Notes on PSTAT 223

Haosheng Zhou

Sept, 2022

Mid-term: Nov. 2, take-home

Final: Dec. 7, 8-11 am in class

BM

Approximation by Random Walk

Set $X_1, .., X_n, ...$ *i.i.d.* with probability $\frac{1}{2}$ taking value $\sqrt{\varepsilon}$ and probability $\frac{1}{2}$ taking value $-\sqrt{\varepsilon}$ with $\lfloor \frac{t}{\varepsilon} \rfloor = n$ and set $S_t = \sum_{i=1}^n X_i$, then by CLT

$$\frac{S_t}{\sqrt{t}} \xrightarrow{d} N(0,1) \tag{1}$$

which gives the informal approximation to the BM (no path regularity mentioned).

BM existence is ensured by Kolmogorov's extension theorem and the Kolmogorov's lemma (the first use marginal distributions to construct the continuous-time stochastic process with the same finite dimensional distributions and the second ensures the regularity of path so that it's continuous).

Property of BM

- BM not differentiable. If differentiable on $[0, T]$ then total variation is finite. Note that with $T = n \cdot \Delta t$

$$\sum_i |B_{t_i} - B_{t_{i-1}}| \sim n\mathbb{E}|B_{\Delta t}| \sim \frac{T}{\Delta t}\sqrt{\Delta t} \to \infty \ (n \to \infty, \Delta t \to 0) \tag{2}$$

- Quadratic variation of BM on $[0, T]$ is just T. Partition $\Delta : 0 = t_0 < t_1 < ... < t_n = T$ with $||\Delta|| = \sup_i |t_i - t_{i-1}|$, then

$$\sum_{i=1}^n |B_{t_i} - B_{t_{i-1}}|^2 \xrightarrow{L_2} T \ (n \to \infty) \tag{3}$$

actually such limit can be lifted to *a.s.* sense.

- Levy's characterization of BM: $B_0 = 0, B_t$ is continuous $a.s., \mathbb{E}(e^{iu(B_t - B_s)}|\mathscr{F}_s) = e^{-\frac{1}{2}u^2(t-s)}$
- Markov Property
- Martingale

Refer to HW 1 for more properties, or GTM 274

Week 2

Ito's Integral

Differential form of SDE:

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dB_t \tag{4}$$

with initial value $X_0 = x_0$. But it's actually not rigorous since dB_t is not well defined. Instead, use the integral form:

$$X_t = x_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s \tag{5}$$

for well definition. Here b is the **drift** coefficient (mean return, controls the speed of evolution), σ is the **diffusion** coefficient (volatility, controls the size of the noise).

Define for $0 \leq S \leq T$, the space of all measurable (actually progressive), adapted and L^2 process.

$$V(S, T) = \{f : [0, +\infty) \times \Omega \to \mathbb{R}\} \tag{6}$$

such that $(t, \omega) \to f_t(\omega)$ is measurable w.r.t. $\mathscr{B}_{\mathbb{R}_+} \times \mathscr{F}$, $f_t \in \mathscr{F}_t$ and $\mathbb{E}\left(\int_S^T f_t^2\, dt\right) < \infty$ (say $f \in L^2([S, T] \times \Omega)$ since $\langle f_t, g_t \rangle = \mathbb{E}\left(\int_S^T f_t g_t\, dt\right)$ is the inner product on such space under $a.s.$ sense). Our goal is to **define the stochastic integral $I(f) = \int_S^T f_t\, dB_t$ for a general process f in such Hilbert space**.

We follow several steps, first consider defining this for a "simple" process and then extend it to general processes. The main thought is to find a **dense subset** of the Hilbert space and define the stochastic integral on such dense subset to prove that it's actually an **isometry**. After that, **extend** it as the isometry on the whole Hilbert space.

Step 0: Consider **elementary process** defined by $\varphi_t(\omega) = \sum_{j=1}^{\infty} e_j(\omega) \mathbb{I}_{[t_j, t_{j+1})}(t)$ **where** $e_j \in \mathscr{F}_{t_j}, e_j \in L^2(\Omega)$ and n, t_i are all fixed. t_j is the j-th **dyadic number** within $[S, T]$, i.e. it is $\frac{j}{2^n}$ if such value is in $[S, T]$. If such value is less than S, $t_j = S$. If such value is greater than T, $t_j = T$. (In simple words, only care about the dyadic partition within $[S, T]$).

Remark. *Note that the elementary process is a generalization of the step function, replacing the fixed constant with a random variable. The important point here is that this random variable here is* ***measurable w.r.t. the left endpoint*** *t_j. Changing the measurability here as the one w.r.t. the midpoint $e_j \in \mathscr{F}_{\frac{t_j + t_{j+1}}{2}}$ results in a different integration scheme.*

Naturally, the stochastic integral for elementary process is defined as

$$\int_S^T \varphi_t\, dB_t = \sum_{j=1}^{\infty} e_j (B_{t_{j+1}} - B_{t_j}) \tag{7}$$

Remark. *Let's show a counterexample here why $e_j \in \mathscr{F}_{t_j}$ can't be removed.*

For $\varphi_t^{(1)}$, take $e_j = B_{t_j}$. By definition, the integral should be $\sum_{j=1}^{\infty} B_{t_j}(B_{t_{j+1}} - B_{t_j})$, its expectation is

$$\mathbb{E}\sum_{j=1}^{\infty} B_{t_j}(B_{t_{j+1}} - B_{t_j}) = \sum_{j=1}^{\infty} \mathbb{E}B_{t_j}(B_{t_{j+1}} - B_{t_j}) \tag{8}$$

$$= \sum_{j=1}^{\infty} \mathbb{E}B_{t_j} \cdot \mathbb{E}(B_{t_{j+1}} - B_{t_j}) = 0 \tag{9}$$

For $\varphi_t^{(2)}$, take $e_j = B_{t_{j+1}}$. By definition, the integral should be $\sum_{j=1}^{\infty} B_{t_{j+1}}(B_{t_{j+1}} - B_{t_j})$, its expectation is

$$\mathbb{E}\sum_{j=1}^{\infty} B_{t_{j+1}}(B_{t_{j+1}} - B_{t_j}) = \sum_{j=1}^{\infty} \mathbb{E}B_{t_{j+1}}(B_{t_{j+1}} - B_{t_j}) \tag{10}$$

$$= \sum_{j=1}^{\infty} \mathbb{E}(B_{t_{j+1}} - B_{t_j})^2 + \mathbb{E}B_{t_j}(B_{t_{j+1}} - B_{t_j}) \tag{11}$$

$$= \sum_{j=1}^{\infty} \mathbb{E}(B_{t_{j+1}} - B_{t_j})^2 = T - S \tag{12}$$

As we can see, a slight change in measurability at the endpoints results in the large change of the integral. This is due to the infinite total variation of BM, and this tells us from another perspective that the Lebesgue-Stieljes integration does not work for BM any longer!

The following lemma shows that such definition is actually an isometry between Hilbert spaces.

Lemma 1. *If φ_t is a bounded elementary process, then $\mathbb{E}\left(\int_S^T \varphi_t\, dB_t\right)^2 = \mathbb{E}\left(\int_S^T \varphi_t^2\, dt\right)$. This means that $||\int_S^T \varphi_t\, dB_t||_{L^2(\Omega)} = ||\varphi_t||_{L^2([S,T]\times\Omega)}$, the* ***Ito's isometry for elementary process****.*

Proof.

$$\mathbb{E}\left(\int_S^T \varphi_t\, dB_t\right)^2 = \mathbb{E}\left(\int_S^T \sum_{j=1}^{\infty} e_j \mathbb{I}_{[t_j,t_{j+1})}\, dB_t\right)^2 \tag{13}$$

$$= \mathbb{E}\left(\sum_{j=1}^{\infty} e_j(B_{t_{j+1}} - B_{t_j})\right)^2 \tag{14}$$

$$= \mathbb{E}\left(\sum_{i,j=1}^{\infty} e_i e_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})\right) \tag{15}$$

$$= \sum_{i,j=1}^{\infty} \mathbb{E}\left(e_i e_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})\right) \tag{16}$$

$$= \sum_{i=1}^{\infty} \mathbb{E}\left(e_i^2 (B_{t_{i+1}} - B_{t_i})^2\right) + 2\sum_{i<j}^{\infty} \mathbb{E}\left(e_i e_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})\right) \tag{17}$$

Note that the interchange of expectation and the infinite sum is the consequence of Fubini's theorem with the fact that φ_t is $a.s.$ bounded, so $\forall i, |e_i| \leq M$ $a.s.$ are uniformly bounded.

$$\sum_{i,j=1}^{\infty} \mathbb{E}(|e_i e_j| \cdot |B_{t_{i+1}} - B_{t_i}| \cdot |B_{t_{j+1}} - B_{t_j}|) \leq M^2 \sum_{i,j=1}^{\infty} \mathbb{E}(|B_{t_{i+1}} - B_{t_i}| \cdot |B_{t_{j+1}} - B_{t_j}|) \tag{18}$$

$$= M^2 \sum_{i=1}^{\infty} \mathbb{E}(B_{t_{i+1}} - B_{t_i})^2 + 2M^2 \sum_{i<j} \mathbb{E}(|B_{t_{i+1}} - B_{t_i}|) \cdot \mathbb{E}(|B_{t_{j+1}} - B_{t_j}|) \tag{19}$$

$$= M^2 \sum_{i=1}^{\infty} (t_{i+1} - t_i) + 2M^2 \sum_{i<j} \frac{2}{\pi} \sqrt{(t_{i+1} - t_i)(t_{j+1} - t_j)} \tag{20}$$

$$\leq M^2(T-S) + \frac{4M^2}{\pi}(T-S)^2 2^{n-1} < \infty \tag{21}$$

Also note that for $i < j$

$$\mathbb{E}\left(e_i e_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})\right) = \mathbb{E}\left[\mathbb{E}\left(e_i e_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j}) | \mathscr{F}_{t_j}\right)\right] \tag{22}$$

$$= \mathbb{E}\left[e_i e_j (B_{t_{i+1}} - B_{t_i}) \mathbb{E}\left(B_{t_{j+1}} - B_{t_j} | \mathscr{F}_{t_j}\right)\right] = 0 \tag{23}$$

and use the independency of e_i and $B_{t_{i+1}} - B_{t_i}$

$$\mathbb{E}\left(\int_S^T \varphi_t \, dB_t\right)^2 = \sum_{i=1}^{\infty} \mathbb{E}\left(e_i^2 (B_{t_{i+1}} - B_{t_i})^2\right) \tag{24}$$

$$= \sum_{i=1}^{\infty} \mathbb{E} e_i^2 \cdot (t_{i+1} - t_i) \tag{25}$$

$$= \mathbb{E} \sum_{i=1}^{\infty} e_i^2 \cdot (t_{i+1} - t_i) \tag{26}$$

$$= \mathbb{E}\left(\int_S^T \varphi_t^2 \, dt\right) \tag{27}$$

the interchange of the expectation and the infinite sum is due to the non-negativity and the Fubini theorem.

□

Remark. *The Ito's isometry provides a link between stochastic integral and Lebesgue-Stieljes integral in that the* L^2 *norm of the stochastic integral* $\int_S^T \varphi_t \, dB_t$ *is equal to the expectation of a Lebesgue-Stieljes integral* $\int_S^T \varphi_t^2 \, dt$ *that only integrates w.r.t. the time.*

To extend the definition of Ito's integral and the Ito's isometry property onto the whole process space $V(S,T)$, the core is to prove that elementary processes are actually a dense subset of $V(S,T)$.

Step 1: Consider any **bounded continuous process** $g_t \in V(S,T)$, it's quite natural to notice that there exist bounded elementary process $\varphi_t^{(n)}$ such that $||g_t - \varphi_t^{(n)}||_{L^2([S,T]\times\Omega)} \to 0$ $(n \to \infty)$. This is done by construction

$$\varphi_t^{(n)} = \sum_{j=1}^{\infty} g_{t_j} \mathbb{I}_{[t_j, t_{j+1})}(t) \tag{28}$$

where t_i are truncated dyadic numbers in the interval $[S,T]$, i.e. $\frac{i}{2^n}$ if it's in the interval and the endpoint if not. The L^2 convergence is ensured by

$$\begin{aligned}
\mathbb{E}\int_S^T \left(g_t - \varphi_t^{(n)}\right)^2 dt &= \mathbb{E}\int_S^T \left(\sum_{j=1}^{\infty}(g_t - g_{t_j})\mathbb{I}_{[t_j,t_{j+1})}(t)\right)^2 dt &(29)\\
&= \mathbb{E}\int_S^T \sum_{j=1}^{\infty}(g_t - g_{t_j})^2\mathbb{I}_{[t_j,t_{j+1})}(t)\, dt &(30)\\
&= \mathbb{E}\sum_{j=1}^{\infty}\int_{t_j}^{t_{j+1}}(g_t - g_{t_j})^2\, dt &(31)
\end{aligned}$$

Since g_t is continuous $a.s.$ for $t \in [S,T]$, it's uniformly continuous, so $\forall \varepsilon > 0, \exists \delta > 0, \forall x, y \in [S,T]$ $s.t.$ $|x-y| < \delta, |g_x - g_y| < \varepsilon$. Now for large enough n, exists ε, δ such that $\delta > \frac{1}{2^n}$, then

$$\begin{aligned}
\mathbb{E}\sum_{j=1}^{\infty}\int_{t_j}^{t_{j+1}}(g_t - g_{t_j})^2\, dt &\le \mathbb{E}\sum_{j=1}^{\infty}\varepsilon^2(t_{j+1} - t_j) &(32)\\
&= \varepsilon^2(T-S) &(33)\\
\mathbb{E}\int_S^T \left(g_t - \varphi_t^{(n)}\right)^2 dt \to 0 \ (n \to \infty) & &(34)
\end{aligned}$$

Remark. *The construction of $\varphi_t^{(n)}$ is just similar to that in the deterministic case (use step functions to approximate continuous bounded functions well enough), where we replace the fixed constant as a $\mathscr{F}_{t_j}$ measurable random variable g_{t_j}.*

Step 2: Extend the definition for all **bounded process** $h_t \in V(S,T)$. We hope to prove that there always exists a bounded continuous process $g_t^{(n)} \in V(S,T)$ such that $||h_t - g_t^{(n)}||_{L^2([S,T]\times\Omega)} \to 0$ $(n \to \infty)$.

The construction uses **the convolution w.r.t. the mollifier**. To be specific, let $\psi_n \ge 0$ be a continuous mollifier on $\mathbb{R}$ such that $\forall x \ge 0, \forall x \le -\frac{1}{n}, \psi_n(x) = 0$ and $\int_{\mathbb{R}} \psi_n(x)\, dx = 1$. Consider

$$g_t^{(n)} = \int_0^t \psi_n(s-t)\cdot h_s\, ds \tag{35}$$

It's then not hard to see that if ω is fixed,

$$|g_t^{(n)}| \le \sup_s |h_s| \cdot \int_0^t \psi_n(s-t)\, ds \le \sup_s |h_s| \tag{36}$$

so $g_t^{(n)}$ is uniformly bounded by the same bound of h_t.

The continuity follows from the fact that

$$\begin{aligned}
\left|g_{t+\Delta t}^{(n)} - g_t^{(n)}\right| &= \left|\int_0^{t+\Delta t} \psi_n(s-t-\Delta t)\cdot h_s\,ds - \int_0^t \psi_n(s-t)\cdot h_s\,ds\right| && (37)\\
&= \left|\int_0^t [\psi_n(s-t-\Delta t) - \psi_n(s-t)]\cdot h_s\,ds\right| + \left|\int_t^{t+\Delta t} \psi_n(s-t-\Delta t)\cdot h_s\,ds\right| && (38)\\
&\le \sup_s |h_s|\cdot\left(\left|\int_0^t [\psi_n(s-t-\Delta t) - \psi_n(s-t)]\,ds\right| + \left|\int_t^{t+\Delta t} \psi_n(s-t-\Delta t)\,ds\right|\right) && (39)\\
&= \sup_s |h_s|\cdot\left(\left|\int_{-t}^0 [\psi_n(u-\Delta t) - \psi_n(u)]\,du\right| + \left|\int_{-\Delta t}^0 \psi_n(u)\,du\right|\right) && (40)
\end{aligned}$$

Note that $\left|\int_{-\Delta t}^0 \psi_n(u)\,du\right| \to 0$ as $\Delta t \to 0$ since the continuity of ψ_n ensures its boundedness on $[-\Delta t, 0]$ and the integration is on a small enough range. For the other term, notice that ψ_n is uniformly continuous on $[-t-\Delta t, 0]$, so $\forall\varepsilon > 0, \exists\delta > 0, \forall u_1, u_2 \in [-t-\Delta t, 0],\ if\ |u_1 - u_2| < \delta, |\psi_n(u_1) - \psi_n(u_2)| < \varepsilon$.

Since we hope to investigate this term as $\Delta t \to 0$, there exists δ such that $\Delta t < \delta$, so

$$\left|\int_{-t}^0 [\psi_n(u-\Delta t) - \psi_n(u)]\,du\right| < \varepsilon t \tag{41}$$

is also small enough.

As a result, we have shown that the convoluted process $g_t^{(n)}$ is continuous and bounded. Now we only have to show the convergence in L^2.

$$\begin{aligned}
||h_t - g_t^{(n)}||_{L^2([S,T]\times\Omega)} &= \mathbb{E}\int_S^T \left(h_t - g_t^{(n)}\right)^2\,dt && (42)\\
&= \mathbb{E}\int_S^T \left(h_t - \int_{t-\frac{1}{n}}^t \psi_n(s-t)\cdot h_s\,ds\right)^2\,dt && (43)\\
&= \mathbb{E}\int_S^T \left(\int_{t-\frac{1}{n}}^t \psi_n(s-t)\cdot (h_t - h_s)\,ds\right)^2\,dt && (44)
\end{aligned}$$

note that if we can prove the property that

$$\int_S^T \left(\int_{t-\frac{1}{n}}^t \psi_n(s-t)\cdot (h_t - h_s)\,ds\right)^2\,dt \to 0\ (n\to\infty) \tag{45}$$

then the L^2 convergence is proved by the bounded convergence theorem since

$$\left|\int_S^T \left(\int_{t-\frac{1}{n}}^t \psi_n(s-t)\cdot(h_t-h_s)\,ds\right)^2 dt\right| \le \left(2\sup_s |h_s|\right)^2 \cdot \int_S^T \left(\int_{t-\frac{1}{n}}^t \psi_n(s-t)\,ds\right)^2 dt \tag{46}$$

$$= \left(2\sup_s |h_s|\right)^2 \cdot (T-S) < \infty \tag{47}$$

Let's now try to prove the fact that

$$\int_S^T \left(\int_{t-\frac{1}{n}}^t \psi_n(s-t)\cdot(h_t-h_s)\,ds\right)^2 dt \to 0 \ (n\to\infty) \tag{48}$$

actually holds. The **Minkowski integral inequality** gives

$$\left[\int_S^T \left(\int_{t-\frac{1}{n}}^t \psi_n(s-t)\cdot(h_t-h_s)\,ds\right)^2 dt\right]^{\frac{1}{2}} \le \int_{t-\frac{1}{n}}^t \left(\int_S^T \psi_n^2(s-t)\cdot(h_t-h_s)^2\,dt\right)^{\frac{1}{2}} ds \tag{49}$$

$$= \int_{-\frac{1}{n}}^0 \psi_n(u)\left(\int_S^T (h_t-h_{u+t})^2\,dt\right)^{\frac{1}{2}} du \tag{50}$$

$$\tag{51}$$

with the fact that the **translation in time h_t to h_{u+t} is continuous**, so $\forall\varepsilon>0, \exists\delta>0$, if $|u|<\delta$, then $|h_t-h_{t+u}|<\varepsilon$. Since $\int_{\mathbb{R}}\psi_n=1$, $\exists n_0$ such that for any $n>n_0$, $\int_{u\le-\delta}\psi_n(u)\,du<\varepsilon$. Split the integral above into two parts:

$$\int_{u>-\delta,u\in[-\frac{1}{n},0]} \psi_n(u)\left(\int_S^T (h_t-h_{u+t})^2\,dt\right)^{\frac{1}{2}} du + \int_{u\le-\delta,u\in[-\frac{1}{n},0]} \psi_n(u)\left(\int_S^T (h_t-h_{u+t})^2\,dt\right)^{\frac{1}{2}} du \tag{52}$$

$$\le \varepsilon\sqrt{T-S}\int_{u>-\delta,u\in[-\frac{1}{n},0]} \psi_n(u)\,du + (2\sup_s|h_s|)\cdot\sqrt{T-S}\int_{u\le-\delta,u\in[-\frac{1}{n},0]} \psi_n(u)\,du \tag{53}$$

$$\le \varepsilon\sqrt{T-S} + (2\sup_s|h_s|)\cdot\sqrt{T-S}\cdot\varepsilon \tag{54}$$

so the conclusion gets proved. This property is exactly the **approximation identity of the mollifier**, which tells us that when the support of the mollifier goes to 0, the convolution converges to the true mollified function in L^p sense.

Remark. *This is a classic technique to use in analysis. First write the difference between the function and the convolution as an integral form, then use inequalities to change the order of the integral and at last tear the integral into two parts. The first part is **near the singularity of the mollifier**, where the **continuity of translation** is used. The second part is **far away from the singularity of the mollifier**, where **the support can be shrunk** such that the integral of the mollifier is always small enough.*

Note that there might be issue proving the measurability of $g_t^{(n)}$ (not verified here)

Step 3: For **general process** $f_t \in V(S,T)$, always exists bounded process $h_t^{(n)} \in V(S,T)$ such that $||f_t - h_t^{(n)}||_{L^2([S,T]\times\Omega)} \to 0$ $(n \to \infty)$. The construction is given by simple truncation of function value that

$$h_t^{(n)} = f_t \wedge n \vee (-n) \tag{55}$$

It's then quite obvious to see that by using the Fubini theorem, the monotone convergence theorem for $(f_t + n)^2 \mathbb{I}_{f_t<-n} \searrow 0$ $(n \to \infty)$ and the monotone convergence theorem for $\mathbb{E}[(f_t + n)^2 \mathbb{I}_{f_t<-n}] \searrow 0$ $(n \to \infty)$:

$$\mathbb{E}\int_S^T \left(f_t - h_t^{(n)}\right)^2 dt = \mathbb{E}\int_S^T (f_t+n)^2 \mathbb{I}_{t:f_t<-n} + (f_t-n)^2 \mathbb{I}_{t:f_t>n}\, dt \tag{56}$$

$$= \mathbb{E}\int_{t\in[S,T],f_t<-n} (f_t+n)^2\, dt + \mathbb{E}\int_{t\in[S,T],f_t>n} (f_t-n)^2\, dt \tag{57}$$

$$= \int_S^T \mathbb{E}[(f_t+n)^2 \mathbb{I}_{f_t<-n}]\, dt + \int_S^T \mathbb{E}[(f_t-n)^2 \mathbb{I}_{f_t>n}]\, dt \to 0\ (n \to \infty) \tag{58}$$

As a result, we have proved that the set of all elementary process is a **dense** subset of $V(S,T)$. For any $f_t \in V(S,T)$, its stochastic integral is defined as the L^2 limit of the stochastic integral of the approximation process $\varphi_t^{(n)}$, i.e.

$$\mathbb{E}\int_S^T \left(f_t - \varphi_t^{(n)}\right)^2 dt \to 0\ (n \to \infty) \tag{59}$$

$$\int_S^T \varphi_t^{(n)}\, dB_t \overset{L^2(\Omega)}{\to} \int_S^T f_t\, dB_t\ (n \to \infty) \tag{60}$$

Remark. *Note that here* $\left\{\int_S^T \varphi_t^{(n)}\, dB_t\right\}_{n=1}^{\infty}$ *has to converge in* L^2 *sense since the set of elementary functions is a dense subset in the Hilbert space and that there's already an isometry on this dense subset. That is to say, if there are Hilbert spaces* H_1, H_2 *with norms* $||\cdot||_1, ||\cdot||_2$*, and* $D \subset H_1$ *is dense with* $f : D \to H_2$ *an isometry, then there exists an extension of* f *denoted* $g : H_1 \to H_2$ *to be an isometry.*

The construction is intuitive

$$\forall x \in H_1, \exists d_n \in D, ||d_n - x||_1 \to 0\ (n \to \infty) \tag{61}$$

$$g(x) \overset{def}{=} \lim_{n\to\infty} g(d_n) \tag{62}$$

to notice that the completeness of Hilbert space ensures that g *is well-defined (such limit exists).*

The Ito's isometry holds for general process in $V(S,T)$. Assume that f_t is a bounded process in $V(S,T)$, then exists a series of bounded elementary processes $\varphi_t^{(n)}$ such that $\varphi_t^{(n)} \overset{L^2([S,T]\times\Omega)}{\to} f_t$ $(n \to \infty)$.

$$\mathbb{E}\left(\int_S^T f_t\, dB_t\right)^2 = \lim_{n\to\infty} \mathbb{E}\left(\int_S^T \varphi_t^{(n)}\, dB_t\right)^2 = \lim_{n\to\infty} \mathbb{E}\left(\int_S^T \left(\varphi_t^{(n)}\right)^2 dt\right) = \mathbb{E}\left(\int_S^T f_t^2\, dt\right) \tag{63}$$

where the first and third equations are the consequences of L^2 convergence in space $L^2(\Omega)$ (space of r.v.) and $L^2([S,T]\times\Omega)$ (space of process) since L^2 convergence implies the convergence of L^2 norms. The second equation comes from the Ito's isometry for bounded elementary process.

Theorem 1. *$\forall f_t \in L^2([S,T]\times\Omega)$ (which means that $\mathbb{E}\int_S^T f_t^2\,dt < \infty$), $f_t \in \mathscr{F}_t$ and f_t progressive (which means that $(t,\omega)\to f_t(\omega)$ is measurable w.r.t. $\mathscr{B}_{[S,T]}\times\mathscr{F}$), then the definition of stochastic integral above gives* ***the Ito's isometry for general process*** *$\mathbb{E}(\int_S^T f_t\,dB_t)^2 = \mathbb{E}\int_S^T f_t^2\,dt < \infty$ a.s..*

Proof. We proved above that any bounded process $f_t \in V(S,T)$ satisfies the Ito's isometry. Now for general $f_t \in V(S,T)$, find a series of bounded process to approximate

$$\exists h_t^{(n)} \in V(S,T), \forall t, |h_t| \le M \tag{64}$$

$$h_t^{(n)} \overset{L^2([S,T]\times\Omega)}{\to} f_t \ (n\to\infty) \tag{65}$$

It's natural that

$$\mathbb{E}\left(\int_S^T f_t\,dB_t\right)^2 = \lim_{n\to\infty}\mathbb{E}\left(\int_S^T h_t^{(n)}\,dB_t\right)^2 = \lim_{n\to\infty}\mathbb{E}\left(\int_S^T \left(h_t^{(n)}\right)^2\,dt\right) = \mathbb{E}\left(\int_S^T f_t^2\,dt\right) \tag{66}$$

□

Example

Compute $\int_0^t B_s dB_s$.

First check if B_s is in the Hilbert space

$$\mathbb{E}\int_0^t B_s^2\,ds = \int_0^t \mathbb{E}B_s^2\,ds = \frac{t^2}{2} < \infty \tag{67}$$

so $B_s|_{s\in[0,t]} \in V(0,t)$.

Now we can first try to follow the definition of stochastic integral. Find elementary process

$$\varphi_s^{(n)} = \sum_{j=1}^{\infty} B_{s_j}\mathbb{I}_{[s_j,s_{j+1})}(s) \tag{68}$$

to approximate BM such that

$$\begin{aligned}
\mathbb{E}\int_0^t \left(\varphi_s^{(n)} - B_s\right)^2 ds &= \mathbb{E}\int_0^t \sum_{j=1}^{\infty} (B_{s_j} - B_s)^2 \mathbb{I}_{[s_j, s_{j+1})}(s)\, ds &(69)\\
&= \mathbb{E}\sum_{j=1}^{\infty} \int_{s_j}^{s_{j+1}} (B_{s_j} - B_s)^2\, ds &(70)\\
&= \sum_{j=1}^{\infty} \int_{s_j}^{s_{j+1}} \mathbb{E}(B_{s_j} - B_s)^2\, ds &(71)\\
&= \sum_{j=1}^{\infty} \int_{s_j}^{s_{j+1}} (s - s_j)\, ds &(72)\\
&= \sum_{j=1}^{\infty} \frac{(s_{j+1} - s_j)^2}{2} \to 0 \ (n \to \infty) &(73)
\end{aligned}$$

That's why the stochastic integral is formulated as

$$\sum_{j=1}^{\infty} B_{s_j}(B_{s_{j+1}} - B_{s_j}) \overset{L^2(\Omega)}{\to} \int_0^t B_s\, dB_s \ (n \to \infty) \tag{74}$$

A transformation gives

$$\begin{aligned}
\sum_{j=1}^{\infty} B_{s_j}(B_{s_{j+1}} - B_{s_j}) &= \sum_{j=1}^{\infty} \frac{B_{s_{j+1}} + B_{s_j}}{2}(B_{s_{j+1}} - B_{s_j}) - \sum_{j=1}^{\infty} \frac{B_{s_{j+1}} - B_{s_j}}{2}(B_{s_{j+1}} - B_{s_j}) &(75)\\
&\to \frac{B_t^2}{2} - \frac{t}{2} \ (n \to \infty) &(76)
\end{aligned}$$

note that the first limit is a telescoping and the second limit comes from the quadratic variation of BM. So we get

$$\int_0^t B_s\, dB_s = \frac{B_t^2}{2} - \frac{t}{2} \tag{77}$$

note that there's another **Ito's correction term** as the quadratic variation so **the chain rule does not hold!**

Wiener Integral

Just to mention, the wiener integral is defined as a special situation for stochastic integral, where the integrand is a deterministic function in time instead of a process. For function $f(s) \in C^1, \int_0^t f^2(s)\, ds < \infty$, define by the integration by parts:

$$\int_0^t f(s)\,dB_s = f\cdot B\Big|_0^t - \int_0^t B_s\,df(s) \tag{78}$$

$$= f(t)\cdot B_t - \int_0^t f'(s)\cdot B_s\,ds \tag{79}$$

This turns the stochastic integral into a Lebesgue-Stieljes integral. Now we want to show that this definition is consistent with that for the stochastic integral set up above. The continuity of f ensures that it's bounded and continuous on $[0,t]$, so it can be easily approximated by elementary process (which, in deterministic case, is just the step function). By the definition above,

$$\int_0^t f(s)\,dB_s = \lim_{n\to\infty}\sum_j f(t_j)(B_{t_{j+1}} - B_{t_j}) \tag{80}$$

for t_j as truncated dyadic numbers $\frac{j}{2^n}$ in $[0,t]$ and the limit is in the $L^2(\Omega)$ sense.

Let's assume that the truly effective dyadic numbers are $0 = t_0^n < t_1^n < ... < t_{p_n}^n \leq t < t_{p_n+1}^n$, so the sum for fixed n is actually

$$\sum_{j=0}^{p_n-1} f(t_j^n)\left(B_{t_{j+1}^n} - B_{t_j^n}\right) + f(t_{p_n}^n)\left(B_t - B_{t_{p_n}^n}\right) \tag{81}$$

$$= f(t_{p_n-1}^n)B_{t_{p_n}^n} - f(t_0^n)B_{t_0^n} - \sum_{j=1}^{p_n-1} B_{t_j^n}(f(t_j^n) - f(t_{j-1}^n)) + f(t_{p_n}^n)\left(B_t - B_{t_{p_n}^n}\right) \tag{82}$$

$$= f(t_{p_n}^n)B_t - \sum_{j=1}^{p_n} B_{t_j^n}(f(t_j^n) - f(t_{j-1}^n)) \tag{83}$$

here we use the Abel's lemma (summation by parts), and set $n\to\infty$ to find that

$$f(t_{p_n}^n)B_t \to f(t)B_t \tag{84}$$

$$\sum_{j=1}^{p_n} B_{t_j^n}(f(t_j^n) - f(t_{j-1}^n)) = \sum_{j=1}^{p_n} B_{t_j^n}\Delta t\frac{f(t_j^n) - f(t_{j-1}^n)}{\Delta t} \tag{85}$$

$$\xrightarrow{L^2} \int_0^t B_s f'(s)\,ds \tag{86}$$

The L^2 convergence comes from the fact that

$$\mathbb{E}\left[\sum_{j=1}^{p_n} B_{t_j^n} f'(t_j^n)\Delta t - \sum_{j=1}^{p_n} B_{t_j^n}\Delta t \frac{f(t_j^n) - f(t_{j-1}^n)}{\Delta t}\right]^2 \tag{87}$$

$$= \mathbb{E}\left[\sum_{j=1}^{p_n} B_{t_j^n}\Delta t \left(f'(t_j^n) - \frac{f(t_j^n) - f(t_{j-1}^n)}{\Delta t}\right)\right]^2 \tag{88}$$

$$\leq \mathbb{E}\left[\sum_{j=1}^{p_n} B_{t_j^n}\Delta t \cdot \varepsilon\right]^2 \tag{89}$$

$$\leq \varepsilon^2 t \tag{90}$$

note that the existence of the uniform ε for all t_j^n is ensured by the continuity of f', and the last equation can be derived by expanding the square.

As a result, we have proved that the Wiener integral is consistent with the definition of general stochastic integral (uniqueness of L^2 limit). After the work, we shall also notice from

$$\sum_j f(t_j)(B_{t_{j+1}} - B_{t_j}) \stackrel{L^2}{\to} \int_0^t f(s)\, dB_s \ (n \to \infty) \tag{91}$$

that the Wiener integral is actually the L^2 limit of a linear combination of independent Gaussian random variables (independent increment), so the Wiener integral is the L^2 limit of Gaussian. Note that L^2 limit of Gaussian must be Gaussian and the expectation and variance of the limit are just the limit of the expectation series and the variance series (this is from the pointwise limit of the characteristic function), so Wiener integral must be Gaussian.

$$B_{t_{j+1}} - B_{t_j} \sim N(0, \Delta t) \tag{92}$$

$$f(t_j)(B_{t_{j+1}} - B_{t_j}) \sim N(0, f^2(t_j)\Delta t) \tag{93}$$

$$\mathbb{E}\left(\int_0^t f(s)\, dB_s\right) = 0 \tag{94}$$

$$Var\left(\int_0^t f(s)\, dB_s\right) = \sum_j f^2(t_j)\Delta t = \int_0^t f^2(s)\, ds \tag{95}$$

or we can conclude from the Ito's isometry that $\mathbb{E}\left(\int_0^t f(s)\, dB_s\right)^2 = \mathbb{E}\left(\int_0^t f^2(s)\, ds\right) = \int_0^t f^2(s)\, ds$. As a result, we get the distribution of the Wiener integral

$$\int_0^t f(s)\, dB_s \sim N\left(0, \int_0^t f^2(s)\, ds\right) \tag{96}$$

Property of Stochastic Integral

The first one is **linearity**.

$$\forall c, d \in \mathbb{R}, \forall f_t, g_t \in V(S,T), \int_S^T (cf_s + dg_s)\, dB_s = c\int_S^T f_s\, dB_s + d\int_S^T g_s\, dB_s \tag{97}$$

Let's prove this with the definition using elementary process to approximate general process.

$$\exists \varphi_t^n, \varphi_t^n \overset{L^2([S,T]\times\Omega)}{\to} f_t, \exists \psi_t^n, \psi_t^n \overset{L^2([S,T]\times\Omega)}{\to} g_t \ (n\to\infty) \tag{98}$$

$$c\varphi_t^n + d\psi_t^n \overset{L^2([S,T]\times\Omega)}{\to} cf_t + dg_t \ (n\to\infty) \tag{99}$$

That's why by the definition of the stochastic integral,

$$\sum_j \left(ce_{t_j}^n(\varphi) + de_{t_j}^n(\psi)\right)\cdot\left(B_{t_{j+1}} - B_{t_j}\right) \overset{L^2}{\to} \int_S^T (cf_s + dg_s)\, dB_s \ (n\to\infty) \tag{100}$$

$$\sum_j ce_{t_j}^n(\varphi)\cdot\left(B_{t_{j+1}} - B_{t_j}\right) \overset{L^2}{\to} c\int_S^T f_s\, dB_s \ (n\to\infty) \tag{101}$$

$$\sum_j de_{t_j}^n(\psi)\cdot\left(B_{t_{j+1}} - B_{t_j}\right) \overset{L^2}{\to} d\int_S^T g_s\, dB_s \ (n\to\infty) \tag{102}$$

and linearity is proved by the uniqueness of L^2 limit. Here $e_{t_j}^n(\varphi)$ is the $\mathscr{F}_{t_j}$ measurable random variable used in the construction of φ_t^n such that $\varphi_t^n = \sum_j e_{t_j}^n(\varphi)\mathbb{I}_{[t_j,t_{j+1})}(t)$

The second one is the **partition of integration area**.

$$\forall S \le U \le T, \forall f_t \in V(S,T), \int_S^T f_s\, dB_s = \int_S^U f_s\, dB_s + \int_U^T f_s\, dB_s \tag{103}$$

which is also natural from the definition of stochastic integral and the approximation of elementary process.

The third property is that **the process of Ito integral $M_t = \int_0^t f_s\, dB_s$ is an L^2 martingale adapted to the filtration generated by BM**. These are all observations directly from the definition using elementary process that

$$\exists \varphi_t^n, \varphi_t^n \overset{L^2([S,T]\times\Omega)}{\to} f_t \ (n\to\infty) \tag{104}$$

$$\sum_j e_{t_j}^n(\varphi)(B_{t_{j+1}} - B_{t_j}) \overset{L^2}{\to} M_t \ (n\to\infty) \tag{105}$$

since $e_{t_j}^n(\varphi) \in \mathscr{F}_{t_j}$, so $\varphi_{t_j}^n(B_{t_{j+1}} - B_{t_j}) \in \mathscr{F}_{t_{j+1}} \subset \mathscr{F}_t$, so the L^2 limit $M_t \in \mathscr{F}_t$ is adapted. $M_t \in L^2$ is also obvious from the Ito's isometry. Now let's prove the martingale property for f_r:

$$\exists \varphi_r^n, \varphi_r^n \overset{L^2([s,t]\times\Omega)}{\rightarrow} f_r \ (n \to \infty) \tag{106}$$

$$\sum_j e^n_{r_j}(\varphi)(B_{r_{j+1}} - B_{r_j}) \overset{L^2}{\rightarrow} \int_s^t f_r \, dB_r \ (n \to \infty) \tag{107}$$

$$\begin{aligned}
\forall s \leq t, \mathbb{E}(M_t - M_s|\mathscr{F}_s) &= \mathbb{E}\left(\int_s^t f_r \, dB_r \middle| \mathscr{F}_s\right) && (108)\\
&= \lim_{n\to\infty} \sum_j \mathbb{E}\left(e^n_{r_j}(\varphi)(B_{r_{j+1}} - B_{r_j}) \middle| \mathscr{F}_s\right) && (109)\\
&= \lim_{n\to\infty} \sum_j \mathbb{E}\left[\mathbb{E}\left(e^n_{r_j}(\varphi)(B_{r_{j+1}} - B_{r_j}) \middle| \mathscr{F}_{r_j}\right) \middle| \mathscr{F}_s\right] && (110)\\
&= \lim_{n\to\infty} \sum_j \mathbb{E}\left[\mathbb{E}\left(e^n_{r_j}(\varphi)(B_{r_{j+1}} - B_{r_j}) \middle| \mathscr{F}_{r_j}\right) \middle| \mathscr{F}_s\right] && (111)\\
&= \lim_{n\to\infty} \sum_j \mathbb{E}\left[e^n_{r_j}(\varphi) \cdot \mathbb{E}\left(B_{r_{j+1}} - B_{r_j} \middle| \mathscr{F}_{r_j}\right) \middle| \mathscr{F}_s\right] && (112)\\
&= 0 && (113)
\end{aligned}$$

note that the appearance of the limit is due to the L^2 convergence to the stochastic integral and the martingale property follows directly from the tower property of C.E. and the fact that BM is itself a martingale. (r_j^n are the truncated dyadic numbers within $[s,t]$ with the grid of partition to be $\frac{1}{2^n}$, for fixed n, the sum w.r.t. j is actually a finite sum)

Path Regularity for Ito Integral

Now that M_t is a martingale, we know that under some special conditions (filtration to be right-continuous and complete, and $t \to \mathbb{E}M_t$ to be right-continuous), a martingale has a modification with **Cadlag sample paths** (refer to GTM 274 P57). To verify those conditions, the completeness of filtration is trivial and the right-continuity of filtration is also satisfied (Blumenthal's 0-1 law of BM), actually this filtration is called **the canonical filtration** of BM. Moreover, $\mathbb{E}M_t = \mathbb{E}M_0 = 0$ so it's continuous (MG property). As a result, the Ito integral process has a Cadlag modification and the one-sided continuity **enables the application of MG inequalities**. (Doob's maximal, Doob's L^p etc.). However, due to the special structure of Ito's integral (the continuity of BM), we can actually show that this process has a modification with **continuous sample paths**.

Theorem 2. *There exists a unique continuous modification of M_t.*

Proof. The uniqueness under indistinguishability directly follows from the continuity of sample path, so only need to prove existence.

To apply the definition of stochastic integral, there exists elementary process $\varphi_s^n \overset{L^2([0,t]\times\Omega)}{\rightarrow} f_s$ $(n\to\infty)$ with

$$\varphi_s^n = \sum_j e_j^n \mathbb{I}_{[t_j^n, t_{j+1}^n)}(s), e_j^n \in \mathscr{F}_{t_j^n} \tag{114}$$

where t_j^n is the truncated dyadic number in $[0,t]$ with grid size $\frac{1}{2^n}$ and consider

$$I_t^n = \int_0^t \varphi_s^n \, dB_s = \sum_j e_j^n \cdot (B_{t_{j+1}^n} - B_{t_j}^n) \tag{115}$$

which is obviously continuous in t for each fixed n. This is due to the uniform continuity of BM on closed intervals. Since I_t^n is itself an Ito integral, it's also an adapted martingale.

Notice that $I_t^n \overset{L^2}{\rightarrow} M_t$ $(n\to\infty)$, to prove the path continuity of M_t, it suffices to prove that I_t^n **converges uniformly on any compact set** $[0,T]$.

Let's take $\forall T\geq 0$ and consider the convergence on $t\in[0,T]$. Our goal is to prove that

$$\forall \varepsilon > 0, \exists N, \forall m,n > N, \sup_{t\leq T} |I_t^n - I_t^m| < \varepsilon \tag{116}$$

so we recall the Borel-Cantelli lemma and hope to prove that $\sup_{t\leq T}|I_t^n - I_t^m| < \varepsilon$ holds eventually. However, the difficulty is that here we have both m and n going to infinity and we also have to deal with ε, so we hope that there is a way for us to turn these three things into the dependency on a same variable going to infinity, a natural thought is to **take a good enough subsequence**.

To do this, note that $I_t^m - I_t^n$ is always a martingale with continuous sample path, so Doob's maximal inequality gives

$$\forall \varepsilon > 0, \mathbb{P}\left(\sup_{t\leq T}|I_t^m - I_t^n| \geq \varepsilon\right) \leq \frac{\mathbb{E}(I_T^m - I_T^n)^2}{\varepsilon^2} = \frac{\mathbb{E}\left(\int_0^T (\varphi_s^m - \varphi_s^n)\, dB_s\right)^2}{\varepsilon^2} \tag{117}$$

$$= \frac{\mathbb{E}\left(\int_0^T (\varphi_s^m - \varphi_s^n)^2\, ds\right)}{\varepsilon^2} \to 0 \ (n\to\infty) \tag{118}$$

by the Cauchy principle of the $L^2([0,T]\times\Omega)$ convergence of φ_s^n. As a result, there exists a subsequence $n_k\to\infty$ such that (a simple construction)

$$\forall k\geq 1, \mathbb{P}\left(\sup_{t\leq T}|I_t^{n_k} - I_t^{n_{k+1}}| \geq 2^{-k}\right) \leq \frac{1}{2^k} \tag{119}$$

Now it's easy to use Borel-Cantelli:

$$\sum_{k\geq 1}\mathbb{P}\left(\sup_{t\leq T}|I_t^{n_k}-I_t^{n_{k+1}}|\geq 2^{-k}\right)<\infty \tag{120}$$

$$\mathbb{P}\left(\sup_{t\leq T}|I_t^{n_k}-I_t^{n_{k+1}}|\geq 2^{-k}\ i.o.\right)=0 \tag{121}$$

so $a.s.$ $\sup_{t\leq T}|I_t^{n_k}-I_t^{n_{k+1}}|\leq 2^{-k}$ eventually for large enough k, which means that $a.s.$ $I_t^{n_k}$ converges uniformly on $[0,T]$ as $k\to\infty$. Since the L^2 limit is unique, it's easy to see that the limit of this subsequence has to be equal to M_t almost surely, so this limit is just a modification of M_t and uniform convergence ensures the continuity of sample path.

$\square$

Remark. *Doob's maximal inequality bounds the tail probability of the tail supreme of the approximation, leading to the existence of a good enough subsequence and the uniform convergence on any compact set. This is a frequently used criterion for* ***proving path continuity: bound the tail probability of the tail supreme, take a good subsequence and show uniform convergence with Borel-Cantelli****.*

Remark. *It's quite obvious that the upcrossing inequality is the key to the Cadlag modification of martingales. However, in order to get continuous modification, the continuity of BM is the key, i.e. if we are integrating w.r.t. a general semi-martingale, the continuity won't necessarily hold.*

From now on, we always assume that the **Ito's process w.r.t. process** f_t

$$M_t=\int_0^t f_s\,dB_s \tag{122}$$

is a continuous martingale.

Week 3

Extension of Ito Integral

There are two main extensions for Ito integral. One is that the filtration can be slightly enlarged. We can choose **the filtration $\mathscr{F}_t$ such that B_t is a $\mathscr{F}_t$-BM and** $f_t \in \mathscr{F}_t$, instead of the one directly generated by the BM, which is $\mathscr{F}_t = \sigma(X_s, 0 \le s \le t)$. The main motivation for this is that when considering the stochastic integral for multi-dimensional BM $B_t = (B_t^1, ..., B_t^n)$, an object of interest would be $\int_S^T B_t^2\, dB_t^1$. If we consider this integral under the old settings, then the filtration $\mathscr{F}_t$ should be the one generated by B_t^1. However, B_t^2 is not necessarily adapted to this filtration.

However, by setting $\mathscr{F}_t = \sigma((B_s^1, ..., B_s^n), 0 \le s \le t)$ (the jointly generated sigma field), $B_t^2 \in \mathscr{F}_t$ is adapted now and let's verify that this enlargement of sigma field still guarantees that B_t^1 is $\mathscr{F}_t$-adapted BM. This is quite obvious since multi-dimensional BM has independent coordinates, so B_t^1 conditioning on $\sigma((B_s^1, ..., B_s^n), 0 \le s \le t)$ is the same as B_t^1 conditioning on $\sigma(B_s^1, 0 \le s \le t)$.

The other extension is made for the Hilbert space $L^2([S,T] \times \Omega)$, in which the processes f_t should satisfy $\mathbb{E} \int_S^T f_t^2\, dt < \infty$. Now we would like to weaken the finiteness of expectation into almost surely finite

$$\mathbb{P}\left(\int_S^T f_t^2\, dt < \infty\right) = 1 \tag{123}$$

As a result, in the following context, the **Ito integral $\int_S^T f_t\, dB_t$ is actually defined for processes f_t and filtration $\mathscr{F}_t$ such that**

$$\begin{cases} B_t \ is\ \mathscr{F}_t - BM \\ (t,\omega) \to f_t(\omega) \in \mathscr{B}_{\mathbb{R}_+} \times \mathscr{F} \\ \forall t, f_t \in \mathscr{F}_t \\ \mathbb{P}\left(\int_S^T f_t^2\, dt < \infty\right) = 1 \end{cases} \tag{124}$$

So how does this generalization work? Actually still by approximation using elementary process but the convergence is expected in a weaker sense (converge in probability). Our task is to find a sequence of elementary process φ_t^n such that

$$\int_S^T |f_t - \varphi_t^n|^2\, dt \xrightarrow{p} 0 \ (n \to \infty) \tag{125}$$

and the Ito integral is formed as the limit in probability

$$\int_S^T \varphi_t^n\, dB_t \xrightarrow{p} \int_S^T f_t\, dB_t \ (n \to \infty) \tag{126}$$

The approximation is basically the same as what we've done above. For continuous bounded process $f_t, \forall t, |f_t| \le$

M, use the endpoint to construct a sequence of elementary processes at the truncated dyadic numbers t_j of $[S,T]$

$$\varphi_t^n = \sum_j f_{t_j} \mathbb{I}_{[t_j,t_{j+1})}(t) \tag{127}$$

and the approximation is well enough by the uniform continuity of f_t on $[S,T]$, for which $|f_t - f_{t_j}|$ is controlled by $\forall \varepsilon' > 0$

$$\forall \varepsilon > 0, \mathbb{P}\left(\int_S^T |f_t - \varphi_t^n|^2 \, dt > \varepsilon\right) = \mathbb{P}\left(\int_S^T \left|\sum_j (f_t - f_{t_j})\mathbb{I}_{[t_j,t_{j+1})}(t)\right|^2 \, dt > \varepsilon\right) \tag{128}$$

$$\leq \mathbb{P}\left(\sum_j \int_{t_j}^{t_{j+1}} (f_t - f_{t_j})^2 \, dt > \varepsilon\right) \tag{129}$$

$$\leq \mathbb{P}\left(\varepsilon'^2 (T-S) > \varepsilon\right) \to 0 \ (n \to \infty) \tag{130}$$

note that here ε is fixed first, and when $n \to \infty$, the partition is fine enough and a smaller ε' can always be found such that it controls $|f_t - f_{t_j}|$ and $\varepsilon'^2(T-S) \leq \varepsilon$. In other words, ε' can depend on n, making it possible to be much smaller than the fixed ε.

For bounded f_t, still use the convolution with the mollifier, and for general f_t, use a bounded truncation to approximate just as done above, but in the sense of convergence in probability. Note that since we are in the Hilbert space any longer, the convergence in probability has to be proved in an explicit way (much work to do). We can also show that the Ito's isometry still holds if $\mathbb{E}\int_S^T f_t^2 \, dt = \infty$.

For all the details, refer to GTM 274.

The price of such extension is to **lose the martingality** and now we can only ensure that **the Ito process is a continuous local martingale** (the continuity of path is maintained). Local martingale X_t is defined in a way that there exists a sequence of stopping time $\tau_n \nearrow \infty \ (n \to \infty)$ such that the stopped process $X_{t \wedge \tau_n}$ is martingale for each n. Note that there's no local MG in discrete time (countably many), so local MG is a special concept only for continuous time. If a process is a discrete-time local MG, then it must be a true MG.

The classical example of a local MG which is not a MG is the inverse Bessel process $X_t = \frac{1}{||B_t||}$, where B_t is a 3-dim BM does not start from origin. A less complicated example can be $M_t = ZB_t, Z \in \mathscr{F}_0, \mathbb{E}|Z| = \infty$.

Stratonovich Integral

In the definition of Ito integral, the left endpoint is always taken in the approximation to ensure the measurability property. The advantage is that Ito integrals are continuous local MG, but the disadvantage is that chain rule fails for Ito integral (quadratic variation as Ito correction term appears).

It's natural to ask if it's possible to take the right endpoint $f_{t_{j+1}}$ or the midpoint $\frac{f_{t_j}+f_{t_{j+1}}}{2}$ on the interval $[t_j, t_{j+1})$ to do the approximation. The **Fisk-Stratonovich integral** is defined as taking $(1-\varepsilon)f_{t_j} + \varepsilon f_{t_{j+1}}$ on each interval $[t_j, t_{j+1})$. When $\varepsilon = 0$, it's just Ito integral. When $\varepsilon = \frac{1}{2}$, it's the **Stratonovich integral**, denoted

$$\int f_s \circ dB_s \tag{131}$$

It satisfies the usual **chain rule** but we lose the martingale property (in the proof of martingality, $f_{t_j} \in \mathscr{F}_{t_j}$ is critical). To see the chain rule, compute the integral for BM:

$$\int_0^t B_s \circ dB_s = \lim_{n\to\infty} \sum_j \frac{B_{t_j} + B_{t_{j+1}}}{2}(B_{t_{j+1}} - B_{t_j}) \tag{132}$$

$$= \frac{1}{2} \lim_{n\to\infty} \sum_j (B_{t_{j+1}}^2 - B_{t_j}^2) = \frac{B_t^2}{2} \tag{133}$$

This works as if we replace B_s with s and conclude that $\int_0^t s\, ds = \frac{t^2}{2}$.

Ito Formula

Ito formula provides a method to know about the behavior of $g(B_t)$, a function of BM or other processes by expanding it into stochastic integrals. The 1-dim Ito formula is formulated as

Theorem 3. *If $g \in C^2 : \mathbb{R} \to \mathbb{R}$, then $dg(B_t) = g'(B_t)\, dB_t + \frac{1}{2} g''(B_t)\, d\langle B, B\rangle_t$, where $\langle B, B\rangle_t$ is the quadratic variation of BM in time interval $[0, t]$ (so $\langle B, B\rangle_t = t$). $\langle M, M\rangle_t$ is generally defined for continuous local MG M_t as the unique increasing process such that $M_t^2 - \langle M, M\rangle_t$ is also a continuous local MG (analogue to Doob's MG decomposition).*

Proof. Let's prove the integral form since the terms are actually defined in the integral form:

$$g(B_t) - g(B_0) = \int_0^t g'(B_s)\, dB_s + \frac{1}{2} \int_0^t g''(B_s)\, ds \tag{134}$$

Let's naturally apply the Taylor expansion with Lagrange remainder with a telescoping form, t_j are truncated dyadic numbers in $[0, t]$

$$g(B_t) - g(0) = \sum_j g(B_{t_{j+1}}) - g(B_{t_j}) \tag{135}$$

$$= \sum_j \left[g'(B_{t_j})(B_{t_{j+1}} - B_{t_j}) + \frac{1}{2} g''(\xi_{t_j})(B_{t_{j+1}} - B_{t_j})^2 \right] \tag{136}$$

for some ξ_{t_j} between $B_{t_j}, B_{t_{j+1}}$.

The first term converges to $\int_0^t g'(B_s)\, dB_s$ in L^2 sense by the definition of Ito integral above directly. Now we prove that the second term converges in L^2 to $\frac{1}{2}\int_0^t g''(B_s)\,ds$:

$$\mathbb{E}\left(\frac{1}{2}\sum_j g''(\xi_{t_j})(B_{t_{j+1}}-B_{t_j})^2-\frac{1}{2}\sum_j g''(B_{t_j})(B_{t_{j+1}}-B_{t_j})^2\right)^2 \tag{137}$$

$$=\frac{1}{4}\mathbb{E}\left(\sum_j [g''(\xi_{t_j})-g''(B_{t_j})](B_{t_{j+1}}-B_{t_j})^2\right)^2 \tag{138}$$

$$=\frac{1}{4}\sum_{i,j}\mathbb{E}[g''(\xi_{t_j})-g''(B_{t_j})](B_{t_{j+1}}-B_{t_j})^2[g''(\xi_{t_i})-g''(B_{t_i})](B_{t_{i+1}}-B_{t_i})^2 \tag{139}$$

$$=\frac{1}{4}\sum_j \mathbb{E}[g''(\xi_{t_j})-g''(B_{t_j})]^2(B_{t_{j+1}}-B_{t_j})^4+\frac{1}{2}\sum_{i<j}\mathbb{E}[g''(\xi_{t_j})-g''(B_{t_j})](B_{t_{j+1}}-B_{t_j})^2\cdot\mathbb{E}[g''(\xi_{t_i})-g''(B_{t_i})](B_{t_{i+1}}-B_{t_i})^2 \tag{140}$$

the uniform continuity of g'' and the uniform continuity of BM path on interval $[0,t]$ ensures that $\forall j, |g''(\xi_{t_j})-g''(B_{t_j})|<\varepsilon$. Notice that $\sum_j \mathbb{E}(B_{t_{j+1}}-B_{t_j})^4=\sum_j 3(t_{j+1}-t_j)^2\to 0\ (n\to\infty)$ and that $\sum_{i<j}(t_{j+1}-t_j)(t_{i+1}-t_i)\leq \frac{(t2^n)^2}{2^{2n}}=t^2<\infty$, so

$$\mathbb{E}\left(\frac{1}{2}\sum_j g''(\xi_{t_j})(B_{t_{j+1}}-B_{t_j})^2-\frac{1}{2}\sum_j g''(B_{t_j})(B_{t_{j+1}}-B_{t_j})^2\right)^2 \tag{141}$$

$$\leq \frac{\varepsilon^2}{4}\sum_j 3(t_{j+1}-t_j)^2+\frac{\varepsilon^2}{2}\sum_{i<j}(t_{j+1}-t_j)(t_{i+1}-t_i)\to 0\ (n\to\infty) \tag{142}$$

Next we prove that $\frac{1}{2}\sum_j g''(B_{t_j})(B_{t_{j+1}}-B_{t_j})^2$ converges to $\frac{1}{2}\int_0^t g''(B_s)\,ds$ in the L^2 sense. Since g'' is continuous and BM has continuous path, $B_s, 0\leq s\leq t$ is bounded, so $|g''(B_{t_j})|\leq M$ for uniform bound M.

$$\mathbb{E}\left(\frac{1}{2}\sum_j g''(B_{t_j})(t_{j+1}-t_j)-\frac{1}{2}\sum_j g''(B_{t_j})(B_{t_{j+1}}-B_{t_j})^2\right)^2 \tag{143}$$

$$=\frac{1}{4}\mathbb{E}\left(\sum_j g''(B_{t_j})[(t_{j+1}-t_j)-(B_{t_{j+1}}-B_{t_j})^2]\right)^2 \tag{144}$$

$$=\frac{1}{4}\mathbb{E}\sum_{i,j} g''(B_{t_j})[(t_{j+1}-t_j)-(B_{t_{j+1}}-B_{t_j})^2]g''(B_{t_i})[(t_{i+1}-t_i)-(B_{t_{i+1}}-B_{t_i})^2] \tag{145}$$

$$=\frac{1}{4}\sum_j \mathbb{E}[g''(B_{t_j})]^2[(t_{j+1}-t_j)-(B_{t_{j+1}}-B_{t_j})^2]^2 \tag{146}$$

$$+\frac{1}{2}\sum_{i<j}\mathbb{E}g''(B_{t_j})[(t_{j+1}-t_j)-(B_{t_{j+1}}-B_{t_j})^2]\cdot\mathbb{E}g''(B_{t_i})[(t_{i+1}-t_i)-(B_{t_{i+1}}-B_{t_i})^2] \tag{147}$$

$$\leq \frac{M^2}{4}\sum_j \mathbb{E}[(t_{j+1}-t_j)-(B_{t_{j+1}}-B_{t_j})^2]^2 \tag{148}$$

$$+\frac{M^2}{2}\sum_{i<j} \mathbb{E}[(t_{j+1}-t_j)-(B_{t_{j+1}}-B_{t_j})^2]\cdot\mathbb{E}[(t_{i+1}-t_i)-(B_{t_{i+1}}-B_{t_i})^2] \tag{149}$$

with calculations

$$\sum_j \mathbb{E}[(t_{j+1}-t_j)-(B_{t_{j+1}}-B_{t_j})^2]^2 \tag{150}$$

$$=\sum_j (t_{j+1}-t_j)^2+\mathbb{E}(B_{t_{j+1}}-B_{t_j})^4-2(t_{j+1}-t_j)\mathbb{E}(B_{t_{j+1}}-B_{t_j})^2 \tag{151}$$

$$=\sum_j (t_{j+1}-t_j)^2+3(t_{j+1}-t_j)^2-2(t_{j+1}-t_j)^2\to 0\ (n\to\infty) \tag{152}$$

$$\tag{153}$$

$$\sum_{i<j} \mathbb{E}[(t_{j+1}-t_j)-(B_{t_{j+1}}-B_{t_j})^2]\cdot\mathbb{E}[(t_{i+1}-t_i)-(B_{t_{i+1}}-B_{t_i})^2] \tag{154}$$

$$=\sum_{i<j}[(t_{j+1}-t_j)-(t_{j+1}-t_j)]\cdot[(t_{i+1}-t_i)-(t_{i+1}-t_i)]\to 0\ (n\to\infty) \tag{155}$$

Combining all these estimations, we have proved that

$$\frac{1}{2}\sum_j g''(\xi_{t_j})(B_{t_{j+1}}-B_{t_j})^2 \xrightarrow{L^2} \frac{1}{2}\int_0^t g''(B_s)\,ds\ (n\to\infty) \tag{156}$$

thus the Ito formula holds.

□

Actually, the Ito formula can be extended in a parametric case.

Theorem 4. *If $g(t,x)\in C^{1,2}$, $dg(t,B_t)=g_t(t,B_t)\,dt+g_x(t,B_t)\,dB_t+\frac{1}{2}g_{xx}(t,B_t)\,d\langle B,B\rangle_t$.*

Proof. The structure of the proof is basically the same. Telescope, use 2-dim Taylor expansion and estimate the sums using integrals.

$$g(t,B_t)-g(0,B_0)=\sum_j g(t_{j+1},B_{t_{j+1}})-g(t_j,B_{t_j}) \tag{157}$$

$$=\sum_j\left[g_t(t_j,B_{t_j})(t_{j+1}-t_j)+g_x(t_j,B_{t_j})(B_{t_{j+1}}-B_{t_j})+\frac{1}{2}g_{xx}(t_j,B_{t_j})(B_{t_{j+1}}-B_{t_j})^2+...\right] \tag{158}$$

The reason we have not written all second order terms is that the other terms vanish

$$
\begin{aligned}
\mathbb{E}\left(\sum_j (t_{j+1}-t_j)(B_{t_{j+1}}-B_{t_j})\right)^2 &= \sum_{i,j} \mathbb{E}(t_{j+1}-t_j)(B_{t_{j+1}}-B_{t_j})(t_{i+1}-t_i)(B_{t_{i+1}}-B_{t_i}) && (159)\\
&= \sum_j (t_{j+1}-t_j)^2 \cdot \mathbb{E}(B_{t_{j+1}}-B_{t_j})^2 && (160)\\
&+ 2\sum_{i<j}(t_{i+1}-t_i)(t_{j+1}-t_j)\cdot \mathbb{E}(B_{t_{j+1}}-B_{t_j})\cdot \mathbb{E}(B_{t_{i+1}}-B_{t_i}) && (161)\\
&= \sum_j (t_{j+1}-t_j)^3 && (162)\\
&\to 0 \ (n\to\infty) && (163)\\
\mathbb{E}\left(\sum_j (B_{t_{j+1}}-B_{t_j})^2\right)^2 &= \sum_{i,j}\mathbb{E}(B_{t_{j+1}}-B_{t_j})^2(B_{t_{i+1}}-B_{t_i})^2 && (164)\\
&= \sum_j \mathbb{E}(B_{t_{j+1}}-B_{t_j})^4 + 2\sum_{i<j}\mathbb{E}(B_{t_{j+1}}-B_{t_j})^2\cdot\mathbb{E}(B_{t_{i+1}}-B_{t_i})^2 && (165)\\
&= \sum_j 3(t_{j+1}-t_j)^2 + 2\sum_{i<j}(t_{j+1}-t_j)^2(t_{i+1}-t_i)^2 \to 0\ (n\to\infty) && (166)
\end{aligned}
$$

similarly, the L^2 convergence still hold:

$$
\begin{aligned}
&\sum_j g_t(t_j,B_{t_j})(t_{j+1}-t_j) \xrightarrow{L^2} \int_0^t g_t(s,B_s)\,ds\ (n\to\infty) && (167)\\
&\sum_j g_x(t_j,B_{t_j})(B_{t_{j+1}}-B_{t_j}) \xrightarrow{L^2} \int_0^t g_x(s,B_s)\,dB_s\ (n\to\infty) && (168)\\
&\sum_j \frac{1}{2}g_{xx}(t_j,B_{t_j})(B_{t_{j+1}}-B_{t_j})^2 \xrightarrow{L^2} \frac{1}{2}\int_0^t g_{xx}(s,B_s)\,d\langle B,B\rangle_s\ (n\to\infty) && (169)
\end{aligned}
$$

□

Remark. *It's easy to see that in notation, we can write $dtdt=0, dtdB_t=0, dB_tdB_t=dt$. This is due to the fact that $f_t=t$ is of finite variation, so the second variation must be 0 and that the cross variation of $f_t=t$ and B_t is 0. This also explains why there are no higher order terms in the Ito formula.*

*An immediate generalization of the Ito formula with a time parameter is that we can replace the time t with **any finite variation process** f_t to let the Ito formula work for things like $g(f_t,B_t)$.*

Example

The first example is to compute

$$d(e^{bt+\sigma B_t}) \tag{170}$$

where b, σ are constants for drift and diffusion, set $g(t,x) = e^{bt+\sigma x}$ and apply the Ito formula

$$d(e^{bt+\sigma B_t}) = be^{bt+\sigma B_t}\,dt + \sigma e^{bt+\sigma B_t}\,dB_t + \frac{\sigma^2}{2}e^{bt+\sigma B_t}\,dt \tag{171}$$

As a result, if define $X_t = e^{bt+\sigma B_t}X_0$, then this X_t is just the solution to the SDE

$$dX_t = \left(b + \frac{\sigma^2}{2}\right)X_t\,dt + \sigma X_t\,dB_t \tag{172}$$

which defines a geometric BM and is closely related to the Black-Scholes model (μ as mean return and σ as volatility)

$$dX_t = \mu X_t\,dt + \sigma X_t\,dB_t \tag{173}$$

Ito formula for Ito Process

The **Ito process** is defined as

$$X_t = x_0 + \int_0^t \psi_s\,ds + \int_0^t \varphi_s\,dB_s \tag{174}$$

$$\psi_s, \varphi_s \in \mathscr{F}_s, \mathbb{E}\int_0^t \varphi_s^2\,ds < \infty, \mathbb{E}\int_0^t |\psi_s|\,ds < \infty \tag{175}$$

a constant x_0 plus the Stieljes integral of a process and plus the stochastic integral of another process. These two processes are both adapted, with φ_s in $L^2([0,t]\times\Omega)$ where the stochastic integral is defined and ψ_s such that the Stieljes integral part has finite expectation.

To go into more details, note that

$$\sum_j \left|\int_0^{t_{j+1}} \psi_s\,ds - \int_0^{t_j} \psi_s\,ds\right| \le \sum_j \int_{t_j}^{t_{j+1}} |\psi_s|\,ds = \int_0^t |\psi_s|\,ds < \infty \ a.s. \tag{176}$$

so the $\int_0^t \psi_s\,ds$ part is a **finite variation process**, i.e. it contributes nothing to the quadratic variation of the whole process. As proved above, the $\int_0^t \varphi_s\,dB_s$ part is a **continuous MG** which is typically not finite variation (if a finite variation process is continuous local MG, it must be constant almost surely). In simple words, **the Ito process is made up of constant part, finite variation part, continuous MG part**.

Theorem 5. *If $g(t,x) \in C^{1,2}$, then $dg(t,X_t) = g_t(t,X_t)\,dt + g_x(t,X_t)\,dX_t + \frac{1}{2}g_{xx}(t,X_t)\,d\langle X,X\rangle_t$ with $d\langle X,X\rangle_t = \varphi_t^2\,dt$.*

Proof. The proof is the same as that for BM above. The only thing to verify now is the quadratic variation of X_t. By previous calculations, only the term $\int_0^t \varphi_s \, dB_s$ contributes to the quadratic variation.

$$\mathbb{E}\sum_j (X_{t_{j+1}} - X_{t_j})^2 = \sum_j \mathbb{E}\left(\int_{t_j}^{t_{j+1}} \varphi_s \, dB_s\right)^2 \tag{177}$$

$$= \sum_j \mathbb{E}\int_{t_j}^{t_{j+1}} \varphi_s^2 \, ds \tag{178}$$

$$= \mathbb{E}\int_0^t \varphi_s^2 \, ds \tag{179}$$

$$\forall t, \mathbb{E}\langle X, X\rangle_t = \mathbb{E}\int_0^t \varphi_s^2 \, ds \tag{180}$$

by the Ito's isometry and the monotone convergence theorem. So it's reasonable to guess that

$$\langle X, X\rangle_t = \int_0^t \varphi_s^2 \, ds \tag{181}$$

The proof can be given in the following sense that when φ_s is a bounded process,

$$\mathbb{E}\left(\sum_j \varphi_{t_j}^2 (B_{t_{j+1}} - B_{t_j})^2 - \sum_j \varphi_{t_j}^2 (t_{j+1} - t_j)\right)^2 \tag{182}$$

$$= \sum_{i,j} \mathbb{E}\left(\varphi_{t_j}^2 [(B_{t_{j+1}} - B_{t_j})^2 - (t_{j+1} - t_j)] \varphi_{t_i}^2 [(B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i)]\right) \tag{183}$$

$$= \sum_j \mathbb{E}\left(\varphi_{t_j}^4 [(B_{t_{j+1}} - B_{t_j})^2 - (t_{j+1} - t_j)]^2\right) \tag{184}$$

$$+ 2\sum_{i<j} \mathbb{E}\left[\varphi_{t_i}^2 [(B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i)] \cdot \mathbb{E}\left(\varphi_{t_j}^2 [(B_{t_{j+1}} - B_{t_j})^2 - (t_{j+1} - t_j)] \middle| \mathscr{F}_{t_i}\right)\right] \tag{185}$$

$$\to 0 \ (n \to \infty) \tag{186}$$

since there are only terms having order no lower than $dtdt, dtdB_t$. Use truncation and dominated convergence theorem to prove the result for general φ_s.

□

Example

Consider the exponential MG of BM:

$$M_t = e^{\int_0^t h_s \, dB_s - \frac{1}{2}\int_0^t h_s^2 \, ds} \tag{187}$$

with h_s to be a bounded process. Specify $M_t = e^{X_t}, X_t = \int_0^t h_s\, dB_s - \frac{1}{2}\int_0^t h_s^2\, ds$ and set $g(x) = e^x$ to get

$$dM_t = e^{X_t}\, dX_t + \frac{1}{2}e^{X_t}\, d\langle X, X\rangle_t \tag{188}$$

recall that $\langle X, X\rangle_t = \int_0^t h_s^2\, ds$, and

$$dX_t = h_t\, dB_t - \frac{1}{2}h_t^2\, dt \tag{189}$$

so

$$dM_t = e^{X_t} h_t\, dB_t \tag{190}$$

This is telling us that this M_t is generally a **continuous local MG**

$$M_t = M_0 + \int_0^t e^{X_s} h_s\, dB_s,\ M_0 = 1 \tag{191}$$

with

$$\langle M, M\rangle_t = \int_0^t e^{2X_s} h_s^2\, ds \tag{192}$$

As a result, if $\forall t, \mathbb{E}\int_0^t e^{2X_s} h_s^2\, ds < \infty$ (h_s is bounded suffices), the continuous local MG satisfies $\forall t, \mathbb{E}\langle M, M\rangle_t < \infty$ so such M_t must be a L^2 MG. In such case, the M_t is a **natural extension of the exponential MG** (in the original setting, h_s is constant but now it can be a bounded process).

Remark. *For continuous local MG M_t, $\forall t, \mathbb{E}\langle M, M\rangle_t < \infty$ is equivalent to M_t being L^2 MG and $\mathbb{E}\langle M, M\rangle_\infty < \infty$ is equivalent to M_t being L^2 bounded MG. For more detailed conditions on M_t being a MG, refer to Kazamaki and Novikov conditions.*

Multi-dimensional Ito Formula

Since Ito process is a more general setting than a function of BM or a function of both time and BM (Ito process is semi-MG), we only describe the Ito formula for Ito process. First set up **the d-dimensional Ito process as the integral w.r.t. m-dimensional BM** as following:

$$X_t^i = x_0^i + \int_0^t \psi_s^i\, ds + \sum_{k=1}^m \int_0^t \varphi_s^{i,k}\, dB_s^k\ (k = 1, ..., m,\ i = 1, 2, ..., d) \tag{193}$$

with the explanation that such Ito process lives in the space $\mathbb{R}^d$ and is constructed by the stochastic integral w.r.t. a m-dimensional BM $B_s = (B_s^1, ..., B_s^m)$. To write it in a more compact form, introduce the notation that

$$X_t = x_0 + \int_0^t \psi_s\, ds + \int_0^t \varphi_s \cdot dB_s \tag{194}$$

$$X_t, x_0, \psi_s \in \mathbb{R}^d, \varphi_s \in \mathbb{R}^{d\times m} \tag{195}$$

here $\varphi_s^{i,k}$ stands for the process φ used to construct the i-th coordinate of Ito process as a stochastic integral w.r.t. the k-th coordinate of m-dimensional BM.

Theorem 6. *If vector-valued function $g : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^p \in C^{1,2}$, then multi-dimensional Ito formula holds*

$$dg^i(t, X_t) = \partial_t g^i(t, X_t)\, dt + \nabla_x g^i(t, X_t) \cdot dX_t + \frac{1}{2} Tr(\varphi_t^T H \varphi_t)\, dt \tag{196}$$

$$= \partial_t g^i(t, X_t)\, dt + \sum_{k=1}^d \partial_{x_k} g^i(t, X_t)\, dX_t^k + \frac{1}{2} \sum_{j,k=1}^d \partial_{x_j, x_k} g^i(t, X_t)\, d\langle X^j, X^k\rangle_t \tag{197}$$

where $\varphi_t \in \mathbb{R}^{d\times m}$ is a matrix and $H_{d\times d}$ is the Hessian of g restricted on its action on $x \in \mathbb{R}^d$ at (t, X_t). The bracket $d\langle X^j, X^k\rangle_t = \sum_{l=1}^m \varphi_t^{j,l} \varphi_t^{k,l}\, dt$.

Proof. The structure of the proof is still exactly the same as it is in the 1-dimensional case. The only two things to be verified is that $dB_t^p\, dB_t^q = 0$ $(p \neq q)$ and $d\langle X^j, X^k\rangle_t = \sum_{l=1}^m \varphi_t^{j,l} \varphi_t^{k,l}\, dt$.

$$\mathbb{E}\left(\sum_j (B_{t_{j+1}}^p - B_{t_j}^p)(B_{t_{j+1}}^q - B_{t_j}^q)\right)^2 \tag{198}$$

$$= \mathbb{E} \sum_{i,j} (B_{t_{j+1}}^p - B_{t_j}^p)(B_{t_{j+1}}^q - B_{t_j}^q)(B_{t_{i+1}}^p - B_{t_i}^p)(B_{t_{i+1}}^q - B_{t_i}^q) \tag{199}$$

$$= \sum_j \mathbb{E}(B_{t_{j+1}}^p - B_{t_j}^p)^2 (B_{t_{j+1}}^q - B_{t_j}^q)^2 + 2\sum_{i<j} \mathbb{E}(B_{t_{j+1}}^p - B_{t_j}^p)(B_{t_{j+1}}^q - B_{t_j}^q)(B_{t_{i+1}}^p - B_{t_i}^p)(B_{t_{i+1}}^q - B_{t_i}^q) \tag{200}$$

$$= \sum_j \mathbb{E}(B_{t_{j+1}}^p - B_{t_j}^p)^2 \cdot \mathbb{E}(B_{t_{j+1}}^q - B_{t_j}^q)^2 + 2\sum_{i<j} \mathbb{E}(B_{t_{j+1}}^p - B_{t_j}^p) \cdot \mathbb{E}(B_{t_{j+1}}^q - B_{t_j}^q) \cdot \mathbb{E}(B_{t_{i+1}}^p - B_{t_i}^p) \cdot \mathbb{E}(B_{t_{i+1}}^q - B_{t_i}^q) \tag{201}$$

$$= \sum_j \mathbb{E}^2(B_{t_{j+1}}^p - B_{t_j}^p)^2 + 2\sum_{i<j} \mathbb{E}^2(B_{t_{j+1}}^p - B_{t_j}^p) \cdot \mathbb{E}^2(B_{t_{i+1}}^p - B_{t_i}^p) \tag{202}$$

$$= \sum_j (t_{j+1} - t_j)^2 \to 0\ (n \to \infty) \tag{203}$$

For the bracket of X_t, it's still true that only $\varphi_s^{i,k}$ contributes to the quadratic variation.

$$\left\langle \sum_{l=1}^{m} \int_0^t \varphi_s^{j,l}\, dB_s^l, \sum_{p=1}^{m} \int_0^t \varphi_s^{k,p}\, dB_s^p \right\rangle_t \tag{204}$$

$$= \sum_{l,p=1}^{m} \left\langle \int_0^t \varphi_s^{j,l}\, dB_s^l, \int_0^t \varphi_s^{k,p}\, dB_s^p \right\rangle_t \tag{205}$$

$$= \sum_{l=1}^{m} \left\langle \int_0^t \varphi_s^{j,l}\, dB_s^l, \int_0^t \varphi_s^{k,l}\, dB_s^l \right\rangle_t = \sum_{l=1}^{m} \int_0^t \varphi_s^{j,l} \varphi_s^{k,l}\, ds \tag{206}$$

by using the property just derived that independent BM has bracket 0. □

Remark. *The isometry property of stochastic integral and the bracket may make it much easier to calculate. For general semi-MG M, N and $f_s, g_s \in L^2([0,t] \times \Omega)$(process for which stochastic integral is well-defined),*

$$\left\langle \int_0^{\cdot} f_s\, dM_s, \int_0^{\cdot} g_s\, dN_s \right\rangle_t = \int_0^t f_s g_s\, d\langle M, N\rangle_s \tag{207}$$

(See GTM 274, P101, Theorem 5.4) .Using this property, all calculations on the brackets of stochastic integrals are trivial.

Applications

The integration by parts is an application of multi-dimensional Ito formula. Consider the 2-dim Ito process constructed using the stochastic integral w.r.t. 2-dim BM.

$$X_t^1 = x_0^1 + \int_0^t \psi_s^1\, ds + \int_0^t \varphi_s^{1,1}\, dB_s^1 + \int_0^t \varphi_s^{1,2}\, dB_s^2 \tag{208}$$

$$X_t^2 = x_0^2 + \int_0^t \psi_s^2\, ds + \int_0^t \varphi_s^{2,1}\, dB_s^1 + \int_0^t \varphi_s^{2,2}\, dB_s^2 \tag{209}$$

calculate $d(X_t^1 X_t^2)$ setting $g(x_1, x_2) = x_1 x_2$ to find

$$d(X_t^1 X_t^2) = X_t^2\, dX_t^1 + X_t^1\, dX_t^2 + \frac{1}{2}\, dX_t^1\, dX_t^2 + \frac{1}{2}\, dX_t^2\, dX_t^1 \tag{210}$$

$$= X_t^2\, dX_t^1 + X_t^1\, dX_t^2 + \varphi_t^{1,1}\varphi_t^{2,1}\, dt + \varphi_t^{1,2}\varphi_t^{2,2}\, dt \tag{211}$$

the **integration by parts** formula.

Another trivial example is to calculate the moments of 1-dim BM B_t (the superscripts here are powers).

To get $\mathbb{E}B_t^2$, we can think about expanding dB_t^2 using Ito formula with $g(x) = x^2$:

$$dB_t^2 = 2B_t\,dB_t + dt \tag{212}$$

$$B_t^2 = 2\int_0^t B_s\,dB_s + t \tag{213}$$

taking expectation on both sides to get:

$$\mathbb{E}B_t^2 = 2\mathbb{E}\int_0^t B_s\,dB_s + t = t \tag{214}$$

note that $\mathbb{E}\int_0^t B_s\,dB_s = 0$ follows from the fact that Ito integral process (denoted M_t above) is MG. As a result, **for any $\int_0^t f_s\,dB_s, f_s \in V(0,t)$, its expectation is always 0.**

Ito formula can also be applied to B_t^4:

$$dB_t^4 = 4B_t^3\,dB_t + 6B_t^2\,dt \tag{215}$$

$$B_t^4 = 4\int_0^t B_s^3\,dB_s + 6\int_0^t B_s^2\,ds \tag{216}$$

taking expectation on both sides to get

$$\mathbb{E}B_t^4 = 6\mathbb{E}\int_0^t B_s^2\,ds = 6\int_0^t s\,ds = 3t^2 \tag{217}$$

with the interchange of expectation and integral ensured by Fubini.

The last example is the **n-dimensional Bessel process for** $n \geq 2$. Consider $B_t = (B_t^1, ..., B_t^n)$ to be n-dimensional BM and $R_t = ||B_t||_2$ is the Euclidean distance to the origin. Ito formula is applied for $g(x) = ||x||_2$.

$$dR_t = \sum_{j=1}^n \frac{B_t^j}{R_t}\,dB_t^j + \frac{1}{2}\sum_{j=1}^n \frac{(R_t)^2 - (B_t^j)^2}{(R_t)^3}\,dt \tag{218}$$

$$= \sum_{j=1}^n \frac{B_t^j}{R_t}\,dB_t^j + \frac{n-1}{2R_t}\,dt \tag{219}$$

Remark. *The C^2 assumption in Ito formula can be weakened (as long as the process almost surely does not touch the singularity it's fine).*

In the example above, since for $n \geq 2$, almost surely BM never hits the origin, the Ito formula still holds. For the same reason, as long as X_t is positive almost surely, we can also apply Ito formula for $\log X_t$ (the situation in solving Black-Scholes model).

Martingale Representation Theorem

The motivation of MG Rep Thm is natural: since we have already proved that the process of Ito integral for good enough $f_s \in L^2([0,t]\times\Omega)$

$$M_t = \int_0^t f_s\,dB_s \tag{220}$$

is a L^2 continuous MG, can we represent any L^2 continuous MG as the Ito integral of some process f_s? A slight detail is that by martingality, $\forall t, \mathbb{E}\int_0^t f_s\,dB_s = 0$. So if the MG M_t we want to represent is not starting from 0 at time 0, at least we shall subtract the starting point, i.e. to find $f_s \in L^2([0,t]\times\Omega)$ such that

$$\forall t, M_t - M_0 = \int_0^t f_s\,dB_s \tag{221}$$

The MG Rep Thm starts with a weakened version, which is the Ito Representation theorem, stating that any L^2 random variable can be represented as the stochastic integral of a process.

Theorem 7. *If $F \in L^2(\mathscr{F}_T)$ for a fixed time T, then unique $\exists f_t \in L^2([0,T]\times\Omega)$ such that $F = \mathbb{E}F + \int_0^T f_s\,dB_s$.* ***(Ito's Representation Theorem)***

Proof. The proof starts by considering a special family of random variables as extended exponential MG of BM

$$F = e^{\int_0^T h(s)\,dB_s - \frac{1}{2}\int_0^T h^2(s)\,ds} \tag{222}$$

for a deterministic function $h(s)$ (such h should be such that F is well-defined).

Let's consider the integral process $Y_t = e^{\int_0^t h(s)\,dB_s - \frac{1}{2}\int_0^t h^2(s)\,ds}$ changing with time t, Ito formula gives

$$X_t = \int_0^t h(s)\,dB_s - \frac{1}{2}\int_0^t h^2(s)\,ds \tag{223}$$

$$dY_t = e^{X_t}\,dX_t + \frac{1}{2}e^{X_t}\,d\langle X,X\rangle_t \tag{224}$$

$$dX_t = h(t)\,dB_t - \frac{1}{2}h^2(t)\,dt \tag{225}$$

$$d\langle X,X\rangle_t = h^2(t)\,dt \tag{226}$$

so change it into the integral form

$$dY_t = Y_t h(t)\,dB_t \tag{227}$$

$$Y_t - Y_0 = \int_0^t Y_s h(s)\,dB_s \tag{228}$$

$$Y_0 = 1 \tag{229}$$

Setting $f_s = Y_s h(s)$ ends the proof since by martingality $\mathbb{E}F = \mathbb{E}Y_0 = 1$, so $F = \mathbb{E}F + \int_0^T Y_s h(s)\,dB_s$.

Now let's notice that

$$span\left\{e^{\int_0^T h(s)\,dB_s - \frac{1}{2}\int_0^T h^2(s)\,ds}, h \in L^2([0,T])\right\} \tag{230}$$

is a dense subset in $L^2(\Omega, \mathscr{F}_T)$ (with h deterministic). By admitting this, a general F can be approximated by linear combinations of the random variables having the form of exponential MG.

$$\exists c_i \in \mathbb{R}, F_n = \sum_{i=1}^{n} c_i M_i \xrightarrow{L^2} F \ (n \to \infty) \tag{231}$$

$$M_i \in \left\{e^{\int_0^T h(s)\,dB_s - \frac{1}{2}\int_0^T h^2(s)\,ds}, h \in L^2([0,T])\right\} \tag{232}$$

linearity ensures that the Ito representation theorem still holds for F_i (note that F_i does not necessarily have the form $e^{\int_0^T h(s)\,dB_s - \frac{1}{2}\int_0^T h^2(s)\,ds}$ for deterministic h)

$$\exists f_s^n, F_n = \mathbb{E}F_n + \int_0^T f_s^n\,dB_s \tag{233}$$

it's natural to think about taking the limit of f_s^n as the process that represents F. So we have to figure out whether this sequence of process converge in the $L^2([0,T] \times \Omega)$ sense.

$$\mathbb{E}(F_m - F_n)^2 = \mathbb{E}\left(\mathbb{E}F_m - \mathbb{E}F_n + \int_0^T f_s^m - f_s^n\,dB_s\right)^2 \tag{234}$$

$$= \mathbb{E}^2(F_m - F_n) + 2\mathbb{E}(F_m - F_n) \cdot \mathbb{E}\left(\int_0^T (f_s^m - f_s^n)\,dB_s\right) + \mathbb{E}\left(\int_0^T (f_s^m - f_s^n)\,dB_s\right)^2 \tag{235}$$

$$= \mathbb{E}^2(F_m - F_n) + \mathbb{E}\left(\int_0^T (f_s^m - f_s^n)\,dB_s\right)^2 \tag{236}$$

notice that $F_n \xrightarrow{L^2} F \ (n \to \infty)$, this is telling us that

$$\mathbb{E}(F_m - F_n)^2 \to 0 \ (m, n \to \infty) \tag{237}$$

to know immediately with Ito's isometry that

$$\mathbb{E}\left(\int_0^T (f_s^m - f_s^n)\,dB_s\right)^2 \to 0 \ (m, n \to \infty) \tag{238}$$

$$\mathbb{E}\int_0^T (f_s^m - f_s^n)^2\,ds \to 0 \ (m, n \to \infty) \tag{239}$$

$$||f_s^m - f_s^n||_{L^2([0,T]\times\Omega)} \to 0 \ (m, n \to \infty) \tag{240}$$

so the completeness of Hilbert space ensures that $f_s^n \overset{L^2([0,T]\times\Omega)}{\to} f_s$ $(n \to \infty)$ converges to some limit process $f_s \in L^2([0,T]\times\Omega)$.

Verify that this limit gives the representation for F.

$$\mathbb{E}F + \int_0^T f_s\, dB_s = \mathbb{E} \lim_{n\to\infty} F_n + \int_0^T \lim_{n\to\infty} f_s^n\, dB_s \tag{241}$$

$$= \lim_{n\to\infty} \mathbb{E}F_n + \lim_{n\to\infty} \int_0^T f_s^n\, dB_s \tag{242}$$

$$= \lim_{n\to\infty} F_n = F \tag{243}$$

where the limit holds in L^2 sense and the second equation is due to the L^2 convergence of F_n and the convergence of stochastic integral.

Eventually, let's show that

$$\int_0^T f_s^n\, dB_s \overset{L^2}{\to} \int_0^T f_s\, dB_s \ (n\to\infty) \tag{244}$$

to complete the proof. By Ito's isometry,

$$\mathbb{E}\left(\int_0^T (f_s^n - f_s)\, dB_s\right)^2 = \mathbb{E}\int_0^T (f_s^n - f_s)^2\, ds \to 0 \ (n\to\infty) \tag{245}$$

and that's the end of the proof for existence (except the dense subset proposition).

For uniqueness, still use Ito's isometry

$$\int_0^T f_s^1\, dB_s = \int_0^T f_s^2\, dB_s \tag{246}$$

$$\int_0^T (f_s^1 - f_s^2)\, dB_s = 0 \tag{247}$$

$$||f_s^1 - f_s^2||_{L^2([0,T]\times\Omega)} = 0 \tag{248}$$

$$f_t^1 = f_t^2 \ a.a. \ (t,\omega) \tag{249}$$

□

Theorem 8. *For any L^2 continuous MG M_t, there always exists $f_t \in L^2([0,T]\times\Omega)$ such that $M_t = M_0 + \int_0^t f_s\, dB_s$.* ***(Martingale Representation Theorem)***

Proof. By Ito's representation theorem, for any fixed time t, always exists $f_s^t \in L^2([0,t]\times\Omega)$ such that

$$M_t - M_0 = \int_0^t f_s^t\, dB_s \tag{250}$$

with the process f_s^t depending on the fixed time point t. The next work is to prove that such process actually doesn't need to depend on t. Since we have not yet used martingality of M_t, try to apply it for $\forall 0\leq t_1\leq t_2$ to get

$$M_{t_1}=\mathbb{E}(M_{t_2}|\mathscr{F}_{t_1}) \tag{251}$$
$$=M_0+\mathbb{E}\left(\int_0^{t_2} f_s^{t_2}\,dB_s\Big|\mathscr{F}_{t_1}\right) \tag{252}$$
$$=M_0+\int_0^{t_1} f_s^{t_2}\,dB_s \tag{253}$$

Compared with the representation of M_{t_1} to conclude

$$\int_0^{t_1} f_s^{t_1}\,dB_s\overset{L^2}{=}\int_0^{t_1} f_s^{t_2}\,dB_s \tag{254}$$

and by Ito's isometry get

$$\mathbb{E}\int_0^{t_1}(f_s^{t_1}-f_s^{t_2})^2\,ds=0 \tag{255}$$
$$f_s^{t_1}=f_s^{t_2}\ a.a.\ (s,\omega) \tag{256}$$

so this f_s^t can be modified such that it does not depend on time t, and the proof is done.

□

A Lemma for Dense Subset

In the proof above, a crucial criterion is that the linear span of the set consisting of exponential MG with deterministic L^2 function $h(s)$, i.e.

$$span\left\{e^{\int_0^T h(s)\,dB_s-\frac{1}{2}\int_0^T h^2(s)\,ds},h\in L^2([0,T])\right\} \tag{257}$$

is a dense subset in $L^2(\Omega,\mathscr{F}_T)$ for fixed time T and deterministic h. Let's prove this lemma here to complete the whole proof for Ito's representation theorem.

The first lemma shows that for fixed time T, $L^2(\Omega,\mathscr{F}_T)$ has a dense subset consisting of all smooth and compactly supported functionals of finitely many time points of BM.

Lemma 2. *For fixed time T,*

$$\{\phi(B_{t_1},...,B_{t_n}):t_i\in[0,T],\phi\in C_0^\infty,\forall n\in\mathbb{N}\} \tag{258}$$

is a dense subset of $L^2(\Omega,\mathscr{F}_T)$.

Proof. $\forall g\in L^2(\Omega,\mathscr{F}_T)$, the main thought is to consider its projection onto the filtration spanned by the time points of BM. Denote $\mathscr{H}_n=\sigma(B_{t_1},...,B_{t_n})$, then $\mathscr{H}_n$ is a filtration with the projection of g to be $X_n=\mathbb{E}(g|\mathscr{H}_n)$.

It's obvious that X_n is a closed MG, so it converges *a.s.* and in L^1 to $X_\infty = \mathbb{E}(g|\mathscr{F}_T) = g$. Since g is L^2, $\sup_n \mathbb{E}^2(g|\mathscr{H}_n) \le \mathbb{E}g^2 < \infty$, by MG converegence theorem, this convergence is actually in L^2. So we just have to prove that $\mathbb{E}(g|\mathscr{H}_n)$ can be approximated in L^2 by the elements in the set.

Now that $\mathbb{E}(g|\mathscr{H}_n) \in \mathscr{H}_n$, so exists g_n Borel measurable such that $\mathbb{E}(g|\mathscr{H}_n) = g_n(B_{t_1}, ..., B_{t_n})$. Now consider compactly supported function $h_n(x) = g_n(x)\mathbb{I}_{||x||<n}$ as approximation of g_n at $(B_{t_1}, ..., B_{t_n})$. To get smoothness, just find a mollifier and take the convolution to conclude.

□

Lemma 3. *For fixed time T,*

$$span\left\{e^{\int_0^T h(s)\,dB_s - \frac{1}{2}\int_0^T h^2(s)\,ds}, h \in L^2([0,T])\right\} \tag{259}$$

for deterministic h is dense in $L^2(\Omega, \mathscr{F}_T)$.

Proof. By Hilbert space theory, to prove that it's a dense subset, just has to prove that its orthogonal complement is trivial. If $g \in L^2(\Omega, \mathscr{F}_T)$ is orthogonal to all elements in this set, want to prove that $g = 0$ *a.s.*.

Such g should satisfy

$$\forall \lambda, G(\lambda) = \int e^{\lambda_1 B_{t_1} + ... + \lambda_n B_{t_n}} g\,d\mathbb{P} = 0 \tag{260}$$

and extend all these variables as complex variables to see that

$$\forall z \in \mathbb{C}^n, G(z) = \int e^{z_1 B_{t_1} + ... + z_n B_{t_n}} g\,d\mathbb{P} \tag{261}$$

is actually holomorphic and is always 0 when all components are real, so G is always 0 on the whole $\mathbb{C}^n$ (isolation of the zeros of holomorphic function).

Let $\hat{\phi}$ be the Fourier transform of ϕ, so for $\forall \phi \in C_0^\infty, \forall y \in \mathbb{R}^n$,

$$\int \phi(B_{t_1}, ..., B_{t_n}) g\,d\mathbb{P} = (2\pi)^{-\frac{n}{2}} \int \hat{\phi}(y) G(iy)\,dy = 0 \tag{262}$$

by replacing ϕ by its inverse Fourier transform, using Fubini to change the integration order and replacing once again with G. (just the form of c.f. and Levy's inversion formula) Since all $\phi(B_{t_1}, ..., B_{t_n})$ form a dense subset, $g = 0$ and the lemma is proved.

□

Week 4

Stochastic Differential Equation

A general SDE in $\mathbb{R}^n$ with initial condition looks like

$$\begin{cases} dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dB_t \\ X_0 = x \end{cases} \tag{263}$$

where $x, X_t \in \mathbb{R}^n, b : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n, \sigma : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ and B_t is m-dim BM. Here b is the **drift coefficient** and σ is the **volatility coefficient**. Notice the difference between SDE and a general Ito process

$$X_t = x + \int_0^t \psi_s\, dB_s + \int_0^t \varphi_s\, ds \tag{264}$$

lies in the fact that b, σ are functions of the unknown process X_t while ψ_s, φ_s are known.

To prove the existence and uniqueness of the solution to SDE under special conditions, let's first state the Grownwall's inequality as a tool.

Theorem 9. ***(Grownwall's inequality)*** *Assume v is defined on interval $[a, +\infty)$ and is continuous with $A, F \in \mathbb{R}$. If $v(t) \leq F + A \int_a^t v(s)ds$, then*

$$v(t) \leq F e^{A(t-a)} \tag{265}$$

Assume v, β are defined on interval $[a, +\infty)$ and v is differentiable in the interior. If $v'(t) \leq \beta(t)\nu(t)$, then

$$v(t) \leq v(a) e^{\int_a^t \beta(s)ds} \tag{266}$$

Proof. Let's first prove the differential form. Set $u(t) = e^{\int_a^t \beta(s)\, ds}$ to see that $u'(t) = u(t)\beta(t)$ with $u(a) = 1, u(t) > 0$. Conside the derivative of their quotient

$$\frac{d}{dt}\frac{v(t)}{u(t)} = \frac{v'(t)u(t) - v(t)u(t)\beta(t)}{u^2(t)} \leq 0 \tag{267}$$

to find $\forall t \geq a$, $\frac{v(t)}{u(t)} \leq \frac{v(a)}{u(a)} = v(a)$. That's why

$$v(t) \leq v(a) e^{\int_a^t \beta(s)\, ds} \tag{268}$$

For the integral form, start by constructing (WLOG assume $A > 0$)

$$u(t) = e^{-A(t-a)} \int_a^t Av(s)\, ds \tag{269}$$

to find that $u(a) = 0$ and

$$u'(t) = Ae^{-A(t-a)}\left(v(t) - \int_a^t Av(s)\,ds\right) \leq AFe^{-A(t-a)} \tag{270}$$

integrate from a to t on both sides to get

$$u(t) - u(a) \leq AF\int_a^t e^{-A(s-a)}\,ds \tag{271}$$

$$u(t) \leq F(1 - e^{-A(t-a)}) \tag{272}$$

now turn back to v and take the derivative to conclude

$$A\int_a^t v(s)\,ds \leq F(e^{A(t-a)} - 1) \tag{273}$$

$$v(t) \leq Fe^{A(t-a)} \tag{274}$$

$\square$

Remark. *A special case is when v is continuous function and $F = 0$, so if $v(t) \leq A\int_0^t v(s)\,ds$, then $v \leq 0$.*

Basically the spirit of Grownwall is that if the derivative of a function $v'(t)$ is bounded by a multiple of its function value $\beta(t)v(t)$, then the function actually has an upper bound which is just the solution to the ODE $v'(t) = \beta(t)v(t)$.

The next theorem states the existence and uniqueness of the solution to a general SDE with some conditions.

Theorem 10. *Fix time $T > 0$ and assume $b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n, \sigma : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are deterministic functions with bounded time variable T and satisfy*

$$\exists c > 0, \forall t \in [0,T], \forall x \in \mathbb{R}^n, |b(t,x)| + |\sigma(t,x)| \leq c(1 + |x|) \tag{275}$$

$$\exists D > 0, \forall t \in [0,T], \forall x, y \in \mathbb{R}^n, |b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq D|x - y| \tag{276}$$

*then the SDE has unique solution in $L^2([0,T] \times \Omega)$ that has continuous sample path. The first condition is called **growth condition** and the second condition is called **Lipschitz condition**.*

Proof. First prove the **uniqueness**. If there are two solutions to the SDE: $X_t, \tilde{X}_t$. Consider plugging them in the function b, σ and form the difference process

$$\alpha_s = b(s, X_s) - b(s, \tilde{X}_s) \tag{277}$$

$$\gamma_s = \sigma(s, X_s) - \sigma(s, \tilde{X}_s) \tag{278}$$

and consider

$$dX_t - d\tilde{X}_t = \alpha_t\,dt + \gamma_t\,dB_t \tag{279}$$

to prove that these two solutions are the same in L^2, turn it into the integral form

$$X_t - \tilde{X}_t = \int_0^t \alpha_s \, ds + \int_0^t \gamma_s \, dB_s \tag{280}$$

and compute the L^2 norm to prove that it converges to 0

$$\begin{aligned}
\mathbb{E}(X_t - \tilde{X}_t)^2 &= \mathbb{E}\left(\int_0^t \alpha_s \, ds + \int_0^t \gamma_s \, dB_s\right)^2 && (281)\\
&\leq 2\mathbb{E}\left(\int_0^t \alpha_s \, ds\right)^2 + 2\mathbb{E}\left(\int_0^t \gamma_s \, dB_s\right)^2 && (282)\\
&= 2\mathbb{E}\left(\int_0^t \alpha_s \, ds\right)^2 + 2\mathbb{E}\left(\int_0^t \gamma_s^2 \, ds\right) && (283)\\
& && (284)
\end{aligned}$$

by the Ito's isometry and recall Cauchy's inequality on L^2 space that $\int_0^t \alpha_s \, ds \leq \sqrt{t \cdot \int_0^t \alpha_s^2 \, ds}$, so

$$\leq 2t \cdot \mathbb{E}\left(\int_0^t \alpha_s^2 \, ds\right) + 2\mathbb{E}\left(\int_0^t \gamma_s^2 \, ds\right) \tag{285}$$

The reason to turn $\int_0^t \alpha_s \, ds$ into $\int_0^t \alpha_s^2 \, ds$ is that the Lipschitz condition would then allow us to bound the square integral by the square difference of two solutions. To see that, notice that D is uniform: $|\alpha_s| \leq D|X_s - \tilde{X}_s|, |\gamma_s| \leq D|X_s - \tilde{X}_s|$

$$\begin{aligned}
&\leq (2D^2 t + 2D^2) \cdot \mathbb{E}\left(\int_0^t (X_s - \tilde{X}_s)^2 \, ds\right) && (286)\\
&= (2D^2 t + 2D^2) \cdot \int_0^t \mathbb{E}(X_s - \tilde{X}_s)^2 \, ds && (287)
\end{aligned}$$

Denote $v(s) = \mathbb{E}(X_s - \tilde{X}_s)^2$, so

$$v(t) \leq (2D^2 + 2D^2) \cdot \int_0^t v(s) \, ds \tag{288}$$

by Grownwall's inequality, conclude that

$$\begin{aligned}
\forall t \in [0,T], v(t) = 0 && (289)\\
\forall t \in [0,T], \mathbb{E}(X_t - \tilde{X}_t)^2 = 0 && (290)\\
\forall t \in [0,T], X_t = \tilde{X}_t \ a.s. && (291)
\end{aligned}$$

Then prove the **existence** of such solution by **Picard iteration**. Similar to that in the ODE theory, construct

$$X_t^0 = x \tag{292}$$

$$X_t^{k+1} = x + \int_0^t b(s, X_s^k)\, ds + \int_0^t \sigma(s, X_s^k)\, dB_s \ (k = 0, 1, ...) \tag{293}$$

we want to prove that X_t^n actually converges in L^2 sense to some limit as n goes to infinity and the limit is just a solution to the SDE. First prove convergence by showing that it's Cauchy.

Notice the fact that

$$\mathbb{E}(X_t^{k+1} - X_t^k)^2 = \mathbb{E}\left(\int_0^t [b(s, X_s^k) - b(s, X_s^{k-1})]\, ds + \int_0^t [\sigma(s, X_s^k) - \sigma(s, X_s^{k-1})]\, dB_s\right)^2 \tag{294}$$

$$\leq 2\mathbb{E}\left(\int_0^t [b(s, X_s^k) - b(s, X_s^{k-1})]\, ds\right)^2 + 2\mathbb{E}\left(\int_0^t [\sigma(s, X_s^k) - \sigma(s, X_s^{k-1})]\, dB_s\right)^2 \tag{295}$$

$$= 2\mathbb{E}\left(\int_0^t [b(s, X_s^k) - b(s, X_s^{k-1})]\, ds\right)^2 + 2\mathbb{E}\left(\int_0^t [\sigma(s, X_s^k) - \sigma(s, X_s^{k-1})]^2\, ds\right) \tag{296}$$

$$\leq 2t \cdot \mathbb{E}\int_0^t [b(s, X_s^k) - b(s, X_s^{k-1})]^2\, ds + 2\mathbb{E}\int_0^t [\sigma(s, X_s^k) - \sigma(s, X_s^{k-1})]^2\, ds \tag{297}$$

$$\leq (2D^2T + 2D^2) \cdot \int_0^t \mathbb{E}(X_s^k - X_s^{k-1})^2\, ds \tag{298}$$

by applying this iteratively, it's easy to see that

$$\mathbb{E}(X_t^{k+1} - X_t^k)^2 \leq (2D^2T + 2D^2) \cdot \int_0^t \mathbb{E}(X_s^k - X_s^{k-1})^2\, ds \tag{299}$$

$$\leq (2D^2T + 2D^2)^2 \cdot \int_0^t \int_0^{s_1} \mathbb{E}(X_{s_2}^{k-1} - X_{s_2}^{k-2})^2\, ds_2\, ds_1 \tag{300}$$

$$\leq ... \tag{301}$$

$$\leq (2D^2T + 2D^2)^k \cdot \int_{0<s_k<...<s_1<t} \mathbb{E}(X_{s_k}^1 - X_{s_k}^0)^2\, ds_k ... \, ds_2\, ds_1 \tag{302}$$

$$\tag{303}$$

The problem turns into getting an upper bound of $\mathbb{E}(X_t^1 - X_t^0)^2$

$$\mathbb{E}(X_t^1 - X_t^0)^2 = \mathbb{E}\left(\int_0^t b(s, X_s^0)\, ds + \int_0^t \sigma(s, X_s^0)\, dB_s\right)^2 \tag{304}$$

$$\leq 2t \cdot \mathbb{E}\int_0^t b^2(s, x)\, ds + 2\mathbb{E}\int_0^t \sigma^2(s, x)\, ds \tag{305}$$

the techniques are exactly the same as above (Cauchy, Ito's isometry...), then apply the growth condition for uniform

c that $|b(t,x)| \le c(1+|x|), |\sigma(t,x)| \le c(1+|x|)$

$$\begin{aligned}
\mathbb{E}(X_t^1 - X_t^0)^2 &\le 2c^2 t \cdot \mathbb{E}\int_0^t (1+|X_s^0|)^2\, ds + 2c^2 \cdot \mathbb{E}\int_0^t (1+|X_s^0|)^2\, ds && (306)\\
&\le 2c^2 t^2 \cdot \mathbb{E}(1+|X_s^0|)^2 + 2c^2 t \cdot \mathbb{E}(1+|X_s^0|)^2 && (307)\\
&\le 2c^2 T^2 \cdot \mathbb{E}(1+|X_s^0|)^2 + 2c^2 T \cdot \mathbb{E}(1+|X_s^0|)^2 && (308)\\
&\le A && (309)
\end{aligned}$$

for some fixed constant A. As a result, the estimation of the upper bound is

$$\begin{aligned}
\mathbb{E}(X_t^{k+1} - X_t^k)^2 &\le A(2D^2T + 2D^2)^k \cdot \int_{0<s_k<\ldots<s_1<t} ds_k \ldots ds_2\, ds_1 && (310)\\
&= A(2D^2T + 2D^2)^k \cdot \frac{t^k}{k!} && (311)
\end{aligned}$$

Now compute the L^2 norm to prove that the sequence is Cauchy for $n < m$ by telescoping:

$$\begin{aligned}
||X_t^m - X_t^n||_{L^2([0,T]\times\Omega)} &\le \sum_{k=n}^{m-1} ||X_t^{k+1} - X_t^k||_{L^2([0,T]\times\Omega)} && (312)\\
&= \sum_{k=n}^{m-1} \sqrt{\int_0^T \mathbb{E}(X_t^{k+1} - X_t^k)^2\, dt} && (313)\\
&\le \sum_{k=n}^{m-1} \sqrt{\int_0^T A(2D^2T + 2D^2)^k \cdot \frac{t^k}{k!}\, dt} && (314)\\
&= \sum_{k=n}^{m-1} \sqrt{\frac{AT^{k+1}(2D^2T + 2D^2)^k}{(k+1)!}} \to 0 \ (m,n\to\infty) && (315)
\end{aligned}$$

since $\sum_{k=0}^{\infty} \sqrt{\frac{AT^{k+1}(2D^2T+2D^2)^k}{(k+1)!}} < \infty$, so it's proved that this Picard sequence is Cauchy and its limit exists in L^2 space (actually it's **uniformly Cauchy in L^2 sense** since the upper bound does not depend is uniform in t)

$$\exists X_t \in L^2([0,T]\times\Omega), X_t^n \overset{L^2([0,T]\times\Omega)}{\to} X_t \ (n\to\infty) \tag{316}$$

To see that this limit is actually a solution to the SDE with continuous sample path, let's prove the convergence of the integrals. Now for $\forall t \in [0,T]$, set $k \to \infty$

$$\begin{aligned}
\mathbb{E}\left(\int_0^t [b(s,X_s^k) - b(s,X_s)]\, ds\right)^2 &\le t \cdot \mathbb{E}\int_0^t [b(s,X_s^k) - b(s,X_s)]^2\, ds && (317)\\
&\le D^2 t \cdot \int_0^t \mathbb{E}(X_s^k - X_s)^2\, ds \to 0 \ (k\to\infty) && (318)
\end{aligned}$$

for the stochastic integral, use Ito's isometry and Lipschitz

$$\mathbb{E}\left(\int_0^t [\sigma(s,X_s^k)-\sigma(s,X_s)]\,dB_s\right)^2 = \mathbb{E}\int_0^t [\sigma(s,X_s^k)-\sigma(s,X_s)]^2\,ds \tag{319}$$

$$\leq D^2\int_0^t \mathbb{E}(X_s^k-X_s)^2\,ds \to 0 \ (k\to\infty) \tag{320}$$

the convergence of $\int_0^t \mathbb{E}(X_s^k-X_s)^2\,ds$ is due to dominated convergence theorem since $\sup_k \mathbb{E}(X_s^k-X_s)^2$ is bounded and does not depend on s since it's uniformly Cauchy, thus integrable on bounded interval $[0,T]$. So we proved that the limit is actually a solution to this SDE.

Note that

$$\mathbb{E}\int_0^t \sigma^2(s,X_s)\,ds \leq c^2\cdot\mathbb{E}\int_0^t (1+|X_s|)^2\,ds < \infty \tag{321}$$

since $X_t\in L^2([0,T]\times\Omega)$, so $\mathbb{E}\int_0^t X_s^2\,ds<\infty$. This tells us that $\int_0^t \sigma(s,X_s)\,dB_s$ is always a continuous local MG, ensuring that there exists a modification of the solution with continuous sample paths.

□

Remark. *Note that this is actually the definition of a* ***strong solution*** *to this SDE. There is always a condition*

$$\mathbb{P}\left(\int_0^T |b(s,X_s)|+\sigma^2(s,X_s)\,ds<\infty\right)=1 \tag{322}$$

added for the general definition of strong solution to ensure the continuous modification, but it's not necessary in the theorem above since we are operating in L^2 space.

Note that the uniqueness of the solution in the theorem above is ***in the sense of modification*** *but not in the sense of indistinguishability. The existence of the solution is in the* ***global*** *sense.*

Example

Consider the SDE

$$\begin{cases} dX_t = X_t^2\,dt \\ X_0 = 1 \end{cases} \tag{323}$$

which is actually an ODE. The unique solution is

$$X_t = \frac{1}{1-t}\ t\in[0,1) \tag{324}$$

notice that $b(t,x)=x^2$ which is not Lipschitz in x and violates the growth condition, so this SDE does not have any global solutions (violates the existence).

Consider the SDE

$$\begin{cases} dX_t = 3X_t^{\frac{3}{2}}\,dt \\ X_0 = 0 \end{cases} \tag{325}$$

which is also an ODE. The solution is

$$X_t = 0/(t-a)^3 \tag{326}$$

so there are multiple solutions in the sense of modification (violates the uniqueness). Since $b(t,x) = 3x^{\frac{3}{2}}$ which is not Lipschitz in x and violates the growth condition.

Example

Consider the Black-Scholes model

$$dX_t = X_t(\mu\,dt + \sigma\,dB_t)\ (\mu,\sigma \in \mathbb{R}) \tag{327}$$

where $b(t,x) = \mu x, \sigma(t,x) = \sigma x$ both Lipschitz in x and satisfy the growth condition. By the theorem, the solution exists globally in L^2 space with a modification with continuous sample paths, also unique in the sense of modification.

To solve it, notice that if there's no stochastic terms, $dX_t = \mu X_t\,dt$ is an ODE with solution $X_t = X_0 \cdot e^{\mu t}$. As a result, consider changing the variables with $Y_t = \log(X_t)$

$$d\log(X_t) = \frac{1}{X_t}\,dX_t - \frac{1}{2}\frac{1}{X_t^2}\,d\langle X,X\rangle_t \tag{328}$$

$$= \mu\,dt + \sigma\,dB_t - \frac{1}{2}\frac{1}{X_t^2}\,d\langle X,X\rangle_t \tag{329}$$

where the bracket can be computed from the intergal from $X_t = X_0 + \mu\int_0^t X_s\,ds + \sigma\int_0^t X_s\,dB_s$ that

$$d\langle X,X\rangle_t = \sigma^2 X_t^2\,dt \tag{330}$$

plug in to get the solution

$$d\log(X_t) = \sigma\,dB_t + \left(\mu - \frac{\sigma^2}{2}\right)\,dt \tag{331}$$

$$\log(X_t) = \log(X_0) + \sigma B_t + \left(\mu - \frac{\sigma^2}{2}\right)t \tag{332}$$

$$X_t = X_0 \cdot e^{\sigma B_t + \left(\mu - \frac{\sigma^2}{2}\right)t} \tag{333}$$

The parameters μ,σ in the model can be estimated using quadratic variation of the observed data $X_t,\ t\in[0,T]$

since

$$\log(X_t) = \log(X_0) + \sigma B_t + \left(\mu - \frac{\sigma^2}{2}\right) t \tag{334}$$

$$\langle \log(X), \log(X) \rangle_t = \sigma^2 t \tag{335}$$

$$\mathbb{E}(\log(X_t) - \log(X_0)) = \left(\mu - \frac{\sigma^2}{2}\right) t \tag{336}$$

As a result, if there's enough data with equal time gap, the empirical quadratic variation over interval $[0,T]$ provides estimation of diffusion σ

$$\hat{\sigma^2} = \frac{\sum_j (\log X_{t_{j+1}} - \log X_{t_j})^2}{T} \overset{L^2, a.s.}{\to} \sigma^2 (\Delta t \to 0) \tag{337}$$

and the drift coefficient μ is always harder to estimate.

Example

Consider the Ornstein-Uhlenbeck process defined by SDE with a deterministic initial value condition

$$\begin{cases} dX_t = \alpha(m - X_t)\, dt + \sigma\, dB_t \\ X_0 = x \end{cases} \tag{338}$$

where m is the mean reversion level and α is the speed of mean reversion. The dynamics described by this SDE is that no matter what value X_t takes, it goes toward m with a stochastic noise of size σ. This model can be used to describe the fluctuation of interest rate around the mean interest rate m with the speed of regression described by α.

To solve the SDE, change the variable to set the regression level to 0: $Z_t = X_t - m$, so the SDE becomes

$$dZ_t = -\alpha Z_t\, dt + \sigma\, dB_t \tag{339}$$

Consider the case where $\sigma = 0$, i.e. there is no stochastic noise, then SDE turn into an ODE with solution

$$Z_t = Z_0 \cdot e^{-\alpha t} \tag{340}$$

turn the constant into a process and assume $Z_t = C_t e^{-\alpha t}$, apply Ito formula and compare to original SDE to get

$$dZ_t = -\alpha Z_t\, dt + e^{-\alpha t}\, dC_t \tag{341}$$

$$dC_t = \sigma e^{\alpha t}\, dB_t \tag{342}$$

$$C_0 = Z_0 = X_0 - m = x - m \tag{343}$$

solve out the **OU process**

$$C_t = x - m + \sigma \int_0^t e^{\alpha s}\, dB_s \tag{344}$$

$$Z_t = (x-m)e^{-\alpha t} + \sigma e^{-\alpha t} \int_0^t e^{\alpha s}\, dB_s \tag{345}$$

$$X_t = (x-m)e^{-\alpha t} + m + \sigma e^{-\alpha t} \int_0^t e^{\alpha s}\, dB_s \tag{346}$$

It's easy to see that since $\int_0^t e^{\alpha s}\, dB_s$ is a Wiener integral, it's a centered Gaussian random variable. In the case where $X_0 = x$ is deterministic, X_t is also Gaussian with expectation $(x-m)e^{-\alpha t} + m$ and variance

$$\mathbb{E}\left(\sigma e^{-\alpha t} \int_0^t e^{\alpha s}\, dB_s\right)^2 = \sigma^2 e^{-2\alpha t}\mathbb{E}\left(\int_0^t e^{2\alpha s}\, ds\right) = \sigma^2 e^{-2\alpha t} \int_0^t e^{2\alpha s}\, ds = \frac{\sigma^2(1-e^{-2\alpha t})}{2\alpha} \tag{347}$$

So the OU process has Gaussian distribution at each time point

$$X_t \sim N\left((x-m)e^{-\alpha t} + m, \frac{\sigma^2(1-e^{-2\alpha t})}{2\alpha}\right) \tag{348}$$

Since it's clear that the L^2 limit of Gaussian random variables is still Gaussian with the mean and variance just the respective limits of the mean and variance sequences (from characteristic function). As a result,

$$X_t \xrightarrow{L^2} N\left(m, \frac{\sigma^2}{2\alpha}\right) \ (t \to \infty) \tag{349}$$

providing the inspiration of taking the Gaussian $N\left(m, \frac{\sigma^2}{2\alpha}\right)$ as the **invariant distribution** of this SDE. The invariant distribution is defined in a way that **if the initial value X_0 is a random variable with the invariant distribution, then according to the dymanics defined by the SDE, at each time the underlying solution still follows such invariant distribution**.

Let's now prove that **the invariant distribution of OU process is Gaussian** $N\left(m, \frac{\sigma^2}{2\alpha}\right)$.

Remark. *The setting for initial value X_0 to be a given random variable is always that X_0 is independent of the whole BM and the filtration is set as ($\vee$ denotes the sigma field generated by the union of two sigma fields)*

$$\mathscr{F}_t = \sigma(B_s, 0 \le s \le t) \vee \sigma(X_0) \tag{350}$$

In general cases, denote X_0 as the initial value of the solution, a random variable. The solution to the SDE is now

$$X_t = (X_0 - m)e^{-\alpha t} + m + \sigma e^{-\alpha t} \int_0^t e^{\alpha s}\, dB_s \tag{351}$$

since X_0 is Gaussian and the Wiener integral is also Gaussian and they are independent, X_t must be Gaussian, only need to calculate the expectation and variance. By previous calculations,

$$\mathbb{E}X_t = (\mathbb{E}X_0 - m)e^{-\alpha t} + m = m \tag{352}$$

$$\begin{aligned} Var(X_t) &= Var((X_0 - m)e^{-\alpha t}) + Var\left(\sigma e^{-\alpha t}\int_0^t e^{\alpha s}\,dB_s\right) && (353)\\ &= e^{-2\alpha t}\frac{\sigma^2}{2\alpha} + \sigma^2 e^{-2\alpha t}\cdot \mathbb{E}\left(\int_0^t e^{\alpha s}\,dB_s\right)^2 && (354)\\ &= e^{-2\alpha t}\frac{\sigma^2}{2\alpha} + \sigma^2 e^{-2\alpha t}\cdot \int_0^t e^{2\alpha s}\,ds && (355)\\ &= \frac{\sigma^2}{2\alpha} && (356)\end{aligned}$$

so the invariant distribution is proved.

The OU process is a **continuous Markov Gaussian process** and it's **generally not a MG**. The autocorrelation function (with invariant initial value condition) is

$$cov(X_t, X_s) = \frac{\sigma^2}{2\alpha}e^{-\alpha|s-t|} \tag{357}$$

The calculation for general initial condition goes like

$$\begin{aligned} \forall s < t, cov(X_t, X_s) &= cov\left((X_0 - m)e^{-\alpha t} + \sigma e^{-\alpha t}\int_0^t e^{\alpha p}\,dB_p, (X_0 - m)e^{-\alpha s} + \sigma e^{-\alpha s}\int_0^s e^{\alpha q}\,dB_q\right) && (358)\\ &= e^{-\alpha(s+t)}Var(X_0) + \sigma^2 e^{-\alpha(s+t)}cov\left(\int_0^t, \int_0^s\right) && (359)\\ &= e^{-\alpha(s+t)}Var(X_0) + \sigma^2 e^{-\alpha(s+t)}cov\left(\int_0^s + \int_s^t, \int_0^s\right) && (360)\\ &= e^{-\alpha(s+t)}Var(X_0) + \sigma^2 e^{-\alpha(s+t)}\mathbb{E}\left(\int_0^s e^{\alpha p}\,dB_p\right)^2 && (361)\\ &= e^{-\alpha(s+t)}Var(X_0) + \sigma^2 e^{-\alpha(s+t)}\int_0^s e^{2\alpha p}\,dp && (362)\\ &= e^{-\alpha(s+t)}Var(X_0) + \frac{\sigma^2}{2\alpha}\left(e^{-\alpha(t-s)} - e^{-\alpha(t+s)}\right) && (363)\\ \forall t, s \geq 0, cov(X_t, X_s) &= e^{-\alpha(s+t)}Var(X_0) + \frac{\sigma^2}{2\alpha}\left(e^{-\alpha|t-s|} - e^{-\alpha(t+s)}\right) && (364)\end{aligned}$$

There are similar problems for estimating parameters m, σ, α. One can notice that

$$\langle X, X\rangle_t = \sigma^2 e^{-2\alpha t}\int_0^t e^{2\alpha s}\,ds = \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\alpha t}\right) \tag{365}$$

Strong solution and Weak Solution

The **strong solution** is defined as the solution $X_t \in \mathscr{F}_t$ adapted to the filtration generated by BM and the initial condition (if it's random)

$$\mathscr{F}_t = \sigma(B_s, 0 \leq s \leq t) \vee \sigma(X_0) \tag{366}$$

The uniqueness and global existence theorem stated for SDE is just proving that for strong solutions.

The **weak solution**, on the other hand, refers to the solution pair $(\tilde{X}_t, \tilde{B}_t, \mathscr{H}_t)$ such that $\tilde{X}_t \in \mathscr{H}_t$ is adapted to a specific filtration and $\tilde{B}_t \in \mathscr{H}_t$ is also BM under such filtration. The important point is that different weak solutions can live in the same or different filtered probability spaces. We will see that the uniqueness for weak solutions cannot be discussed in the pathwise sense unless two different weak solutions are living in the same probability space. Instead, the uniqueness is in the sense of finite-dimensional distribution.

Example

Consider SDE

$$\begin{cases} dX_t = dB_t \\ X_0 = 0 \end{cases} \tag{367}$$

and let B_t^1, B_t^2 be two BM living in two different probability spaces $(\Omega^1, \mathscr{F}^1, \mathscr{F}_t^1, \mathbb{P}^1)$ and $(\Omega^2, \mathscr{F}^2, \mathscr{F}_t^2, \mathbb{P}^2)$. For example, $(\Omega^1, \mathscr{F}^1, \mathscr{F}_t^1, \mathbb{P}^1) = ([0,1], \mathscr{B}_{[0,1]}, \mathscr{F}_t^1, \lambda), (\Omega^2, \mathscr{F}^2, \mathscr{F}_t^2, \mathbb{P}^2) = ([3,4], \mathscr{B}_{[3,4]}, \mathscr{F}_t^2, \lambda)$, both probability space equipped with Lebesgue measure but the sample space has no intersections.

By the definition of weak solution, $(\tilde{X}_t = B_t^1, \tilde{B}_t = B_t^1, \mathscr{F}_t = \sigma(B_s^1, 0 \leq s \leq t)), (\tilde{X}_t = B_t^2, \tilde{B}_t = B_t^2, \mathscr{F}_t = \sigma(B_s^2, 0 \leq s \leq t))$ are both weak solutions to this SDE. In this situation, $\mathbb{P}\left(B_t^1 = B_t^2\right)$ is not well-defined so no pathwise uniqueness argument can be made.

However, if we consider now the BM B_t^3 defined on a filtered probability space $(\Omega^3, \mathscr{F}^3, \mathscr{F}_t^3, \mathbb{P}^3)$, pairs $(\tilde{X}_t = B_t^3, \tilde{B}_t = B_t^3, \mathscr{F}_t = \sigma(B_s^3, 0 \leq s \leq t))$ and $(\tilde{X}_t = -B_t^3, \tilde{B}_t = -B_t^3, \mathscr{F}_t = \sigma(B_s^3, 0 \leq s \leq t))$ are both weak solutions, but now they are in the same filtered probability space and pathwise uniqueness arguments are well-defined. However,

$$\mathbb{P}\left(B_t^3 = -B_t^3\right) = \mathbb{P}\left(B_t^3 = 0\right) = 0 \tag{368}$$

so the sample paths of these two weak solutions has 0 probability of looking the same. Although the pathwise argument generally cannot work for weak solutions, the uniqueness in the sense of finite-dimensional distribution works since $(B_{t_1}^3, ..., B_{t_d}^3) \overset{d}{=} (-B_{t_1}^3, ..., -B_{t_d}^3)$.

Week 5

Example

Consider the following **Tanaka's equation**

$$\begin{cases} dX_t = sign(X_t)\, dB_t \\ X_0 = x \end{cases} \tag{369}$$

By **Tanaka's formula**, one knows that

$$d|B_t| = sgn(B_t)dB_t + dL_t \tag{370}$$

$$L_t = \lim_{\varepsilon\to 0} \frac{1}{2\varepsilon} \{s \in [0,t] : |B_s| < \varepsilon\} \tag{371}$$

where L_t is the local time of BM at 0.

Let's prove that this equation **has no strong solution, has a unique weak solution in the distribution sense and has no pathwise uniqueness**.

Firstly, assume that X_t is a strong solution. This means that X_t satisfies the SDE and is also adapted to the filtration $\mathscr{F}_t$ generated by BM B_t. Then

$$d\langle X, X\rangle_t = sign^2(X_t)\, dt = dt \tag{372}$$

and the Levy's characterization of BM tells us that X_t is a BM. This is actually telling us the uniqueness of weak solution, since each solution shall have the same finite-dimensional distribution as that of BM. To see the non-existence of strong solution, we need a **measurability argument**. Tanaka's formula for X_t implies

$$|X_t| = |x| + \int_0^t sign(X_s)\, dX_s + L_t^X \tag{373}$$

$$= |x| + B_t + L_t^X \tag{374}$$

$$B_t = |X_t| - |x| - L_t^X \tag{375}$$

where $|X_t| - |x| - L_t^X \in \mathscr{F}_t^{|X|} = \sigma(|X_s|, 0 \le s \le t)$ so $\mathscr{F}_t^B \subset \mathscr{F}_t^{|X|}$. However, by the definition of strong solution, $\mathscr{F}_t^X \subset \mathscr{F}_t^B$. As a result,

$$\forall t \ge 0, \mathscr{F}_t^X \subset \mathscr{F}_t^{|X|} \tag{376}$$

This can't be true since for a BM, the filtration generated by its absolute value process is a strict subset of the filtration generated by itself, a contradiction!

Remark. *An analogue is that x always contains more information than $|x|$. By knowing x one can compute $|x|$ but one can never infer x by knowing the value of $|x|$. Note that the measurability argument only holds for $\overset{a.s.}{=}$ but not* *for* $\overset{d}{=}$.

To be rigorous, $|X|$ *is a function of* X*, so* $\mathscr{F}_t^{|X|} \subset \mathscr{F}_t^X$*. However, since* X *is BM which is not trivial, there exists event like* $\{T_1 \leq t\} \in \mathscr{F}_t^X$ *where* T_1 *is the first hitting time to 1 of* X_t*, but* $\{T_1 \leq t\} \notin \mathscr{F}_t^{|X|}$*. Another example is* $\{X_t > 0\}$*, if it is in* $\mathscr{F}_t^{|X|}$*, then since* $\{|X_1| = 1\} \in \mathscr{F}_t^{|X|}$*, their intersection is* $\{X_1 = 1\} \in \mathscr{F}_t^{|X|}$ *which is a contradiction since* $\{1\}$ *is not symmetric w.r.t. 0.*

The following theorem characterizes BM by a calculation of the quadratic variation, continuity of sample paths is even not required as long as X_t has the structure of local martingale.

Theorem 11. *If* X_t *is local martingale with* $X_0 = 0$ *adapted to filtration* $\mathscr{F}_t$*, then* X *is BM under filtration* $\mathscr{F}_t$ *iff* $X_t^2 - t$ *is* $\mathscr{F}_t$ *adapted continuous local martingale iff* $\forall t \geq 0, \langle X, X \rangle_t = t$ ***(Levy's characterization of BM)***

Let's then construct a weak solution to the Tanaka's equation. Since we already know that X_t has to be BM, set $X_t = x + \hat{B}_t$ with $\hat{B}$ to be any BM. Define

$$\tilde{B}_t = \int_0^t sign(X_s)\, dX_s \tag{377}$$

then the pair $(X_t, \tilde{B}_t)$ is the solution to the SDE because $dX_t = d\hat{B}_t$ and $d\tilde{B}_t = sign(X_t)\, dX_t$, so $dX_t = sign(X_t)\, d\tilde{B}_t$ satisfies the SDE and $\langle \tilde{B}, \tilde{B} \rangle = \langle X, X \rangle_t = t$, proving that $\tilde{B}$ is a BM.

Remark. *Although there is no way to find a solution* X_t *adapted to the filtration generated by* B_t*, which is the BM in the SDE, there is a way to specify another BM* $\tilde{B}_t$ *and* X_t *such that they have some connections and work as the solution to this SDE. The "weak" refers to replacing a general BM with a specific chosen BM in the SDE (allows connections with the constructed solution).*

We can also see that the weak solution to this SDE has no pathwise uniqueness (the uniqueness for strong solution). The definition of **pathwise uniqueness** is that the solution has pathwise uniqueness if any two solutions in the same filtered probability space almost surely have the same path.

Assume that $(X_t, \tilde{B}_t)$ is a weak solution pair of Tanaka's equation, define $\tau = \inf\{t \geq 1, X_t = 0\}$ be the first hitting time after time 1 that hits 0. Define the reflected process

$$\tilde{X}_t = \begin{cases} X_t & t \leq \tau \\ -X_t & t > \tau \end{cases} \tag{378}$$

to find that the pair $(\tilde{X}_t, \tilde{B}_t)$ is still a weak solution.

$$x + \int_0^t sign(\tilde{X}_s)\, d\tilde{B}_s = x + \int_0^{t\wedge\tau} sign(X_s)\, d\tilde{B}_s - \int_{t\wedge\tau}^t sign(X_s)\, d\tilde{B}_s \tag{379}$$

$$= x + \int_0^{t\wedge\tau} dX_s - \int_{t\wedge\tau}^t dX_s \tag{380}$$

$$= 2X_{t\wedge\tau} - X_t = \tilde{X}_t \tag{381}$$

since $X_\tau = 0$. However, $\tilde{X}_t$ and X_t have totally different sample paths, so pathwise uniqueness fails.

The following theorem asserts the reason why pathwise uniqueness is crucial to consider in solving SDE.

Theorem 12. *If a SDE has weak solution and the solution has pathwise uniqueness, then its strong solution exists.* ***(Yamada-Watanabe)***

For Tanaka's equation, since weak solution exists, the non-existence of strong solution directly implies the failure of pathwise uniqueness.

Ito Diffusion

Let's consider the **Ito diffusion**, which is defined by a time-homogeneous SDE

$$\begin{cases} dX_t = b(X_t)\,dt + \sigma(X_t)\,dB_t \ (t \geq s) \\ (X_t \in \mathbb{R}^n, B_t \in \mathbb{R}^m, b : \mathbb{R}^n \to \mathbb{R}^n, \sigma : \mathbb{R}^n \to \mathbb{R}^{n\times m}) \\ X_s = x \end{cases} \tag{382}$$

that is a n-dim process generated by m-dim BM with given initial condition at time s. In addition, it is assumed that b, σ are both Lipschitz on $\mathbb{R}^n$.

Remark. *Ito diffusion refers to* ***the drift and diffusion coefficient*** b, σ ***not depending on time*** t ***and the Lipschitz condition*** *is here to ensure the existence and uniqueness of the strong solution of such SDE (proved above). Note that the growth condition is naturally satisfied since Lipschitz condition implies*

$$\exists D, \forall x, y \in \mathbb{R}^n, ||b(x) - b(y)|| + ||\sigma(x) - \sigma(y)|| \leq D||x - y|| \tag{383}$$

so the growth condition

$$\exists C, \forall x \in \mathbb{R}^n, ||b(x)|| + ||\sigma(x)|| \leq C(1 + ||x||) \tag{384}$$

is always true. If the SDE has a general form (b, σ depends on time t), then such process is an Ito process. (distinguish)

Next we talk about **properties** of such Ito diffusion process. The first property is the time homogeneity naturally implied by this structure. Let's denote $X_{s+h}^{s,x}$ as the solution to the SDE above with initial condition $X_s = x$, so it's obvious that

$$X_{s+h}^{s,x} = x + \int_s^{s+h} b(X_u^{s,x})\,du + \int_s^{s+h} \sigma(X_u^{s,x})\,dB_u \tag{385}$$

change the variables by $v = u - s$ to drag the initial condition to time 0 to get

$$X_{s+h}^{s,x} = x + \int_0^h b(X_{v+s}^{s,x})\,dv + \int_0^h \sigma(X_{v+s}^{s,x})\,dB_{v+s} \tag{386}$$

$$= x + \int_0^h b(X_{v+s}^{s,x})\,dv + \int_0^h \sigma(X_{v+s}^{s,x})\,d(B_{v+s} - B_s) \tag{387}$$

note that here we make use of the fact that

$$\int_0^h \sigma(X_{v+s}^{s,x})\,dB_s = 0 \tag{388}$$

since the integral is w.r.t. v but not s. The motivation is to add another part to B_{s+v} such that it becomes another BM. As a result, for BM $\tilde{B}_v = B_{v+s} - B_s$ due to the Markov property,

$$X_{s+h}^{s,x} = x + \int_0^h b(X_{v+s}^{s,x})\,dv + \int_0^h \sigma(X_{v+s}^{s,x})\,d\tilde{B}_v \tag{389}$$

Notice that $X_t^{0,x}$, which is the solution to the SDE with initial condition $X_0 = x$ satisfies

$$X_t^{0,x} = x + \int_0^t b(X_s^{0,x})\,ds + \int_0^t \sigma(X_s^{0,x})\,dB_s \tag{390}$$

by comparing the two equations above, we conclude that

$$(X_{s+t}^{s,x}, \tilde{B}_t), (X_t^{0,x}, B_t) \tag{391}$$

are both weak solutions to the same SDE for Ito diffusion. Since the Lipschitz condition of original SDE holds, the weak solution has uniqueness in distribution. That's why we conclude

$$\forall s \geq 0, \forall x \in \mathbb{R}^n, \left\{X_{s+t}^{s,x}\right\}_{t\geq 0} \overset{d}{=} \left\{X_t^{0,x}\right\}_{t\geq 0} \tag{392}$$

the finite-dimensional distribution of the Ito diffusion process is the same as long as the initial starting point x is the same regardless of the initial time s. This is called **time-homogeneity** of Ito diffusion process.

Here we have made use of the theorem that if the existence and uniqueness condition of a SDE holds, weak solution is unique in the sense of distribution.

Theorem 13. *Consider SDE with initial condition*

$$\begin{cases} dX_t = b(t, X_t)\,dt + \sigma(t, X_t)\,dB_t \\ X_0 = x \end{cases} \tag{393}$$

If the existence and uniqueness condition holds (growth condition and Lipschitz condition), any two weak solutions have the same distribution.

Proof. Assume $(\hat{X}_t, \hat{B}_t, \hat{\mathscr{H}_t}), (\tilde{X}_t, \tilde{B}_t, \tilde{\mathscr{H}_t})$ are two weak solution pairs, then by the existence of strong solution, construct strong solutions for filtration $\hat{\mathscr{H}_t}$ and BM $\hat{B}_t$ denoted $\hat{Z}_t$ and strong solution for filtration $\tilde{\mathscr{H}_t}$ and BM $\tilde{B}_t$ denoted $\tilde{Z}_t$. By the uniqueness of strong solution in the sense of modification, $\forall t, \hat{X}_t = \hat{Z}_t, \tilde{X}_t = \tilde{Z}_t$ *a.s.* To prove that $\hat{X}_t \overset{d}{=} \tilde{X}_t$, only need to prove that $\hat{Z}_t \overset{d}{=} \tilde{Z}_t$.

Recall that the strong solutions $\hat{Z}_t, \tilde{Z}_t$ are constructed as the L^2 limit of the Picard iteration sequence

$$\hat{Z}_t^n \overset{L^2}{\to} \hat{Z}_t \ (n \to \infty) \tag{394}$$

$$\tilde{Z}_t^n \overset{L^2}{\to} \tilde{Z}_t \ (n \to \infty) \tag{395}$$

since the underlying SDE of the construction of $\hat{Z}_t^n, \tilde{Z}_t^n$ are the same and the initial conditions are also the same

$$\forall n, \hat{Z}_t^n \overset{d}{=} \tilde{Z}_t^n \tag{396}$$

their L^2 limits should also be the same. As a result, the weak solutions in the same probability space have the same distribution.

□

Remark. *These two weak solutions don't have to be in the same probability space since distribution is only determined by the distribution function. Note that $\hat{Z}_t, \hat{X}_t$ are in the same probability space, that's why we can use the uniqueness argument to conclude that at any time they are equal almost surely.*

The second property is the **Markov property**. With the statement that **for any bounded Borel measurable $f : \mathbb{R}^n \to \mathbb{R}$ and $\forall t, h \geq 0$,**

$$\mathbb{E}_x(f(X_{t+h})|\mathscr{F}_t) = \mathbb{E}_{X_t} f(X_h) \tag{397}$$

where $\mathscr{F}_t$ is the filtration generated by BM B_t in the SDE and $\mathbb{E}_x$ denotes the expectation under the condition that the process starts from x.

This comes from the observation that

$$\forall t, h \geq 0, X_{t+h}^{0,x} = X_t^{0,x} + \int_t^{t+h} b(X_u^{0,x}) \, du + \int_t^{t+h} \sigma(X_u^{0,x}) \, dB_u \tag{398}$$

$$X_{t+h}^{t,X_t} = X_t + \int_t^{t+h} b(X_u^{t,X_t}) \, du + \int_t^{t+h} \sigma(X_u^{t,X_t}) \, dB_u \tag{399}$$

the uniqueness of strong solution tells us that

$$\forall t, h \geq 0, X_{t+h}^{0,x} = X_{t+h}^{t,X_t} \ a.s. \tag{400}$$

As a result,

$$\mathbb{E}(f(X_{t+h}^{t,X_t})|\mathscr{F}_t) = \mathbb{E}(f(X_h^{0,x}))|_{x=X_t} \tag{401}$$

$$= \mathbb{E}_x(f(X_{t+h}^{t,X_t})) \tag{402}$$

Diffusions

Markov property(all diffusion process). $\forall f \in C_b(\mathbb{R}^n), t, h, \mathbb{E}(f(X_{t+h})|\mathscr{F}_t) = \mathbb{E}(f(X_{t+h})|X_t)$. Flow property: consider starting from 0 ($X_0^{0,x}$), to $t+h$ ($X_{t+h}^{0,x}$). Consider stopping at time t and restart to time $t+h$(first to $X_t^{0,x}$ then to $X_h^{0,X_t^{0,x}}$). So the flow property is stating $X_h^{0,X_t^{0,x}} = X_{t+h}^{0,x}$ $a.s.$. Claim that $\forall s \geq t, X_s = \varphi(X_t^{0,x}, \tilde{B}_{u,u\leq s})$ where $\tilde{B}_s = B_{t+s} - B_t$ is a new BM.

Strong Markov property $\forall f \in C_b(\mathbb{R}^n), h, \tau < \infty \ a.s., \mathbb{E}(f(X_{\tau+h})|\mathscr{F}_\tau) = \mathbb{E}(f(X_{\tau+h})|X_\tau)$ followed by strong Markov of BM.