Stochastic Calculus and Applications

Lectured by Dr. Roland Bauerschmidt, Typed by Jiwoon Park

Lent 2019

Stochastic Calculus
∃ website : http://www.statslab.cam.ac.uk/∼rb812/teaching/sc2019/
Roland Bauerschmidt
Example Classes given by Daniel Heydecker, location and time is posted on the website.

(18th January, Friday)

1 Introduction

1.1 Motivation

Ordinary differential equations, $\dot{x}(t) = F(x(t))$, are fundamental in analysis.

A stochastic differential euqation(SDE) is given in form $\dot{x}(t) = F(x(t)) + \eta(t)$, where η is a random noise.
What should η be?

- For $|t - s| \gg 0$, $\eta(t)$ and $\eta(s)$ should be essentially independent.
- Our idealisation would be to assume $\eta(t)$ and $\eta(s)$ should be independent for any $t \neq s$. If we just assume this independence, then there is a lot of freedom of the choice of η, and there is also an issue of existence of such function η.

Such an η exists, the **White Noise**, but it is only a radom generalised function (random Schwartz distribution). But, even if $F = 0$, to make sense of

$$\dot{x} = \eta, \quad i.e.\ x(t) - x(0) = \int_0^t \eta(s) ds$$

deterministically, η *should at least be signed measure.* However,

- White noise is *not* a random signed mesure
- If the above equation holds, for any $0 = t_0 < t_2 < \cdots$, the increments

$$x(t_i) - x(t_{i-1}) = \int_{t_{i-1}}^{t_i} \eta(s) ds$$

  should be independent and their variance should be proportional to $|t_i - t_{i-1}|$. Hence, it is reasonable to think that x should be a Browninan motion.

In which sense can we make sense of this?

Stochastic calculus is similar to ordinary calculus. First half of the course would be devoted in developing the theorems and examples to justify the stochastic integral, but once we have the tool, it is not very difficult to use it.

1.2 The Wiener Integral

The Wiener integral is the first model where the stochastic calculus makes sense most easily.

Definition) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Then $S \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a **Gaussian space** if S is a closed linear subspce and any $X \in S$ is a centred Gaussian random variable.

Example : Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which a sequence of independent random variables $X_i \sim N(0,1)$ is defined. Then the X_i are an orthonormal system in $L^2(\Omega, \mathcal{F}, \mathbb{P})$:

$$\mathbb{E}(X_i X_j) = 0 \quad \text{for } i \neq j \text{ and } \quad \mathbb{E}(X_i^2) = 1$$

and $S = \overline{\text{span}\{X_i\}}$ is a Gaussian space. (*Exercise* : the limit in L^2 of Gaussian random variables is Gaussian.)

Proposition) Let H be a separable Hilbert space and $(\Omega, \mathcal{F}, \mathbb{P})$ as in the example. Then there is an isomtery $I : H \to S$. In particular, for every $f \in H$, there is a random variable $I(f) \in S$ such that

$$I(f) \sim N(0, (f,f)_H) \quad \text{and} \quad \mathbb{E}(I(f)I(g)) = (f,g)_H$$

Moreover, $I(\alpha f + \beta g) = \alpha I(f) + \beta I(g)$ a.s.

proof) The proof would be done by identifying the basis. Let $(e_i)_{i=1}^{\infty}$ be an orthonormal basis for H. For $f \in H$, set

$$I(f) = \sum_{i=1}^{\infty} (f, e_i) X_i \in L^2(\Omega, \mathcal{F}, \mathbb{P})$$

Then the limit exists in L^2 since $(\sum_{i=1}^{k}(f,e_i)X_i)_k$ is a Cauchy sequence in L^2 :

$$\mathbb{E}((\sum_{i=1}^{k}(f,e_i)X_i - \sum_{i=1}^{l}(f,e_i)X_i)^2) \leq \sum_{i=k}^{l} |(f,e_i)|^2 \to 0 \quad \text{as } k,l \to \infty$$

since $f \in H$. [In fact, this convergence is also made almost surely, becuase $k \mapsto \sum_{i=1}^{k}(f,e_i)X_i$ is also a martingale, bounded in L^2, hence L^2-martigale convergence theorem applies.] The map I is an isometry since it maps the orthonormal basis (e_i) to the orthonormal system (X_i) in $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

(End of proof) $\square$

Definition) A Gaussian **white noise** on $\mathbb{R}_+$ is an isometry WN from $L^2(\mathbb{R}_+)$ into a Gaussian space. For $A \subset \mathbb{R}_+$ Borel, write $WN(A) = WN(1_A)$.

Proposition)

(1) For $A \subset \mathbb{R}_+$, Borel, $|A| < \infty$, $WN(A) \sim N(0, |A|)$.

(2) For $A, B \subset \mathbb{R}_+$ Borel, $A \cap B = \phi$ then $WN(A)$ and $WN(B)$ are independent.

(3) If $A = \cup_{i=1}^{\infty} A_i$ for disjoint sets A_i with $|A_i| < \infty$, $|A| < \infty$, then

$$WN(A) = \sum_{i=1}^{\infty} WN(A_i) \quad \text{in } L^2 \text{ and a.s.} \quad \cdots\cdots\cdots (\star)$$

proof)

(1) This holds since $(1_A, 1_A) = |A|$

(2) This holds since $\mathbb{E}[WN(A)WN(B)] = (1_A, 1_B) = 0$.

(3) Let $M_n = \sum_{i=1}^n WN(A_i)$. Then this is a martingale bounded in L^2 (the boundedness comes from $\mathbb{E}(M_n^2) = \sum_{i=1}^n \mathbb{E}[WN(A_i)^2] = \sum_{i=1}^n |A_i| \leq |A|$). Thus $\sum_{i=1}^\infty WN(A_i)$ converges a.s. and in L^2.

   Similarly, we also have $\mathbb{E}((WN(A) - M_n)^2) \to 0$ so $(\star)$ holds.

Remark : The proposition seems to indicate that the white noise is a random measure $A \in \mathcal{B}(\mathbb{R}_+) \mapsto WN(\omega, A)$, where $\omega \in \Omega$, but it is *not*! In $(\star)$, the event $E \subset \Omega$ of ω for which $(\star)$ holds depends on the sets A_i.

However, stochastic calculus is formulated in such a way that this still is used as a useful notion of convergence, hence can be used like a random measure.

For $t \geq 0$, define the Brownian motion as $B_t = WN([0, t])$, just like the integration of white noise from 0 to t.

(21st January, Monday)

Last time : we examined the white noise $WN(A) \sim N(0, |A|)$ for $A \subset \mathbb{R}_+$ a measurable set. For $t \geq 0$, define $B_t = WN([0, t])$.

Fact : For any $t_1, \cdots, t_n$, the vector $(B_{t_i})_{i=1}^n$ is jointly Gaussian and $\mathbb{E}(B_s B_t) = s \wedge t$ for all $s, t \geq 0$. Moreover, $B_0 = 0$ a.s. and $B_t - B_s$ is independent from $\sigma(B_r : r \leq s)$, $B_t - B_s \sim N(0, t - s)$ for $t \geq s$.

Remark : (See *Advanced Probablity*) There is a modification of (B_t) such that $t \mapsto B_t$ is continuous, a.s.

Example : Let $f \in L^2(\mathbb{R}_+)$ be a step function, $f = \sum_{i=1}^n f_i 1_{[t_i, t_{i+1}]}$, $t_i < t_{i+1}$. Then

$$WN(f) = \sum_{i=1}^n f_i (B_{t_{i+1}} - B_{t_I})$$

This motivates the *notation* $WN(f) = \int f(s) dB_s$. If $(B_t)_t$ was a function of finite variation for a.e. $\omega \in \Omega$, the last line could be defined as a Lebesgue-Stieltjes Integral. Though unfortunately, it is not.

(So we will decompose B_t into a part with finite variation and a semi-martingale part to jusify this integral.)

1.3 The Lebesgue-Stieltjes Integral

For an interval $T \subset \mathbb{R}$, we always use the Borel σ-algebra.

Definition) Let $T > 0$.

- A **signed measure** μ on $[0, T]$ is the difference of two finite positive measure $\mu_\pm$ on $[0, T]$ with disjoint support.
- The decomposition $\mu = \mu_+ - \mu_-$ is called the **Hahn-Jordan decomposition of** μ.
- The **total variation** of a signed measure $\mu = \mu_+ - \mu_-$ is the positive meuasre $|\mu| = \mu_+ + \mu_-$.

Proposition) *(Hahn-Jordan)* For any positive measures μ_1, μ_2 on $[0,T]$(we do not require them to have disjoint support), there is a signed measure μ s.t. $\mu = \mu_1 - \mu_2$.

proof) Let $\nu = \mu_1 + \mu_2$. Since μ_1, μ_2 are absolutely continuous with respect to ν, by the *Radon-Nikodym Theorem*, there are Borel functions $f_i \geq 0 (i = 1, 2)$ on $[0,T]$ such that

$$\mu_i(dt) = f_i(t)\nu(dt)$$

Let $f(t) = f_1(t) - f_2(t)$. Then

$$\begin{aligned}(\mu_1 - \mu_2)(dt) = f(t)\nu(dt) &= f(t)^+\nu(dt) - f(t)^-\nu(dt) \\ &= \mu_+(dt) - \nu_-(dt)\end{aligned}$$

where $f(t)^+ = f(t) \vee 0$, $f(t)^- = -f(t) \wedge 0$ are the positive/negative parts of f. This gives the required decomposition into disjoint measures.

(End of proof) $\square$

Definition) Let $T \geq 0$.

- A function $a : [0,T] \to \mathbb{R}$ is **càdlàg**, for which we also write $a \in D([0,T])$, if $a(t_+) = a(t)$ for all t and $a(t_-)$ exists for all t.

  Here, $a(t_\pm) = \lim_{s \to 0^\pm} a(t+s)$.

- The **total variation** of a function $a : [0,T] \to \mathbb{R}$ is

$$v_a(0,T) = \sup \left\{ \sum_{i=1}^{n} |a(t_i) - a(t_{i-1})| \ : \ 0 = t_0 < t_1 < \cdots < t_n = T \right\}$$

- A function $a : [0,T] \to \mathbb{R}$ is of **bounded variation**, for which we write $a \in BV([0,T])$, if $v_a(0,T) < \infty$.

We find in many instances of Stochastic Calculus(and in general probability), that rather than trying to work with a measure, it is often more useful to work with distribution functions.

Proposition)

(i) Let μ be a signed measure on $[0,T]$. Then $a(t) = \mu([0,t])$ is càdlàg and $|\mu|((0,t]) = v_a(0,t)$ (i.e. $|\mu|([0,T])) = |a(0)| + v_a(0,t)$)

In particular, $a \in BV([0,T])$.

(ii) Let $a : [0,T] \to \mathbb{R}$ be càdlàg of bounded variation. The there is a signed measure μ such that $a(t) = \mu([0,t])$.

Note that, **Fact :** (From Prob.& Meas.) The map $v \mapsto f$, $f(t) = \nu([0,T])$ (distribution function) is a bijection from finite positive measures on $[0,T]$ to increasing right-continuous functions.

proof of Prop.)

(i) Let $\mu = \mu_+ - \mu_-$ be the Hahn-Jordan decomposition of μ. Letting $\mu_+([0,t]) = a_+(t)$, $\mu_-([0,t]) = a_-(t)$, we have $a(t) = \mu_+([0,t]) - \mu_-([0,t]) = a_+(t) - a_-(t)$ is càdlàg since $a_\pm$ are increasing right-continuous function.

Claim : $v_a(0,t) \leq |\mu|([0,t])$.

: For any subdivision $0 = t_0 < t_1 < \cdots < t_n = t$, one has

$$\sum_{i=1}^{n} |a(t_i) - a(t_{i-1})| = \sum_{i=1}^{n} |\mu((t_{i-1}, t_i])| \leq |\mu|((t_{i-1}, t_i])$$

and so $v_a(0,t) \leq |\mu|((0,t])$.

Claim : For any nested sequence of partitions $0 = t_0^{(m)} < t_1^{(m)} < \cdots < t_{n_m}^{(m)} = t$ with $\max_i |t_i^{(m)} - t_{i-1}^{(m)}| \to 0$, one has

$$|\mu|((0,t]) = \lim_{m\to\infty} \sum_{i=1}^{n_m} |a(t_i^{(m)}) - a(t_{i-1}^{(m)})|$$

In particular, $v_a(0,t) \geq |\mu|((0,T])$.

: Consider the probability measure $P(ds) = \frac{|\mu|(ds)}{|\mu|((0,t])}$ on $(0,t]$. Let $\mathcal{F}_m = \sigma((t_{i-1}^{(m)}, t_i^{(m)}] : 1 \leq i \leq n_m)$. Then $\mathcal{F}_{m+1} \supset \mathcal{F}_m$. Let $X = \frac{d\mu}{d|\mu|} = 1_{\text{supp}(\mu_+)} - 1_{\text{supp}(\mu_-)}$ and let $X_m = \mathbb{E}(X|\mathcal{F}_m)$. Then we may write, for $s \in (t_{i-1}^{(m)}, t_i^{(m)}]$,

$$X_m(s) = \frac{\mu((t_{i-1}^{(m)}, t_i^{(m)}])}{|\mu|((t_{i-1}^{(m)}, t_i^{(m)}])} = \frac{a(t_i^{(m)}) - a(t_{i-1}^{(m)})}{|\mu|((t_{i-1}^{(m)}, t_i^{(m)}])}$$

Since (X_m) is a bounded martingale, we also have $X_m \to Y$ in L^1 and a.s., for some random variable Y. Since $\bigvee \mathcal{F}_m = \mathcal{B}((0,t])$ ($\bigvee \mathcal{F}_m = \sigma(\cup_m \mathcal{F}_m)$), it follows that $X = Y$ a.s., and therefore $\mathbb{E}[|X_m|| \to \mathbb{E}[|Y|] = \mathbb{E}[|X|] = 1$, which is equivalent to

$$\frac{1}{|\mu|((0,t])} \sum_{i=1}^{n_m} |a(t_i^{(m)}) - a(t_{i-1}^{(m)})| \to 1$$

which is the claim.

(23rd January, Wednesday)

From Last Time :
Proposition)

(i) Let μ be a signed measure on $[0,T]$. Then $a(t)$ is càdlàg and $|\mu|((0,t]) = v_a(0,t)$.

(ii) Let $a : [0,T] \to \mathbb{R}$ be càdlàg and of bounded variation. then there is a signed measure μ such that $\mu([0,t])$.

proof of (ii)) Let a be as in the statement (ii). Define $a_\pm(t) = \frac{1}{2}(v_a(0,t) \pm a(t))$. First note that both $a_\pm$ are positive. We would like to define μ as the difference of these two.

Claim : $a_\pm$ are both increasing.

: Let $0 = t_0 < t_1 < \cdots < t_n = t$ be a subdivision of $[0,t]$ and $s > t$. Then $t_0 < \cdots < t_n = t < s$ is a subdivision of $[0,s]$. We pick $t_0 < \cdots < t_n$ such that $\sum_{i=1}^{n} |a(t_i) - a(t_{i=1}))| \geq v_a(0,t) - \epsilon$. Then

$$\begin{aligned} 2a_\pm(s) = v_a(0,s) \pm a(s) &\geq \sum_{i=1}^{n} |a(t_i) - a(t_{i-1})| + |a(s) - a(t)| \pm a(s) \\ &\geq v_a(0,t) - \epsilon \pm a(t) = a_\pm(t) - \epsilon \end{aligned}$$

This holds for any $\epsilon > 0$, so we have $a_\pm(s) \geq a_\pm(t)$ for all $\epsilon > 0$, and therefore $a_\pm$ are increasing.

Claim : v_a is right-continuous.

Exercise.

The second claim implies that $a_\pm$ is right-continuous, so $a_\pm(t) = \tilde{\mu}_\pm([0,t])$ for some finite positive measure $\tilde{\mu}_\pm$.

Let $\mu = \tilde{\mu}_+ - \tilde{\mu}_-$. Then μ is a signed measure and $a(t) = a_+(t) - a_-(t) = \mu([0,t])$.

(End of proof) □

We are finally ready to define the Lebesgue-Stieltjes integral. Once we define this, we will be able to define stochastic integrals for stochastic processes of finite variation.

Having done this class of distributions that are integrable with respect to will naturally extend to more general cases, *e.g.* Brownian motion as the distribution of the white noise.

Definition) Let $a : [0,T] \to \mathbb{R}$ be càdlàg of bounded variation, and let μ be the associated signed measures. For $h \in L^1([0,T], |\mu|)$, the **Lebesgue-Stieltjes integral** is defined by

$$\int_s^t h(u)da(u) = \int_{(s,t]} h(u)\mu(du), \quad 0 \le s < t \le T$$
$$\int_s^t h(s)|da(u)| = \int_{(s,t]} h(u)|\mu|(du)$$

We also write $(h \cdot a)(t) = \int_0^t h(s)da(s)$ for $0 \le t \le T$.

Fact : Let $a : [0,T] \to \mathbb{R}$ be càdlàg and BV (bounded variation), $h \in L^2([0,T], |da|)$. Then

$$\left| \int_0^t h(s)da(s) \right| \le \int_0^t |h(s)||da(s)|$$

and the function $h \cdot a : [0,T] \to \mathbb{R}$ is càdlàg and BV with signed measures $h(s)da(s)$, $|h(s)da(s)| = |h(s)||da(s)|$.

Proposition) Let $a : [0,T] \to \mathbb{R}$ be càdlàg and BV. Let $h : [0,T] \to \mathbb{R}$ be *left-continuous* and bounded. Then

$$\int_0^t h(s)da(s) = \lim_{m\to\infty} \sum_{i=1}^{n_m} h(t_{i-1}^{(m)})(a(t_i^{(m)}) - a(t_{i-1}^{(m)})), \quad t \le T$$
$$\int_0^t h(s)|da(s)| = \lim_{m\to\infty} \sum_{i=1}^{n_m} h(t_{i-1}^{(m)})\left|a(t_i^{(m)}) - a(t_{i-1}^{(m)})\right|$$

for a sequence of subdivisions $0 = t_0^{(m)} < \cdots < t_{n_m}^{(m)} = t$ with $\max_i |t_i^{(m)} - t_{i-1}^{(m)}| \to 0$ as $m \to \infty$.

[Note that we have a corresponding version of the proposition if we exchange the left continuity with right continuity. Also we do not lose much from assuming continuity.]

proof) Set $h_m(0) = 0$, $h_m(s) = h(t_{i-1}^{(m)})$ if $s \in [t_{i-1}^{(m)}, t_i^{(m)}]$. Then $h(s) = \lim_{m\to\infty} h_m(s)$ by left-continuity. Hence,

$$\sum_{i=1}^{n_m} h(t_{i-1}^{(m)})(a(t_i^{(m)}) - a(t_{i-1}^{(m)})) = \int_{(0,t]} h_m(s)\mu(ds) \xrightarrow{m\to\infty} \int h(s)\mu(ds)$$

where the first equality follows from $\int_y^x da(s) = a(x) - a(y)$ and the limit follows from dominated convergence theorem.

The claim about $|da(s)|$ is left as an exercise. (For nested subdivisions, proceed as in the proof of the previous proposition.)

If we are dealing with functions $\mathbb{R}_{\geq 0} \to \mathbb{R}$, it is reasonable to define the local notion of bounded variation.

Definition) A function $a : [0, \infty) \to \mathbb{R}$ is of **finite variation (FV)** if $a|_{[0,T]} \in BV([0,T])$ for all $T > 0$.

2 Semimartingales

A semimartingale is sum of a finite variation process and a local martingale. It turns out that semimartingales are stable under stochastic integrals, and hence is going to be at the centre of our study - and the most general object of stochastic integral.

From now on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{\geq 0}, \mathbb{P})$ is a filtered probability space.

Definition) A **càdlàg adapted process** X is a map $\Omega \times [0, \infty) \to \mathbb{R}$ such that

(i) X is *càdlàg, i.e.* $X(\omega, \cdot) \to \mathbb{R}$ is càdlàg for all $\omega \in \Omega$.

(ii) X is *adapted, i.e.* $X_t = X(\cdot, t)$ is $\mathcal{F}_t$-measurable for all t.

Notation : write $X \in \mathcal{F}$ to denote that a random variable X is measurable with respect to the sigma algebra $\mathcal{F}$.

2.1 Finite variation process

Definition)

(i) A càdlàg adapted process A is a **finite variation process** if $A(\omega, \cdot) : [0, \infty) \to \mathbb{R}$ has finite variation for all $\omega \in \Omega$.

(ii) The **total variation process** V associated to a FV process A is $V_t = \int_0^t |dA_s|$.

Fact : The total variation process V of càdlàg adapted process A is also càdlàg adapted and it is also increasing.

proof) That V is càdlàg and increasing follows from the properties in deterministic settings of *Section 1.3.* To show that V is adapted, let $0 = t_0^{(m)} < \cdots < t_{n_m}^{(m)} = t$ be nested sequence of subdivisions of $[0, t]$ with $\lim_{m\to\infty} \max_i |t_i^{(m)} - t_{i-1}^{(m)}| = 0$. Then (*Section 1.3.*)

$$V_t = \lim_{m\to\infty} \sum_{i=1}^{n,} |A_{t_i^{(m)}} - A_{t_{i-1}^{(m)}}|$$

This is in $\mathcal{F}_t$ since A is a adapted. So V is adapted.

(End of proof) $\square$

(25th January, Friday)

Definition) Let A be a *finite variation process* and let H be a process such that $\forall \omega \in \Omega$ and $\forall t \geq 0 :$ $\int_0^t |H_s(\omega)||dA_s(\omega)| < \infty$. Then define a process $(H \cdot A_t)_{t \geq 0}$ by

$$(H \cdot A)_t = \int_0^t H_s dA_s$$

But for $H \cdot A$ to be adapted, we need an extra condition :

Definition) The **predictable (or previsible) σ-algebra** $\mathcal{P}$ is the σ-algebra on $\Omega \times [0,\infty)$ generated by the sets

$$E \times (s,t], \quad E \in \mathcal{F}_s,\ s < t$$

A process $H : \Omega \times [0,\infty) \to \mathbb{R}$ is **predictable** if it is $\mathcal{P}$-measurable.

If a finite variation process is predictable with its integral finite, then $H \cdot A$ is also a finite variation process, in particular adapted (see a proposition below).

Definition) A process $H : \Omega \times [0,\infty) \to \mathbb{R}$ is **simple**, $H \in \mathcal{E}$, if $H(\omega,t) = \sum_{i=1}^n H_{i-1}(\omega) 1_{(t_{i-1},t_i]}(t)$ for bounded random variables $H_{i-1} \in \mathcal{F}_{t_{i-1}}$ and $0 = t_0 < t_1 < \cdots < t_n$.

Fact : *Simple processes and their pointwise limits are predictable.*

Fact : *Adapted left-continuous processes are predictable*

proof) Let H be adapted, left-continuous. Then $H^n \to H$ as $n \to \infty$ where

$$H_t^n = \sum_{i=1}^{2^n n} H_{(i-1)2^{-n}} 1_{((i-1)2^{-n}, i2^{-n}]}(t) \wedge n$$

Since H is adapted, H^n is simple. Thus H is predictable as a pointwise limit.

(End of proof) $\square$

Fact : *Let H be predictable. Then $H_t \in \mathcal{F}_{t^-}$ where $\mathcal{F}_{t^-} = \sigma(\mathcal{F}_s : s < t)$. (See Example Sheet #1)*

Fact : *Let X be adapted càdlàg. Then $X_{t^-} = \lim_{s \to t^-} X_s$ is left-continuous, predictable.*

proof) The fact that X_{t^-} is left-continuous follows from the fact that X is càdlàg. Then X_{t^-} is $\mathcal{F}_{t^-}$-measurable, hence is adpated. It follows that X_{t^-} is predictable from a previous fact.

Examples

- Brownian motion is predictable since continuous.
- A Poisson process (N_t) is *not* predictable since $N_t \notin \mathcal{F}_{t^-}$. More generally, jump processes are not predictable.

Proposition) Let A be a finite variation process, and let H be a predictable process such that $\int_0^t |H_s||dA_s| < \infty$ for all t and ω. Then $H \cdot A$ is also a finite variation process.

proof) By *Section 1.3*, for every $\omega \in \Omega$, $(H \cdot A)(\omega, \cdot)$ is of finite variation and càdlàg. So we only need to show that $H \cdot A$ is adapted.

Consider first $H(\omega,t) = 1_{(u,v]}(t) 1_E(\omega)$, $u < v$, $E \in \mathcal{F}_u$. Then

$$(H \cdot A)(\omega,t) = 1_E(\omega)(A(\omega, t \wedge v) - A(\omega, t \wedge u))$$

Then this is surely in $\mathcal{F}_t$.

We extend this argument using monotone-class-type argument : Let $\Pi = \{E \times (u, v] : E \in \mathcal{F}_u, u < v\} \subset \Omega \times [0, \infty)$. Clearly, Π is a π-system (closed under intersection and nonempty), and Π generates the predictable σ-algebra by definition. Now let $\mathcal{V} = \{H : \Omega \times [0, \infty) : H \cdot A \text{ is adapted}\}$. Then $1 \in \mathcal{V}$, $1_H \in \mathcal{V}$ for $H \in \Pi$, and if $0 \le H_n \in \mathcal{V}$ with $H_n \nearrow H$ then $H \in \mathcal{V}$ since measurability is closed under pointwise limits. Thus $\mathcal{V}$ is a *monotone class*. By the *measure class theorem(or monotone class theorem)*, $\mathcal{V}$ contains all bounded predictable processes.

The general case follows by approximating H by bounded H^n with $|H^n| \le |H|$.

(End of proof) □

2.2 Local Martingale

From now on, we assume that $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$ satisfies the *usual conditions*(see *Advanced Probability*).

- $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets.
- $(\mathcal{F}_t)$ is right-continuous, *i.e.* $\mathcal{F}_t = \mathcal{F}_{t^+} = \bigcap_{s>t} \mathcal{F}_s$ for all $t \ge 0$.

Theorem) *(Optional Stopping Theorem, OST)* Let X be a càdlàg adapted integrable process. Then the following are equivalent :

(i) X is a martingale : $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ a.s. for all $t \ge s$.

(ii) for all stopping times T, S with T bounded, one has $X_T \in L^1$ and

$$\mathbb{E}[X_T|\mathcal{F}_S] = X_{T\wedge S} \quad \text{a.s.}$$

(iii) for all stopping times T, the process X^T where $X_t^T = X_{t\wedge T}$ is a martingale.

(iv) for all bounded stopping times T, one has $X_T \in L^1$ and $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ a.s.

For X uniformly integrable, (iii) and (iv) hold for all (unbounded) stopping times T.

The following definition does not look very bad and just technical, but turns out to be very useful. Most stochastic processes we will be dealing with fall in to the class of local martingale :

Definition) *A* càdlàg adapted process X is a **local martingale** if there are stopping times T_n such that $T_n \nearrow \infty$ as $n \to \infty$ and X^{T_n} is a martingale for every n. The sequence $(T_n)_n$ is said to **reduce** X.

Example :

(i) Every martingale is a local martingale (Take $T_n = n$ and use OST).

(ii) Let (B_t) be a standard Brownian motion on $\mathbb{R}^3$. Then $(X_t)_{t\ge 1} = (1/|B_t|)_{t\ge 1}$ is a local martingale, but not a martingale.

: First X cannot be a martingale since (see *Advanced Probability*)

$$\sup_{t\ge 1} \mathbb{E}X_t^2 < \infty, \quad \mathbb{E}X_t \to 0$$

To see that it is a local martingale, recall that for $f \in C_b^2(\mathbb{R}^3)$,

$$f(B_t) - f(B_0) - \frac{1}{2}\int_0^t \triangle f(B_s)ds =: M_s^f$$

is a martingale. We would like to choose $f(x) = \frac{1}{|x|}$ so that $X_t = f(B_t)$, but f is not bounded at 0. Choose $f_n \in C_b^2$ such that $f_n(x) = \frac{1}{|x|}$ for $|x| > \frac{1}{n}$. Let $T_n = \inf\{t \geq 1 : |B_t| < \frac{1}{n}\}$. Then

$$X_t^{T_n} - X_1^{T_n} = f_0(B_{t\wedge T_n}) - f_n(B_1) = M_{t\wedge T_n}^{f_n}$$

since $\triangle f_n = \triangle(1/|x|)$ for $|x| > 1/n$. So X^{T_n} is a martingale.

To conclude that X is a local martingale, it only remains to check that $T_n \to \infty$ a.s. Let $S_m = \inf\{t \geq 1 : |B_n| > m\}$. Since X^{T_n} is a bounded martingale (bounded by n), we see from *Optional Stopping Theorem* that

$$\mathbb{E}X_{T_n\wedge S_m} = \mathbb{E}X_{S_m}^{T_n} = \mathbb{E}X_1 < \infty$$

But also

$$\begin{aligned} &\mathbb{E}X_{T_n\wedge S_m} = n\mathbb{P}[T_n < S_m] + \frac{1}{m}\mathbb{P}[T_n \geq S_m] = \Big(n - \frac{1}{m}\Big)\mathbb{P}[T_n < S_m] + \frac{1}{m} \\ \Rightarrow \ &\mathbb{P}[T_n < \infty] = \lim_{m\to}\mathbb{P}[T_n M < S_m] = \frac{\mathbb{E}X_1}{n} \\ \Rightarrow \ &\mathbb{P}[\lim_n T_n < \infty] = 0 \end{aligned}$$

(28th January, Monday)

Proposition) Let X be a *local martingale* and $X_t \geq 0$ for all $t \geq 0$. Then X is a supermartingale.

proof) Let (T_n) be a reducing sequence for X. Then

$$\begin{aligned} \mathbb{E}[X_t|\mathcal{F}_s] = \mathbb{E}[\lim_{n\to\infty} X_{t\wedge T_n}|\mathcal{F}_s] &\leq \liminf_{n\to\infty}\mathbb{E}[X_{t\wedge T_n}|\mathcal{F}_s] \quad \text{by } \textit{conditional Fatou} \\ &= \liminf_{n\to\infty} X_{s\wedge T_n} = X_s \quad \text{a.s.} \end{aligned}$$

(End of proof) $\square$

There are also instances where local martingales become true martingales. There are in fact a more general settings in which local martingales are true martingales, but we just state a simple version that we are going to use in practice.

Proposition) Let X be a local martingale and suppose that there is $Z \in L^1$ such that $|X_t| \leq Z$ for all $t \geq 0$. Then X is a martingale. In particular, bounded local martingales are martingales.

proof) Let (T_n) be a reducing sequence for X. Let S be a stopping time. Then

$$\mathbb{E}X_0 = \mathbb{E}X_0^{T_n} = \mathbb{E}X_S^{T_n} = \mathbb{E}X_{T_n\wedge S}$$

since $|X_{T_n\wedge S}| \leq Z$, we have $X_{T_n\wedge S} \to X_S$ in L^1 as $n \to \infty$, so $\mathbb{E}X_0 = \mathbb{E}X_S$ for any stopping time S, which is one of the equivalent conditions for X to be a martingale (by *Optional Stopping Theorem*).

(End of proof) $\square$

Fact : Let X be a *continuous adapted process* with $X_0 = 0$. Then

$$S_n = \inf\{t \geq 0 : |X_t| = n\}$$

are stopping times and $S_n \nearrow \infty$ as $n \to \infty$.

proof) $(S_n)_{n\in\mathbb{N}}$ are stopping times since

$$\{S_n \leq t\} = \{\sup_{s\leq t} |X_s| \geq n\} = \bigcap_{k\in\mathbb{N}} \bigcup_{s\leq t, s\in\mathbb{Q}} \{|X_s| > n - \frac{1}{k}\} \in \mathcal{F}_t$$

were the second equality comes from continuity of X. Also, $S_n \nearrow \infty$ since, for every $\omega \in \Omega$, $|X_s|$ is bounded on any bounded interval of s (by continuity).

(End of proof) □

Proposition) Let X be a continuous local martingale with $X_0 = 0$. Then the sequence (S_n) defined above reduces X.

proof) Let (T_k) be a reducing sequence for X. By OST, $X^{T_k \wedge S_n}$ is a martingale, so X^{S_n} is also a local martingale. But $|X^{S_n}| \leq n$, so X^{S_n} is a bounded local martingale for each n, hence X^{S_n} is a martingale. So (S_n) reduces X.

(End of proof) □

Theorem) Let X be a *continuous local martingale* with $X_0 = 0$. If X is also a *finite variation process*, then $X_t = 0$ for all $t \geq 0$ a.s.

proof) Let $S_n = \inf\{t \geq 0 : \int_0^t |dX_s| = n\}$. Call $v_t = \int_0^t |dX_s|$, the total variation process of X. Note S_n is a stopping time, $S_n \nearrow \infty$, and X^{S_n} is a local martingale (by *Optional Stopping Theorem*, stopped local martingale is also local martingale). Also X^{S_n} is bounded since

$$|X_t^{S_n}| \leq \int_0^{t\wedge S_n} |dX_s| \leq n$$

and therefore X^{S_n} is a martingale using the previous proposition. Hence (S_n) reduces X.

Let $0 = t_0 < t_1 < \cdots < t_n = t$ be a subdivision of $[0, t]$. Then, since X^{S_n} is a martingale,

$$\begin{aligned}
\mathbb{E}(X_t^{S_n})^2 &= \sum_{i=1}^{k} \mathbb{E}\Big((X_{t_i}^{S_n} - X_{t_{i-1}}^{S_n})^2\Big) \quad \text{(cross terms vanish)} \\
&\leq \mathbb{E}\Big(\max_i |X_{t_i}^{S_n} - X_{t_{i-1}}^{S_n}| \sum_{i=1}^{k} |X_{t_i}^{S_n} - X_{t_{i-1}}^{S_n}|\Big) \\
&\to 0 \quad \text{as } \max_i |t_i - t_{i-1}| \to 0
\end{aligned}$$

by continuity of X and since $\sum_{i=1}^k |X_{t_i}^{S_n} - X_{t_{i-1}}^{S_n}| \leq \int_0^{t\wedge S_n} |dX_s| \leq n$. Hence $\mathbb{E}(X_t^{S_n})^2 = 0$ and in particular

$$X_{t\wedge S_n} = 0 \quad \text{a.s. } t \geq 0$$

so $X_t = 0$ a.s. for all $t \geq 0$, and therefore $X_t = 0$ for all $t \geq 0$ a.s.

(End of proof) □

(30th January, Wednesday)

2.3 L^2 bounded martingales

We now turn our attention to a specific class of martingales. L^2-bounded martingales form a *Hilbert space*, and therefore we can use various structural properties of Hilbert spaces to develop our theory.

Definition) Let

$$M^2 = \{X : \Omega \times [0,\infty) \to \mathbb{R} : X \text{ is a càdlàg martingale, } \sup_{t\geq 0} \mathbb{E}[X_t^2] < \infty\}/\sim$$

$$M_c^2 = \{X \in M^2 : X(\omega,\cdot) \text{ is continuous for every } \omega \in \Omega\}/\sim$$

where $\sim$ means that indistinguishable processes are identified. Moreover, set

$$\left\|X\right\|_{M^2} = \left(\sup_t \mathbb{E}[X_t^2]\right)^{1/2} = \left(\mathbb{E}[X_\infty^2]\right)^{1/2}$$

Here recall that if $X \in M^2$ then

- $X_t \to X_\infty$ a.s. and in L^2.
- (X_t^2) is a submartingales (by Jensen's inequality), so $t \mapsto \mathbb{E}X_t^2$ is increasing, so $\mathbb{E}[X_\infty^2] = \sup_t \mathbb{E}[X_t^2]$.
- Doob's L^2 inequality implies $\mathbb{E}[\sup_t X_t^2] \leq 4\mathbb{E}[X_\infty^2]$.

In particular, $\left\|X\right\|_{M^2} = 0$ implies $X = 0$. This makes $\left\|\cdot\right\|_{M^2}$ a norm (the other properties are clear).

In fact, $(X,Y)_{M^2} = \mathbb{E}[X_\infty Y_\infty]$ is an inner product on M^2 that induces $\left\|\cdot\right\|_{M^2}$ - to see this, it is sufficient to have parallelogram identity, but this is clear as $\mathbb{E}[(A_\infty + B_\infty)^2] + \mathbb{E}[(A_\infty - B_\infty)^2] = 2\mathbb{E}[(A_\infty)^2] + 2\mathbb{E}[(B_\infty)^2]$.

Proposition) M^2 is a *Hilbert space* and M_c^2 is a closed subspace.

proof) We need to prove that M^2 is complete. Thus let $(X^n) \subset M^2$ be a Cauchy sequence, *i.e.* $\mathbb{E}[X_\infty^n - X_\infty^m]^2 \to 0$ as $n,m \to \infty$. By passing to a subsequence, we may assume that $\mathbb{E}[X_\infty^n - X_\infty^{n-1}]^2 \leq 2^{-n}$ and it suffices to show that the subsequence converges to see that the original sequence converges. To this end,

$$\begin{aligned}\mathbb{E}\Big[\sum_{n=1}^\infty \sup_{t\geq 0} |X_t^n - X_t^{n-1}|\Big] &\leq \sum_{n=1}^\infty \mathbb{E}\Big[\sup_{t\geq 0} |X_t^n - X_t^{n-1}|^2\Big]^{1/2} \quad \text{(Cauchy-Schwarz)}\\ &\leq \sum_{n=1}^\infty 2\mathbb{E}\Big[X_\infty^n - X_\infty^{n-1}\Big]^{1/2} \leq \sum_{n=1}^\infty 2\cdot 2^{-n/2} < \infty\end{aligned}$$

so $\sum_{n=1}^\infty \sup_{t\geq 0} |X_t^n - X_t^{n-1}| < \infty$ a.s., and hence (X^n) is a Cauchy sequence in $(D[0,\infty), \|\cdot\|_\infty)$ a.s. So we have that $\left\|X^n - X\right\|_\infty \to 0$ for some $X \in D[0,\infty)$ a.s. Set $X = 0$ outside this almost sure event. Then $X \in D[0,\infty)$ for every $\omega \in \Omega$. We know X is a uniform limit of càdlàg functions hence is càdlàg. It still remains to check that the converging point X is in M^2.

♡ **Claim :** $\mathbb{E}[\sup_{t\geq 0} |X^n - X|^2] \to 0$, and in particular, X is bounded in L^2.

: We have

$$\begin{aligned}\mathbb{E}\Big[\sup_{t\geq 0} |X^n - X|^2\Big] = \mathbb{E}\Big[\lim_{m\to\infty} \sup_t |X_t^n - X_t^m|^2\Big] &\leq \liminf_{m\to\infty} \mathbb{E}\Big[\sup_t |X_t^n - X_t^m|^2\Big] \quad \text{(Fatou)}\\ &\leq \liminf_{m\to} 4\mathbb{E}(X_\infty^n - X_\infty^m) \quad \text{(Doob)}\end{aligned}$$

This just converges to 0 as $(X^n)_n$ was a Cauchy sequence.

♡ **Claim :** X is a martingale.

: We have

$$\left\|\mathbb{E}[X_t|\mathcal{F}_s] - X_s\right\|_{L^2} \leq \left\|\mathbb{E}[X_t - X_t^n|\mathcal{F}_s]\right\|_{L^2} + \left\|X_s^n - X_s\right\|_{L^2}$$
$$\leq \left\|X_t - X_t^n\right\|_{L^2} + \left\|X_s^n - X_s\right\|_{L^2} \leq 2\mathbb{E}[\sup_{s\geq 0}|X_s^n - X_s|^2]^{1/2} \to 0$$

where the second inequality follows from Jensen's inequality.

Thus $X \in M^2$ and $X^n \to X$ in M^2, so we have shown that M^2 is complete.

Clearly, M_c^2 is a subspace. It is a complete (thus closed) by exactly the same argument with $D[0,\infty)$ replaced by $C[0,\infty)$.

(End of proof) □

2.4 Quadratic Variation

The quadratic variation of a process provides a measure of the rate of martingale "oscillation", or "diffusion".

Definition) For a sequence of processes (X^n) and a process X, we say $X^n \to X$ **ucp (uniformly on compact intervals in probability)** *iff*

$$\mathbb{P}\left[\sup_{s\in[0,t]}|X_s^n - X_s| > \epsilon\right] \xrightarrow{n\to 0} 0 \quad \forall t > 0, \forall \epsilon > 0$$

Theorem) Let M be a *continuous* local martingale. Then there exists a unique (up to indistinguishability) *continuous adapted increasing process* $\langle M\rangle = \left(\langle M\rangle_t\right)_t$ such that (is uniquely characterized by) $\langle M\rangle_0 = 0$ and $M^2 - \langle M\rangle$ is a continuous local martingale.

Moreover, with $0 = t_0^m < t_1^m < \cdots$ given by $t_i^m = 2^{-m}i$,

$$\langle M\rangle_t^{(m)} \xrightarrow{\text{ucp}} \langle M\rangle_t \quad \text{where } \langle M\rangle_t^{(m)} = \sum_{i=1}^{\lfloor 2^m t\rfloor}(M_{t_i} - M_{t_{i-1}})^2$$

[In fact, the convergence is true for all locally finite subdivision of $[0,\infty)$ *with* $\max_i|t_i^m - t_{i-1}^m| \to 0$ *as* $m \to \infty$*.]*

Definition) $\langle M\rangle$ as in the theorem is the **quadratic variation** of M.

Example : Let B be a standard Brownian motion. Then $B_t^2 - t$ is a martingale, so $\langle B\rangle_t = t$, since the quadratic variation is uniquely characterized by the fact that $M^2 - \langle M\rangle$ is a martingale and $\langle M\rangle_0 = 0$.

proof of Theorem) By replacing M_t by $M_t - M_0$, it suffices to assume that $M_0 = 0$.

♡ **Uniqueness**

: Suppose that (A_t) and (B_t) both obey the conditions required for $\langle M\rangle$. Then $A_t - B_t = (M_t^2 - B_t) - (M_t^2 - A_t)$. Here, $A_t - B_t$ is of finite variation since is a difference of two continuous increasing processes, and $(M_t^2 - B_t), (M_t^2 - A_t)$ are continuous local martingales. So this shows $A - B = 0$ a.s by an earlier result on local martingales (the last theorem of **Section 2.2**).

The rest of the theorem would be proved after proving a lemma.

(1st February, Friday)

Let us first prove the theorem in the reduced setting of (uniformly) bounded martingales.

Lemma) *(bounded case)* The theorem is true under the additional assumption $|M_t| \le C$ for all (ω, t), $M_t = M_{t\wedge T}$ for C, T deterministic constants.

proof) Let X be an another martingale defined by

$$X_t^m = \sum_{i=1}^{\lfloor 2^m T\rfloor} M_{(i-1)2^{-m}}(M_{i2^{-m}\wedge t} - M_{(i-1)2^{-m}\wedge t})$$

Then X^m is bounded martingale.

♡ **Claim :** With $\langle M\rangle^{(m)}$ defined as in the theorem, $\langle M\rangle^{(m)}_{k2^{-m}} = M^2_{k2^{-m}} - 2X^m_{k2^{-m}}$

: Just see that

$$\begin{aligned}\langle M\rangle^{(m)}_{k2^{-m}} =& \sum_{i=1}^{k}(M_{i2^{-m}} - M_{(i-1)2^{-m}})^2 \\ =& \sum_{i=1}^{k}(M^2_{i2^{-m}} - M^2_{(i-1)2^{-m}}) - 2M_{(i-1)2^{-m}}(M_{i2^{-m}} - M_{(i-1)2^{-m}}) \\ =& -2X^m_{k2^{-m}} + \sum_{i=1}^{k}(M^2_{i2^{-m}} - M^2_{(i-1)2^{-m}}) = M^2_{k2^{-m}-2X^m_{k2^{-m}}}\end{aligned}$$

Also observe that $\langle M\rangle^{(m)}_t$ is increasing on $t \in \{2^{-m}i : i \in \mathbb{N}\}$ (from its definition).

♡ **Claim :** $(X^m) \subset M_c^2$ is a Cauchy sequence.

: For $m' > m$,

$$X^{m'}_\infty - X^m_\infty = \sum_{i=1}^{\lfloor 2^{m'}T\rfloor} \Big(M_{(i-1)2^{-m'}} - M_{\lfloor (i-1)2^{m-m'}\rfloor 2^{-m}}\Big)\Big(M_{i2^{-m'}} - M_{(i-1)2^{-m'}}\Big)$$

and by martingale property, cross terms are independent, so

$$\begin{aligned}\mathbb{E}\Big[(X^{m'}_\infty - X^m_\infty)^2\Big] =& \sum_{i=1}^{\lfloor 2^{m'}T\rfloor} \mathbb{E}\Big[(M_{(i-1)2^{-m'}} - M_{\lfloor (i-1)2^{m-m'}\rfloor 2^{-m}})^2(M_{i2^{-m'}} - M_{(i-1)2^{-m'}})^2\Big] \\ \le& \sup_{|s-t|\le 2^{-m}} |M_t - M_s|^2 \\ \le& \mathbb{E}\Big(\sup_{|t-s|\le 2^{-m}} |M_t - M_s|^2 \sum_{i=1}^{\lfloor 2^{m'}T\rfloor} (M_{i2^{-m'}} - M_{(i-1)2^{-m'}})^2\Big) \\ \le& \mathbb{E}\Big(\sup_{|s-t|\le 2^{-m}} |M_t - M_s|^4\Big)^{1/2}\mathbb{E}\Big((\langle M\rangle^{(m')}_T)^2\Big)^{1/2} \quad \text{(Cauchy-Schwarz)}\end{aligned}$$

Here, we have $|M_t - M_s|^4 \le (2C)^4$ and $\sup_{|t-s|\le 2^{-m}} |M_t - M_s| \to 0$ by uniform continuity of M on $[0,T]$, so $\mathbb{E}(\sup_{|t-s|\le 2^{-m}} |M_t - M_s|^4) \to \infty$ as $m \to \infty$ by *dominated convergence theorem.*

Also, the second term $\mathbb{E}\Big((\langle M\rangle_T^{(m')})^2\Big)$ can be bounded. Indeed,

$$
\begin{aligned}
&\mathbb{E}\Big[\Big(\sum_{i=1}^n (M_{i2^{-m}} - M_{(i-1)2^{-m}})^2\Big)^2\Big] \\
&= \sum_{i=1}^n \mathbb{E}\Big((M_{i2^{-m}} - M_{(i-1)2^{-m}})^4\Big) + 2\sum_{i=1}^n \mathbb{E}((M_{i2^{-m}} - M_{(i-1)2^{-m}})^2 \sum_{k=i+1}^n (M_{k2^{-m}} - M_{(k-1)2^{-m}})^2) \\
&= \sum_{i=1}^n \mathbb{E}\Big((M_{i2^{-m}} - M_{(i-1)2^{-m}})^4\Big) + 2\sum_{i=1}^n \mathbb{E}((M_{i2^{-m}} - M_{(i-1)2^{-m}})^2 (M_{n2^{-m}} - M_{i2^{-m}})^2) \\
&\le (2C)^2 \sum_{i=1}^n \mathbb{E}\Big((M_{i2^{-m}} - M_{(i-1)2^{-m}})^2\Big) + 2\sum_{i=1}^n \mathbb{E}((M_{i2^{-m}} - M_{(i-1)2^{-m}})^2 (2C)^2) \\
&\le 12C^2 \sum_{i=1}^n \mathbb{E}((M_{i2^{-m}} - M_{(i-1)2^{-m}})^2) \\
&= 12C^2 \mathbb{E}(M_{n2^{-m}} - M_0)^2 \le 12C^4
\end{aligned}
$$

Putting these altogether, we conclude that $\mathbb{E}\Big((X_\infty^{m'} - X_\infty^m)^2\Big) \to 0$ as $m, m' \to \infty$. Hence (X^m) is a Cauchy sequence in M_c^2.

As $(X^m) \subset M_c^2$ is a Cauchy sequence, we see that there exists $X \in M_c^2$ such that $X^m \to X$ in M_c^2. Since $X^m \to X$ in M_c^2, in particular $\Big\|\sup_t |X_t^m - X_t|\Big\|_{L^2} \to 0$, and hence $\sup_t |X_t^{m_l} - X_t| \to 0$ a.s. for some subsequence $(m_l)_l$. So set

$$
\langle M\rangle_t = \begin{cases} M_t^2 - 2X_t & \text{on the event of probability 1 on which the a.s. convergence holds} \\ 0 & \text{on the complementary of the event} \end{cases}
$$

Then $\langle M\rangle$ is continuous and adapted since M and X are. So $M^2 - \langle M\rangle - 2X$ is a martingale since X is. Also $\langle M\rangle$ is increasing since $M^2 - 2X^m$ is increasing on $\{i2^{-m} : i \in \mathbb{N}\}$ and the convergence $M^2 - 2X^m \to \langle M\rangle$ is uniform almost surely.

But to check the summation formula for $\langle M\rangle$, we still have to prove $\langle M\rangle^{(m)} \to \langle M\rangle$.

♡ **Claim :** $\langle M\rangle^{(m)} \to \langle M\rangle$ ucp.

: To see this, recall $\langle M\rangle_t^{(m)} = \langle M\rangle_{\lfloor 2^m t\rfloor 2^{-m}}^{(m)} = M^2_{\lfloor 2^m t\rfloor 2^{-m}} - 2X^m_{\lfloor 2^m t\rfloor 2^{-m}}$, and $\sup_t |X_t^m - X_t| \to 0$ in L^2 and thus in probability. Hence,

$$
\begin{aligned}
\sup_t |\langle M\rangle_t - \langle M\rangle_t^{(m)}| \le & \sup_t |M^2_{\lfloor 2^m t\rfloor 2^{-m}} - M_t^2| + \sup_t |X_{\lfloor 2^m t\rfloor 2^{-m}} - X_t| \\
& + \sup_t |X^m_{\lfloor 2^m t\rfloor 2^{-m}} - X_{\lfloor 2^m t\rfloor 2^{-m}}|
\end{aligned}
$$

but each term converges to 0 in probability - two first terms converges a.s. by uniform continuity and the third one by convergence in L^2.

(End of proof) □

(4th February, Monday)

Last time :

Theorem) Let M be a continuous local martingale. Then there exists a unique continuous adapted increasing process $\langle M\rangle$ such that $\langle M\rangle_0 = 0$ and $M^2 - \langle M\rangle$ is a continuous local martingale. Also,

$$\langle M\rangle_t^{(m)} = \sum_{i=1}^{\lfloor 2^n t\rfloor} (M_{2^{-m}t} - M_{2^{-m}(i-1)})^2 \xrightarrow{ucp} \langle M\rangle_t$$

We proved this theorem for the case $|M| \le C$, $M_{t\wedge T} = M_t$ for deterministic constants C, T.

Lemma) Suppose M is a continuous local martingale for which $\langle M\rangle$ exists. Let T be a stopping time. Then $\langle M^T\rangle$ exists and is given by $\langle M^T\rangle_t = \langle M\rangle_{t\wedge T}$ (up to indistinguishability).

proof) Since $M_t^2 - \langle M\rangle_t$ is a continuous local martingale, so by *Optional Stopping Theorem*, $M_{t\wedge T}^2 - \langle M\rangle_{t\wedge T} = (M^T)_t^2 - \langle M\rangle_{t\wedge T}$ also is. By uniqueness of quadratic variation process, we have $\langle M^T\rangle = \langle M\rangle_{t\wedge T}$.

(End of proof) $\square$

Now we prove the main theorem of the section.

proof of the theorem, continued) Let M be a continuous local martingale. We have assumed, without losing generality, that $M_0 = 0$. Let

$$T_n = \inf\{t \ge 0 : |M_t| \ge n\}, \quad S_n = T_n \wedge n$$

Then $S_n \nearrow \infty$ and M^{S_n} is a bounded martingale such that $(M^{S^n})_{t\wedge n} = (M^{S_n})_t$. By the special case proved last time, $\langle M^{S_n}\rangle$ exists.

By the previous lemma,

$$\langle M^{S_n}\rangle_t = \langle M^{S_{n+1}}\rangle_{t\wedge S_n}$$

Thus there is a continuous process $\langle M\rangle$ such that $\langle M\rangle_{t\wedge S_n}$ and $\langle M^{S_n}\rangle$ are indistinguishable for all $n \in \mathbb{N}$.

Clearly, $\langle M\rangle$ is increasing since the $\langle M^{S_n}\rangle$ are, and $(M^2 - \langle M\rangle)^{S_n}$ is a martingale for all n. Since $S_n \nearrow \infty$, thus $M^2 - \langle M\rangle$ is a local martingale.

♡ **Claim :** $\langle M\rangle^{(m)} \xrightarrow{ucp} \langle M\rangle$.

: by the bounded case, $\langle M^{S_n}\rangle^{(m)} \xrightarrow{ucp} \langle M^{S_n}\rangle$ as $m \to \infty$ for every n. Hence

$$\mathbb{P}\Big[\sup_{t\le T}\Big|\langle M\rangle_t^{(m)} - \langle M\rangle_t\Big| > \epsilon\Big] \le \mathbb{P}(S_n < T) + \mathbb{P}\Big[\sup_{t\le T}\Big|\langle M^{S_n}\rangle_t^{(m)} - \langle M^{S_n}\rangle_t\Big| > \epsilon\Big]$$
$$\to 0 \quad \text{as } n, m \to \infty$$

So we are done.

(End of proof) $\square$

Fact : Let M be a continuous local martingale with $M_0 = 0$. Then $M \equiv 0$ *iff* $\langle M\rangle = 0$.

proof) If $\langle M\rangle = 0$ then $M^2 - \langle M\rangle = M^2$ is a continuous local martinglae and non-negative, hence is a supermartingale. Thus $\mathbb{E}[M_t^2] \le \mathbb{E}[M_0^2] = 0$ for all t. This implies $M \equiv 0$.

(End of proof) $\square$

Proposition) Let $M \in M_c^2$ with $M_0 = 0$. Then $M^2 - \langle M \rangle$ is a *uniformly integrable* martingale and

$$\left\|M\right\|_{M^2} = \Big(\mathbb{E}\big[\langle M \rangle_\infty\big]\Big)^{1/2}$$

In particular, the norm only depends on the quadratic variation.

proof) We will show that $\langle M \rangle_\infty \in L^1$. Once we show this,

$$|M_t^2 - \langle M \rangle_t| \leq \sup_{s \geq 0} M_s^2 + \langle M \rangle_\infty \quad \forall t \geq 0$$

$\sup_{s\geq 0} M_s^2 + \langle M \rangle_\infty =: Z$ is in L^1 by Doob's L^2-inequality and assumptioin for $\langle M \rangle_\infty$. Thus $M^2 - \langle M \rangle$ is a continuous local martingale bounded by $Z \in L^1$. In particular, $M^2 - \langle M \rangle$ is a uniformly integrable martignale.

♡ **Claim :** $\langle M \rangle_\infty \in L^1$.

: Let $S_n = \inf\{t \geq 0 : \langle M \rangle_t > n\}$. Then $S_n \nearrow \infty$, S_n is a stopping time, and $\langle M \rangle_{t \wedge S_n} \leq n$. So

$$M_{t\wedge S_n}^2 - \langle M \rangle_{t \wedge S_n} \leq \sup_{s \geq 0} M_s^2 + n \in L^1 \quad \text{(by Doob's inequality)}$$

so $M_{t\wedge S_n}^2 - \langle M \rangle_{t \wedge S_n}$ is a *true* martingale. Hence

$$\mathbb{E}[M_{t\wedge S_n}^2] = \mathbb{E}\langle M \rangle_{t\wedge S_n}$$

Taking $t \to \infty$, has

$$\mathbb{E}[M_{t\wedge S_n}^2] \to \mathbb{E}[M_{S_n}^2] \quad \text{by Dominated convergence, as } \mathbb{E}[\sup_t M_t^2] < \infty$$
$$\mathbb{E}\langle M \rangle_{t\wedge S_n} \to \mathbb{E}\langle M \rangle_{S_n} \quad \text{by Dominated convergence, as } \sup_t \langle M \rangle_{t\wedge S_n} \leq n$$

Taking $n \to \infty$, has

$$\mathbb{E}M_{S_n}^2 \to \mathbb{E}M_\infty^2 \quad \text{by Dominated convergence}$$
$$\mathbb{E}\langle M \rangle_{S_n} \to \mathbb{E}\langle M \rangle_\infty \quad \text{by monotone convergence}$$

hence $\mathbb{E}\langle M \rangle_\infty = \mathbb{E}M_\infty^2 = \left\|M\right\|_{M^2} < \infty$.

(End of proof) □

2.5 Covariation

Definition) For M and N continuous local martingales, define,

$$\langle M, N \rangle = \frac{1}{4}\Big(\langle M + N \rangle - \langle M - N \rangle\Big)$$

The process $\langle M, N \rangle = (\langle M, N \rangle_t)_t$ is the **covariation** or bracket of M and N.

Proposition)

(i) $\langle M, N \rangle$ is the unique (up to indistinguishability) finite variation process such that $MN - \langle M, N \rangle$ is a continuous local martingale.

(ii) We have $\langle M,N\rangle_t^{(m)} \xrightarrow{ucp} \langle M,N\rangle_t$ where

$$\langle M,N\rangle_t^{(m)} = \sum_{i=1}^{\lfloor 2^m t\rfloor} (M_{i2^{-m}} - M_{(i-1)2^{-m}})(N_{i2^{-m}} - N_{(i-1)2^{-m}})$$

(iii) The mapping $M, N \mapsto \langle M,N\rangle$ is bilinear and symmetric.

(iv) For every stopping time T, $\langle M^T, N^T\rangle_t = \langle M^T, N\rangle_t = \langle M,N\rangle_{T\wedge t}$.

(v) If $M, N \in M_c^2$ with $M_0 = N_0 = 0$, then $M_T N_t - \langle M,N\rangle$ is a uniformly integrable martingale and

$$(M,N)_{M^2} = \mathbb{E}\langle M,N\rangle_\infty$$

proof) Prove exactly as for $M = N$.

Example : Let B and B' be independent Brownian motions (adapted with respect to the same filtration). Then BB' is a martingale (by independence), so $\langle B, B'\rangle = 0$.

Let $B'' = \rho B + \sqrt{1-\rho^2}B'$ for some $\rho \in [0,1]$. Then B'' is also a Brownian motion, and by bilinearity,

$$\langle B, B''\rangle_t = \rho\langle B,B\rangle_t + \sqrt{1-\rho^2}\langle B,B'\rangle_t = \rho t$$

(6th February, Wednesday)

Proposition) *(Kunita-Watanabe inequality)* Let M and N be continuous local martingale,s and let H and K be measurable processes. Then a.s.

$$\int_0^\infty |H_s||K_s||d\langle M,N\rangle_s| \le \Big(\int_0^\infty |H_s|^2 d\langle M\rangle_s\Big)^{1/2}\Big(\int_0^\infty |K_s|^2 d\langle N\rangle_s\Big)^{1/2} \quad \cdots\cdots\cdots \text{ (KW)}$$

proof) The proof is just done by applying Cauchy-Schwarz inequality multiple times. Write $\langle M,N\rangle_s^t = \langle M,N\rangle_t - \langle M,N\rangle_s$.

♡ **Claim 1 :** For all $0 \le s < t$, $|\langle M,N\rangle_s^t| \le \sqrt{\langle M,M\rangle_s^t}\sqrt{\langle N,N\rangle_s^t} \quad \cdots\cdots\cdots \ (\star)$.

: By continuity, we can assume that s and t are dyadic rationals. Then indeed,

$$\begin{aligned}
|\langle M,N\rangle_s^t| &= \lim_{n\to\infty}\Big|\sum_{i=2^n s+1}^{2^n t}(M_{2^{-n}i} - M_{2^{-n}(i-1)})(N_{2^{-n}i} - N_{2^{-n}(i-1)})\Big| \\
&\le \lim_{n\to\infty}\Big(\sum_{i=2^n s+1}^{2^n t}(M_{2^{-n}i} - M_{2^{-n}(i-1)})^2\Big)^{1/2}\Big(\sum_{i=2^n s+1}^{2^n t}(N_{2^{-n}i} - N_{2^{-n}(i-1)})^2\Big)^{1/2} \quad \text{(CS)} \\
&= \sqrt{\langle M,M\rangle_s^t}\sqrt{\langle N,N\rangle_s^t}
\end{aligned}$$

Now fix an event such that $(\star)$ holds for all $s < t$ (holds for rationals and thus for all irrationals as well).

♡ **Claim 2 :** $\int_s^t |d\langle M,N\rangle| \le \sqrt{\langle M,M\rangle_s^t}\sqrt{\langle N,N\rangle_s^t}$.

: Indeed, for any subdivision $s = t_0 < \cdots < t_n = t$,

$$\begin{aligned}\sum_{i=1}^{n} \left|\langle M, N\rangle_{t_{i-1}}^{t_i}\right| &\le \sum_{i=1}^{n} \sqrt{\langle M, M\rangle_{t_{i-1}}^{t_i}}\sqrt{\langle N, N\rangle_{t_{i-1}}^{t_i}} \quad \text{(Claim 1)} \\ &\le \Big(\sum_{i=1}^{n}\langle M, M\rangle_{t_{i-1}}^{t_i}\Big)^{1/2}\Big(\sum_{i=1}^{n}\langle N, N\rangle_{t_{i-1}}^{t_i}\Big)^{1/2} \\ &= \sqrt{\langle M, M\rangle_s^t}\sqrt{\langle N, N\rangle_s^t}\end{aligned}$$

The cliam follows by taking the supremum over all subdivisions.

♡ **Claim 3 :** For all bounded Borel sets $B \subset [0, \infty)$, $\int_B \left|d\langle M, N\rangle_s\right| \le \sqrt{\int_B |d\langle M\rangle_s|}\sqrt{\int_B |d\langle N\rangle_s|}$.

: For B a finite union of intervals, this agian follows from Cauchy-Schwarz as above. Also the inequality is preserved under taking disjoint union of Borel sets, and therefore by Dynkin's lemma, we seet that the inequality holds for any Borel set.

♡ **Claim 4 :** (KW) holds for all simple functions $H = \sum_{l=1}^{L} h_l 1_{B_l}$ and $K = \sum_{l=1}^{L} k_l 1_{B_l}$ where the B_l are bounded disjoint Borel sets.

: just do computation,

$$\begin{aligned}\int_0^\infty |H_s K_s||d\langle M, N\rangle_s| &\le \sum_l |h_l k_l| \int_{B_l} |d\langle M, N\rangle_s| \\ &\le \sum_l |h_l k_l| \Big(\int_{B_l} d\langle M\rangle_s\Big)^{1/2}\Big(\int_{B_l} d\langle N\rangle_s\Big)^{1/2} \\ &\le \Big(\sum_l h_l^2 \int_{B_l} d\langle M\rangle_s\Big)^{1/2}\Big(\sum_l k_l^2 \int_{B_l} d\langle N\rangle_s\Big)^{1/2} \\ &= \Big(\int H_s^2 d\langle M\rangle_s\Big)^{1/2}\Big(\int K_s^2 d\langle N\rangle_s\Big)^{1/2}\end{aligned}$$

Finally, having the claim, approximate general H, K by simple functions as above and use monotone class argument to complete the proof.

(End of proof) □

2.6 Semimartingales

Definition) A **(continuous) semimartingale** is a (continuous) adapted process X such that

$$X = X_0 + M + A$$

with $X_0 \in \mathcal{F}_0$, M is a (continuous) local martingale with $M_0 = 0$ and A is a finite variation process with $A_0 = 0$.

Remark : The decomposition is unique (up to indistinguishability).

Definition) Let $X = X_0 + M + A$ and $X' = X_0' + M_0' + A_0'$ be continuous semimartingales. Set

$$\langle X\rangle = \langle M\rangle, \quad \langle X, X'\rangle = \langle M, M'\rangle$$

Exercise : We agian have limit expression

$$\langle X, Y\rangle_t^{(m)} = \sum_{i=1}^{\lfloor 2^m t\rfloor} (X_{i2^{-m}} - X_{(i-1)2^{-m}})(Y_{i2^{-m}} - Y_{(i-1)2^{-m}}) \xrightarrow{ucp} \langle X, Y\rangle_t$$

3 The Itô integral

3.1 Simple processes

Definition) The space of **simple processes** $\mathcal{E}$ consists of $H : \Omega \times [0, \infty) \to \mathbb{R}$ that can be written as

$$H(\omega, t) = \sum_{i=1}^{n} H_{i-1}(\omega) 1_{(t_{i-1}, t_i]}(t)$$

for bounded random variabes $H_{i-1} \in \mathcal{F}_{t_{i-1}}$ and $0 \le t_0 < \cdots < t_n$.

Definition) For $M \in M_c^2$ and $H \in \mathcal{E}$. Set **(Itô integral for simple processes)**

$$(H \cdot M)_t = \sum_{i=1}^{n} H_{i-1}(M_{t_i \wedge t} - M_{t_{i-1} \wedge t})$$

We also write $\int_0^t H_s dM_s = (H \cdot M)_t$.
[Note, there is no necessity in picking H_{i-1} in the integral. In fact, there are different choices for this. For example, summation over $(H_{i-1} + H_i)/2$ is called the Stratonovich integral]

Proposition) Let $M \in M_c^2$ and $H \in \mathcal{E}$. Then $H \cdot M \in M_c^2$ and

$$\Big\| H \cdot M \Big\|_{M^2}^2 = \mathbb{E}\Big(\int_0^\infty H_s^2 d\langle M \rangle_s \Big) \quad \text{(Itô isometry for simple process)}$$

[Observe that the right hand side makes sense for more general family of martingales - so this proposition plays a crucial role in extending the Itô integral.]

proof)

♡ **Claim :** $H \cdot M$ is a martingale in M_c^2.

Let $X_t^i = H_{i-1}(M_{t_i \wedge t} - M_{t_{i-1} \wedge t})$. Then $H \cdot M = \sum_{i=1}^n X^i$ and it suffices to show that $X^i \in M_c^2$. Indeed,

- for $s \ge t_{i-1}$, $\mathbb{E}(X_t^i | \mathcal{F}_s) = H_{i-1}(\mathbb{E}(M_{t \wedge t_i} | \mathcal{F}_s) - M_{t_{i-1}}) = H_{i-1}(M_{t_i \wedge s} - M_{t_{i-1}}) = X_s^i$
- for $s < t_{i-1}$, $\mathbb{E}(X_t^i | \mathcal{F}_s) = \mathbb{E}(H_{i-1} \mathbb{E}(M_{t \wedge t_i} - M_{t \wedge t_{i-1}} | \mathcal{F}_{t_{i-1}}) | \mathcal{F}_s) = 0 = X_s^i$.

Also, $\Big\| X^i \Big\|_{M^2} \le 2 \Big\| H \Big\|_\infty \Big\| M \Big\|_{M^2} < \infty$, so $X \in M_c^2$.

(8th February, Friday)

Last time : $H \in \mathcal{E} \Leftrightarrow H(\omega, t) = \sum_{i=1}^n H_{i-1} 1_{(t_{i-1} - t_i]}$, $H_{i-1} \in \mathcal{F}_{t_{i-1}}$ bounded. Then

$$\int_0^t H_s dM_s = (H \cdot M)_t = \sum_{i=1}^{n} H_{i-1}(M_{t_i \wedge t} - M_{t_{i-1} \wedge t}) \quad \text{if } H \in \mathcal{E}$$

We let $X_t^i = H_{i-1}(M_{t_i \wedge t} - M_{t_{i-1} \wedge t})$.

Proposition) Let $M \in M_c^2$ and $H \in \mathcal{E}$. Then $H \cdot M \in M_c^2$ and

$$\Big\| H \cdot M \Big\|_{M^2}^2 = \mathbb{E}\Big(\int_0^\infty H_s^2 d\langle M \rangle_s \Big) \quad \cdots\cdots\cdots \ (\star)$$

proof continued)

Claim : $(\star)$ holds.

: We have, for $j > i$,

$$\mathbb{E}X^i_\infty X^j_\infty = \mathbb{E}\Big(H_{i-1}(M_{t_i} - M_{t_{i-1}})H_{j-1}\mathbb{E}(M_{t_j} - M_{t_{j-1}}|\mathcal{F}_{t_{i-1}})\Big) = 0$$

$$\Rightarrow \quad \big\|H \cdot M\big\|^2_{M^2} = \sum_{i=1}^m \big\|X^i\big\|^2_{M^2} = \sum_{i=1}^n \mathbb{E}(X^i_\infty)^2$$

Here, $\mathbb{E}(X^i_\infty)^2 = \mathbb{E}(H^2_{i-1}\mathbb{E}\Big((M_{t_i} - M_{t_{i-1}})^2|\mathcal{F}_{t_{i-1}})\Big)$ while $\mathbb{E}(M^2_{t_i} + M^2_{t_{i-1}} - 2M_{t_i}M_{t_{i-1}}|\mathcal{F}_{t_{i-1}}) = \mathbb{E}(M^2_{t_i} - M^2_{t_{i-1}}|\mathcal{F}_{t_{i-1}}) = \mathbb{E}(\langle M\rangle_{t_i} - \langle M\rangle_{t_{i-1}}|\mathcal{F}_{t_{i-1}})$ so

$$\mathbb{E}(X^i_\infty)^2 = \mathbb{E}(H^2_{i-1}\Big(\langle M\rangle_{t_i} - \langle M\rangle_{t_{i-1}})\Big) = \mathbb{E}\int_{t_{i-1}}^{t_i} H^2_s d\langle M\rangle_s$$

and therefore $\big\|H \cdot M\big\|^2_{M^2} = \mathbb{E}\int_0^\infty H^2_s d\langle M\rangle_s$

(End of proof) □

Proposition) Let $M \in M^2_c$ and let $H \in \mathcal{E}$. Then

$$\langle H \cdot M, N\rangle = H \cdot \langle M, N\rangle, \quad \forall N \in M^2_c$$

i.e. $\langle \int_0^\cdot H_s dM_s, N\rangle = \int_0^\cdot H_s d\langle M, N\rangle_s$.

proof) Let $H \cdot M = \sum_i X^i$ as previously. Then

$$\langle X^i, N\rangle_t = H_{i-1}\langle M_{t_i\wedge\cdot} - M_{t_{i-1}\wedge\cdot}, N\rangle_t = H_{i-1}(\langle M, N\rangle_{t_i\wedge t} - \langle M, N\rangle_{t_{i-1}\wedge t})$$

so

$$\langle H \cdot M, N\rangle_t = \int_0^t H_s d\langle M, N\rangle_s = (H \cdot \langle M, N\rangle)_t$$

(End of proof) □

3.2 Itô isometry

We extend our definition of Itô integral to a larger class of processes usinig Itô isometry. During the way, the definition of integral in the usual sense does not work anymore, but would find that the commutativity of Itô integral and the covariation is preserved under this extension, hence we would be able to use this as an alternative defining property for Itô integral.

Definition) For fixed $M \in M^2_c$, define $L^2(M)$ to be the space of equivalence classes of *predictable* $H : \Omega \times [0, \infty) \to \mathbb{R}$ such that

$$\big\|H\big\|_{L^2(M)} = \big\|H\big\|_M = \mathbb{E}\Big(\int_0^\infty H^2_s d\langle M\rangle_s\Big)^{1/2} < \infty$$

For $H, K \in L^2(M)$, set

$$(H, K)_{L^2(M)} = (H, K)_M = \mathbb{E}\Big(\int_0^\infty H_s K_s d\langle M\rangle_s\Big)$$

which is finite because of *Kunita-Watanabe inequality.*

Fact : $L^2(M) = L^2(\Omega \times [0, \infty), \mathcal{P}, d\mathbb{P}d\langle M\rangle)$ is a Hilbert space. (Recall $\mathcal{P}$ is the previsible σ-algebra)

Proposition) Let $M \in M^2_c$. Then $\mathcal{E}$, the space of simple processes, is dense in $L^2(M)$.

proof) Since $L^2(M)$ is a Hilbert space (in particular, is complete) it suffices to show that if $(K, H)_M = 0$ for all $H \in \mathcal{E}$ then $K = 0$.

So assume that $(K, H)_M = 0$ for all $H \in \mathcal{E}$ and set $X_t = \int_0^t K_s d\langle M \rangle_s$. X is a well-defined finite variation process with $X_t \in L^1$ for all t since

$$\mathbb{E}\int_0^t |K_s| d\langle M \rangle_s \leq \mathbb{E}\Big(\int_0^\infty K_s^2 d\langle M \rangle_s\Big)^{1/2} \mathbb{E}\Big(\langle M \rangle_\infty\Big)^{1/2} < \infty$$

where we used that $K \in L^2(M)$ and let $M \in M_c^2$.

♣ **Claim :** X is a continuous martingale.

: Let $0 \leq s < t$, $F \in \mathcal{F}_s$, $H = F 1_{(s,t]} \in \mathcal{E}$, F bounded. By assumption,

$$0 = (K, H)_M = \mathbb{E}\Big(F \int_s^t K_u d\langle M \rangle_u\Big) = \mathbb{E}(F(X_t - X_s)) \quad \forall s < t, \forall F \in \mathcal{F}_s \text{ bounded}$$

so $\mathbb{E}(X_t | \mathcal{F}_s) = X_s$ almost surely, *i.e.* X is a (continuous) martingale.

Since X is a finite variation process and a continuous martingale, we see $X \equiv 0$. So $K_u = 0$ for $d\langle M \rangle$-a.e. U, a.s. Hence $K = 0$ in $L^2(M)$.

(End of proof) □

Theorem/Definition) Let $M \in M_c^2$. Then

(i) The map $H \in \mathcal{E} \mapsto H \cdot M \in M_c^2$ extends uniquely to an isometry $L^2(M) \to M_c^2$, the *Itô isomtery.*

(ii) $H \cdot M$ is the unique martingale in M_c^2 such that

$$\langle H \cdot M, N \rangle = H \cdot \langle M, N \rangle, \quad \forall N \in M_c^2$$

$(H \cdot M)_t = \int_0^t H_s dM_s$ is then called the **Itô integral** of H with respect to M.

proof)

(i) For $H \in \mathcal{E}$, we have seen that $\Big\|H \cdot M\Big\|_{M^2} = \Big\|H\Big\|^2_{L^2(M)}$. Since $\mathcal{E} \subset L^2(M)$ is dense and M_c^2 is a Hilbert space, it follows that the map $H \mapsto H \cdot M$ extends uniquely to all of $L^2(M)$ and the extension is also an isometry.

(ii) Again, we have seen that $\langle H \cdot M, N \rangle = H \cdot \langle M, N \rangle$ holds for $H \in \mathcal{E}$. Given $H \in L^2(M)$, choose $(H^n)_n \subset \mathcal{E}$ such that $H^n \to H$ in $L^2(M)$. Then $H^n \cdot M \to H \cdot M$ by (i). We will justify

$$\begin{aligned}\langle H \cdot M, N \rangle_\infty &= \lim_{n\to\infty} \langle H^n \cdot M, N \rangle_\infty \quad \text{in } L^1 \\ &= \lim_{n\to\infty} (H^n \cdot \langle M, N \rangle)_\infty \\ &= (H \cdot \langle M, N \rangle)_\infty \quad \text{in } L^1\end{aligned}$$

- the equalities holds by the Kunita-Watanabe inequality.

(11th February, Monday)

proof continued)

♣ **Claim :** $\langle H\cdot M, N\rangle = \lim_{n\to\infty}\langle H^n\cdot M, N\rangle$ in L^1 and $\lim_{n\to\infty} H^n\cdot\langle M,N\rangle = H\cdot\langle M,N\rangle$ in L^1

: One has by Kunita-Watanabe inequality,

$$\begin{aligned}\mathbb{E}|\langle (H-H^n)\cdot M, N\rangle_\infty| \leq &\Big(\mathbb{E}\langle H\cdot M - H^n\cdot M\rangle_\infty\Big)^{1/2}\Big(\mathbb{E}\langle N\rangle_\infty\Big)^{1/2} \\ =&\Big\|H\cdot M - H^n\cdot M\Big\|_{M^2}\Big\|N\Big\|_{M^2} \\ =&\Big\|H-H^n\Big\|_{L^2(M)}\Big\|N\Big\|_{M^2}\to 0\end{aligned}$$

Also again by Kunita-Watanabe,

$$\mathbb{E}((H-H^n)\cdot\langle M,N\rangle)_\infty \leq \Big\|H-H^n\Big\|_{L^2(M)}\Big\|N\Big\|_{M^2}\to 0$$

Thus $\langle H\cdot M, N\rangle_\infty = (H\cdot\langle M,N\rangle)_\infty$ for all $N\in M_c^2$. Replacing N by the stopped martingale N^t gives

$$\langle H\cdot M, N\rangle_t = \langle H\cdot M, N\rangle_t = \langle H\cdot M, N^t\rangle_\infty = (H\cdot\langle M, N^t\rangle)_\infty = (H\cdot\langle M,N\rangle)_t$$

Next, **Uniqueness** : assume that $X\in M_c^2$ also satisfies $\langle X,N\rangle = H\cdot\langle M,N\rangle = \langle H\cdot M, N\rangle$ for all $N\in M_c^2$. Then $\langle H\cdot M - X, N\rangle = 0\quad \forall N\in M_c^2$. Since $H\cdot M - X\in M_c^2$, we have

$$\langle H\cdot M - X, H\cdot M - X\rangle = 0$$

and therefore $\Big\|H\cdot M - X\Big\|_{M^2} = 0$ implying $X = H\cdot M$.

(End of proof) □

So we

Corollary) If T is a stopping time, then

$$(1_{[0,T]}H)\cdot M = (H\cdot M)^T = H\cdot M^T$$

proof) For any $N\in M_c^2$, has

$$\langle (H\cdot M)^T, N\rangle_t = \langle H\cdot M, N\rangle_{t\wedge T} = (H\cdot\langle M,N\rangle)_{t\wedge T} = \Big(1_{[0,T]}H\cdot\langle M,N\rangle\Big)_t$$

and by uniqueness of such integral, has $(H\cdot M)^T = H1_{[0,T]}\cdot M$.

The next equality $(H\cdot M)^T = H\cdot M^T$ follows from completely analogous argument.

(End of proof) □

Corollary) $\langle H\cdot M, K\cdot N\rangle = (HK)\cdot\langle M,N\rangle$, *i.e.*

$$\langle\int_0^\cdot H_s dM_s, \int_0^\cdot K_s dN_s\rangle_t = \int_0^t H_sK_s d\langle M,N\rangle_s$$

proof) Has

$$\langle H\cdot M, K\cdot N\rangle = H\cdot\langle M, K\cdot N\rangle = H\cdot(K\cdot\langle M,N\rangle) = (HK)\cdot\langle M,N\rangle$$

where the last equaility follows from associativity of finite variational integrals, *i.e.* $hkdf = h(kdf) = hd(k\circ f)$.

(End of proof) □

Corollary) One has, if $t > u$,

$$\mathbb{E}\Big(\int_0^t H_s dM_s\Big) = 0$$
$$\mathbb{E}\Big(\int_0^t H_s dM_s | \mathcal{F}_u\Big) = \int_0^u H_s dM_s$$
$$\mathbb{E}\Big(\int_0^t H_s dM_s \int_0^t K_s dN_s\Big) = \mathbb{E}\Big(\int_0^t H_s K_s d\langle M, N\rangle_s\Big)$$

proof) $H \cdot M$ and $(H \cdot M)(K \cdot N) - \langle H \cdot M, K \cdot N\rangle$ are martingales starting at 0.

(End of proof) □

Corollary) *(Associativity of Itô integral)* Let $H \in L^2(M)$. Then $KH \in L^2(M)$ *iff* $K \in L^2(H \cdot M)$ and then

$$(KH) \cdot M = K \cdot (H \cdot M)$$

proof) Since $H^2 \cdot \langle M\rangle = \langle H \cdot M\rangle$, has

$$\mathbb{E}\Big(\int_0^\infty K_s^2 H_s^2 d\langle M\rangle_s\Big) = \mathbb{E}\Big(\int_0^\infty K_s^2 d\langle H \cdot M\rangle_s\Big)$$

so $HK \in L^2(M)$ is equivalent to having $K \in L^2(H \cdot M)$.

For $N \in M_c^2$, has

$$\langle (KH) \cdot M, N\rangle_t = \Big((KH) \cdot \langle M, N\rangle\Big)_t = \Big(K \cdot (H \cdot \langle M, N\rangle)\Big)_t$$

and

$$\langle K \cdot (H \cdot M), N\rangle_t = (K \cdot \langle H \cdot M, N\rangle)_t = (K \cdot (H \cdot \langle M, N\rangle))_t$$

so $KH \cdot M = K \cdot (H \cdot M)$ by uniqueness.

(End of proof) □

3.3 Extension to local martingales

Definition) Let M be a *continuous local martingale.* Define $L^2_{loc}(M)$ to be the space of (up to euivalence classes) predictable H such that

$$\text{a.s.,} \quad \forall t \geq 0, \ \int_0^t H_s^2 d\langle M\rangle_s < \infty$$

Theorem) Let M be a continuous local martingale.

(i) For every $H \in L^2_{loc}(M)$, there is a unique (up to indistiguishability) contiuous local martingale $H \cdot M$ with $(H \cdot M)_0 = 0$ such that

$$\langle H \cdot M, N\rangle = H \cdot \langle M, N\rangle \quad \forall N \text{ continuous local martingale}$$

(ii) If $H \in L^2_{loc}(M)$ and K is predictable then $K \in L^2_{loc}(H \cdot M)$ *iff* $HK \in L^2_{loc}(M)$ and then

$$H \cdot (K \cdot M) = (HK) \cdot M$$

(iii) If T is a stopping time,

$$(1_{[0,T]}H)\cdot M = (H\cdot M)^T = H\cdot M^T$$

Finally, if $M\in M_c^2$ and $H\in L^2(M)$ then the definition is consistent with then previous one.

(13th February, Wednesda)

Theorem) Let M be a continuous local martingale.

(i) For every $H\in L^2_{loc}(M)$, there is a unique (up to indistiguishability) contiuous local martingale $H\cdot M$ with $(H\cdot M)_0 = 0$ such that

$$\langle H\cdot M, N\rangle = H\cdot\langle M,N\rangle \quad \forall N \text{ continuous local martingale}$$

(ii) If $H\in L^2_{loc}(M)$ and K is predictable then $K\in L^2_{loc}(H\cdot M)$ *iff* $HK\in L^2_{loc}(M)$ and then

$$H\cdot(K\cdot M) = (HK)\cdot M$$

(iii) If T is a stopping time,

$$(1_{[0,T]}H)\cdot M = (H\cdot M)^T = H\cdot M^T$$

Finally, if $M\in M_c^2$ and $H\in L^2(M)$ then this definition is consistent with the previous one.

proof) Once we prove (i), the other two follow using the exactly same arguement. So we just prove (i).

(i) Assume $M_0 = 0$ and $\int_0^t H_s^2 d\langle M\rangle_s < \infty$ for all (t,ω) (by setting $H=0$ where this fails). Set $S_n = \inf\{t\geq 0 : \int_0^t (1+H_s^2)d\langle M\rangle_s > n\}$. Note that S_n is a stopping times, $S_n\nearrow\infty$ as $n\to\infty$ and

$$\langle M^{S_n}, M^{S_n}\rangle_t = \langle M,M\rangle_{t\wedge S_n} \leq n$$

Hence M^{S_n} are in M_c^2 and $\int_0^\infty H_s^2 d\langle M^{S_n}\rangle_s = \int_0^{S_n} H_s^2 d\langle M\rangle_s \leq n$ so $H\in L^2(M^{S_n})$. So the stochastic integral $H\cdot M^{S_n}$ is already defined for each n, and we can let

$$(H\cdot M)^{S_n} = \left(H\cdot M^{S_m}\right)^{S_n} \quad \text{for } m>n$$

(since "stopping commutes with stochastic integral"). So there is a unique (up to indistinguishability) process denoted $H\cdot M$ such that

$$(H\cdot M)^{S_n} = H\cdot M^{S_n} \quad \forall n\in\mathbb{N}$$

Then $H\cdot M$ is adapted continuous, and it is a local martingale since the $(H\cdot M)^{S_n} = (H\cdot M^{S_n})$ are martingales and $S_n\to\infty$.

♣ **Claim :** $\langle H\cdot M, N\rangle = H\cdot\langle M,N\rangle$ for any continuous local martingale N.

: Assume that $N_0 = 0$ and set $S'_n = \inf\{t \geq 0 : |N_t| > n\}$, $T_n = S_n \wedge S'_n$. Then $N^{S'_n}$ is in M_c^2 and

$$\begin{aligned} \langle H \cdot M, N \rangle^{T_n} &= \langle (H \cdot M)^{S_n}, N^{S'_n} \rangle \\ &= \langle H \cdot M^{S_n}, N^{S'_n} \rangle \\ &= H \cdot \langle M^{S_n}, N^{S'_n} \rangle \\ &= H \cdot \langle M, N \rangle^{T_n} \\ &= (H \cdot \langle M, N \rangle)^{T_n} \end{aligned}$$

Since $T_n \nearrow \infty$, thus $\langle H \cdot M, N \rangle = H \cdot \langle M, N \rangle$.

Uniqueness follows as before. Also, (ii) and (iii) follow from (i) as before.

For the last statement, ff $M \in M_c^2$ and $H \in L^2(M)$ then $H \cdot M \in M_c^2$ by (i) which shows that $\langle H \cdot M \rangle_\infty = (H^2 \cdot \langle M \rangle)_\infty$ and thus $\left\| H \cdot M \right\|_{M^2}^2 = \mathbb{E}\langle H, M \rangle_\infty < \infty$. The uniqueness statement in the equivalent of (i) from L^2-bounded cases shows consistency with previous definition. *(End of proof)* □

3.4 Extension to Semimartingales

Definition) A process H is **locally bounded** if

$$\text{a.s., } \forall t > 0, \quad \sup_{s \leq t} |H_s| < \infty$$

In particular, any continuous process is locally bounded.

Fact : If H is locally bounded and predictable and if A is a finite variation process,

$$\forall t > 0, \quad \int_0^t H_s |dA_s| < \infty \quad \text{a.s.}$$

In particular, for such H, and M a continuous local martingale, it follows that $H \in L^2_{loc}(M)$.

Definition) Let $X = X_0 + M + A$ be a continuous semimartingale, and let H be a predictable locally bounded process. Then the **Itô integral** $H \cdot X$ **is the continuous semimartingale** is

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

where $H \cdot M$ is the integral defined in the previous section and $H \cdot A$ is the finite variation integral. We write $(H \cdot X)_t = \int_0^t H_s dX_s$.

Proposition) *(Stochastic Dominated Convergence Theorem, Stochastic DCT)* Let X be a continuous semimartingale, and let H be locally bounded predictable process and let K be a predictable non-negative process. Let $t > 0$ and assume that

(i) $H_s^n \xrightarrow{n \to \infty} H_s$ for all $s \in [0, t]$.

(ii) $|H_s^n| \leq K_s$ for all $s \in [0, t]$ and $n \in \mathbb{N}$.

(iii) $\int_0^t K_s^2 d\langle M \rangle + \int_0^t K_s |dA_s| < \infty$ (where $X = X_0 + M + A$). *[This condition is always true if K is locally bounded]*

Then $\int_0^t H_s^n dX_s \xrightarrow{ucp} \int_0^t H_s dX_s$ as $n \to \infty$.

proof) Let $X = X_0 + X + A$. For the finite variation part A, the *usual DCT* implies

$$\int_0^t H_s^n dA_a \to \int_0^t H_s dA_s$$

Set $T_m = \inf\{t \geq 0 : \int_0^t K_s^2 d\langle M \rangle_s > n\}$. Then

$$\mathbb{E}\Big[\Big(\int_0^{t\wedge T_m} H_s^n dM_s - \int_0^{t\wedge T_m} H_s dM_s\Big)^2\Big] \leq \mathbb{E}\Big[\int_0^{t\wedge T_m} (H_s^n - H_s)^2 d\langle M \rangle_s\Big] \to 0$$

by DCT. Since $T_m \wedge t = t$ for m large enough almost surely, $\mathbb{P}(T_m \wedge t = t) \to 1$ and this implies the convergence in probability for fixed t by Markov's inequality. For convergence ucp(uniformly on compact domains in probability), use Doob's inequality to see that

$$\begin{aligned}\mathbb{E}\Big[\sup_{t\in[0,u]} \Big(\int_0^{t\wedge T_m} H_s^n dM_s - \int_0^{t\wedge T_m} H_s dM_s\Big)^2\Big] \leq& 4 \sup_{t\in[0,u]} \mathbb{E}\Big[\int_0^{t\wedge T_m} (H_s^n - H_s)^2 d\langle M \rangle_s\Big] \\ =& 4\mathbb{E}\Big[\int_0^{u\wedge T_m} (H_s^n - H_s)^2 d\langle M \rangle_s\Big] \to 0\end{aligned}$$

and again use Markov's inequality. *(End of proof)* □

Corollary) Let X be a continuous semimartingale, and let H be a locally bounded adapted left-continuous process. Then for any subdivision $0 = t_0^{(m)} < \cdots < t_{n_m}^{(m)} = t$ of $[0,t]$ with $\max_i |t_i^{(m)} - t_{i-1}^{(m)}| \to 0$ as $m \to \infty$, has :

$$\lim_{m\to\infty} \sum_{i=1}^{n_m} H_{t_{i-1}^{(m)}} \Big(X_{t_i^{(m)}} - X_{t_{i-1}^{(m)}}\Big) = \int_0^t H_s dX_s$$

where the convergence is made ucp. *[Taking the left end-point $t_{i-1}^{(m)}$ is important, and is consistent with the choice of Itô integral. Different choice corresponds to what the integral means.]*

proof) Exactly as in the finite variation case (using stochastic DCT).

(15th February, Friday)

Last time : $\int_0^t H_s dX_s = \lim_{m\to\infty} \sum_{i=1}^{n_m} H_{t_{i-1}^{(m)}}(X_{t_i^{(m)}} - X_{t_{i-1}^{(m)}})$.

Remark : Suppose H is continuous. Unlike the case that X is of finite variation, it is essential here that H is evaluated at the left end point.

Example :

$$\lim_{m\to\infty} \sum_{i=1}^{n_m} X_{t_{i-1}^{(m)}}(X_{t_i^{(m)}} - X_{t_{i-1}^{(m)}}) = \int_0^t X_s dX_s$$

but

$$\begin{aligned}&\lim_{m\to\infty} \sum_{i=1}^{n_m} X_{t_i^{(m)}}(X_{t_i^{(m)}} - X_{t_{i-1}^{(m)}}) \\ =& \lim_{m\to\infty} \sum_{i=i}^{n_m} X_{t_{i-1}^{(m)}}(X_{t_i^{(m)}} - X_{t_{i-1}^{(m)}}) + \lim_{m\to\infty} \sum_{i=1}^{n_m} (X_{t_i^{(m)}} - X_{t_{i-1}^{(m)}})^2 \\ =& \int_0^t X_s dX_s + \langle X, X \rangle_t\end{aligned}$$

For example, if $X = B$ is the standard Brownian motion, then $\langle X, X\rangle_t = t$.

Remark : The choice of the left endpoint gives the Itô integral. There is another common choice called the **Stratonovich integral** which is defined by

$$\int_0^t X_s \partial Y_s = \int_0^t X_s dY_s + \frac{1}{2}\langle X, Y\rangle_t$$

It thus corresponds to the approximation

$$\sum_{i=1}^{n_m} \frac{1}{2}(X_{t_i^{(m)}} + X_{t_{i-1}^{(m)}})(Y_{t_i^{(m)}} - Y_{t_{i-1}^{(m)}})$$

Note that $\int_0^\cdot X_s \partial Y_s$ is generally not a local martingale.

3.5 Itô formula

Theorem) *(Integration by parts)* Let X and Y be continuous semimartingales. Then a.s,

$$X_tY_t - X_0Y_0 = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y\rangle_t$$

The last term $\langle X, Y\rangle$ is called the **Itô correction.**

Remark : In terms of the Stratonovich integral, a.s,

$$X_tY_t - X_0Y_0 = \int_0^t X_s dY_s + \int_0^t Y_s dX_s$$

proof of the theorem) Clearly,

$$\begin{aligned}
& X_tY_t - X_sY_s = X_s(Y_t - Y_s) + Y_s(X_t - X_s) + (X_t - X_s)(Y_t - Y_s) \\
\Rightarrow \quad & X_{k2^{-m}}Y_{k2^{-m}} - X_0Y_0 = \sum_{i=1}^{k}(X_{i2^{-m}}Y_{i2^{-m}} - X_{(i-1)2^{-m}}Y_{(i-1)2^{-m}}) \\
&= \sum_{i=1}^{k}\Big(X_{(i-1)2^{-m}}(Y_{i2^{-m}} - Y_{(i-1)2^{-m}}) + Y_{(i-1)2^{-m}}(X_{i2^{-m}} - X_{(i-1)2^{-m}}) \\
&\qquad + (X_{i2^{-m}} - X_{(i-1)2^{-m}})(Y_{i2^{-m}} - Y_{(i-1)2^{-m}})\Big)
\end{aligned}$$

Thus for $t \in 2^{-n} \in \mathbb{N}$, by taking $m \to \infty$,

$$X_tY_t - X_0Y_0 = (X \cdot Y)_t + (Y \cdot X)_t + \langle X, Y\rangle_t \quad \text{(by previous approximation results)}$$

Finally, for $t \in \mathbb{R}$ use continuity to conclude.

(End of proof) $\square$

Theorem) *(Itô formula)* Let $X^1, \cdots, X^p$ be continuous semimartingales, and let $f \in C^2(\mathbb{R}^p; \mathbb{R})$. Then, writing $X = (X^1, \cdots, X^p)$, a.s,

$$f(X_t) = f(X_0) + \sum_{i=1}^{p}\int_0^t \frac{\partial f}{\partial x^i}(X_s)dX_s^i + \frac{1}{2}\sum_{i,j=1}^{p}\int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s)d\langle X^i, X^j\rangle_s \quad \cdots\cdots\cdots (\star)$$

[Note that if X is a local martingale, then the second term in the RHS is a continuous local martingale and the third term in the RHS is a finite variation process, hence $f(X_t)$ is a continuous semimartingale - so we end up with a semimartingale even if we want to just work with martingales.]

Informally, we may write

$$df(X_t) = \sum_{i=1}^{p} \frac{\partial f}{\partial x^i}(X_t)dX_t^i + \frac{1}{2}\sum_{i,j=1}^{p} \frac{\partial^2 f}{\partial x^i \partial x^j}(X_t)d\langle X^i, X^j\rangle_t$$

proof) For f constant, $(\star)$ is obvious.

♠ **Claim :** Assume $(\star)$ holds for some f. Then it also holds for g defined by $g(x) = x^k f(x)$, for $x = (x^1, \cdots, x^p) \in \mathbb{R}^p$.

: Application of *integration by parts*(theorem above) with $X = X^k$ and $Y = f(X)$ gives

$$g(X_t) - g(X_0) = \int_0^t X_s^k df(X_s) + \int_0^t f(X_s)dX_s^k + \langle X^k, f(X)\rangle_t = \text{(A)+(B)+(C)}$$

By $(\star)$ for f, and $(H \cdot (K \cdot X)) = (HK) \cdot X$,

$$\begin{aligned}
\text{(A)} &= \int_0^t X_s^k df(X_s) = (X^k \cdot f(X))_t \\
&= \Big(X^k \cdot \Big(f(X_0) + \sum_{i=1}^{p}(\frac{\partial f}{\partial x^i} \cdot X^i)_t + \frac{1}{2}\sum_{i,j=1}^{p}\Big(\frac{\partial^2 f}{\partial x_i \partial x_j} \cdot \langle X^i, X^j\rangle\Big)\Big)\Big)_t \\
&= \sum_{i=1}^{p}\int_0^t X_s^k \frac{\partial f}{\partial x^i}(X_s)dX_s^i + \frac{1}{2}\sum_{i,j=1}^{p}\int_0^t X_s^k \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s)d\langle X^i, X^j\rangle_s
\end{aligned}$$

By $(\star)$ for f and $\langle X, H \cdot Y\rangle = H \cdot \langle X, Y\rangle$, and noting that finite variation part has no contribution in covariance,

$$\text{(C)} = \langle X^k, f(X)\rangle_t = \sum_{i=1}^{p}\int_0^t X_s^k \frac{\partial f}{\partial x^i}(X_s)d\langle X^k, X^i\rangle_s$$

Putting these together,

$$g(X_t) = g(X_0) + \sum_{i=1}^{p}\int_0^t \frac{\partial g}{\partial x^i}(X_s)dX_s^i + \frac{1}{2}\sum_{i,j=1}^{p}\int_0^t \frac{\partial^2 g}{\partial x^i \partial x^j}(X_s)d\langle X^i, X^j\rangle_s$$

Having the claim, by induction, $(\star)$ holds for all polynomials.

For a continuous local martingale X^i, write $X^i = X_0^i + M^i + A^i$, with M^i a continuous local martingale and A^i a finite variation process.

♠ **Claim :** $(\star)$ holds for $f \in C^2$ if $|X_t^i(\omega)| \le n$, and $\int_0^t |dA_s^i| \le n$ for all (t, ω).

: By the *Weierstrass approximation theorem*, there are polynomials p_k such that

$$\sup_{|k| \le n}\Big(|f(x) - p_k(x)| + |Df(x) - Dp_k(x)| + |D^2 f(x) - D^2 p_k(x)|\Big) \le \frac{1}{k}$$

Taking limits, in probability,

$$\begin{aligned}
&f(X_t) - f(X_0) = \lim_{k\to\infty}(p_k(X_t) - p_k(X_0)) \\
&\int_0^t \frac{\partial f}{\partial x^i}(X_s)dX_s^i = \lim_{k\to\infty}\int_0^t \frac{\partial p_k}{\partial x^i}(X_s)dX_s^i \quad \text{by } \textit{stochastic DCT} \\
&\int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s)d\langle X_s^i, X_s^j\rangle = \lim \int_0^\infty \frac{\partial^2 p_k}{\partial x^i \partial x^j}(X_s)d\langle X^i, X^j\rangle_s \quad \text{by DCT}
\end{aligned}$$

(18th February, Monday)

We were proving,
Theorem) *(Itô formula)* Let $X^1, \cdots, X^p$ be continuous semimartingales, and let $f \in C^2(\mathbb{R}^p;\mathbb{R})$. Then, writing $X = (X^1, \cdots, X^p)$, a.s,

$$f(X_t) = f(X_0) + \sum_{i=1}^{p} \int_0^t \frac{\partial f}{\partial x^i}(X_s) dX_s^i + \frac{1}{2}\sum_{i,j=1}^{p} \int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s) d\langle X^i, X^j\rangle_s \quad \cdots\cdots\cdots \ (\star)$$

proof continued) Last time we proved $(\star)$ under the additional hypothesis that

$$\max_i |X_t^i| \leq n, \quad \max_i \int_0^t |dA_s^i| \leq n$$

For the general case, let $T_n = \inf\{t \geq 0 : \max_| X_t^i| \geq n, \max_i \int_0^t |dA_s^i| \geq n\}$. Then by the case we already know,

$$\begin{aligned} f(X_t^{T_n}) =& f(X_0^{T_n}) + \sum_{i=1}^{p} \int_0^t \frac{\partial f}{\partial x^i}(X_s^{T_n}) d(X^{T_n})_s^i + \frac{1}{2}\sum_{i,j=1}^{p} \int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s^{T_n}) d\langle (X^{T_n})^i, (X^{T_n})^j\rangle_s \\ =& f(X_0) + \int_0^{t\wedge T_n} \frac{\partial f}{\partial x^i}(X_s) dX_s^i + \int_0^{t\wedge T_n} \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s) d\langle X^i, X^j\rangle_s \end{aligned}$$

Take $n \to \infty$ then we are done.

(End of proof) □

If you are lost in the way getting here, this is a good point to jump back in - most of the utility of *Stochastic Calculus* comes from this Itô formula, and does not care much about how we arrived here.

Remark : In terms of the *Stratonovich integral*,

$$f(X_t) = f(X_0) + \sum_{i=1}^{p} \int_0^t \frac{\partial f}{\partial x^i}(X_s) dX_s^i$$

This is the reason why Stratonovich integral is sometimes more useful than Itô integral. However, it is hard to verify if the result is a martingale or not in this setting, so you will eventually need to pass on to Itô integral in order to do martingale computations.

Summary of calculation rules for the Itô integral :

Let us adopt the notations

$$Z_t - Z_0 = \int_0^t H_s dX_{S_t} \quad \Leftrightarrow \quad dZ_t = H_t dX_t$$
$$Z_t - Z_0 = \langle X, Y\rangle_t = \int_0^t d\langle X, Y\rangle_s \quad \Leftrightarrow \quad dZ_t = dX_t dY_t$$

Then, :

"Associativity" $H_t(K_t dX_t) = (H_t K_t) dX_t$, $\quad$ (*i.e.* $H \cdot (K \cdot X) = (HK) \cdot X$)
"Kunita-Watanabe equality" $H_t dX_t dY_t = (H_t dX_t) dY_t$, $\quad$ (*i.e.* $H \cdot \langle X, Y\rangle = \langle H \cdot X, Y\rangle$)
"Itô formula" $df(X_t) = \sum_i \frac{\partial f}{\partial x^i} dX_t^i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial x^i \partial x^j}(X_t) dX_t^i dX_t^j$

For example, if $X = B$ is a Brownian motion, then

$$df(B_t) = f'(B_t) dB_t + \frac{1}{2} f''(B_t) dt, \quad i.e.\ (dB_t)^2 = dt$$

4 Applications to Brownian Motion and Martingales

4.1 Lévy's characterisation of Brownian motion

Theorem) Let $X = (X^1, \cdots, X^p)$ be continuous local martingales. Suppose $X_0 = 0$ and that $\langle X^i, X^j \rangle_t = \delta_{ij} t$ for all $t \geq 0$. Then X is a standard p-dimensional Brownaim motion. That is, *the covariation singles out the Brownian motion.*

proof) It suffices to show that for all $0 \leq s < t$, (i) $X_t - X_s$ is independent of $\mathcal{F}_s$ and (ii) $X_t - X_s \sim N(0, (t-s)id_{p\times p})$.

Both properties follow form the following claim :

$$\mathbb{E}\Big(e^{i\theta\cdot(X_t - X_s)}|\mathcal{F}_s\Big) = e^{-\frac{1}{2}|\theta|^2(t-s)} \quad \text{for all } \theta \in \mathbb{R}^p,\ s < t \quad \cdots\cdots\cdots\ (\diamondsuit)$$

Indeed, $(\diamondsuit)$ implies $\mathbb{E}(e^{i\theta\cdot(X_t-X_s)}) = \exp(-\frac{1}{2}|\theta|^2(t-s))$ thus $X_t - X_s \sim N(0,(t-s)id)$. To show independence, consider $A \in \mathcal{F}_s$ with $\mathbb{P}(A) > 0$. Then $X_t - X_s \sim N(0,(t-s)id)$ under $\mathbb{P}(\cdot|A) = \mathbb{P}(\cdot \cap A)/\mathbb{P}(A)$. Hence

$$\mathbb{E}(1_A f(X_t - X_s)) = \mathbb{P}(A)\mathbb{E}(f(X_t - X_s)) \quad \forall f \text{ measurable},$$

i.e. A is independent of $X_t - X_s$. The same is trivial if $\mathbb{P}(A) = 0$. Thus, $X_t - X_s$ is independent of $\mathcal{F}_s$.

To show $(\diamondsuit)$, fix $\theta \in \mathbb{R}^p$ and set $Y = \theta \cdot X_t = \sum_{i=1}^p \theta^i X_t^i$. Then

$$\langle Y \rangle_t = \langle Y, Y \rangle_t = \sum_{i,j=1}^p \theta^i \theta^j \langle X^i, X^j \rangle_t = \sum_{i=1}^p (\theta^i)^2 t = |\theta|^2 t$$

where the third equality follows from the assumption in the statement of the theorem. Let $Z_t = e^{iY_t + \frac{1}{2}\langle Y \rangle_t} = e^{iY_t + \frac{1}{2}|\theta|^2 t}$. By Itô formula with $f(x) = e^x$ and $\overline{X} = iY + \frac{1}{2}\langle Y \rangle$,

$$\begin{aligned} dZ_t =& df(X_t) = Z_t\Big(d\overline{X}_t + \frac{1}{2}d\langle \overline{X} \rangle_t\Big) \\ =& Z_t(i \cdot dY_t + \frac{1}{2}d\langle Y \rangle_t - \frac{1}{2}d\langle Y \rangle_t) = iZ_t dY_t \end{aligned}$$

(verify these calculations - might be faulty.) In particular, Z is a continuous local martingale. Since Z is bounded on every bounded interval, Z is in fact a martingale (when we are checking martingale property, we only do computations on bounded times, so it suffices to have absolute bound on each bounded interval). So $\mathbb{E}(Z_t|\mathcal{F}_s) = Z_s$ and hence

$$\mathbb{E}\Big(e^{i\theta\cdot(X_t - X_s)}|\mathcal{F}_s\Big) = e^{-\frac{1}{2}|\theta|^2(t-s)}$$

which was the claim.

(End of proof) □

This theorem can be of intrinsic interest, but is also interesting in terms of applications. We often construct a stochastic process not directly from a Brownian motion, but in different ways, for example by Wiener measure. Then this characterization verifies if this process is a Brownian motion or not.

4.2 Dubins-Schwarz Theorem

Theorem) Let M be a continuous local martingale with $M_0 = 0$ and $\langle M\rangle_\infty = \infty$ a.s. Let $T_s = \inf\{t \geq 0 : \langle M\rangle_t > s\}$ be the right-continuous inverse of $\langle M\rangle$.

$$B_s = M_{T_s}, \quad \mathcal{G}_s = \mathcal{F}_{T_s}$$

Then T_s is an $(\mathcal{F}_t)$ stopping times, $\langle M\rangle_{T_s}$ for all $s \geq 0$, B is a $(\mathcal{G}_s)_{s\geq 0}$-Brownian motion and

$$M_t = B_{\langle M\rangle_t},$$

i.e. M is a random time change of a Brownian motion.

(20th February, Wednesday)

Before we prove the theorem, we need a lemma.

Lemma) Let M be a continuous local martingale. Almost surely for all $u < v$, M is constant on $[u, v]$ *iff* $\langle M\rangle$ is constant on $[u, v]$.

proof) By continuity, it suffices to prove that for any fixed u, v, a.s.

$$\{M_t = M_u \ \forall t \in [u, v]\} = \{\langle M\rangle_u = \langle M\rangle_v\}$$

Let $N_t = M_t - M_{t\wedge u} = \int_{t\wedge u}^t dM_s$. Then $\langle N\rangle_t = \langle M\rangle_t - \langle M\rangle_{t\wedge u}$ (by *Kunita-Watanabe* formula). For any ϵ, let $T_\epsilon = \inf\{t \geq 0 : \langle N\rangle_t > \epsilon\}$. Then $N^{T_\epsilon} \in M_c^2$ since $\langle N^{T_\epsilon}\rangle \leq \epsilon$ and $\mathbb{E}((N_t^{T_\epsilon})^2) = \mathbb{E}(\langle N^{T_\epsilon}\rangle_t) \leq \epsilon$. So

$$\mathbb{E}\Big(N_t^2 \mathbf{1}_{\{\langle M\rangle_v = \langle M\rangle_u\}}\Big) = \mathbb{E}\Big(\mathbf{1}_{\langle N\rangle_v = 0} N_{t\wedge T_\epsilon}^2\Big) \leq \mathbb{E}(N_{t\wedge T_\epsilon}^2) \leq \epsilon$$

so $N_t = 0$ a.s. on $\{\langle M\rangle_v = \langle M\rangle_u\}$ for any $t \in [u, v]$. Hence we have shown that a.s. $\langle M\rangle_u = \langle M\rangle_v$ implies M is constant on $[u, v]$.

The other direction is implied by the approximation formula for the quadratic variation. Without loss of generality, put $u = 0 < v$ and let $\langle M\rangle_v^{(m)} = \sum_{i=1}^{2^m} \Big(M_{i\cdot 2^{-m}v} - M_{(i-1)\cdot 2^{-m}v}\Big)^2$, then $\langle M\rangle_v^{(m)} = 0$ whenever $M \equiv M_0$ on $[0, v]$. So for any $\epsilon > 0$,

$$\begin{aligned}&\mathbb{P}\Big(|\langle M\rangle_v - \langle M\rangle_v^{(m)}| \geq \epsilon, \ M_t = M_0 \ \forall t \in [0, v]\Big) \\ =&\mathbb{P}\Big(|\langle M\rangle_v| \geq \epsilon \ M_t = M_0 \ \forall t \in [0, v]\Big) \to 0 \quad \text{as } m \to \infty\end{aligned}$$

so in fact $\mathbb{P}\Big(|\langle M\rangle_v| \geq \epsilon \ M_t = M_0 \ \forall t \in [0, v]\Big) = 0$ for any $\epsilon > 0$, which concludes the proof.

[Note that we will never use the last implication]

(End of proof) □

Theorem) *(Dubins-Schwarz)* Let M be a continuous local martingale with $M_0 = 0$ and $\langle M\rangle_\infty = \infty$ a.s. Let $T_s = \inf\{t \geq 0 : \langle M\rangle_t > s\}$ be the right-continuous inverse of $\langle M\rangle$. Let

$$B_s = M_{T_s}, \quad \mathcal{G}_s = \mathcal{F}_{T_s}$$

Then T_s is an $(\mathcal{F}_t)$ stopping time, $\langle M\rangle_{T_s} = s$ for all $s \geq 0$, B is a $(\mathcal{G}_s)_{s\geq 0}$-Brownian motion and

$$M_t = B_{\langle M\rangle_t}$$

proof) Since $\langle M \rangle$ is continuous and adapted, the T_s are stopping times and $T_s < \infty$ a.s. Redefine $T_s = 0$ if $\langle M \rangle_\infty < \infty$. Note that T_s is still a stopping times since $(\mathcal{F}_s)$ is complete, hence $\mathcal{F}_0$ includes all events of probability 0.

♡ **Claim :** $(\mathcal{G}_s)$ is a filtration obeying the *usual conditions. [note, we have constructed our thoery on the implicit assumption that each filtration obeys the usual condition. So should* $(\mathcal{G}_s)$*]*

: For $A \in \mathcal{G}_s = \mathcal{F}_{T_s}$ and $s < t$,

$$A \cap \{T_t \leq u\} = A \cap \{T_s \leq u\} \cap \{T_t \leq u\} \in \mathcal{F}_u$$

so $A \in A_{T_t} = \mathcal{G}_t$. Hence $(\mathcal{G}_s)$ is a filtration.

Right-continuity and completeness are descended from the properties $(\mathcal{F}_t)$ and right-continuity of $t \mapsto T_t$.

♡ **Claim :** B is adapted to $(\mathcal{G}_s)$.

: Recall, from *Advanced Probability*, that if X is càdlàg and T is a stopping time then $X_T \mathbf{1}_{T<\infty} \in \mathcal{F}_T$. Apply this with $X = M$, $T = T_s$ and $\mathcal{F}_T = \mathcal{G}_s$. This gives $B_t \in \mathcal{G}_s$.

♡ **Claim :** B is continuous.

: T_s is càdlàg in s, so $B_s = M_{T_s}$ is càdlàg and thus right-continuous.
Also to check B is left continuous, observe B is left-continuous at $s \Leftrightarrow B_s = B_{s^-} \Leftrightarrow M_{T_s} = M_{T_{s^-}}$ where $T_{s^-} = \inf\{t \geq 0 : \langle M \rangle_t = s\}$. We divide into two cases,

- if $T_s = T_{s^-}$ then B is left-continuous.
- if $T_s > T_{s^-}$ then $\langle M \rangle$ is constant on $[T_{s^-}, T_s]$. Hence $M_{T_s} = M_{T_{s^-}}$ holds by the previous lemma, *i.e.* B is left-continuous.

By these two cases, we see that B is left-continuous

♡ **Claim :** B is a continuous martingale with respect to $(\mathcal{G}_s)$ and $\langle B \rangle_s = s$ for all $s \geq 0$.

: Let $0 \leq r < s$. Then $\langle M^{T_s} \rangle_\infty = \langle M \rangle_{T_s} = s$ (here we used the fact that $\langle M^{T_s} \rangle_t$ tends to ∞ a.s.). So $M^{T_s} \in \mathcal{M}_c^2$, and so $(M^2 - \langle M \rangle)^{T_s}$ is a uniformly integrable martingale. Now *Optional Stopping Theorem* implies

$$\mathbb{E}(B_s|\mathcal{G}_r) = \mathbb{E}(M_\infty^{T_s}|\mathcal{F}_{T_r}) = M_{T_r} = B_r$$
$$\mathbb{E}(B_s^2 - s|\mathcal{G}_r) = \mathbb{E}\Big((M^2 - \langle M \rangle)_\infty^{T_s}|\mathcal{F}_{T_r}\Big) = M_{T_r}^2 - \langle M \rangle_{T_r} = B_r^2 - r$$

By the claim and the *Lévy's characterisation* of Brownian motions, it follows that B is a Brownian motion.

(End of proof) □

Caution! : Schwarz in *Dubins-Schwarz* does not have a 't' in his name.

4.3 Girsanov's Theorem

This is a next application of Itô's formula.

Example : Let $X \sim N(0, C)$. be an n-dimensional Gaussian vector with positive definite covariance matrix $C = (C_{ij})_{i,j=1}^n$ and mean 0. Then

$$\mathbb{E}(f(X)) = \Big(\det \frac{M}{2\pi}\Big)^{1/2} \int_{\mathbb{R}^n} f(x) e^{-\frac{1}{2}x\cdot Mx} dx, \quad (M = C^{-1})$$

Let $a \in \mathbb{R}^n$. Then

$$\begin{aligned}\mathbb{E}(f(X+a)) =& \Big(\det \frac{M}{2\pi}\Big)^{1/2} \int_{\mathbb{R}^n} f(x) e^{-\frac{1}{2}(x-a)\cdot M(x-a)} dx \\ =& \Big(\det \frac{M}{2\pi}\Big)^{1/2} \int_{\mathbb{R}^n} f(x) e^{-\frac{1}{2}x\cdot Mx - \frac{1}{2}a\cdot Ma + x\cdot Ma} dx =: \Big(\det \frac{M}{2\pi}\Big)^{1/2} \int_{\mathbb{R}^n} f(x) e^{-\frac{1}{2}x\cdot Mx} Z dx\end{aligned}$$

where $Z := \exp(\frac{1}{2}a \cdot Ma + x \cdot Ma)$, so $\mathbb{E}(f(X+a)) = \mathbb{E}(f(X)Z)$. Thus if $\mathbb{P}$ denotes the distribution of X, then the measure $\mathbb{Q}$ with

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = Z$$

is that of a $N(0, C)$ Gaussian vector, *i.e.* of $X + a$.

Example : Let B be a standard Brownian motion with $B_0 = 0$. Fix finitely many times $0 = t_0 = t_1 < \cdots < t_n$. Then $(B_{t_i})_{i=1}^n$ is a centred Gaussian vector with

$$\mathbb{E}(f((B_{t_i})_{i=1}^n)) = \text{const.} \times \int_{\mathbb{R}^n} f(x) e^{-\frac{1}{2}\sum_{i=1}^n \frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}}} dx_1 \cdots dx_n$$

Let $h : \mathbb{R}_+ \to \mathbb{R}$ be a deterministic function. Then

$$\mathbb{E}(f((B+h)_{t_i})) = \mathbb{E}(Zf(B_{t_i}))$$

with $Z = \exp\Big(-\frac{1}{2}\sum_{i=1}^n \frac{(h_{t_i} - h_{t_{i-1}})^2}{t_i - t_{i-1}} + \sum_{i=1}^n \frac{(h_{t_i} - h_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})}{t_i - t_{i-1}}\Big)$.

The Girsanov's Theorem is essentially a generalized version of these with the summation replaced by integration with more careful treatment.

(22nd February, Friday)

Definition) Let L be a continuous local martingale with $L_0 = 0$. Then the **stochastic exponential** of L is

$$\mathcal{E}(L)_t = e^{L_t - \frac{1}{2}\langle L\rangle_t}$$

Fact : $Z = \mathcal{E}(M)$ is a continuous local martingale and it satisfies

$$dZ_t = Z_t dL_t$$

i.e. $Z_t = 1 + \int_0^t Z_s dL_s$.

proof) By Itô's formula applied to $X = L - \frac{1}{2}\langle L\rangle$ and $f(x) = e^x$,

$$dZ_t = df(X_t) = Z_t(dL_t - \frac{1}{2}d\langle L\rangle_t + \frac{1}{2}d\langle L\rangle_t) = Z_t dL_t$$

Since L is a continuous local martingale, so is $Z \cdot L$ since $Z = 1 + Z \cdot L$, hence Z is a continuous local martingale.

(End of proof) □

Theorem) *(Girsanov)* Let L be a continuous local martingale with $L_0 = 0$. Suppose that $\mathcal{E}(L)$ is a *UI(uniformly integrable)* martingale. Define a probability measure $\mathbb{Q}$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}(L)_\infty$$

If M is a continuous local martingale with respect to $\mathbb{P}$, then $\tilde{M} = M - \langle M, L\rangle$ is a continuous local martingale with respect to $\mathbb{Q}$.

Remark : The quadratic variation does not change, $\langle M\rangle = \langle \tilde{M}\rangle$. This follows, for example, from

$$\langle M\rangle_t = \lim_{n\to\infty} \sum_{i=1}^{\lfloor 2^m t\rfloor} (M_{2^{-m}i} - M_{2^{-m}(i-1)})^2 \quad \mathbb{P}\text{-a.s. along a subsequence}$$

and $\mathbb{Q}$-null sets are identical to $\mathbb{P}$-null sets, so the limit is also true in $\mathbb{Q}$-a.s. sense.

proof) Let $T_n = \inf\{t \geq 0 : |\tilde{M}_t| > n\}$. Then T_n is a stopping time, and $\mathbb{Q}(T_n \nearrow \infty) = \mathbb{P}(T_n \nearrow \infty) = 1$ by continuity of $\tilde{M}$ and that $\mathbb{P} \sim \mathbb{Q}$. Thus it suffices to show that $\tilde{M}^{T_n}$ is a continuous local martingale with respect to $\mathbb{Q}$ for all n. Let $Y_t = M_t^{T_n} - \langle M^{T_n}, L\rangle_t$ and $Z_t = \mathcal{E}(L)_t$.

♡ **Claim :** $(ZY)_t$ is a continuous local martingale with respect to $\mathbb{P}$.

: we just check using Itô formula, that

$$\begin{aligned}
d(ZY) =& Y_t dZ_t + Z_t dY_t + d\langle Z, Y\rangle_t \\
=& (M_t^{T_n} - \langle M^{T_n}, L\rangle_t) Z_t dL_t + Z_t(dM_t^{T_n} - d\langle M^{T_n}, L\rangle) + Z_t d\langle L, M^{T_n}\rangle_t \\
=& (M_t^{T_n} - \langle M^{T_n}, L\rangle_t) Z_t dL_t + Z_t dM_t^{T_n}
\end{aligned}$$

where we used $d\langle Z, Y\rangle = d\langle Z, M^{T_n}\rangle = Zd\langle L, M^{T_n}\rangle$ since $dZ = ZdL$ and $\langle Z \cdot L, M^{T_n}\rangle = Z \cdot \langle L, M^{T_n}\rangle$. Thus $d(ZY)$ is a sum of stochastic differentials with respect to local martingales. Hence ZY is a continuous local martingale.

♡ **Claim :** ZY is a uniformly integrable martingale with respect to $\mathbb{P}$.

: this follows from the fact that $Z = \mathcal{E}(L)$ is by assumption a uniformly integrable martingale, together with the fact that Y is bounded (recall, by definition that $|Y| \leq n$) (but be careful that if Z was just a uniformly integrable local martingale, then this is not necessarily true so we need a bit more). Indeed, a local martingale M is a martingale *iff*

$$\forall t, \quad \mathcal{X}_t = \{M_T : T \text{ is a stopping time with } T \leq t\} \text{ is UI}$$

: (verifying this statement was an Exercise. Or see online lecture notes p21-22) Forward implication follows from the fact that, for any $t > 0$, $\{\mathbb{E}[X_t|\mathcal{G}] : \mathcal{G} \subset \mathcal{F}_t$ a sub-σ-algebra$\}$ is uniformly integrable (see *Advanced Probability* or *Example Sheet #1*).

To see the backward implication, first note that for any bounded stopping time T, say $T \leq t_0$ a.s, the process $(X_t^T)_t$ is uniformly integrable and therefore $\mathbb{E}[X_t^T|\mathcal{F}_s] = X_s^T$ whenever $s \leq t \leq t_0$. Then the result follows from a part of the *Optional Stopping Theorem* stating that X is a martingale *iff* the process X^T is a martingale for any bounded stopping time T.

Since being UI is preserved under multiplication by bounded random variables, this implies the claim.

♡ **Claim :** Y is a martingale with respect to $\mathbb{Q}$ (recall $\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}(L)_\infty = Z_\infty$).

: First observe that

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}(Y_t - Y_s|\mathcal{F}_s) =&\mathbb{E}^{\mathbb{P}}(Z_\infty Y_t - Z_\infty Y_s|\mathcal{F}_s) \\ =&\mathbb{E}^{\mathbb{P}}\Big(\mathbb{E}^{\mathbb{P}}(Z_\infty Y_t - Z_\infty Y_s|\mathcal{F}_t)|\mathcal{F}_s\Big) = \mathbb{E}^{\mathbb{P}}(Z_t Y_t - Z_s Y_s|\mathcal{F}_s) = 0\end{aligned}$$

Since $Y = (M - \langle M, L\rangle)^{T_n}$ is a $\mathbb{Q}$-martingale and $T_n \nearrow \infty$ a.s., thus $M - \langle M, L\rangle$ is a $\mathbb{Q}$-local martingale.

Having these claims, the proof is complete.

(End of proof) □

To apply the theorem, we need to verify that the exponential martingale is uniformly integrable. One criterion for doing this is the following, yet not the most general one.

Proposition) Suppose that $\langle L\rangle$ is bounded, say $\langle L\rangle_\infty \le C$. Then $\mathcal{E}(L)$ is a UI martingale.

proof) It suffices to show that $\sup_t L_t$ has Gaussian tail :

$$\mathbb{P}\Big(\sup_{t\ge 0} L_t \ge a\Big) \le e^{-a^2/2C}$$

Indeed, then

$$\begin{aligned}\mathbb{E}\Big(\sup_t \mathcal{E}(L)_t\Big) \le& \mathbb{E}\Big(\exp(\sup_t L_t)\Big) \quad (\text{used } \mathcal{E}(L) = e^{L - \frac{1}{2}\langle L\rangle} \le e^L) \\ =&\int_0^\infty \mathbb{P}\Big(\exp(\sup L_t) \ge \lambda\Big)d\lambda \\ =&\int_0^\infty \mathbb{P}\Big(\sup_t L_t \ge \log\lambda\Big)d\lambda \\ \le& 1 + \int_1^\infty e^{-\frac{(\log\lambda)^2}{2C}}d\lambda < \infty\end{aligned}$$

so $\mathcal{E}(L)_t$ is bounded by the random variable $\sup_t \mathcal{E}(L)_t \in L^1$ and thus $\mathcal{E}(L)$ is uniformly integrable.

(Checking the tail bound is an exercise, on the Example Sheet)

♡ **Claim :** Let M be a continuous local martingale with $M_0 = 0$. Then for any $a, b > 0$,

$$\mathbb{P}\Big(\sup_{t\ge 0} M_t \ge a, \langle M\rangle_\infty \le b\Big) \le \exp\Big(-\frac{a^2}{2b}\Big)$$

: Let $T = \inf\{t \ge 0 : M_t = a\}$, then $\{M_\infty^T = a\} \supset \{\sup_t M_t > a\}$. Fix $\lambda \in \mathbb{R}_{>0}$, and let

$$Z_t = \exp\Big(\lambda M_t^T - \frac{1}{2}\lambda^2\langle M\rangle_t^T\Big)$$

By *Ito's formula*, Z is a continuous local martingale and $Z_t \le e^{\lambda a}$ for all $t \ge 0$, so is bounded. Hence, Z is in fact a true martingale. Now

$$\begin{aligned} 1 = Z_0 = \mathbb{E}Z_\infty &\ge \mathbb{E}(Z_\infty \mathbf{1}_{\sup_t M_t > a, \langle M \rangle_\infty \le b}) \\ &\ge \mathbb{E}(\exp(\lambda a - \frac{1}{2}\lambda^2 b)\mathbf{1}_{\sup_t M_t > a, \langle M \rangle_\infty \le b}) \end{aligned}$$

since upon the event $\{\sup_t M_t > a\}$, we always have $M^T_\infty = a$. Hence

$$\mathbb{P}(\sup_t M_t > a, \langle M \rangle_\infty \le b) \le e^{-\lambda a + \frac{1}{2}\lambda^2 b}$$

Optimizing over λ, we get

$$\mathbb{P}(\sup M_t > a, \langle M \rangle_\infty \le b) \le e - \frac{a^2}{2b}$$

and

$$\mathbb{P}(\sup_t M_t \ge a, \langle M \rangle_\infty \le b) \le \inf\{\mathbb{P}(\sup_t M_t > a', \langle M \rangle_\infty \le b); a' < a\} \le e^{-a^2/2b}$$

as desired.

(End of proof) $\square$

In the proof, we had to check tedious details using techniques of measure theory and martingale theory, but the real core of the proof of the theorem in fact comes from the examples displayed at the end of the last lecture. We just need to see the finite dimensional cases to see how the proof in general works.

(25th February, Monday)

(A comment from last lecture) **Proposition)** Let M be a continuous local martingale with $M_0 = 0$. Then $M \in M^2_c$ *iff* $\mathbb{E}\langle M \rangle_\infty < \infty$ and then $M^2 - \langle M \rangle$ is a UI martingale and $\|M\|_{M^2} = (\mathbb{E}\langle M \rangle_\infty)^{1/2}$.

proof) See *Example Sheet #1.*

There is a more general criterion for $\mathcal{E}(M)$ to be uniformly integrable.

Theorem) *(Novikov)* Let M be a continuous local martingale with $M_0 = 0$. Then $\mathbb{E}(e^{\frac{1}{2}\langle M \rangle_\infty}) < \infty$ implies that $\mathcal{E}(M)$ is a UI martingale.

Corollary) *(corollary of Girsanov's Theorem)* Let B be a standard Brownian motion (under $\mathbb{P}$) and let L be a continuous local martingale with $L_0 = 0$ such that $\mathcal{E}(L)$ is a UI martingale. Then $\tilde{B} = B - \langle B, L \rangle$ is a standard Brownian motion under the measure $\mathbb{Q}$ where

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \mathcal{E}(L)_\infty$$

proof) By *Girsanov's Theorem*, $\tilde{B}$ is a continuous local martingale. Moreover, $\langle \tilde{B} \rangle_t = \langle B \rangle_t = t$. By *Lévy's characterisation*, $\tilde{B}$ is a standard Brownian motion.

(End of proof) $\square$

Why is this useful? Consider the following (informal) example.

Example : Consider the SDE (this is yet to be defined), for a fixed time $T < \infty$,

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

We can construct a solution as follows. Let X be a standard Brownian motion under $\mathbb{P}$. Let

$$L_t = \int_0^{t\wedge T} b(X_s)dX_s$$

Assume that $\mathcal{E}(L)$ is a UI martingale. Then

$$X_t - \langle X, L\rangle_t = X_t - \int_0^{t\wedge T} b(X_s)d\langle X\rangle_s = X_t - \int_0^{t\wedge T} b(X_s)ds$$

is a standard Brownian motion under $\mathbb{Q}$ given by $d\mathbb{Q}/d\mathbb{P} = \mathcal{E}(L)_\infty$. Thus if we call this Brownian motion $\tilde{B}$ then the last equation is written by

$$X_t - \int_0^{t\wedge T} b(X_s)ds = \tilde{B}_t$$

So instead of solving the equation $dX_t = b(X_t)dt + dB_t$, we can just start with a Brownian motion in the changed measure $\mathbb{Q}$ and change the measure back to $\mathbb{P}$ to see the distribution of X.

When is $\mathcal{E}(L)$ is a UI martingale? We have

$$\langle L\rangle_\infty = \int_0^T b(X_s)^2 ds$$

So it is sufficient if b is a bounded function.

4.4 The Cameron-Martin formuala

This section is about a different application of Girsanov's Theorem.

Definition) The **(classical/canonical) Wiener space** $(W, \mathcal{W}, P)$ is given by $W = C(\mathbb{R},\mathbb{R})$, $\mathcal{W} = \sigma(X_t : t \geq 0)$ where $X_t : W \to \mathbb{R}$ is given by $X_t(w) = w(t)$ for $w \in W$ and P is the unique probability measure on $(W, \mathcal{W})$ such that (X_t) is a standard Brownian motion. X is also called the **canonical version of Brownian motion**. (This is a Banach space)

Definition) The **Cameron-Martin space** is

$$\mathcal{H} = \{h \in W : h(t) = \int_0^t g(s)ds \quad \text{for some } g \in L^2(\mathbb{R}_+)\}$$

For $h \in \mathcal{H}$, the function $\dot{h} = g$ is the weak derivative of g.

Exercise : $\mathcal{H}$ is a Hilbert space with inner product

$$(h, f)_{\mathcal{H}} = \int_0^\infty \dot{h}(s)\dot{f}(s)ds$$

The dual space of $\mathcal{H}$ can be identified with

$$\mathcal{H}^* = \{\mu \in \mathcal{M}(\mathbb{R}_+) : \int_0^\infty (s\wedge t)\mu(ds)\mu(dt) = (\mu,\mu)_{\mathcal{H}^*} < \infty, \mu(\{0\}) = 0\}$$

in the sense that for any $l : \mathcal{H} \to \mathbb{R}$ bounded and linear, there is $\mu \in \mathcal{H}$ such that $l(h) = \int_0^\infty h(t)\mu(dt)$ and vice-versa.

Remark : We would like to think of a Brownian motion as the standard Gaussian measure on $\mathcal{H}$. This measure does not exist. But the next theorem shows it almost does.

Theorem) *(Cameron-Martin)* Let $h \in \mathcal{H}$ and define P^h by (P^h is going to be a canonical measure on the Wiener space)

$$P^h(A) = P\Big(\{w \in W : w + h \in A\}\Big)$$

for $A \in \mathcal{W}$. Then the measure P^h is absolutely continuous with respect to the Wiener measure P and

$$\frac{dP^h}{dP} = \exp\left(\int_0^\infty \dot{h}(s)dX_s - \frac{1}{2}\int_0^\infty \dot{h}(s)^2 ds\right)$$

proof) Apply *Girsanov's Theorem* with $L_t = \int_0^t \dot{h}(s)dX_s$. Since $\langle L\rangle_\infty = \int_0^\infty \dot{h}^2 ds = \|h\|_{\mathcal{H}}^2 < \infty$. $\mathcal{E}(L)$ is a UI martingale. Then

$$\mathcal{E}(L)_\infty = \exp\left(\int_0^\infty \dot{h}(s)dX_s - \frac{1}{2}\int_0^\infty \dot{h}(s)^2 ds\right)$$

so the rest is as in previous examples.

(End of proof) $\square$

(See a reference book for deeper analytical viewpoint)

5 Stochastic Differential Equations

5.1 Notions of Solutions

In *Section 1*, we considered the SDE(Stochastic Differntial Equation) $\dot{x}(t) = F(x(t)) + \eta(t)$, where η is a white noise. Since the integral of white noise should be interpreted as a Brownian motion, it is reasonable to interpret this SDE as

$$X_t - X_0 = \int_0^t F(X_s)ds + B_t \quad \Leftrightarrow \quad dX_t = F(X_t)dt + dB_t$$

Note that $\Leftrightarrow$ holds because it is defined as a notation.

(27th February, Wednesday)

Definition) Let $d, m \in \mathbb{N}$, $B : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be locally bounded Borel functions. The **stochastic diffrential equation (SDE)** $E(\sigma, b)$ is

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

The SDE $E(\sigma, b)$ together with the initial condition $X_0 = x \in \mathbb{R}^d$ is denoted $E_x(\sigma, b)$.

Definition) A **(weak) solution** to the SDE $E(\sigma, b)$ consists of

- a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ obeying the *usual conditions*;

- an m-dimensional $(\mathcal{F}_t)$-Brownian motion B;
- an $(\mathcal{F}_t)$-adapted continuous process X with values in $\mathbb{R}^d$ such that

$$X_t = X_0 + \int_0^t \sigma(s, X_s)dB_s + \int_0^t b(s, X_s)ds.$$

Definition) For a **strong solution** to $E(\sigma, b)$, we specify the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a Brownian motion B, and choose $(\mathcal{F}_t)_t$ to be the *completed filtration* induced by B. A strong solution is an $(\mathcal{F}_t)$-adapted continuous process X as in the definition of weak solution. That is, it satisfies

$$X_t = X_0 + \int_0^t \sigma(s, X_s)dB_s + \int_0^t b(s, X_s)ds.$$

This means one can think of a strong solution as a function of the Brownian motion.

The main difference in strong solution from weak solution is that the filtration and the Brownian motion is given *a priori*. As the names suggest, weak solutions are easier to find but strong solutions have better properties. They are both useful.

Definition) For the SDE $E(\sigma, b)$ we say that there is

- **weak uniqueness** or **uniqueness in law** if all solutions to $E_x(\sigma, b)$ has the same law.
- **pathwise uniqueness** if, for $(\Omega, \mathcal{F}, \mathcal{F}_t, \Omega, \mathbb{P})$ and B fixed, all solutions with the same initial conditions are indistinguishable. *[This does not mean that the filtration and B are given as in the definition of a strong solution. This just means that if two weak solutions have the same attached filtration and Brownian motion, then they are indistinguishable.]*

Pathwise uniqueness implies weak uniqueness, but this is not obvious (called Yamada-Watanabe theorem, see below).

Example : *(Tanaka)* The SDE

$$dX_t = \text{sign}(X_t)dB_t, \quad X_0 = x \quad \cdots\cdots\cdots \text{(TK)}$$

where $\text{sign}(x) = 1$ if $x > 0$, $\text{sign}(x) = -1$ if $x \leq 0$, has a weak solution that is unique in law, but it is pathwise uniqueness does not hold.

proof) Let X be a one-dimensional Brownian motion with $X_0 = x$. Set $\tilde{B}_t = \int_0^t \text{sign}(X_s)dX_s$ then

$$x + \int_0^t \text{sign}(X_s)d\tilde{B}_s = x + \int_0^t \text{sign}(X_s)^2 dX_s = x + (X_t - X_0) = X_t$$

so $dX_t = \text{sign}(X_t)d\tilde{B}_t$ with $X_0 = x$.

Moreover, $\tilde{B}$ is a Brownian motion since it is a continuous local martingale with

$$\langle \tilde{B} \rangle_t = \int_0^t \text{sign}(X_t)^2 d\langle X \rangle_t = \langle X \rangle_t = t,$$

so $\tilde{B}$ and X are both standard Brownian motion by *Lévy's characterisation*.

By the same argument, in fact, any solution is a standard Brownian motion. Therefore weak uniqueness holds.

To show that pathwise uniqueness fails, at least when $x = 0$, we will show that if X is a solution with $X_0 = 0$ then $-X$ is a solution with the same Brownian motion B. Indeed,

$$-X_t = -\int_0^t \operatorname{sign}(X_s)dB_s = \int_0^t \operatorname{sign}(-X_s)dB_s + 2\int_0^t 1_{X_s=0}dB_s$$

Let $N_t = 2\int_0^t 1_{X_s=0}dB_s$. It will be sufficient to prove that $N = 0$ to see that $-X_t$ is also a solution.

♡ **Claim :** N is indistinguishable from 0.

: Clearly, N is a continuous local martingale and $\langle N\rangle_t = 4\int_0^t 1_{X_s=0}ds = 0$ a.s. since the zero set of Brownian motion has Lebesgue measure 0 a.s.(*why?*). So $N = 0$ up to indistinguishability.

So $-X$ also solves (TK).

(End of proof) □

Remark : X above is not a strong solution.

Theorem) *(Pathwise uniqueness for SDEs with Lipschitz coefficients)* Suppose that b and σ are *locally Lipschitz* (in space variable), *i.e.*, for each $n > 0$, there exists $K_n > 0$ such that for all $|x|, |y| \le n$, $t \ge 0$, has

$$\begin{aligned} &|b(t,x) - b(t,y)| \le K_n|x-y| \quad \text{and} \\ &|\sigma(t,x) - \sigma(t,y)| \le K_n|x-y|. \end{aligned}$$

Then pathwise uniqueness holds for $E(\sigma, b)$.

proof) Let X and X' be two solutions to $E(\sigma, b)$ defined on the same probability space such that $X_0 = X'_0$. Let $T_n = \inf\{t \ge 0 : |X_t| > n \text{ or } |X'_t| > n\}$ and

$$f_n(t) = \mathbb{E}\Big(|X_{t\wedge T_n} - X'_{t\wedge T_n}|^2\Big)$$

By continuity of X and X', it suffices to show that for all n and all t, one has $f_n(t) = 0$.

♡ **Claim :** $f_n(t) = 0$ for all $n, t > 0$.

: By *Itô's formula,*

$$\begin{aligned} |X_{t\wedge T_n} - X'_{t\wedge T_n}|^2 = &\int_0^{t\wedge T_n} 2(X_s - X'_s)\cdot(b(X_s) - b(X'_s))ds \\ &+ \int_0^{t\wedge T_n} 2(X_s - X'_s)\cdot(\sigma(X_s) - \sigma(X'_s))dB_s + \int_0^{t\wedge T_n} |\sigma(X_s) - \sigma(X'_s)|^2 ds \end{aligned}$$

Since $(X_s - X'_s)(\sigma(X_s) - \sigma(X'_s))$ is bounded on $s < T_n$, the second term is a martingale with expectation 0, so

$$\begin{aligned} &\mathbb{E}\Big(|X_{t\wedge T_n} - X'_{t\wedge_n}|^2\Big) \le (2K_n + K_n^2)\int_0^t \mathbb{E}\Big(|X_{s\wedge T_n} - X'_{s\wedge T_n}|^2\Big)ds \\ \Rightarrow \quad & f_n(t) \le (2K_n + K_n^2)\int_0^t f_n(s)ds \end{aligned}$$

By *Grönwall's inequality* (see below), we have $f_n(t) \le f_n(0)e^{(2K_n+K_n^2)t} = 0$.

(End of proof) □

(1st March, Friday)

Gronwall's Lemma) (on *Example Sheet #3*) Let $T > 0$ and let $f : [0,T] \to \mathbb{R}$ be non-negative *bounded* Borel function. Assume $f(t) \le a + b\int_0^t f(s)ds$ for all $t \le T$. Then

$$f(t) \le ae^{bt} \quad \text{for all } t \le T$$

Remark : The proof of the uniqueness proves (for locally Lipschitz coefficients) processes defined up to the time T must agree up to time T.

5.2 Strong existence for Lipschitz coefficients

Recall, we denote $E(\sigma, b)$ for $dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t$.

Theorem) Assume b and σ are globally Lipschitz, *i.e.* there is $K > 0$ such that for all $x, y \in \mathbb{R}^d$, $t \ge 0$,

$$|b(t,x) - b(t,y)| \le K|x-y|, \quad |\sigma(t,x) - \sigma(t,y)| \le K|x-y|$$

For any $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ (obeying usual condition), any $(\mathcal{F}_t)$-Brownian motion B, any $x \in \mathbb{R}$, there is a unique strong solution to $E_x(\sigma, b)$.

proof) To simplify notation, we assume $d = m = 1$. Define

$$F(X)_t = x + \int_0^t \sigma(s, X_s)dB_s + \int_0^t b(s, X_s)ds$$

Then X is a strong solution to $E_x(\sigma, b)$ if $F(X) = X$. To find such a fixed point, we use *Picard iteration.* Fix $T > 0$. For X continuous, adapted process, set

$$|||X|||_T = \mathbb{E}\Big(\sup_{t\in[0,T]} |X_t|^2\Big)^{1/2}$$

Then $B = \{X : \Omega \times [0,T] \to \mathbb{R}, |||X|||_T < \infty\}$ is a Banach space.

♣ **Claim :** $|||F(X) - F(Y)|||_T^2 \le (2T+8)K^2 \int_0^T |||X - Y|||_t^2\, dt$.

: Just estimate $|||F(X) - F(Y)|||_T$,

$$|||F(X) - F(Y)|||_T^2 \le 2\mathbb{E}\Big(\sup_{t\le T}\Big|\int_0^t \big(b(s,X_s) - b(s,Y_s)\big)ds\Big|^2\Big)$$
$$+ 2\mathbb{E}\Big(\sup_{t\le T}\Big|\int_0^t \big(\sigma(s,X_s) - \sigma(s,Y_s)\big)dB_s\Big|^2\Big) =: \text{(A)+(B)}$$

Then

$$\begin{aligned}\text{(A)} &\le 2T\mathbb{E}\Big(\sup_{t\le T}\int_0^t |b(s,X_s) - b(s,Y_s)|^2 ds\Big) \quad \text{(Cauchy-Schwarz)}\\ &\le 2TK^2 \int_0^T |||X - Y|||_T^2\, dt\end{aligned}$$

and

$$\begin{aligned}\text{(B)} &\le 8\mathbb{E}\Big(\int_0^T |\sigma(s,X_s) - \sigma(s,Y_s)|^2 ds\Big) \quad \text{(Doob's } L^2\text{)}\\ &\le 8K^2 \int_0^T |||X - Y|||_t\, dt\end{aligned}$$

♣ **Claim :** $\left|\!\left|\!\left| F(0) \right|\!\right|\!\right|_T < \infty$.

: the argument is the same as putting 0 for $F(Y)$ above,

$$F(0)_t = x + \int_0^t b(s,0)ds + \int_0^t \sigma(s,0)dB_s$$

$$\Rightarrow \quad \left|\!\left|\!\left| F(0) \right|\!\right|\!\right|_T^2 \le 3\Big(|x|^2 + \left|\!\left|\!\left| \int_0^t b(s,0)ds \right|\!\right|\!\right|_T^2 + \left|\!\left|\!\left| \int_0^t \sigma(s,0)dB_s \right|\!\right|\!\right|_T^2 \Big)$$

$$\le 3\Big(|x|^2 + T\int_0^T b(s,0)^2 ds + 4\int_0^T |\sigma(s,0)|^2 ds \Big) < +\infty$$

Now use *Picard iteration* : Let $X_t^0 = 0$ for all t and set $X^{i+1} = F(X^i)$ so by the first claim, for some $C > 0$,

$$\left|\!\left|\!\left| X^{i+1} - X^i \right|\!\right|\!\right|_T^2 \le CT\int_0^T \left|\!\left|\!\left| X^i - X^{i-1} \right|\!\right|\!\right|_t^2 dt \le (CT)^2 \int_0^T \int_0^t \left|\!\left|\!\left| X^{i-1} - X^{i-2} \right|\!\right|\!\right|_s ds$$

$$\le \cdots\cdots \le \frac{(CT)^i}{i!} \left|\!\left|\!\left| X^1 - X^0 \right|\!\right|\!\right|_T = \frac{(CT)^i}{i!} \left|\!\left|\!\left| F(0) \right|\!\right|\!\right|_T$$

Therefore,

$$\sum_{i=1}^{\infty} \left|\!\left|\!\left| X^i - X^{i-1} \right|\!\right|\!\right|_T^2 < \infty \quad \forall T > 0$$

and hence X^i converges uniformly on $[0,T]$ a.s. for all $T > 0$. Say X is the limit. Then $F(X) = X$.

By uniqueness, solutions up to different times T must agree when both are defined. Hence we can extend them to all of $[0,\infty)$.

(End of proof) □

The following proposition provides a (rough) estimate on the dependence of the solution on the initial condition.

Proposition) Under the assumptions of the theorem, let X^x be the solution with initial condition $X_0^x = x$. Then for any $p \ge 2$,

$$\mathbb{E}\Big(\sup_{s\le t} |X_t^x - X_s^y|^p \Big) \le C_p |x-y|^p e^{C_p (t\vee 1)^p t}$$

(4th March, Monday)

To prove the proposition, we need :

Lemma) *(Burkholder-Davis-Gundy (BDG) inequality)* For every real $p > 0$, there exists $C_p > 0$ depending only on p such that, for every continuous local martingale M with $M_0 = 0$ and every stopping time T,

$$\mathbb{E}[\sup_{0\le s\le T} (M^s)^p] \le C_p \mathbb{E}[\langle M \rangle_T^{p/2}]$$

For proof for the case $p \ge 2$, see *Example sheet #2*. For the case $p < 2$, a good reference would be "Brownian motions, Martingales, and Stochastic Calculus" by Jean-Francois Le Gall.

Note that we will only use the $p \ge 2$ case.

Proposition) Under the assumptions as in the theorem, let X^x be the solution with initial condition $X_0^x = x \in \mathbb{R}^d$. That is,

$$X_t^x = x + \int_0^t \sigma(r, X_r^x) dB_r + \int_0^t b(r, X_r^x) dx.$$

Then for $p \geq 2$,

$$\mathbb{E} \sup_{s \leq t} |X_s^x - X_s^y|^p \leq C_p |x-y|^p \exp(C_p (t \vee 1)^p t)$$

proof) For simplicity, assume $d = m = 1$. Fix $x, y \mathbb{R}^d$ and let $T_n = \inf\{t \geq 0 : |X_t^x| > n \text{ or } |X_t^y| > n\}$. Since $|a+b+c|^p \leq 3^{p-1}(|a|^p + |b|^p + |c|^p)$,

$$\begin{aligned}
\mathbb{E}\Big(\sup_{s \leq t} |X_{s\wedge T_n}^x - X_{s \wedge T_n}^y|^p\Big) &\leq 3^{p-1}\bigg[|x-y|^p + \mathbb{E}\Big(\sup_{s \leq t}\Big|\int_0^{s \wedge T_n} \Big(\sigma(r, X_r^x) - \sigma(r, X_r^y)\Big) dB_r\Big|^p\Big) \\
&\quad + \mathbb{E}\Big(\sup_{s \leq t}\Big|\int_0^{s \wedge T_n} \Big(b(r, X_r^x) - b(r, X_r^y)\Big) dr\Big|^p\Big)\bigg] \\
&=: 3^{p-1}(|x-y|^p + \text{(A)} + \text{(B)})
\end{aligned}$$

By *BDG inequality*,

$$\begin{aligned}
\text{(A)} &\leq C_p \mathbb{E}\Big(\int_0^{t \wedge T_n} \Big(\sigma(r, X_r^x) - \sigma(r, X_r^y)\Big)^2 dr\Big)^{p/2} \\
&\leq C_p t^{\frac{p}{2}-1} \mathbb{E}\Big(\int_0^{t \wedge T_n} \Big|\sigma(r, X_r^x) - \sigma(r, X_r^y)\Big|^p dr\Big) \quad \text{(Hölder)} \\
&\leq C_p t^{\frac{p}{2}-1} \int_0^t \mathbb{E}\Big|\sigma(t \wedge T_n, X_{r \wedge T_n}^x) - \sigma(t \wedge T_n, X_{r \wedge T_n}^y)\Big|^p dr \\
&\leq C_p t^{\frac{p}{2}-1} \int_0^t \mathbb{E}\Big(K|X_{r\wedge T_n}^x - X_{r \wedge T_n}^y|^p\Big) dr
\end{aligned}$$

Also, only using Hölder,

$$\begin{aligned}
\text{(B)} &\leq t^{p-1} \mathbb{E}\Big(\int_0^t \Big|b(r \wedge T_n, X_{r \wedge T_n}^x) - b(r \wedge T_n, X_{r \wedge T_n}^y)\Big|^p dr\Big) \\
&\leq t^{p-1} \mathbb{E} \int_0^t K|X_{r \wedge T_n}^x - X_{r \wedge T_n}^y|^p dr
\end{aligned}$$

In summary,

$$f_n(t) := \mathbb{E}(\sup_{s \leq t} |X_{s \wedge T_n}^x - X_{s \wedge T_n}^y|^p) \leq 3^{p-1}|x-y|^p + \tilde{C}_p (t \vee 1)^{p/2} t^{\frac{p}{2}} \int_0^t \mathbb{E}(\sup_{s \leq t} |X_{s \wedge T_n}^x - X_{s \wedge T_n}^y|^p) dr$$

Note that f_n is bounded on any interval $[0, T]$ for $n < \infty$. By *Grönwall's inequality*,

$$f_n(t) \leq 3^{p-1}|x-y|^p \exp(\tilde{C}_p (t \vee 1)^{p/2} t^{\frac{p}{2}+1})$$

By Fatou, taking $n \to \infty$, we obtain the claimed inequality.

(End of proof) $\square$

Strong solution can be considered functions of Brownian motion in the following sense. Recall the *(d-dimensional) Wiener space* $(W^d, \mathcal{W}^d, P^d)$ where

$$W^d = C(\mathbb{R}_+, \mathbb{R}^d), \quad \mathcal{W} = \sigma(X_t^i :\in \mathbb{R}_+, i = 1, \cdots, d), \quad \text{where } X_t(w) = w(t) \text{ for } w \in W^d$$

and P^d is the probability measure on $(W^d, \mathcal{W}^d)$ such that $(X_t)_{t\geq 0}$ is a standard Brownian motion with $X_0 = 0$.

The space $C(\mathbb{R}_+, \mathbb{R}^d)$ can be given the topology of uniform convergence on compact intervals. This topology is induced by the metric

$$d(w, \tilde{w}) = \sum_{k=1}^{\infty} \alpha_k (\|w - \tilde{w}\|_{L^\infty([0,t];\mathbb{R}^d)} \wedge 1)$$

for any sequence $(\alpha_k) \subset \mathbb{R}_+$ with $\sum_{k=1}^{\infty} \alpha_k = 1$.

Remark : This metric makes $C(\mathbb{R}_+\mathbb{R}^d)$ a complete separable metric space (a so called *Polish space*).

Theorem) Under the assumptions of the last theorem (strong solution for Lipschitz coefficients), for $x \in \mathbb{R}^d$, there exists maps

$$F_x : W^m = C(\mathbb{R}_+, \mathbb{R}^m) \to W^d = C(\mathbb{R}_+, \mathbb{R}^d)$$

measurable with respect to the completion of $\mathcal{W}^m$ on W^m and w.r.t. $\mathcal{W}^d$ on W^d such that

(i) $\forall t \geq 0$, $F_x(w)_t$ is a measurable function of $\sigma(w(s) : s \leq t)$ for P^d-a.s. $w \in W^m$.

(ii) $\forall w \in C(\mathbb{R}_+, \mathbb{R}^m) : x \in \mathbb{R}^d \mapsto F_x(w) \in C(\mathbb{R}_+, \mathbb{R}^d)$ is continuous.

(iii) $\forall x \in \mathbb{R}^d$, $\forall (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions, every $(\mathcal{F}_t)$-Brownian motion $\hat{B}$ with $\hat{B}_0 = 0$, the unique solution to $E_x(\sigma, b)$ is $\hat{X}_t = F_x(\hat{B})_t$.

(iv) In the set-up of (iii), if U is $\mathcal{F}_0$-measurable, then $F_U(\hat{B})_t$ is the unique solution to $E(\sigma, b)$ with $X_0 = U$.

(Such F is called the **Itô map**.)

proof) For simplicity, assume $d = m = 1$ in the notation. Let

$$\mathcal{G}_t = \sigma(w(s) : 0 \leq s \leq t) \vee \mathcal{N}, \quad \mathcal{G} = \mathcal{G}_\infty$$

where $\mathcal{N}$ are the P-null sets. Then by the existence theorem applied to $(W^n, \mathcal{G}, (\mathcal{G}_t), P^d)$, $B_t(w) = w(t)$, there is a unique solution X^x to $E_x(\sigma, b)$.

Let $p \geq 2$, that is to be specified later. By the last proposition and $d(w, \tilde{w}) = \sum_k \alpha_k (\sup_{s\leq k} |w(s) - \tilde{w}(s)| \wedge 1)$ with (α_k) to be chosen,

$$\begin{aligned}
\mathbb{E}\Big(d(X^x, X^y)^p\Big) &\leq \mathbb{E}\Big(\Big(\sum_k \alpha_k \sup_{s\leq k} |X_s^x - X_s^y|\Big)^p\Big) \\
&\leq \sum_{k=1}^{\infty} \alpha_k \mathbb{E}\Big(\sup_{s\leq k} |X_s^x - X_s^y|^p\Big) \quad \text{(Jensen)} \\
&\leq C_p |x - y|^p \sum_k \alpha_k \exp(C_p k^{p+1}) \\
&\leq \tilde{C}_p |x - y|^p \quad \cdots\cdots\cdots \ (\dagger)
\end{aligned}$$

where the last inequality follows by choosing (α_k) appropriately. A version of *Kolmogorov's continuity criterion* applies to processes in a complete metric space indexed by $\mathbb{R}^d$ if $(\dagger)$ holds with $p > d$. Applying this to $(X^x, x \in \mathbb{R}^d)$, there is a modification $(\tilde{X}^x, x \in \mathbb{R}^d)$ that is continuous in $x \in \mathbb{R}^d$. We set $F_x(w) = \tilde{X}^x(w) = (\tilde{X}_t^x(w))_{t\geq 0}$.

(6th March, Wednesday)

Last time : X^x was strong solution with respect to filtration induced by B on the canonical space and $\tilde{X}^x$ was a continuous (in x) modification by Kolmogorov continuity theorem, and let $F_x(w) = \tilde{X}^x(w)$ for $w \in C(\mathbb{R}_+, \mathbb{R}^m)$.

proof continued) The construction from last lecture proves point (ii).

For (i), we observe $w \mapsto F_x(w)_t = \tilde{X}^x_t(w)$. We have $\tilde{X}^x_t(w) = X^x_t(x)$ for P-a.e. w. But X^x_t measurable with respect to $\sigma(w(s) : s \leq t)$ completed by null sets. Hence $\tilde{X}^x_t$ also is $\mathcal{G}_t$-measurable (recall, $\mathcal{G}_t$ was the completed filtration of the Brownian motion).

To show (iii), fix $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ and $\hat{B}$, we set

$$\hat{X} = F_x(\hat{B})_t$$

Since F_x maps into $C(\mathbb{R}_+, \mathbb{R}^d)$, $\hat{X}$ is continuous in t. Since $F_x(\hat{B})_t$ coincides a.s. with a measurable function of $(\hat{B}_t : 0 \leq s \leq t)$ and $(\mathcal{F}_t)$ in complete it follows that $\hat{X}$ is adapted. By definition,

$$\begin{aligned}
\tilde{X}_t &= x + \int_0^t \sigma(s, \tilde{X}_s) dB_s + \int_0^t b(s, \tilde{X}_s) ds \\
&= x + \lim_{m\to\infty} \sum_{i=1}^{\lfloor 2^m t \rfloor} \sigma(s, \tilde{X}_{(i-1)2^{-m}})(B_{i2^{-m}} - B_{(i-1)2^{-m}}) + \int_0^\infty b(s, \tilde{X}_s) ds
\end{aligned}$$

Since the limit converges in P^m-probability, we also have convergence P^m-a.s. along a subsequence. Since $F_x(w)_t = \tilde{X}^x_t(w)$ along this subsequence,

$$\begin{aligned}
F_x(w)_t = x + \lim_{m\to\infty} &\sum_{m=1}^{\lfloor 2^m t \rfloor} \sigma(s, F_x(w)_{(i-1)2^{-m}})(w(i2^{-m}) - w((i-1)2^{-m})) \\
&+ \int_0^t b(s, F_x(w)_s) ds \quad \text{for } P^m\text{-a.s. } w \in W^m
\end{aligned}$$

Since $\hat{B}$ has law P^m on W^m, we can substitute $w = \hat{B}$ and then revert the approximation of the stochastic integral to get

$$\hat{X}_t = x + \int_0^t \sigma(s, \hat{X}_s) d\hat{B}_s + \int_0^t b(s, \hat{X}_s) ds$$

as desired. Refer to a reference for the proof of point (iv).

(End of proof) □

[Note that, this technical procedure is necessary because we are changing from one probability space to the other. Since stochastic integrals not only refers to a path of a Brownian motion but also refers to a larger part of the probability space (as we consider convergence in the construction of integrals), we have to be careful when we are changing the probability space.]

Corollary) The solutions to $E_x(\sigma, b)$ can be constructed for all $x \in \mathbb{R}^d$ simultaneously such that a.s. X^x is continuous in the initial condition.

proof) Direct from the theorem.

5.3 Some examples of SDEs

Geometric Brownian motion

For $w \in C(\mathbb{R}_+, \mathbb{R})$, define $F_x(w)$ by

$$[F_x(w)](t) = x \exp\Big(\sigma(w(t)) + (\mu - \frac{\sigma^2}{2})t\Big)$$

If B is a standard Brownian motion with $B_0 = 0$, then $X_t = F_x(B)_t$ satisfies

$$dX_t = \sigma X_t dB_t + \mu X_t dt, \quad X_0 = x \quad \cdots\cdots\cdots \ (*)$$

On the other hand, if we choose w to be any smooth path, then $x_t = F_x(w)_t$ satisfies the ODE

$$dx_t = \sigma x_t dw_t + x_t\Big(\mu - \frac{\sigma^2}{2}\Big)dt, \quad \text{where } x_0 = x$$

Thus the *Itô map* F satisfies the 'wrong' equation on smooth paths! The process solving $(*)$ is called *Geometric Brownian motion.*

The Ornstein-Uhlenbeck process

Let $\lambda > 0$. The *Ornstein-Uhlenbeck process* is the (unique) solution to

$$dX_t = -\lambda X_t dt + dB_t$$

To solve this SDE, apply *Itô's formula* to $e^{\lambda t} X_t$:

$$\begin{aligned} & d(e^{\lambda t} X_t) = e^{\lambda t} dX_t + \lambda e^{\lambda t} X_t dt = e^{\lambda t} dB_t \\ \Leftrightarrow \quad & e^{\lambda t} X_t - X_0 = \int_0^t e^{\lambda s} dB_s \\ \Leftrightarrow \quad & X_t = e^{-\lambda t} X_0 + \int_0^t e^{-\lambda(t-s)} D_b \end{aligned}$$

The last term $\int_0^t e^{-\lambda(t-s)} dB_s$ integrates a deterministic function in the Brownian motion, so it can be thought as a Wiener integral.

Fact : If $X_0 = x$ is fixed (or if X_0 is Gaussian), then (X_t) is a *Gaussian process*, *i.e.* $(X_{t_i})_{i=1}^n$ is jointly Gaussian, for $0 = t_0 < t_1 < \cdots < t_n$.

proof) Exercise - use of Wiener integral simplifies the proof.

A Gaussian process is determined by its mean and its covariance.

Fact : If $X_0 = x$, then $\mathbb{E}(X_t) = e^{-\lambda t} x$, $\mathrm{Cov}(X_t, X_s) = \frac{1}{2\lambda}(e^{-\lambda|t-s|} - e^{-\lambda|t+s|})$.

proof) Clearly, $\mathbb{E}X_t = e^{-\lambda t}\mathbb{E}X_0 + \mathbb{E}\int_0^t e^{-\lambda(t-s)} dB_s = e^{-\lambda t}\mathbb{E}X_0$. Also, by *Itô isometry*,

$$\begin{aligned} \mathrm{Cov}(X_t, X_s) &= \mathbb{E}\Big((X_t - \mathbb{E}X_t)(X_s - \mathbb{E}X_s)\Big) \\ &= \mathbb{E}\Big(\int_0^t e^{-\lambda(t-u)} dB_u \int_0^s e^{-\lambda(s-u)} dB_u\Big) \\ &= \int_0^\infty \mathbf{1}_{u<t} e^{-\lambda(t-u)} \mathbf{1}_{u<s} e^{-\lambda(s-u)} du \\ &= e^{-\lambda(t+s)} \int_0^{t+s} e^{2\lambda u} du = \frac{1}{2\lambda} e^{-\lambda(t+s)} (e^{2\lambda(s\wedge t)} - 1) \end{aligned}$$

(End of proof) □

Corollary) $X_t \sim N(e^{-\lambda t}x, \frac{1-e^{-2\lambda t}}{2\lambda})$

Fact : If $X_0 \sim N(0, \frac{1}{\lambda})$, then $X_t \sim N(0, \frac{1}{2\lambda})$ for all $t > 0$, and X_t is a *stationary* Gaussian process with $\mathrm{Cov}(X_s, X_t) = \frac{1}{2\lambda}e^{-\lambda|t-s|}$

(8th March, Friday)

5.4 Local Solutions

Proposition) *(Local Itô formula)* Let $X = (X^1, \cdots, X^d)$ be semimartingales. Let $U \subset \mathbb{R}^d$ be open, and let $f : U \to \mathbb{R}^d$ be C^2. Set $T = \inf\{t \geq 0 : X_t \notin U\}$. Then for all $t < T$,

$$f(X_t) = f(X_0) + \sum_{i=1}^{d} \frac{\partial f}{\partial x^i}(X_s)dX_s^i + \frac{1}{2}\frac{\partial^2 f}{\partial x^i \partial x^j}(X_s)d\langle X^i, X^j\rangle_s$$

proof) Apply Itô's formula to X^{T_n} where $T_n = \inf\{t \geq 0 : \mathrm{dist}(X_t, U^c) \leq \frac{1}{n}\}$. Observe that $T_n \nearrow T$ as $n \to \infty$.

(End of proof) $\square$

Example : Let B be a standard Brownian motion with $B_0 = 1$ (in dimension 1). Taking $U = (0, \infty)$, $f(x) = \sqrt{x}$ gives

$$\sqrt{B_t} = 1 + \frac{1}{2}\int_0^t B_s^{-1/2}dB_s - \frac{1}{8}\int_0^t B_s^{-3/2}ds$$

for $t < T = \inf\{t \geq 0 : B_t = 0\}$.

Theorem) Let $U \subset \mathbb{R}^d$ be open and $b : \mathbb{R}_+ \times U \to \mathbb{R}^d$ and $\sigma : \mathbb{R}_+ \times U \to \mathbb{R}^{d \times m}$ be *locally Lipschitz continuous*. Then for every $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$, a Brownian motion B adapted to this filtration, and every $x \in U$, there exists a stopping time T such that, for $t < T$,

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s$$

where T is such that for all $K \subset U$ compact, we have $\sup\{t < T : X_t \in K\} < T$.
Such T is called the **explosion time**.

proof) Fix $K_n \subset K$ compact such that $K_{n+1} \supset K_n$ and $\bigcup_n K_n = U$. One can find b_n and a_n defined on all of $\mathbb{R}^d$ such that

$$b_n|_{K_n} = b|_{K_n}, \quad \text{and } \sigma_n|_{K_n} = \sigma|_{K_n}$$

and such that b_n and σ_n are *globally Lipschitz continuous.* Hence there is a unique solution X^n to $E_x(\sigma_n, b_n)$ for each n. Let $T_n = \inf\{t \geq 0 : X_t^n \notin K_n\}$. By uniqueness, X^{n+1} also solves $E_x(\sigma_n, b_n)$ up to time T_n. Thus $X_t^{n+1} = X_t^n$ for $t < T_n$ and we can define X_t for $t < T = \sup_n T_n$ by requiring that $X_t = X_t^N$ for $t < T_n$.

♠ **Claim :** Let K be compact. Then on $\{T < \infty\}$,

$$\text{a.s.} \quad \sup\{t < T : X_t \in K\} < T$$

: Let L be another compact set such that $K \subset \mathrm{int}(L) \subset L \subset U$. Let $\varphi : U \to \mathbb{R}$ be C^∞ such that $\varphi|_K = 1$ and $\varphi|_{\mathrm{int}(L)^c} = 0$. Let

$$\begin{aligned} &R_1 = \inf\{t \geq 0 : X_t \notin L\}, \quad S_n = \inf\{t \geq R_n : X_t \in K\} \\ &R_n = \inf\{t \geq S_{n-1} : X_t \notin L\} \end{aligned}$$

Let N be the number of crossings of X from $\mathrm{int}(L)^c$ to K. Then on $\{T \leq t, N \geq n\}$,

$$\begin{aligned} n = \sum_{k=1}^n (1-0) &= \sum_{k=1}^n (\varphi(X_{S_k}) - \varphi(X_{R_k})) \\ &= \int_0^t \sum_{k=1}^n \mathbf{1}_{(R_k, S_k]}(s)(D\varphi(X_s) \cdot dX_s + \frac{1}{2} D^2\varphi(X_s) d\langle X \rangle_s) \\ &= \int_0^t (H_s^n dB_s + \tilde{H}_s^n ds) =: Z_t^n \end{aligned}$$

with H^n and $\tilde{H}^n$ are predictable and *bounded* independently of n. So

$$\begin{aligned} & n^2 \mathbf{1}_{\{T \leq t, N \geq n\}} = (Z_t^n)^2 \\ \Rightarrow \quad & \mathbb{P}(T \leq t, N \geq n) = \frac{\mathbb{E}(Z_t^n)^2}{n^2} \leq \frac{C(t)}{n^2} \end{aligned}$$

and hence $\mathbb{P}(T \leq t, N = \infty) = 0$, and in particular $\mathbb{P}(T < \infty, N = \infty) = 0$, which implies the claim.

(End of proof) □

Example : Consider the SDEs

$$dX_t^i = -\nabla_i H(X_t) dt + dB_t^i, \quad X_0 = x$$

Assume that there are $a \geq 0$, $b \geq 0$ such that

$$x \cdot \nabla H(x) \geq -a|x|^2 - b$$

Then, the SDE has a global solution, *i.e.* $T = \infty$ a.s.

proof) Let $T_n = \inf\{t \geq 0 : |X_t|^2 > n\}$. Then by *Itô's formula* to X^{T_n},

$$\begin{aligned} \mathbb{E}|X_{t \wedge T_n}|^2 &= \mathbb{E}|X_0|^2 - \mathbb{E}\Big(2 \int_0^{t \wedge T_n} X_s \cdot \nabla H(X_s) ds - t \wedge T_n\Big) \\ &\leq \mathbb{E}|X_0|^2 + 2a\mathbb{E}\Big(\int_0^{t \wedge T_n} |X_s|^2 ds\Big) + (1+2b)\mathbb{E}(t \wedge T_n) \\ &\leq \mathbb{E}|X_0|^2 + (1+2b)t + 2a \int_0^t \mathbb{E}|X_{s \wedge T_n}|^2 ds \end{aligned}$$

By *Gronwall's lemma*,

$$\mathbb{E}|X_{t \wedge T_n}|^2 \leq (\mathbb{E}|X_0^2| + (1+2b)t)e^{2at}$$

If $\mathbb{P}(T < \infty) > 0$, then for sufficiently large t, $|X_{t \wedge T_n}|^2 \to \infty$ as $n \to \infty$ with positive probability, so it follows that $\mathbb{P}(T < \infty) = 0$.

(End of proof) □

(11th March, Monday)

6 Applications to PDEs and Markov Processes

6.1 Probabilistic representations of solutions to PDEs

Exercise : Let $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be (locally) bounded Borel functions and let $x \in \mathbb{R}^d$. Assume that X is a solution to $E_x(\sigma, b)$. Then for every $f \in C^1(\mathbb{R}_+) \otimes C^2(\mathbb{R}^d)$,

$$M_t^f = f(t, X_t) - f(0, X_0) - \int_0^t \Big(\frac{\partial}{\partial s} + L\Big) f(s, X_s) ds$$

is a continuous local martingale where

$$Lf(y) = \frac{1}{2}\sum_{i,j=1}^d a_{ij}(y) \frac{\partial^2 f}{\partial y^i \partial y^j} + \sum_{i=1}^d b_i(y) \frac{\partial f}{\partial y^i}$$

where $a(y) = \sigma(y)\sigma(y)^T \in \mathbb{R}^{d\times d}$.

Definition) The L is called the **(infinitesimal) generator** of X.

Example :

- $dX = dB$, a Brownian motion has $L = \frac{1}{2}\triangle$.
- $dX = -Xdt + dB$, an Orstein-Uhlenbeck process has $L = \frac{1}{2}\triangle - x \cdot \nabla$.

Drichlet-Poisson problem

Let $U \subset \mathbb{R}^d$, $U \neq \phi$ be open and bounded. Given $f \in C(\overline{U})$ and $g \in C(\partial U)$, a (DP) asks to find $u \in C^2(\overline{U}) = C^2(U) \cap C(\overline{U})$ such that

$$\begin{cases} -Lu(x) = f(x) & \text{for } x \in U \\ u(x) = g(x) & \text{for } x \in \partial U \end{cases} \quad \cdots\cdots\cdots \text{ (DP)}$$

This is called a **Poisson problem** if $f = 0$ and called a **Dirichlet Problem** if $g = 0$.

Definition) $a : \overline{U} \to \mathbb{R}^{d\times d}$ is **uniformly elliptic** if there is $c > 0$ such that

$$\xi^T a(x) \xi \geq c|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^d, x \in \overline{U}$$

Theorem) Assume that U has a smooth boundary, that a and b are Hölder continuous functions, and that a is uniformly elliptic. Then for every Hölder continuous $f : \overline{U} \to \mathbb{R}$ and every continuous $g : \partial U \to \mathbb{R}$, (DP) has a solution.
[See PDE textbooks. Can also use probabilistic method to prove this.]

Theorem) Let $U \subset \mathbb{R}^d$ be open, bounded and non-empty. Let b and σ be bounded measurable, assume $a = \sigma\sigma^T$ is uniformly elliptic and let u be the solution of (DP) with coefficients σ and b. Let $x \in U$, let X be a solution to $E_x(\sigma, b)$. Let $T_U = \inf\{t \geq 0 : X_t \notin U\}$. Then $\mathbb{E}[T_U] < \infty$ and

$$u(x) = \mathbb{E}_x\Big(u(X_{T_U}) - \int_0^{T_U} Lu(X_s) ds\Big) = \mathbb{E}_x\Big(g(X_{T_U}) + \int_0^{T_U} f(X_s) ds\Big)$$

proof) Let $U_n = \{x \in U : \mathrm{dist}(x, \partial U) > \frac{1}{n}\}$, $T_n = \inf\{t \geq 0 : X_t \notin U_n\}$. There are $u_n \in C_b^2(\mathbb{R}^d)$ such that $u|_{U_n} = u_n|_{U_n}$. Then

$$M_t^n = (M^{U_n})_t^{T_n} = u_n(X_{t\wedge T_n}) - u_n(X_0) - \int_0^{t\wedge T_n} Lu_n(X_s)ds$$

is a continuous local martingale, bounded for $t \leq t_0$ for any $t_0 > 0$, so a martingale. Hence $u(x) = u_n(x)$ for $x \in U$, n large enough so that $x \in U_n$ and

$$\begin{aligned} u(x) = u_n(x) &= \mathbb{E}\Big(u_n(X_{t\wedge T_n}) - \int_0^{t\wedge T_n} Lu_n(X_s)\Big) \\ &= \mathbb{E}\Big(u(X_{t\wedge T_n}) + \int_0^{t\wedge T_n} f(X_s)ds\Big) \end{aligned}$$

To take the limit $t \wedge T_n \to T_U$ we will need $\mathbb{E}T_U < \infty$. To see this, let v be a solution to (DP) with $f(x)1$ and $g(x) = 0$ for all x. Then

$$\mathbb{E}(t \wedge T_n) = \mathbb{E}\Big(\int_0^{t\wedge T_n} 1ds\Big) = v(x) - \mathbb{E}(v(X_{w\wedge T_n})) \leq 2\|v\|_\infty < \infty$$

By monotone convergence and since $T_n \wedge t \nearrow T_U$, has

$$\mathbb{E}(T_U) = \lim_{t\to\infty} \lim_{n\to\infty} \mathbb{E}(t \wedge T_n) \leq 2\|v\|_\infty < \infty$$

♡ **Claim:** $u(x) = \mathbb{E}(u(X_{T_U}) + \int_0^{T_U} f(X_s)ds)$

: Since $t \wedge T_n \nearrow T_U$ as $n \to \infty$ and $t \to \infty$, and since

$$\mathbb{E}\Big(\int_0^{T_U} |f(X_s)|ds\Big) \leq \|f\|_\infty \mathbb{E}[T_U] \leq C < \infty.$$

The *Dominated convergence theorem* implies

$$\mathbb{E}\Big(\int_0^{t\wedge T_n} f(X_s)ds\Big) \to \mathbb{E}\Big(\int_0^{T_U} f(X_s)ds\Big).$$

Since u is continuous on $\overline{U}$, also by DCT,

$$\mathbb{E}(u(X_{t\wedge T_n})) \to \mathbb{E}(u(X_{T_U}))$$

This completes the proof.

(End of proof) □

A similar method can also be used not only to prove the existence of solution but to find the soltuion. (not going to do this here).

Cauchy Problem

Given $f \in C_b^2(\mathbb{R}^d)$, find $u \in C(\mathbb{R}_+) \otimes C^2(\mathbb{R}^d)$ such that

$$\begin{cases} \frac{\partial u}{\partial t} = Lu & \text{on } (0,\infty) \times \mathbb{R}^d \\ u(0,\cdot) = f & \text{on } \mathbb{R}^d \end{cases} \quad \cdots\cdots\cdots \text{(CP)}$$

where L is given as above.

Theorem) For $f \in C_b^2(\mathbb{R}^d)$, there exists a solution to (CP).
[Again, refer to a standard PDE texts, such as Evans.]

Theorem) Let u be a (bounded) solution to (CP). Let $x \in \mathbb{R}^d$, let X be any solution to $E_x(\sigma, b)$, $0 \le s \le t$, then

$$\mathbb{E}(f(X_t)|\mathcal{F}_s) = u(t-s, X_s)$$

In particular,

$$\mathbb{E}_x f(X_t) = u(t,x)$$

proof) Let $g(s,x) = u(t-s,x)$ (time runs backward). Then

$$\Big(\frac{\partial}{\partial s} + L\Big) g(s,x) = 0$$

so $M^g = g(s, X_s) - g(s,x)$ is a (true) martingale, so

$$u(t-s, X_s) = g(s, X_s) = \mathbb{E}(g(t, X_t)|\mathcal{F}_s) = \mathbb{E}(u(0, X_t)|\mathcal{F}_s) = \mathbb{E}(f(X_t)|\mathcal{F}_s)$$

(End of proof) □

(13th March, Wednesday)

Theorem) *(Feynman-Kac formula)* Let L, b, σ as before. Let $f \in C_b^2(\mathbb{R}^d)$, $V \in C_b(\mathbb{R}^d)$ and suppose that $u : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ satisfies

$$\begin{cases} \frac{\partial u}{\partial t} = Lu + Vu & \text{on } \mathbb{R}_+ \times \mathbb{R}^d \\ u(0, \cdot) = f & \text{on } \mathbb{R}^d \end{cases}$$

(Vu here is just a pointwise multiplication). Let X be a solution to $E_x(\sigma, b)$ for some $x \in \mathbb{R}^d$. Then for all $t \ge 0$,

$$u(t,x) = \mathbb{E}_x\Big[f(X_t) \exp\Big(\int_0^t V(X_s) ds\Big)\Big]$$

[This result resembles the form of path integral solution of Schrödinger's equation. This is the reason why this is called the 'Feynman'-Kac formula. However, proving the same with Schrödinger's equation (where there is $-i$ in front of $\frac{\partial u}{\partial t}$) is much more difficult as we lose positivity of L.]

proof) Let $E_t = \exp(\int_0^t V(X_s)ds)$. For $s < t$, set $M_s = u(t-s, X_s)E_s$, then

$$\begin{aligned} dM_s &= -\frac{\partial}{\partial t} u(t-s, X_s) E_s ds + \nabla u(t-s, X_s) E_s \sigma_s dB_s + Lu(t-s, X_s) E_s ds + u(t-s, X_s) V(X_s) E_s ds \\ &= \Big(-\frac{\partial}{\partial t} + L + V(X_s)\Big) u(t-s, X_s) E_s ds + d(\text{martingale}) \\ &= d(\text{martingale}) \end{aligned}$$

Thus M is a continuous local martingale on $[0,t]$. by assumption, M is also bounded, so a martingale. Hence

$$u(t,x) = M_0 = \mathbb{E}_x M_t = \mathbb{E}_x u(0, X_t) E_t = \mathbb{E}(f(X_t)E_t)$$

(End of proof) □

6.2 Markov property

Let $B(\mathbb{R}^d)$ be the Banach space of **bounded Borel functions** on $\mathbb{R}^d$, with $\|f\| = \sup_{x\in\mathbb{R}^d} |f(x)|$ for $f \in B(\mathbb{R}^d)$.

Definition)

(i) A collection of bounded linear operators Q_t on $B(\mathbb{R}^d)$ is a **transition semigroup** if $Q_t f \geq 0$ if $f \geq 0$ (pointwise), $Q_t\mathbf{1} = \mathbf{1}$, $\|Q_t\| \leq 1$, and

$$Q_{t+s} = Q_t Q_s \quad \forall t, s \geq 0$$

(ii) An $(\mathcal{F}_t)$-adapted process X is a **Markov process** with transition semigroup $(Q_t)_t$ if

$$\mathbb{E}(f(X_{s+t})|\mathcal{F}_s) = Q_t f(X_s) \quad \forall s, t \geq 0, \ f \in B(\mathbb{R}^d)$$

Theorem) Let $b : \mathbb{R}^d \to \mathbb{R}$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be *Lipschitz* (this can be weakened). Assume X is a solution to $E(\sigma, b)$ on some $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and B. Then $X = (X_t)_{t\geq 0}$ is a Markov process with semigroup

$$Q_t f(x) = \mathbb{E}(f(X_t^x)) = \int f(F_x(w)_t) P^m(dw)$$

where X_t^x is an arbitrary solution to $E_x(\sigma, b)$, and F_x is the *Itô solution map*, P^m is the Wiener measure.

proof) Let X be a solution to $E(\sigma, b)$.

♡ **Claim :** $\mathbb{E}(f(X_{t+s})|\mathcal{F}_s) = Q_t f(X_s)$

: By definition of X,

$$\begin{aligned} X_t &= X_0 + \int_0^t \sigma(X_u) dB_u + \int_0^t b(X_u) du \quad \forall t \geq 0 \\ \Rightarrow \quad X_{t+s} &= X_s + \int_s^{t+s} \sigma(X_u) dB_u + \int_s^{t+s} b(X_u) du \quad \forall s, t \geq 0 \end{aligned}$$

Set $X'_t = X_{s+t}$, $\mathcal{F}'_t = \mathcal{F}_{s+t}$, $B'_t = B_{s+t} - B_s$. Then $(\Omega, \mathcal{F}, (\mathcal{F}'), P)$ is another filtered probability space obeying the usual conditions, and B' is a $(\mathcal{F}'_t)$-Brownian motion with $B'_0 = 0$. Then

$$X'_t = X'_0 + \int_0^t \sigma(X'_u) dB'_u + \int_0^t b(X'_u) du$$

(Justification of $\int_s^{s+t} \sigma(X_u) dB_u = \int_0^t \sigma(X'_u) dB'_u$ uses approximation of integrals - this should not be treated naively!). Thus X' solves $E(\sigma, b)$ with $X'_0 = X_s$. By the theorem about the solution map, we have $X' = F_{X_s}(B')$ a.s. So

$$\mathbb{E}(f(X_{s+t})|\mathcal{F}_s) = \mathbb{E}(f(X'_t)|\mathcal{F}_s) = \mathbb{E}(f(F_{X_s}(B')_t)|\mathcal{F}_s) = \int f(F_{X_s}(w)_t) P^m(dw) = Q_t f(X_s)$$

Also,

$$Q_{t+s} f(x) = \mathbb{E}(f(X_{t+s}^x)) = \mathbb{E}\Big(\mathbb{E}(f(X_{t+s}^x)|\mathcal{F}_s)\Big) = \mathbb{E}(Q_t f(X_s^x)) = Q_s Q_t f(x)$$

so $(Q_t)_{t\geq 0}$ indeed forms a transition semigroup.

(End of proof) □

Definition) Let Q_t be the transition semigroup.

(i) A probability measure μ on $\mathbb{R}^d$ is **invariant** under (Q_t) if

$$\int Q_t f(x)\mu(dx) = \int f(x)d\mu(x) \quad \forall f \in B(\mathbb{R}^d)$$

(ii) A probability measure μ is **reversible** with respect to (Q_t) if

$$\int g(x)Q_t f(x)\mu(dx) = \int f(x)Q_t g(x)\mu(dx)$$

(check this.)

Fact : Reversibility of μ implies it is invariant. (Take $g = 1$ and use $Q_t 1 = 1$.)

Example : Consider the transition semigroup associated to the SDE, with suitable on H,

$$dX_t = -\frac{1}{2}\nabla H(X_t)dt + dB_t$$

(Note that, if taking $H(x) = \lambda|x|^2$, this gives an Orstein-Uhlenbeck process.) Then the measure $\mu(dx) = \frac{1}{Z}e^{-H(x)}dx$, where $Z = \int e^{-H(x)}dx$ is reversible for $(*)$.

An application, if you want to sample from the measure $\mu(dx) = \frac{1}{Z}e^{-H(x)}dx$, then we can simulate the SDE for a suitably long time and how the distribution is made after long time. (called Markov chain Monte-Carlo simulation)

Lemma) Assume that the explosion time for $(*)$ is infinite. Then for $f : C([0,T],\mathbb{R}^)d \to \mathbb{R}$,

$$\mathbb{E}\Big(f(X|_{[0,T]})\Big) = \mathbb{E}^{\mathrm{BM}}\Big[f(X|_{[0,T]})\exp\Big(\frac{1}{2}H(X_0) - \frac{1}{2}H(X_T) - \int_0^T(\frac{1}{8}|\nabla H|^2 - \frac{1}{4}\triangle H)(X_s)ds\Big)\Big]$$

(where $\mathbb{E}^{\mathrm{BM}}$ takes average over law under which X is a Brownian motion with same initial condition.)