Stochastic Analysis

LU Shan

Preface

These notes accompany a sequel to my earlier study of stochastic processes, and grow out of the course on *Stochastic Analysis* taught by Prof. Jieliang HONG at SUSTech. After completing my exchange semester at SUSTech in 2025 Fall, I returned to my home institution. Yet each Wednesday I still travel to Shenzhen to sit in the lectures as an auditing student, and only later, often with some delay, rewrite the blackboard arguments into a form that can be revisited, tested, and refined.

If my first notebook began as a struggle to keep pace with martingales, Markov structure, and Brownian motion, this one begins at the point where the continuum becomes genuinely unforgiving: the moment we try to integrate with respect to a path of infinite variation. Stochastic analysis, to me, is the discipline of rebuilding real analysis under randomness, first by learning exactly *why* the Riemann-Stieltjes intuition fails for Brownian motion, and then by replacing it with Ito integration, quadratic variation, and the semimartingale viewpoint.

This is the mathematical backbone behind modern financial mathematics: the insistence that "fair" price dynamics should behave like martingales, and trading strategies should be modeled by predictable integrands whose stochastic integrals preserve that structure. The primary references for this notebook are the classic texts of Karatzas Shreve and Rogers Williams. I will follow their standards of rigor while keeping the narrative as close as possible to what I found most helpful when learning.

Finally, I am grateful to Prof. HONG for a lecture style that makes difficult concepts feel approachable without sacrificing precision. Any errors, omissions, or infelicities in exposition are entirely my own. To the reader who is also returning to these topics after feeling overwhelmed the first time: may these notes offer a steadier path into Ito calculus, and a reminder that the beauty of stochastic analysis lies not merely in its abstraction, but in its ability to turn uncertainty into structure.

Contents

0 Motivation 3

1 Preliminary 5

2 Stochastic Integral 28

3 Brownian Motion & PDE 29

4 Stochastic Differential Equations 30

0 Motivation

To initiate, recall the **Riemann-Stieltjes** Integral:

Definition 0.1. If f is continuous and g is of **finite variation**, then:

$$\int_a^b f(x)\mathrm{d}g(x)$$

is well defined to be, with $a \leq x_0 < x_1 < \cdots < x_n \leq b$ and $\xi_i \in [x_i, x_{i+1}]$:

$$\lim_{n\to\infty} \sum_{i=0}^{n-1} f(\xi_i)\left(g(x_{i+1}) - g(x_i)\right)$$

, and known as **Riemann-Stieltjes** Integral.

Then the main goal of stochastic analysis is to re-walk through the real analysis by rigorously defining the **Ito Integral** even for **stochastic processes**:

$$(H \cdot X)_t = \int_0^t H_s \mathrm{d}X_s$$

The original rule for Riemann-Stieltjes fails if we choose integrator to be Brownian motion.

Theorem 0.1. *For a standard brownian motion,* $\mathsf{B} = \{\mathsf{B}_t, t \geq 0\}$*:*

$$\sum_{k=1}^{2^n} \left(\mathsf{B}_{\frac{k}{2^n}} - \mathsf{B}_{\frac{k-1}{2^n}}\right)^2 \xrightarrow{a.s.} 1$$

Proof. On dyadic partition, for each fixed n, define, $\forall\ 1 \leq k \leq 2^n$:

$$\Delta_k^{(n)} := \mathsf{B}_{k/2^n} - \mathsf{B}_{(k-1)/2^n} \implies Q_n := \sum_{k=1}^{2^n} \left(\mathsf{B}_{k/2^n} - \mathsf{B}_{(k-1)/2^n}\right)^2 = \sum_{k=1}^{2^n} \left(\Delta_k^{(n)}\right)^2$$

Since B is a SBM, the random variables $\Delta_1^{(n)}, \dots, \Delta_{2^n}^{(n)}$ are independent and

$$\Delta_k^{(n)} \sim N(0, 2^{-n}) \implies \mathsf{E}\left(\Delta_k^{(n)}\right)^2 = 2^{-n}$$

Therefore, the expectation of Q_n:

$$\mathsf{E}Q_n = \sum_{k=1}^{2^n} \mathsf{E}\left(\Delta_k^{(n)}\right)^2 = 2^n \cdot 2^{-n} = 1$$

Besides, by simple manipulation, the variance is given by, using independence:

$$\mathrm{Var}(Q_n) = \mathsf{E}\left(Q_n - 1\right)^2 = \sum_{k=1}^{2^n} \mathrm{Var}\left[\left(\Delta_k^{(n)}\right)^2\right] = 2^n \cdot 2 \cdot 2^{-2n} = 2^{1-n} \longrightarrow 0$$

Now fix $\varepsilon > 0$, by Chebyshev's inequality,

$$\mathsf{P}\big(|Q_n - 1| > \varepsilon\big) \le \frac{\mathsf{E}\,(Q_n - 1)^2}{\varepsilon^2} = \frac{2^{1-n}}{\varepsilon^2} \implies \sum_{n=1}^{\infty} \mathsf{P}\big(|Q_n - 1| > \varepsilon\big) \le \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} 2^{1-n} < \infty$$

By the Borel–Cantelli lemma,

$$\mathsf{P}\big(|Q_n - 1| > \varepsilon \text{ i.o.}\big) = 0.$$

Since this holds for every $\varepsilon = 1/m$, $m \in \mathbb{N}$, we conclude that

$$Q_n \to 1 \qquad \text{a.s.}$$

This completes the proof. □

Using above easy-to-prove theorem, we first to be precise about **finite variation**:

Definition 0.2. The **Finite Variation** of a funtion $g : [a, b] \mapsto \mathbb{R}$ is defined to be:

$$V_g([a, b]) = \int_a^b |\mathrm{d}g(x)| = \sup_n \sum_{i=1}^{n} |g(x_{i+1}) - g(x_i)|$$

Now, we could justify that brownian motion fails in having infinite variation:

Corollary 0.2.

$$V_{\mathsf{B}}([0, 1]) = \infty, \ \ a.s.$$

Proof. Firstly, we notice the following inequality:

$$\begin{aligned}\sum_{k=1}^{2^n} \left(\mathsf{B}_{\frac{k}{2^n}} - \mathsf{B}_{\frac{k-1}{2^n}}\right)^2 &\le \sup_{1\le k\le 2^n} \left|\mathsf{B}_{\frac{k}{2^n}} - \mathsf{B}_{\frac{k-1}{2^n}}\right| \cdot \sum_{k=1}^{2^n} \left|\mathsf{B}_{\frac{k}{2^n}} - \mathsf{B}_{\frac{k-1}{2^n}}\right| \\ &\le V_{\mathsf{B}}([0,1]) \cdot \sup_{1\le k\le 2^n} \left|\mathsf{B}_{\frac{k}{2^n}} - \mathsf{B}_{\frac{k-1}{2^n}}\right|\end{aligned}$$

Then if $V_{\mathsf{B}}([0, 1]) < \infty$, then by continuity of B, we have:

$$\sum_{k=1}^{2^n} \left(\mathsf{B}_{\frac{k}{2^n}} - \mathsf{B}_{\frac{k-1}{2^n}}\right)^2 \longrightarrow 0$$

, which is a contradiction, so $V_{\mathsf{B}}([0, 1]) = \infty$ with probability 1. □

However, Ito found that, if we put some restriction on integrand and integrator, such as:

$$\begin{aligned} H_s &= \text{previsible/ predictable} \\ X_s &= \text{semimartingale} \end{aligned} \implies (H \cdot X)_t = \int_0^t H_s \mathrm{d}X_s \text{ is well defined.}$$

Example 0.1. Let $H_n \in \mathcal{F}_{n-1}$, and $M = \{M_n, n \in \mathbb{N}\}$ be a martingale, then:

$$(H \cdot M)_n = \sum_{k=1}^{n} H_k \left(M_k - M_{k-1}\right)$$

is a martingale, which is known as discrete **martingale transformation**.

Besides rigorously defining the integral for general case, we also want the result integration keeps some desirable property, like **martingale**, this intuition comes from the financial application, where if the integral can't keep martingale, we could replicate an **arbitrage portfolio**. Then next section is to introduce some toolkits and also know the definition of **predictable, semimartingale**.

1 Preliminary

Given a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, a stochastic process $X = \{X_t, t \geq 0\}$ is a collection of r.v.s taking values in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ formally, it is a **measurable function**:

$$\begin{aligned} X :& (t, \omega) \mapsto X_t(\omega) \\ & [0, \infty) \times \Omega \mapsto \mathbb{R}^d \end{aligned}$$

1.1 Sample Paths

Now if we fix ω, then the function:

$$t \mapsto X_t(\omega)$$

is called a **sample path**. Based on this concept, we could provide some useful definitions:

Definition 1.1. For two stochastic processes X and Y, we say Y is a **modification** of X if $\forall\, t \geq 0$, $\mathsf{P}(X_t = Y_t) = 1$

Remark 1.1. To be more precise, $\mathsf{P}(\{\omega : X_t(\omega) = Y_t(\omega)\}) = 1$

Most of time, modification is enough for us to inherit some property like continuity, but in order for the smoothness of knowledge and to get better understanding, we present a "slightly" stronger concept:

Definition 1.2. We say X and Y are **indistinguishable** if

$$\mathsf{P}\left(X_t = Y_t, \forall\, t \geq 0\right) = 1$$

Then Y is a **version** of X.

The reason why we say it is a slightly stronger version, is that due to the countable addivity, we could guarantee all the rational points to be the same, and if we have the "continuity" of the sample path we could approximate other points by rational points, then achieve indistinguishable, as following example gives:

Example 1.1. Let Y be a modification of X and suppose X and Y both have a.s. **right continuous** sample paths, then X and Y are indistinguishable.

Definition 1.3. A function f is **right-continuous** if:

$$f(x) = f(x^+) = \lim_{t \downarrow x} f(t)$$

Proof. Since $\forall\, t \geq 0$, $\mathsf{P}(X_t \neq Y_t) = 0$, then we get:

$$\mathsf{P}\left(\bigcup_{q \in \mathbb{Q}^+} X_q \neq Y_q\right) \leq \sum_{q \in \mathbb{Q}^+} \mathsf{P}(X_q \neq Y_q) = 0 \implies \mathsf{P}\left(\bigcap_{q \in \mathbb{Q}^+} X_q = Y_q\right) = 1$$

For two **right-continuous** functions f, g if:

$$f(q) = g(q), \forall\, q \in \mathbb{Q}^+ \implies f(t) = g(t), \forall\, t \geq 0$$

Therefore, using this result from **measure theory**, we land in:

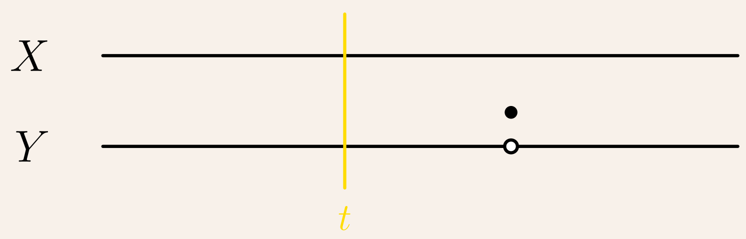
$$\mathsf{P}\left(X_t = Y_t, \forall\, t \geq 0\right) = \mathsf{P}\left(\bigcap_{t \geq 0} X_t = Y_t\right) = 1$$

Hence, we completed the proof. □

Please notice from above example 1.1 that we restrict to a.s. right-continuous sample path, so that we could land in approximation, but the following opposite example shows us the difference of modification and indistinguishable process.

Example 1.2. Let $T \sim \text{Exp}(1)$, define $X_t \equiv 0, \forall\, t \geq 0$, and another one:

$$Y_t = \begin{cases} 0, & t \neq T \\ 1, & t = T \end{cases}$$



Then we know:

$$\forall\, t \geq 0, \mathsf{P}(X_t \neq Y_t) = \mathsf{P}(Y_t = 1) = \mathsf{P}(t = T) = 0$$

so X is a modification of Y but **not a version of**, which you should verify.

1.2 Filtration

After talking about process, as processes go the information also increases, then to model this intuitive behavior, we introduce **filtration**: Let $\{\mathcal{F}_t, t \geq 0\}$ be a filtration of $(\Omega, \mathcal{F})$ such that $\forall\, 0 \leq s < t < \infty$:

$$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$$

Typically, we just adapt to the so called **natural filtration**:

Example 1.3. Let $X = \{X_t, t \geq 0\}$ be a stochastic process, define:

$$\mathcal{F}_t^X = \sigma\left(X_s, 0 \leq s \leq t\right)$$

, then $\{\mathcal{F}_t^X, t \geq 0\}$ is a filtration.

From definition we could see, natural filtration is the minimum filtration that contains the information of the corresponding process, and also the adaptivity, i.e. $X_t \in \mathcal{F}_t^X$.

Example 1.4. Let X be a process, every sample paths of which is **RCLL**, let:

$$A = \{\omega : X(\omega) \text{ is continuous on } [0, t_0)\}$$

for some fixed $t_0 > 0$, then $A \in \mathcal{F}_{t_0}^X$.

Definition 1.4. The sample path of a process is **RCLL** if its sample path is **Right-Continuous with Left Limit**, i.e.:

$$f(x) = f(x^+),\ \ f(x^-)\ \exists$$

Proof. Since we already have right continuous, then what remain is show it meets with left limit:

$$A = \{\omega : \forall\, s \in (0, t_0), X_{s^-}(\omega) = X_s(\omega)\}$$

Similar trick, to rewrite event A into "rational number version":

$$A = \{\omega : \forall\, q \in (0, t_0) \cap \mathbb{Q}, X_{q^-}(\omega) = X_q(\omega)\}$$

It boils down into countable cases, to numerate $(0, t_0) \cap \mathbb{Q} = \{q_m, m \geq 1\}$:

$$\begin{aligned} A &= \left\{\omega : \forall\, q_m, \forall\, k \geq 1, \exists\, N \geq 1, \forall\, n \geq N, \left|X_{q_m - \frac{1}{n}}(\omega) - X_{q_m}(\omega)\right| < \frac{1}{k}\right\} \\ &= \bigcap_{m=1}^{\infty} \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n=N}^{\infty} \left\{\left|X_{q_m - \frac{1}{n}}(\omega) - X_{q_m}(\omega)\right| < \frac{1}{k}\right\} \end{aligned}$$

Notice each $\left\{\left|X_{q_m - \frac{1}{n}}(\omega) - X_{q_m}(\omega)\right| < \frac{1}{k}\right\} \in \mathcal{F}_{t_0}^X$, then $A \in \mathcal{F}_{t_0}^X$. □

Notice that example 1.4 requires every sample path to be RCLL, but if we lessen a little bit, even into a.s. RCLL, the conclusion may fail as following example suggests:

Example 1.5. Let X be a process whose sample paths are **RCLL** a.s., with all the same setup as example 1.4, then it is possible that $A \notin \mathcal{F}_{t_0}^X$.

Proof. WLOG, let $t_0 = 1$, and U be a non-measurable r,v. such that:

$$\mathsf{P}(U = 1) = 0$$

Define our X by:

$$X_t(\omega) = \begin{cases} 0, & U(\omega) \neq 1 \\ \begin{cases} 0, & t \neq \frac{1}{2} \\ 1, & t = \frac{1}{2} \end{cases}, & U(\omega) = 1 \end{cases}$$

By our construction, we know that:

$$A = \{\omega : X(\omega) \text{ is continuous on } [0, t_0)\} = \Omega / \{U = 1\}$$

, so if $\{U = 1\} \notin \mathcal{F}_{t_0}^X$, then $A \notin \mathcal{F}_{t_0}^X$. □

From example 1.5, we see the key of the proof is to construct a zero-measure set such that it is not in our natural filtration, which is too tricky and hopefully be forbidden, then it gives us the motivation to consider broader class of filtration:

Definition 1.5. A filtration $\{\mathcal{F}_t, t \geq 0\}$ is said to satisfy the **usual condition** if:

(a) **Right-continuous**: $\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$.

(b) $\mathcal{F}_0$ contains all the **P-negligible** set in $\mathcal{F}$.

To complete the gap, we further provide the definitions of two:

Definition 1.6. An event A is **P-negligible** if $\exists\, N \in \mathcal{F}$, s.t. $A \subseteq N, \mathsf{P}(N) = 0$.

Another one is "left/right limit" of filtration:

Definition 1.7. We say a filtration is **right-continuous** if $\mathcal{F}_t = \mathcal{F}_{t+}$, where:

$$\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s, \mathcal{F}_{t-} = \sigma\left(\bigcup_{s<t} \mathcal{F}_s\right)$$

After introducing information flow, we need to relate it with process:

Definition 1.8. The stochastic process X is **adapted** to the filtration $\{\mathcal{F}_t, t \geq 0\}$ if $\forall\, t \geq 0$, $X_t \in \mathcal{F}_t$. Clearly, X is adapted to the **natural filtration**.

Adaptivity is too loose, and impedes understanding, thus we consider **progressively measurable**.

Definition 1.9. The stochastic process X is called **progressively measurable**

w.r.t $\{\mathcal{F}_t, t \geq 0\}$ if $\forall\, t \geq 0$, the map following is measurable.

$$(s,\omega) \mapsto X_s(\omega) : ([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t) \mapsto \big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\big)$$

Remark 1.2. To compare, adaptivity is to require measurability to:

$$(s,\omega) \mapsto X_s(\omega) : ([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}) \mapsto \big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\big)$$

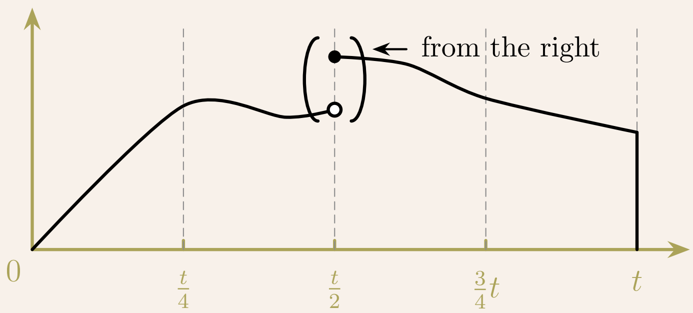
Intuitively speaking, progressively measurable is to require the information is "strictly" enough at this each time, to see the relations:

$$\text{adapted} \supseteq \text{progressively measurable} \supseteq \text{previsible}$$

Most of cases, we don't differentiate them, since it is really close as below example shows:

Example 1.6. If X is adapted to filtration $\{\mathcal{F}_t, t \geq 0\}$ and every sample path is right-continuous (same for left-continuous), then X is also progressively measurable.

Proof. Since right continuity, $\forall\, t > 0, \forall\, n \geq 1, \forall\, 0 \leq k \leq 2^n - 1$, we construct:



$$X_s^{(n)}(\omega) = X_{\frac{k+1}{2^n}t}(\omega), \forall\, \frac{kt}{2^n} < s \leq \frac{(k+1)t}{2^n}$$

, then $\lim_{n\to\infty} X_s^{(n)}(\omega) = X_s(\omega)$ for all $(s,\omega) \in [0,t] \times \Omega$. Recall that limit perserve our measurability, then it suffices to prove:

$$(s,\omega) \mapsto X_s^{(n)}(\omega) : ([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t) \mapsto \big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\big)$$

is measurable, however, this is clearly true since, $\forall\, A \in \mathcal{B}(\mathbb{R}^d)$:

$$\{(s,\omega) : X_s^{(n)}(\omega) \in A\} = \bigcup_{k=0}^{2^n-1} \left(\frac{k}{2^n}t, \frac{k+1}{2^n}t\right] \times \left\{X_{\frac{k+1}{2^n}t} \in A\right\} \in \mathcal{B}([0,t]) \otimes \mathcal{F}_t$$

Therefore, we completed the proof. □

1.3 Stopping Times

Stopping times are really important subject of interest, since we already known that original brownian motion has infinite variation, but if we use stopping times to replace the original time, then to construct **local martingale**, which can be used as integrator.

Definition 1.10. A random time T is a **stopping time** of $\{\mathcal{F}_t, t \geq 0\}$ if $\forall\, t \geq 0$, $\{T \leq t\} \in \mathcal{F}_t$, another one is **optional time** if $\forall\, t \geq 0$, $\{T < t\} \in \mathcal{F}_t$.

Intuitively speaking, stopping time is a r.v. such that its property can be checked at each time (i.e. the information ($\{T \leq t\}$) of it contains within time t ($\mathcal{F}_t$)), and optional time is by time t but not included. However, they don't have too much difference as below proposition suggests:

Proposition 1.1. *If T is a stopping time, then also a optional time. Conversely, if the filtration is **right-continuous**, then stopping time is same as optional time.*

Proof. First one is easy to see, we focus on latter equivalence statement:

$\Longrightarrow$ The forward direction is simple, as we notice below:

$$\{T < t\} = \bigcup_{n=1}^{\infty} \left\{T \leq t - \frac{1}{n}\right\}$$

, as each $\{T \leq t - 1/n\} \in \mathcal{F}_{t-1/n} \subseteq \mathcal{F}_t$, then $\{T < t\} \in \mathcal{F}_t$.

$\Longleftarrow$ For backward direction, it is non-trival, $\forall\, m \geq 1$:

$$\{T \leq t\} = \bigcap_{n=m}^{\infty} \left\{T < t + \frac{1}{n}\right\}$$

, as each $\{T < t + 1/n\} \in \mathcal{F}_{t+1/n} \subseteq \mathcal{F}_{t+1/m}$, then $\{T \leq t\} \in \mathcal{F}_{t+1/m}$.

Lastly, we finished the proof by the arbitrariness of m, then $\{T \leq t\} \in \mathcal{F}_{t+} = \mathcal{F}_t$. □

Next, after we defined the stopping time, let us see some properties:

Proposition 1.2. *If T and S are stopping times, then:*

$$T \wedge S,\ T \vee S,\ T + S$$

are all stopping times.

Proof. The first two are easy once we seperate the event properly:

(1) For $T \wedge S$, we could consider as below:

$$\{T \wedge S \leq t\} = \{T \leq t\} \bigcup \{S \leq t\} \in \mathcal{F}_t$$

, as T, S are stopping times, then $\{T \leq t\}, \{S \leq t\} \in \mathcal{F}_t$.

(2) For $T \vee S$, we do the similar seperation, $\{T \vee S \leq t\} = \{T \leq t\} \bigcap \{S \leq t\}$.

(3) $T+S$ is tricky, where we should transfer using closure under complement:

$$\{T+S\leq t\}\in\mathcal{F}_t \Longleftrightarrow \{T+S>t\}\in\mathcal{F}_t$$

After we do this transformation, we could seperate this by:

$$\{T+S>t\}=\{T=0,S>t\}\bigcup\{S=0,T>t\}\bigcup\{T\geq t,S>0\}\bigcup\{0<T<t,T+S>t\}$$

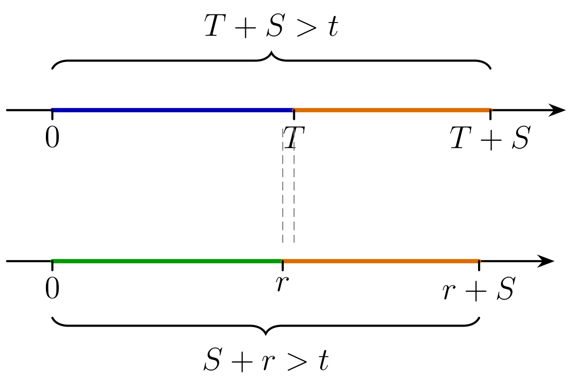
For the first one, it is easy to see:

$$\{T=0,S>t\}=\{T=0\}\bigcap\{S>t\}=\{T=0\}\bigcap\{S\leq t\}^c\in\mathcal{F}_t$$

Then using similar method, we could easily check middle two, what remains is:

Lemma.

$$\{0<T<t,T+S>t\}=\bigcup_{\substack{0<r<t\\ r\in\mathbb{Q}^+}}\{r<T<t,S>t-r\}$$



Proof. of Lemma:

$\Longleftarrow$ Backward direction is easy, as $S>t-r>t-T$, hence $S+T>t$.

$\Longrightarrow$ Forward direction relies on the dense of rational number, where since $S+T>t$, then by dense of rational numbers, $\exists\, r\in(0,T)$, s.t. $S+r>t$.

Last one can be visualised as above by moving r really close to T. □

With the lemma above, we could seperate event as before, then complete the proof. □

Lastly, if we have a set of stopping times, then the sup of it is also a stopping time:

Proposition 1.3. *If $\{T_n,n\geq 1\}$ are s.t., then $\sup_{n\geq 1}T_n$ is a s.t.*

Proof. This proof is simple and straightforward, $\forall\, t\geq 0$:

$$\left\{\sup_{n\geq 1}T_n\leq t\right\}=\bigcap_{n=1}^{\infty}\{T_n\leq t\}\in\mathcal{F}_t$$

Then we complete the proof. □

Since we already have s.t. as a kind of time, it is reasonable to ask its σ algebra:

Definition 1.11. The stopping σ-algebra is:

$$\mathcal{F}_T = \left\{ A \in \mathcal{F} : \forall\, t \geq 0, A \bigcap \{T \leq t\} \in \mathcal{F}_t \right\}$$

Intuitively speaking, above definition 1.11 is to say for every event in stopping σ algebra, given the time t we are at, we could know whether it happens or not.

Proposition 1.4. *If $T \equiv t_0$, then $\mathcal{F}_T = \mathcal{F}_{t_0}$.*

Proof. Firstly, we observe that $\forall\, t \geq 0$:

$$\{T \leq t\} = \begin{cases} \emptyset, & t_0 > t \\ \Omega, & t_0 \leq t \end{cases}$$

$\Longleftarrow$ Then we firstly check the backward direction, $\forall\, A \in \mathcal{F}_{t_0}$, $\forall\, t \geq 0$:

(1) If $t \geq t_0$, then:

$$A \cap \{T \leq t\} = A \cap \Omega = A \in \mathcal{F}_{t_0} \subseteq \mathcal{F}_t$$

(2) If $t < t_0$, then:

$$A \cap \{T \leq t\} = A \cap \emptyset = \emptyset \in \mathcal{F}_t$$

Therefore, $A \in \mathcal{F}_T$ as definition states.

$\Longrightarrow$ Forward direction is similar, as $\forall\, A \in \mathcal{F}_T$, and $\forall\, t \geq 0$, then:

$$A \cap \{T \leq t\} \in \mathcal{F}_t \implies A = A \cap \Omega = A \cap \{t_0 \leq t_0\} = A \cap \{T \leq t_0\} \in \mathcal{F}_{t_0}$$

Therefore, we completed the proof. □

Besides, we will also be interested into the relationship of a set of s.t.s,

Proposition 1.5. *$\forall$ stopping times T and S, $\forall\, A \in \mathcal{F}_S$, we have:*

$$A \cap \{S \leq T\} \in \mathcal{F}_T$$

In particular, if $S \leq T$ for all $\omega \in \Omega$, then:

$$\mathcal{F}_S \subseteq \mathcal{F}_T$$

Proof. The key idea is to use the condition that $A \in \mathcal{F}_S$, so that:

$$A \cap \{S \leq t\} \in \mathcal{F}_t$$

And in order to show, $A \cap \{S \leq T\} \in \mathcal{F}_T$, we need $\forall\, t > 0$,

$$A \cap \{S \leq T\} \cap \{T \leq t\} \in \mathcal{F}_t$$

Notice here, since on $\{S \leq T\} \cap \{T \leq t\}$:

(a) $\{S \leq T\} \cap \{T \leq t\} = \{S \leq T \leq t\} \implies S \leq t$.

(b) $\{S \le T\} = \{S \wedge t \le T \wedge t\}$.

Then, it suffice to show, $\forall\, t > 0$:

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

The first lives in $\mathcal{F}_t$ followed from the fact that $A \in \mathcal{F}_S$, and second comes from the fact that T is a s.t., it remains to show, $\{S \wedge t \le T \wedge t\} \in \mathcal{F}_t$:

Lemma. *Let T and S be two s.t.s, then $\forall\, t \ge 0$:*

$$\{S \wedge t \le T \wedge t\} \in \mathcal{F}_t$$

Proof. of Lemma
This follows the similar structure of proposition 1.2's lemma to decompose:

$$\{S \wedge t \le T \wedge t\} = \bigcup_{r \in \mathbb{Q}^+ \cap [0,t]} \{S \wedge t \le r\} \cap \{T \wedge t \ge r\}$$

Using the closure of complement of σ-algebra to show its conclusion. □

Therefore, we completed the proof for general situation, and in particular, since $S \le T$ for all $\omega \in \Omega$, then $\{S \le T\} = \Omega$, so $\forall\, A \cap \Omega = A \in \mathcal{F}_S, A \in \mathcal{F}_T \implies \mathcal{F}_S \subseteq \mathcal{F}_T$. □

To reiterate, the reason why we introduce the stopped time $(T \wedge t)$ is to relate the original s.t. with t, so that we could manipulate on it with relationship pf $\mathcal{F}_t$.

Proposition 1.6. *Let S and T be two s.t.s, then:*

$$\mathcal{F}_{S \wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$$

, and also:

$$\{T < S\}, \{S < T\}, \{T \le S\}, \{S \le T\}, \{S = T\} \in \mathcal{F}_S \cap \mathcal{F}_T$$

Proof. The intuition is clear, for σ-algebra for s.t.s which is the smallest of two, is certainly the intersection of two σ-algebra of two s.t.s Let's prove them in order:

(a) Firstly, consider the inclusion:

$\implies$ From proposition 1.5, we have:

$$\begin{cases} S \wedge T \le S \\ S \wedge T \le T \end{cases} \implies \begin{cases} \mathcal{F}_{S \wedge T} \subseteq \mathcal{F}_S \\ \mathcal{F}_{S \wedge T} \subseteq \mathcal{F}_T \end{cases} \implies \mathcal{F}_{S \wedge T} \subseteq \mathcal{F}_S \cap \mathcal{F}_T$$

$\impliedby$ For backward direction, $\forall\, A \in \mathcal{F}_S \cap \mathcal{F}_T$, $\forall\, t > 0$, we need to show:

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

, above is easy to show by distribution rule and the fact that $A \in \mathcal{F}_S, \mathcal{F}_T$:

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

(b) Similarly, from proposition 1.5, we have, $\forall\, A \in \mathcal{F}_S$, we have:

$$A \cap \{S \le T\} \in \mathcal{F}_T \overset{\text{Choose } A=\Omega}{\Longrightarrow} \{S \le T\} \in \mathcal{F}_T \implies \{S \le T\}^c = \{S > T\} \in \mathcal{F}_T$$

It remains to show, $\{S < T\} \in \mathcal{F}_T$, since others can be shown by interchanging the place of them, which we will see it later, here notice that, in order to introduce the relations of S and T, we construct a stopped s.t. by defining $R := S \wedge T$, then:

$$\{S < T\} = \{R < T\}, R \in \mathcal{F}_R \subseteq \mathcal{F}_T \implies \{S < T\} = \{R < T\} \in \mathcal{F}_T$$

Lastly, we use the power of interchanging the orders,

$$\{T \le S\}, \{T > S\}, \{S > T\} \in \mathcal{F}_S$$

, so that other parts can all be expressed by their complements:

(i) For $\{S \le T\} \in \mathcal{F}_T$, we use $\{S > T\} \in \mathcal{F}_S$:

$$\{S \le T\} = \{S > T\}^c \in \mathcal{F}_S \implies \{S \le T\} \in \mathcal{F}_S \cap \mathcal{F}_T$$

(ii) For $\{S < T\} \in \mathcal{F}_T$, we use $\{T \le S\} \in \mathcal{F}_S$:

$$\{S < T\} = \{T \ge S\}^c \in \mathcal{F}_S \implies \{S < T\} \in \mathcal{F}_S \cap \mathcal{F}_T$$

(iii) Lastly, for $\{T = S\}$, we use (i) and (ii) $\{S \ge T\} = \{S < T\}^c \in \mathcal{F}_S \cap \mathcal{F}_T$:

$$\{S = T\} = \{S \le T\} \cap \{S \ge T\} \in \mathcal{F}_S \cap \mathcal{F}_T$$

Then we completed the proof. □

Remark 1.3. A short summary of the properties for s.t.s:

(1) For s.t. T, $T \in \mathcal{F}_T$, i.e. $\forall\, t > 0$, $\{T < t\} \in \mathcal{F}_T$.

(2) For two s.t.s, S and T, if $S \le T$, then $\mathcal{F}_S \subseteq \mathcal{F}_T$.

(3) For s.t. T, and $\forall\, t > 0$, then $T \wedge t \in \mathcal{F}_{T\wedge t} \subseteq \mathcal{F}_t$.

(4) For two s.t.s S and T, then $\mathcal{F}_{S\wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$.

Lastly, we need to justify its validity of changing time with s.t.:

> **Theorem 1.7.** *Let $X = \{X_t, t \ge 0\}$ be a progressively measurable process, and T be a s.t., then X_T defined on $\{T < +\infty\}$ is $\mathcal{F}_T$-measurable. And the stopped process $\{X_{T\wedge t}, t \ge 0\}$ is also progressively measurable.*

Proof. The two statements is equivalent, since if we want to show, X_T defined on $\{T < +\infty\}$ is $\mathcal{F}_T$-measurable, then it is equal to show, $\forall\, t \ge 0$, and $\forall\, B \in \mathcal{B}(\mathbb{R}^d)$:

$$\{X_T \in B\} \cap \{T \le t\} \in \mathcal{F}_t \Longleftrightarrow \{X_{T\wedge t} \in B\} \cap \{T \le t\} \in \mathcal{F}_t$$

Then it remains to show, $X_{T\wedge t} \in \mathcal{F}_t$, by observing:

$$(s, \omega) \mapsto (T(\omega) \wedge s, \omega)$$

is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable, and since the process is progressively measurable:

$$(s, \omega) \mapsto X_s(\omega)$$

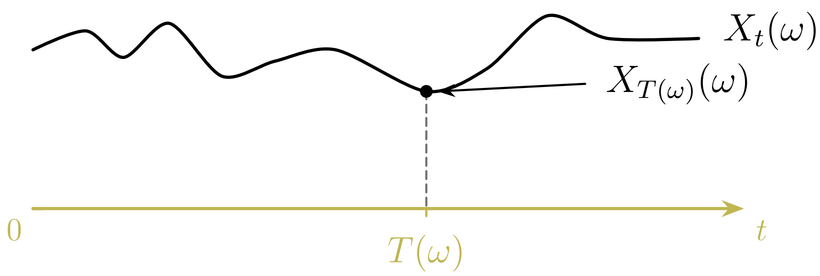
is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable, then by composition rule:

$$(s, \omega) \mapsto (T(\omega) \wedge s, \omega) \mapsto (X_{T(\omega) \wedge s}, \omega)$$

is also $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable, then we completed the proof. □

Remark 1.4. To remind, we interpret the stochastic process replaced time with s.t. (or random time), by following composite functions and figure:

$$\omega \mapsto (T(\omega), \omega) \mapsto X_{T(\omega)}(\omega)$$



1.4 Continuous-time Martingale

Martingale sits at an important role of this course, and also stochastic process, for this session, we mainly focus on how to transfer the classical discrete time martingale properties (where most of reader should learn from classical stochastic process like Rick. Durrett's book) to continuous time version, and most powerful tool for proving convergence.

Definition 1.12. A martingale is an adapted stochastic process $X = \{X_t, t \geq 0\}$ such that:

(a) $\forall\, t \geq 0, \mathsf{E}|X_t| < \infty$.

(b) $\forall\, t \geq s \geq 0, \mathsf{E}\left(X_t \mid \mathcal{F}_s\right) = X_s$

Remark 1.5. A simple extension to sub/super-martingale:

(1) If definition 1.12's (b) change to:

$$\mathsf{E}\left(X_t \mid \mathcal{F}_s\right) \geq X_s$$

, then it is called submartingale.

(2) If definition 1.12's (b) change to:

$$\mathsf{E}\left(X_t \mid \mathcal{F}_s\right) \leq X_s$$

, then it is called supermartingale.

Following we presents some important theorem some of them without proof:

Theorem 1.8. *Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a convex (concave up) function, if $X = \{X_t, t \geq 0\}$ is a martingale, then $\varphi(X) = \{\varphi(X_t), t \geq 0\}$ is a submartingale.*

Proof. By Jensen's inequality:

$$|\mathsf{E} X_t| \leq \mathsf{E}\,|X_t| \implies |\mathsf{E}\,(X_t \mid \mathcal{F}_s)| \leq \mathsf{E}\,(|X_t| \mid \mathcal{F}_s)$$

Hence, by assuming integrability and adaptivity, it remains:

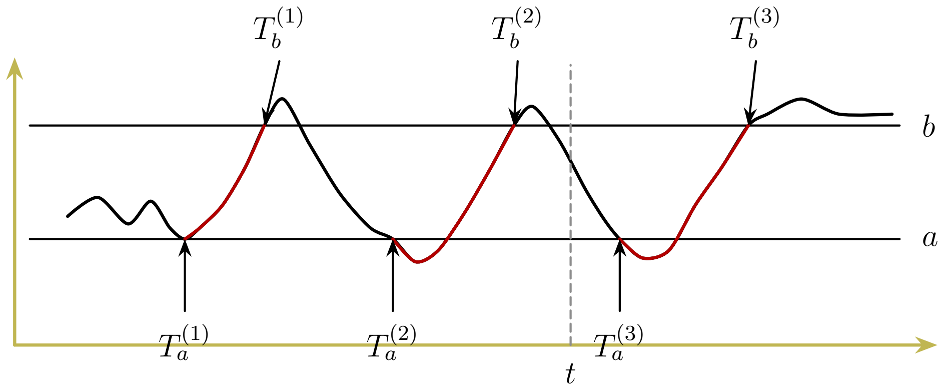
$$\mathsf{E}\,(\varphi(X_t) \mid \mathcal{F}_s) \geq \varphi\,(\mathsf{E}\,(X_t \mid \mathcal{F}_s)) = \varphi(X_s)$$

Then we completed the proof. □

Remark 1.6. Some common convex function we used are, where $p \geq 1$.

$$|X_t|, \quad X_t^2, \quad (X_t)^+ = \max(X_t, 0), \quad |X_t|^p$$

Next important theorem is **continuous-time Doob's upcrossing inequality**:



As the figure shows, we could denote the i th upcrossings (or downcrossings) time to be the first time, after the i-1 th downcrossings (or upcrossings), more formally:

$$\begin{aligned} T_b^0 &= -\infty \\ T_a^{(1)} &= \inf\left\{t \geq \left(T_b^{(0)}\right)^+ : X_t \leq a\right\} \\ T_b^{(1)} &= \inf\left\{t \geq T_a^{(1)} : X_t \geq b\right\} \end{aligned}$$

After the initial setup, for $\forall\, \ell \geq 2$:

$$\begin{aligned} T_a^{(\ell)} &= \inf\left\{t \geq T_b^{(\ell-1)} : X_t \leq a\right\} \\ T_b^{(\ell)} &= \inf\left\{t \geq T_a^{(\ell-1)} : X_t \geq b\right\} \end{aligned}$$

Then we focus on the # of upcrossing in an interval $[0, t]$ for given two bounds $(a < b)$:

$$U_{ab}^X[0, t] = \max\left\{k \geq 0 : T_b^{(k)} \leq n\right\}$$

Now, we could present our most important theorem about submartingale:

Theorem 1.9. ***(Continuous-time Doob's Upcrossings Inequality)*** *Let $X = \{X_t, t \geq 0\}$ be a right-continuous submartingale, then for all $T > 0$:*

$$\mathsf{E}U_{ab}^{X}[0,T] \leq \frac{\mathsf{E}(X_T - a)^+}{b-a}$$

Proof. The idea is to how to apply the discrete-time Doob's upcrossing inequality, where we assume to be known, since we presume all readers have studied basic stochastic process. Then we could try to discretise the continuous time, for each $n \geq 1$, we let:

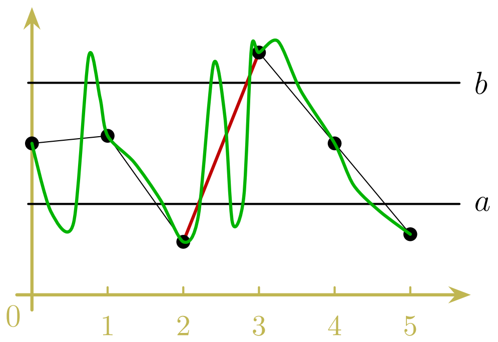
$$D_n = \mathbb{Z}/2^n = \left\{0, \frac{\pm 1}{2^n}, \frac{\pm 2}{2^n}, \frac{\pm 3}{2^n}, \cdots \right\}$$

Consider the upcrossings of X on $[0,T] \cap D_n$, where:

$$\{X_{j/2^n}, j \geq 0\}$$

, is **discrete-time submartingale**, then as showed as the figure below, we have:

$$U_{ab}^{X}([0,T] \cap D_n) \leq U_{ab}^{X}[0,T]$$



Clearly if we discretise it finer, it will be nondecreasing, then:

$$U_{ab}^{X}([0,T] \cap D_n) \uparrow U_{ab}^{X}[0,T]$$

Therefore, by using discrete-time upcrossing inequality, i.e.

$$U_{ab}^{X}([0,T] \cap D_n) \leq \frac{1}{b-a}\mathsf{E}(X_{T_n} - a)^+ \leq \frac{1}{b-a}\mathsf{E}(X_T - a)^+$$

, where $T_n := \frac{\lfloor T \cdot 2^n \rfloor}{2^n}$, let $n \to \infty$ by limiting theory, we finished the proof. □

Just as we said, it is powerful tool to show the convergence for submartingale:

Proposition 1.10. *If $X = \{X_t, t \geq 0\}$ is a right-continuous submartingale, and:*

$$\sup_{t \geq 0} \mathsf{E}X_t^+ < \infty$$

, then $\exists\, X \in L^1$ such that, as $t \longrightarrow \infty$:

$$X_t \xrightarrow{a.s.} X$$

Proof. Here, we ignore the procedure of showing existence, just focus on convergence:

$$\begin{aligned}
\mathsf{E}U_{ab}^{X}[0,\infty) &= \lim_{n\to\infty} \mathsf{E}U_{ab}^{X}[0,n] \\
&\le \limsup_{n\to\infty} \frac{1}{b-a}\mathsf{E}\left(X_n - a\right)^+ \\
&\le \limsup_{n\to\infty} \frac{1}{b-a}\left(\mathsf{E}X_n^+ + |a|\right) < \infty
\end{aligned}$$

, above is provided by theorem 1.9 and condition, hence:

$$\forall a < b, U_{ab}^{X}[0,,\infty) < \infty, \text{ a.s.} \quad \Longrightarrow \quad \text{w.p.1 } \forall\, a < b \in \mathbb{Q}, U_{ab}^{X}[0,\infty) < \infty$$

, this is given by the fact that countable additivity of our measure, therefore:

$$\liminf_{t\to\infty} X_t < a \le b < \limsup_{t\to\infty} X_t$$

does not occur, which is equivalent to say:

$$\liminf_{t\to\infty} X_t = \limsup_{t\to\infty} X_t$$

, then we completed the proof. □

To connect relations of different convergence modes, we need one more definition:

Definition 1.13. A family of r.v.s $\{X_t, t \ge 0\}$ is **uniformly integrable** (u.i.) if:

$$\lim_{M\to\infty} \sup_{t\ge 0} \mathsf{E}|X_t|\mathbf{1}_{\{|X_t|>M\}} = 0$$

Remark 1.7. Using Markov Inequality, if $\sup_{t\ge0} \mathsf{E}|X_t|^p < \infty$, for some $p > 1$, then X is u.i.

With **uniformly integrable**, we could present the following equvalence:

Theorem 1.11. *If $X_n \xrightarrow{p.} X$, then the following are equivalent:*

(a) $X_n \xrightarrow{L^1} X$.

(b) $\mathsf{E}|X_n| \longrightarrow \mathsf{E}|X|$.

(c) $\{X_n, n \ge 0\}$ *is u.i.*

Proof. We omit the proof since it is not our focus. □

Last important property is **last elements**, before that we need:

Proposition 1.12. *Let $Z \in L^1$, then following is u.i.*

$$\{\mathsf{E}\left(Z \mid \mathcal{G}\right) : \mathcal{G} \text{ is a sub } \sigma\text{-field of } \mathcal{F}\}$$

Proof. Proof omitted.

Remark 1.8. In particular, if we pick $\mathcal{G}$ to be filtration, i.e.:

$$\{\mathsf{E}(Z \mid \mathcal{F}_t), t \geq 0\}$$

, is u.i. and also a **martingale**.

□

With the help of u.i., we could show following relations, (last element), providing convenience in proving **Optional Sampling Theorem** (OST).

Proposition 1.13. *A right-continuous martingale* $\{X_t, t \geq 0\}$ *is u.i. iff:*

$$\exists\, X_\infty \in L^1,\ \ s.t.\ \ X_t = \mathsf{E}(X_\infty \mid \mathcal{F}_t)$$

Proof. Firstly, we start with the inclusion:

$\Longrightarrow$ Firstly, by u.i., we have $\sup_{t\geq 0} \mathsf{E}|X_t| < \infty$, then through proposition 1.10:

$$\exists\, X_\infty,\ \text{s.t.}\ \ X_t \xrightarrow{a.s.} X_\infty$$

, more important, we know it is L^1 since:

$$\mathsf{E}|X_\infty| = \mathsf{E} \lim_{t\to\infty} |X_t| \leq \mathsf{E} \liminf_{t\to\infty} \mathsf{E}|X_t| < \infty$$

, where last inequality comes from **Fatou's Lemma**, next to prove:

$$\forall\, t \geq 0, X_t = \mathsf{E}(X_\infty \mid \mathcal{F}_t)$$

is equivalent, by the definition of conditional expectation, to show $\forall\, A \in \mathcal{F}_t$:

$$\mathsf{E}X_t \mathbf{1}_A = \mathsf{E}X_\infty \mathbf{1}_A$$

Common as usual, we approximate it by sending limit to infinity, for any $s \geq 0$:

$$\mathsf{E}(X_{t+s} \mid \mathcal{F}_t) = X_t \implies \mathsf{E}X_{t+s}\mathbf{1}_A = \mathsf{E}X_t\mathbf{1}_A$$

For limiting validity, we recall that $u.i. + a.s. \implies L^1$:

$$|\mathsf{E}X_{t+s}\mathbf{1}_A - \mathsf{E}X_\infty\mathbf{1}_A| \leq \mathsf{E}|X_{t+s} - X_\infty| \longrightarrow 0$$

Then by sending $s \to \infty$ we complete the fisrt part of the proof.

$\Longleftarrow$ This direction immediately follows from proposition 1.12.

Then we completed the proof by showing both directions. □

After all the preparation, we commence the OST:

Theorem 1.14. ***(Optional Sampling Theorem)***
Let $\{X_t, t \geq 0\}$ *be a right-continuous martingale, let* $S \leq T$ *be two s.t.s suppose:*

(a) $\{X_t, t \geq 0\}$ *is u.i.*

(b) $\exists$ *constant* $K > 0$*, s.t.* $S, T \leq K$

, either above two satisfied, then we have:

$$\mathsf{E}(X_T \mid \mathcal{F}_S) = X_S$$

Proof. In order to use proposition 1.13, we consider:

(a) Then by proposition 1.10, we choose $Z = X_\infty$

(b) We directly pick the "last element", i.e. $Z = X_K$

In either case, we have: $X_t = \mathsf{E}(Z \mid \mathcal{F}_t)$, then it suffices to show:

$$X_T = \mathsf{E}(Z \mid \mathcal{F}_T)$$

Since provided the above satified, by symmetry, we could interchanged it to S, so that:

$$X_S = \mathsf{E}(Z \mid \mathcal{F}_S)$$

Hence, by tower rule and proposition 1.5, then:

$$\mathsf{E}(X_T \mid \mathcal{F}_S) = \mathsf{E}[\mathsf{E}(Z \mid \mathcal{F}_T) \mid \mathcal{F}_S] = \mathsf{E}(Z \mid \mathcal{F}_S) = X_S$$

Now, we could firstly consider the discrete cases:

(1) If T can be decomposed into discrete cases, i.e. $T \in \{t_1, t_2, \cdots\}$, then $\forall\ A \in \mathcal{F}_T$:

$$\mathsf{E}X_T\mathbf{1}_A = \mathsf{E}Z\mathbf{1}_A = \sum_{n=1}^{\infty} \mathsf{E}X_{t_n}\mathbf{1}_{\{T=t_n\}}\mathbf{1}_A$$

$$\text{Because } \begin{cases} A \cap \{T = t_n\} \in \mathcal{F}_{t_n} \\ \mathsf{E}(Z \mid \mathcal{F}_{t_n}) = X_{t_n} \end{cases}$$

$$= \sum_{n=1}^{\infty} \mathsf{E}Z\mathbf{1}_{A\cap\{T=t_n\}} = \mathsf{E}Z\mathbf{1}_A$$

(2) For general T, we discretise from right (for the sake of right-continuous), then define:

$$T_k = \frac{\lfloor 2^k \cdot T \rfloor + 1}{2^k}$$

, clearly $T_k \downarrow T$, then $\forall\ A \in \mathcal{F}_T \subseteq \mathcal{F}_{T_k}$:

$$\mathsf{E}(Z \mid \mathcal{F}_{T_k}) = X_{T_k} \implies \mathsf{E}Z\mathbf{1}_A = \mathsf{E}X_{T_k}\mathbf{1}_A$$

Sending the $k \to \infty$ to see $\mathsf{E}Z\mathbf{1}_A = \mathsf{E}X_T\mathbf{1}_A$, which is provided by:

$$\begin{cases} X_{T_k} \xrightarrow{a.s.} X_T \text{ by right-continuous} \\ \{X_{T_k}\} \text{ is u.i. by proposition 1.12} \end{cases} \implies X_{T_k} \xrightarrow{L^1} X_T$$

Hence, we completed the proof, by the help of dicretisation. □

By closing subsections for martingale, we presents some useful inequality:

Theorem 1.15. *(Doob's Maximal Inequality)*
Let $X = \{X_t, t \geq 0\}$ be a ***right-continuous submartingale****, then $\forall\ \lambda > 0$:*

$$\mathsf{P}\left(\sup_{0\leq s\leq t} X_s > \lambda\right) \leq \frac{1}{\lambda}\mathsf{E}X_t^+$$
$$\mathsf{P}\left(\inf_{0\leq s\leq t} X_s < -\lambda\right) \leq \frac{1}{\lambda}\left(\mathsf{E}X_t^+ - \mathsf{E}X_0\right)$$

Proof. Since two of the statements are similar, we only show first one, $\forall\ \lambda > 0$, define:

$$T = \inf\{t \geq 0 : X_t > \lambda\} \implies \mathsf{P}\left(\sup_{0\leq s\leq t} X_s > \lambda\right) = \mathsf{P}(T \leq t)$$

, then consider a stopped process, $\{X_{t\wedge T}, t \geq 0\}$, notice that:

$$\begin{aligned}\lambda\mathsf{P}\left(\sup_{0\leq s\leq t} X_s > \lambda\right) &= \lambda\mathsf{P}(T \leq t) = \lambda\mathsf{E}\mathbf{1}_{\{T\leq t\}} \leq \mathsf{E}X_T\mathbf{1}_{\{T\leq t\}} \\ &= \mathsf{E}X_{t\wedge T}\mathbf{1}_{\{T\leq t\}} = \mathsf{E}X_{t\wedge T}^+\mathbf{1}_{\{T\leq t\}} \leq \mathsf{E}X_{t\wedge T}^+ \leq \mathsf{E}X_t^+\end{aligned}$$

Hence, we completed proof, by transfering original question to new stopping time. □

The above theorem is to consider the tail probability, and next one is for moment:

Theorem 1.16. *Let $X = \{X_t, t \geq 0\}$ be submartingale, then $\forall\ p > 1$:*

$$\mathsf{E}\left(\sup_{0\leq s\leq t} |X_s|\right)^p \leq \left(\frac{p}{p-1}\right)^p \mathsf{E}\,|X_t|^p$$

Proof. Similar to previous inequality, here we need more dedicate tail probability estimation, where we first define $Y = \sup_{0\leq s\leq t}|X_s|$, then:

$$\begin{aligned}\lambda\mathsf{P}\left(Y \geq \lambda\right) &\leq \mathsf{E}\,|X_{T\wedge t}|\,\mathbf{1}_{\{T\leq t\}} = \mathsf{E}\,|X_{T\wedge t}| - \mathsf{E}\,|X_{T\wedge t}|\,\mathbf{1}_{\{T>t\}} \\ &\leq \mathsf{E}\,|X_t| - \mathsf{E}\,|X_t|\,\mathbf{1}_{\{T>t\}} = \mathsf{E}\,|X_t| - \mathsf{E}\,|X_t|\,\mathbf{1}_{\{Y<\lambda\}} = \mathsf{E}\,|X_t|\,\mathbf{1}_{\{Y\geq\lambda\}}\end{aligned}$$

Then, using basic calculus to expand the moment:

$$\begin{aligned}\mathsf{E}Y^p &= \int_0^\infty p\lambda^{p-1}\mathsf{P}(Y \geq \lambda)\mathrm{d}\lambda \leq \int_0^\infty p\lambda^{p-2}\mathsf{E}\,|X_t|\,\mathbf{1}_{\{Y\geq\lambda\}}\mathrm{d}\lambda \\ &= \mathsf{E}\,|X_t| \int_0^\infty p\lambda^{p-2}\mathbf{1}_{\{Y\geq\lambda\}}\mathrm{d}\lambda = \mathsf{E}\,|X_t| \int_0^Y p\lambda^{p-2}\mathrm{d}\lambda \\ &= \frac{p}{p-1}\mathsf{E}\,|X_t|\,Y^{p-1} \leq \frac{p}{p-1}\left(\mathsf{E}\,|X_t|^p\right)^{1/p}\left(\mathsf{E}Y^p\right)^{(p-1)/p}\end{aligned}$$

The last inequality is due to holder's inequality, then we arrange terms to get:

$$\left(\mathsf{E}Y^p\right)^{1/p} \leq \frac{p}{p-1}\left(\mathsf{E}\,|X_t|^p\right)^{1/p}$$

For unbounded case, we stopped the process by $Y \wedge k$, and then sending k to inifinity, hence we could easily completed the proof. □

1.5 The Doob-Meyer Decomposition

Recall the discretised version of **Doob Decomposition**:

$$\underbrace{X_n}_{\text{Submartingale}} = \underbrace{M_n}_{\text{Martingale}} + \underbrace{A_n}_{\substack{\text{Predictable}\\\text{Increasing}}}$$

The intuition is simple, from above construction:

$$\mathsf{E}\left(M_{n+1} \mid \mathcal{F}_n\right) = M_n \implies \mathsf{E}\left(X_{n+1} - A_{n+1} \mid \mathcal{F}_n\right) = X_n - A_n$$

Using the predictability of $\{A_n, n \in \mathbb{N}\}$, i.e. $A_n \in \mathcal{F}_{n-1}$:

$$\mathsf{E}\left(X_{n+1} \mid \mathcal{F}_n\right) - A_{n+1} = X_n + A_n \implies A_{n+1} - A_n = \mathsf{E}\left(X_{n+1} \mid \mathcal{F}_n\right) - X_n \geq 0$$

Since $\{X_n, n \in \mathbb{N}\}$ is submartingale, then in general, it is increasing, hence A_n is to capture the additional increasing parts, and martingale can be seen as "white noise", by iterating, and by default to set $A_0 = 0$, we have:

$$A_n = \sum_{k=0}^{n-1} \left[\mathsf{E}\left(X_{k+1} \mid \mathcal{F}_k\right) - X_k\right]$$

Hence, this subsection is to delve into such decomposition with continuous setting:

$$\underbrace{X_t}_{\text{Submartingale}} = \underbrace{M_t}_{\text{Martingale}} + \underbrace{A_t}_{\substack{\text{Predictable}\\\text{Increasing}}}$$

, where assuming $\{X_t, t \geq 0\}$ is **right-continuous** submartingale. First is to consider increasing process in continuous setting:

Definition 1.14. An adapted process $A = \{A_t, t \geq 0\}$ is called **increasing** if:

(a) $A_0 = 0$ a.s.

(b) $t \mapsto A_t$ is non-decreasing and right-continuous for ω-a.e.

(c) $\mathsf{E}A_t < \infty$, $\forall\, t \geq 0$.

Remark 1.9. Note that for integrability condition is a little different from (c):

$$\mathsf{E}A_\infty = \mathsf{E} \lim_{t\to\infty} A_t < \infty$$

Besides increasing, other than predictable, we introduce natural:

Definition 1.15. An increasing process $A = \{A_t, t \geq 0\}$ is called **natural** if $\forall$ bounded, right continuous martingale, $M = \{M_t, t \geq 0\}$, then $\forall\, t \geq 0$:

$$\mathsf{E} \int_0^t M_s \mathrm{d}A_s = \mathsf{E} \int_0^t M_{s-} \mathrm{d}A_s$$

Remark 1.10. It is easy to see , if $s \mapsto A_s$ is a.s. continuous, then A is natural, since $M_s - M_{s-} \neq 0$ only for jump points, for a.s. continuity, there are only finite of them.

It seems wired to propose this definition, by thinking from discretised version:

$$A_n \in \mathcal{F}_{n-1} \Longleftrightarrow \text{predictable}$$
$$A_t \in \mathcal{F}_{t-} \Longleftrightarrow \text{"predictable"}$$

Hence, we first check its discrete-time version:

Definition 1.16. An adapted and increasing sequence, $A = \{A_n, n \in \mathbb{N}\}$ is called **natural** if $\forall$ bounded martingale, $M = \{M_n, n \in \mathbb{N}\}$, then, $\forall\, n \in \mathbb{N}$:

$$\mathsf{E}\left(M_n A_n\right) = \mathsf{E}\left[\sum_{k=1}^{n} M_{k-1}\left(A_k - A_{k-1}\right)\right]$$

Then we are eager to see the connection between **natural** and **predictable**:

Proposition 1.17. *For in creasing sequence, A is predictable iff it is natural.*

Proof. The proof is direct, mainly by dedicate algebra:

$\Longrightarrow$ To prove, it is natural, it is same to show for every martingale M:

$$\mathsf{E}\left(M_n A_n\right) = \mathsf{E}\left[\sum_{k=1}^{n} M_{k-1}\left(A_k - A_{k-1}\right)\right] = \sum_{k=1}^{n}\left[\mathsf{E}M_{k-1}A_k - \mathsf{E}M_{k-1}A_{k-1}\right]$$

Now we could rearrange things to put minus term to left hand side:

$$\sum_{k=1}^{n}\mathsf{E}M_{k-1}A_{k-1} + \mathsf{E}M_nA_n = \sum_{k=1}^{n}\mathsf{E}M_{k-1}A_k$$

Since $A_0 = 0$ and M_nA_n serves as final term of first summation, then:

$$\sum_{k=1}^{n}\mathsf{E}M_kA_k = \sum_{k=1}^{n}\mathsf{E}M_{k-1}A_{k-1} \implies \sum_{k=1}^{n}\mathsf{E}\left(M_k - M_{k-1}\right)A_k = 0$$

From this equivalent statement, we plug in the predictability, if $A_n \in \mathcal{F}_{n-1}$:

$$\mathsf{E}A_n(M_n - M_{n-1}) = \mathsf{E}\left[A_n\mathsf{E}\left(M_n - M_{n-1} \mid \mathcal{F}_{n-1}\right)\right] = 0$$

Then each term of the summation is zero, we completed this direction.

$\Longleftarrow$ Assume A is natural, and $\forall$ bounded martingale M, from abovem, then:

$$\mathsf{E}A_n\left(M_n - M_{n-1}\right) = 0,\ \forall\, n \geq 1$$

Then goal is to show predictability, i.e. $A_n \in \mathcal{F}_{n-1}$, or $A_n = \mathsf{E}\left(A_n \mid \mathcal{F}_{n-1}\right)$ a.s., now since A is natural, then for every bounded M, we notice that:

$$\mathsf{E}M_n\left[A_n - \mathsf{E}\left(A_n \mid \mathcal{F}_{n-1}\right)\right] = \mathsf{E}A_nM_{n-1} - \mathsf{E}\left[M_n\mathsf{E}\left(A_n \mid \mathcal{F}_{n-1}\right)\right]$$
$$= \mathsf{E}M_{n-1}\left[A_n - \mathsf{E}\left(A_n \mid \mathcal{F}_{n-1}\right)\right] + \mathsf{E}\left[\mathsf{E}\left(A_n \mid \mathcal{F}_{n-1}\right)\left(M_n - M_{n-1}\right)\right]$$
$$= \underbrace{\mathsf{E}M_{n-1}A_n - \mathsf{E}\mathsf{E}\left(A_nM_{n-1} \mid \mathcal{F}_{n-1}\right)}_{=0} + \mathsf{E}\left[\mathsf{E}\left(A_n \mid \mathcal{F}_{n-1}\right)\underbrace{\mathsf{E}\left(M_n - M_{n-1} \mid \mathcal{F}_{n-1}\right)}_{=0}\right] = 0$$

Hence, by the arbitrarity, for each fixed n, we choose:

$$\xi_n = \operatorname{sgn}\left(A_n - \mathsf{E}\left(A_n \mid \mathcal{F}_{n-1}\right)\right)$$

, and define our martingale as:

$$M_k = \begin{cases} \mathsf{E}\left(\xi_n \mid \mathcal{F}_k\right) & 0 \le k \le n-1 \\ \xi_n & k \ge n \end{cases}$$

, which is bounded, and martingle, from this well-designed martingale, we know:

$$0 = \mathsf{E}M_n\left[A_n - \mathsf{E}\left(A_n \mid \mathcal{F}_{n-1}\right)\right] = \mathsf{E}\left|A_n - \mathsf{E}\left(A_n \mid \mathcal{F}_{n-1}\right)\right|$$

, above gives us:

$$A_n = \mathsf{E}\left(A_n \mid \mathcal{F}_{n-1}\right) \text{ a.s.} \quad \Longrightarrow \quad A_n \in \mathcal{F}_{n-1}$$

Hence, we completed the proof by showing both directions. $\square$

Then next thing is to bridge the discrete one with continuous one:

$$\mathsf{E}\left(M_n A_n\right) = \mathsf{E}\left[\sum_{k=1}^{n} M_{k-1}\left(A_k - A_{k-1}\right)\right] \Longleftrightarrow \mathsf{E}\int_0^t M_s \mathrm{d}A_s = \mathsf{E}\int_0^t M_{s-} \mathrm{d}A_s$$

Hence, use the traditional discretisation method, by partition, $\Pi = \{t_0, t_1, \cdots, t_n\} \subseteq [0, t]$ with $0 = t_0 < t_1 < t_2 < \cdots < t_n = t$ and by setting $\|\Pi\| = \max_{1 \le k \le n} |t_k - t_{k-1}|$:

$$\int_0^t M_s \mathrm{d}A_s = \lim_{\|\Pi\| \to 0} \int_0^t M_s^{\Pi} \mathrm{d}A_s$$

, where by right continuous, we chose our discrete martingale to be:

$$M_s^{\Pi} = \sum_{k=1}^{n} M_{t_k} \mathbf{1}_{(t_{k-1}, t_k]}(s)$$

Then, by this setup, for each Π, we could see:

$$\begin{aligned} \mathsf{E}\int_0^t M_s^{\Pi} \mathrm{d}A_s &= \mathsf{E}\left[\sum_{k=1}^{n} M_{t_k}\left(A_{t_k} - A_{t_{k-1}}\right)\right] = \mathsf{E}M_{t_n}A_{t_n} + \sum_{k=1}^{n-1} \mathsf{E}M_{t_k}A_{t_k} - \sum_{k=2}^{n} \mathsf{E}M_{t_k}A_{t_{k-1}} \\ &= \mathsf{E}M_t A_t + \sum_{k=1}^{n-1} \mathsf{E}M_{t_k}A_{t_k} - \sum_{k=1}^{n-1} \mathsf{E}M_{t_{k+1}}A_{t_k} \\ &= \mathsf{E}M_t A_t - \sum_{k=1}^{n-1} \mathsf{E}\left[A_{t_k} \underbrace{\mathsf{E}\left(M_{t_{k+1}} - M_{t_k} \mid \mathcal{F}_{t_k}\right)}_{=0}\right] = \mathsf{E}M_t A_t \end{aligned}$$

Hence, we see the connection between discrete version and continuous one, next:

Definition 1.17. Let $S = \{$ stopping times T with $T < \infty$ a.s. $\}$, we say a right-continuous $X_t, t \geq 0$ is of **class D** if $\{X_T, T \in S\}$ is u.i., furthermore, to fix $a > 0$, let $S_a = \{$ stopping times T with $Tlea$ a.s. $\}$, then $\{X_t, t \geq 0\}$ is of **class DL** if $\{X_T, T \in S_a\}$ is u.i. for every $a > 0$.

Then Doobs and Meyer found membership of class DL is sufficient for decomposition:

Theorem 1.18. ***(Doob-Meyer Decomposition)***
If $\{X_t, t \geq 0\}$ is of ***class DL****, and a submartingale, then it admits, $\forall\, t \geq 0$:*

$$X_t = M_t + A_t$$

, with $\{M_t, t \geq 0\}$ is martingale, $\{A_t, t \geq 0\}$ is increasing process. In particular, if we resrtict $\{A_t, t \geq 0\}$ to be natural, then such decomposition is unique.

Proof. Let us only consider the special cases, where we admit A is natural:

(1) For uniqueness, we assume it admits two different decomposition:

$$X_t = \begin{cases} M_t + A_t \\ M_t' + A_t' \end{cases} \implies B_t := M_t - M_t' = A_t' - A_t$$

, where $\{B_t, t \geq 0\}$ is martingale and of bounded variation , also natural, from the definition and exercise before, $\forall$ bounded continuous martingale $\{\xi_t, t \geq 0\}$:

$$\mathsf{E}\xi_t(A_t - A_t') = \mathsf{E}\int_0^t \xi_{s-}(\mathrm{d}A_s - \mathrm{d}A_s') = \mathsf{E}\int_0^t \xi_{t-}\mathrm{d}B_s = \lim_{n\to\infty} \mathsf{E}\sum_{j=1}^{m_n} \xi_{t_{j-1}^{(n)}}\left(B_{t_j^{(n)}} - B_{t_{j-1}^{(n)}}\right)$$

, where $\Pi_n = \{t_0^{(n)}, t_1^{(n)}, \cdots, t_{m_n}^{(n)}\} \subseteq [0, t]$, for each term:

$$\mathsf{E}\xi_{t_{t_{j-1}}^{(n)}}\left(B_{t_j^{(n)}} - B_{t_{j-1}^{(n)}}\right) = \mathsf{E}\left[\xi_{t_{t_{j-1}}^{(n)}} \underbrace{\mathsf{E}\left(B_{t_j^{(n)}} - B_{t_{j-1}^{(n)}} \mid \mathcal{F}_{t_{j-1}^{(n)}}\right)}_{=0}\right] = 0$$

, then by similar trick, we obtained:

$$\mathsf{E}\,|A_t - A_t'| = 0 \implies A_t = A_t' \text{ a.s. } \forall\, t > 0 \implies A_t = A_t' \text{ a.s. } \forall\, t \in \mathbb{Q}^+$$

Lastly, by right-continuity, we achieved $A_t = A_t'$ a.s. $\forall\, t \geq 0$.

(2) For existence, fix $a > 0$, it suffices to construct the decomposition on $[0, a]$, consider:

$$Y_t := X_t - \mathsf{E}(X_a \mid \mathcal{F}_t) \leq 0$$

Then $\{Y_t, 0 \leq t \leq a\}$ is right-continuous submartingale.Consider dyadic partition:

$$\Pi_n = \left\{t_0^{(n)}, t_1^{(n)}, \cdots, t_{2^n}^{(n)}\right\}, \quad t_j^{(n)} = \frac{j}{2^n}a, \quad 0 \leq j \leq 2^n$$

For each $n \geq 1$, the discrete-time submartingale admits the Doob decomposition

$$Y_{t_j^{(n)}} = M_{t_j^{(n)}}^{(n)} + A_{t_j^{(n)}}^{(n)}, \qquad 0 \leq j \leq 2^n$$

where $A_0^{(n)} = 0$, and for $1 \leq j \leq 2^n$,

$$A_{t_j^{(n)}}^{(n)} = \sum_{k=0}^{j-1} \mathsf{E}\Big(Y_{t_{k+1}^{(n)}} - Y_{t_k^{(n)}} \,\Big|\, \mathcal{F}_{t_k^{(n)}}\Big)$$

In particular, $\{A_{t_j^{(n)}}^{(n)}\}_{0 \leq j \leq 2^n}$ is predictable and increasing, notice:

$$Y_a = X_a - \mathsf{E}(X_a \mid \mathcal{F}_a) = 0 \implies 0 = Y_a = M_a^{(n)} + A_a^{(n)}$$

Hence, we may rewrite this into:

$$M_{t_j^{(n)}}^{(n)} = \mathsf{E}\big(M_a^{(n)} \mid \mathcal{F}_{t_j^{(n)}}\big) = -\mathsf{E}\big(A_a^{(n)} \mid \mathcal{F}_{t_j^{(n)}}\big) \implies Y_{t_j^{(n)}} = A_{t_j^{(n)}}^{(n)} - \mathsf{E}\big(A_a^{(n)} \mid \mathcal{F}_{t_j^{(n)}}\big)$$

We now claim that the family $\{A_a^{(n)}\}_{n \geq 1}$ is u.i.(you may verifty). Hence, by the Dunford-Pettis compactness criterion, there exist an integrable random variable A_a and a subsequence $\{A_a^{(n_k)}\}_{k \geq 1}$ such that $A_a^{(n_k)} \xrightarrow{L^1} A_a$, i.e.

$$\lim_{k \to \infty} \mathsf{E}(\xi A_a^{(n_k)}) = \mathsf{E}(\xi A_a)$$

for every bounded random variable ξ. To simplify notation, we relabel this subsequence again by $\{A_a^{(n)}\}$, now define

$$A_t := Y_t + \mathsf{E}(A_a \mid \mathcal{F}_t), \qquad 0 \leq t \leq a$$

where we choose its right-continuous modification. Likewise, for each n, define

$$A_t^{(n)} := Y_t + \mathsf{E}(A_a^{(n)} \mid \mathcal{F}_t), \qquad 0 \leq t \leq a$$

again with the right-continuous version. Then, for every $t \in \Pi_n$, this agrees with the discrete compensator constructed above, let

$$\mathbb{T} := \bigcup_{n=1}^{\infty} \Pi_n$$

By the weak L^1 stability of conditional expectation, for every $t \in \mathbb{T}$ and every bounded random variable ξ, we have

$$\lim_{n \to \infty} \mathsf{E}\big(\xi A_t^{(n)}\big) = \mathsf{E}(\xi A_t)$$

Now take $s, t \in \mathbb{T}$ with $0 \leq s < t \leq a$, and let ξ be bounded and nonnegative. For all large n, both s, t belong to Π_n, hence

$$\mathsf{E}\big[\xi(A_t - A_s)\big] = \lim_{n \to \infty} \mathsf{E}\big[\xi(A_t^{(n)} - A_s^{(n)})\big] \geq 0$$

because $\{A_u^{(n)}\}_{u \in \Pi_n}$ is increasing. Choosing $\xi = \mathbf{1}_{\{A_s > A_t\}}$, we obtain

$$A_s \leq A_t \quad \text{a.s.}$$

Since $\mathbb{T}$ is countable and by right-continuity, it is then nondecreasing on the whole interval $[0, a]$. Therefore, $\{A_t, 0 \le t \le a\}$ is an increasing process. Next:

$$Y_t = A_t - \mathsf{E}(A_a \mid \mathcal{F}_t) = X_t - \mathsf{E}(X_a \mid \mathcal{F}_t) \implies X_t = A_t + \mathsf{E}(X_a - A_a \mid \mathcal{F}_t)$$

Define $M_t := \mathsf{E}(X_a - A_a \mid \mathcal{F}_t)$, then it is a right-continuous martingale, and

$$X_t = M_t + A_t$$

It remains to show that A is natural. Let $\{\xi_t, 0 \le t \le a\}$ be any bounded right-continuous martingale. For each n, since $\{A^{(n)}_{t^{(n)}_j}\}_{0 \le j \le 2^n}$ is predictable on Π_n:

$$\begin{aligned} \mathsf{E}(\xi_a A^{(n)}_a) &= \mathsf{E}\left[\sum_{j=1}^{2^n} \xi_{t^{(n)}_{j-1}} \left(A^{(n)}_{t^{(n)}_j} - A^{(n)}_{t^{(n)}_{j-1}}\right)\right] = \mathsf{E}\left[\sum_{j=1}^{2^n} \xi_{t^{(n)}_{j-1}} \left(Y_{t^{(n)}_j} - Y_{t^{(n)}_{j-1}}\right)\right] \\ &= \mathsf{E}\left[\sum_{j=1}^{2^n} \xi_{t^{(n)}_{j-1}} \left(A_{t^{(n)}_j} - A_{t^{(n)}_{j-1}}\right)\right] \end{aligned}$$

Letting $n \to \infty$, the left-hand side converges to $\mathsf{E}(\xi_a A_a)$ by weak convergence:

$$\mathsf{E}(\xi_a A_a) = \mathsf{E}\int_0^a \xi_s \mathrm{d}A_s = \mathsf{E}\int_0^a \xi_{s-} \mathrm{d}A_s$$

, so A is natural. Hence, on $[0, a]$, we have constructed

$$X_t = M_t + A_t, \qquad 0 \le t \le a,$$

, with M a right-continuous martingale, and A a natural increasing process.

Then we completed this long proof. □

2 Stochastic Integral

3 Brownian Motion & PDE

4 Stochastic Differential Equations