product space $([0, \infty) \times \Omega, \mathcal{B}^+ \otimes \mathcal{A})$ into $(S, \mathcal{S})$ is a measurable map (here $\mathcal{B}^+$ is the Borel σ-field on $[0, \infty)$).
$\{\mathcal{A}_t,\, t \geq 0\}$ denotes the natural filtration of the process $\{X_t,\, t \geq 0\}$. Then, $\mathcal{A} = \bigvee_t \mathcal{A}_t$. For each $t \geq 0$, the **shift operator** $\theta_t : \Omega \to \Omega$ is defined as $(\theta_t\, \omega)(\cdot) = \omega(t + \cdot)$, $\omega \in \Omega$. Then, for each $s \geq 0$,

$$\{X_t \circ \theta_s,\, t \geq 0\} \text{ is the post-}s\text{ process } \{X_{s+t},\, t \geq 0\}$$

and $\theta_s : (\Omega, \mathcal{A}) \to (\Omega, \mathcal{A})$ is measurable. Also, the sets $X_t^{-1}(B)$, $B \in \mathcal{S}, t \geq 0$, generate $\mathcal{A}$), clearly by definition.

Remark 3.1. *Let us set up some notations to be used frequently,*

1. $C_b(S)$ *- the class of all bounded real-valued continuous functions on* S.
2. $b(\mathcal{E})$*- the class of all bounded real* $\mathcal{E}$*-measurable functions on* E*, for a measurable space* $(E, \mathcal{E})$*, .*
3. $b(\mathcal{S})$ *- the class of bounded real* $\mathcal{S}$*-measurable functions on* S
4. $b(\mathcal{A})$ *- the class of bounded real* $\mathcal{A}$*-measurable functions on* Ω.
5. $b(\mathcal{A}_s)$*) - the class of bounded real* $\mathcal{A}_s$*-measurable functions on* Ω.

Definition 3.1. *(Markov Process) A family* $\mathbb{M} = (\mathbf{P}_x, x \in S)$ *of probabilities on* $(\Omega, \mathcal{A})$ *defines a* ***Markov process with state space*** S *if it satisfies:*

(a) $\mathbf{P}_x(X_0 = x) = 1$ *for all* $x \in S$

(b) The map $x \to \mathbf{P}_x(A)$ *is* $\mathcal{S}$*-measurable , for each* $A \in \mathcal{A}$

(c) (Markov Property) For each $x \in S$ *and each* $s \geq 0$,

$$P_x(G \cap \theta_s^{-1}(A)) = \int_G P_{X_s(\omega)}(A) d\mathbf{P}_x(\omega)\, \forall\, G \in \mathcal{G}, A \in \mathcal{A}$$

3.1 Transition Probabilities

Definition 3.2. *Let $\mathbb{M} = (\mathbf{P}_x, x \in S)$ be a Markov process with state space S. For $t > 0$, $x \in S$ and $E \in \mathcal{S}$, let us define,*

$$p(t, x, E) = \mathbf{P}_x(X_t \in E)$$

The quantities $\{p(t, x, E) : t > 0, x \in S, E \in \mathcal{S}\}$ are called ***transition probabilities*** *of the Markov process. $p(t, x, E)$ represents the* ***probability of transition from a state*** *x* ***to a set of states*** *E* ***in time*** *t.*

Clearly, for fixed $t > 0$ and $x \in S$, the set function $p(t, x, \cdot)$ is a probability measure on $\mathcal{S}$ and represents the distribution of the S-valued random variable X_t under $\mathbf{P}_x$ (that is, $\mathbf{P}_x X_t^{-1}$).
The map $x \to p(t, x, E)$ is $\mathcal{S}$-measurable , for each fixed $t > 0$ and $E \in \mathcal{S}$. Also, $p(t, x, E)$ is jointly measurable in (t, x) for each fixed $E \in \mathcal{S}$.

Chapman-Kolmogorov Equations:

$$p(s + t, x, E) = \int p(t, y, E)p(s, x, dy) \,\forall\, s, t > 0, x \in S, E \in \mathcal{S}$$

This equation essentially captures the Markov property and is also the heart of the semigroup theory of Markov processes.

Theorem 3.1. *A Markov process is uniquely determined by its transition probabilities.*

Proof. Let $\mathbb{M}^1 = (\mathbf{P}^1_x,\, x] \in S)$ and $\mathbb{M}^2 = (\mathbf{P}^2_x,\, x] \in S)$ be two Markov processes on a state space S and let p^1 and p^2 be the associated transition probabilities. We have to show that if

$$p^1(t, x, E) = p^2(t, x, E),\ \forall\, t > 0, x \in S, E \in \mathcal{S}\ \cdots\ (*),$$

then, $\mathbf{P}^1_x = \mathbf{P}^2_x$ on $\mathcal{A}$, for all $x \in S$. Fix any $0 < t_1 < \cdots < t_n$ and observe that, if $(*)$ holds, then by the fact that, for $0 < t_1 < \cdots < t_n$, $x \in S$ and $E_1, \cdots, E_n \in \mathcal{S}$,

$$\mathbf{P}_x\big(X_{t_1} \in E_1, \cdots, X_{t_n} \in E_n\big) = \int_{E_1} \cdots \int_{E_n} p(t_n - t_{n-1}, x_{n-1}, dx_n) \cdots p(t_2 - t_1, x_1, dx_2)p(t_1, x, dx_1),$$

we will have, $\mathbf{P}^1_x\big(X_{t_1} \in E_1, \cdots, X_{t_n} \in E_n\big) = \mathbf{P}^2_x\big(X_{t_1} \in E_1, \cdots, X_{t_n} \in E_n\big)$, for all $x \in S$.
Since sets of the form, $\{X_{t_1} \in E_1, \cdots, X_{t_n} \in E_n\}$ with $E_1 \in \mathcal{S}, \cdots, E_n \in \mathcal{S}$, form a semi-field that generates $\sigma\big(\{X_{t_1}, \cdots, X_{t_n}\}\big)$, we conclude that if $(*)$ holds then

$$\mathbf{P}^1_x = \mathbf{P}^2_x \ \text{ on } \ \sigma\big(\{X_{t_1}, \cdots, X_{t_n}\}\big),\ \text{ for all } x \in S$$

But this is true for all choices of $0 < t_1 < \cdots < t_n$ and $\mathcal{F} = \bigcup \big\{\sigma\big(\{X_{t_1}, \cdots, X_{t_n}\}\big) : 0 < t_1 < \cdots < t_n,\ n \geq 1\big\}$ is a field that generates $\sigma\big(\{X_t,\ t > 0\}\big)$. Thus if $(*)$ holds then,

$$\mathbf{P}^1_x = \mathbf{P}^2_x \ \text{ on } \ \sigma\big(\{X_t,\ t > 0\}\big),\ \text{ for all } x \in S$$

Now, since $\mathbf{P}^1_x(X_0 = x) = \mathbf{P}^2_x(X_0 = x) = 1$, for all $x \in S$, if $(*)$ holds, then

$$\mathbf{P}^1_x = \mathbf{P}^2_x \ \text{ on } \ \mathcal{A} = \sigma\big(\{X_t,\ t \geq 0\}\big),\ \text{ for all } x \in S$$

This completes the proof . □

Theorem 3.2. *Let $(\mathbf{P}_x,\ x \in S)$ be a family of probabilities on $(\Omega, \mathcal{A})$ and let $p(t, x, E) = \mathbf{P}_x(X_t \in E)$, for $t > 0$, $x \in S$ and $E \in \mathcal{S}$. If $p(t, x, E)$ satisfy the properties (1), (2) and (3) above, then $\mathbb{M} = (\mathbf{P}_x, x \in S)$ is a Markov process.*

3.2 Transition Semigroup

Let $\mathbb{M} = (\mathbf{P}_x, x \in S)$ be a Markov process with state space S and let $p(t, x, E), t > 0, x \in S, E \in \mathcal{S}$ be its transition probabilities. For $t > 0$ and $f \in b(\mathcal{S})$, let us define a map $T_t f$ on S by

$$T_t f(x) = \int f(y)p(t, x, dy) = \mathbf{E}_x(f(X_t)), x \in S$$

For $f \in b(\mathcal{S})$, the map $(t, x) \to T_t f(x)$ is jointly measurable.
In particular, for $t > 0$ and $f \in b(\mathcal{S})$, we have $T_t f \in b(\mathcal{S})$. Thus , for $t > 0$, T_t defines a map on $b(\mathcal{S})$ into $b(\mathcal{S})$.
For $t = 0$, $T_0 f(x) = \mathbf{E}_x(f(X_0)) = f(x)$, that is , T_0 is the identity map on $b(\mathcal{S}$.

Properties of T_t:

- (Linearity) $T_t : b(\mathcal{S}) \to b(\mathcal{S})$ is linaer , for each t.
- (Positivity) T_t is positive for each t,that is, for $f \in b(\mathcal{S}), f \geq 0, T_f \geq 0$.
- (Constant-preserving) $T_t 1 = 1$ for each t, where "1" denotes the constant function 1.
- (Contraction) T_t is a contraction for each t, that is, for $f \in b(\mathcal{S}), ||T_t f|| \leq ||f||$
- $\{T_t, t \geq 0\}$ is a semigroup, that is, $T_{s+t} = T_s T_t$.
- (Weak Continuity) $T_t : b(\mathcal{S}) \to b(\mathcal{S})$ for each t is weakly continuous, that is, if $f \in b(\mathcal{S}), ||f_n|| \leq M \, \forall \, n \geq 1$ and $f_n \to f$ pointwise, then $T_t f_n \to T_t f$ pointwise.
- For $f \in C_b(S), t \to T_t f(x)$ continuous for each $x \in S$.
- If $f \in b(\mathcal{S})$ is continuous at x, then $\lim_{t \downarrow 0} T_t f(x) = f(x)$

Thus, the operators $\{T_t\}$ form a semigroup of operators on the Banach space $b(\mathcal{S})$ and the semigroup uniquely determines the Markov process.

3.2.1 Green Operators /Resolvent Operators

The formal Laplace transform of the semigroup $\{T_t\}$,

$$R_\alpha = \int_0^\infty e^{-\alpha t} T_t dt, \alpha > 0$$

is known as the **Green Operators/Resolvent Operators**.
More specifically, for $f \in b(\mathcal{S})$ and $x \in S$, we define,$R_\alpha f(x) = \int_0^\infty e^{-\alpha t} T_t f(x) dt, \alpha > 0$

Properties of Green Operator:

- $R_\alpha f \in b(\mathcal{S})$, for $f \in b(\mathcal{S})$, indeed, $||R_\alpha f|| \leq \alpha^{-1} ||f||$
- $R_\alpha : b(\mathcal{S}) \to b(\mathcal{S})$ is **linear** , for each $\alpha > 0$.
- R_α for each $\alpha > 0$, is a **positive** operator,that is, if $f \in b(\mathcal{S}), f \geq 0$, then $R_\alpha f \geq 0$
- $R_\alpha : b(\mathcal{S}) \to b(\mathcal{S})$ is weakly continuous .
- If $f \in b(\mathcal{S})$ is continuous at x,then $\alpha R_\alpha f(x) \to f(x)$ as $\alpha \to 0$.

3.2.2 Green Measure

For $\alpha > 0, x \in S, E \in \mathcal{S}$,

$$R(\alpha, x, E) = (R_\alpha I_E)(x) = \int_0^\infty e^{-\alpha t} p(t, x, E) dt$$

$R(\alpha, x, \cdot)$ is a finite measure on $\mathcal{S}$ for each $\alpha > 0$ and $x \in S$ (in fact, $R(\alpha, x, S) = \alpha^{-1}$).
The measure $R(\alpha, x, \cdot)$ is called the **Green measure**.

3.2.3 Resolvent Equation:

$R_\alpha - R_\beta + (\alpha - \beta) R_\alpha R_\beta = 0$

Consequences :

- (Semigroup Property of $\{T_t\}$) $T_s T_t = T_{s+t} = T_t T_s$
- $R_\alpha R_\beta = R_\beta R_\alpha$
- $T_t R_\alpha = R_\alpha T_t$

- Range, $\mathcal{R}_\alpha = \{R_\alpha f : f \in b(\mathcal{S})\}$ and Null space, $\eta_\alpha = \{f \in b(\mathcal{S}) : R_\alpha f = 0\}$
  Now , $\mathcal{R}_\alpha$ have the same range and same null space for all $\alpha > 0$, that is, $\mathcal{R}_\alpha = \mathcal{R}_\beta$ and $\eta_\alpha = \eta_\beta \,\forall\, \alpha \neq \beta$.

The last consequence plays a very important role in defining the infinitesimal generator of a Markov Process.

3.3 Infinitesimal Generator

Let us denote the common range space of the Green operators by $\mathcal{R}$ and the common null space of the Green operators by η. $\mathcal{R} \cap \eta = \{0\}$.
If $g \in \mathcal{R}$, then $\forall\, \alpha > 0, \exists f \in b(\mathcal{S})$ s.t. $\mathcal{R}_\alpha f = g$. Such a f is denoted as $\mathcal{R}_\alpha^{-1} g$, depends on α, i.e., $f = \mathcal{R}_\alpha^{-1} g$ will not be unique, but will be unique only upto $(mod\eta)$.

We define an operator A_α on $\mathcal{R}$, as

$$A_\alpha g = \alpha g - \mathcal{R}_\alpha^{-1} g, g \in \mathcal{R} \text{ for each } \alpha > 0$$

Now, for $\alpha, \beta > 0, \alpha \neq \beta, A_\alpha \equiv A_\beta$, that is, A_α does not depend on α. Hence we denote A_α by A. This operator A with domain $\mathcal{D}(A) = \mathcal{R}$ and defined upto $mod\eta$ is called the **infinitesimal generator** or simply the **generator** of the underlying Markov process.
$A : \mathcal{R} \to b(\mathcal{S})$ is **linear** and will usually be an **unbounded** operator.

Theorem 3.3. *The generator of a Markov process uniquely determines the Markov process.*

Definition 3.3. *(Strong Markov Property) A Markov Process $M = (\mathbf{P}_x, x \in S)$ is said to have the **Strong Markov Property**, if , for any finite optional time τ,*

$$\mathbf{P}_x(G \cap \theta_\tau^{-1}(A)) = \int_G \mathbf{P}_{X_\tau}(A) d\mathbf{P}_x, \,\forall\, x \in S, G \in \mathcal{A}_{\tau+}, A \in \mathcal{A}$$

Remark: **Not every** continuous path Markov process satisfies the strong Markov property. Thus we are going to describe an interesting property of the semigroup that works as a sufficient condition for strong Markov property.
We know , for $f \in C_b(S), T_t f(x)$ is always continuous in t, but it may not be continuous in x. This leads to the following definition.

Definition 3.4. *(Feller Property) A Markov process is said to have the Feller Property if the associated transition operators T_t map $C_b(S)$ into $C_b(S)$. In that case, the Markov process is called a **Feller process**.*

Theorem 3.4. *Every Feller process has the Strong Markov Property.*

Apart from having the Strong markov Property, Feller processes ave another very important advantage, for a general continuous path Markov process, the Transition operators (T_t) and the Green operators $(\mathcal{R}_\alpha)$ are all bounded linear operators on the Banach space $b(\mathcal{S})$. But for a Feller process, we can consider the smaller Banach space $C_b(S)$ (with the same sup-norm) and **restrict** the operators $\{T_t, t \geq 0\}$ and $\{\mathcal{R}_\alpha, \alpha > 0\}$ to $C_b(S)$.
By the definition of Feller property, T_t, for each $t \geq 0$, is a bounded ,linear operator on $C_b(S)$ into $C_b(S)$. Thus , for $f \in C_b(S)$, $T_t f \in C_b(S)$ and $\mathcal{R}_\alpha f \in C_b(S)$.

Let us now consider the Green operators restricted to $C_b(S)$ and consider the range and null space of these **restricted Green operators**. Using the resolvent equation, one can again see that the restricted Green operators all have the same range and null space. Let us denote the common range and null space of the restricted Green operators by $\tilde{\mathcal{R}}$ and $\tilde{\eta}$ respectively, that is , $\tilde{\mathcal{R}} = \{\mathcal{R}_\alpha f : f \in C_b(S)\}$ and $\tilde{\eta} = \{f \in C_b(S) : \mathcal{R}_\alpha f = 0\}$.
Now, $\tilde{\eta} = \{0\}$, i.e., the restricted null space is trivial.

An important consequence is that , for $g \in \tilde{\mathcal{R}}$, there is a **unique** $f \in C_b(S)$ such that $\mathcal{R}_\alpha f = g$, that is, $\mathcal{R}_\alpha^{-1} : \tilde{\mathcal{R}} \to C_b(S)$ is uniquely defined. Thus , the **restricted generator** $Ag = \alpha g - \mathcal{R}_\alpha^{-1} g, g \in \tilde{\mathcal{R}}$ is a uniquely defined linear operator $A : \tilde{\mathcal{R}} \to C_b(S)$.

Theorem 3.5. *For a Feller process , the restricted generator uniquely determines the process.*

Two results for Strong Markov Property

Theorem 3.6. *Let $(\mathbf{P}_x, x \in S)$ be a Strong Markov process and τ a finite optional time.*

(a) Then, for any $f \in b(\mathcal{S})$ and $\alpha > 0$,

$$\mathcal{R}_\alpha f(x) = E_x(\int_0^\tau e^{-\alpha t} f(X_t)dt) + E_x(e^{-\alpha t}\mathcal{R}_\alpha f(X_\tau))$$

(b) (Dynkin's Formula) If $E_x(\tau) < \infty$, then for any $g \in \mathcal{D}(A)$,

$$E_x(\int_0^\tau Ag(X_t)dt) = E_x(g(X_\tau) - g(x)$$

Remark: The Dynkin's Formula generates a large class of continuous martingles associated to a Strong Markov process.

3.4 General theory of Markov Processes on Brownian Motion

In this section , we return to Brownian motion and examine our general theory on this special Markov process. We will now view Brownian motion as a Markov process in the canonical framework. The state space now is $S = \mathbb{R}$ with $\mathcal{S} = \mathcal{B}$ and the Markov process is $\mathbb{M} = (\mathbf{P}_x, x \in \mathbb{R})$ where $\mathbf{P}_x$, for each $x \in \mathbb{R}$, represents the "law" of Brownian motion starting at x (that is, the probability induced on $(\Omega, \mathcal{A}) = (C, \mathcal{C})$ by a Brownian motion starting at x). Thus, $\mathbb{M} = (\mathbf{P}_x, x \in \mathbb{R})$ is a family of probabilities on $(\Omega, \mathcal{A}) = (C, \mathcal{C})$ such that, for each $x \in \mathbb{R}$, the coordinate process $\{X_t, t \geq 0\}$ on $(\Omega, \mathcal{A})$ is, under $\mathbf{P}_x$, a Brownian motion starting at x. From the definition of Brownian motion starting at x, it follows that the $\mathbf{P}_x$-distribution of X_t is $\mathrm{N}(x, t)$, for $t > 0$.
Thus the transition probabilities are given by :

$$p(t, x, E) = \frac{1}{\sqrt{2\pi t}} \int_E e^{-\frac{(y-x)^2}{2t}}\, dy, \;\text{ for }\; t > 0, x \in \mathbb{R}, E \in \mathcal{B}$$

It then follows that, for $f \in b(\mathcal{B})$ and $t > 0$,

$$T_t f(x) = \mathbf{E}_x\big(f(X_t)\big) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} f(y) e^{-\frac{(y-x)^2}{2t}}\, dy \;\cdots\; (1)$$

Using (1), we will see later that Brownian motion is not just a Feller process, but it actually satisfies an even stronger property.Next, the Green operators for the Brownian motion are

$$R_\alpha f(x) = \int_0^\infty e^{-\alpha t} \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} f(y) e^{-\frac{(y-x)^2}{2t}}\, dy\, dt, \text{ for } f \in b(\mathcal{B}) \text{ and } \alpha > 0.$$

We now proceed to simplify the above formula for $R_\alpha f(x)$ in a way that will help us not only to clearly identify the range of the Green operators but also to identify the generator of the Brownian motion.
Firstly, $\frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} |f(y)| e^{-\frac{(y-x)^2}{2t}}\, dy \leq \|f\|,\ \forall\, t > 0$ implies

$$\int_0^\infty \int_{-\infty}^{\infty} \Big|e^{-\alpha t} \frac{1}{\sqrt{2\pi t}} f(y) e^{-\frac{(y-x)^2}{2t}}\Big|\, dy\, dt \leq \|f\|/\alpha < \infty$$

The above means that we may interchange the order of integration in the formula for $R_\alpha f(x)$ to get

$$R_\alpha f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(y) \int_0^\infty \frac{1}{\sqrt{t}}\, e^{-\left(\alpha t + \frac{(y-x)^2}{2t}\right)}\, dt\, dy, \text{ for all } f \in b(\mathcal{B}) \text{ and } \alpha > 0.$$

The next step is to evaluate the inner integral explicitly.
Denoting $I_1(y) = \int_0^\infty \frac{1}{\sqrt{t}}\, e^{-\left(\alpha t + \frac{(y-x)^2}{2t}\right)}\, dt,\ y \in (-\infty, \infty)$, we have $\quad R_\alpha f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(y) I_1(y)\, dy \;\cdots\; (2)$
The substitution $u = \sqrt{t}$ in $I_1(y)$ gives

$$I_1(y) = 2\int_0^\infty e^{-\left(\alpha u^2 + \frac{(y-x)^2}{2u^2}\right)} du = 2e^{-\sqrt{2\alpha}\,|y-x|} \int_0^\infty e^{\left(au - \frac{b}{u}\right)^2} du, \text{ where } a = \sqrt{\alpha} \;\text{ and }\; b = |y-x|/\sqrt{2}$$

For $a > 0$, $b > 0$, let us denote $I_2 = \int_0^\infty e^{\left(au - \frac{b}{u}\right)^2} du$. The substitution $v = \frac{b}{au}$ gives $I_2 = \int_0^\infty \frac{b}{av^2}\, e^{\left(av - \frac{b}{v}\right)^2} dv$.
The above two give $I_2 = \frac{1}{2}\int_0^\infty \left(1 + \frac{b}{au^2}\right) e^{\left(au - \frac{b}{u}\right)^2} du$.The substitution $z = au - \frac{b}{u}$ finally yields

$$I_2 = \frac{1}{2a}\int_{-\infty}^{\infty} e^{-z^2} dz = \frac{\sqrt{\pi}}{2a}$$

Plugging the value of the integral I_2 with $a = \sqrt{\alpha}$ in the formula for $I_1(y)$, we get $I_1(y) = \sqrt{\frac{\pi}{\alpha}} e^{-\sqrt{2\alpha}\,|y-x|}$
Using this in (2) gives us the following simple formula for the Green operators of Brownian motion :

$$R_\alpha f(x) = \frac{1}{\sqrt{2\alpha}} \int_{-\infty}^{\infty} f(y) e^{-\sqrt{2\alpha}\,|y-x|}\, dy \quad \cdots \quad (3)$$

This shows that for Brownian motion, Green measures $R(\alpha, x, \cdot)$ on $\mathcal{B}$ are absolutely continuous with respect to Lebesgue measure and have density given by,

$$r_{\alpha,x}(y) = \frac{1}{\sqrt{2\alpha}}\, e^{-\sqrt{2\alpha}\,|y-x|}, \quad y \in (-\infty, \infty)$$

From (3), one can get a simple description of the range $\mathcal{R}$ of the Green operators of Brownian motion and also a clear identification of the generator A on $\mathcal{D}(A) = \mathcal{R}$. However, to keep things slightly simpler, we will first see that Brownian motion is a Feller process and then consider only the restricted Green operators on $C_b(\mathbb{R})$. Actually we are going to see that Brownian motion satisfies a property stronger than Feller property. Recall that for Brownian motion, T_t, $t > 0$ are given by

$$T_t f(x) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} f(y) e^{-\frac{(y-x)^2}{2t}}\, dy, \text{ for } f \in b(\mathcal{B}),\ x \in \mathbb{R},\ t > 0.$$

Using DCT, it is easy to show that $T_t f(x)$ is **continuous in** x, for any $f \in b(\mathcal{B})$ and $t > 0$. This means that $T_t f \in C_b(\mathbb{R})$ for all $f \in b(\mathcal{B})$, that is,

$$T_t : b(\mathcal{B}) \to C_b(\mathbb{R}), \text{ for all } t > 0.$$

When a Markov process satsifies $T_t : b(\mathcal{B}) \to C_b(\mathbb{R})$, for all $t > 0$, we say that it has the **Strong Feller property**.

We have thus seen that Brownian motion has Strong Feller property, clearly stronger than the Feller property. Since Brownian motion is a Feller process, we may and will consider the Green operators R_α **restricted** to the Banach space $C_b(\mathbb{R})$.

In that case, $R_\alpha : C_b(\mathbb{R}) \to \widetilde{\mathcal{R}}$, for all $\alpha > 0$, is one-one and onto, as seen before, where

$$\widetilde{\mathcal{R}} = \{R_\alpha f : f \in C_b(\mathbb{R})\}, \text{ for all } \alpha > 0.$$

In case of Brownian motion, the formula (3) for Green operators implies that $\widetilde{\mathcal{R}}$ **consists precisely of all functions** g given by

$$g(x) = \frac{1}{\sqrt{2\alpha}} \int_{-\infty}^{\infty} f(y) e^{-\sqrt{2\alpha}\,|y-x|}\, dy, \text{, for } \alpha > 0 \text{ and an } f \in C_b(\mathbb{R}).$$

Further, for each $g \in \widetilde{\mathcal{R}}$ (and $\alpha > 0$), the $f \in C_b(\mathbb{R})$ as above **is unique**, so that $R_\alpha^{-1} g = f$ and hence the restricted generator is given by

$$Ag = \alpha g - f,\ g \in \widetilde{\mathcal{R}}$$

To get a clearer identification of the class $\widetilde{\mathcal{R}}$, we will start with an $f \in C_b(\mathbb{R})$ and an $\alpha > 0$ and observe some properties of the function g given by the above formula.

For $f \in C_b(\mathbb{R})$, $\alpha > 0$, the function $g = R_\alpha f$ equals

$$g(x) = \frac{e^{-\sqrt{2\alpha}x}}{\sqrt{2\alpha}} \int_{-\infty}^{x} f(y) e^{\sqrt{2\alpha}y} dy + \frac{e^{\sqrt{2\alpha}x}}{\sqrt{2\alpha}} \int_{x}^{\infty} f(y) e^{-\sqrt{2\alpha}y} dy$$

Since $f \in C_b(\mathbb{R})$, it follows from Fundamental Theorem of Calculus that g is differentiable everywhere and its derivative $g' \in C_b(\mathbb{R})$ is given by

$$g'(x) = -e^{-\sqrt{2\alpha}x} \int_{-\infty}^{x} f(y) e^{\sqrt{2\alpha}y} dy + e^{\sqrt{2\alpha}x} \int_{x}^{\infty} f(y) e^{-\sqrt{2\alpha}y} dy$$

It follows once again that g' is differentiable everywhere and $g'' \in C_b(\mathbb{R})$ is given by

$$g''(x) = \sqrt{2\alpha} e^{-\sqrt{2\alpha}x} \int_{-\infty}^{x} f(y) e^{\sqrt{2\alpha}y} dy$$
$$+ \sqrt{2\alpha} e^{\sqrt{2\alpha}x} \int_{x}^{\infty} f(y) e^{-\sqrt{2\alpha}y} dy - 2f(x) = 2\alpha g(x) - 2f(x)$$

What we have shown, therefore, is that

$$g \in \widetilde{\mathcal{R}} \implies g'' \in C_b(\mathbb{R}) \text{ and } g'' = 2\alpha\, g - 2R_\alpha^{-1} g$$

Conversely, if $g \in C_b(\mathbb{R})$ is such that $g'' \in C_b(\mathbb{R})$, then, by taking $f = \alpha\, g - \frac{1}{2} g'' \in C_b(\mathbb{R})$, one can easily see that $g = R_\alpha f$ and hence $g \in \widetilde{\mathcal{R}}$. All these lead to the following final result.

Theorem 3.7. *For the Brownian motion, the range of the Green operators restricted to the Banach space* $C_b(\mathbb{R})$ *equals*

$$\widetilde{\mathcal{R}} = \{g \in C_b(\mathbb{R}) : g'' \in C_b(\mathbb{R})\}$$

and the restricted generator A *is given by*

$$Ag = \tfrac{1}{2} g'',\quad g \in \widetilde{\mathcal{R}} = \widetilde{\mathcal{D}}(A)$$

If one is interested in the **full generator** for Brownian motion, that is, in considering the Green operators on the full Banach space $b(\mathcal{B})$ and defining the generator on the full range $\mathcal{R}$, then, with some work, one can derive the following descriptions.
Firstly, the common null space of the Green operators can be shown to be

$$N = \{f \in b(\mathcal{B}) : f = 0 \text{ a.e.}\}$$

Next, with a little work, the common range of the Green operators on $b(\mathcal{B})$ can be shown to equal

$$\mathcal{R} = \{g \in b(\mathcal{B}) : g \text{ abs cts, } g' \text{ abs cts, } g'' \in b(\mathcal{B})\}$$

These lead to the following description of the **full generator** for the Brownian motion.

Theorem 3.8. *The full generator of the Brownian motion is given by*

$$\mathcal{D}(A) = \{g \in b(\mathcal{B}) : g \textit{ abs cts, } g' \textit{ abs cts, } g'' \in b(\mathcal{B})\}$$

and

$$Ag = \tfrac{1}{2}g'' \quad \textit{a.e.,} \quad \textit{for } g \in \mathcal{D}(A)$$

Chapter 4

Wiener Integral

Let $\{B_t, t \geq 0\}$ be a standard brownian motion (SBM) on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$.
Let us consider the class of integrands $L^2([0,T])$, defined as,

$$L^2([0,T]) = L^2([0,T], \mathcal{B}([0,T]), \lambda) = \{f : f \text{ measurable on } [0,T], \int_0^T f^2(t)df < \infty\}$$

Now, let us consider the L^2 space of square intgrable random variables on $(\Omega, \mathcal{A}, \mathbf{P})$, defined as,

$$\mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P}) = \{X : X \text{ measurable on } \Omega, \int_\Omega X^2 dP < \infty\}$$

$$= \{X \text{ is a random variable }, \int_\Omega X^2 dP < \infty\}$$

Note that, both $L^2([0,T])$ and $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P})$ are Hilbert spaces with the usual L^2 norm.
Let us define another class of functions as,

$$\mathcal{E} = \{\sum_{i=1}^{n} c_i 1_{(t_{i-1},t_i]} : n \geq 1, 0 \leq t_0 < t_1 < ... < t_n \leq T\}$$

4.1 Definite integral

Let us fix $T \geq 0$. In this section, we aim to define the integral of $L^2[0,T]$ functions with respect to the standard brownian motion, that is to define

$$\int_0^T f_s dB_s \text{ for } f \in L^2([0,T])$$

We denote this integral by $I_T(f)$, i.e.,

$$I_T(f) = \int_0^T f_s dB_s \text{ for } f \in L_2([0,T])$$

Theorem 4.1. *$\mathcal{E}$ is a vector subspace of $L^2[0,T]$*

Proof. Let us consider a function $f \in \mathcal{E}$. Then , for some $n \geq 1$ and $0 \leq t_0 < t_1 < t_2 < ... < t_n \leq T$,

$$f = \sum_{i=1}^n c_i 1_{(t_{i-1},t_i]}$$

Hence, $f = \sum_{i=1}^n c_i 1_{(t_{i-1},t_i]} = c_1 1_{(t_0,t_1]} + c_2 1_{(t_1,t_2]} + ... + c_n 1_{(t_{n-1},t_n]}$, which is measurable.
Now, $||f||_{L^2[0,T]} = (\int_0^T f^2(t)dt)^{1/2} = [\int_{t_0}^{t_1} f^2(t)dt + \int_{t_1}^{t_2} f^2(t)dt + ... + \int_{t_n-1}^{t_n} f^2(t)dt]^{1/2} = [\int_{t_0}^{t_1} {c_1}^2 dt + \int_{t_1}^{t_2} {c_2}^2 dt + ... + \int_{t_n-1}^{t_n} {c_n}^2 dt]^{1/2} = [{c_1}^2(t_1 - t_0) + ... + {c_n}^2(t_n - t_{n-1})]^{1/2} = [\sum_{i=1}^n c_i^2(t_i - t_{i-1}]^{1/2} < \infty$.
Since, $||f||_{L^2[0,T]} < \infty$ and f measurable, $f \in L^2[0,T]$. Hence $\mathcal{E} \subseteq L^2[0,T]$.
Now, let

$$f = \sum_{i=1}^n c_i 1_{(t_{i-1},t_i]}, n \geq 1, 0 \leq t_0 < t_1 < ... < t_n \leq T$$

and

$$g = \sum_{j=1}^m d_j 1_{(s_{j-1},s_j]}, m \geq 1, 0 \leq s_0 < s_1 < ... < s_m \leq T$$

be two functions in $\mathcal{E}$.
Then for $A, B \in \mathbb{R}$, $Af + Bg = A\sum_{i=1}^n c_i 1_{(t_{i-1},t_i]} + B\sum_{j=1}^m d_j 1_{(s_{j-1},s_j]} = \sum_{i=1}^n Ac_i 1_{(t_{i-1},t_i]} + \sum_{j=1}^m Bd_j 1_{(s_{j-1},s_j]} = \sum_{i=1}^n \sum_{j=1}^m (Ac_i + Bd_j) 1_{(t_{i-1} \vee s_{j-1}, t_i \wedge s_j]} \in \mathcal{E}$. Thus , $\mathcal{E}$ ia a vector subspace of $L^2[0,T]$ $\square$

Lemma 4.1. *Suppose $(\Omega, \mathcal{A}, \mu)$ is a finite measure space. Let $\mathcal{F}$ be a field generating $\mathcal{A}$. Then the class of simple functions of the form $g = \sum_{i=1}^n c_i 1_{F_i}, F_i \in \mathcal{F}$ is dense in $L^2(\Omega, \mathcal{A}, \mu)$.*

Proof. Let us consider a function $f \in L^2(\Omega, \mathcal{A}, \mu)$. Let $\epsilon > 0$ be arbitrary. We need to show that

$$\exists \text{ a } g = \sum_{i=1}^{n} c_i 1_{F_i}, F_i \in \mathcal{F}, \ \forall\, i, \text{ such that, } ||g - f||_{L^2} < \epsilon.$$

We know that, for such a f there will exist a sequence $\{h_n\}$ of simple functions in $L^2(\Omega, \mathcal{A}, \mu)$ with $|h_n| \leq |f|$, $\forall\, n$ and $h_n \to f$ in L^2.Hence we get a simple function, $h = \sum_{j=1}^{m} d_j 1_{A_j}, A_j \in \mathcal{A}$ such that,

$$||h - f||_{L^2} < \epsilon/2 \ \cdots \ (*)$$

Now, using a very well-known result from measure theory, we have that , if $\mathcal{F}$ is a field generating $\mathcal{A}$, then $\forall\, A \in \mathcal{A}$, $\forall\, \epsilon > 0, \exists F \in \mathcal{F} : \sqrt{\mu(A \triangle F)} < \epsilon$. Let us consider the simple function $h = \sum_{j=1}^{m} d_j 1_{A_j}, A_j \in \mathcal{A}$. Then for each such j, we will get a $F_j \in \mathcal{F}$ such that,

$$\mu(A_j \triangle F_j) < \epsilon/2m|d_j|$$

Consider $g = \sum_{j=1}^{m} d_j 1_{F_j}$.

Now, $||g - h||_{L^2} = || \sum_{j=1}^{m} d_j (1_{A_j} - 1_{F_j})||_{L^2} \leq \sum_{j=1}^{m} |d_j| ||1_{A_j} - 1_{F_j}||_{L^2} = \sum_{j=1}^{m} |d_j| \sqrt{\int_\Omega |1_{A_j} - 1_{F_j}|^2 d\mu} = \sum_{j=1}^{m} |d_j| \sqrt{\int_\Omega |1_{A_j \triangle F_j}| d\mu} = \sum_{j=1}^{m} |d_j| \sqrt{\mu(A_j \triangle F_j)} < \sum_{j=1}^{m} |d_j| \epsilon/2m|d_j| = \epsilon/2 \ \cdots \ (**)$

Combining $(*)$ and $(**)$, we get

$$||g - f||_{L^2} \leq ||g - h||_{L^2} + ||h - f||_{L^2} < \epsilon/2 + \epsilon/2 = \epsilon$$

Hence, for a $f \in L^2(\Omega, \mathcal{A}, \mu)$, we get a $g = \sum_{i=1}^{n} c_i 1_{F_i}, F_i \in \mathcal{F}\, \forall\, i$ such that $||g - f||_{L^2} < \epsilon$, thus proving that the class of simple functions of the form $g = \sum_{i=1}^{n} c_i 1_{F_i}, F_i \in \mathcal{F}$ is dense in $L^2(\Omega, \mathcal{A}, \mu)$. □

Theorem 4.2. *$\mathcal{E}$ is dense in $L^2[0, T]$.*

Proof. This can be proved using simple analogy to the previous lemma.

In the previous lemma, let us consider $\Omega = [0, T]$, $\mathcal{A} = \mathcal{B}[0, T]$, $\mu = \lambda$. Since [0,T] is a finite interval and lebesgue measure of a finite interval is finite, $([0, T]\mathcal{B}[0, T], \lambda)$ is a finte measure space.Let $\mathcal{S} = \{(s, t], 0 \leq s \leq t \leq T\}$ form a semifield and the finite disjoint union of sets from $\mathcal{S}$ forms a field, $\mathcal{F}$ that generates $\mathcal{A} = \mathcal{B}[0, T]$. Hence, by lemma 4.1, the class of functions of the form,

$$g = \sum_{i=1}^{n} c_i 1_{F_i}, F_i \in \mathcal{F}$$

is dense in $L^2([0, T], \mathcal{B}[0, T], \lambda)$. Thus, it is enough to show that $g \in \mathcal{E}$.

We prove it by the method of induction. For n=2, let $F_1 = (s_1, t_1] \cup (s_2, t_2] \cup \cdots \cup (s_k, t_k]$, disjoint intervals and $F_2 = (s'_1, t'_1] \cup (s'_2, t'_2] \cup \cdots \cup (s'_l, t'_l]$, disjoint intervals, where $0 \leq s_1 \leq s_2 \leq \cdots \leq t_{k-1} \leq s_k \leq t_k \leq T$ and $0 \leq s'_1 \leq s'_2 \leq \cdots \leq t'_{l-1} \leq s'_l \leq t'_l \leq T$. Then

$$g = c_1 1_{F_1} + c_2 1_{F_2} = \sum_{i=1}^{k} c_1 1_{(s_i, t_i]} + \sum_{j=1}^{l} c_2 1_{(s'_j, t'_j]} = \sum_i \sum_j (c_1 + c_2) 1_{(s_i \vee s'_i, t_i \wedge t'_i]} \in \mathcal{E}$$

Thus, for n=2, $g = \sum_{i=1}^{n} c_i 1_{F_i} \in \mathcal{E}$.

Let $g \in \mathcal{E}$ for n=k. Now, for n=k+1, $g = \sum_{i=1}^{k+1} c_i 1_{F_i} = \sum_{i=1}^{k} c_i 1_{F_i} + c_{k+1} 1_{F_{k+1}}$. Let $F_{k+1} = (s''_1, t''_1] \cup (s''_2, t''_2] \cup \cdots \cup (s''_m, t''_m]$, disjoint intervals and $0 \leq s''_1 \leq s''_2 \leq \cdots \leq t''_{m-1} \leq s''_m \leq t''_m \leq T$. Then $c_{k+1} 1_{F_{k+1}} = \sum_{p=1}^{m} c_{k+1} 1_{(s''_p, t''_p]} \in \mathcal{E}$. Since $\mathcal{E}$ is a vector space and $\sum_{i=1}^{k} c_i 1_{F_i} \in \mathcal{E}$ and $c_{k+1} 1_{F_{k+1}} \in \mathcal{E}$,

$$g = \sum_{i=1}^{k+1} c_i 1_{F_i} = \sum_{i=1}^{k} c_i 1_{F_i} + c_{k+1} 1_{F_{k+1}} \in \mathcal{E}$$

Hence, by the principle of mathematical induction, $g \in \mathcal{E}$. Thus completes the proof. □

Let us define a map I_T on $\mathcal{E}$ by,

$$I_T(f) = I_T(\sum_{i=1}^{n} c_i 1_{(t_{i-1}, t_i]}) = \sum_{i=1}^{n} c_i (B_{t_i} - B_{t_{i-1}}), f \in \mathcal{E}$$

Clearly , for all $f \in \mathcal{E}$, $I_T(f) \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P})$.

Theorem 4.3. *I_T is a linear isometry on $\mathcal{E}$.*

Proof. Let $f = \sum_{i=1}^{n} c_i 1_{(t_{i-1}, t_i]}, n \geq 1, 0 \leq t_0 < t_1 < ... < t_n \leq T$ and $g = \sum_{j=1}^{m} d_j 1_{(s_{j-1}, s_j]}, m \geq 1, 0 \leq s_0 < s_1 < ... < s_m \leq T$ be two functions in $\mathcal{E}$. Then for $A, B \in [0, T]$, $(Af + Bg) \in \mathcal{E}$, since $\mathcal{E}$ is a vector space. Thus I_T is defined for (Af+Bg). Hence, $I_T(Af + Bg) = I_T(A \sum_{i=1}^{n} c_i 1_{(t_{i-1}, t_i]} + B \sum_{j=1}^{m} d_j 1_{(s_{j-1}, s_j]}) = I_T(\sum_{i=1}^{n} Ac_i 1_{(t_{i-1}, t_i]} + \sum_{j=1}^{m} Bd_j 1_{(s_{j-1}, s_j]}) = I_T(\sum_{i=1}^{n} \sum_{j=1}^{m} (Ac_i + Bd_j) 1_{(t_{i-1} \vee s_{j-1}, t_i \wedge s_j]}) = \sum_{i=1}^{n} \sum_{j=1}^{m} (Ac_i + Bd_j)(B_{t_i \wedge s_j} - B_{t_{i-1} \vee s_{j-1}}) = AI_T(f) + BI_T(g)$, thus I_T is linear on $\mathcal{E}$.

Now, $||I_T(f)||_{\mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P})} = [E((I_T(f))^2)]^{1/2} = [E((\sum_{i=1}^{n} c_i (B_{t_i} - B_{t_{i-1}}))^2)]^{1/2} = [E(\sum_{i=1}^{n} c_i^2 (B_{t_i} - B_{t_{i-1}})^2) + 2E(\sum_{i<j} c_i c_j (B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})) + 2E(\sum_{i>j} c_i c_j (B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}}))]^{1/2}$

Here, $E(B_{t_i} - B_{t_{i-1}}) = E(B_{t_j} - B_{t_{j-1}}) = 0$, since for $s \geq 0, t > 0, B_{s+t} - B_s \sim N(0, t)$

Hence, $||I_T(f)||_{\mathcal{L}^2} = [E(\sum_{i=1}^{n} c_i^2 (B_{t_i} - B_{t_{i-1}})^2)]^{1/2} = [\sum_{i=1}^{n} c_i^2 E((B_{t_i} - B_{t_{i-1}})^2)]^{1/2} = [\sum_{i=1}^{n} c_i^2 Var((B_{t_i} - B_{t_{i-1}}))]^{1/2} = [\sum_{i=1}^{n} c_i^2 (t_i - t_{i-1})]^{1/2} = ||f||_{L^2[0,T]}$. The last two steps follow from the fact that $(B_{t_i} - B_{t_{i-1}}) \sim N(0, t_i - t_{i-1})$. $\mathcal{E}$. This completes the proof. □

Let us recall a well-known result from functional analysis.

Lemma 4.2. *Let H_1, H_2 be Hilbert spaces and A be a linear subspace of H_1. Then if $I : A \to H_2$ is a linear isometry on A and A is dense in H_1, I has a unique extension to an isometry on H_1*

Here, $L^2([0,T])$, $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P})$ are Hilbert spaces and $\mathcal{E}$ is a linear subspace of $L^2([0,T])$. Also $I_T : \mathcal{E} \to \mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P})$ is a linear isometry on $\mathcal{E}$ and $\mathcal{E}$ is dense in $L^2([0,T])$. Thus by lemma 4.2, I_T has a unique extension to an isometry $I_T : L^2([0,T]) \to \mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P})$. This isometry is the Wiener Integral of $f \in L^2([0,T])$ and is written as $\int_0^T f_s dB_s$.
In this section, we fixed a $T > 0$ and defined only the definite integral $I_T(f) = \int_0^T f_s dB_s$. In the next section, we are going to discuss the indefinite Wiener integral.

4.2 Indefinite integral

From what we have done so far, we could actually take any $f \in L^2([0,\infty))$ and define $I_t(f) = \int_0^t f_s dB_s$, for each $t \geq 0$, but in the general theory of integration, it is usually desirable that an indefinite integral $\int_0^t f(s)ds$ should be continuous in t. So, in order to make sure that the integral defined is continuous, we first aim to define the integral for each $t \in [0,T]$ as follows,

$$I_t(f) = \int_0^t f(s)dB_s \text{ for each } t \in [0,T]$$

Using the linear isometry I_T defined on $L^2[0,T]$, we can write the indefinite integral as,

$$I_t(f) = \int_0^t f(s)dB_s = \int_0^T f1_{(0,t]}(s)dB_s = I_T(f1_{(0,t]}), \text{ for each } t \in [0,T].$$

Given a function $f \in L^2[0,T]$, is it possible to choose , for every $t \in [0,T]$, a version of $I_T(f1_{(0,t]})$, say, Y_t, such that, $t \to Y_t(\omega)$ is a continuous function for almost every ω?
Let us consider a function $f \in \mathcal{E}$. Thus $f = \sum_{i=1}^n c_i 1_{(t_{i-1},t_i]}$, for $n \geq 1$ and $0 \leq t_0 < t_1 < \cdots < t_n \leq T$
Then, for $t \in (t_{i-1}, t_i]$,

$$f1_{(0,t]} = c_1 1_{(t_0,t_1]} + c_2 1_{(t_1,t_2]} + \cdots + c_{i-1} 1_{(t_{i-2},t_{i-1}]} + c_i 1_{(t_{i-1},t]} + 0 + 0 + \cdots = \sum_{i=1}^n c_i 1_{(t_{i-1}\wedge t, t_i \wedge t]}$$

Hence, for $t \in (t_{i-1}, t_i]$,

$$I_t(f) = \int_0^t f(s)dB_s = I_T(f1_{(0,t]}) = I_T(\sum_{i=1}^n c_i 1_{(t_{i-1}\wedge t, t_i\wedge t]}) = \sum_{i=1}^n c_i (B_{t_i \wedge t} - B_{t_{i-1}\wedge t}) = c_1(B_{t_1} - B_{t_0}) + c_2(B_{t_2} - B_{t_1}) + \cdots + c_i(B_t - B_{t_{i-1}}) + c_{i+1}(B_t - B_t) + \cdots = c_1(B_{t_1} - B_{t_0}) + c_2(B_{t_2} - B_{t_1}) + \cdots + c_i(B_t - B_{t_{i-1}}) + 0 + 0 + \cdots = c_1(B_{t_1} - B_{t_0}) + c_2(B_{t_2} - B_{t_1}) + \cdots + c_i(B_t - B_{t_{i-1}})$$

Thus, for $f \in \mathcal{E}, I_t(f) = c_1(B_{t_1} - B_{t_0}) + c_2(B_{t_2} - B_{t_1}) + \cdots + c_i(B_t - B_{t_{i-1}})\ if\ t_{i-1} < t \leq t_i$
Since, brownian path is continuous, i.e., $t \to B_t(\omega)$ is continuous in t for almost every ω, we have, for $f \in \mathcal{E}, t \to I_t(f)(\omega)$ is continuous for almost every ω.
Now, when we extend this definition of integral from $f \in \mathcal{E}$ to $f \in L^2[0,T]$ via isometry, using lemma 2, can we retain this continuity?

Theorem 4.4. *For $f \in \mathcal{E}$,say, $f = \sum_{i=1}^n c_i 1_{(t_{i-1},t_i]}, n \geq 1$ and $0 \leq t_0 < t_1 < \cdots < t_n \leq T, I_t(f) = \int_0^t f(s)dB_s = I_T(f1_{(0,t]}) = \sum_{i=1}^n c_i(B_{t_i\wedge t} - B_{t_{i-1}\wedge t})$, then,*

(a) For every $\omega \in \Omega, t \to \int_0^t f_s dB_s$ is continuous on [0,T].

(b) For each $t \in [0,T], \int_0^t f_s dB_s$ is a square integrable random variable.

(c) For each $t \in [0,T], \int_0^t f_s dB_s$ is $\mathcal{A}_t$-measurable.

(d) $\{I_t(f), t \in [0,T]\}$ is a martingale.

Proof. (a) For $f \in \mathcal{E}$, we have $\int_0^t f(s)dB_s = I_T(f1_{(0,t]}) = \sum_{i=1}^n c_i(B_{t_i\wedge t} - B_{t_{i-1}\wedge t})$. Since brownian paths are continuous on [0,T] for every $\omega \in \Omega, t_i \wedge t \to B_{t_i \wedge t}(\omega)\ and\ t_{i-1} \to B_{t_{i-1}\wedge t}(\omega)$ are continuous in t for every $\omega \in \Omega$. Thus, for every $\omega \in \Omega, t \to \int_0^t f_s dB_s(\omega) = \int_0^t f_s dB_s$ is continuous on [0,T]
(b) For each $t \in [0,T], [\int_\Omega (\int_0^t f_s dB_s)^2 dP]^{1/2} = [E((\int_0^t f_s dB_s)^2)]^{1/2} = [E((\sum_{i=1}^n c_i(B_{t_i\wedge t} - B_{t_{i-1}\wedge t}))^2)]^{1/2} = [E(\sum_{i=1}^n c_i^2(B_{t_i\wedge t} - B_{t_{i-1}\wedge t})^2) + 2E(\sum_{i<j} c_i c_j (B_{t_i\wedge t} - B_{t_{i-1}\wedge t})(B_{t_j\wedge t} - B_{t_{j-1}\wedge t})) + 2E(\sum_{i>j} c_i c_j (B_{t_i\wedge t} - B_{t_{i-1}\wedge t})(B_{t_j\wedge t} - B_{t_{j-1}\wedge t}))]^{1/2}$
Here, $E(B_{t_i\wedge t} - B_{t_{i-1}\wedge t}) = E(B_{t_j\wedge t} - B_{t_{j-1}\wedge t}) = 0$, since for $s \geq 0, t > 0, B_{s+t} - B_s \sim N(0,t)$

Hence, $[\int_\Omega(\int_0^t f_s dB_s)^2 dP]^{1/2} = [E(\sum_{i=1}^n c_i^2(B_{t_i\wedge t} - B_{t_{i-1}\wedge t})^2)]^{1/2} = [\sum_{i=1}^n c_i^2 E((B_{t_i\wedge t} - B_{t_{i-1}\wedge t})^2)]^{1/2} = [\sum_{i=1}^n c_i^2 Var((B_{t_i\wedge t} - B_{t_{i-1}\wedge t}))]^{1/2} = [\sum_{i=1}^n c_i^2(t_i\wedge t - t_{i-1}\wedge t)]^{1/2} = ||f1_{(0,t]}||_{L^2[0,T]} < \infty$, since $f1_{(0,t]} \in L^2[0,T]$. The last two steps follow from the fact that $(B_{t_i\wedge t} - B_{t_{i-1}\wedge t}) \sim N(0, t_i\wedge t - t_{i-1}\wedge t)$.
Hence proved.
(c) We have, $\int_0^t f(s)dB_s = \sum_{i=1}^n c_i(B_{t_i\wedge t} - B_{t_{i-1}\wedge t})$. Now, $t_i \wedge t \le t$ and $t_{i-1}\wedge t \le t$, hence $\mathcal{A}_{t_i\wedge t} \subseteq \mathcal{A}_t$ and $\mathcal{A}_{t_{i-1}\wedge t} \subseteq \mathcal{A}_t$. Also $B_{t_i\wedge t}$ is $\mathcal{A}_{t_i\wedge t}$-measurable and $B_{t_{i-1}\wedge t}$ is $\mathcal{A}_{t_{i-1}\wedge t}$-measurable. Thus both $B_{t_i\wedge t}$ and $B_{t_{i-1}\wedge t}$ are $\mathcal{A}_t$-measurable. Hence , by definition,$\int_0^t f_s dB_s$ is $\mathcal{A}_t$-measurable
(d) Fix any $t_0 \in [0,T]$. Consider $\{B_{t\wedge t_0}, t\in[0,T]\}$ Let s¡t. Then, three cases are possible,
Case 1: For $0 \le s < t \le t_0 \le T, E(B_{t\wedge t_0}|\mathcal{A}_s) = E(B_t|\mathcal{A}_s) = B_s = B_{(s\wedge t_0)}$
Case 2: For $0 \le s \le t_0 \le t \le T, E(B_{t\wedge t_0}|\mathcal{A}_s) = E(B_{t_0}|\mathcal{A}_s) = B_s = B_{(s\wedge t_0)}$.
Case 3: For $0 \le t_0 \le s < t \le T, E(B_{t\wedge t_0}|\mathcal{A}_s) = E(B_{t_0}|\mathcal{A}_s) = B_{t_0} = B_{(s\wedge t_0)}$.
Hence, in all the three cases, $E(B_{t\wedge t_0}|\mathcal{A}_s) = B_{(s\wedge t_0)}$. Since $t_0 \in [0,T]$ is arbitrary, we have $E(B_{t_i\wedge t}|\mathcal{A}_s) = B_{t_i\wedge s}$ and $E(B_{t_{i-1}\wedge t}|\mathcal{A}_s) = B_{t_{i-1}\wedge s}$. Thus , from the definition of $I_t(f)$, we have, $E(I_t(f)|\mathcal{A}_s) = E(I_s(f))$.
Hence , from (b) , (c) and $E(I_t(f)|\mathcal{A}_s) = E(I_s(f))$, we can say that, $\{I_t(f), t\in[0,T]\}$ is a martingale.
□

We conclude that, for $f \in \mathcal{E}$, the process $\{I_t(f) = \int_0^t f_s dB_s : 0 \le t \le T\}$ is an $\mathcal{A}_t$-adapted, square integrable martingale with $t \to I_t(f)(\omega)$ continuous on [0,T] for every $\omega \in \Omega$.

Theorem 4.5. *For every $f \in L^2[0,T]$, there exists a process $\{Y_t, 0 \le t \le T\}$, such that $\forall\, t \in [0,T]$,*

(a) $Y_t = \int_0^t f_s dB_s$
(b) Y_t *is $\mathcal{A}_t$-measurable for each $t \in [0,T]$.*
(c) $\exists$ *a* $\mathbf{P}$*-null set N, such that , $\forall\, \omega \notin N, t \to Y_t(\omega)$ is continuous on [0,T].*

Further, such a process is ***unique*** *upto a* $\mathbf{P}$*-null set , i.e., if $\{Z_t, 0 \le t \le T\}$ is any process satisfying (a), (b) and (c), then $\exists$ a* $\mathbf{P}$*-null set set N such that , for $\omega \notin N, Z_t(\omega) = Y_t(\omega)$ for all t.*

Proof. (Existence) Let us consider a function $f \in L^2[0,T]$. Choose a sequence $f_n \in \mathcal{E}$, such that, $f_n \to f$ in $L^2[0,T]$ and $||f_{n+1} - f_n||_{L^2[0,T]} \le 4^{-n}$ for all n. Let $Y_n(t) = I_t(f_n)$. Let $A_n = \{\omega : \sup_{0\le t\le T}|I_t(f_{n+1})(\omega) - I_t(f_n)(\omega)| > 4^{-n}\}$. Then , by Chebyshev's inequality, since $I_t(f)$ is square integrable, $\mathbf{P}(A_n) \le \frac{1}{4^{-n}}E((\sup_{0\le t\le T}|I_t(f_{n+1}) - I_t(f_n)|)^2) \le 4.\frac{1}{4^{-n}}E(|I_T(f_{n+1} - I_T(f_n)|^2) = 4.\frac{1}{4^{-n}}||f_{n+1} - f_n||^2_{L^2[0,T]} \le 4.\frac{1}{4^{-n}}.(4^{-n})^2 = \frac{1}{4^{(n-1)}} < \infty\, \forall\, n$.
Hence, by Borel-Cantelli lemma, which says that, $\sum_n \mathbf{P}(A_n) < \infty$, implies, $\mathbf{P}(\cap_n \cup_{k\ge n} A_k) = 0$, we get, $\mathbf{P}(\cap_n \cup_{k\ge n}\{\omega : \sup_{0\le t\le T}|I_t(f_{k+1}) - I_t(f_k)| > 4^{-k}\}) = 0$.
Let $\cap_n \cup_{k\ge n}\{\omega : \sup_{0\le t\le T}|I_t(f_{k+1})(\omega) - I_t(f_k)(\omega)| > 4^{-k}\}$=N, say, that is, $\mathbf{P}$(N)=0.
Now, $\omega \notin N \implies \omega \in \cup_n \cap_{k\ge n}\{\omega : \sup_{0\le t\le T}|I_t(f_{k+1})(\omega) - I_t(f_k)(\omega)| \le 4^{-k}\} \implies \exists\, n$, such that, for all, $k \ge n, \sup_{0\le t\le T}|I_t(f_{k+1})(\omega) - I_t(f_k)(\omega)| \le 4^{-k}$. Hence, $\sum_{k\ge n}|I_t(f_{k+1})(\omega) - I_t(f_k)(\omega)| \le \sum_{k\ge n} 4^{-k} = 4^{-n}.\frac{4}{3}$. Thus, $\{I_t(f_n)(\omega)\}_n$ is uniformly Cauchy on [0,T].
This implies that, $\exists$ a $\mathbf{P}$-null set N, such that, for $\omega \notin N, \{I_t(f_n)(\omega)\}_n$ is uniformly convergent on [0.T].
Hence, for $\omega \notin N, \lim_{n\to\infty} I_t(f_n)(\omega)$ is continous on [0.T].
Since $f_n \to f$ in $L^2[0,T], f_n 1_{(0.t]} \to f1_{(0,t]}$ in $L^2[0,T]$. This implies that, $I_t(f_n) \to I_t(f)$ in $\mathcal{L}^2_T$. Hence, $I_t(f)$ is continuous in [0,T] for $\omega \notin N$. This proves the existence of such a process satisfying (a), (b), (c).
(Uniqueness) Suppose $\{Y_t, t\in[0,T]$ and $\{Z_t, t\in[0,T]\}$ both satisfy (a), (b) and (c).
Then , for any $t \in [0,T], Y_t = \int_0^t f_s dB_s = Z_t$, which implies that, $\exists$ a $\mathbf{P}$-null set such that, for $\omega \notin N_t, Y_t(\omega) = Z_t(\omega)$.
Take $N' = \bigcup_{r\in\mathbf{Q}\cap[0,T]} N_r$, where each N_r is a $\mathbf{P}$-null set. Then $\mathbf{P}(N') \le \sum_{r\in\mathbf{Q}\cap[0,T]}\mathbf{P}(N_r) = 0$.
Hence, N' is a $\mathbf{P}$-null set. Thus $\omega \notin N' \implies \omega \notin N_r$, for any rational $r \in [0,T] \implies Y_r(\omega) = Z_R(\omega)$ for any rational $r \in [0,T]$. Hence, by (c), $Y_t(\omega) = Z_t(\omega)\,\forall\, t \in [0,T]$. This proves the uniqueness of such a process.
□

This process $\{Y_t = I_t(f), t\in[0,T]\}$ is the Wiener Integral of $f \in L^2[0,T]$, $I_t(f) = \int_0^t f_s\, dB_s$ for each $t \in [0,T]$.
Now, since we have proved the continuity of the indefinite integral for functions in $L^2[0,T]$, we can extend the definition of Wiener integral for functions in $L^2[0,\infty)$.
Thus theorem 4.5 can be restated for $f \in L^2[0,\infty)$, that is , the process $\{Y_t = I_t(f), t\in[0,\infty)\}$ is the Wiener Integral of $f \in L^2[0,\infty)$, $I_t(f) = \int_0^t f_s\, dB_s$ for each $t \in [0,\infty)$

Definition 4.1. *(Wiener Integral) Let $\{B_t, t \geq 0\}$ be a standard brownian motion (SBM) on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$. For any $f \in L^2[0,\infty)$, the linear isometry $I_t : L^2[0,\infty) \to \mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P})$, defined as,*

$$I_t(f) = \int_0^t f_s dB_s \text{ for any } t \in [0,\infty)$$

such that,the process $\{I_t(f), t \in [0,\infty)\}$ is $\mathcal{A}_t$-measurable , $t \to I_t(f)(\omega)$ is continuous in $[0,\infty)$ for almost every $\omega \in \Omega$ and it is unique upto a $\mathbf{P}$-null set, is known as the Wiener Integral of the function $f \in L^2[0,\infty)$ with respect to the SBM for any $t \in [0,\infty)$.

Chapter 5

Ito Integral

Let $\{B_t, t \geq 0\}$ be a standard brownian motion (SBM) on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$
Let us consider the L^2 space of square intgrable random variables on $(\Omega, \mathcal{A}, \mathbf{P})$, defined as,

$$\mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P}) = \{X : X \text{ measurable on } \Omega, \int_\Omega X^2 dP < \infty\}$$
$$= \{X \text{ is a random variable }, \int_\Omega X^2 dP < \infty\}$$

Unlike Wiener integral which are only defined for deterministic integrands, Ito integral aims to define it for stochastic integrands..
Fix $T \geq 0$. Let us consider the class of integrands $\mathcal{L}^2_T$, defined as,

$$\mathcal{L}^2_T = \{f(s, \omega) : [0, T] \times \Omega \to \mathcal{R}, \text{ such that,}$$

(i) f is jointly measurable in (s, ω), that is, $f : [0, T] \times \Omega \to \mathcal{R}$ is measurable w.r.t. $\mathcal{B}[0, T] \otimes \mathcal{A}$.
(ii) f is adapted,that is, for each $t \in [0, T], \omega \to f(t, \omega)$ is $\mathcal{A}_t$-measurable.
(iii) $||f||^2_{\mathcal{L}^2_T} = \int_\Omega \int_0^T |f(s, \omega)|^2 ds d\mathbf{P} < \infty\}$.

The usual L^2 space on $([0, T] \times \Omega)$ is defined as,

$$L^2([0, T] \times \Omega, \mathcal{B}[0, T] \otimes \mathcal{A}, \lambda \otimes \mathbf{P}) = \{f(s, \omega) : [0, T] \times \Omega \to \mathcal{R}, \text{ such that,}$$

(i) f is jointly measurable in (s, ω), that is, $f : [0, T] \times \Omega \to \mathcal{R}$ is measurable w.r.t. $\mathcal{B}[0, T] \otimes \mathcal{A}$.
(ii) $||f||^2_{\mathcal{L}^2_T} = \int_\Omega \int_0^T |f(s, \omega)|^2 ds d\mathbf{P} < \infty\}$.

From the definition of both the spaces, $\mathcal{L}^2_T$ and $L^2([0, T] \times \Omega)$, it can be clearly seen that, they differ by only one condition.
Clearly $\mathcal{L}^2_T \subseteq L^2([0, T] \times \Omega)$. In fact $\mathcal{L}^2_T$ is a closed subspace of $L^2([0, T] \times \Omega)$, will be proved later.

Lemma 5.1. *Let $(\Omega, \mathcal{A}, \mu)$ be a finite measure space. Fix $p \geq 1$. If $\{f_n\}$ is a sequence in $L^p(\Omega, \mathcal{A}, \mu)$, such that, $f_n \to f$ in L^p, then $\exists$ a subsequence $\{f_{n_k}\}$, such that, $f_{n_k} \to f$ $\mu - a.e.$*

Proof. For any $\lambda > 0$,

$$\mu(|f_n - f| > \lambda) \leq \frac{1}{\lambda^p} \int |f_n - f|^p d\mu, \text{ by Chebyshev's inequality}$$

Thus, $\mu(|f_n - f| > \lambda) = \mu(|f_n - f|^p > \lambda^p) = \int 1_(|f_n - f|^p > \lambda^p) d\mu \leq \frac{1}{\lambda^p} \int |f_n - f|^p d\mu$.
Since, by hypothesis, $\int |f_n - f|^p d\mu \to 0, \exists$ a subsequence $1 \leq n_1 < n_2 < n_3 < \cdots < n_k < \cdots \uparrow \infty$, such that, $\forall\, n \geq n_k$

$$\mu(|f_n - f| > 2^{-k}) \leq 2^{kp} \int |f_n - f|^p d\mu < 2^{kp}.4^{-kp}, \text{ since } \int |f_n - f|^p d\mu \to 0$$

Thus, $\mu(|f_n - f| > 2^{-k}) < 2^{-kp}, \ \forall\, n \geq n_k$. We want to show, $f_{n_k} \to f$ $\mu - a.e.$
Now, the above inequality will imply,

$$\sum_k \mu(|f_{n_k} - f| > 2^{-k}) < \sum_k 2^{-kp}$$

Hence,

$$\sum_{k \geq K} \mu(|f_{n_k} - f| > 2^{-k}) = \mu(\cup_{k \geq K} \{|f_{n_k} - f| > 2^{-k}\}) \to 0 \text{ as } K \uparrow \infty$$

Therefore, $\mu(\cap_{k\uparrow\infty}\cup_{k\geq K}\{|f_{n_k}-f|>2^{-k}\})=0$.
Now, if $\omega\notin\cap_{k\uparrow\infty}\cup_{k\geq K}\{|f_{n_k}-f|(\omega)>2^{-k}\}$, $\exists K$, such that, $\forall\, k\geq K, |f_{n_k}(\omega)-f(\omega)|\leq 2^{-k}$,which implies, $f_{n_k}(\omega)\to f(\omega)$ for a.e. ω, that is, $f_{n_k}\to f$, $\mu-a.e.$ Hence proved. □

Theorem 5.1. *$\mathcal{L}^2_T$ is a closed subspace of $L^2([0,T]\times\Omega,\mathcal{B}[0,T]\otimes\mathcal{A},\lambda\otimes\mathbf{P})$.*

Proof. Suppose $f_n\in\mathcal{L}^2_T$ and $||f_n-f||\to 0$. We need to show, $f\in\mathcal{L}^2_T$. Clearly, f satisfies condition (i) and (iii) for being in the space $\mathcal{L}^2_T$, as $f\in L^2([0,T]\times\Omega)$. So, it is enough to show that f satisfies condition (ii), i.e.,for each $t\in[0,T],\omega\to f(t,\omega)$ is $\mathcal{A}_t$-measurable.
Now, by hypothesis, $f_n\to f$ in L^2 , hence, by lemma 3,

$$\exists \text{ a subsequence } \{n_k\}, \text{ such that, } f_{n_k}(s,\omega)\to f(s,\omega) \text{ for a.e. } (s,\omega)$$

Let A=$\{(s,\omega):f(s,\omega)\neq\lim_k f_{n_k}(s,\omega)\}$. So, $\lambda\otimes\mathbf{P}(A)=0$ and,

$$\lambda\otimes\mathbf{P}(A)=\int\mathbf{P}(A_t)dt, \text{ where } A_t=\{\omega:f(t,\omega)\neq\lim_k f_{n_k}(t,\omega)\}$$

Hence ,$\int\mathbf{P}(A_t)dt=0$, that is, for a.e. $t\in[0,T],\mathbf{P}(A_t)=0$, that is $\mathbf{P}(\{\omega:f(t,\omega)\neq\lim_k f_{n_k}(t,\omega))\}=0$
Thus, for a.e. $t\in[0,T], f(t,\omega)=\lim_k f_{n_k}(t,\omega)$, $\mathbf{P}-a.e.\ \cdots\ (*)$
Hence, for a.e. t, $f(t,\cdot)$ is $\mathcal{A}_t$-measurable. Setting $f(t,\omega)\equiv 0$ for the set of t for which (*) fails, we get, for each $t\in[0,T],\omega\to f(t,\omega)$ is $\mathcal{A}_t$-measurable. Hence proved. □

Theorem 5.2. *(Lusin's theorem) Let $(\mathbb{R},\mathcal{B},\lambda)$ be a measure space. Let $L^2(\mathbb{R},\mathcal{B},\lambda)=\{f:\int_{\mathbb{R}}f^2d\lambda<\infty\}$ be a space of functions. Then the set of continuous functions on $\mathbb{R}$ with compact support is dense in $L^2(\mathbb{R},\mathcal{B},\lambda)$*

Proof. Let $f\in L^2(\mathbb{R},\mathcal{B},\lambda)$ and $\epsilon>0$ be given. We need to find a continuous function $g:\mathbb{R}\to\mathbb{R}$ with compact support converging to f. Now , for such a f,we can always find a sequence of simple functions in $L^2(\mathbb{R},\mathcal{B},\lambda)$ converging to f. Let us consider a simple function, $\sum_{i=1}^n c_i 1_{B_i}$ with $\lambda(B_i)<\infty\,\forall\, 1\leq i\leq n$. Then,

$$||f-\textstyle\sum_{i=1}^n c_i 1_{B_i}||_2<\epsilon/3\ \cdots\ (*)$$

Further, we can get a function $\sum_j d_j 1_{(a_j,a_{j+1}]}$ such that,

$$||\textstyle\sum_i c_i 1_{B_i}-\sum d_j 1_{(a_j,a_{j+1}]}||_2<\epsilon/3\ \cdots\ (**)$$

Now, to complete the proof , we aim to get a continuous function $g:\mathbb{R}\to\mathbb{R}$ with compact support , such that

$$||\textstyle\sum d_j 1_{(a_j,a_{j+1}]}-g||_2<\epsilon/3\ \cdots\ (***),$$

so that by combining (*),(**),(***), we get the required result that $||f-g||_2<\epsilon$.
We claim to find, for each j, a continuous function g_j on $\mathbb{R}$ with compact support , such that,

$$||1_{(a_j,a_{j+1}]}-g_j||_2<\delta,\ for\ some\ \delta>0\ \cdots\ (****)$$

Let us define $g_j:\mathbb{R}\to\mathbb{R}$, continuous with compact support in the compact interval $[a_j,a_{j+1}]$. Thus (****) will hold for such g_j, for each $j, 1\leq j\leq n$. Take $g=\sum_j d_j g_j$. Then,

$$||\textstyle\sum d_j 1_{(a_j,a_{j+1}]}-g||_2=||\sum d_j 1_{(a_j,a_{j+1}]}-\sum_j d_j g_j||_2\leq\sum|d_j|||1_{(a_j,a_{j+1}]}-g_j||_2<\delta\sum|d_j|, \text{ by (****)}$$

Now, (***) will hold, if $\delta\sum|d_j|<\epsilon/3$, that is, if $\delta<\frac{\epsilon}{3\sum|d_j|}$, which is >0.
Thus we have proved the result for an arbitrary continuous function g on $\mathbb{R}$ with compact support. Hence it holds for the set of all such functions. Hence proved. □

Lemma 5.2. *Consider $L^p(\mathbb{R},\mathcal{B},\lambda)$ for any $p\geq 1$. For any real number x and f^p, define*

$$T_xf:\mathbb{R}\to\mathbb{R}\ \ by\ T_xf(y)=f(x+y),y\in\mathbb{R}$$

Then,
(a) $f\in L^p\implies T_xf\in L^p$ and $||T_xf||_p=||f||_p\,\forall\, x\in\mathbb{R},\,\forall\, f\in L^p$
(b) $x\in\mathbb{R},x_n\to x\implies T_{x_n}f\to T_xf\,\forall\, f\in L^p$

Proof. (a) Let $f\in L^p(\mathbb{R},\mathcal{B},\lambda)$ and x $\in\mathbb{R}$, then $(T_xf)^{-1}(B)=\{y:T_xf(y)\in B\}=\{y:f(x+y)\}=f^{-1}(B)-x\in\mathcal{B}\ \forall\, B\in\mathcal{B}$, Thus T_xf is measurable. Now, fix $x\in\mathbb{R}$, then , $\theta_x:\mathbb{R}\to\mathbb{R}$, defined as ,

$$\theta_x(y)=x+y \text{ is a measurable map.}$$

Now, $\theta_x:(\mathbb{R},\mathcal{B},\lambda)\to(\mathbb{R},\mathcal{B},\mu)$, where μ is the induced measure,

$$\mu(B)=\lambda(\theta_x^{-1}(B)), \text{ for , } B\in\mathcal{B}$$

That is, $\mu(B) = \lambda(\{y : x + y\}) = \lambda(B - x) = \lambda$. Thus, by change of variable theorem, we have,

$$||T_x f||_p = \int |T_x f|^p d\lambda = \int_{\mathbb{R}} |f \circ \theta_x|^p d\lambda = \int_{\mathbb{R}} |f|^p d\lambda \theta_x^{-1} = \int_{\mathbb{R}} |f|^p d\lambda = ||f||_p$$

This completes the proof

(b) Given a f$\in L^p(\mathbb{R}, \mathcal{B}, \lambda)$, by Lusin's theorem, we get a continuous function $g : \mathbb{R} \to \mathbb{R}$ with compact support [-K,K], such that, $||f - g||_p < \epsilon$.

Then ,

$$||T_{x_n} f - T_x f||_p \leq ||T_{x_n} f - T_{x_n} g||_p + ||T_x f - T_x g||_p + ||T_{x_n} g - T_x g||_p$$

Thus, $||T_{x_n} f - T_x f||_p = 2||f - g||_p + ||T_{x_n} g - T_x g||_p < 2\epsilon + ||T_{x_n} g - T_x g||_p$, by (a) and since $||f - g||_p < \epsilon$

To complete the proof, we need to show,

$$||T_{x_n} g - T_x g||_p \to 0$$

Since g is continuous, $T_{x_n} g(y) = g(y + x_n) \to g(y + x) = T_x g(y)$, for all y. Again, g has compact support [-K,K], so $g(y) = 0$ if $|y| > K$. Now, since $x_n \to x, \exists M > 0$, such that, $|x_n|, |x| \leq M \,\forall\, n$. Thus,

$$T_x g(y) = 0, if\ |y + x| > K \text{ and } T_{x_n} g(y) = 0, if\ |y + x_n| > K$$

Hence , $T_{x_n} g$ and $T_x g$ vanishes outside [-(K+M),(K+M)], that is, $T_{x_n} g$ and $T_x g$ are bounded on [-(K+M),(K+M)] and zero outside it. Thus,

$$|T_{x_n} g| \leq C.1_{[-(K+M),(K+M)]} \in L^p \text{ , measurable}$$

Since, both $T_{x_n} g$ and $T_x g$ are measurable, proved in part (a), by DCT,

$$||T_{x_n} g - T_x g||_p \to 0$$

This completes the proof. □

Let us define a class of functions as,

$$\mathcal{E}_T = \{f : ([0, T] \times \Omega) \to \mathbb{R} : f(s, \omega) = \textstyle\sum_{i=1}^n \xi_i(\omega) 1_{(t_{i-1}, t_i]}(s), where\ 0 \leq t_0 < t_1 < \cdots < t_n \leq T\ and\ \xi_1, \xi_2, \ldots, \xi_n \text{ are bounded random variables with } \xi_i \text{ measurable w.r.t. } \mathcal{A}_{t_{i-1}}\}$$

5.1 Definite Integral

We aim to define the integral of $\mathcal{L}_T^2$ functions with respect to the standard brownian motion, that is to define

$$\int_0^T f_s dB_s, \text{ for } f \in \mathcal{L}_T^2$$

We denote this integral by $I_T(f)$, i.e.,

$$I_T(f) = \int_0^T f_s dB_s, \text{ for } f \in \mathcal{L}_T^2$$

Theorem 5.3. *$\mathcal{E}_T$ is a vector subspace of $\mathcal{L}_T^2$.*

Proof. Let us consider a function $f \in \mathcal{E}_T$. Then, $f = \sum_{i=1}^n \xi_i(\omega) 1_{(t_{i-1}, t_i]}(s)$, Now ,

(i) for each i, $\xi_i(\omega) 1_{(t_{i-1}, t_i]}(s)$ is jointly measurable in (s, ω), and hence so is their finite sum.

(ii) for $t \in (t_{i-1}, t_i]$, $f(t, \omega) = \xi_i(\omega)$, which is $\mathcal{A}_{t_{i-1}}$-measurable and $\mathcal{A}_{t_{i-1}} \subset \mathcal{A}_t, i.e.\ , \omega \to f(t, \omega)$ is $\mathcal{A}_t$ measurable for each t.

(iii) $\int_\Omega \int_0^T |f(s, \omega)|^2 ds d\mathbf{P} = \sum_{i=1}^n E(|\xi_i|^2)(t_{i-1} - t_i) < \infty$. Hence $f \in \mathcal{L}_T^2$. Thus $\mathcal{E}_T \subseteq \mathcal{L}_T^2$.

Now, let g be another function in $\mathcal{E}_T$. Then,

$$g = \textstyle\sum_{j=1}^m \eta_j(\omega) 1_{(s_{j-1}, s_j]}(s),$$

for some $m \geq 1$ and $0 \leq s_0 < s_1 < s_2 < ... < s_n \leq T$ and $\eta_1, \ldots, \eta_n$ are bounded random varriables with η_j measurable w.r.t. $\mathcal{A}_{t_j}$.Then for A,B $\in \mathbb{R}$,

$$\begin{aligned} Af + Bg &= A \textstyle\sum_{i=1}^n \xi_i(\omega) 1_{(t_{i-1}, t_i]}(s) + B \sum_{j=1}^m \eta_j(\omega) 1_{(s_{j-1}, s_j]}(s) \\ &= \textstyle\sum_{i=1}^n A\xi_i(\omega) 1_{(t_{i-1}, t_i]}(s) + \sum_{j=1}^m B\eta_j(\omega) 1_{(s_{j-1}, s_j]}(s) \\ &= \textstyle\sum_{i=1}^n \sum_{j=1}^m (A\xi_i(\omega) + B\eta_j(\omega)) 1_{(t_{i-1} \vee s_{j-1}, t_i \wedge s_j]}(s), \end{aligned}$$

where, $(A\xi_i(\omega) + B\eta_j(\omega))$ is measurable measurable w.r.t. $\mathcal{A}_{(t_{i-1} \vee s_{j-1}}$. Thus , $Af + bg \in \mathcal{E}_T$.

Hence $\mathcal{E}_T$ is a vector subspace of $\mathcal{L}_T^2$. □

Theorem 5.4. *$\mathcal{E}_T$ is dense in $\mathcal{L}_T^2$.*

Proof. Let $f \in \mathcal{L}_T^2$. Since any $f \in \mathcal{L}_T^2$ can be approximated (in $\mathcal{L}_T^2$ by a bounded function, we assume , without loss of generality, that f is bounded. Now $f : [0, T] \times \Omega \to \mathbb{R}$. Put $f(s, \omega) = 0$ for $s \notin [0, T]$, $\forall\, \omega$.

Thus, for every ω, $f(\cdot,\omega)$ is a bounded function on $\mathbb{R}$, such that

$$\int_\Omega \int_{\mathbb{R}} f^2(s,\omega) ds d\mathbf{P} < \infty$$

Let us define a function, $\phi_n(v) = \frac{[2^n v]}{2^n}$, $v \in \mathbb{R}$. Here, $[2^n v] = k \Leftrightarrow k \leq 2^n v < k+1 \Leftrightarrow \frac{k}{2^n} \leq v < \frac{k+1}{2^n}$ and thus, $\phi_n(v) = \frac{k}{2^n}$, $v \in [\frac{k}{2^n}, \frac{k+1}{2^n})$. Thus,

$$\phi_n(v) \leq v < \phi_n(v) + \frac{1}{2^n},$$

which implies , $\phi_n(v) \to v$ as $n \to \infty$ $\forall\, v \in \mathbb{R}$. For $\mathbf{P}$-a.e. ω, $f(\cdot,\omega) \in L^2(\mathbb{R})$, implies, for $\mathbf{P}$-a.e. ω

$$\int_{\mathbb{R}} |f(s+\phi_n(v),\omega) - f(s+v,\omega)|^2 ds \to 0 \;\; \forall\, v \in \mathbb{R}$$

Since f bounded, by DCT, we get,

$$\int_{v=-1}^{T+1} \int_\Omega \int_{\mathbb{R}} |f(s+\phi_n(v),\omega) - f(s+v,\omega)|^2 ds \; d\mathbf{P} \; dv \to 0 \text{ as } n \to \infty$$

$$\Rightarrow \int_{\mathbb{R}} \int_{v=-1}^{T+1} \int_\Omega |f(s+\phi_n(v),\omega) - f(s+v,\omega)|^2 d\mathbf{P} \; dv \; ds \to 0 \text{ as } n \to \infty, \; by \; Fubini$$

$$\Rightarrow \int_0^1 \int_{v=-1}^{T+1} \int_\Omega |f(s+\phi_n(v),\omega) - f(s+v,\omega)|^2 d\mathbf{P} \; dv \; ds \to 0 \text{ as } n \to \infty.$$

Now, if $h_n(s) = \int_{v=-1}^{T+1} \int_\Omega |f(s+\phi_n(v),\omega) - f(s+v,\omega)|^2 d\mathbf{P} \; dv$, $s \in [0,1]$, $\int_0^1 |h_n(s)| ds \to 0$, which implies, $\exists$ a subsequence n_k,s.t., $h_{n_k}(s) \to 0$ for a.e. $s \in [0,1]$. Thus , we can pick one $s_0 \in [0,1], say, s_0 = \alpha$,s.t., $h_{n_k}(\alpha) \to 0$, which gives,

$$\int_{v=-1}^{T+1} \int_\Omega |f(\alpha+\phi_{n_k}(v),\omega) - f(\alpha+v,\omega)|^2 d\mathbf{P} \; dv \to 0 \; as \; k \to \infty$$

Put $s = \alpha + v$, $0 \leq \alpha \leq 1$. Then, $\int_{s=\alpha-1}^{\alpha+T+1} \int_\Omega |f(\alpha+\phi_{n_k}(s-\alpha),\omega) - f(s,\omega)|^2 d\mathbf{P} \; ds \to 0 \; as \; k \to \infty$

Note that, $\alpha - 1 \leq 0$ and $\alpha + T + 1 \geq T$,i.e., $[0,T] \subseteq [\alpha-1, \alpha+T+1]$. Hence, we have

$$\int_0^T \int_\Omega |f(\alpha+\phi_{n_k}(s-\alpha),\omega) - f(s,\omega)|^2 d\mathbf{P} \; ds \to 0 \; as \; k \to \infty$$

Let $g_k(s,\omega) = f(\alpha+\phi_{n_k}(s-\alpha),\omega)$, for $(s,\omega) \in [0,T] \times \Omega$. So, now it is enough to show that, $g_k \in \mathcal{E}_T$, to prove the claim. $\square$

Let us define a map I_T on $\mathcal{E}_T$ by,

$$I_T(f) = I_T(\textstyle\sum_{i=1}^n \xi_i 1_{(t_{i-1},t_i]}) = \sum_{i=1}^n \xi_i (B_{t_i} - B_{t_{i-1}}), f \in \mathcal{E}_T$$

For all $f \in \mathcal{E}_T$, $I_T(f)$ is a square integrable random variable, that is, $I_T(f) \in \mathcal{L}^2(\Omega,\mathcal{A},\mathbf{P})$.

Theorem 5.5. *$I_T(\cdot) : \mathcal{E}_T \to \mathcal{L}^2(\Omega,\mathcal{A},\mathbf{P})$ is a linear isometry on $\mathcal{E}_T$*

Proof. Let $f, g \in \mathcal{E}_T$. Then, $\exists\, 0 \leq t_0 < t_1 < \cdots < t_n \leq T$, such that,

$$f = \textstyle\sum_{i=1}^n \xi_i 1_{(t_{i-1},t_i]} \text{ and } g = \sum_{i=1}^n \eta_i 1_{(t_{i-1},t_i]}$$

Then, for $\alpha, \beta \in \mathbb{R}, \alpha f + \beta g = \sum_{i=1}^n (\alpha\xi_i + \beta\eta_i) 1_{(t_{i-1},t_i]}$.
Thus, $I_T(\alpha f + \beta g) = \sum_{i=1}^n (\alpha\xi_i + \beta\eta_i)(B_{t_i} - B_{t_{i-1}}) = \alpha \sum_{i=1}^n \xi_i (B_{t_i} - B_{t_{i-1}}) + \beta \sum_{i=1}^n \eta_i (B_{t_i} - B_{t_{i-1}}) = \alpha I_T(f) + \beta I_T(g)$. Hence $I_T(\cdot)$ is a linear map on $\mathcal{E}_T$ to $\mathcal{L}^2(\Omega,\mathcal{A},\mathbf{P})$.
Next, we want to proof that $I_T(\cdot) : \mathcal{E}_T \to \mathcal{L}^2(\Omega,\mathcal{A},\mathbf{P})$ is an isometry, i.e., to proof ,

$$\int_\Omega \int_0^T (f.g) ds \; d\mathbf{P} = E(I_T(f) I_T(g)) = \int_\Omega I_T(f) I_T(g) d\mathbf{P}$$

Now,$\int_\Omega \int_0^T (f.g) ds \; d\mathbf{P} = \int_\Omega \int_0^T \sum_{i=1}^n \xi_i \eta_i 1_{(t_{i-1}-t_i]} ds \; d\mathbf{P} + \int_\Omega \int_0^T \sum\sum_{i \neq j} \xi_i \eta_j 1_{(t_{i-1},t_i]} 1_{(t_{j-1},t_j]} ds \; d\mathbf{P} = \int_\Omega \int_0^T \sum_{i=1}^n \xi_i \eta_i 1_{(t_{i-1}-t_i]} ds \; d\mathbf{P} + 0$, since $(t_{i-1}, t_i]$ and $(t_{j-1}, t_j]$ are disjoint, so one of the indicator functions will be zero for any i,j.Thus,

$$\int_\Omega \int_0^T \textstyle\sum_{i=1}^n \xi_i \eta_i 1_{(t_{i-1}-t_i]} ds \; d\mathbf{P} = \sum_{i=1}^n E(\xi_i \eta_i)(t_i - t_{i-1}) \; \cdots \; (*)$$

Again, $I_T(f) = \sum_{i=1}^n \xi_i (B_{t_i} - B_{t_{i-1}})$, $I_T(g) = \sum_{i=1}^n \eta_i (B_{t_i} - B_{t_{i-1}})$, so, $I_T(f) I_T(g) = \sum_{i=1}^n \xi_i \eta_i (B_{t_i} - B_{t_{i-1}})^2 + \sum\sum_{i \neq j} \xi_i \eta_j (B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})$. So, for $i \neq j$, say $t_{i-1} < t_{j-1}$,then,

$$\begin{aligned} E(\xi_i \eta_j (B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})) &= E(\xi_i \eta_j (B_{t_i} - B_{t_{i-1}}) E((B_{t_j} - B_{t_{j-1}}) | \mathcal{A}_{t_{j-1}}) \\ &= E(\xi_i \eta_j (B_{t_i} - B_{t_{i-1}}) E(B_{t_j} - B_{t_{j-1}}) = 0 \end{aligned}$$

Since, $(\xi_i \eta_j (B_{t_i} - B_{t_{i-1}}))$ are $A_{t_{j-1}}$-measurable, and , $(B_{t_j} - B_{t_{j-1}})$ is independent of $A_{t_{j-1}}$. Now,
$E(\xi_i \eta_i (B_{t_i} - B_{t_{i-1}})^2) = E(\xi_i \eta_i E((B_{t_i} - B_{t_{i-1}})^2 | \mathcal{A}_{t_{i-1}})) = E(\xi_i \eta_i) E((B_{t_i} - B_{t_{i-1}})^2) = E(\xi_i \eta_i)(t_i - t_{i-1})$
Thus,

$$E(I_T(f) I_T(g)) = \textstyle\sum_{i=1}^n E(\xi_i \eta_i)(t_i - t_{i-1})$$

Hence, comparing with (*) completes the proof. $\square$

Remark 5.1. $\mathcal{L}_T^2$, $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P})$ *are Hilbert spaces and* $\mathcal{E}_T$ *is a linear subspace of* $\mathcal{L}_T^2$*. Also* $I_T : \mathcal{E}_T \to \mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P})$ *is a linear isometry on* $\mathcal{E}_T$ *and* $\mathcal{E}_T$ *is dense in* $\mathcal{L}_T^2$*. Hence, by lemma 4.2,* I_T *has a unique extension to an isometry* $I_T : \mathcal{L}_T^2 \to \mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P})$*.This isometry is the Ito Integral of* $f \in \mathcal{L}_T^2$ *and is written as* $\int_0^T f_s dB_s$*.*

In this section, we fixed a $T > 0$ and defined only the definite integral $I_T(f) = \int_0^T f_s dB_s$. In the next section, we are going to discuss the indefinite Ito integral.

5.2 Indefinite Integral

Let us define the space $\mathcal{L}_\infty^2$ as,

$$\mathcal{L}_\infty^2 = \{f(s,\omega) : [0,\infty) \times \Omega \to \mathbb{R}\}, \text{ such that,}$$

(i) f is jointly measurable in (s, ω), that is, $f : [0,\infty) \times \Omega \to \mathcal{R}$ is measurable w.r.t. $\mathcal{B}[0,\infty) \otimes \mathcal{A}$.
(ii) f is adapted,that is, for each $t \in [0,\infty), \omega \to f(t,\omega)$ is $\mathcal{A}_t$-measurable.
(iii) $||f||^2_{\mathcal{L}_\infty^2} = \int_\Omega \int_0^T |f(s,\omega)|^2 ds\, d\mathbf{P} < \infty$, $\forall\, T > 0\}$.

From what we have done so far, we could actually take any $f \in \mathcal{L}_\infty^2$ and define $I_t(f) = \int_0^t f_s dB_s$, for each $t \geq 0$, but as in the Wiener Integral case, continuity also becomes an issue here. So, in order to make sure that the indefinite integral is continuous, we first aim to define the integral for each $t \in [0,T]$ as follows,

$$I_t(f) = \int_0^t f_s\, dB_s \text{ for each } t \in [0,T]$$

Using the linear isometry I_T defined on $\mathcal{L}_T^2$, we can write the indefinite integral as,

$$I_t(f) = \int_0^t f(s)dB_s = \int_0^T f1_{(0,t]}(s)dB_s = I_T(f1_{(0,t]}), \text{ for each } t \in [0,T].$$

With reference to Theorem 4.5, we state the next result.

Theorem 5.6. *For every* $f \in \mathcal{L}_T^2$*, there exists a process* $\{Y_t, 0 \leq t \leq T\}$*, such that* $\forall\, t \in [0,T]$*,*

(a) $Y_t = \int_0^t f_s dB_s$
(b) Y_t *is* $\mathcal{A}_t$*-measurable for each* $t \in [0,T]$*.*
(c) $\exists$ *a* $\mathbf{P}$*-null set N, such that ,* $\forall\, \omega \notin N, t \to Y_t(\omega)$ *is continuous on [0,T].*

Further, such a process is ***unique*** *upto a* $\mathbf{P}$*-null set , i.e., if* $\{Z_t, 0 \leq t \leq T\}$ *is any process satisfying (a), (b) and (c), then* $\exists$ *a* $\mathbf{P}$*-null set set N such that , for* $\omega \notin N, Z_t(\omega) = Y_t(\omega)$ *for all t.*

This process $\{Y_t = I_t(f), t \in [0,T]\}$ is the Ito Integral of $f \in \mathcal{L}_T^2$, $I_t(f) = \int_0^t f_s\, dB_s$ for each $t \in [0,T]$. Now, since we have proved the continuity of the indefinite integral for functions in $\mathcal{L}_T^2$, we can extend the definition of Ito integral for functions in $\mathcal{L}_\infty^2$. Thus, for $f \in \mathcal{L}_\infty^2$ theorem 5.6 can be restated as,

Theorem 5.7. *For every* $f \in \mathcal{L}_\infty^2$*, there exists a process* $\{Y_t, t \in [0,\infty)\}$*, such that* $\forall\, t \in [0,\infty)$*,*

(a) $Y_t = \int_0^t f_s dB_s = I_t(f)$
(b) Y_t *is* $\mathcal{A}_t$*-measurable for each* $t \in [0,\infty)$*.*
(c) $\exists$ *a* $\mathbf{P}$*-null set N, such that ,* $\forall\, \omega \notin N, t \to Y_t(\omega)$ *is continuous on [0,∞).*

Further, such a process is ***unique*** *upto a* $\mathbf{P}$*-null set , i.e., if* $\{Z_t, t \in [0,\infty)\}$ *is any process satisfying (a), (b) and (c), then* $\exists$ *a* $\mathbf{P}$*-null set set N such that , for* $\omega \notin N, Z_t(\omega) = Y_t(\omega)$ *for all t.*

Theorem 5.8. $\{Y_t, t \in [0,\infty)\}$ *is a square integrable martingale w.r.t.* $(\mathcal{A}_t, t \in [0,\infty))$

Proof. We need to show that, (i) Y_t is $\mathcal{A}_t$-measurable and (ii) for $s < t, E(Y_t|\mathcal{A}_s) = Y_s$.
(i) Suppose $Y_t = \int_0^t f_s dB_s, t \in [0,\infty),$ *where* $f \in \mathcal{L}^2$. Given any $t \geq 0$ and any $s < t$, fix $T > 0$, such that, $s,t \in [0,T]$. Now, $f \in \mathcal{L}^2 \Rightarrow f \in \mathcal{L}_T^2$, and since $\mathcal{E}_T$ is dense in $\mathcal{L}_T^2$,

$$\exists \text{ a sequence } f_n, n \geq 1 \text{ in } \mathcal{E}_T, \text{ such that, } f_n \to f \text{ in } \mathcal{L}_T^2 \ \cdots\ (*)$$

Now, if $f \in \mathcal{E}_T$, it can be shown that, both (i) and (ii) hold.
So, for $f_n \in \mathcal{E}_T$, $I_t(f_n)$ is $\mathcal{A}_t$ measurable and $E(I_t(f_n)|\mathcal{A}_s) = I_s(f_n)$. Now, by theorem 5.7 c, from (*), we have $I_t(f_n) \to I_t(f)$ in $\mathcal{L}^2(\Omega, \mathcal{A}_t, \mathbf{P})$. Hence, for some subsequence n_k,

$$I_t(f_{n_k}) \to I_t(f) \text{ with probability } 1$$

Thus, $I_t(f)$ is $\mathcal{A}_t$-measurable . Hence proved.

(ii) To prove the claim, we need to show, for any $G \in \mathcal{A}_s$

$$\int_G I_t(f)\, d\mathbf{P} = \int_G I_s(f)\, d\mathbf{P}$$

Now, for $f_n \in \mathcal{E}_T$, we know, $\int_G I_t(f_n)\, d\mathbf{P} = \int_s (f_n)\, d\mathbf{P}\ \forall\, n$. Also, $I_t(f_n) \to I_t(f)$ and $I_s(f_n) \to I_s(f)$ in L^2. Now,

$$|\int_G I_t(f_n)\, d\mathbf{P} - \int_G I_t(f)\, d\mathbf{P}| \leq \int |I_t(f_n) - I_t(f)| 1_G\, d\mathbf{P} \leq ||I_t(f_n) - I_t(f)||.||1_G||_2 \to 0 \text{ in } L^2$$

Thus, $\int_G I_t(f_n)\, d\mathbf{P} \to \int_G I_t(f)\, d\mathbf{P}$ in L^2. Similarly, $\int_G I_s(f_n)\, d\mathbf{P} \to \int_G I_s(f)\, d\mathbf{P}$ in L^2. This proves the claim. $\square$

Thus , $Y_t \in \mathcal{L}^2(\Omega, \mathcal{A}_t, \mathbf{P})$, for each $t \in [0, \infty)$. This process $\{Y_t = I_t(f), t \in [0, \infty)\}$ is the Ito Integral of $f \in \mathcal{L}^2_\infty$, $I_t(f) = \int_0^t f_s\, dB_s$ for each $t \in [0, \infty)$

Definition 5.1. *(Ito Integral) Let $\{B_t, t \geq 0\}$ be a standard brownian motion (SBM) on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$. For any $f \in \mathcal{L}^2_\infty$, the linear isometry $I_t : \mathcal{L}^2_\infty \to \mathcal{L}^2(\Omega, \mathcal{A}, \mathbf{P})$, defined as,*

$$I_t(f) = \int_0^t f_s dB_s \textit{ for any } t \in [0, \infty)$$

such that,the process $\{I_t(f), t \in [0, \infty)\}i$ is $\mathcal{A}_t$-measurable , $t \to I_t(f)(\omega)$ is continuous in $[0, \infty)$ for almost every $\omega \in \Omega$ and it is unique upto a $\mathbf{P}$-null set, is known as the Ito Integral of the function $f \in \mathcal{L}^2_\infty$ with respect to the SBM for any $t \in [0, \infty)$.

Using the Doob-Meyer Decomposition, we get the following lemma.

Lemma 5.3. *If $\{M_t, t \in [0, \infty)\}$ is a square integrable martingale w.r.t. $(\mathcal{A}_t, t \in [0, \infty))$ with continuous paths, then $\exists$ a process $\{A_t, t \in [0, \infty)\}$ with continuous and non-decreasing paths , such that,$\{M_t^2 - A_t, t \in [0, \infty)\}$ is a martingale.*

This A_t is the Quadratic variation of Y_t and is written as $< Y >_t$.
Let us consider,

$$A_t(\omega) = \int_0^t f^2(s, \omega) ds, t \in [0, \infty)\ \cdots\ (*)$$

Note that,here,

(a) $t \to A_t(\omega)$ is continuous on [0,∞)]

(b) $t \to A_t(\omega)$ is non-decreasing in t

(c) For each t, A_t is $\mathcal{A}_t$-measurable.

Looking at the lemma 5.3, we give the following proposition.

Proposition 5.1. *Let $Y_t = I_t(f) = \int_0^t f(s) dB_s, t \in [0, \infty) \forall f \in \mathcal{L}_2$,that is, $\{Y_t, t \in [0, \infty)\}$ is a continuous square integrable martingale w.r.t. $(\mathcal{A}_t)$. . Then, $\{Y_t^2 - \int_0^t f_s^2 ds, t \in [0, \infty)\}$ is a martingale w.r.t. $(\mathcal{A}_t)$.*

Thus, the quadratic variation of Y_t , is given by, $< Y >_t = \int_0^t f_s^2 ds, t \in [0, \infty)$.

Chapter 6

Ito's Formula

In ordinary calculus, the usual way to compute the Riemann-Stieltjes integral of a function f with respect to a bounded variation function a is by reversing the process of differentiation, namely, if we know that f is the derivative of a function F,say, so that$I(t) = \int_0^t (\alpha(s))d\alpha = \int_0^t F'(\alpha(s))d\alpha$, then it directly follows that$I(t) = F(\alpha(t))F(\alpha(0))$ by fundamental theorem of integral calculus. But in Ito's theory of integration, the derivative of a stochastic process has not been defined and integration is not defined as a reverse process of differentiation. So the above method of computation cannot be applied verbatim in this setup. Ito's formula, however, provides a way of computing stochastic integral.

Theorem 6.1. *(Ito's formula I) Let $\{B_t : t \in [0,\infty)$ be a SBM and $\phi \in C_2(\mathbb{R})$. Then,*

$$\phi(B_t) - \phi(0) = \int_0^t \phi'(B_s)\, dB_s + \tfrac{1}{2}\int_0^t \phi''(B_s)ds, \ \ \forall\, t \geq 0$$

Here , the term $\frac{1}{2}\int_0^t \phi''(B_s)ds$ is called Ito correction.

Usual calculus stops with the first term. But in the case of Ito Integrals, that is not correct, so the 2nd terms has to be there. That term is called the Ito correction. This is the simplest version of Ito's formula. We will see other versions of Ito's formula and prove one of them.

Lemma 6.1. *Let $Y_t = \int_0^t f_s\, dB_s,\ t \in [0,\infty)$, $f \in \mathcal{L}^2$, for any sequence of partition $\{\pi_n\}$, as $||\pi_n|| \downarrow 0$*

$$\textstyle\sum_i (Y_{t_i} - Y_{t_{i-1}})^2 \to \int_0^t f_s^2\, ds \text{ in } L^2 \text{ almost surely}$$

Theorem 6.2. *(Ito's formula II) Let $f \in \mathcal{L}^2$. Then , $Y_t = \int_0^t f_s dB_s$ is the Ito integral of f with respect to $\{B_t\}$. Then, for $\phi \in C_2(\mathbb{R})$, we have,*

$$\phi(Y_t) - \phi(Y_u) = \int_u^t \phi'(Y_s) f_s dB_s + \tfrac{1}{2}\int_u^t \phi''(Y_s) f_s^2 ds$$

Proof. Let us take a partition π, $u = t_0 < t_1 < \cdots < t_n = t$, then we can write,

$$\phi(Y_t) - \phi(Y_u) = \textstyle\sum_{i=1}^n (\phi(Y_{t_i}) - \phi(Y_{t_{i-1}}))$$

Now, by Taylor expansion, for any ξ_i lies between Y_{t_i} , $Y_{t_{i-1}}$,

$$\textstyle\sum_{i=1}^n (\phi(Y_{t_i}) - \phi(Y_{t_{i-1}})) = \sum_{i=1}^n \phi'(Y_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}}) + \frac{1}{2}\sum_{i=1}^n \phi''(\xi_i)(Y_{t_i} - Y_{t_{i-1}})^2$$
$$= \textstyle\sum_i \phi'(Y_{t_{i-1}}) \int_{t_{i-1}}^{t_i} f_s\, dB_s + \frac{1}{2}\sum_{i=1}^n \phi''(\xi_i)(Y_{t_i} - Y_{t_{i-1}})^2 = T_1 + \frac{1}{2}T_2, \text{ say}$$

For the partition π, let, $T_1 = \int_u^t g_\pi(s) dB_s$, where, $g_\pi(s,\omega) = \sum_i \phi'(Y_{t_{i-1}})\, f_s(\omega)\, 1_{(t_{i-1},t_i]}(s)$.
Let $\phi'(Y_s(\omega)).f_s(\omega) = g(s,\omega)$. Then,

$$\int_\Omega \int_u^t |g_\pi(s,\omega) - g(s,\omega)|^2\, ds\, d\mathbf{P} = \int_\Omega \textstyle\sum_i \int_{t_{i-1}}^{t_i} (\phi'(Y_s) - \phi'(Y_{t_{i-1}}))^2 f_s^2\, ds\, d\mathbf{P} \to 0$$

Thus, $g_\pi(s,\omega) \to g(s,\omega)$ in L^2, hence $T_1 \to \int_u^t \phi'(Y_s) f_s dB_s \ \cdots$ (1)
Now, $T_2 = \sum_{i=1}^n \phi''(\xi_i)(Y_{t_i} - Y_{t_{i-1}})^2 = \sum_i \phi''(Y_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}})^2 + \sum_i (\phi''(\xi_i) - \phi''(Y_{t_{i-1}}))(Y_{t_i} - Y_{t_{i-1}})^2$.
Here, $\sum_i (\phi''(\xi_i) - \phi''(Y_{t_{i-1}}))(Y_{t_i} - Y_{t_{i-1}})^2 \to 0$.
Let $T_3 = \sum_i \phi''(Y_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}})^2 = \sum_i \phi''(Y_{t_{i-1}}) \int_{t_{i-1}}^{t_i} f_s^2\, ds + \sum_i \phi''(Y_{t_{i-1}})((Y_{t_i} - Y_{t_{i-1}})^2 - \int_{t_{i-1}}^{t_i} f_s^2)$.
Let $T_5 = \sum_i \phi''(Y_{t_{i-1}}) \int_{t_{i-1}}^{t_i} f_s^2\, ds \to \int_u^t \phi''(Y_s) f_s^2\, ds$, as $||\pi|| \downarrow 0$. for all ω. $\cdots$ (2)
Let $T_6 = \sum_i \phi''(Y_{t_{i-1}})((Y_{t_i} - Y_{t_{i-1}})^2 - \int_{t_{i-1}}^{t_i} f_s^2) \to 0$ by lemma 6.1. Thus,

$$\phi(Y_t) - \phi(Y_u) = T_1 + \tfrac{1}{2}(T_5 + T_6 + T_4) \text{ all except } T_1 \text{ and } T_5 \text{ goes to 0.}$$

Hence, from (1) and (2), we have,

$$\phi(Y_t) - \phi(Y_u) = \int_u^t \phi'(Y_s) f_s dB_s + \frac{1}{2}\int_u^t \phi''(Y_s) f_s^2 \, ds$$

This completes the proof. □

Theorem 6.3. *(Ito's formula III) Let $f \in \mathcal{L}^2$ and $g \in L^1$. Then $Y_t = \int_0^t f_s dB_s$ and $X_t = \int_0^t g_s ds$. Then for a function $\phi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, such that $\phi(x, y) \in C_{2,1}(\mathbb{R})$, then,*

$$\phi(X_t, Y_t) - \phi(X_0, Y_0) = \int_0^t \frac{\partial \phi}{\partial x}(X_s, Y_s) \, dX_s + \int_0^t \frac{\partial \phi}{\partial y}(X_s, Y_s) \, dY_s + \frac{1}{2}\int_0^t \frac{\partial^2 \phi}{\partial x^2}(X_s, Y_s) f_s^2 \, ds$$

Ito's formula is a very important tool used to give probabilistic proofs of various problems in mathematics.

Chapter 7

Applications

The solutions to many problems of elliptic and parabolic partial differential equations can be represented as expectations of stochastic functionals. Such representations allow one to infer properties of these solutions and, conversely, to determine the distributions of various functionals of stochastic processes by solving related partial differential equation problems.
In this section, we give probabilistic proofs of some problems of PDE theory.

7.1 Dirichlet problem

In this section, we treat the Dirichlet problem of finding a function which is harmonic in a given region and assumes specified boundary values. One can use Brownian motion to characterize those Dirichlet problems for which a solution exists, to construct a solution, and to prove uniqueness. The Laplacian appearing in the Dirichlet problem is the simplest elliptic operator.

7.1.1 Harmonic functions

The connection between Brownian motion and harmonic functions is profound . We take this connection as our first illustration of the interplay between probability theory and analysis.

Definition 7.1. *Let D be a domain in $\mathbb{R}^2$. A function $u : D \to \mathbb{R}$ is called harmonic in D, if $u \in C_2(\mathbb{R})$ and*

$$\triangle u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \text{ in } D$$

We will prove shortly, a harmonic function has the mean-value property. It is this mean-value property which introduces Brownian motion in a natural way into the study of harmonic functions.

Lemma 7.1. *If $A \in \mathbb{R}^2$ is any bounded set, then for any $x \in A$, $\mathbf{P}(x + B_t \in A^c$, for some $t)=1$*

Lemma 7.2. *If A, any bounded set in $\mathbb{R}$, is borel, then $\tau_A = inf\{t > 0 : x + B_t \in A^c\}$ is a stopping time.*

Thus lemma 7.1 can be restated as ,

$$\mathbf{P}_x(\tau_A < \infty) = 1$$

Definition 7.2. *(Mean-value property) A function u on D, is said to have mean value property if, for any $x \in D$ any $r > 0$, such that, $\overline{B}(x, r) \subset D$,*

$$u(x) = \int_{\partial\overline{B}(x,r)} u(y)d\mu(y) = \frac{1}{2\pi}\int_{\partial\overline{B}(x,r)} u(y)dy$$

Theorem 7.1. *If u is harmonic in D, then u has the mean-value property in D.*

Proof. Let $x = (x_1, x_2) \in D$ with $\overline{B}(x, r) \subset D$, for $r > 0$. Let us consider the process $\{x + B_t, t \geq 0\}$. Since u is harmonic in D, $u \in C_2(D)$. Hence , for $\tau_{\overline{B}(x,r)} = \tau$, by Ito's formula , we have

$$u(x+B_{t\wedge\tau}) = u(x)+\int_0^{t\wedge\tau} \frac{\partial u}{\partial x_1}(B_s)\, dB_1(s)+\int_0^{t\wedge\tau} \frac{\partial u}{\partial x_2}(B_s)\, dB_2(s)+\frac{1}{2}\int_0^{t\wedge\tau} \frac{\partial^2 u}{\partial x_1^2}(B_s)\, ds+\frac{1}{2}\int_0^{t\wedge\tau} \frac{\partial^2 u}{\partial x_2^2}(B_s)\, ds$$

Here, $B_t = (B_1, B_2)$, where each B_i are independent SBMs and $x + B_t = (x_1 + B_1, x_2 + B_2)$.
Now, since u is harmonic in D,

$$\triangle u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} = 0$$

Thus, the above equality is reduced to,

$$u(x + B_{t\wedge\tau}) = u(x) + \int_0^{t\wedge\tau} \frac{\partial u}{\partial x_1}(B_s)\, dB_1(s) + \int_0^{t\wedge\tau} \frac{\partial u}{\partial x_2}(B_s)\, dB_2(s)$$

Now, taking expectation on both sides, we get, $E(u(x+B_{t\wedge\tau})) = u(x)$, the other terms vanish since they stochastic integrands, hence martingales starting at 0, so there expectation is 0. Now let $t \to \infty$, then $t \wedge \tau \to \tau$ and thus $u(x+B_{t\wedge\tau}) \to u(x+B_\tau)$. Now $u(x+B_{t\wedge\tau})$ is bounded for all x, by $sup_{y\in\overline{B}(x,r)}|u(y)|$.
Hence, using DCT, we get,

$$E(u(x + B_{t\wedge\tau})) \to E(u(x + B_\tau))$$

Now, $x + B_\tau \in \partial\overline{B}(x,r)$ and is uniformly distributed on the boundary. Since measure on a circular domain is the normalized arc length measure, $E(u(x + B_\tau)) = \int_{\partial B(x,r)} u(y)\frac{1}{2\pi}dy = \frac{1}{2\pi}\int_{\partial B(x,r)} u(y)dy$. Thus, we get, $u(x) = \frac{1}{2\pi}\int_{\partial B(x,r)} u(y)dy$, that is u has mean value property in D. Hence proved. $\square$

Theorem 7.2. *If $u : D(\subset \mathbb{R}) \to R$ and has the mean value property, then $u \in C_2(\mathbb{R})$ and harmonic.*

In our further discussion, we consider D to be a bounded domain.

7.1.2 The Dirichlet Problem

Let D be a bounded domain in $\mathbb{R}^2$ and $f : \overline{D} \to \mathbb{R}$ a continuous function, then the **Dirichlet Problem** is to find a continuous function $u : \overline{D} \to \mathbb{R}$ such that u is harmonic in D and takes on boundary values specified by f, that is , $u \in C_2(D)$ and

$$\triangle u = 0 \text{ in D and } u = f \text{ on } \partial D$$

Such a function , when it exists , will be called a *solution to the Dirichlet problem (D,f).*

The power of the probabilistic method is demonstrated by the fact that we can immediately give a very likely solution to (D,f), stated in the following proposition.

Proposition 7.1. *Let $\{B_t, t \geq 0\}$ be a 2-dimensional SBM. For $x \in \overline{D}$, let $\{B_t^x, t \geq 0\}$ be the process $B_t^x = x + B_t, t \geq 0$, starting at x. Let $\tau_D^x = inf\{t > 0 : x + B_t^x \in D^c\}$. Then $B^x_{\tau_D^x} \in D^c$. Then,*

$$u(x) = E(f(B^x_{\tau_D^x})),\ x \in \overline{D} \text{ is a solution of the Dirichlet problem.}$$

Remark 7.1. *Harmonicity and twice differentiabiality of the function u on D can be shown by proving the mean value property using the Markov property theorem 7.2.*

The issue arises with proving continuity of u in $\overline{D}$, that is, upto and including ∂D. It turns out that it depends on the *regularity* of ∂D, which can be defined as follows.

Definition 7.3. *(Regularity) Any $x \in \partial D$ is said to be regular if $\mathbf{P}(\tau^x_{D^c} = 0) = 1$.*

In the one-dimensional case every point of ∂D is regular and the Dirichlet problem is always solvable, the solution being piecewise linear. For higher dimension, regularity of any $x \in \partial D$ is sufficient to obtain a probabilistic solution of the Dirichlet problem.

We now look at a probabilistic interpretation of the heat equation.

7.2 Heat equation

Let $f : \mathbb{R} \to \mathbb{R}$ borel measurable. Let $u : (0, \infty) \times \mathbb{R} \to \mathbb{R}$ in $C_2(\mathbb{R})$ .Then, the heat equation is given by,

$$\frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2}\ ,\ u(0, x) = f(x)$$

The physical interpretation here, is to imagine the real line as an insulated rod. Here, f(x) denotes temperature at location x at time 0 and u(t,x) denotes temperature at location x at time t. Then from the usual theory of physics, it can be seen that, u must satisfy the above parabolic PDE. The Laplacian appearing in the heat equation is the simplest parabolic operator.

Let $\{B_t, t \geq 0\}$ be a SBM.Let $x \in \mathbb{R}$, then $x + B_t$ is a brownian motion starting at x. Then, the density of $x + B_t$ in $y \in \mathbb{R}$ is given by,

$$p(t,x,y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}\frac{(y-x)^2}{t}}$$

Fix $y \in \mathbb{R}$, then it can be computed that,

$$\frac{\partial p(t,x,y)}{\partial t} = \frac{1}{2}\frac{\partial^p(t,x,y)}{\partial x^2}$$

Proposition 7.2. *Let $\{B_t, t \geq 0\}$ be a SBM.Let $x \in \mathbb{R}$, then $x + B_t$ is a brownian motion starting at x. Then, the density in $y \in \mathbb{R}$ is given by, $p(t,x,y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}\frac{(y-x)^2}{t}}$. Then,*

$$u(t,x) = \int f(y)p(t,x,y)dy = E(f(x+B_t)$$

gives a solution of the heat equation.

A **sufficient condition** for this solution to exist is that,

$$\int_{\mathbb{R}} e^{-\alpha x^2}|f(x)|dx < \infty \text{ for some } \alpha > 0.$$

Thus, if f is bounded and continuous , $u(t,x) = E(f(x+B_t)$ **is the only bounded solution.**

7.3 Other parabolic PDEs

Some other parabolic PDEs which have probabilistic solution are,

1. For $g : [0,\infty) \times \mathbb{R} \to \mathbb{R}, f : \mathbb{R} \to \mathbb{R}$,the PDE

   $$\frac{\partial u}{\partial t} = \frac{1}{2}\triangle u + g \ , \ u(0,x) = f(x)$$

   has solution, $u(t,x) = E(f(x+B_t)) + E(\int_0^t g(t-s, x+B_s)\ ds)$

2. (Feynman-Kac) For $g : [0,\infty) \times \mathbb{R} \to \mathbb{R}, f : \mathbb{R} \to \mathbb{R}$, $c : \mathbb{R} \to \mathbb{R}$, the PDE

   $$\frac{\partial u}{\partial t} = \frac{1}{2}\triangle u + c.u \ , \ u(0,x) = f(x)$$

   has solution, $u(t,x) = E(f(x+B_t)exp(\int_0^t c(x+B_s)\ ds))$

3. For $g : [0,\infty) \times \mathbb{R} \to \mathbb{R}, f : \mathbb{R} \to \mathbb{R}$, $b : \mathbb{R} \to \mathbb{R}$, the PDE

   $$\frac{\partial u}{\partial t} = \frac{1}{2}\triangle u + b\triangle u \ , \ u(0,x) = f(x)$$

   has solution, $u(t,x) = E(f(x+B_t)exp(\int_0^t b(x+B_s)\ dB_s - \frac{1}{2}\int b^2(x+B_s)\ ds))$