MATH64101 Stochastic Calculus

Zhengbo Zhou*

2022-2023 Academic year Semester 1

*Department of Mathematics, University of Manchester, Manchester, M13 9PL, England (zhengbo.zhou@postgrad.manchester.ac.uk).

Contents

1 Wiener Process (Brownian Motion)

1.1 Basic Definition

Brief history of Brownian motion:

1. It was observed by Robert Brown in 1828.
2. In 1905, Einstein explained it in his studies in heat diffusion.
3. In 1923, Wiener proved the existence of a stochastic process satisfying the Einstein's 1905 postulates.

Definition 1.1. A stochastic process

$$B = (B_t)_{t\geq 0}$$

defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is called a Wiener process (or a Brownian Motion Process) if the following four conditions are satisfied:

1. $t \mapsto B_t$ is continuous from $\mathbb{R}_+$ to $\mathbb{R}$, and $B_0 = 0$ (both $\mathbb{P}$-a.s);
2. B has stationary increments, i.e.

   $$B_t - B_s \sim B_{t-s} \sim N\left((t-s)\mu, (t-s)\sigma^2\right)$$

   for any $0 \leq s < t$.
3. B has independent increments, i.e. $B_{t_1-t_0}, B_{t_2-t_1}, \ldots, B_{t_n-t_{n-1}}$ are independent random variables for any choice of $0 \leq t_0 < t_1 < \cdots < t_n$ with $n \geq 1$.
4. $B_t \sim N\left(\mu t, \sigma^2 t\right)$ for every $t > 0$ where $\mu \in \mathbb{R}$ and $\sigma > 0$ are given and fixed constants.

RECALL

$\mathbb{P}$-a.s.: Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. An event $E \in \mathscr{F}$ happens almost surely if $\mathbb{P}(E) = 1$. Equivalently, E happens almost surely if the probability of not happening E is zero: $\mathbb{P}(E^C) = 0$.
$A \sim B$ implies A and B are equal and have the same distribution, and it is same as $A \sim B$. Namely, they have the same distribution functions.

Remark 1.2.

- It can be proved (using characteristic functions, i.e. Fourier transform) that:

$$(1) + (2) + (3) \longrightarrow (4) \quad \text{with } \sigma \geq 0.$$

- If $\mu = 0$ and $\sigma^2 = 1$ in (4), then B is said to be a *Standard Brownian Motion (SBM)*.
- Note that (2) + (3) is the key hypothesis/postulate introduced and discussed by Einstein (1905) in his derivation of the diffusion equaiton. Also, (3) breaks down for a physical BM if $\Delta t_i = t_i - t_{i-1}$ very small ($\ll 1/\beta \approx 10^{-8}$ sec).
- The value $B_t(\omega)$, with $\omega \in \Omega$ given and fixed, represents the position of a BM in $\mathbb{R}$ at time t.
- The stochastic process satisfying (2) + (3) and have *right-continuous* sample path (trajectories) with *left limit* are called *Lèvy Process*. [The natural question to classify all processes satisfying Einstein's postulate (2) + (3) was initially studies by de Finetti (1929), then Kolmogorov (1932) and finally Lèvy (1934)]
- Thus the BM process B is a Lèvy process, and it can be shown that: B is the only Lèvy process with continuous sample path. (Lèvy process can have jumps, but BM is continuous)

1.2 Existence of BM

Can we construct a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and $B = (B_t)_{t\geq 0}$ such that (1), (2) and (3) exists.

- It can be proved (by Wiener in 1923) that the BM $B = (B_t)_{t\geq 0}$ exist. Namely, there exists a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and a stochastic process $B = (B_t)_{t\geq 0}$ defined on it such that (1)–(4) above hold. (with $\mu \in \mathbb{R}$ and $\sigma > 0$ given and fixed)
- There are several ways to prove the existence
  - Using the "Kolmogorov Consistency Theorem" and the "Kolmogorov Continuity Theorem" (1930).
  - Using that B can be approximated by "random walk" and exploiting "weak convergence" techniques. (Bachelier 1900)
  - Using "Hilbert space theory" and exploiting the "Gaussian property" of B, e.g. if $\xi_0, \xi_1, \ldots,$ are iid as $N(0,1)$, then:

$$B_t = \xi_0 \frac{t}{\sqrt{\pi}} + \frac{2}{\sqrt{\pi}} \sum_{k=1}^{\infty} \frac{\sin(kt)}{k} \xi_k, \quad 0 \leq t \leq \pi$$

    defines a SBM $B = (B_t)_{0\leq t\leq \pi}$. (Wiener 1934)

Note that the general BM $X = (X_t)_{t\geq 0}$ can be expressed in terms of a standard BM $B = (B_t)_{t\geq 0}$ as follows

$$X_t = \mu t + \sigma B_t, \quad (\mu \in \mathbb{R}, \sigma > 0),$$

where μ is the drift coefficient and σ is the diffusion coefficient.

Theorem 1.3. *Given* $0 < t_1 < t_2 < \cdots < t_n$*,* $n \geq 1$*,*

$$\begin{aligned} B_{t_1} &\sim N(0, t_1) \\ B_{t_2} - B_{t_1} &\sim N(0, t_2 - t_1) \\ &\vdots \\ B_{t_n} - B_{t_{n-1}} &\sim N(0, t_n - t_{n-1}) \end{aligned}$$

are independent if and only if

$$\left(B_{t_1}, B_{t_2}, \dots, B_{t_n}\right) \sim N\left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \left[t_i \wedge t_j\right]_{i,j=1}^{n}\right)$$

1.3 Standard Brownian Motion (SBM)

Theorem 1.4. *Let* $B = (B_t)_{t\geq 0}$ *be a SBM defined on* $(\Omega, \mathscr{F}, \mathbb{P})$*. Then each of the following processes is a SBM as well:*

1. *Strong Markov Property.*

$$(B_{T+t} - B_T)_{t\geq 0} \quad (T \text{ is stopping time of } B)$$

2. *Time-reversal*

$$(B_{T-t} - B_T)_{t\in[0,T]} \quad (T \text{ is stopping time of } B)$$

3. *Reflection property*

$$(-B_t)_{t\geq 0}$$

4. *Time-inversion*

$$(tB_{1/t})_{t\geq 0} := \begin{cases} tB_{1/t} & \text{if } t > 0 \\ 0 & \text{if } t = 0. \end{cases}$$

5. *Brownian scaling*

$$\left(\frac{B_{\rho t}}{\sqrt{\rho}}\right)_{t\geq 0}, \quad (\rho > 0)$$

Proposition 1.5 (The law of large number for BM)**.** *If* $B = (B_t)_{t\geq 0}$ *is a SBM defined on* $(\Omega, \mathscr{F}, \mathbb{P})$*, then*

$$\lim_{t\to\infty} \frac{B_t}{t} = 0, \quad \mathbb{P} - a.s$$

1.4 Stopping Time

Remark 1.6. When we say in (a) and (b), T is a stopping time of B, we mean that T is a stopping time with respect to the natural filtration generated by B, $(\mathscr{F}_t^B)_{t\geq 0}$, where $(\mathscr{F}_t^B)_{t\geq 0} = \sigma(B_s \mid 0 \leq s \leq t)$ is the smallest σ-algebra on Ω which makes each $(B_s \mid 0 \leq s \leq t)$ measurable for $0 \leq s \leq t$.

Note. Knowing $\mathscr{F}_t^B$ means that we know the sample path $s \to B_s$ for $s \in [0, t]$ and vice versa.

Definition 1.7 (Stopping time)**.** T is a stopping time w.r.t. $(\mathscr{F}_t^B)_{t\geq 0}$ if and only if:

$$\{T \leq t\} \in \mathscr{F}_t^B$$

for all $t \geq 0$. This T is a function from Ω into $\mathbb{R}_+$ or $\overline{\mathbb{R}}_+ = [0, \infty]$. Here $\{T \leq t\} \equiv \{\omega \in \Omega \mid T(\omega) \leq t\}$.

Let

$$\mathscr{F}_{t+}^B = \bigcap_{s>t} \mathscr{F}_t^B, \quad \mathscr{F}_t^B \subsetneqq \mathscr{F}_{t+}^B$$

for $t \geq 0$. Note that each $\mathscr{F}_{t+}^B$ is a sigma-algebra and that $\mathscr{F}_t^B \subseteq \mathscr{F}_{t+}^B$ for all $t \geq 0$. This inclusion is strict for B.

Example 1.8. Let us define $T : \Omega \to \mathbb{R}_+$ as

$$T = \inf\{t \geq 0 \mid B_t = 1\}.$$

is a stopping time with respect to $(\mathscr{F}_t^B)_{t\geq 0}$.

Example 1.9. The first hitting time to the set $(1, \infty)$:

$$\tau_{(1,\infty)} := \inf\{t > 0 \mid B_t > 1\}$$

is a stopping time w.r.t. $(\mathscr{F}_{t+}^B)_{t\geq 0}$.

Theorem 1.10 (The law of iterated logarithm for BM)**.** *If $B = (B_t)_{t\geq 0}$ is a SBM defined on $(\Omega, \mathscr{F}, \mathbb{P})$, then:*

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t \log\log t}} = 1, \quad \mathbb{P}\text{-}a.s,$$

$$\liminf_{t\to\infty} \frac{B_t}{\sqrt{2t \log\log t}} = -1, \quad \mathbb{P}\text{-}a.s,$$

$$\limsup_{t\to 0} \frac{B_t}{\sqrt{2t \log\log(1/t)}} = 1, \quad \mathbb{P}\text{-}a.s,$$

$$\liminf_{t\to 0} \frac{B_t}{\sqrt{2t \log\log(1/t)}} = -1, \quad \mathbb{P}\text{-}a.s,$$

Proposition 1.11 (SBM is nowhere differentiable)**.** *Let $B = (B_t)_{t\geq 0}$ be a SBM defined on $(\Omega, \mathscr{F}, \mathbb{P})$. Then there exist $Z \in \mathscr{F}$ with $\mathbb{P}(Z) = 0$ such that for each $\omega \in \Omega \setminus Z$, then function $t \mapsto B_t(\omega)$ is nowhere differentiable on $\mathbb{R}_+$.*

1.5 Quadratic Variation of Standard Brownian Motion

Definition 1.12. A function $f : \mathbb{R}_+ \to \mathbb{R}$ is said to be of bounded (or finite) variation if

$$\mathcal{V}(f; [0,t]) := \sup \left\{ \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| : 0 \leq t_0 < t_1 < \cdots < t_{n-1} < t_n \leq t \right\}$$

is finite. We say that $\mathcal{V}(f; [0,t])$ is the total variation of f on $[0,t]$.

Example 1.13.

- Continuous functions are not necessarily of bounded variation (BV) as the following example show:
$$f(t) = \begin{cases} t\sin(1/t) & \text{for } t > 0 \\ 0 & \text{for } t = 0 \end{cases}$$
- If $f \in C^1[0,t]$, then f is of BV on $[0,t]$.
- If $f(t) = \int_0^t g(s)\mathrm{d}s$ for $t \geq 0$ (where $g \in L^1$), then f is of BV and we have
$$\mathcal{V}(f; [0,t]) = \int_0^t |g(s)|\mathrm{d}s, \quad (t \geq 0)$$
which is the arc length of its graph.

Setting the following:

$$\mathcal{V}^+(f; [0,t]) = \frac{\mathcal{V}(f; [0,t]) + f(t) - f(0)}{2}$$
$$\mathcal{V}^-(f; [0,t]) = \frac{\mathcal{V}(f; [0,t]) - f(t) + f(0)}{2}$$

for $t \geq 0$, then $t \mapsto \varphi(t) := \mathcal{V}^+(f; [0,t])$ and $\psi(t) := \mathcal{V}^-(f; [0,t])$ are increasing function on $\mathbb{R}_+$ such that $\varphi(0) = \psi(0) = 0$. Then,

$$f(t) = f(0) + \varphi(t) - \psi(t),$$
$$\mathcal{V}(f; [0,t]) = \varphi(t) + \psi(t)$$

for $t \geq 0$. Then function $\mathcal{V}^+$ is called the positive variation and the function $\mathcal{V}^-$ is called the negative variation of f.

This shows that f is of BV if and only if f is the difference of two increasing functions. It follows that each f of BV is differentiable at almost all points in $\mathbb{R}^+$ with respect to the Lebesgue measure. And now we can state the following consequence of Proposition 1.11.

Corollary 1.14 (SBM is of unbounded variation)**.** *Let $B = (B_t)_{t\geq 0}$ to be a SBM defined on $(\Omega, \mathscr{F}, \mathbb{P})$. Then there exist $Z \in \mathscr{F}$ with $\mathbb{P}(Z) = 0$ such that for each $\omega \in \Omega \setminus Z$, the function $t \mapsto B_t(\omega)$ is of unbouned variation on each interval $[s,t] \subseteq \mathbb{R}_+$. In other words, for every $\omega \in \Omega \setminus Z$ and any $[s,t] \neq \emptyset \subseteq \mathbb{R}^+$, we have*

$$\sup\left\{\sum_{i=1}^{n}\left|B_{t_i}(\omega) - B_{t_{i-1}}(\omega)\right| : s \leq t_0 < \cdots < t_n \leq t\right\} = \infty. \tag{1.1}$$

Proposition 1.15 (Quadratic variation of SBM)**.** *Let $B = (B_t)_{t\geq 0}$ be a SBM defined on $(\Omega, \mathscr{F}, \mathbb{P})$. Consider*

$$S_n = \sum_{k=1}^{N_n}\left(B_{t_k^{(n)}} - B_{t_{k-1}^{(n)}}\right)^2$$

where $0 = t_0^{(n)} < t_1^{(n)} < \cdots < t_{N_n-1}^{(n)} < t_{N_n}^{(n)} = t$ is a subdivision of $[0,t]$ such that:

$$\delta_n := \max_{1\leq k\leq N_n}\left(t_k^{(n)} - t_{k-1}^{(n)}\right) \to 0, \tag{1.2}$$

as $n \to \infty$. Then:

$$S_n \to t \quad \text{in } L^2(\mathbb{P}) \text{ (and in } \mathbb{P}\text{-probability)} \tag{1.3}$$

as $n \to \infty$, where t is the length of the interval.

Moreover, if the rate of convergence in (1.2) is sufficiently fast to imply:

$$\sum_{n=1}^{\infty}\delta_n < \infty,$$

then:

$$S_n \to t, \quad \mathbb{P}\text{-}a.s.$$

as $n \to \infty$.

Proof. Note that,

$$B_{t_k}^{(n)} - B_{t_{k-1}}^{(n)} \underbrace{\sim}_{\substack{\text{stationary}\\\text{increment}}} B_{t_k^{(n)} - t_{k-1}^{(n)}} \underbrace{\sim}_{\substack{\text{Brownian}\\\text{scaling}}} \sqrt{t_k^{(n)} - t_{k-1}^{(n)}}B_1.$$

Hence

$$S_n = \sum_{k=1}^{N_n}\left(B_{t_k^{(n)}} - B_{t_{k-1}^{(n)}}\right)^2 \sim \sum_{k=1}^{N_n}\left(t_k^{(n)} - t_{k-1}^{(n)}\right)Z_k,$$

where Z_k is independent and identically distributed and $B_1^2 \sim \chi^2(1)$. Denote $d_k = t_k^{(n)} - t_{k-1}^{(n)}$. It follows that:

$$\mathbb{E}\left[(S_n - t)^2\right] = \mathbb{E}\left[\left(\sum_{k=1}^{N_n}d_k Z_k - \sum_{k=1}^{N_n}d_k\right)^2\right]$$

Using $\mathbb{E}[Z_k] = \mathbb{E}[B_1^2] = \mathbb{V}[B_1] - \mathbb{E}[B_1]^2 = 1$, we have

$$\begin{aligned}
\mathbb{E}\left[(S_n - t)^2\right] &= \mathbb{E}\left[\left(\sum_{k=1}^{N_n} d_k Z_k - \sum_{k=1}^{N_n} d_k \mathbb{E}[Z_k]\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_{k=1}^{N_n} d_k Z_k - \mathbb{E}\left[\sum_{k=1}^{N_n} d_k Z_k\right]\right)^2\right] \\
&= \mathbb{V}\left[\sum_{k=1}^{N_n} d_k Z_k\right] = \sum_{k=1}^{N_n} d_k^2 \mathbb{V}[Z_k]
\end{aligned}$$

Moreover, using $\mathbb{V}[Z_k] = \mathbb{E}[Z_k^2] - \mathbb{E}[Z_k]^2 = \mathbb{E}[B_1^4] - \mathbb{E}[B_1^2]^2 = 2$, we have

$$\begin{aligned}
\mathbb{E}\left[(S_n - t)^2\right] &= 2\sum_{k=1}^{N_n} d_k^2 = 2\sum_{k=1}^{N_n}\left(t_k^{(n)} - t_{k-1}^{(n)}\right)^2 \\
&\le 2\delta_n \sum_{k=1}^{N_n}\left(t_k^{(n)} - t_{k-1}^{(n)}\right) = 2\delta_n t \longrightarrow 0 \quad \text{as } t \longrightarrow \infty,
\end{aligned}$$

which completes the proof of (1.3).

Note that:

$$\mathbb{E}\left[\sum_{n=1}^{\infty}(S_n - t)^2\right] = \sum_{n=1}^{\infty}\mathbb{E}\left[(S_n - t)^2\right] \le \sum_{n=1}^{\infty} 2t\delta_n = 2t\sum_{n=1}^{\infty}\delta_n < \infty.$$

It follows that $\sum_{n=1}^{\infty}(S_n - t)^2 < \infty \quad \mathbb{P}\text{-}a.s.$ and hence $(S_n - t)^2 \longrightarrow 0 \quad \mathbb{P}\text{-}a.s.$, which is the same thing as $S_n \longrightarrow t \quad \mathbb{P}\text{-}a.s$ as $n \longrightarrow \infty$. The proof is now complete. □

2 Stochastic Integral Process

2.1 Motivation and Notation

The Chapman-Kolmogorov equation or Kolmogorov forward equation (Fokker Planck Equation) states

$$p_t = -(\mu p)_x + (Dp)_{xx}$$

where μ is the drift coefficient and $D = \sigma^2/2$ where σ is the diffusion coefficient. Then we would like to find

$$\mathbb{P}(X_t \in B) = \int_B p(t,x)\mathrm{d}x \tag{2.1}$$

The Kolmogorov problem stated in 1930s said:

How to construct a $(X_t)_{t\geq 0}$ such that Equation (2.1) holds for all t and all Borel set B given $p(t,x)$?

The idea provided by Itô said the process X should solve the following SDE:

$$\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t.$$

- $B = (B_t)_{t\geq 0}$ as the standard Brownian motion.
- $\mathscr{F}_t = \overline{\mathscr{F}_t^B} = \sigma\left(\mathscr{F}_t^B \cup \mathcal{N}\right)$,
  where $\mathscr{F}_t^B = \sigma(B_s \mid 0 \leq s \leq t)$ and $\mathcal{N} = \{N \in \mathscr{F} \mid \mathbb{P}(N) = 0\}$.
- $H = (H_t)_{t\geq 0}$ is the stochastic process.

2.2 Simple Process

Question. We would like to define

$$\int_0^t H_s(\omega)\mathrm{d}B_s(\omega) = ? \tag{SI}$$

since $\mathrm{d}B_s(\omega)$ can not be defined as a classic Lebesgue-Stieltjes integral since $s \mapsto B_s(\omega)$ is of unbounded variation.

Example 2.1. A natural idea is to define Eq (SI) first for simple process.

$$H_t(\omega) = \sum_{i\geq 1} h_{i-1}(\omega)\mathbb{1}_{(t_{i-1},t_i]}(t). \tag{2.2}$$

The geometric interpretation of the simple process can be seen from Figure 1.

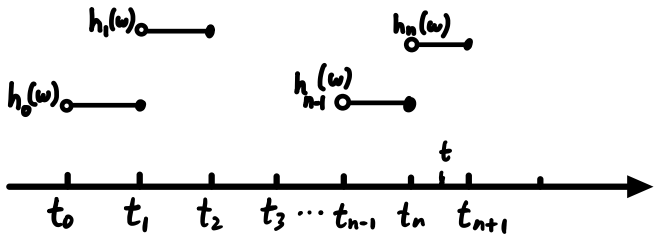

Figure 1: The visualization of the simple process (2.2).

Based on graph, we can write the integral as

$$\int_0^t H_s \mathrm{d}B_s = \sum_{i=1}^{n} h_{i-1} \underbrace{\left(B_{t_i} - B_{t_{i-1}}\right)}_{\text{Increment of SBM}} + h_n \underbrace{\left(B_t - B_{t_n}\right)}_{\text{Increment from } t_n \text{ to } t}$$

where $t_0, t_1, \ldots, t_n$ are partition points inside the interval $[0, t)$.

The problem arises: $h_{i-1}(\omega)$ may depends on the final point of the interval, then $h_n(B_t - B_{t_n})$ will need the information from t_{n+1} as well.

Example 2.2. Consider the interval $[0, 1]$, and $t_i = i/n$ for $0 \le i \le n$. Then we have the following two definitions of the simple process's integral

$$H_t^n = \sum_{i=1}^{n} B_{t_{i-1}} \mathbb{1}_{(t_{i-1}, t_i]}(t), \quad \text{Approximation using left node}$$

$$K_t^n = \sum_{i=1}^{n} B_{t_i} \mathbb{1}_{(t_{i-1}, t_i]}(t). \quad \text{Approximation using right node}$$

Then we have

$$\begin{aligned}
\mathbb{E}\left[\int_0^1 H_s^n \mathrm{d}B_s\right] &= \sum_{i=1}^{n} \mathbb{E}\left[B_{t_{i-1}}(B_{t_i} - B_{t_{i-1}})\right] = 0 \\
\mathbb{E}\left[\int_0^1 K_s^n \mathrm{d}B_s\right] &= \sum_{i=1}^{n} \mathbb{E}\left[B_{t_i}(B_{t_i} - B_{t_{i-1}})\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[(B_{t_i} - B_{t_{i-1}})^2 + B_{t_{i-1}}(B_{t_i} - B_{t_{i-1}})\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[(B_{t_i} - B_{t_{i-1}})^2\right] = \sum_{i=1}^{n} \mathbb{E}\left[(B_{1/n})^2\right] = n \cdot 1/n = 1.
\end{aligned}$$

Thus, although $H^n \approx B$ and $K^n \approx B$ on $[0, 1]$ for large n, we see that

$$\int_0^1 H_s^n \mathrm{d}B_s \not\approx \int_0^1 K_s^n \mathrm{d}B_s,$$

no matter how large n is.

Let us now describe the class $\mathcal{H}_2([0,t])$ of process H for which the Itô integral will be defined:

1. H is measurable on $[0,t] \times \Omega$, i.e. $(s, \omega) \mapsto H_s(\omega)$ is measurable with respect to

$$\mathscr{B}([0,t]) \otimes \mathscr{F} \text{ and } \mathscr{B}(\mathbb{R}),$$

   where $\mathscr{B}([0,t]) \otimes \mathscr{F}$ is defined as $\sigma\left(\{B \times F \mid B \in \mathscr{B}([0,t]), F \in \mathscr{F}\}\right)$.

2. H is $(\mathscr{F}_s)_{0 \le s \le t}$-adapted. Namely, H_s is $\mathscr{F}_s$ measurable for each $0 \le s \le t$.
3. $\mathbb{E}\left[\int_0^t H_s^2 \mathrm{d}s\right] < \infty$.

Let $H \in \mathcal{H}_2([0,t])$ be given. We will now show how to define the Itô's integral with respect to SBM, i.e.

$$\int_0^t H_s(\omega) \mathrm{d}B_s(\omega) = ?$$

Idea: We first define for simple process H in $\mathcal{H}_2([0,t])$, then we show that each $H \in \mathcal{H}_2([0,t])$ can be approximated by a sequence of simple process H^n in $\mathcal{H}_2([0,t])$ for $n \ge 1$. Finally $\int_0^t H_s \mathrm{d}B_s$ can be defined as a limit of $\int_0^t H_s^n \mathrm{d}B_s$ as $n \to \infty$.

Definition 2.3 (Simple Process)**.** A process $H \in \mathcal{H}_2([0,t])$ is called *simple* (or elementary), if

$$H_t(\omega) = \sum_{i=1}^{\infty} h_{i-1}(\omega) \mathbb{1}_{(t_{i-1}, t_i]}(t),$$

where $h_{i-1}(\omega)$ is $\mathscr{F}_{i-1}$ measurable.

Let $\mathcal{S}([0,t])$ denote all simple processes in $\mathcal{H}_2([0,t])$.

Definition 2.4 (The Itô's integral of simple process)**.** If $H \in \mathcal{S}([0,t])$ then,

$$\int_0^t H_s \mathrm{d}B_s = \sum_{i=1}^{n} h_{i-1}(B_{t_i} - B_{t_{i-1}}) + h_n(B_t - B_{t_n}).$$

We have the following important observation

Lemma 2.5 (The Itô's isometry). *If $H \in \mathcal{S}([0,t])$ is bounded, i.e.*

$$\sup_{(s,\omega)\in[0,t]\times\Omega} |H_s(\omega)| < \infty,$$

then we have

$$\mathbb{E}\left[\left|\int_0^t H_s \mathrm{d}B_s\right|^2\right] = \mathbb{E}\left[\int_0^t |H_s|^2 \mathrm{d}s\right]$$

Proof. We have:

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^t H_s \mathrm{d}B_s\right)^2\right] &= \mathbb{E}\left[\left(\sum_{i=1}^n h_{i-1}(B_{t_i} - B_{t_{i-1}}) + h_n(B_t - B_{t_n})\right)^2\right] \\
&= \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}\left[h_{i-1}h_{j-1}(B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})\right] \\
&+ 2\sum_{i=1}^n \left[\underbrace{h_{i-1}h_n(B_{t_i} - B_{t_{i-1}})}_{\mathcal{F}_{t_n} \text{ measurable}} \underbrace{(B_t - B_{t_n})}_{\perp \mathcal{F}_{t_n}}\right] \\
&+ \mathbb{E}\left[h_n^2(B_t - B_{t_n})^2\right]
\end{aligned} \tag{2.3}$$

We can first focus on the double sum:

$$i < j \quad \Rightarrow \quad \mathbb{E}\left[\underbrace{h_{i-1}h_{j-1}(B_{t_i} - B_{t_{i-1}})}_{\mathcal{F}_{t_{j-1}} \text{ measurable}} \underbrace{(B_{t_j} - B_{t_{j-1}})}_{\perp \mathcal{F}_{t_{j-1}}}\right] = 0;$$

$$i > j \quad \Rightarrow \quad \mathbb{E}\left[\underbrace{h_{i-1}h_{j-1}(B_{t_j} - B_{t_{j-1}})}_{\mathcal{F}_{t_{i-1}} \text{ measurable}} \underbrace{(B_{t_i} - B_{t_{i-1}})}_{\perp \mathcal{F}_{t_{i-1}}}\right] = 0;$$

Hence the Eq. (2.3) becomes

$$\begin{aligned}
&= \sum_{i=1}^n \mathbb{E}\left[h_{i-1}^2 \left(B_{t_i} - B_{t_{i-1}}\right)^2\right] + \mathbb{E}\left[h_n^2(B_t - B_{t_n})^2\right] \\
&= \sum_{i=1}^n \mathbb{E}\left[h_{i-1}^2\right] \mathbb{E}[(B_{t_i} - B_{t_{i-1}})^2] + \mathbb{E}[h_n^2]\mathbb{E}[(B_t - B_{t_n})^2] \\
&= \sum_{i=1}^n \mathbb{E}[h_{t_{i-1}}^2](t_i - t_{i-1}) + \mathbb{E}[h_n^2](t - t_n) \\
&= \mathbb{E}\left[\int_0^t H_s^2 \mathrm{d}s\right]
\end{aligned}$$

□

2.3 Relax $\mathcal{S}([0,t])$ to $\mathcal{H}_2([0,t])$

The idea now is to use the isometry and extend the Itô integral from $\mathcal{S}([0,t])$ to $\mathcal{H}_2([0,t])$. Here are three steps

Step 1. Let $H \in \mathcal{H}_2([0,t])$ be continuous [1] and bounded [2]. Then there exist $(H^n)_{n\geq 1}$ in $\mathcal{S}([0,t])$ such that

$$\mathbb{E}\left[\left(\int_0^t \left|H_s - H_s^n\right|^2 \mathrm{d}s\right)\right] \to 0 \quad n \to \infty.$$

Proof. Define

$$H_s^n(\omega) = \sum_{i\geq 1} H_{t_{i-1}^n}(\omega)\mathbb{1}_{(t_{i-1}^n, t_i^n]}(s)$$

for $0 \leq s \leq t$. Here

$$\max_{1\leq i\leq m_n}(t_i^n - t_{i-1}^n) \to 0, \quad \text{as } n \to \infty.$$

Then $H^n \in \mathcal{S}([0,t])$ for each $n \geq 1$.

By dominated convergence

$$\int_0^t \left|H_s(\omega) - H_s^n(\omega)\right|^2 \mathrm{d}s \to 0, \quad \text{as } n \to \infty$$

Since $s \to H_s(\omega)$ is continuous for all $\omega \in \Omega$, hence by bounded convergence

$$\mathbb{E}\left[\int_0^t \left|H_s - H_s^n\right|^2 \mathrm{d}s\right] \to 0 \quad \text{as } n \to \infty.$$

□

Step 2. Let $H \in \mathcal{H}_2([0,t])$ be bounded. Then there exist a continuous and bounded $H^n \in \mathcal{H}_2([0,t])$ for $n \geq 1$ such that

$$\mathbb{E}\left[\int_0^t \left|H_s - H_s^n\right|^2 \mathrm{d}s\right] \to 0 \quad \text{as } n \to \infty.$$

Proof. We need the following general result on the *convolution approximation*.

Lemma 2.6 (A Convolution approximation)**.** *Let $F : \mathbb{R} \to \mathbb{R}$ belongs to L^2, i.e.*

$$\int_{\mathbb{R}} |F(x)|^2 \, \mathrm{d}x < \infty.$$

Let $\Omega : \mathbb{R} \to \mathbb{R}_+$ satisfies

[1] $s \mapsto B_s(\omega)$ is continuous on $[0,t]$

[2] $\sup_{(s,\omega)\in[0,t]\times\Omega} |H_s(\omega)| < \infty$

(a) Ω *is* C^0, *i.e. continuous,*

(b) $\mathrm{supp}(\Omega) = [0,1]$,

(c) $\int_{\mathbb{R}} \Omega(x)\,\mathrm{d}x = 1.$

Define

$$F^n(x) = \int_{\mathbb{R}} F\left(x - \frac{y}{n}\right)\Omega(y)\,\mathrm{d}y,$$

then we have

- $x \longrightarrow F^n(x)$ *is continuous.*
- $F^n \longrightarrow F$ *in* L^2 *as* $n \to \infty$.

Then we can return to our proof of step 2.

Since $H \in \mathcal{H}_2([0,t])$ we have H is bounded

$$\mathbb{E}\left[\int_0^t H_s^2 \mathrm{d}s\right] < \infty,$$

and hence $s \mapsto H_s(\omega) \in L^2([0,t])$ for each $\omega \in \Omega \setminus N$ where $\mathbb{P}(N) = 0$. Extend $s \mapsto H_s(\omega)$ on $\mathbb{R}$ by setting $H_s(\omega) = 0$ if $s < 0$ or $s > t$. Then, $s \to H_s(\omega) \in L^2(\mathbb{R})$, and we may define

$$H_s^n(\omega) = \int_{\mathbb{R}} H_{s-t/n}(\omega)\Omega(r)\,\mathrm{d}r \quad (s \in \mathbb{R}).$$

Then by the Lemma 2.6, we know that

- $s \longrightarrow H_s^n(\omega)$ is continuous on $\mathbb{R}$,
- $\int_{\mathbb{R}} \left|H_s(\omega) - H_s^n(\omega)\right|^2 \mathrm{d}s \longrightarrow 0 \quad (n \to \infty)$

Clearly, each $H_s^n(\cdot)$ is $\mathscr{F}_s$ measurable for $s \in [0,t]$, so that

$$H^n = (H_s^n)_{0\le s\le t} \in \mathcal{H}_2([0,t]).$$

Moreover since H is bounded, so is H^n, and from (ii), we see that

$$\int_0^t \left|H_s(\omega) - H_s^n(\omega)\right|^2 \mathrm{d}s \longrightarrow 0 \quad \text{as } n \to \infty$$

Hence by dominated convergence

$$\mathbb{E}\left(\int_0^t \left|H_s - H_s^n\right|^2 \mathrm{d}s\right) \longrightarrow 0 \quad \text{as } n \to \infty.$$

The proof is complete. □

Step 3. Let $H \in \mathcal{H}_2([0,t])$, then there exist a bounded $H^n \in \mathcal{H}_2([0,t])$ for $n \geq 1$ such that:

$$\mathbb{E}\left(\int_0^t \left|H_s - H_s^n\right|^2 \mathrm{d}s\right) \to 0 \quad \text{as } n \to \infty$$

Proof. Put:

$$H_s^n(\omega) = \begin{cases} -n & \text{if } H_s(\omega) < -n \\ H_s(\omega) & \text{if } -n \leq H_s(\omega) \leq n \\ n & \text{if } H_s(\omega) > n \end{cases}$$

for $s \in [0,t]$ and $\omega \in \Omega$. The claim then follows by dominated convergence. □

We are now ready to complete the construction of the Itô integral of $H \in \mathcal{H}_2([0,t])$.

Definition 2.7 (The Itô integral)**.** Let $H \in \mathcal{H}_2([0,t])$, then the Itô integral of H on $[0,t]$ with respect to B is defined by

$$\int_0^t H_s \, \mathrm{d}B_s = L^2 - \lim_{n\to\infty} \int_0^t H_s^n \, \mathrm{d}B_s \tag{$**$}$$

where $(H^n)_{n\geq 1}$ is any sequence from $[0,t])$ such that

$$\mathbb{E}\left(\int_0^t \left|H_s - H_s^n\right|^2 \mathrm{d}s\right) \to 0 \quad \text{as } n \to \infty. \tag{$*$}$$

Remark 2.8.

(i) Such a sequence exists by step (1) to step (3)

Step (1) There exist $(H^n)_{n\geq 0} \in \mathcal{S}([0,t])$ such that H^n can approximate $H \in \mathcal{H}_2([0,t])$ where H is continuous and bounded.

Step (2) There exist a continuous and bounded $H^n \in \mathcal{H}_2([0,t])$ such that H^n can approximate bounded $H \in \mathcal{H}_2([0,t])$.

Step (3) There exist a bounded $H^n \in \mathcal{H}_2([0,t])$ such that H^n can approximate $H \in \mathcal{H}_2([0,t])$.

(ii) If H^n and K^n are two sequences in $\mathcal{S}([0,t])$ such that ($*$) holds for both, then,

$$L^2 - \lim_{n\to\infty} \int_0^t H_s^n \, \mathrm{d}B_s = L^2 - \lim_{n\to\infty} \int_0^t K_s^n \, \mathrm{d}B_s$$

Proof.

$$
\begin{aligned}
&\mathbb{E}\left[\left|\int_0^t H_s^n \,\mathrm{d}B_s - \int_0^t K_s^n \,\mathrm{d}B_s\right|^2\right]\\
=&\mathbb{E}\left[\left|\int_0^t H_s^n - K_s^n \,\mathrm{d}B_s\right|^2\right]\\
=&\mathbb{E}\left[\int_0^t \left|H_s^n - K_s^n\right|^2 \,\mathrm{d}s\right] \quad \text{Itô isometry}\\
\leq &2\mathbb{E}\left[\int_0^t \left|H_s^n - H_s\right|^2 \,\mathrm{d}s\right] + 2\mathbb{E}\left[\int_0^t \left|H_s - K_s^n\right|^2 \,\mathrm{d}s\right]\\
&\to 0.
\end{aligned}
$$

□

(iii) The L^2–limit in ($**$) exists for any $(H^n)_{n\geq 1}$ in $\mathcal{S}([0,t])$ satisfying ($*$), because the latter forms a Cauchy sequence in $L^2(\mathbb{P})$.

Corollary 2.9 (The Itô isometry)**.** *If* $H \in \mathcal{H}_2([0,t])$*, then:*

$$
\mathbb{E}\left[\left|\int_0^t H_s \,\mathrm{d}B_s\right|^2\right] = \mathbb{E}\left[\int_0^t |H_s|^2 \,\mathrm{d}s\right].
$$

Corollary 2.10. *If* H^n *and* H *belongs to* $\mathcal{H}([0,t])$ *for* $n \geq 1$ *and*

$$
\mathbb{E}\left[\int_0^t \left|H_s^n - H_s\right|^2 \,\mathrm{d}s\right] \to 0, \quad \text{as } n \to \infty,
$$

then we have

$$
\int_0^t H_s^n \,\mathrm{d}B_s \to \int_0^t H_s \,\mathrm{d}B_s \quad \text{in } L^2 \text{ as } n \to \infty.
$$

Example 2.11. Let us compute

$$
\int_0^t B_s \,\mathrm{d}B_s.
$$

Put

$$
\begin{aligned}
&H_s^n(\omega) = \sum_{i\geq 1} B_{t_{i-1}}(\omega)\mathbb{1}_{(t_{i-1},t_i]}(s) \quad 0 \leq s \leq t,\\
&\delta_n := \max_{1\leq i\leq n}(t_i - t_{i-1}) \to 0 \quad \text{as } n \to \infty\text{, and let } t_n = t.
\end{aligned}
$$

Then

$$\mathbb{E}\left[\int_0^t \left|B_s - H_s^n\right|^2 \,\mathrm{d}s\right] = \mathbb{E}\left[\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \left|B_s - B_{t_{i-1}}\right|^2 \,\mathrm{d}s\right]$$

$$= \sum_{i=1}^n \int_{t_{i-1}}^{t_i} (s - t_{i-1}) \,\mathrm{d}s$$

$$= \sum_{i=1}^n \frac{1}{2}(t_i - t_{i-1})^2 \leq \delta_n t_n/2 \longrightarrow 0 \quad \text{as } n \longrightarrow \infty.$$

By Corollary 2.10, we have

$$\int_0^t B_s \,\mathrm{d}s = L^2 - \lim_{n\to\infty} \int_0^t H_s^n \,\mathrm{d}B_s = L^2 - \lim_{n\to\infty} \sum_{i=1}^n B_{t_{i-1}}(B_{t_i} - B_{t_{i-1}}).$$

Now, note the following:

$$B_t^2 = \sum_{i=1}^n B_{t_i}^2 - B_{t_{i-1}}^2 = \sum_{i=1}^n \left\{(B_{t_i} - B_{t_{i-1}})^2 + 2B_{t_{i-1}}(B_{t_i} - B_{t_{i-1}})\right\}$$

$$= \underbrace{\sum_{i=1}^n (B_{t_i} - B_{t_{i-1}})^2}_{\to t \text{ (Quadratic convergence)}} + 2\underbrace{\sum_{i=1}^n B_{t_{i-1}}(B_{t_i} - B_{t_{i-1}})}_{\int_0^t B_s \,\mathrm{d}B_s \text{ in } L^2}$$

Hence

$$\int_0^t B_s \,\mathrm{d}B_s = \frac{1}{2}B_t^2 - \frac{1}{2}t.$$

2.4 Continuous Modification

Let H and K belongs to $\mathcal{H}^2([0,t])$, and let $0 \leq t_1 < t_2 \leq t$ be given and fixed. Then

1. $$\int_0^{t_2} H_s \,\mathrm{d}B_s = \int_0^{t_1} H_s \,\mathrm{d}B_s + \int_{t_1}^{t_2} H_s \,\mathrm{d}B_s$$

2. $$\int_0^t (aH_s + bK_s) \,\mathrm{d}B_s = a\int_0^t H_s \,\mathrm{d}B_s + b\int_0^t K_s \,\mathrm{d}B_s \quad \forall a, b \in \mathbb{R}$$

3. $$\mathbb{E}\left[\int_0^t H_s \,\mathrm{d}B_s\right] = 0$$

4.
$$\int_0^t H_s \, \mathrm{d}B_s \text{ is } \mathscr{F}_t \text{ measurable}$$

Let $H \in \mathcal{H}_2([0,t])$ for each $t \geq 0$, and set $I_t = \int_0^t H_s \, \mathrm{d}B_s$ for every $t \geq 0$. In this way, we have obtained a stochastic process $I = (I_t)_{t\geq 0}$. The next theorem shows that I admits a *continuous modification*.

Lemma 2.12 (Doob's Maximal Inequality)**.** *If $M = (M_t)_{t\geq 0}$ is a continuous martingale, then*
$$\mathbb{P}\left(\max_{0\leq t\leq T} |M_t| \geq \lambda\right) \leq \frac{1}{\lambda^p} \mathbb{E}\left[|M_T|^p\right]$$
for each $p \geq 1, \lambda > 0$ and $T > 0$.

Theorem 2.13. *There exist a continuous modification of the process $I = (I_t)_{t\geq 0}$, i.e. there exist a continuous process $J = (J_t)_{t\geq 0}$ such that $\mathbb{P}(I_t = J_t) = 1$ for each $t \geq 0$.*

Proof. Let
$$H_t^n(\omega) = \sum_{i\geq 1} h_{i-1}(\omega) \mathbb{1}_{(t_{i-1},t_i]}(t), \quad 0 \geq t \geq T$$
be element of $\mathcal{S}([0,t])$ such that
$$\mathbb{E}\left[\int_0^T \left|H_t^n - H_t\right|^2 \mathrm{d}t\right] \to 0 \quad \text{as } n \to \infty.$$
Put
$$I_t^n = \int_0^t H_s^n \, \mathrm{d}B_s \qquad I_t = \int_0^t H_s \, \mathrm{d}B_s.$$
Then clearly $t \longrightarrow I_t^n$ is continuous on Ω. Moreover, the process $I^n = (I_t^n)_{t\geq 0}$ is a martingale with respect to $(\mathscr{F}_t)_{t\geq 0}$. Here we will present a proof of $\mathbb{E}[I_t^n \mid \mathscr{F}_s] = I_s^n$.

Proof of Martingale. Suppose $s < t$ and $t_k < s < t_{k+1} < \cdots < t_\ell < t < t_{\ell+1}$, then
$$\begin{aligned}
\mathbb{E}[I_t^n \mid \mathscr{F}_s] &= \mathbb{E}\left[I_s^n + \int_s^t H_r^n \, \mathrm{d}B_r \mid \mathscr{F}_s\right] \\
&= I_s^n + \mathbb{E}\left[\int_s^t H_r^n \, \mathrm{d}B_r \mid \mathscr{F}_s\right] \\
&= I_s^n + \mathbb{E}\left[h_k(B_{t_{k+1}} - B_s) \mid \mathscr{F}_s\right] \quad \cdots (1) \\
&+ \mathbb{E}\left[\sum_{i=k+2}^{\ell} h_{i-1}(B_{t_i} - B_{t_{i-1}}) \mid \mathscr{F}_s\right] \quad \cdots (2) \\
&+ \mathbb{E}\left[h_\ell(B_t - B_{t_\ell}) \mid \mathscr{F}_s\right] \quad \cdots (3)
\end{aligned}$$

For part (1), h_k is $\mathscr{F}_{t_k} \subseteq \mathscr{F}_s$ measurable. Also, $(B_{t_{k+1}} - B_s)$ is independent of $\mathscr{F}_s$, hence we have

$$\mathbb{E}\left[h_k(B_{t_{k+1}} - B_s) \mid \mathscr{F}_s\right] = h_k \mathbb{E}\left[B_{t_{k+1}} - B_s\right] = 0.$$

For part (2) and (3), we use the property that if $\mathscr{F}_s \subseteq \mathscr{F}_{i-1}$, then in conditional expectation $\mathscr{F}_s$ and $\mathscr{F}_{i-1}$ commutes.

$$\mathbb{E}\left[h_{i-1}(B_{t_i} - B_{t_{i-1}}) \mid \mathscr{F}_s\right] = \mathbb{E}\left[\mathbb{E}\left[h_{i-1}(B_{t_i} - B_{t_{i-1}}) \mid \mathscr{F}_{i-1}\right] \mid \mathscr{F}_s\right] = 0.$$

This completes the proof. □

Hence we have $I^n - I^m = (I_t^n - I_t^m)_{t\geq 0}$ is also a martingale with respect to $(\mathscr{F}_t)_{t\geq 0}$. By Doob's Maximal Inequality, we therefore get

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\leq t\leq T}\left|I_t^m - I_t^n\right| > \varepsilon\right) &\leq \frac{1}{\varepsilon^2}\mathbb{E}\left[\left|I_T^m - I_T^n\right|^2\right] \\
&= \frac{1}{\varepsilon^2}\mathbb{E}\left[\int_0^T \left|H_t^m - H_t^n\right|^2 \mathrm{d}t\right] \quad \text{Itô isometry} \\
&\to 0 \quad \text{as } n, m \to \infty.
\end{aligned}$$

Hence we can choose $n_k \to \infty$ as $k \to \infty$ such that

$$\mathbb{P}\left(\sup_{0\leq t\leq T}\left|I_t^{n_{k+1}} - I_t^{n_k}\right| > 2^{-k}\right) \leq 2^{-k} \quad k \geq 1.$$

By the Borel-Cantelli lemma, it follows

$$\mathbb{P}\left(\underbrace{\sum_{0\leq t\leq T}\left|I_t^{n_{k+1}} - I_t^{n_k}\right| > 2^{-k} \text{ for infinitely many } k}_{N_T = \text{event happened}}\right) \to 0.$$

(**Note:** There only exist finitely many ω such that $\sum_{0\leq t\leq T}\left|I_t^{n_{k+1}} - I_t^{n_k}\right| > 2^{-k}$ happened. We collect these elements into a set N_T) Hence for each $\omega \in \Omega \setminus N_T$ with $\mathbb{P}(N_T) = 0$, there exists $k_0 = k_0(\omega) \geq 1$ large enough such that

$$\sup_{0\leq t\leq T}\left|I_t^{n_{k+1}} - I_t^{n_k}\right| \leq 2^{-k} \quad \forall k \geq k_0.$$

This shows that $I_t^{n_k}(\omega) \to J_t(\omega)$ on $[0, T]$ as $k \to \infty$, where $J_t(\omega)$ is continuous on $[0, T]$. Moreover, since $I_t^{n_k} \to I_t$ in $L^2(\mathbb{P})$ as $k \to \infty$, we must have $I_t = J_t$ $\mathbb{P}$--a.s. (due to completeness of the space $(L^2(\mathbb{P}), \|\cdot\|)$), and the proof is complete. □

Remark 2.14. $J = (J_t)_{t\geq 0}$ can be chosen to satisfy:

- J_t is $\mathscr{F}_t$ measurable.
- $t \mapsto J_t$ is $\mathbb{P}$-a.s continuous, i.e. there exist $N \in \mathscr{F}_\infty$ with $\mathbb{P}(N) = 0$, such that $t \mapsto J_t(\omega)$ is continuous $\forall \omega \in \Omega \setminus N$.
- $\mathbb{P}(I_t = J_t) = 1$.

We note that, if $H_s \in \mathcal{S}([0,t])$ is $\mathscr{F}_s$ measurable for each $s \in [0,t]$, then the Itô stochastic integral is also $\mathscr{F}_t$ measurable. This also holds if from the beginning we replace $\mathscr{F}_t$ by $\mathscr{F}_t^B$.

Proof of Remark 2.14. Set

$$J_t(\omega) = \begin{cases} \lim_{k\to\infty} I_t^{n_k}(\omega) & \omega \in \Omega \setminus N_t \\ 0 & \omega \in N_t \end{cases}$$

Note, $N_t \subseteq N_T$. □

Remark 2.15. From now on, we will always assume that

$$\left(\int_0^t H_s \,\mathrm{d}B_s\right)_{t\geq 0}$$

is the continuous modification, i.e. continuous and $\mathscr{F}_t$ measurable (adapted).

Corollary 2.16. *Let $H \in \mathcal{H}_2 := \bigcup_{T>0} \mathcal{H}_2([0,T])$. Then,*

$$M_t := \int_0^t H_s \,\mathrm{d}B_s$$

is a martingale with respect to $(\mathscr{F}_t)_{t\geq 0}$. Moreover,

$$\mathbb{P}\left(\sup_{0\leq t\leq T} |M_t| \geq \lambda\right) \leq \frac{1}{\lambda^2}\mathbb{E}\left[\int_0^T H_s^2 \,\mathrm{d}s\right]$$

for each $\lambda > 0$ and $T > 0$.

Proof. Recall from the proof of Theorem 2.13, if $I_t^n \to I_t$ in L^2, then

$$\mathbb{E}[I_t^n \mid \mathscr{F}_s] \to \mathbb{E}[I_t \mid \mathscr{F}_s] \quad \text{in } L^2,$$

We know that, I_t^n is a martingale, hence

$$\mathbb{E}[I_t^n \mid \mathscr{F}_s] = I_s^n.$$

Since $I_s^n \to I_s$ in L^2, therefore taking the L^2-limit of the equation, we have

$$\mathbb{E}[I_t \mid \mathscr{F}_s] = I_s.$$

Therefore we have proved that M_t is indeed a martingale. To get the inequality, we use Lemma 2.12.

$$\begin{aligned}\mathbb{P}\left(\sup_{0\le t\le T}\left|\int_0^t H_s \, \mathrm{d}B_s\right| \ge \lambda\right) &\le \frac{1}{\lambda^2}\mathbb{E}\left[\left|\int_0^T H_s \, \mathrm{d}B_s\right|^2\right] \\ &= \frac{1}{\lambda^2}\mathbb{E}\left[\left(\int_0^T |H_s| \, \mathrm{d}s\right)^2\right] \quad \text{By Itô isometry} \\ &\le \frac{1}{\lambda^2}\mathbb{E}\left[\int_0^T |H_s|^2 \, \mathrm{d}s\right], \quad \text{by Jensen's inequality.}\end{aligned}$$

□

2.5 Relax $\mathcal{H}_2([0,t])$ to $\mathcal{D}([0,t])$

Recall that $H \in \mathcal{H}_2([0,t])$ if

(1) H is $\mathscr{B}([0,t]) \otimes \mathscr{F}$ measurable.

(2) H_s is $\mathscr{F}$ measurable for each $0 \le s \le t$.

(3) $\mathbb{E}\left[\int_0^t H_s^2 \, \mathrm{d}s\right] < \infty$.

Firstly, (2) can be relaxed as follows:

(2)': there exist a filtration (not necessarily natural filtration) $(\mathscr{G}_s)_{0\le s\le t}$ such that

a. $(B_s)_{0\le s\le t}$ is a martingale with respect to $(\mathscr{G}_s)_{0\le s\le t}$.

b. H_s is $\mathscr{G}_s$ measurable for each $0 \le s \le t$.

Note. From now on, we does not enforce that the integrand is $\mathscr{F}_t$ measurable where $\mathscr{F}_t$ is the natural filtration with respect to B_t. Instead, any filtration $(\mathscr{G}_s)_{0\le s\le t}$ that makes B_t a martingale and makes H_s $\mathscr{G}_s$ measurable is suffice.

Example 2.17. (B^1, B^2) as a two dimensional BM, where B^1 and B^2 are independent. How can we evaluate

$$\int_0^t B_s^1 \, \mathrm{d}B_s^2.$$

The natural filtration of B^2 enlarged with a null set as required by the definition of $\mathcal{H}_2([0,t])$, then B^1 will not be adapted to this filtration since B^1 and B^2 are independent. Hence the integral is not evaluable. However, (2)' can be satisfied by taking $(\mathscr{G}_s)_{0\le s\le t}$ be the natural filtration of both B^1 and B^2, i.e. the natural filtration generated by this two-dimensional BM.

Secondly, (3) can be relaxed to

$$(3)' \qquad \mathbb{P}\left(\int_0^t H_s^2 < \infty\right) = 1.$$

Clearly, if (3) holds, then (3)' will definitely holds.

Let us denote $\mathcal{D}([0,t])$ the class of process H satisfying (1), (2)' and (3)'. Then, if $H \in \mathcal{D}([0,t])$, then there exist a simple process $(H^n)_{n\geq 1}$ in $\mathcal{D}([0,t])$ such that

$$\int_0^t \left|H_s - H_s^n\right|^2 \,\mathrm{d}s \to 0 \quad \text{in } \mathbb{P}\text{-probability.}$$

Moreover,

$$\int_0^t H_s^n \,\mathrm{d}B_s \to L_t \text{ in } \mathbb{P}\text{-probability,} \quad L_t := \int_0^t H_s \,\mathrm{d}B_s,$$

which is the Itô integral.

CONVERGENCE

Definition 2.18 (Convergence in Probability)**.** A sequence $X_1, X_2, \ldots$ of random variables converges in probability to a random variable X, written as

$$X_n \xrightarrow{\mathbb{P}} X,$$

if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| \geq \varepsilon) = 0,$$

or equivalently

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| < \varepsilon) = 1.$$

Definition 2.19 (Convergence almost surely)**.** A sequence $X_1, X_2, \ldots$ of random variables converges in almost surely to a random variable X, written as

$$X_n \xrightarrow{\text{a.s.}} X,$$

if for every $\varepsilon > 0$,

$$\mathbb{P}(\lim_{n\to\infty} |X_n - X| < \varepsilon) = 1,$$

This condition is stronger than the convergence in Probability.

Definition 2.20 (Convergence in distribution)**.** $X_1, X_2, \ldots$ are sequence of random variables with corresponding PDFs $\pi_{X_1}(x), \pi_{X_2}(x), \ldots$. If $\pi_X(x)$ is a distribution function of the random variable X, and

$$\lim_{n\to\infty} \pi_{X_n}(x) = \pi_X(x)$$

at all points for which $\pi_X(x)$ is continuous, then X_n is said to have a limiting random variable X, and X_n converges to X in distribution, i.e. $X_n \xrightarrow{D} X$.

Remark 2.21.

1. From the proof of Theorem 2.13, we can say there always exists a continuous modification of the process $\int_0^t H_s \, \mathrm{d}B_s$.
2. $(\int_0^t H_s \, \mathrm{d}B_s)$ is not a martingale generally, but a local martingale only.

A natural desire when having

$$M_t = \int_0^t H_s \, \mathrm{d}B_s$$

is to consider

$$\int_0^t K_s \, \mathrm{d}M_s$$

where K is a given process. Or even more generally, we want to define

$$\int_0^t K_s \, \mathrm{d}X_s$$

where $X = M + A$, here M is a stochastic integral defined above and A is of BV.

Key example will be solving the stochastic differential equation

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t.$$

Solving this SDE is equivalent as solving the following integral equation

$$X_t = X_0 + \int_0^t \mu(X_s) \, \mathrm{d}s + \int_0^t \sigma(X_s) \, \mathrm{d}B_s \coloneqq A_t + M_t.$$

Now we would like to define the Itô stochastic integral not only for SBM but for martingales, local martingales and semimartingales.

3 Continuous Semimartingale

We turn to the next step of defining a stochastic (Itô) integral with respect to a continuous (square-integrable)$^{(\mathrm{I})}$ (local)$^{(\mathrm{II})}$ martingale $M = (M_t)_{t\geq 0}$:

$$\int_0^t H_s \, \mathrm{d}M_s.$$

Throughout, $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ is a filtered probability space where

$$\mathcal{N} = \{N \in \mathscr{F} \mid \mathbb{P}(N) = 0\} \subseteq \mathscr{F}_t, \quad \forall t \geq 0.$$

3.1 Finite Quadratic Variation

Lemma 3.1. *Let $M = (M_t)_{t\geq 0}$ is a continuous martingale and $t \mapsto M_t$ is of BV (recall, bounded variation), then $M_t = M_0$ $\mathbb{P}$-a.s for all $t \geq 0$.*

Given a process $X = (X_t)_{t\geq 0}$ and a subdivision Δ consising of $0 = t_0 < t_1 < t_2 < \cdots$ (with finitely many t_is in each $[0, t]$), we define

$$T_t^{\Delta}(X) = \sum_{i=1}^{n} \left(X_{t_i} - X_{t_{i-1}}\right)^2 + (X_t - X_{t_n})^2 \tag{3.1}$$

where $t_n < t \leq t_{n+1}$.

Definition 3.2. X is said to be of finite quadratic variation, if there exist a process $\langle X, X\rangle$, such that

$$T_t^{\Delta}(X) \longrightarrow \langle X, X\rangle_t \quad \text{in } \mathbb{P}\text{-probability, for all } t \geq 0.$$

as

$$|\Delta| = \max_{1\leq i\leq n} (t_i - t_{i-1}) \longrightarrow 0.$$

We have the following *main result*:

Theorem 3.3.

1. *A continuous bounded martingale M is of finite quadratic variation.*
2. *The process $\langle M, M\rangle$ is the unique continuous increasing adapted process which is* 0 *at* 0*, such that*
$$M^2 - \langle M, M\rangle$$
is a martingale. (Here M^2 is the square of the process M)

Example 3.4. Recall Example 2.11. B_t is a martingale, then $B_t^2 - t$ is also a martingale.

To enlarge the scope of the preceding result, we will need

Proposition 3.5. *If M is a continuous bounded martingale, then*

$$\langle M^\tau, M^\tau \rangle = \langle M, M \rangle^\tau$$

for every stopping time τ. Here M^τ means "stopping the process M at τ", where $M = (M_t)_{t\geq 0}$, τ is a stopping time and

$$M^\tau = (M_t^\tau)_{t\geq 0}, \quad M_t^\tau := M_{t\wedge\tau}.$$

3.2 Continuous Local Martingale

Notice that, Theorem 3.3 does not cover the case of BM. To do so, and much more, we introduce the following.

Definition 3.6. An adapted continuous process X with $X_0 \in L^1(|X_0| < \infty$, otherwise $X := X - X_0)$ is a *local martingale* if there exists a sequence of stopping time $(\tau_n)_{n\geq 1}$ such that

(i) $\tau_n \uparrow \infty$ as $n \longrightarrow \infty$.

(ii) $X^{\tau_n} (= (X_{t\wedge\tau_n})_{t\geq 0})$ is a martingale for all $n \geq 1$.

Remark 3.7.

1. Setting $\tau_n := \tau_n \wedge \sigma_n$ with

$$\sigma_n = \inf\{t > 0 \mid |X_t| \geq n\}.$$

   we see that it is no restriction to assume in (ii) that X^{τ_n} is a bounded martingale.

2. When (ii) holds, we say that τ_n reduces X.

3. We denote $\mathcal{M}_{\text{loc}}^C$ be all the continuous local martingale.

4. Each martingale is a local martingale, however, local martingale are much more general than martingale!

Theorem 3.8 (The Quadratic Variation of a Continuous Local Martingale)**.** *If $M \in \mathcal{M}_{\text{loc}}^C$, then there exists a unique continuous increasing adapted process $\langle X, X \rangle$ which is 0 at 0 such that*

$$M^2 - \langle X, X \rangle$$

is a continuous local martingale, i.e. $M^2 - \langle X, X \rangle \in \mathcal{M}_{\text{loc}}^C$. Moreover, for each $t > 0$ and any sequence of subdivision $(\Delta_n)_{n\geq 1}$ of $[0, t]$ such that $|\Delta_n| \longrightarrow 0$ as $n \longrightarrow \infty$, we have

$$\sup_{0\leq s\leq t} \left| T_s^{\Delta_n}(M) - \langle M, M \rangle_s \right| \longrightarrow 0 \quad \text{in } \mathbb{P}\text{-probability}$$

as $n \longrightarrow \infty$. The definition of $T_s^{\Delta_n}(M)$ can be found in (3.1).

THEOREM 4.9 OF [2]

Let $M = (M_t)_{t\geq 0}$ be a continuous local martingale. There exist an unique increasing process denoted by $(\langle M, M\rangle_t)_{t\geq 0}$, such that $M_t^2 - \langle M, M\rangle_t$ is a continuous local martingale.
Furthermore, for every fixed $t > 0$, if $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$ is an increasing sequence of subdivision of $[0, t]$ with mesh tending to 0, we have

$$\langle M, M\rangle_t = \lim_{n\to\infty} \sum_{i=1}^{p_n} (M_{t_i^n} - M_{t_{i-1}}^n)^2$$

in probability. The process $\langle M, M\rangle$ is called the quadratic variation of M.

Definition 3.9. The process $\langle M, M\rangle$ is called the increasing process with M, the increasing process of M, or the quadratic variation (process) of M.

We further extend the previous result by 'polarization'.

Theorem 3.10. *If $M, N \in \mathcal{M}_{\text{loc}}^C$, then there exist a unique continuous* ***BV*** *adapted process $\langle M, N\rangle$ which is 0 at 0 such that:*

$$MN - \langle M, N\rangle$$

is a continuous local martingale. Moreover, for each $t > 0$ and any sequence of subdivision $(\Delta_n)_{n\geq 1}$ of $[0, t]$ such that $|\Delta_n| \to 0$ as $n \to \infty$, we have:

$$\sup_{0\leq s\leq t} \left|T_s^{\Delta_n}(M, N) - \langle M, N\rangle_s\right| \to 0 \quad \text{in } \mathbb{P}\text{-probability}$$

as $n \to \infty$, where we set

$$T_s^{\Delta_n}(M, N) = \sum_{i=1}^{n} (M_{t_i}^s - M_{t_{i-1}}^s)(N_{t_i}^s - N_{t_{i-1}}^s),$$

for $0 \leq t_0 < t_1 < \cdots < t_n \leq t$ from Δ_n, $(n \geq 1)$.

Proposition 3.11. *If $M, N \in \mathcal{M}_{\text{loc}}^C$, then*

$$\langle M^\tau, N^\tau\rangle = \langle M, N^\tau\rangle = \langle M, N\rangle^\tau,$$

for every stopping time τ.

Note:

- $M^\tau N^\tau - \langle M, N\rangle^\tau \in \mathcal{M}_{\text{loc}}^C$,
- $M^\tau(N - N^\tau) \in \mathcal{M}_{\text{loc}}^C$, and
- $M^\tau N^\tau - \langle M, N\rangle^\tau + M^\tau(N - N^\tau) = M^\tau N - \langle M, N\rangle^\tau \in \mathcal{M}_{\text{loc}}^C$.

3.3 Continuous Semimartingale

The following inequality will be useful in defining the stochastic integral; in particular

$$\mathrm{d}\langle M, N\rangle \ll \mathrm{d}\langle M, M\rangle$$

for $M, N \in \mathcal{M}_{\text{loc}}^{C}$.

Absolutely continuous

Let μ, ν be two measures on a σ-algebra $\mathscr{G}$, then ν is *absolutely continuous* with respect to μ if for any $A \in \mathscr{G}$ such that $\mu(A) = 0$, we have $\nu(A) = 0$. We denote this relationship as $\nu \ll \mu$.

Recall that $H = (H_t)_{t\geq 0}$ is measurable if $(t, \omega) \to H_t(\omega)$ is $\mathscr{B}(\mathbb{R}_+) \times \mathscr{F}$ measurable.

Proposition 3.12. *For $M, N \in \mathcal{M}_{\text{loc}}^{C}$ and H, K be measurable process, we have*

$$\int_0^t |H_s||K_s||\,\mathrm{d}\langle M, N\rangle_s| \leq \left(\int_0^t H_s^2\,\mathrm{d}\langle M, M\rangle_s\right)^{1/2} \left(\int_0^t K_s^2\,\mathrm{d}\langle N, N\rangle_s\right)^{1/2}$$

with $\mathbb{P}$-a.s. for each $0 \leq t \leq \infty$.

Corollary 3.13 (The Kunitu-Watanabe inequality)**.** *For $M, N \in \mathcal{M}_{\text{loc}}^{C}$, H, K are measurable process, $p, q > 1$ with $p^{-1} + q^{-1} = 1$, we have*

$$\mathbb{E}\left[\int_0^\infty |H_s||K_s|\,|\mathrm{d}\langle M, N\rangle_s|\right] \leq \left\|\left(\int_0^\infty H_s^2\,\mathrm{d}\langle M, M\rangle_s\right)^{1/2}\right\|_p \left\|\left(\int_0^\infty K_s^2\,\mathrm{d}\langle N, N\rangle_s\right)^{1/2}\right\|_q$$

Holder's inequality

For $p, q > 1$ and $p^{-1} + q^{-1} = 1$, we have

$$\mathbb{E}[|XY|] \leq \left(\mathbb{E}\left[|X|^p\right]\right)^{1/p} \left(\mathbb{E}\left[|Y|^q\right]\right)^{1/q}.$$

To simplify the notation, we write $\|X\|_p = \mathbb{E}\left[|X|^p\right]^{1/p}$. For the special case $p = q = 2$, it is the Cauchy-Schwartz inequality.

Definition 3.14. A **continuous semimartingale** is a continuous process X which can be written as

$$X = M + A$$

where $M \in \mathcal{M}_{\text{loc}}^{C}$ and A is a continuous adapted process of BV. This decomposition is unique.

Proposition 3.15. *A continuous semimartingale $X = M + A$ has a finite quadratic variation, where*

$$\langle X, X\rangle = \langle M, M\rangle.$$

Proof. If $\varDelta$ is a subdivision of $[0, t]$ consisting of points $0 = t_0 < t_1 < \cdots < t_n = t$, note that

$$\left|\sum_i (M_{t_i} - M_{t_{i-1}})(A_{t_i} - A_{t_{i-1}})\right| \le \sup_i \left|M_{t_i} - M_{t_{i-1}}\right| \cdot \mathcal{V}(A; [0, t]).$$

The first term on the right hand side tends to 0 as $|\varDelta| \to 0$ by continuity of M, and the second term is finite since A is of bounded variation on $[0, t]$.

Likewise

$$\lim_{|\varDelta|\to 0} \sum_i (A_{t_i} - A_{t_{i-1}})^2 = 0.$$

It follows that

$$\begin{aligned}\lim_{|\varDelta|\to 0} \sum_i (X_{t_i} - X_{t_{i-1}})^2 &= \lim_{|\varDelta|\to 0} \sum_i \left((M_{t_i} - M_{t_{i-1}}) + (A_{t_i} - A_{t_{i-1}})\right)^2 \\ &= \lim_{|\varDelta|\to 0} \sum_i \left(M_{t_i} - M_{t_{i-1}}\right)^2 \quad \text{in } \mathbb{P}\text{-probability.}\end{aligned}$$

□

Definition 3.16. If $X = M + A$ and $Y = N + B$ are continuous semimartingale, we define the bracket of X and Y by

$$\langle X, Y\rangle = \langle M, N\rangle = \frac{1}{4}\left(\langle X + Y, X + Y\rangle - \langle X - Y, X - Y\rangle\right). \tag{3.2}$$

Obviously,

$$\langle X, Y\rangle_t = \mathbb{P}\text{-}\lim_{|\varDelta|\to 0} \sum_i (X_{t_i} - X_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}}).$$

3.3.1 From $\mathcal{M}_{2,\infty}^{C,0}$ to $\mathcal{M}_{\mathrm{loc}}^{C}$

In order to define the stochastic (Itô) integral:

$$\int_0^t H_s \,\mathrm{d}M_s,$$

we will introduce two spaces of process:

$$L^2(M) \ni H, \quad \mathcal{M}_{2,\infty}^{C} \ni M.$$

The definition will be extended to $M \in \mathcal{M}_{\mathrm{loc}}^{C}$ by localization (also to more general H).

Firstly, let $\mathcal{M}_{2,\infty}^{C}$ denote the class of continuous martingale M that are L^2-bounded, i.e.

$$\sup_{t\geq 0}\mathbb{E}[M_t^2] < \infty.$$

By Doob's maximal inequality,

$$M \in \mathcal{M}_{2,\infty}^{C} \longrightarrow M_\infty^* := \sup_{t\geq 0}|M_t| \in L^2$$

which means M is uniformly integrable.

Let M_∞ be the L^2 limit of M_t as $t \longrightarrow \infty$, namely,

$$L^2 - \lim_{t\longrightarrow\infty} M_t = M_\infty, \quad \mathbb{P}\text{-a.s.},$$

then $M_t = \mathbb{E}[M_\infty \mid \mathscr{F}_t]$.

Proposition 3.17. *$\mathcal{M}_{2,\infty}^{C}$ is a Hilbert space with the inner product*

$$(M,N)_{2,\infty} = \mathbb{E}[M_\infty N_\infty] = \lim_{n\longrightarrow\infty}\mathbb{E}\,[M_t N_t].$$

Therefore, the norm is defined by $\|M\|_{2,\infty}^2 = (M,M)_{2,\infty}$,

$$\|M\|_{2,\infty} = \sqrt{\mathbb{E}[M_\infty^2]} = \lim_{t\longrightarrow\infty}\sqrt{\mathbb{E}[M_t^2]}$$

for $M \in \mathcal{M}_{2,\infty}^{C}$.

Proposition 3.18. *Let $M \in \mathcal{M}_{\mathrm{loc}}^{C}$,*

$$M \in \mathcal{M}_{2,\infty}^{C} \quad \Leftrightarrow \quad M_0 \in L^2,\ \mathbb{E}\,[\langle M,M\rangle_\infty] < \infty.$$

In this case, $M^2 - \langle M,M\rangle$ is uniformly integrable and we have

$$\begin{aligned}\mathbb{E}\left[M_\tau^2 - M_\sigma^2 \mid \mathscr{F}_\sigma\right] &= \mathbb{E}\left[(M_\tau - M_\sigma)^2 \mid \mathscr{F}_\sigma\right]\\ &= \mathbb{E}\,[\langle M,M\rangle_\tau - \langle M,M\rangle_\sigma \mid \mathscr{F}_\sigma]\end{aligned}$$

for any pair of stopping time $\sigma \leq \tau$.

Corollary 3.19. *If $M \in \mathcal{M}_{2,\infty}^{C,0} = \{M \in \mathcal{M}_{2,\infty}^{C} \mid M_0 = 0\}$, then*

$$\|M\|_{2,\infty} = \sqrt{\mathbb{E}[\langle M,M\rangle_\infty]} = \left\|(\langle M,M\rangle_\infty)^{1/2}\right\|_2.$$

Secondly, given $M \in \mathcal{M}^C_{2,\infty}$, let $L^2(M)$ denote the space of all progressively measurable processes H such that

$$\|H\|^2_{2,M} = \mathbb{E}\left[\int_0^\infty H_s^2 \,\mathrm{d}\langle M, M\rangle_s\right] < \infty.$$

(progressively measurable means H is measurable with respect to $[0,t] \times \Omega$ for any t fixed)

Define a finite measure

$$\mathbb{P}_M : \mathscr{B}(\mathbb{R}_+) \times \mathscr{F}_\infty \to \mathbb{R}_+$$

by

$$\mathbb{P}_M(A) = \mathbb{E}\left[\int_0^\infty \mathbb{1}_A \,\mathrm{d}\langle M, M\rangle_s\right]$$

Then we can see that $L^2(M)$ is nothing else but the space of all $\mathbb{P}_M$-square integrable progressively measurable functions from $\mathbb{R}_+ \times \Omega$ to $\mathbb{R}$. Hence

- $(L^2(M), \|\cdot\|_{2,M})$ is a Hilbert space.
- In particular, $L^2(M)$ includes all bounded continuous adapted processes.

Theorem 3.20. *Let $M \in \mathcal{M}^C_{2,\infty}$, then for each $H \in L^2(M)$, there exist a unique $H \cdot M \in \mathcal{M}^{C,0}_{2,\infty}$ such that*

$$\langle H \cdot M, N\rangle = H \cdot \langle M, N\rangle$$

for every $N \in \mathcal{M}^C_{2,\infty}$. The mapping $H \to H \cdot M$ is an isometry from $L^2(M)$ into $\mathcal{M}^{C,0}_{2,\infty}$.

Definition 3.21. The martingale $H \cdot M$ is called the stochastic (or Itô) integral of H with respect to M and is also denoted by

$$\left(\int_0^t H_s \,\mathrm{d}M_s\right)_{t \geq 0}.$$

Lemma 3.22. *If $M \in \mathcal{M}^C_{\text{loc}}$, then $\langle M, M\rangle = 0$ if and only if $M_t = M_0$ $\mathbb{P}$-a.s. for all $t \geq 0$.*

Lemma 3.23. *A Càdlàg (right continuous with left limits) adapted process X is a martingale if and only if for each bounded stopping time tau, we have:*

i. $X_\tau \in L^1$,

ii. $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$.

Remark 3.24.

1. We stress that the fact that the stochastic integral $H \cdot M$ vanishes at 0, i.e. $(H \cdot M)_0 = 0$.

2. The key reason for calling $H \cdot M$ a stochastic integral will now be explained. Let $\mathscr{E}$ denote the space of elementary processes, i.e. the processes H which can be written as
$$H_t(\omega) = \sum_{i \geq 1} h_{i-1}(\omega) \mathbb{1}_{(t_{i-1}, t_i]}(t)$$
where $h_{i-1}(\omega)$ is $\mathscr{F}_{t_{i-1}}$ measurable and uniformly bounded, and $0 = t_0 < t_1 < \cdots$, and $\lim_{n \to \infty} t_n \to \infty$.
3. If $M \in \mathcal{M}^C_{2,\infty}$, then $\mathscr{E} \subseteq L^2(M)$. For $H \in \mathscr{E}$, we define the *elementary stochastic integral* $H \cdot M$ by setting:
$$(H \cdot M)_t = \sum_{i=1}^{n} h_{i-1}(M_{t_i} - M_{t_{i-1}}) + h_n(M_t - M_{t_n})$$
whenever $t_n < t \leq t_{n+1}$.
4. It is easy to see that $H \cdot M \in \mathcal{M}^{C,0}_{2,\infty}$. Moreover, considering subdivisions $\varDelta$ including the t_is, it can be proved using the definition of the brackets that for any $N \in \mathcal{M}^C_{2,\infty}$, we have
$$\langle H \cdot M, N \rangle = H \cdot \langle M, N \rangle$$
5. Consequently, the elementary stochastic integral coincides with the stochastic integral constructed in Theorem 3.20

Below, we assume that $M \in \mathcal{M}^C_{2,\infty}$,

Proposition 3.25. *If $H \in L^2(M)$ and $K \in L^2(H \cdot M)$, then $KH \in L^2(M)$ and we have*
$$K \cdot (H \cdot M) = (KH) \cdot M.$$

The next result show how stochastic integration behaves with respect to stopping time; this will be important to enlarge the scope of its definition from $\mathcal{M}^C_{2,\infty}$ to $\mathcal{M}^C_{\text{loc}}$.

Proposition 3.26. *If τ is a stopping time, then we have:*
$$H \cdot M^\tau = (H \mathbb{1}_{[0,\tau]}) \cdot M = (H \cdot M)^\tau.$$

We now extend the definition of $\int_0^t H_s \, \mathrm{d}M_s$ to $M \in \mathcal{M}^C_{\text{loc}}$, and to more general H than $L^2(M)$.

Definition 3.27. If $M \in \mathcal{M}^C_{\text{loc}}$, we define $L^2_{\text{loc}}(M)$ to be the space of all progressively measurable process H for which there exist stopping time $\tau_n \uparrow +\infty$ as $n \to \infty$ such that
$$\mathbb{E}\left[\int_0^{\tau_n} H_s^2 \, \mathrm{d}\langle M, M \rangle_s\right] < \infty, \quad \forall n \geq 1. \tag{$\star$}$$

Note that ($\star$) is equivalent to
$$\mathbb{P}\left(\int_0^t H_s^2 \, \mathrm{d}\langle M, M \rangle_s < \infty\right) = 1, \quad \forall t \geq 0. \tag{$\star\star$}$$

Theorem 3.28. *Let* $M \in \mathcal{M}^C_{\text{loc}}$*, then for each* $H \in L^2_{loc}(M)$*, there exists a unique* $H \cdot M \in \mathcal{M}^{C,0}_{loc}$ *such that*

$$\langle H \cdot M, N \rangle = H \cdot \langle M, N \rangle$$

for every $N \in \mathcal{M}^C_{\text{loc}}$*.*

Definition 3.29. The continuous local martingale $H \cdot M$ is called the stochastic (Itô) integral of H with respect to M and also is denoted by

$$\left(\int_0^t H_s \, \mathrm{d}M_s \right)_{t \geq 0}$$

Notice that, here $M_s \in \mathcal{M}^C_{\text{loc}}$ and the whole process $\in \mathcal{M}^C_{\text{loc}}$.

3.3.2 From $\mathcal{M}^C_{\text{loc}}$ to Continuous Semimartingale

Definition 3.30. A progressively measurable process H is *locally bounded* if there exist a sequence of stopping time $\tau_n \uparrow \infty$ as $n \to \infty$, and there exist $C_n \geq 0$ for all $n \geq 1$ such that $|H^{\tau_n}| \leq C_n$ for each $n \geq 1$.

Remark 3.31.

- A continuous adapted process is locally bounded.
- Locally bounded processes are always belongs to $L^2_{\text{loc}}(M)$ for each $M \in \mathcal{M}^C_{\text{loc}}$.

Definition 3.32. If H is locally bounded and $X = M + A$ is a continuous semimartingale, then the stochastic integral of H with respect to X is the continuous semimartingale

$$H \cdot X = H \cdot M + H \cdot A.$$

where $H \cdot M$ is the stochastic (Itô) integral of Theorem 3.28 and $H \cdot A$ is the pathwise LS integral with respect to $\mathrm{d}A$. The semimartingale $H \cdot X$ is also written as

$$\left(\int_0^t H_s \, \mathrm{d}X_s \right)_{t \geq 0}.$$

Here if $X_s \in \mathcal{M}^C_{\text{loc}}$, i.e. $A = 0$, then the whole integral process $\in \mathcal{M}^C_{\text{loc}}$. Similarly, If $M = 0$, then $X = A$ is of BV (Definition 3.14), then the whole integral process is a continuous process of BV.

Proposition 3.33. *The mapping* $H \to H \cdot X$*, where* H *is locally bounded and* X *is a continuous semimartingale, satisfies the following properties:*

1. $K \cdot (H \cdot X) = (KH) \cdot X$. *(Proposition 3.25)*
2. $(H \cdot X)^\tau = (H \mathbb{1}_{[0,\tau]}) \cdot X = H \cdot X^\tau$ *for every stopping time* τ*. (Proposition 3.26)*

3. *If $H \in \mathscr{E}$, then*

$$(H \cdot X)_t = \sum_{i=1}^{n} h_{i-1}(X_{t_i} - X_{t_{i-1}}) + h_n(X_t - X_{t_n})$$

whenever $t_n < t \leq t_{n+1}$, where h_{i-1} is $\mathscr{F}_{t_{i-1}}$ measurable and uniformly integrable.

We turn to a very important property of the stochastic integral

Theorem 3.34 (Stochastic Dominated Convergence Theorem)**.** *Let X be a continuous semimartingale and let $(H^n)_{n\geq 1}$ be a sequence of locally bounded process satisfying:*

(i) $H^n \to 0$ pointwise on $\mathbb{R}_+ \times \Omega$, i.e.

$$H_t^n(\omega) \to 0, \quad \textit{for all } t \geq 0 \textit{ and } \omega \in \Omega, \textit{ as } n \to \infty.$$

(ii) There exist a locally bounded process H such that $|H^n| \leq H, \quad \forall n \geq 1$, i.e.

$$\left|H_t^n(\omega)\right| \leq H_t(\omega) \quad \textit{for all } t \geq 0 \textit{ and } \omega \in \Omega, \textit{ as } n \to \infty.$$

Notice that, H does not depends on n, and we only require it to be locally bounded and dominate H^n.

Then $H^n \cdot X \to 0$ in $\mathbb{P}$-probability uniformly on each compact interval, i.e.

$$\mathbb{P}\text{-}\lim_{n\to\infty} \sup_{0\leq s\leq t} |(H^n \cdot X)_s| = 0, \quad \forall t \geq 0. \tag{SDCT}$$

Remark 3.35. The above theorem states that: The limit of the stochastic integral is the integral of the limit under some conditions. Equation (SDCT) can be viewed as

$$\mathbb{P}\text{-}\lim_{n\to\infty} \sup_{0\leq s\leq t} \left| \int_0^s H_s^n \, \mathrm{d}X_s \right| = \sup_{0\leq s\leq t} \left| \int_0^s \mathbb{P}\text{-}\lim_{n\to\infty} H_s^n \, \mathrm{d}X_s \right| = 0.$$

Proposition 3.36. *If X is continuous semimartingale, and H is locally bounded and left continuous, then:*

$$\int_0^t H_s \, \mathrm{d}X_s = \mathbb{P}\text{-}\lim_{n\to\infty} \sum_{\Delta_n} H_{t_{i-1}}(X_{t_i} - X_{t_{i-1}})$$

for each sequence of subdivisions Δ_n of $[0, t]$ such that $|\Delta_n| \to 0$ as $n \to \infty$.

4 Stochastic Integration

We derive the celebrated Itô's formula, which shows that the image of one or several continuous semimartingales under a smooth function is still a continuous semimartingale, whose canonical decomposition is given in terms of stochastic integrals. Itô's formula is the main technical tool of stochastic calculus, and we discuss several important applications of this formula, including Lèvy's theorem characterizing Brownian motion as a continuous local martingale with quadratic variation process equal to t, and the Burkholder-Davis-Gundy inequalities. The end of the section is devoted to Girsanov's theorem, which deals with the stability of the notions of a martingale and a semimartingale under an absolutely continuous change of probability measure.

4.1 Construction of Itô Formula

Assumption: Throughout, $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ is a filtered probability space where

$$\mathcal{N} = \{N \in \mathscr{F} \mid \mathbb{P}(N) = 0\} \subseteq \mathscr{F}_t \ (\forall t \geq 0).$$

Itô formula states that a "smooth" function of a continuous semimartingale is a continuous semimartingale and provides the decomposition of the latter.

Theorem 4.1 (Itô (1994); Kunite & Watanabe (1967))**.** *Let $X = M + A$ be a continuous semimartingale, and let $F : \mathbb{R} \to \mathbb{R}$ be a C^2 function*[3]*. Then we have:*

$$\begin{aligned} F(X_t) &= F(X_0) + \int_0^t F'(X_s)\, \mathrm{d}X_s + \frac{1}{2}\int_0^t F''(X_s)\, \mathrm{d}\langle X, X\rangle_s \\ &= F(X_0) + \int_0^t F'(X_s)\, \mathrm{d}M_s + \int_0^t F'(X_s)\, \mathrm{d}A_s + \frac{1}{2}\int_0^t F''(X_s)\, \mathrm{d}\langle M, M\rangle_s, \end{aligned} \tag{4.1}$$

for all $t \geq 0$.

Note. A_s and $\langle M, M\rangle_s$ are of BV ($\langle M, M\rangle_s$ is an increasing function, hence it is of BV). Therefore, in the equation (4.1), the first, third and fourth term are of BV term, and the second term is the continuous local martingale part (which makes the $F(X_t)$ becomes a continous semimartingale).

Proof. We divide the proof in three steps.

Step 1 (Localization): (Prove that WLOG, all the terms appearing in the Itôś formula can be assumed to be bounded)

$$X = M_t + A_t = (M_t - M_0) + (A_t - A_0) + (M_0 + A_0) := \widetilde{M}_t + \widetilde{A}_t + X_0.$$

We see that there is no loss of generality to assume in the sequal that $M_0 = A_0 = 0$.

[3]Twice continuously differentiable

We introduce the stopping time:

$$\tau = \begin{cases} 0 & \text{if } |X_0| > n, \\ \inf\{t > 0 : |M_t| \geq n \text{ or } \widetilde{A}_t \geq n \text{ or } \langle M, M\rangle_t \geq n\} & \text{if } |X_0| \leq n \end{cases}$$

where $\inf(\phi) = \infty$ and $\widetilde{A}_t = A_t^+ + A_t^-$ is the total variation of $s \mapsto A_s$ on $[0, t]$. (i.e. $A_t = A_t^+ - A_t^-$ where $t \mapsto A_t^+$ and $t \mapsto A_t^-$ are increasing continuous functions that start from zero). Then $\tau \uparrow \infty$ as $n \to \infty$. Thus, if we can estabilish (4.1) for the stopped process X^{τ_n}, then we obtain (4.1) for the general process X by letting $k \to \infty$. We may thus assume that $X_0, M, \widetilde{A}$ and $\langle M, M\rangle$ are all bounded by a common constant K. Under this assumption, we have that $|X| \leq 3K$, so that the values of F outside $[-3K, 3K]$ are irrelevant. We may then WLOG assume that F has compact support which implies also that F, F' and F'' are bounded.

Step 2 (Taylor Expansion): Fix $t > 0$ and let $\varDelta_n = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ be a subdivision of $[0, t]$. A Taylor expansion yields:

$$\begin{aligned} F(X_t) - F(X_0) &= \sum_{i=1}^{n} F(X_{t_i}) - F(X_{t_{i-1}}) \\ &= \sum_{i}^{n} F'(X_{t_{i-1}})(X_{t_i} - X_{t_{i-1}}) + \frac{1}{2}\sum_{i=1}^{n} F''(\xi_i)(X_{t_i} - X_{t_{i-1}})^2 \\ &=: I_t(\varDelta_n) + \frac{1}{2} J_t(\varDelta_n), \end{aligned}$$

where $\xi_i = X_{t_{i-1}} + \theta_i(X_{t_i} - X_{t_{i-1}})$, for some $0 \leq \theta_i \leq 1$, i.e. ξ_i lies between $X_{t_{i-1}}$ and X_{t_i}. Notice that, here $X_{t_{i-1}}$ is not necessarily smaller than X_{t_i}.

By Proposition 3.36, we know that:

$$I_t(\varDelta_n) \to \int_0^t F'(X_s)\,\mathrm{d}X_s \quad \text{in } \mathbb{P}\text{-probability} \tag{i}$$

as $|\varDelta_n| \to 0$, since $F'(X) = (F'(X_s))_{s\geq 0}$ is locally bounded and continuous.

Step 3 (The quadratic variation term)($*$): (We would like to prove that $F''(\xi_i) \to F''(X_s)$ and $(X_{t_i} - X_{t_{i-1}})^2 \to \langle M, M\rangle_s$)

Note that $J_t(\varDelta_n)$ can be written:

$$J_t(\varDelta_n) = J_t^1(\varDelta_n) + J_t^2(\varDelta_n) + J_t^3(\varDelta_n)$$

where we set

$$\begin{aligned} J_t^1(\varDelta_n) &= \sum_{i=1}^{n} F''(\xi_i)\left(M_{t_i} - M_{t_{i-1}}\right)^2, \\ J_t^2(\varDelta_n) &= \sum_{i=1}^{n} 2F''(\xi_i)\left(M_{t_i} - M_{t_{i-1}}\right)\left(A_{t_i} - A_{t_{i-1}}\right), \\ J_t^3(\varDelta_n) &= \sum_{i=1}^{n} F''(\xi_i)\left(A_{t_i} - A_{t_{i-1}}\right)^2. \end{aligned}$$

First note that:

$$\left|J_t^2(\Delta_n)\right| \le 2\,\|F''\|_\infty \cdot \underbrace{\max_{1\le i\le n}\left|M_{t_i}-M_{t_{i-1}}\right|}_{\to 0} \cdot \widetilde{A}_t \to 0, \tag{ii}$$

$$\left|J_t^3(\Delta_n)\right| \le \|F''\|_\infty \cdot \underbrace{\max_{1\le i\le n}\left|A_{t_i}-A_{t_{i-1}}\right|}_{\to 0} \cdot \widetilde{A}_t \to 0, \tag{iii}$$

$\mathbb{P}$-a.s. and thus $\mathbb{P}$-probability as well as $|\Delta_n| \to 0$.

It remain to estabilish:

$$J_t^1(\Delta_n) \to \int_0^t F''(X_s)\,\mathrm{d}\langle M, M\rangle_s \quad \text{in } \mathbb{P}\text{-probability}$$

as $|\Delta_n| \to 0$. If we finish this, then by (i)–(iii), it follows that (4.1) holds $\mathbb{P}$-a.s. for $t > 0$ given and fixed, but also for all $t \ge 0$ outside a single $\mathbb{P}$-null set. Noting that both the right-hand side and the left-hand side of (4.1) define continuous process of t.

Define

$$\widetilde{J}_t^1(\Delta_n) = \sum_{i=1}^n F''(X_{t_{i-1}})(M_{t_i}-M_{t_{i-1}})^2.$$

At this point, you are allowed to choose the end point, X_{t_i}, instead of $X_{t_{i-1}}$, and this will give the same results in (∗∗). However, (i) will not necessarily correct, since it is derived from the Itô construction which specifically using the left-hand point.

Based on this choice of endpoint, we have,

$$\left|J_t^1(\Delta_n) - \widetilde{J}_t^1(\Delta_n)\right| \le \underbrace{\max_{1\le i\le n}\left|F''(\xi_i)-F''(X_{t_{i-1}})\right|}_{\to 0} \times \underbrace{\sum_{i=1}^n \left(M_{t_i}-M_{t_{i-1}}\right)^2}_{\to \langle M,M\rangle_t \text{ in } \mathbb{P}\text{-probability}}$$

as $|\Delta_n| \to 0$. The latter term is by Theorem 3.3. Thus, it is enough to show that

$$\widetilde{J}_t^1(\Delta_n) \to \int_0^t F''(X_s)\,\mathrm{d}\langle M, M\rangle_s \quad \text{in } \mathbb{P}\text{-probability}, \tag{$**$}$$

as $|\Delta_n| \to 0$. The latter one is the usual LS-integral, hence we can introduce the approximating sum for this integral as follows:

$$K_t^1(\Delta_n) = \sum_{i=1}^n F''(X_{t_{i-1}})\left(\langle M, M\rangle_{t_i} - \langle M, M\rangle_{t_{i-1}}\right).$$

It is enough to show that

$$\left|\widetilde{J}_t^1(\Delta_n) - K_t^n(\Delta_n)\right| \to 0 \quad \text{in } \mathbb{P}\text{-probability} \tag{$***$}$$

as $|\varDelta_n| \to 0$.

For this, note that we have:

$$
\begin{aligned}
&\mathbb{E}\left[\left|\widetilde{J}_t^1(\varDelta_n) - K_t^1(\varDelta_n)\right|^2\right] \\
&= \mathbb{E}\left[\left(\sum_{i=1}^n F''(X_{t_{i-1}})\left((M_{t_i} - M_{t_{i-1}})^2 - (\langle M, M\rangle_{t_i} - \langle M, M\rangle_{t_{i-1}})\right)\right)^2\right] \\
&\overset{(\perp)}{=} \sum_{i=1}^n \mathbb{E}\left[(F''(X_{t_{i-1}}))^2 \left((M_{t_i} - M_{t_{i-1}})^2 - (\langle M, M\rangle_{t_i} - \langle M, M\rangle_{t_{i-1}})\right)^2\right] \\
&\le \|F''\|_\infty^2 \sum_{i=1}^n \left\{\mathbb{E}\left[(M_{t_i} - M_{t_{i-1}})^4\right] + \mathbb{E}\left[(\langle M, M\rangle_{t_i} - \langle M, M\rangle_{t_{i-1}})^2\right]\right\} \\
&\le \|F''\|_\infty^2 \Bigg(\mathbb{E}\Bigg[\underbrace{\max_{1\le i\le n}(M_{t_i} - M_{t_{i-1}})^2}_{\to 0 \text{ (continuous)}} \underbrace{\sum_{i=1}^n (M_{t_i} - M_{t_{i-1}})^2}_{\to \langle M,M\rangle_t \text{ in } \mathbb{P}\text{-probability}}\Bigg] \\
&\qquad + \mathbb{E}\Bigg[\underbrace{\max_{1\le i\le n}(\langle M, M\rangle_{t_i} - \langle M, M\rangle_{t_{i-1}})}_{\to 0} \underbrace{\sum_{i=1}^n (\langle M, M\rangle_{t_i} - \langle M, M\rangle_{t_{i-1}})}_{=\langle M,M\rangle_t}\Bigg]\Bigg) \\
&\to 0, \quad \text{by dominant convergence theorem.}
\end{aligned}
$$

as $|\varDelta_n| \to 0$, where $(\perp)$ above follows by Remark 4.2 below. This estabilishes $(***)$ and completes the proof. □

The following facts about martingales are useful to know, note that (ii) below has used as $(\perp)$ in the proof above.

Remark 4.2. Let M be a square integrable martingale (i.e. $\mathbb{E}[M_t^2] < \infty$ for all $t \ge 0$), and let $0 \le s < t \le u < v$ be given and fixed. Then we have:

(0) $\mathbb{E}[M_v - M_u \mid \mathscr{F}_t] = 0$, so that $\mathbb{E}\left[(M_v - M_u)(M_t - M_s)\right] = 0$.

(i) $\mathbb{E}[(M_v - M_u)^2 \mid \mathscr{F}_t] = \mathbb{E}[M_v^2 - M_u^2 \mid \mathscr{F}_t] = \mathbb{E}[\langle M, M\rangle_v - \langle M, M\rangle_u \mid \mathscr{F}_t]$.

(ii)

$$\mathbb{E}\Big[\underbrace{(M_v - M_u)^2 - (\langle M, M\rangle_v - \langle M, M\rangle_u)}_{:=Z_{uv}} \mid \mathscr{F}_t\Big] = 0,$$

so that we have $\mathbb{E}[Z_{uv}Z_{st}] = 0$ which is used in the proof of Theorem 4.1, and

$$\mathbb{E}[X_u Z_{uv} X_s Z_{st}] \overset{(\perp)}{=} 0.$$

(iii) $\mathbb{E}\big[\underbrace{(M_v^2 - \langle M,M\rangle_v) - (M_u^2 - \langle M,M\rangle_u)}_{:=Y_{uv}} \mid \mathscr{F}_t\big] = 0$, so that $\mathbb{E}[Y_{uv}Y_{st}] = 0$.

Details of (ii)

(ii) is still true if we replace $\mathscr{F}_t$ by $\mathscr{F}_u$, then pre-multiply X_u which is $\mathscr{F}_u$ measurable and post-multiply $X_s Z_{st}$ which are all $\mathscr{F}_u$ measurable, then take the expectation we have the desired $(\perp)$. We used the following properties of the conditional expectation:

- Pulling out known factors: if X is $\mathscr{F}$ measurable, then $\mathbb{E}[XY \mid \mathscr{F}] = \mathbb{E}[Y \mid \mathscr{F}]X$.
- Law of total expectation: $\mathbb{E}[\mathbb{E}[X \mid \mathscr{F}]] = \mathbb{E}[X]$.

We then have the following multidimensional extension of Theorem 4.1.

Theorem 4.3 (Multidimensional Itô formula). *Let $X = M + A$ be a continuous semimartingale with values in $\mathbb{R}^d$, i.e. $X_t^i = M_t^i + A_t^i$ where $M_t^i \in \mathcal{M}_{\text{loc}}^C$ and A_t^i is a continous adapted process of BV for $1 \le i \le d$, and let $F : \mathbb{R}^d \to \mathbb{R}$ be a C^2 function. Then we have:*

$$\begin{aligned} F(X_t) =& F(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X_s)\, \mathrm{d}X_s^i \\ &+ \frac{1}{2}\sum_{i=1}^d \sum_{i=1}^d \int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}(X_s)\, \mathrm{d}\langle X^i, X^j\rangle_s \\ =& F(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X_s)\, dM_s^i + \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X_s)\, dA_s^i \\ &+ \frac{1}{2}\sum_{i=1}^d \sum_{i=1}^d \int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}(X_s)\, \mathrm{d}\langle M^i, M^j\rangle_s \end{aligned} \tag{4.2}$$

for all $t \ge 0$.

Remark 4.4.

1. If some X^i in Theorem 4.3 is of BV, then $x_i \mapsto F(x_1, \ldots, x_i, \ldots, x_d)$ needs only to be C^1 (not C^2) for the formula (4.2) to hold.
2. In particular, if $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is $C^{1,2}$,[4] then:

$$\begin{aligned} F(A_t, X_t) =& F(A_0, X_0) + \int_0^t F_a(A_s, X_s)\, \mathrm{d}A_s \\ &+ \int_0^t F_x(A_s, X_s)\, \mathrm{d}X_s + \frac{1}{2}\int_0^t F_{xx}(A_s, X_s)\, \mathrm{d}\langle X, X\rangle_s \end{aligned} \tag{4.3}$$

[4] First derivative w.r.t. the first argument is exist and continuous; first and second derivatives w.r.t. the second argument are exist and continuous.

where A is a continuous adapted process of BV, and X is a continuous semimartingale.

3. In particular, if $F : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ is $C^{1,2}$, then:

$$
\begin{aligned}
F(t, X_t) =& F(0, X_0) + \int_0^t F_s(s, X_s)\, \mathrm{d}s \\
&+ \int_0^t F_x(s, X_s)\, \mathrm{d}X_s + \frac{1}{2}\int_0^t F_{xx}(s, X_s)\, \mathrm{d}\langle X, X\rangle_s
\end{aligned}
\tag{4.4}
$$

where X is a continuous local martingale.

4.2 Examples of Itô Formula

Example 4.5. $X = B, F(x) = x^2$, then

$$
\underbrace{F(B_t)}_{B_t^2} = \underbrace{F(B_0)}_{0} + \int_0^t \underbrace{F'(B_s)}_{2B_s}\, \mathrm{d}B_s + \frac{1}{2}\int_0^t \underbrace{F''(B_s)}_{2}\, \mathrm{d}\underbrace{\langle B, B\rangle_s}_{ds}
$$

which yields $B_t^2 = 2\int_0^t B_s\, \mathrm{d}B_s + t$ which is a semimartingale.

Example 4.6. $X = B, F(t, x) = x_0 \mathrm{e}^{(\sigma x - \sigma^2 t/2)}$, then:

$$
\begin{aligned}
F(t, B_t) = \underbrace{F(0, B_0)}_{x_0} &+ \cancel{\int_0^t \underbrace{F_t(s, B_s)}_{-\sigma^2 F(s,B_s)/2}\, \mathrm{d}s} + \int_0^t \underbrace{F_x(s, B_s)}\, \sigma F(s, B_s)\, \mathrm{d}B_s \\
&+ \cancel{\frac{1}{2}\int_0^t \underbrace{F_{xx}(s, B_s)}_{\sigma^2 F(s,B_s)}\, \underbrace{\mathrm{d}\langle B, B\rangle_s}_{ds}} \\
&= x_0 + \sigma \int_0^t F(s, B_s)\, \mathrm{d}B_s \in \mathcal{M}_{\mathrm{loc}}^C.
\end{aligned}
$$

Setting $X_t := \mathcal{E}_t^{\sigma}(B) := x_0 \mathrm{e}^{(\sigma B_t - \sigma^2 t/2)}$, we see that:

$$
X_t = x_0 + \sigma \int_0^t X_s\, \mathrm{d}B_s. \tag{i}
$$

This (by defintion) is equivalently as:

$$
\mathrm{d}X_t = \sigma X_t\, \mathrm{d}B_t, \quad (X_0 = x_0). \tag{i'}
$$

Example 4.7. Let $X = B$ and $F(t, x) = x_0 \mathrm{e}^{\sigma x + (\mu - \frac{\sigma^2}{2})t}$, then:

$$
F(t, B_t) = \cdots = x_0 + \mu \int_0^t F(s, B_s)\, \mathrm{d}s + \sigma \int_0^t F(s, B_s)\, \mathrm{d}B_s.
$$

Setting $X_t = \mathrm{e}^{\sigma B_t + (\mu - \frac{\sigma^2}{2})t}$, we see that

$$X_t = x_0 + \mu \int_0^t X_s \, \mathrm{d}s + \sigma \int_0^t X_s \, \mathrm{d}B_s, \tag{i}$$

or equivalently

$$\mathrm{d}X_t = \mu X_t \, \mathrm{d}t + \sigma X_t \, \mathrm{d}B_t, \quad (X_0 = x_0).$$

This process is called the *geometric Brownian motion*. It is used to model a stock price.

Example 4.8. Let X solve the stochastic differential equation (SDE):

$$dX_t = \mu(X_t) \, \mathrm{d}t + \sigma(X_t) \, \mathrm{d}B_t, \quad (X_0 = x_0). \tag{i}$$

This by definition is equivalent to:

$$X_t = x_0 + \int_0^t \mu(X_s) \, \mathrm{d}s + \int_0^t \sigma(X_s) \, \mathrm{d}B_s.$$

Let $F : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ be a $C^{1,2}$ function. Then:

$$\begin{aligned} F(t, X_t) = F(0, X_0) &+ \int_0^t F_s(s, X_s) \, \mathrm{d}s + \int_0^t F_x(s, X_s) \, \mathrm{d}X_s \\ &+ \frac{1}{2} \int_0^t F_{xx}(s, X_s) \, \mathrm{d}\langle X, X \rangle_s. \end{aligned} \tag{ii}$$

Firstly, note by Proposition 3.26 (1) that we have

$$\int_0^t F_x(s, X_s) \, \mathrm{d}X_s = \int_0^t F_x(s, X_s) \mu(X_s) \, \mathrm{d}s + \int_0^t F_x(s, X_s) \sigma(X_s) \, \mathrm{d}B_s \tag{iii}$$

using (i) above. On the other hand, by Proposition 3.15 we have $\langle X, X \rangle = \langle M, M \rangle$ where we set

$$M_t = \int_0^t \sigma(X_s) \, \mathrm{d}B_s.$$

To compute $\langle M, M \rangle$, let us note that by the very definition of the stochastic integral (see Theorem 3.28) that with $M = \sigma(X) \cdot B$ we have

$$\langle M, M \rangle = \langle \sigma(X) \cdot B, \sigma(X) \cdot B \rangle = \sigma(X) \cdot \langle B, \sigma(X) \cdot B \rangle = \sigma^2(X) \cdot \langle B, B \rangle.$$

Recalling that $\langle B, B \rangle_t = t$, we see that

$$\langle M, M \rangle_t = \int_0^t \sigma^2(X_s) \, \mathrm{d}s.$$

It follows that:

$$\int_0^t F_{xx}(s, X_s)\,\mathrm{d}\langle X, X\rangle_s = \int_0^t F_{xx}(s, X_s)\,\mathrm{d}\langle M, M\rangle_s = \int_0^t F_{xx}(s, X_s)\sigma^2(X_s)\,\mathrm{d}s. \qquad (iv)$$

Inserting (iii) and (iv) back in (ii) we get:

$$F(t, X_t) = F(0, X_0) + \int_0^t \left(F_s + \mu F_x + \frac{\sigma}{2}F_{xx}\right)(s, X_s)\,\mathrm{d}s + \int_0^t F_x(s, X_s)\sigma(X_s)\,\mathrm{d}B_s,$$

where we get

$$\mathbb{L}_X F = \frac{\partial F}{\partial t} + \mu\frac{\partial F}{\partial x} + \frac{\sigma}{2}\frac{\partial^2 F}{\partial x^2}.$$

This $\mathbb{L}_X$ is called the *infinitesimal generator of the Markov process* $((t, X_t))_{t\geq 0}$.

Example 4.9 (Exponential (local) martingale)**.** Let $M \in \mathcal{M}^C_{\mathrm{loc}}$ and let $F : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ be a $C^{1,2}$ function such that:

$$F_v + \frac{1}{2}F_{xx} = 0,$$

then $F(\langle M, M\rangle, M) \in \mathcal{M}^C_{\mathrm{loc}}$.

In particular, taking $F(v, x) = \mathrm{e}^{\lambda x - \lambda^2 v/2}$ for $\lambda > 0$ fixed, we see that:

$$\mathcal{E}^\lambda(M) = \mathrm{e}^{\lambda M - \frac{\lambda^2}{2}\langle M,M\rangle}$$

is a continuous local martingale. Take $M = B$, we have Example 4.6.

Proof. Set $A_t = \langle M, M\rangle_t$ and $X_t = M_t$ in the formula (4.3) above, we got

$$\begin{aligned} F(\langle M, M\rangle_t, M_t) = F(0, M_0) &+ \cancel{\int_0^t F_v(\langle M, M\rangle_s, M_s)\,\mathrm{d}\langle M, M\rangle_s} \\ &+ \int_0^t F_x(\langle M, M\rangle_s, M_s)\,\mathrm{d}M_s + \cancel{\frac{1}{2}\int_0^t F_{xx}(\langle M, M\rangle_s, M_s)\,\mathrm{d}\langle M, M\rangle_s}, \end{aligned}$$

proving the claim. □

Example 4.10 (Integration by parts formula)**.** If X and Y are continuous semimartingales, then:

$$X_t Y_t = X_0 Y_0 + \int_0^t X_s\,\mathrm{d}Y_s + \int_0^t Y_s\,\mathrm{d}X_s + \langle X, Y\rangle_t. \qquad (i)$$

In particular, we have:

$$X_t^2 = X_0^2 + 2\int_0^t X_s\,\mathrm{d}X_s + \langle X, X\rangle_t. \qquad (ii)$$

Proof. (i) follows by (4.2) above with $F(x, y) = xy$. (ii) follows either from (i) with $X = Y$, or by (4.1) above with $F(x) = x^2$. □

Review Session

- The equation:

$$X_t = X_0 + \int_0^t \mu(X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}B_s, \quad (X_0 = x_0). \tag{i}$$

  If μ, σ and B are given, then X_t is implicitly given. However, the next equation:

$$\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t, \quad (X_0 = x_0),$$

  is just the short hand notation for (i). This is make sense in classical integration and differentiation due to Newton-Leibniz Formula, but the above equation does not make sense in stochastic world since stochastic differentiation does not exist.

  Hence our stochastic differentials are nothing else but the integrals, and they are exactly the same with different notations.

- If some component of F in the Itô formula (4.2) is of BV, then that component is only C^1 required.
- $\langle X, Y\rangle = 0$ if either X or Y is of BV.

4.3 P. Lévy Characterization Theorem

We now turn to a significant application of the Itô formula.

Recall that standard Brownian motion B is a continuous (local) martingale with $\langle B, B\rangle_t = t$ for all $t \geq 0$. Suppose that we are given a continuous local martingale X such that $\langle X, X\rangle_t = t$ for all $t \geq 0$,

Question. Can we conclude that X is a standard Brownian motion?

The answer is affirmative by Theorem 4.13. In order to present a multidimensional case, we first recall some definitions.

Recall. $B = (B^1, \ldots, B^d)$ is a standard d-dimensional Brownian motion with respect to the filtration $(\mathscr{F}_t)_{t\geq 0}$ if we have

(i) B is a standard d-dimensional Brownian motion, i.e. each B^i is a standard BM for $1 \leq i \leq d$ and $B^1, \ldots, B^d$ are independent.

(ii) B is adapted to $(\mathscr{F}_t)_{t\geq 0}$, i.e. B_t is $\mathscr{F}_t$ measurable for any fixed t.

(iii) $B_t - B_s \perp \mathscr{F}_s$ for every $0 \leq s < t$.

Notice that, (i) implies that $B_t - B_s \sim N(0, (t-s)I_d)$ for every $0 \leq s < t$.

Proposition 4.11. *Let $B = (B^1, \ldots, B^d)$ be a standard d-dimensional BM with respect to $(\mathscr{F}_t)_{t\geq 0}$. Then we have*

$$\langle B^i, B^j\rangle_t = \delta_{ij} \cdot t = \begin{cases} t & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

for $i, j \in \{1, \ldots, d\}$.

Proof. There is no restriction to assume that $d = 2$. Recall from (3.2) that: if $M, N \in \mathcal{M}^C_{\text{loc}}$, then

$$\langle M, N\rangle = \frac{1}{4}\left[\langle M+N, M+N\rangle - \langle M-N, M-N\rangle\right].$$

Since the process:

$$B^1 + B^1 \ \& \ B^1 - B^2$$

are equally distributed as a BM with drift 0 and diffusion coefficient $\sqrt{2}$, it follows

$$\langle B^1 + B^2, B^1 + B^2\rangle_t = \langle B^1 - B^2, B^1 - B^2\rangle_t = 2t.$$

This clearly estabilish the clain for $i \neq j$. Moreover, the process B^1 and B^2 satisfy

$$\langle B^1, B^1\rangle_t = \langle B^2, B^2\rangle_t = t$$

and the proof is complete. □

THE PROCESS $B^1 + B^2$

By Definition 1.1 (4), we have for B^1_t and B^2_t are SBM, $B^1_t, B^2_t \sim N(0,t)$. Let $X_t = B^1_t + B^2_t$, then X_t is a Brownian motion.

$$\mathbb{E}[X_t] = 0, \quad \mathbb{V}[X_t] = \mathbb{V}[B^1_t] + \mathbb{V}[B^2_t] = 2t,$$

therefore, $X_t \sim N(0, 2t)$. Since X_t is also a Brownian motion, this is equivalent as $X_t = \sqrt{2}B_t$.

CHARACTERISTIC FUNCTION (FOURIER TRANSFORMATION)

Given $(\Omega, \mathscr{F}, \mathbb{P})$ as a probability space. Here let

$$\langle u, Z\rangle = \sum_{i=1}^{d} u_i Z_i, \quad \|u\|^2 = (\langle u, u\rangle =) \sum_{i=1}^{d} |u_i|^2.$$

- Let $X : \Omega \to \mathbb{R}$ be a random variable, then $\varphi_X : \mathbb{R} \to \mathbb{C}$ is the characteristic function (Fourier transform) of X, where

$$\varphi_X(t) := \mathbb{E}\left[e^{itX}\right] := \mathbb{E}\left[\cos(tX)\right] + i\mathbb{E}\left[\sin(tX)\right].$$

- Let $X : \Omega \to \mathbb{R}^n$ be a random variable, then $\phi_X : \mathbb{R}^n \to C$ can be defined similarly as

$$\phi_X(t) := \mathbb{E}\left[e^{i\langle t, X\rangle}\right] := \mathbb{E}\left[\cos(\langle t, X\rangle)\right] + i\mathbb{E}\left[\sin(\langle t, X\rangle)\right]$$

where $t = (t_1, \dots, t_n) \in \mathbb{R}^n$.

- $\mathbb{E}\left[e^{i\langle u, Z\rangle} \mid \mathscr{G}\right]$ is the conditional characteristic function.

Lemma 4.12. *Let Z be a random vector defined on $(\Omega, \mathscr{F}, \mathbb{P})$, with values in $\mathbb{R}^d$, and let $\mathscr{G} \subset \mathscr{F}$ be a σ-algbra, suppose that:*

$$\mathbb{E}\left[e^{i\langle u, Z\rangle} \mid \mathscr{G}\right] = e^{-\frac{\sigma^2}{2}\|u\|^2} \tag{i}$$

for each $u \in \mathbb{R}^d$. Then $Z \sim N(0, \sigma^2 I_d)$ and $Z \perp \mathscr{G}$.

Proof. WLOG, let us assume that $d = 1$. Then (i) reads:

$$\mathbb{E}\left[e^{iuZ} \mid \mathscr{G}\right] = e^{-\frac{\sigma^2}{2}u^2}$$

or in other hands:

$$\int_G e^{iuZ} d\mathbb{P} = e^{-\frac{\sigma^2}{2}u^2}\mathbb{P}(G), \quad (\forall G \in \mathscr{G}). \tag{i''}$$

Choosing $G = \Omega$ in (i''), it follows that $Z \sim N(0, \sigma^2)$ by the uniqueness theorem for characteristic functions[5] which proves the first part.

Let us write (i'') for G and G^C from $\mathscr{G}$:

$$\int_G e^{iuZ} d\mathbb{P} = e^{-\frac{\sigma^2}{2}u^2}\mathbb{P}(G) \quad \Rightarrow \quad \int_\Omega e^{iuZ} e^{iv} \mathbb{1}_G d\mathbb{P} = e^{-\frac{\sigma^2}{2}u^2} e^{iv}\mathbb{P}(G) \tag{ii}$$

$$\int_{G^C} e^{iuZ} d\mathbb{P} = e^{-\frac{\sigma^2}{2}u^2}\mathbb{P}(G^C) \tag{iii}$$

[5] $X \sim N(\mu, \sigma^2)$ if and only if $\varphi_X(t) = e^{i\mu t - \frac{\sigma^2}{2}t^2}$ for all $t \in \mathbb{R}$

where $v \in \mathbb{R}$. Adding (*ii*) and (*iii*) we get,

$$\underbrace{\int_\Omega e^{iuZ+iv\mathbb{1}_G} d\mathbb{P}}_{\varphi_{(Z,\mathbb{1}_G)}(u,v)} = \underbrace{e^{-\frac{\sigma^2}{2}u^2}}_{\varphi_Z(u)} \left[\underbrace{e^{iv}\mathbb{P}(G) + \mathbb{P}(G^C)}_{\varphi_{\mathbb{1}_G}(v)} \right]$$

for all $u, v \in \mathbb{R}$. By the well-known theorem for characteristic functions[6] We can conclude that $Z \perp \mathbb{1}_G$, i.e. $Z \perp \mathscr{G}$. □

Theorem 4.13 (P. Lévy (1948))**.** *Let $M = (M^1, \ldots, M^d)$ be a continuous process satisfy*

(i) M is adapted to $(\mathscr{F}_t)_{t\geq 0}$,

(ii) $M^i \in \mathcal{M}^{C,0}_{loc}$, and

(iii) $\langle M^i, M^j\rangle = \delta_{ij} \cdot t$,

for all $i, j \in \{1, \ldots, d\}$. ***Then*** *$M$ is a standard d-dimensional Brownian motion with respect to $(\mathscr{F}_t)_{t\geq 0}$.*

Proof. Given $0 \leq s < t$, in order to prove M is a standard d-dimensional Brownian motion, it is sufficient to prove:

$$M_t - M_s \perp \mathscr{F}_s \quad \text{independent increment,} \tag{iv}$$
$$M_t - M_s \sim N(0, (t-s)I_d) \quad \text{stationary increment.} \tag{v}$$

For this, in view of Lemma 4.12, it is sufficient to show:

$$\mathbb{E}\left[e^{i\langle u, M_t - M_s\rangle} \mid \mathscr{F}_s\right] = e^{-\frac{1}{2}(t-s)\|u\|^2} \tag{vi}$$

for each $u = (u_1, \ldots, u_d) \in \mathbb{R}^d$. This condition is equivalent as

$$\int_A e^{i\langle u, M_t - M_s\rangle} d\mathbb{P} = \int_A e^{-\frac{1}{2}(t-s)\|u\|^2} d\mathbb{P} \quad \text{for all } A \in \mathscr{F}_s.$$

Since the right-hand side contains constant, we can rewrite this as

$$\int_\Omega e^{i\langle u, M_t - M_s\rangle}\mathbb{1}_A d\mathbb{P} = e^{-\frac{1}{2}(t-s)\|u\|^2}\int_A d\mathbb{P}.$$

By the definition of expectation, we have

$$\mathbb{E}\left[e^{i\langle u, M_t - M_s\rangle}\mathbb{1}_A\right] = e^{-\frac{1}{2}(t-s)\|u\|^2}\mathbb{P}(A). \tag{condition}$$

[6] $\varphi_{(X,Y)} = \varphi_X \cdot \varphi_X \Leftrightarrow X \perp Y$.

For $u \in \mathbb{R}^d$ given and fixed, define a function $F : \mathbb{R}^d \to \mathbb{C}$ by setting

$$F(x) = \mathrm{e}^{\mathrm{i}\langle u,x\rangle}.$$

Applying Itô's formula to the real and imaginary parts of F, we get

$$\mathrm{e}^{\mathrm{i}\langle u,M_t\rangle} = \mathrm{e}^{\mathrm{i}\langle u,M_s\rangle} + \mathrm{i}\sum_{j=1}^{d}\left\{u_j \int_s^t \mathrm{e}^{\mathrm{i}\langle u,M_v\rangle}\mathrm{d}M_v^j\right\} - \frac{1}{2}\sum_{j=1}^{d}\left\{u_j^2\right\}\int_s^t \mathrm{e}^{\mathrm{i}\langle u,M_v\rangle}\mathrm{d}v, \tag{$*$}$$

by noting that $F_{x_j} = \mathrm{i}u_j F(x)$ and $F_{x_j x_k} = -u_j u_k F(x)$ and using (iii) from given.

ITÔ FORMULA ON ANY $[s,t] \in \mathbb{R}_+$

Suppose $s \leq t$. Then we apply the Itô formula on $F(X_s)$ and $F(X_t)$:

$$F(X_s) = F(X_0) + \int_0^s F'(X_v)\,\mathrm{d}X_v + \frac{1}{2}\int_0^s F''(X_v)\,\mathrm{d}\langle X,X\rangle_v,$$
$$F(X_t) = F(X_0) + \int_0^t F'(X_v)\,\mathrm{d}X_v + \frac{1}{2}\int_0^t F''(X_v)\,\mathrm{d}\langle X,X\rangle_v.$$

Subtract the first equation out of the first equation we have

$$F(X_t) = F(X_s) + \int_s^t F'(X_v)\,\mathrm{d}X_v + \frac{1}{2}\int_s^t F''(X_v)\,\mathrm{d}\langle X,X\rangle_v.$$

Since $\langle M^j, M^j\rangle_v = v$, we have $M^j \in \mathcal{M}_2^C$.

DETAILS

By Proposition 4.16, we have if $\langle M^j, M^j\rangle_v = v$ for any fixed $v \geq 0$, then M^j is a continuous martingale. We would like to show it is also a square integrable martingale.

Proof. By Col. 4.15, for any $M \in \mathcal{M}_{\mathrm{loc}}^{C,0}$, then we have the BDG inequality:

$$\mathbb{E}\left[\max_{0\leq t\leq\tau}|M_t|^p\right] \leq c_p\mathbb{E}\left[\langle M, M\rangle_\tau^{p/2}\right],$$

for any $p \geq 1$, and τ is any stopping time. Take $\tau = k$ for some $k > 0$ (any determinsitic time $\tau(\omega) \equiv k$ is a stopping time under any sigma-algebra), then we have

$$\mathbb{E}\left[\max_{0\leq t\leq k}|M_t|^p\right] \leq c_p\mathbb{E}\left[\langle M, M\rangle_k^{p/2}\right].$$

Then we have

$$\mathbb{E}\left[|M_t|^2\right] \leq \mathbb{E}\left[\max_{0\leq t\leq k}|M_t|^2\right] \leq c_2\mathbb{E}\left[\langle M, M\rangle_k\right] = c_2 \cdot k < \infty.$$

Therefore we have M^j is also a square integrable martingale. □

Thus, the real and imaginary part of the integral process $\left(\int_0^t e^{i\langle u, M_v\rangle}dM_v^j\right)_{t\geq 0}$ are not only in $\mathcal{M}_{\mathrm{loc}}^C$ but also in $\mathcal{M}_2^C$ (Remark 3.24, 4). Hence:

$$\mathbb{E}\left[\int_s^t e^{i\langle u, M_v\rangle}dM_v^j \mid \mathscr{F}_s\right] = 0$$

for each $1 \leq j \leq d$.

With $A \in \mathscr{F}_s$, let us multiply ($*$) by $e^{-i\langle u, M_s\rangle}\mathbb{1}_A$ and take the expectation on both sides. This gives

$$\mathbb{E}\left[e^{i\langle u, M_t - M_s\rangle}\mathbb{1}_A\right] = \mathbb{P}(A) - \frac{1}{2}\|u\|^2\int_s^t \mathbb{E}\left[e^{i\langle u, M_v - M_s\rangle}\mathbb{1}_A\right]dv.$$

This is an integral equation for the function:

$$\varphi(t) = \mathbb{E}\left[e^{i\langle u, M_t - M_s\rangle}\mathbb{1}_A\right], \quad \text{with } s, A \text{ fixed}$$

and reads

$$\varphi(t) = \mathbb{P}(A) - \frac{1}{2}\|u\|^2\int_s^t \varphi(v)dv.$$

Take $\partial/\partial t$, we see that it translate into a differential equation with initial condition

$$\varphi'(t) = -\frac{1}{2}\|u\|^2 \varphi(t), \quad \varphi(s) = \mathbb{P}(A).$$

from where we see that the unique solution is

$$\varphi(t) = \mathbb{P}(A)\mathrm{e}^{-\frac{1}{2}\|u\|^2(t-s)}$$

i.e.

$$\mathbb{E}\left[\mathrm{e}^{\mathrm{i}\langle u, M_t - M_s\rangle}\mathbb{1}_A\right] = \mathbb{P}(A)\mathrm{e}^{-\frac{1}{2}(t-s)\|u\|^2}$$

proving (condition) and thus proving the claim. □

4.4 The Burkholder-Davis-Gundy Inequality

The next question will be: **What are the *sufficient conditions* for a continuous local martingale to be a martingale?** An answer can be obtained using the following important result.

Theorem 4.14. *Let $M \in \mathcal{M}_{loc}^{C,0}$ and $p > 0$. Then we have:*

$$c_p\mathbb{E}\left[\langle M, M\rangle_\infty^{p/2}\right] \le \mathbb{E}\left[\sup_{t\ge 0}|M_t|^p\right] \le C_p\mathbb{E}\left[\langle M, M\rangle_\infty^{p/2}\right]$$

for some constants c_p and C_p not dependent on M.

Corollary 4.15. *Let $M \in \mathcal{M}_{loc}^{C,0}$, τ be a stopping time, and $p > 0$. Then we have*

$$c_p\mathbb{E}\left[\langle M, M\rangle_\tau^{p/2}\right] \le \mathbb{E}\left[\sup_{0\le t\le\tau}|M_t|^p\right] \le C_p\mathbb{E}\left[\langle M, M\rangle_\tau^{p/2}\right]$$

where c_p and C_p are the same as above.

In particular, if $H \in L^2_{\text{loc}}(M)$, then:

$$c_p\mathbb{E}\left[\left(\int_0^\tau H_s^2 \mathrm{d}\langle M, M\rangle_s\right)^{p/2}\right] \le \mathbb{E}\left[\max_{0\le t\le\tau}\left|\int_0^t H_s \mathrm{d}M_s\right|^p\right] \le C_p\mathbb{E}\left[\left(\int_0^\tau H_s^2 \mathrm{d}\langle M, M\rangle_s\right)^{p/2}\right]$$

where c_p and C_p are the same as above.

We can now give an answer to the question above.

Proposition 4.16. *Let $M \in \mathcal{M}_{loc}^{C,0}$. If $\mathbb{E}\left[\langle M, M\rangle_t^{1/2}\right] < \infty, \forall t > 0$, then M is a continuous martingale.*

Proof. The fact that $M \in \mathcal{M}_{\mathrm{loc}}^{C,0}$ implies that $(M_{t\wedge\tau_n})_{t\geq 0}$ is a martingale for some stopping time $\tau_n \uparrow \infty$ as $n \to \infty$. Hence for $s < t$, we have

$$\mathbb{E}\left[M_{t\wedge\tau_n} \mid \mathscr{F}_s\right] = M_{s\wedge\tau_n}, \quad (\forall n \geq 1). \tag{limit}$$

We would like to pass the limit of $n \to \infty$ to this equation. By the BDG inequality,

$$\mathbb{E}\left[\sup_{n\geq 1}|M_{t\wedge\tau_n}|\right] \leq \mathbb{E}\left[\max_{0\leq s\leq t}|M_s|\right] \leq c_1\mathbb{E}\left[\langle M, M\rangle_t^{1/2}\right] < \infty$$

which shows that $(M_{t\wedge\tau_n})_{n\geq 1}$ is uniformally integrable. Let $n \to \infty$ in (limit) and using the Dunford-Pettis theorem (conditional), we obtain

$$\mathbb{E}[M_t \mid \mathscr{F}_s] = M_s,$$

proves the claim. □

$\mathbb{E}[\langle M, M\rangle_t] < \infty$ IMPLIES $\mathbb{E}[\langle M, M\rangle_t^{1/2}] < \infty$

By Jensen's Inequality: $G(\mathbb{E}[X]) \leq \mathbb{E}[G(X)]$ where G is a convex function. Let $G(x) = x^2$, then $(E[X])^2 \leq \mathbb{E}[X^2]$. Taking $X = \langle M, M\rangle_t^{1/2}$, then we have

$$\left(\mathbb{E}\left[\langle M, M\rangle_t^{1/2}\right]\right)^2 \leq \mathbb{E}\left[\langle M, M\rangle_t\right].$$

Therefore $\mathbb{E}[\langle M, M\rangle_t] < \infty$ implies $\mathbb{E}[\langle M, M\rangle_t^{1/2}] < \infty$. Notice that, $\mathbb{E}[\langle M, M\rangle_t] < \infty$ is the sufficient condition but not a necessary condition.

Proposition 4.17. *Let $M \in \mathcal{M}_{\mathrm{loc}}^{C}$, $H \in L^2_{loc}(M)$, and τ be a stopping time. Then we have:*

$$\mathbb{E}\left[\sqrt{\int_0^\tau H_s^2 \,\mathrm{d}\langle M, M\rangle_s}\right] < \infty \quad \Rightarrow \quad \mathbb{E}\left[\int_0^\tau H_s \,\mathrm{d}M_s\right] = 0.$$

Equivalently, if we have an integral process $H \cdot M$, then

$$\mathbb{E}\left[\langle H\cdot M, H\cdot M\rangle_\tau^{1/2}\right] < \infty \quad \Rightarrow \quad \mathbb{E}\left[(H\cdot M)_\tau\right] = 0.$$

Proof. Set $I_t = \int_0^t H_s \,\mathrm{d}M_s$, then $I_t \in \mathcal{M}_{\mathrm{loc}}^{C,0}$. Then by the BDG inequality, we get:

$$\mathbb{E}\left[\sup_{t\geq 0}|I_{t\wedge\tau}|\right] = \mathbb{E}\left[\max_{0\leq t\leq\tau}|I_t|\right] \leq \mathbb{E}\left[\left(\int_0^\tau H_s^2 \,\mathrm{d}\langle M, M\rangle_s\right)^{1/2}\right] < \infty$$

from where we see that the local martingale $(I_{t\wedge\tau})_{t\geq 0}$ is uniformally integrable and uniformally bounded by $\sup_{t\geq 0}|I_{t\wedge\tau}|$, and thus $(I_{t\wedge\tau})_{t\geq 0}$ a martingale. Hence $\mathbb{E}[I_{t\wedge\tau}] = 0$, and by uniformally integrability, the Dunford-Pettis theorem can be used. Therefore $\mathbb{E}[I_\tau] = 0$ follows by letting $t \to \infty$.

MORE THOUGHTS

From $\mathbb{E}\left[\sup_{t\geq 0}|I_{t\wedge\tau}|\right] < \infty$ implies $\sup_{t\geq 0}|I_{t\wedge\tau}|$ is uniformally integrable. Also, taking $\eta = \tau$ as another stopping time, we have

$$\left|I_{t\wedge\tau\wedge\eta}\right| = |I_{t\wedge\tau}| \leq \sup_{t\geq 0}|I_{t\wedge\tau}| \in L^1(\mathbb{P}),$$

hence $I_{t\wedge\tau\wedge\eta}$ is uniformally bounded by $\sup_{t\geq 0}|I_{t\wedge\tau}|$ and η is optional for $I_{t\wedge\tau}$. By uniformally boundedness (inequality), we have $I_{t\wedge\tau}$ is a martingale. Therefore, by optional sampling theorem

$$\mathbb{E}[I_{t\wedge\tau}] = \mathbb{E}[I_{t\wedge\tau\wedge\eta}] = \mathbb{E}[I_{t\wedge\tau\wedge 0}] = 0.$$

Since $I_{t\wedge\tau}$ is uniformally integrable in t, hence we can use the Dunford-Pettis theorem and

$$\mathbb{E}[I_\tau] = \mathbb{E}\left[\lim_{t\to\infty} I_{t\wedge\tau}\right] = \lim_{t\to\infty}\mathbb{E}[I_{t\wedge\tau}] = 0.$$

□

Remark 4.18. Recall that uniformally boundedness implies uniformally integrable, but not vice versa. Therefore, we need the uniformally boundedness, otherwise we can construct a counterexample such that a uniformally integrable local martingale is not a martingale.

4.5 Change of Measure

We turn to the following problem. Let $B = (B_t)_{t\geq 0}$ be a standard Brownian motion on $(\Omega, \mathscr{F}, \mathbb{P})$. Given a 'good' process $H = (H_s)_{s\geq 0}$, consider the following process:

$$\widetilde{B}_t = B_t - \int_0^t H_s \, \mathrm{d}s, \quad (t \geq 0).$$

Question. Does there exist a probability measure $\widetilde{\mathbb{P}}$ on $(\Omega, \mathscr{F})$ such that $\widetilde{B} = (\widetilde{B})_{t\geq 0}$ is a standard Brownian motion under $\widetilde{\mathbb{P}}$?

The answer is yes! It is an important result called Girsanov (measure-change) theorem. Before we pass to the formulation of the Girsanov theorem, let us present a simple illustration.

Example 4.19. Let $Z_1, \ldots, Z_n$ be iid with distribution $N(0,1)$ defined on $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\mu_1, \ldots, \mu_n \in \mathbb{R}$ and consider

$$\widetilde{Z}_1 = Z_1 - \mu_1, \ldots, \widetilde{Z}_n = Z_n - \mu_n.$$

Define a new probability measure $\widetilde{\mathbb{P}}$ by:

$$d\widetilde{\mathbb{P}} = \exp\left(\sum_{i=1}^{n} \mu_i z_i - \frac{1}{2}\sum_{i=1}^{n} \mu_i^2\right) d\mathbb{P}.$$

Then we have

$$\begin{aligned}
&\widetilde{\mathbb{P}}(Z_1 \in dz_1, \ldots, Z_n \in dz_n) \\
&= \exp\left(\sum_{i=1}^{n} \mu_i z_i - \frac{1}{2}\sum_{i=1}^{n} \mu_i^2\right) \mathbb{P}\left(Z_1 \in dz_1, \ldots, Z_n \in dz_n\right) \\
&= \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} (z_i - \mu_i)^2\right) dz_1 \cdots dz_n.
\end{aligned}$$

Thus, under $\widetilde{\mathbb{P}}$, we have that $Z_1, \ldots, Z_n$ are independent and $Z_i \sim N(\mu_i, 1)$ for $1 \le i \le n$ or in other hand, we have that $\widetilde{Z}_1, \ldots, \widetilde{Z}_n$ are iid with distribution $N(0,1)$ under $\widetilde{\mathbb{P}}$.

Example 4.20. Suppose we have the following probability space $\Omega = [0,1]$, $\mathscr{F} = \mathscr{B}([0,1])$, and $\mathbb{P}$ is the Lebsegue measure. Define $\widetilde{\mathbb{P}}(\{0\}) = 1/2$ and $\widetilde{\mathbb{P}}(\{1\}) = 1/2$. Here $\widetilde{\mathbb{P}} : \mathscr{F} \mapsto [0,1]$ is a *probability measure*: $\forall F \in \mathscr{F}$,

$$\widetilde{\mathbb{P}}(F) = \widetilde{\mathbb{P}}(F \cup \{0,1\}) = \begin{cases} 0 & \text{if } 0 \notin F \text{ and } 1 \notin F, \\ 1/2 & \text{if } 0 \in F \text{ and } 1 \notin F \textbf{ OR } 0 \notin F \text{ and } 1 \in F, \\ 1 & \text{if } 0 \in F \text{ and } 1 \in F. \end{cases}$$

Let $X : \Omega \to \mathbb{R}$ is a random variable, defined by

$$X(\omega) = \omega, \quad \forall \omega \in \Omega.$$

Then what is $\mathfrak{Law}(X \mid \mathbb{P})$ and $\mathfrak{Law}(X \mid \widetilde{\mathbb{P}})$? It is easy to see that $\mathfrak{Law}(X \mid \mathbb{P}) = U(0,1)$ and $\mathfrak{Law}(X \mid \widetilde{\mathbb{P}})$ is a two point law with

$$\mathfrak{Law}(X \mid \widetilde{\mathbb{P}}) = \begin{pmatrix} 0 & 1 \\ 1/2 & 1/2 \end{pmatrix}.$$

These two examples shows *law of same random variable under different measure can be quite different*. The Girsanov theorem may be viewed as an extension of this fact from the discrete time to the continuous time settings.

Assumption.

- Throughout, $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\ge 0}, \mathbb{P})$ is a filtered probability space where

$$\mathcal{N} = \{N \in \mathscr{F} \mid \mathbb{P}(N) = 0\} \subset \mathscr{F}_t, \quad (\forall t \ge 0).$$

- $B = (B^1, \dots, B^d)$ is a standard d-dimensional Brownian motion with respect to $(\mathscr{F}_t)_{t\geq 0}$.

$H = (H^1, \dots, H^d)$ is a given process such that each $H^i \in L^2_{\mathrm{loc}}(B^i)$, i.e. H^i is progressively measurable and satisfies:

$$\mathbb{P}\left(\int_0^t \left(H_s^i\right)^2 \,\mathrm{d}s < \infty\right) = 1, \quad (\forall t \geq 0).$$

Then $(\int_0^t H_s^i \,\mathrm{d}B_s^i)_{t\geq 0}$ is well-defined and belongs to $\mathcal{M}^{C,0}_{\mathrm{loc}}$ for each $1 \leq i \leq d$. We set:

$$\mathcal{E}_t^H(B) := \exp\left(\sum_{i=1}^d \int_0^t H_s^i \,\mathrm{d}B_s^i - \frac{1}{2}\int_0^2 \|H_s\|^2 \,\mathrm{d}s\right) \tag{i}$$

for $t \geq 0$. By Itô formula:

$$\mathcal{E}_t^H(B) = 1 + \sum_{i=1}^d \int_0^t \mathcal{E}_s^H(B) H_s^i \,\mathrm{d}B_s^i. \tag{$\star$}$$

This shows that $\mathcal{E}_t^H(B) \in \mathcal{M}^C_{\mathrm{loc}}$ with $\mathcal{E}_0^H(B) = 1$.

However, the question now will be *under which condition, $\mathcal{E}_t^H(B)$ will be a martingale?*

Proposition 4.21 (Novikov's condition(1972))**.** *Assume that $(H_s^i)_{0\leq s\leq T} \in L^2_{\mathrm{loc}}(B^i)$. If*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \|H_s\|^2 \,\mathrm{d}s\right)\right] < \infty,$$

then $\mathcal{E}_t^H(B)$ is a martingale for all $0 \leq t \leq T$.

Remark 4.22. *Kazamaki's condition* is also sufficient for $\mathcal{E}^H(B)$ to be a true martingale. Kazamaki's condition is more general than Novikov's condition.

If $\mathcal{E}_t^H(B)$ is a martingale, then we have the consequence that

$$\mathbb{E}[\mathcal{E}_t^H(B)] = 1, \quad (\forall t \geq 0),$$

by using the optional sampling theorem. ($\mathbb{E}[M_\tau] = \mathbb{E}[M_0] = 1$ for all $t \geq 0$ given and fixed, we can choose $\tau = t$) So that with $T > 0$ given and fixed, we can define a probability measure $\widetilde{\mathbb{P}}_T$ on $\mathscr{F}_T$ by

$$\widetilde{\mathbb{P}}_T(A) = \mathbb{E}[\mathbb{1}_A \mathcal{E}_T^H(B)], \quad (A \in \mathscr{F}_T). \tag{ii}$$

Notice (ii) is only valid if $\mathcal{E}_t^H(B)$ is a true martingale, i.e.

$$\widetilde{\mathbb{P}}_T(\Omega) = \mathbb{E}[\mathbb{1}_\Omega \mathcal{E}_T^H(B)] = \mathbb{E}[\mathcal{E}_T^H(B)] = 1.$$

The martingale property of $\mathcal{E}^H(B)$ implies that the family of probability measure $\{\widetilde{\mathbb{P}}_t \mid t \geq 0\}$ satisfies the following consistency condition:

$$\widetilde{\mathbb{P}}_T(A) = \widetilde{\mathbb{P}}_t(A)$$

for each $A \in \mathscr{F}_t$ with any $0 \leq t \leq T$. I.e. Two different measures are identical under the smaller sigma-algebra.

In order to prove Theorem 4.25, we need the following two lemmas. Let us use the following notation:

$$\int_\Omega \mathrm{d}\widetilde{\mathbb{P}}_T = \widetilde{\mathbb{E}}_T.$$

Lemma 4.23. *Assume that $\mathcal{E}^H(B)$ is a martingale. We have for Z which is $\mathscr{F}_t$ measurable and satisfies $\widetilde{\mathbb{E}}_T[|Z|] < \infty$, then*

$$\widetilde{\mathbb{E}}_T[Z \mid \mathscr{F}_s] = \frac{1}{\mathcal{E}_s^H(B)} \mathbb{E}[Z\mathcal{E}_t^H(B) \mid \mathscr{F}_s]$$

for $0 \leq s \leq t \leq T$.

Proof.

Before proving Lemma 4.23, we have the following important observation from (ii): We can rewrite $\widetilde{\mathbb{P}}_T(A) = \mathbb{E}\left[\mathbb{1}_A \mathcal{E}_T^H(B)\right]$ as the integral form

$$\int_\Omega \mathbb{1}_A \,\mathrm{d}\widetilde{\mathbb{P}}_T = \int_\Omega \mathbb{1}_A \mathcal{E}_T^H(B)\,\mathrm{d}\mathbb{P}.$$

Differentiate both sides, we have

$$\mathrm{d}\widetilde{\mathbb{P}}_T = \mathcal{E}_T^H(B)\,\mathrm{d}\mathbb{P}. \tag{CM}$$

for any $T > 0$ given and fixed.

For $A \in \mathscr{F}_s$, we have

$$\begin{aligned}
&\widetilde{\mathbb{E}}_T\left[\mathbb{1}_A \frac{1}{\mathcal{E}_s^H(B)} \mathbb{E}\left[Z\mathcal{E}_t^H(B) \mid \mathscr{F}_s\right]\right] \\
=&\widetilde{\mathbb{E}}_s\left[\mathbb{1}_A \frac{1}{\mathcal{E}_s^H(B)} \mathbb{E}\left[Z\mathcal{E}_t^H(B) \mid \mathscr{F}_s\right]\right] \quad \widetilde{\mathbb{P}}_T = \widetilde{\mathbb{P}}_s \text{ on } \mathscr{F}_s \\
=&\int_\Omega \mathbb{1}_A \frac{1}{\mathcal{E}_s^H(B)} \mathbb{E}\left[Z\mathcal{E}_t^H(B) \mid \mathscr{F}_s\right] \mathrm{d}\widetilde{\mathbb{P}}_s \\
=&\int_\Omega \mathbb{1}_A \frac{1}{\mathcal{E}_s^H(B)} \mathbb{E}\left[Z\mathcal{E}_t^H(B) \mid \mathscr{F}_s\right] \mathcal{E}_s^H(B)\,\mathrm{d}\mathbb{P} \quad \text{using (CM)} \\
=&\mathbb{E}\left[\mathbb{1}_A \mathbb{E}\left[Z\mathcal{E}_t^H(B) \mid \mathscr{F}_s\right]\right] = \mathbb{E}[\mathbb{1}_A Z\mathcal{E}_t^H(B)] = \widetilde{\mathbb{E}}_t[\mathbb{1}_A Z] = \widetilde{\mathbb{E}}_T[\mathbb{1}_A Z].
\end{aligned}$$

□

Denote by $\mathcal{M}^{C,0}_{\mathrm{loc},T}$ the class of continuous local martingales $(M_t)_{0\le t\le T}$ with respect to $(\mathscr{F}_t)_{0\le t\le T}$ on $(\Omega, \mathscr{F}_T, \mathbb{P}_T)$ (where $\mathbb{P}_T = \mathbb{P}\mid_{\mathscr{F}_T}$) such that $M_0 = 0$. Also, denote by $\widetilde{\mathcal{M}}^{C,0}_{\mathrm{loc},T}$ the class of continuous local martingales $(M_t)_{0\le t\le T}$ with respect to $(\mathscr{F}_t)_{0\le t\le T}$ on $(\Omega, \mathscr{F}_T, \widetilde{\mathbb{P}}_T)$ (where $\widetilde{\mathbb{P}}_T$ is defined by (ii)) such that $M_0 = 0$.

Lemma 4.24. *Assume that $\mathcal{E}^H(B)$ is a martingale. If $M \in \mathcal{M}^{C,0}_{\mathrm{loc},T}$, then the process*

$$\widetilde{M}_t := M_t - \sum_{i=1}^d \int_0^t H^i_s \,\mathrm{d}\langle M, B^i\rangle_s, \quad (0 \le t \le T)$$

belongs to $\widetilde{\mathcal{M}}^{C,0}_{\mathrm{loc},T}$. Moreover, if we have another process

$$\widetilde{N}_t := N_t - \sum_{i=1}^d \int_0^t H^i_s \,\mathrm{d}\langle N, B^i\rangle_s \quad (0 \le t \le T)$$

from $\widetilde{\mathcal{M}}^{C,0}_{\mathrm{loc},T}$ for some $N \in \mathcal{M}^{C,0}_{\mathrm{loc},T}$, then

$$\langle \widetilde{M}, \widetilde{N}\rangle_t = \langle M, N\rangle_t \quad (0 \le t \le T)$$

both $\mathbb{P}_T$-a.s. and $\widetilde{\mathbb{P}}_T$-a.s..

Proof. By localization, we may assume M and N are bounded martingales, with bounded $\langle M, M\rangle$ and $\langle N, N\rangle$, as well as that $\mathcal{E}^H(B)$ and $(\sum_{i=1}^d \int_0^t (H^i_s)^2 \,\mathrm{d}s)_{t\ge 0}$ are bounded. By Proposition 3.12, we get

$$\begin{aligned}\left|\int_0^t H^i_s \,\mathrm{d}\langle M, B^i\rangle_s\right|^2 &\le \int_0^t 1 \,\mathrm{d}\langle M, M\rangle_s \int_0^t (H^i_s)^2 \,\mathrm{d}\langle B^i, B^i\rangle_s \\ &= \langle M, M\rangle_t \int_0^t (H^i_s)^2 \,\mathrm{d}s\end{aligned}$$

which shows that the process $\widetilde{M}$ is also bounded. The integration by part formula (Example 4.10) gives

$$\mathcal{E}^H_t(B)\widetilde{M}_t = \int_0^t \mathcal{E}^H_s(B)\,\mathrm{d}\widetilde{M}_s + \int_0^t \widetilde{M}_s \,\mathrm{d}\mathcal{E}^H_s(B) + \langle \mathcal{E}^H(B), \widetilde{M}\rangle_t.$$

Notice the following,

$$\begin{aligned}\mathrm{d}\widetilde{M}_s &= \mathrm{d}M_s - \cancel{\sum_{i=1}^d H^i_s \,\mathrm{d}\langle M, B^i\rangle_s}, \\ \mathrm{d}\mathcal{E}^H_s(B) &= \sum_{i=1}^d \mathcal{E}^H_s(B) H^i_s \,\mathrm{d}B^i_s \quad \text{by } (\star), \\ \langle \mathcal{E}^H(B), \widetilde{M}\rangle_t &= \cancel{\sum_{i=1}^d \int_0^t \mathcal{E}^H_s(B) H^i_s \,\mathrm{d}\langle B^i, M\rangle_s}.\end{aligned}$$

This implies $\mathcal{E}_t^H(B)\widetilde{M}_t$ is a martingale under $\mathbb{P}_T$. Therefore by Lemma 4.23, we get

$$\begin{aligned}\widetilde{\mathbb{E}}_T[\widetilde{M}_t \mid \mathscr{F}_s] &= \frac{1}{\mathcal{E}_s^H(B)}\mathbb{E}[\widetilde{M}_t\mathcal{E}_t^H(B) \mid \mathscr{F}_s] \\ &= \frac{1}{\mathcal{E}_s^H(B)}\widetilde{M}_s\mathcal{E}_s^H(B) = \widetilde{M}_s.\end{aligned}$$

for each $0 \le s \le t \le T$. It follows that $\widetilde{M} \in \widetilde{\mathcal{M}}_{\mathrm{loc},T}^{C,0}$ as claimed.

The rest of the proof are omitted, but you can view the proof as the following: Since $\widetilde{M}_t$ and $\widetilde{N}_t$ are continuous semimartingales under $(\Omega, \mathscr{F}, \widetilde{\mathbb{P}}_T)$, therefore the BV part are irrelevant under quadratic covariation, i.e. $\langle \widetilde{M}, \widetilde{N}\rangle_t = \langle M, N\rangle_t$. □

Theorem 4.25 (Girsanov(1960), Cameron & Martin (1944))**.** *Assume that $\mathcal{E}^H(B)$ is a martingale. Define a process $\widetilde{B} = (\widetilde{B}^1, \dots, \widetilde{B}^d)$ by setting:*

$$\widetilde{B}_t^i = B_t^i - \int_0^t H_s^i \,\mathrm{d}s, \quad (1 \le i \le d).$$

Then $(\widetilde{B})_{0\le t\le T}$ is a standard d-dimensional Brownian motion with respect to $(\mathscr{F}_t)_{0\le t\le T}$ on the probability space $(\Omega, \mathscr{F}_T, \widetilde{\mathbb{P}}_T)$ where $\widetilde{\mathbb{P}}_T$ is defined by (ii) *above and $T > 0$ is given and fixed.*

Proof. In order to show that $\widetilde{B} = (\widetilde{B}^1, \dots, \widetilde{B}^d)$ is a standard d-dimensional Brownian motion by Theorem 4.13, we need to verify:

1. $\widetilde{B}$ is adapted to $(\mathscr{F}_t)_{t\ge 0}$,
2. $\widetilde{B}^i \in \widetilde{\mathcal{M}}_{\mathrm{loc},T}^{C,0}$,
3. $\langle \widetilde{B}^i, \widetilde{B}^j\rangle_t = \delta_{ij}\cdot t$, for all $0 \le t \le T$,

for all $i, j \in \{1, \dots, d\}$.

Clearly the first point is satisfied. For the second point, we use the first part of Lemma 4.24 with $M = B^j$. We know that $B^j \in \mathcal{M}_{\mathrm{loc},T}^{C,0}$, then we have for $j \in \{1, \dots, d\}$

$$\widetilde{B}_t^j = B_t^j - \sum_{i=1}^d \int_0^t H_s^i \,\mathrm{d}\langle B^j, B^i\rangle_s \quad (0 \le t \le T).$$

Since $(B^1, \dots, B^d)$ is a d-dimensional Brownian motion, therefore $\langle B^j, B^i\rangle_t = \delta_{ij}\cdot t$, i.e.

$$\widetilde{B}_t^j = B_t^j - \int_0^t H_s^j \,\mathrm{d}s \in \widetilde{\mathcal{M}}_{\mathrm{loc},T}^{C,0}, \quad 1 \le j \le d,$$

which shows the second point is satisfied. For the third part, we can use the second part of Lemma 4.24 with $M = B^i$ and $N = B^j$. This implies

$$\langle \widetilde{B}^i, \widetilde{B}^j\rangle_t = \langle B^i, B^j\rangle_t = \delta_{ij}\cdot t$$

as claimed. This completes the proof of the theorem. □

5 Representation of Martingales

Throughout $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is a filtered probability space where $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous, i.e. $\mathcal{F}_{t+} = \mathcal{F}_t$ for each $t \geq 0$.

5.1 Change of Time

The following important result shows that *each continuous local martingale is a time-changed Brownian motion*. Note that the 'new clock' of the Brownian motion is exactly the quadratic variation of the martingale.

Theorem 5.1 (Dambis (1965), Dubius & Schwarz (1965))**.** *Let $M \in \mathcal{M}_{\text{loc}}^{C,0}$ such that $\langle M, M\rangle_\infty = \infty$. Consider*

$$\tau_t = \inf\{s > 0 \mid \langle M, M\rangle_s > t\}.$$

Then $B_t = M_{\tau_t}$ is a standard Brownian motion with respect to $(\mathcal{F}_{\tau_t})_{t\geq 0}$ and $M_t = B_{\langle M,M\rangle_t}$ for all $t \geq 0$.

Remark 5.2.

- There is a *multidimensional analogue* of Theorem 5.1 due to Knight (1971).
- Even if $\langle M, M\rangle_\infty < \infty$, the result of the theorem still holds if we are allowed to 'enlarge' the probability space.

5.2 Martingales Adapted to Brownian Filtrations

We will now consider the following problem:

Question. If Z is a measurable random variable with respect to $\mathcal{F}_T^B$, can we conclude that

$$Z = z + \int_0^T H_s \,\mathrm{d}B_s$$

for some H?

The answer is yes! Moreover, if $Z \in L^2$, then H is 'unique'. It has the following informal intepretation (in analogue with classic calculus):

$$H = \frac{dZ}{dB}$$

We begin by collecting a few facts on "Brownian filtrations".

1. *The natural filtration $\mathcal{F}_t^B := \sigma(B_t \mid 0 \leq s \leq t)$:*

(a) is left-continuous, i.e. $\mathscr{F}_{t-}^B = \mathscr{F}_t^B$, $\forall t$.

(b) is NOT right-continuous, i.e. $\mathscr{F}_t^B \subsetneq \mathscr{F}_{t+}^B$, $\forall t$.

2. *The augmented filtration* $\widetilde{\mathscr{F}_t}^B := \sigma(\mathscr{F}_t^B \cup \widetilde{\mathcal{N}}_\infty^B)$ where

$$\widetilde{\mathcal{N}}_\infty^B = \{N_0 \subseteq \Omega \mid N_0 \subseteq N, N \in \mathscr{F}_\infty^B, \mathbb{P}(N) = 0\}.$$

(c) is continuous and complete. (Complete means whenever we have a null set $N \in \widetilde{\mathcal{N}}_\infty$, then all subset of N belongs to $\widetilde{\mathcal{N}}_\infty$)

Denote by $\mathcal{S}_T$ the space of all simple functions $f : \mathbb{R}_+ \to \mathbb{R}$ of the form

$$f = \sum_{j=1}^n \lambda_j \mathbb{1}_{(t_{j-1},t_j]}$$

where $\lambda_j \in \mathbb{R}$ and $0 \leq t_0 < t_1 < \cdots < t_n \leq T$ with $n \geq 1$.

For each $f \in \mathcal{S}_T$, we know that

$$\mathcal{E}_t^f = \exp\left(\int_0^t f(s)\,\mathrm{d}B_s - \frac{1}{2}\int_0^t f^2(s)\,\mathrm{d}s\right), \quad (0 \leq t \leq T)$$

is a martingale.

Lemma 5.3. *The set* $\{\mathcal{E}_T^f \mid f \in \mathcal{S}_T\}$ *is* total *in* $L^2(\widetilde{\mathscr{F}}_T^B, \mathbb{P})$. *Total means that the family of all (finite) linear combinations of* $\mathcal{E}_T^f (f \in \mathcal{S}_T)$ *is dense.*

Theorem 5.4 (Itô-Clark Theorem, [1, 1996, Thm 6.2])**.** *Let Z be from* $L^2(\widetilde{\mathscr{F}}_T^B, \mathbb{P})$. *Then there exist a unique* $H \in L^2(B)$ *such that:*

$$Z = \mathbb{E}[Z] + \int_0^T H_s\,\mathrm{d}B_s \tag{i}$$

We extend Theorem 5.4 to local martingale (not necessarily continuous). We observe that the result below establish the following remarkable feature of the Brownian filtration $(\widetilde{\mathscr{F}}_t^B)_{t\geq 0}$: *There is no discontinuous martingale with respect to* $(\widetilde{\mathscr{F}}_t^B)_{t\geq 0}$*!*

Theorem 5.5 (Martingale Representation Theorem)**.** *Let* $M = (M_t)_{t\geq 0}$ *be a right-continuous local martingale with respect to* $(\widetilde{\mathscr{F}}_t^B)_{t\geq 0}$. *There exist* $H = (H_s)_{s\geq 0} \in L^2_{loc}(B)$ *such that*

$$M_t = c + \int_0^t H_s\,\mathrm{d}B_s$$

for all $t \geq 0$. *(In particular, we see that M must be $\mathbb{P}$-a.s. continuous!) Moreover, if M is L^2-bounded (i.e.* $\sup_{t\geq 0} \mathbb{E}[M_t^2] < \infty$*), then H is unique in* $L^2(B)$.

Example 5.6 (Running Maximum of the Brownian Motion)**.** Let B be a standard Brownian motion, consider

$$S_1 = \max_{0\le t\le 1} B_t.$$

Clearly, $S_1 \in L^2(\widetilde{\mathscr{F}}_1^B, \mathbb{P})$, so that by Theorem 5.4, there exist a unique $H \in L^2(B)$ such that:

$$S_1 = a + \int_0^1 H_s \, \mathrm{d}B_s$$

where $a = \mathbb{E}[S_1] = (\mathbb{E}[|B_1|], \text{since } S_1 \sim |B_1|) = \sqrt{2/\pi}$. (For this, you may refer to Link)

Question. Is it possible to determine H explicitly?

For this, we firstly note that we have:

$$\begin{aligned}
\mathbb{E}[S_1 \mid \mathscr{F}_t^B] &\overset{*}{=} S_t + \mathbb{E}\left[\left(\max_{t\le s\le 1} B_s - S_t\right)^+ \mid \mathscr{F}_t^B\right] \\
&= S_t + \mathbb{E}\left[\left(\underbrace{\max_{t\le s\le 1}(B_s - B_t)}_{\perp \mathscr{F}_t^B,\ \sim S_{1-t}} - \underbrace{(S_t - B_t)}_{\mathscr{F}_t^B \text{ measurable}}\right)^+ \mid \mathscr{F}_t^B\right] \\
&= S_t + \mathbb{E}\left[\left(S_{1-t} - (z-x)\right)^+\right] |_{z=S_t, x=B_t}
\end{aligned}$$

The equation $(*)$ comes from the following argument: $\mathbb{E}[S_1 \mid \mathscr{F}_t^B]$ is the best prediction of S_1 based on what we know up to time t. If S_1 is attained before or on time t, then $\max_{t\le s\le 1} B_s \le S_1$, therefore $(\max_{t\le s\le 1} B_s - S_t) \le 0$. Then the left-hand side will be equal to S_t which is already the maximum. If S_1 is attained after time t, then $\max_{t\le s\le 1} B_s > S_t$, therefore S_t cancelled, therefore we have $\mathbb{E}[S_1 \mid \mathscr{F}_t^B] = \mathbb{E}[\max_{t\le s\le 1} B_s \mid \mathscr{F}_t^B]$.

Using further the formula:

$$\mathbb{E}[(X-c)^+] = \int_c^\infty \mathbb{P}(X > z) \, \mathrm{d}z \tag{5.1}$$

To understand (5.1), we start from the left-hand side,

$$\int_c^\infty \mathbb{P}(X > z)\, \mathrm{d}z = \int_c^\infty \mathbb{E}[\mathbb{1}_{X>z}] = \mathbb{E}\Big(\int_c^\infty \mathbb{1}_{X>z} \, \mathrm{d}z\Big)$$

Notice that X cannot go below c, otherwise $\mathbb{1}_{X>z} = 0$ for all z. If X is equal or above c, then $\int_c^\infty \mathbb{1}_{X>z} \, \mathrm{d}z$ is integrating 1 over c to X, hence

$$\int_c^\infty \mathbb{1}_{X>z}\, \mathrm{d}z = (X-c)^+$$

then the (5.1) holds.

we see that the preceding identity reads:

$$\mathbb{E}[S_1 \mid \mathscr{F}_t^B] = S_t + \int_{S_t - B_t}^{\infty} (1 - F_{1-t}(z))\,\mathrm{d}z := f(t, B_t, S_t), \tag{i}$$

where we set

$$F_{1-t}(z) = \mathbb{P}(S_{1-t} \le z) = \mathbb{P}(|B_{1-t}| \le z) = \mathbb{P}\left(|B_1| \le \frac{z}{\sqrt{1-t}}\right).$$

We know that $B_1 \sim N(0,1)$, therefore we have

$$F_{1-t}(z) = \Phi\left(\frac{z}{\sqrt{1-t}}\right) - \left(1 - \Phi\left(\frac{z}{\sqrt{1-t}}\right)\right) = 2\Phi\left(\frac{z}{\sqrt{1-t}}\right) - 1$$

Applying Itô's formula to $f(t, B_t, S_t)$ and using that the left-hand side in (i) is a continuous local martingale, we find that

$$\mathbb{E}[S_1 \mid \mathscr{F}_t^B] = a + \int_0^t \frac{\partial f}{\partial x}(s, B_s, S_s)\,\mathrm{d}B_s$$

as a non-trivial continuous martingale cannot have paths of bounded variation. Since

$$\frac{\partial f}{\partial x}(t, x, s) = 1 - F_{1-t}(s - x),$$

we see that

$$\mathbb{E}[S_1 \mid \mathscr{F}_t^B] = a + \int_0^t [1 - F_{1-s}(S_s - B_s)]\,\mathrm{d}B_s.$$

Finally, letting $t = 1$, we get

$$S_1 = a + \int_0^t 2\left(1 - \Phi\left(\frac{z}{\sqrt{1-t}}\right)\right)\,\mathrm{d}B_s,$$

where

$$H_s = 2\left(1 - \Phi\left(\frac{z}{\sqrt{1-t}}\right)\right) \quad (0 \le s \le 1).$$

This completes the calculuation. (i.e. $\mathrm{d}S_1/\mathrm{d}B = H$)

6 Stochastic Differential Equations

6.1 Examples of SDE

Let $B = (B_t)_{t\geq 0}$ be a standard Brownian motion. Recall from Example 4.8 that the process $X = (X_t)_{t\geq 0}$ solves the stochastic differential equation:

$$dX_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t, \quad (X = x_0) \tag{i}$$

if and only if

$$X_t = x_0 + \int_0^t \mu(X_s)\,\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}B_s \tag{ii}$$

for $t \geq 0$. We begin our exposition by giving a list of best known stochastic differential equations.

Example 6.1 (Brownian motion with drift)**.**

$$\mathrm{d}X_t = \mu\mathrm{d}t + \sigma\mathrm{d}B_t, \quad (\mu \in \mathbb{R}, \sigma > 0)$$

has the solution

$$X_t = x + \mu t + \sigma B_t, \quad x \in \mathbb{R}.$$

Example 6.2 (Geometric Brownian motion)**.**

$$\mathrm{d}X_t = \mu X_t\mathrm{d}t + \sigma X_t\mathrm{d}B_t \quad (\mu \in \mathbb{R}, \sigma > 0)$$

has the solution

$$X_t = x\exp\left(\sigma B_t + \left(\mu - \frac{\sigma^2}{2}t\right)\right), \quad x > 0.$$

We can see this is the exponential martingale if $\mu = 0$.

Example 6.3 (Bessel process)**.**

$$\mathrm{d}X_t = \frac{\alpha - 1}{2X_t}\mathrm{d}t + \mathrm{d}B_t, \quad (\alpha \geq 0)$$

has the solution for $\alpha = d \in \mathbb{N}$

$$X_t = \left(\sum_{i=1}^{d}(x_i + B_t^i)^2\right)^{1/2}$$

where $B = (B^1, \ldots, B^d)$ is standard d-dimensional Brownian motion.

Example 6.4 (Squared Bessel process)**.**

$$\mathrm{d}X_t = \alpha \mathrm{d}t + 2\sqrt{X_t}\mathrm{d}B_t, \quad (\alpha \geq 0)$$

has the solution for $\alpha = d \in \mathbb{N}$

$$X_t = \sum_{i=1}^{d} (x_i + B_t^i)^2$$

where $B = (B^1, \ldots, B^d)$ is standard d-dimensional Brownian motion.

Example 6.5 (Ornstein-Uhlenbeck Process)**.**

$$\mathrm{d}X_t = -\beta X_t \mathrm{d}t + \sigma \mathrm{d}B_t, \quad (\beta > 0, \sigma > 0).$$

Here, X_t describes the velocity of the Brownian motion. It has the closed form solution.

$$X_t = \mathrm{e}^{-\beta t}\left(x + \sigma \int_0^t \mathrm{e}^{\beta s}\, \mathrm{d}B_s\right), \quad (x \in \mathbb{R}).$$

OU process is essentially the Brownian motion with a drag force towards the real line as shown in the Figure 2.

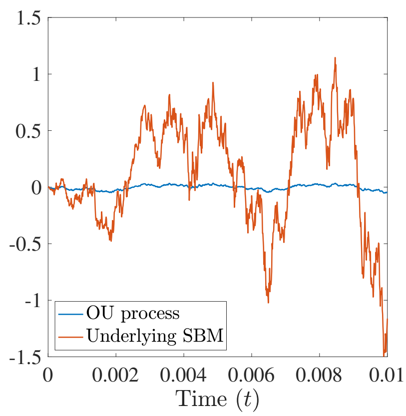

Figure 2: Sample path of the OU process with $\beta = 1$ and $\sigma = 1$, and its underlying SBM.

Example 6.6 (Branching diffusion)**.**

$$\mathrm{d}X_t = \mu X_t \mathrm{d}t + \sigma\sqrt{X_t}\mathrm{d}B_t, \quad (\mu \in \mathbb{R}, \sigma > 0)$$

has no closed form solution.

Example 6.7 (Brownian Bridge)**.**

$$\mathrm{d}X_t = \frac{b - X_t}{T - t}\mathrm{d}t + \mathrm{d}B_t \quad (0 \le t \le T)$$

will gives $X_0 = a$ and $X_{T-} = b$. It has two realizations for $a = b = 0$:

$$X_t = B_t - \frac{t}{T}B_T \quad (0 \le t \le T)$$
$$X_t = (T - t)\frac{B_t}{T(T - t)} \quad (0 \le t \le T).$$

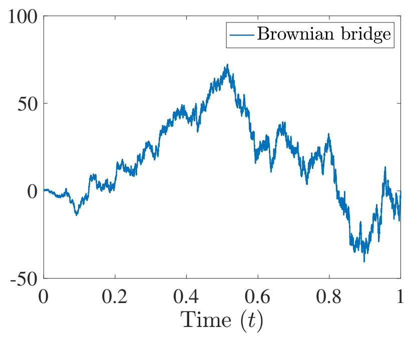

Figure 3: Realization of Brownian bridge in $[0, 1]$ with $B_0 = B_1 = 0$.

Example 6.8 (Shiryaev-Robert Equation)**.**

$$\mathrm{d}X_t = (c_1 + \mu X_t)\mathrm{d}t + (c_2 + \sigma X_t)\mathrm{d}B_t$$

has the solution

$$X_t = S_t\left(x + (c_1 - \sigma c_2)\int_0^t \frac{\mathrm{d}s}{S_s} + c_2\int_0^t \frac{\mathrm{d}B_s}{S_s}\right)$$

where S_t is the geometric Brownian motion

$$S_t = \exp\left(\sigma B_t + \left(\mu - \sigma^2/2\right)t\right).$$

Example 6.9 (Sequential testing equation)**.**

$$\mathrm{d}X_t = \gamma X_t(1 - X_t)\mathrm{d}B_t \quad (\gamma > 0).$$

This equation arises when we would like to observing either a Brownian motion with zero drift or with drift 1. Our task is to stop the observation as soon as possible and with the maximum probability to determine which stochastic process we are observing.

Example 6.10 (Quickest detection equation)**.**

$$\mathrm{d}X_t = \lambda(1 - X_t)\mathrm{d}t + \sigma X_t(1 - X_t)\mathrm{d}B_t.$$

This equation arises when we are observing a Brownian motion with no drift, at a random/unobservable time, a drift appear, and our task is to stop as soon as possible.

Example 6.11 (The Wright-Fisher equation)**.**

$$\mathrm{d}X_t = \varrho X_t(1 - X_t)\mathrm{d}t + \sqrt{X_t(1 - X_t)}\mathrm{d}B_t, \quad (\varrho \geq 0), \tag{6.1}$$

$$\mathrm{d}X_t = -\gamma X_t\mathrm{d}t + \sqrt{X_t(1 - X_t)}\mathrm{d}B_t, \quad (\gamma \geq 0). \tag{6.2}$$

For (6.1), $X_t \in [0, 1]$ represents the fraction of A-type gene within a population having both A-type and B-type genes subject to mutation (from one type to another). For (6.2), it describes the case of *one-way-mutation* only.

6.2 Existence and Uniqueness of the SDE

Throughout $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ is a filtered probability space where $\mathcal{N} = \{N \in \mathscr{F} \mid \mathbb{P}(N) = 0\} \subseteq \mathscr{F}_t$ for all $t \geq 0$. Let $B = (B_t)_{t\geq 0}$ be a standard m-dimensional Brownian motion with respect to $(\mathscr{F}_t)_{t\geq 0}$. Consider the *Stochastic Differential Equation*

$$\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t \quad (X_0 = x) \tag{1}$$

or equivalently

$$X_t = x + \int_0^t \mu(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s \tag{2}$$

where $\mu : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^{m\times n}$ are measurable functions. Also, X_s is n-dimensional and B_s is m-dimensional. Therefore, if we write

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad X_t = \begin{bmatrix} X_t^1 \\ \vdots \\ X_t^n \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix}, \quad B_t = \begin{bmatrix} B_t^1 \\ \vdots \\ B_t^n \end{bmatrix}, \quad \sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nm} \end{bmatrix},$$

where $\mu_i : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ and $\sigma_{ij} : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$. Then the Eq. (2) defines a sequence of SDE as shown below

$$\begin{cases} X_t^1 = x_1 + \int_0^t \mu_1(s, X_s)\,\mathrm{d}s + \sum_{j=1}^m \int_0^t \sigma_{1j}(s, X_s)\,\mathrm{d}B_s^j, \\ X_t^2 = x_2 + \int_0^t \mu_2(s, X_s)\,\mathrm{d}s + \sum_{j=1}^m \int_0^t \sigma_{2j}(s, X_s)\,\mathrm{d}B_s^j, \\ \vdots \\ X_t^n = x_n + \int_0^t \mu_n(s, X_s)\,\mathrm{d}s + \sum_{j=1}^m \int_0^t \sigma_{nj}(s, X_s)\,\mathrm{d}B_s^j. \end{cases}$$

We will now state the basic existence and uniqueness result for the SDE (1) when μ and σ are Lipschitz functions in the state variable[7]. This result is due to Itô (1946). It may be viewed as an extension of the well-known technique (*Picard iteration*) in the theory of ordinary differential equations.

This note is devoted to the investigation of a stochastic integral equation:

(1) $\quad x(t,\omega)=c+\int_0^t a(\tau,x(\tau,\omega))\,d\tau+\int_0^t b(\tau,x(\tau,\omega))\,d_\tau g(\tau,\omega),$

which is closely related to the researches of Markoff process by many authors, especially by S. Bernstein,[2] A. Kolmogoroff,[3] and W. Feller.[4]

Theorem. Let $a(t,x)$ ***and*** $b(t,x)$ ***be continuous in*** (t,x) ***and satisfy***

(2) $\quad |a(t,x)-a(t,y)|\leqq A\,|x-y|$, **(3)** $|b(t,x)-b(t,y)|\leqq B\,|x-y|$, ***where*** $0\leqq t\leqq 1$ ***and*** $-\infty<x,y<\infty$. ***Then the integral cquation (1) has one and only one continuous (in t with P-measure 1) solution.***

Figure 4: Original text from Itô in [3, 1946, Thm. 1]

Recall that a function $f:\mathbb{R}^n\to\mathbb{R}^n$ is said to be Lipschitz if there exist a constant $C>0$ such that

$$|f(x)-f(y)|\le C|x-y|,\quad \forall x,y\in\mathbb{R}^n.$$

Theorem 6.12 (Itô 1946)**.** *Suppose that μ and σ satisfy*

$$|\mu(t,x)|+|\sigma(t,x)|\le C(1+|x|) \qquad \text{(Linear growth)}$$

$$|\mu(t,x)-\mu(t,y)|+|\sigma(t,x)-\sigma(t,y)|\le D|x-y| \qquad \text{(Lipschitz condition)}$$

for all $x,y\in\mathbb{R}^n$ and all $t\in[0,T]$, with some constant $C,D>0$, where $T>0$ is given and fixed. Then the SDE (1) has a unique solution $X=(X_t)_{t\ge 0}$ satisfying:

1. *X_t is $\mathscr{F}_t$ measurable $\forall t\in[0,T]$.*
2. *$t\mapsto X_t$ is continuous on $[0,T]$ $\mathbb{P}$-a.s..*
3. $\mathbb{E}[\int_0^T |X_t|^2\,\mathrm{d}t]<\infty$.

Remark 6.13.

1. The solution X formed above is called a *strong solution* because B is given in advance and the solution X constructed is adapted to the filtration given.
2. If μ and σ are given and we are allowed to construct B together with a filtration, we speak of a weak solution.
3. The uniqueness obtained above is called strong or pathwise uniqueness. Weak uniqueness means that any two solutions are identical in law.

[7] X_s, in contrast with the time variable t.

7 Application in Option Pricing Theory (Optional)

Option pricing theory constitutes a central part of the modern mathematical finance. Stochastic calculus embodies a central tool in the modern option pricing theory. The present section aims at highlighting the fact by discussing a fundamental example in option pricing theory. This example serves a motivation of further development of the course.

7.1 The Black–Scholes model

Consider the Black–Scholes model for a financing market consisting a *riskless* bank account with value $B = (B_t)_{t\geq 0}$, and a *risky* stock with value $S = (S_t)_{t\geq 0}$. The equations which govern B–S model are respectively given by:

$$\mathrm{d}B_t = rB_t\,\mathrm{d}t, \tag{7.1}$$
$$\mathrm{d}S_t = \mu S_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t, \tag{7.2}$$

where $r > 0$ is the *interest rate*, $\mu \in \mathbb{R}$ is the *appreciation rate*, $\sigma > 0$ is the *volatility coefficient*, and $W = (W_t)_{t\geq 0}$ is a standard Brownian motion defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Notice that, we can compute the solution of (7.1) and (7.2) explicitly.

- The bank account is *deterministic* and is given by

  $$B_t = B_0 \mathrm{e}^{rt}, \tag{7.3}$$

  where $B_0 > 0$.

- The stock value (price) is *random* and is given as the geometric Brownian motion:

  $$S_t = S_0 \exp\left(\sigma W_t + \left(\mu - \frac{\sigma^2}{2}\right)t\right), \tag{7.4}$$

  where $S_0 > 0$. (Recall Example 4.7)

By convention, we assume $B_0 = 1$ in (7.3).

Remark 7.1. In order to reach the central point of our exposition as simple as possible, we will drop some regularity assumptions (on measurability, integrability and etc.) in the sequal.

7.2 American Options

Given a reward process $f = (f_t)_{0\leq t\leq T}$, consider the option of American type as a contract between the seller and a buyer that entitles the buyer to exercise the option at any (stopping) time $\tau \in [0, T]$ and receive the payment f_τ from the seller.

After selling the option at a price x, the seller has at disposal *self-financing strategies* $\Pi = (\beta_t, \gamma_t)$ with (non-negative) consumption $C = (C_t)$ which, after starting with $X_0^{\Pi,C} = x$, at time t brings him/her the (non-negative) value:

$$\begin{aligned} X_t^{\Pi,C} &= \beta_t B_t + \gamma_t S_t \\ &= x + \int_0^t \beta_s \, \mathrm{d}B_s + \int_0^t \gamma_s \, \mathrm{d}S_s - C_t, \quad \text{Self-financing with consumption.} \end{aligned} \tag{7.5}$$

if evaluate at time 0, or equivalently, the discounted (real) value

$$Y_t^{\Pi,C} = \frac{X_t^{\Pi,C}}{B_t} = \beta_t + \gamma_t \frac{S_t}{B_t} \tag{7.6}$$

if evaluate at time t.

The central question about such an option contract are:

(A) What is the "fair price" x?

(B) What is the optimal strategy $\Pi^* = (\beta_t^*, \gamma_t^*)$ with consumption $C^* = (C_t^*)_{t \geq 0}$, and what is the optimal exercise time τ_*?

7.2.1 *Fair price (No-arbitrage price)*

A self-financing strategy $\Pi = (\beta_t, \gamma_t)$ with consumption $C = (C_t)$ is called a *hedge* (with respect to x and f given), if $X_0^{\Pi,C} = x$ and

$$X_t^{\Pi,C} \geq f_t, \quad \mathbb{P}\text{-a.s.} \tag{7.7}$$

for all $t \in [0, T]$. The minimal x, denoted by $V^*(f)$, for which there exists a hedge with respect to f is called the "fair price" of the option (*the no-arbitrage price*).

In order to determine the fair price $V^*(f)$, and answer the question (A) and (B) above, the following facts and observations are essential.

- Firstly, the requirement (7.7) is invariant under a change of measure from $\mathbb{P}$ to $\widetilde{\mathbb{P}}$, as long as $\widetilde{\mathbb{P}}$ and $\mathbb{P}$ are equivalent in the sense that $\widetilde{\mathbb{P}}$ and $\mathbb{P}$ have the same null set in $\mathscr{F}_t := \sigma(W_s \mid s \leq t)$ for $t \in [0, T]$. More precisely, this means that $\widetilde{\mathbb{P}}(N) = 0$ if and only if $\mathbb{P}(N) = 0$ whenever N belongs to $\mathscr{F}_T$.

  By the Girsanar theorem, there exist such a measure on $\mathscr{F}_T$ satisfying yet another desired property described below. The measure $\widetilde{\mathbb{P}}$ is given by

$$\mathrm{d}\widetilde{\mathbb{P}} = \underbrace{\exp\left(-\frac{\mu - r}{\sigma} W_T - \frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 T\right)}_{:= Z_T} \mathrm{d}P \tag{7.8}$$

which is understood in the sense that $\widetilde{\mathbb{P}}(F) = \int_F Z_T \, d\mathbb{P}$ for $F \in \mathscr{F}_T$.

Under $\widetilde{\mathbb{P}}$, the process $\widetilde{W}_t = W_t + ((\mu - r)/\sigma)\, t$ is a standard Brownian motion (by the Girsanar theorem), and the process (7.6) admits the following representation (obtained by Itôś formula):

$$Y_t^{\Pi,C} = Y_0^{\Pi,C} + \int_0^t \sigma\gamma_s \frac{S_s}{B_s}\, d\widetilde{W}_s - \int_0^t \frac{dC_s}{B_s}, \tag{7.9}$$

where the first integral defines a (continuous) local martingale and the second integral defines an increasing process (which is zero if $C_s = 0$ for $s \in [0, t]$). For this reason, the measure $\widetilde{\mathbb{P}}$ is called *an equivalent martingale measure*.

In fact, such a measure $\widetilde{\mathbb{P}}$ is unique in the present setting for B and S. Its *existence* guarantees that the market is *arbitrage free*[8] in the sense that for every reward process f, there is a hedge $\Pi = (\beta_t, \gamma_t)$ satisfying (7.7) for some x large enough; its uniqueness is expressed by saying that the market is complete in the sense that equality holds in (7.7) at the optimal exercise time τ_*. When specialised to the case when the exercise is only allowed at time T (the option is then said to be of *European type* and the consumption C_t in (7.5) is set to be zero), these two statements constitutes the *Fundamental Theorem of Asset Pricing*: $\widetilde{\mathbb{P}}$ **(equivalent martingale measure) exist if and only if the market is arbitrage free; the $\widetilde{\mathbb{P}}$ is unique if and only if the market is complete.**

- Secondly, since we want a minimal x for which there is a hedge $\Pi = (\beta_t, \gamma_t)$ with consumption $C = (C_t)$ for which the value $X_t^{\Pi,C}$ satisfying (7.7) (or the value $Y_t^{\Pi,C}$ satisfying $Y_t^{\Pi,C} \geq f_t/B_t$ with $\widetilde{\mathbb{P}}$-a.s. for all $t \in [0, T]$) are "minimal" in some sense.

- Thirdly, it is well-known in "Optimal Stopping Theory" that the *smallest supermartingale* (under $\widetilde{\mathbb{P}}$) which dominates f_t/B_t in $\mathbb{P}$-a.s. on $[0, T]$ is given as the *Snell envelope*:

$$Y_t^* = \underset{\tau \in [t,T]}{\operatorname{esssup}}\, \widetilde{\mathbb{E}}\,[f_\tau/B_\tau \mid \mathscr{F}_t] \tag{7.10}$$

for $t \in [0, T]$, where the essential supremum is taken over all stopping times τ taking values in $[t, T]$.

By the *Doob-Meyer decomposition* (a well-known result in martingale theory) we have:

$$Y_t^* = Y_0^* + M_t^* - A_t^* \tag{7.11}$$

where $M^* = (M_t^*)$ is a (continuous) local martingale (under $\widetilde{\mathbb{P}}$), and $A^* = (A_t^*)$ is an increasing process, both started at zero. Moreover, by the ItôClark theorem, we have:

$$M_t^* = \int_0^t \alpha_s \, d\widetilde{W}_s \tag{7.12}$$

[8]We cannot create money from nothing

for some process $\alpha = (\alpha_t)$.

Note: We formulate Y_t^* process, then we decompose it into M_t^* and A_t^* by Doob-Meyer decomposition, then we find α_s by stochastic integration through the underlying $\widetilde{W}_t$ process.

- Finally, the Snell envelope (Y_t^*) is known to be a martingale before hitting f_t/B_t, and thus during this period $A_t^* = 0$ as $\mathbb{P}$-a.s..

7.2.2 *Rational/Optimal Performance*

From the argument just presented, it is evident that the optimal $\Pi^* = (\beta_t^*, \gamma_t^*)$, $C^* = (C_t^*)$ and τ_* are obtained by identifying:

$$\int_0^t \sigma\gamma_s \frac{S_s}{B_s}\, \mathrm{d}\widetilde{W}_s = M_t^*, \tag{7.13}$$

$$\int_0^t \frac{\mathrm{d}C_s}{B_s} = A_t^*, \tag{7.14}$$

This gives the following *explicit answers* to the question (A) and (B) stated above:

$$V^*(f) = \sup_{0\le\tau\le T} \widetilde{\mathbb{E}}\left[f_\tau/B_\tau\right], \quad \text{taking } t = 0 \text{ at (7.10)} \tag{7.15}$$

$$\beta_t^* = Y_t^* - \frac{\alpha_t}{\sigma}, \quad \text{From (7.6) by writing } \alpha_s = \sigma\gamma_s S_s/B_s \tag{7.16}$$

$$\gamma_t^* = \frac{\alpha_t}{\sigma}\frac{B_t}{S_t}, \quad \text{From (7.12) and (7.13)} \tag{7.17}$$

$$C_t^* = \int_0^t B_s\, \mathrm{d}A_s^*. \tag{7.18}$$

$$\tau_* = \inf\left\{t \in [0,T] \mid Y_t^* = f_t/B_t\right\}. \tag{7.19}$$

The equation (7.18) can be derived from (7.14) by

$$\frac{\mathrm{d}C_t}{B_t} = \mathrm{d}A_t^*, \quad \Rightarrow \quad \mathrm{d}C_t = B_t\, \mathrm{d}A_t^*, \quad \Rightarrow \quad C_t = \int_0^t B_s\, \mathrm{d}A_s^*.$$

Moreover, we have

The stopping time τ_* is optimal in (7.15) (i.e. the supremum is attained at τ_*), and τ_* is pointwise ($\widetilde{\mathbb{P}}$-a.s.) the *smallest stopping time* satisfying this property.

From (7.19) and the martingale property of Snell's envelope noted above, we see that if the *buyer acts rationally* and exercise the option at time τ_*, then *there will be no consumption* for the seller, i.e.

$$C_t^* = 0 \quad \text{for } t \in [0,T] \text{ as } \mathbb{P}\text{-a.s..}$$

Thus, the "fair price" $V^*(f)$ is indeed "fair" from this standpoint as well.

Finally, since $\mathfrak{Law}(S(\mu) \mid \widetilde{\mathbb{P}}) = \mathfrak{Law}(S(r) \mid \mathbb{P})$, it follows from (7.15) that the "fair price" $V^*(f)$ does not depend on μ (neither does the optimal strategy $\Pi^* = (\beta_t^*, \gamma_t^*)$ with consumption $C^* = (C_t^*)$, nor does the optimal stopping time τ_*). The latter property was the key to the success of the option pricing mechanism as the buyer and the seller *did not have to agree* on the actual value of μ (which is difficult to determine/estimate).

Remark 7.2. It should be noted that the process $Y^* = (Y_t^*)$ (and therefore M^* and A^* too) is computable (at least in principle) and known a priori before the option contract has been signed (in much the same way as the "fair price" itself). (This is important since $\Pi^* = (\beta_t^*, \gamma_t^*)$ with consumption $C^* = (C_t^*)$ is expressed in terms of Y^*.)

In fact, the problem of an explicit computation of the Snell envelope $Y^* = (Y_t^*)$ is closely linked to the problem of computing the "fair price" $V^*(f)$.

Solving the optimal stopping problem (7.15) in a *Markovian* setting, we get the "fair price" $V^*(f)$ as a function of the *initial position* of the underlying Markov process. Composing this function with the Markov process itself, we obtain (Y_t^*) as the *smallest supermartingale* which dominates the gain process (f_t/B_t) (To solve (7.15) one usually formulates a *free-boundary problem*).

7.3 European Option

The only difference (in comparison with American options) is that *the buyer can exercise the option only at time T* (not before).

The problem is simplier, since there is no need to introduce consumption, i.e. $C_t \equiv 0$ for all $t \in [0, T]$. Thus (7.9) holds without the final integral (w.r.t. $\mathrm{d}C_s$), and (7.10) holds without esssup and with f_T/B_T in place of f_τ/B_τ. In (7.11), there is no A_t^* term, i.e. it will be martingale, and (7.12) and (7.13) hold as well.

The formula (7.16) and (7.17) are the same, and the "fair price" (7.15) is given by:

$$V(f) = \widetilde{\mathbb{E}}\left[f_T/B_T\right]. \tag{7.20}$$

Specialising to $f_T = (S_T - K)^+$ (call option), the expression for (7.20) is called the **Black–Scholes formula**. It is the best known result of "Mathematical Finance".

References

[1] Richard Durrett. *Stochastic Calculus: A Practical Introduction*. Probability and Stochastics Series. CRC Press, Boca Raton, 1996. ISBN 978-0-8493-8071-6. (Cited on p. 59.)

[2] Jean-François Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*. Springer London, Limited, 2016. ISBN 9783319310893. (Cited on p. 28.)

[3] Kiyosi Itô. On a stochastic integral equation. *Proceedings of the Japan Academy*, 22 (1-4):32–35, 1946. (Cited on p. 66.)