Martingales and Stochastic Calculus

Alex Rutar*
University of Waterloo

Winter 2020†

*arutar@uwaterloo.ca
†Last updated: March 9, 2020

Contents

I. Stochastic Calculus

1 Measure Theory for Probability

Definition. Given a measure space $(\Omega, \mathcal{F}, \mathbb{P})$, a measurable function $X : \Omega \to \mathbb{R}$ is called a **random variable**. If X is a random variable, then we define the **distribution** of X to be the measure Borel measure μ on $\mathbb{R}$ given by $\mu(E) = \mathbb{P}(X^{-1}(E))$.

1.1 Conditional Expectation

1.1 Theorem. (Kolmogorov) *Suppose $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G} \subset \mathcal{F}$ is a sub-σ-algebra. Then there exists some $Z \in L^1(\Omega, \mathcal{G}, \mathbb{P})$ such that for each $A \in \mathcal{G}$*

$$\int_A X d\,\mathbb{P} = \int_A Z d\,\mathbb{P}.$$

Moreover, if $\tilde{Z} \in L^1(\Omega, \mathcal{G}, \mathbb{P})$ satisfies the above constraint, then $\mathbb{P}\{x \in \Omega : \tilde{Z}(x) = Z(x)\} = 1$.

Definition. In the context of the above theorem, we write $Z = \mathbb{E}[X|\mathcal{G}]$ and call $\mathbb{E}[X|\mathcal{G}]$ a **conditional expectation** with respect to $\mathcal{G}$. Certainly $\mathbb{E}[X|\mathcal{G}]$ nee not be pointwise unique.

1.2 Theorem. (Properties of Conditional Expectation) *Suppose X, Y are random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra. Then*

1. *$\mathbb{E}[X|\mathcal{G}] \geq 0$ a.s. whenever $X \geq 0$ a.s.*
2. *For $\alpha, \beta \in \mathbb{R}$ and $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{E}[\alpha X + \beta Y|\mathcal{G}] = \alpha\,\mathbb{E}[X|\mathcal{G}] + \beta\,\mathbb{E}[Y|\mathcal{G}]$ a.s.*
3. *If $XY \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and Y is $\mathcal{G}$–measurable, then $\mathbb{E}[XY|\mathcal{G}] = Y\,\mathbb{E}[X|\mathcal{G}]$*
4. *If $\mathcal{H} \subset \mathcal{G}$ is a σ-algebra, then $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$ a.s.*
5. *$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$*
6. *If $\mathcal{H}$ is a σ-algebra which is independent of the σ–algebra $\sigma\{\sigma\{X\}, \mathcal{G}\}$, then $\mathbb{E}[X|\sigma\{\mathcal{G}, \mathcal{H}\}] = \mathbb{E}[X|\mathcal{G}]$. In particular, $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ a.s. whenever $\sigma\{X\}$ and $\mathcal{H}$ are independent.*

1.2 Stochastic Processes

Definition. A **stochastic process** $X = \{X_t\}_{t \in \mathcal{T}}$ is a collection of random variables defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Typically, we assume that $\mathcal{T} = \mathbb{Z}_{\geq 0}$ or $\mathcal{T} = \mathbb{R}_{\geq 0}$ and equip $\mathcal{T}$ with the order topology Intuitively, we expect t to be a discrete or continuous time parameter. Given some $\omega \in \Omega$, the map $t \mapsto X_t(\omega)$ is called a **realization** or **path** of this process. One of the goals of this section is to treat $\{X_t\}_{t \geq 0}$ as a random element in some path space, equipped with a proper σ–algebra and probability.

Definition. Let $\Phi \subseteq L^p(X, \mathcal{M}, \mu)$. We then say that Φ is **uniformly L^p** if to every $\epsilon > 0$, there exists some $\delta > 0$ such that for any $A \in \mathcal{M}$

$$\int_A |f|^p \,\mathrm{d}\mu < \epsilon$$

whenever $\mathbb{P}(A) < \delta$ and $f \in \Phi$.

In the probability setting, we equivalently require that

$$\lim_{c\to\infty} \sup_{X\in\Phi} \mathbb{E}[|X| : |X| \geq c] = 0.$$

We say that a stochastic process is $\boldsymbol{L^p}$ if $\{X_t\}_{t\in\mathcal{T}} \subset L^p$.

Consider $X_t(\omega)$ as a function $X : \mathcal{T} \times \Omega \to \mathbb{R}$ equipped with the product $\sigma-$algebra.

Definition. The **distribution** of a stochastic process is the collection of all its finite-dimensional distributions, i.e. the collection of all the distributions of $\{X_{t_1}, \dots, X_{t_k}\}$ for and $k \in \mathbb{N}$ and $t_i \in \mathcal{T}$.

There are a number of ways to say that two processes X and Y are equivalent, which we organize in decreasing order of strength.

Definition. Let $X = \{X_t\}_{t\in\mathcal{T}}$ and $Y = \{Y_t\}_{t\in\mathcal{T}}$ be stochastic processes.

- We say that X and Y are called **indistinguishable** if almost all their sample paths agree; in other words,

$$\mathbb{P}(\omega \in \Omega : X_t(\omega) = Y_t(\omega) \text{ for all } t \in \mathcal{T}) = 1.$$

- We say that Y is a **modification** of X if for each $t \in \mathcal{T}$ we have $\mathbb{P}(\omega \in \Omega : X_t(\omega) = Y_t(\omega)) = 1$.
- Finally, X and Y are said to have the **same distribution** if all the finite dimensional distributions agree. In other words, if for all $n \in \mathbb{N}$ and $t_1 < \cdots < t_n$, with $t_i \in \mathcal{T}$, we have $(X_{t_1}, \dots, X_{t_n}) \overset{d}{=} (Y_{t_1}, \dots, Y_{t_n})$.

Example. Let X be a continuous stochastic process and N a Poisson point process on $[0, \infty)$. Then define

$$Y_t := \begin{cases} X_t & : t \notin N \\ X_t + 1 & : t \in N \end{cases}$$

Thus $\mathbb{P}(X_t = Y_t) = 1$ for all t, so X is a modification of Y. However, $\mathbb{P}(X_t = Y_t, t \geq 0) = 0$, so that X and Y are not indistinguishable.

A filtration formalizes the idea of "information acquired over time".

Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A **filtration** is a non-decreasing family $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$ of sub-σ-algebras of $\mathcal{F}$ so that $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for any $s < t$ in $\mathcal{T}$. We write $F_\infty = \sigma(\bigcup_{t\in\mathcal{T}} \mathcal{F}_t)$.

Let $\{X_t\}_{t\in\mathcal{T}}$ be a stochastic process.

Definition. The **filtration generated by $\boldsymbol{\{X_t\}_{t\in\mathcal{T}}}$** is $\{\sigma(\{X_s : s \leq t\})\}_{t\in\mathcal{T}}$. In other words $\mathcal{F}_t$ is the smallest σ-algebra which makes X_s measurable for all $s \leq t$ in $\mathcal{T}$. A stochastic process $\{X_t\}_{t\in\mathcal{T}}$ is called **adapted** to a filtration $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$ if X_t is $\mathcal{F}_t$-measurable for every $t \in \mathcal{T}$.

Clearly, the filtration generated by $\{X_t\}_{t\in\mathcal{T}}$ is the smallest filtration which makes $(X_t)_{t\in\mathcal{T}}$ adapted.

Definition. A filtration $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$ is said to satisfy the "usual condition" if

(i) it is right-continuous: $\lim_{s\to t^+} := \bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_t$, and
(ii) $\mathcal{F}_0$ contains all the $\mathbb{P}-$null events in $\mathcal{F}$.

2 Martingale Theory

2.1 Stopping Times

Consider a filtered space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathcal{T}}, \mathbb{P})$, i.e. a probability space equipped with a filtration.

Definition. A random time $N : \Omega \to \mathcal{T}$ is called a **stopping time** if $N^{-1}([0,t]) \in \mathcal{F}_t$ for each $t \in \mathcal{T}$.

The philosophy here is that we know that a stopping time happens when it happens.

Example. (i) Constants are trivial stopping times.
(ii) Last hit a constant before some fixed bound is not a stopping time

2.1 Proposition. *If T, S are stopping times,* $\min\{T,S\}$, $\max\{T,S\}$, *and $T+S$ are stopping times.*

Proof
- We have $\min\{T,S\}^{-1}([0,t]) = T^{-1}([0,t]) \cap S^{-1}([0,t]) \in \mathcal{F}_t$.
- Similarly, we have $\max\{T,S\}^{-1}([0,t]) = T^{-1}([0,t]) \cup S^{-1}([0,t]) \in \mathcal{F}_t$.
- For $T+S$, we have that

$$\begin{aligned}(T+S)^{-1}((t,\infty)) = &\Big(T^{-1}(\{0\}) \cap S^{-1}((t,\infty))\Big) \cup S^{-1}((t,\infty)) \\ &\cup \Big(T^{-1}((0,t]) \cap (T+S)^{-1}((t,\infty))\Big)\end{aligned}$$

where $T^{-1}(\{0\}) \cap S^{-1}((t,\infty)) \in \mathcal{F}_t$ and $S^{-1}((t,\infty)) \in \mathcal{F}_t$. To finish the proof, since $\mathcal{T}$ has a countable dense subset Q, we have

$$T^{-1}((0,t]) \cap (T+S)^{-1}((t,\infty)) = \bigcup_{r\in Q\cap(0,t)} S^{-1}((r,t]) \cap T^{-1}((t-r,\infty))$$

where the expressions in the union are certainly $\mathcal{F}_t$-measurable. ∎

Definition. The ***σ*–algebra generated by a stopping time *T*** is the set of all events A for which $A \cap T^{-1}([0,t]) \in \mathcal{F}_t$ for every $t \in \mathcal{T}$.

Intuitively, this is the information you collect until the stopping time. Note that this σ-algebra is not the same as the σ-algebra generated by the random variable T. It is left as an exercise to the reader to verify that the above collection is indeed a σ-algebra.

We write $X_{T\wedge t}$ is a random variable evaluated at time $T \wedge t$ (or T); in other words, $(X_{T\wedge t})(\omega) = X_{T\wedge t}(\omega)$. Then $\{X_{T\wedge t}\}_{t\in\mathcal{T}}$, or X^T, is a stochastic process stopped at time t.

2.2 Doob's Upcrossing Inequality

Definition. Consider a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathcal{T}}, \mathbb{P})$. A $\{\mathcal{F}_t\}_{t\in\mathcal{T}}$–adapted proces $\{X_t\}_{t\in\mathcal{T}}$ is said to be a **submartingale** if
(i) for all $t \in \mathcal{T}$, $X_t \in L^1(\Omega, \mathcal{F}, \mathcal{P})$, and
(ii) for all $s < t$ where $s, t \in \mathcal{T}$,

$$\mathbb{E}(X_t|\mathcal{F}_s) \geq X_s$$

almost surely.

and is said to be a **supermartingale** if condition (ii) above is replaced with (ii') for all $s < t$ where $s, t \in \mathcal{T}$,

$$\mathbb{E}(X_t | \mathcal{F}_s) \leq X_s$$

almost surely.

Then $\{X_t\}_{t \in \mathcal{T}}$ is a **martingale** if it is both a submartingale and supermartingale.

If X is a submartingale and $0 \leq s < t$ are fixed times, then $\mathbb{E}(X_0) \leq \mathbb{E}(X_s) \leq \mathbb{E}(X_t)$. Similar statements hold in the (super)martingale cases as well. One of the goals of martingale theory is to extend these results with respect to stopping times, rather than with respect to fixed times.

Definition. Let $X = \{X_n\}_{n \in \mathbb{Z}^+}$ be a discrete time process and fix levels $a < b$ with $a, b \in \mathbb{R}$. Then the **number of upcrossings** of $[a, b]$ by X before time N with respect to the event $\omega \in \Omega$, denoted by $U_N^X([a, b], \omega)$, is the maximal $k \in \mathbb{Z}^+$ such that there exists times $0 \leq s_1 < t_1 < s_2 < t_2 < \cdots < s_k < t_k \leq N$ such that for all i, $X_{s_i}(\omega) \leq a$ and $X_{t_i}(\omega) \geq b$.

Note that $U_N^X([a, b]) : \Omega \to \mathbb{Z}^+$ given by $\omega \mapsto U_N^X([a, b], \omega)$ is a random variable.

Definition. A process $C = \{C_n\}_{n \in \mathbb{Z}^+}$ is called **previsible** if C_n is $\mathcal{F}_{n-1}$–measurable for all $n \geq 1$. Suppose in addition that $\{X_n\}_{n \in \mathbb{Z}^+}$ is a discrete time process. Then the **martingale transform** of X by C, denoted by $C \cdot X$, is defined by

$$(C \cdot X)_n = \begin{cases} \sum_{k=1}^n C_k (X_k - X_{k-1}) & : n > 0 \\ 0 & : n = 0 \end{cases}$$

2.2 Lemma. *Suppose C is a bounded or L^2 previsible process. Then*

(i) Let C be non-negative and X a supermartingale. Then $C \cdot X$ is a supermartingale which is null at 0.

(ii) Let X be a martingale. Then $C \cdot X$ is a martingale which is null at 0.

Proof We first treat the case where C is bounded.

(i) We have by properties of conditional expectation and non-negativity of C

$$\begin{aligned} \mathbb{E}[(C \cdot X)_n - (C \cdot X)_{n-1} | \mathcal{F}_{n-1}] &= \mathbb{E}[C_n (X_n - X_{n-1}) | \mathcal{F}_{n-1}] \\ &= C_n \cdot \mathbb{E}[X_n - X_{n-1} | \mathcal{F}_{n-1}] \\ &\leq 0 \end{aligned}$$

where the second line follows since C is previsible. Integrability follows immediately since C is bounded and X_n is L^1 for each $n \in \mathbb{Z}^+$.

(ii) Consider $C + k$ where k is a constant and $k \geq |C_n(w)|$ for all n and w.

(iii) Similar, but the integrability is now guaranteed by Hölder's inequality. ■

We now have the following result:

2.3 Proposition. (Doob's Upcrossing Inequality) *(i) Let X be a supermartingale. Then*

$$(b - a) \cdot \mathbb{E}[U_N^X([a, b])] \leq \mathbb{E}[(X_N - a)^-].$$

(ii) Let X be a submartingale. Then

$$(b - a) \cdot \mathbb{E}[U_N^X([a, b])] \leq \mathbb{E}[(X_N - a)^+].$$

Proof We prove (i); the proof of (ii) is analgous. Define a process $\{C_n\}_{n\in\mathbb{Z}^+}$ by

$$\begin{aligned}
C_0 &:= 0\\
C_1 &:= \chi_{X_0^{-1}((-\infty,a))}\\
C_n &:= \chi_{C_{n-1}^{-1}(\{1\})}\cdot\chi_{X_{n-1}^{-1}((-\infty,b])}+\chi_{C_{n-1}^{-1}(\{0\})}\cdot\chi_{X_{n-1}^{-1}((-\infty,a))}.
\end{aligned}$$

Certainly C is non-negative and bounded; previsibility follows since C_n is a characteristic function depending only on preimages of X_k for $k<n$. Thus set $Y = C\cdot X$, i.e. $Y_n = \sum_{k=1}^n C_k\cdot(X_k - X_{k-1})$ almost surely; by Lemma 2.2, Y is a supermartingale.

Since each finished upcrossing increases the value of Y by at least $b-a$, we have for any $\omega\in\Omega$

$$Y_N(\omega)\geq (b-a)U_N^X([a,b],\omega)-(X_n-a)^-(\omega)$$

where $(X_N-a)^-(\omega)$ is the upper bound for the "loss" due to the last unfinished upcrossing. Since $\mathbb{E}(Y_1)\leq 0$, we have $\mathbb{E}(Y_N)\leq\mathbb{E}(Y_1)\leq 0$. Thus

$$(b-a)\cdot\mathbb{E}[U_N^X([a,b])]\leq\mathbb{E}[(X_n-a)^-]$$

as required. ■

2.3 Optional Sampling Theorems

2.4 Theorem. *Suppose $\{X_t\}_{t\in\mathcal{T}}$ is uniformly L^2. Then X_t converge in L^2 to a limit X_∞.*

Proof Note that we have the result for both discrete and continuous time martingales. By discretization, it is clear that it suffices to rove the result for discrete case.

We have the following orthogonality between the increments of a martingale $\{X_n\}_{n=0}^\infty$: if $n_1 < n_2 \leq n_2 < n_4$, then

$$\mathbb{E}[(X_{n_2}-X_{n_1})(X_{n_4}-X_{n_3})]=0.$$

This result follows by conditioning on $\mathcal{F}_{n_3}$ and applying the law of total expectation.

Now the proof proceeds. Set $Y_n := X_n - X_{n-1}$, so

$$\|X_n\|_2^2=\mathbb{E}(X_n^2)=\sum_{i=1}^n\|Y_i\|_2^2$$

and $\sum_{i=0}^n\|Y_n\|_2^2\leq B$ for all n, so that $\sum_{i=0}^\infty\|Y_n\|_2^2\leq B$. Thus $\{X_n\}_{n=0}^\infty$ is Cauchy in $L^2(\Omega,\mathcal{F},\mathbb{P})$. ■

2.5 Theorem. (Optional Sampling for Bounded Stopping Times in Discrete Times) *Let $\{X_n\}_{n=0}^\infty$ be a $(\Omega,\mathcal{F},\{\mathcal{F}_n\}_{n=0}^\infty,\mathbb{P})$ supermartingale and S,T be $\{\mathcal{F}_n\}$–stopping times such that $0\leq S\leq T\leq N$ for some constant $N<\infty$. Then X_T is integrable and $\mathbb{E}(X_T|\mathcal{F}_s)\leq X_s$ almost surely.*

Proof Notice that

$$|X_T|\leq|X_0|+|X_1|+\cdots+|X_N|$$

so that $\mathbb{E}(|X_T|)<\infty$. To prove that $\mathbb{E}(X_t|\mathcal{F}_s)\leq X_s$ a.s., it suffices to prove that $\mathbb{E}(X_T;A):=\int_A X_T\,d\mathcal{P}\leq\int_A X_s\,d\mathcal{P}=:\mathbb{E}(X_s;A)$ for all $A\in\mathcal{F}_s$. Assuming this, then

$$\mathbb{E}(\mathbb{E}(X_T|\mathcal{F}_s)-X_s;A)\leq 0$$

for all $A \in \mathcal{F}_s$, so we may take $A = A_0 := \{\mathbb{E}(X_T|\mathcal{F}_s) - X_s > 0\}$, so that $\mathbb{P}(A_0) = 0$.

Let's prove the required statement. First note that

$$\sum_{n=1}^{N} \chi_{\{S<n\leq T\}}(X_n - X_{n-1}) = X_T - X_S$$

and taking expectation over A on both sides

$$\begin{aligned}\mathbb{E}(X_T - X_S; A) &= \sum_{n=1}^{N} \mathbb{E}(\chi_{\{S<n\leq T\}}(X_n - X_{n-1}); A) \\ &= \sum_{n=1}^{N} \mathbb{E}(X_n - X_{n-1}; A \cap \{S < n \leq T\})\end{aligned}$$

But $A \cap \{S < n \leq T\} = A \cap \{S \leq n-1\} \cap \{n-1 < T\} \in \mathcal{F}_{n-1}$. Thus

$$\mathbb{E}(X_n - X_{n-1}; A \cap \{S < n \leq T\}) = \mathbb{E}(\mathbb{E}(X_n - X_{n-1}|\mathcal{F}_{n-1}); A \cap \{S < n \leq T\}) \leq 0$$

so the required statement holds. ∎

Definition. Let $\{X_n\}_{n=0}^{\infty}$ be a $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n=0}^{\infty}, \mathbb{P})$ a supermartingale. We say that $\{X_n\}_{n=0}^{\infty}$ is **closed** by a random variable X_∞ if X_∞ is $\mathcal{F}_\infty$–measurable and $X_n \geq \mathbb{E}(X_\infty|\mathcal{F}_n)$ almost surely for all $n = 0, 1, \ldots$.

Similar statements hold with $X_n \leq \mathbb{E}(X_\infty|\mathcal{F}_n)$ for a submartingale, or equality with a martingale.

2.6 Proposition. *Suppose that $\{X_n\}_{n=0}^{\infty}$ is a $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n=0}^{\infty}, \mathbb{P})$ a non-negative supermartingale, and $X_\infty = 0$. If $S, T : \Omega \to \overline{\mathbb{Z}^+}$ are $\{\mathcal{F}_n\}$–stopping times, $S \leq T$, then*

1. $\mathbb{E}(X_T) < \infty$
2. $\mathbb{E}(X_T|\mathcal{F}_s) \leq X_s$

Proof 1. Note that $X_T \leq \liminf_{n\to\infty} X_{T\wedge n}$ where $T \wedge n$ and 0 are two bounded stopping times. Thus $\mathbb{E}(X_{T\wedge n}) \leq \mathbb{E}(X_0)$ for all $n = 0, 1, \ldots$. Thus by Fatou's lemma

$$\begin{aligned}\mathbb{E}(X_T) &\leq \mathbb{E}(\liminf_{n\to\infty} X_{T\wedge n}) \leq \liminf_{n\to\infty} \mathbb{E}(X_{T\wedge n}) \\ &\leq \mathbb{E}(X_0) < \infty.\end{aligned}$$

2. Let $A \in \mathcal{F}_s$. For $n = 0, 1, \ldots,$,

$$\begin{aligned}\mathbb{E}(X_T; A \cap \{T \leq n\}) &= \mathbb{E}(X_{T\wedge n}; A \cap \{T \leq n\}) \\ &\leq \mathbb{E}(X_{T\wedge n}; A \cap \{S \leq n\}) \\ &\leq \mathbb{E}(X_{S\wedge n}; A \cap \{S \leq n\})\end{aligned}$$

Note that $S \wedge n$ and $T \wedge n$ are two bounded stopping times with $S \wedge n \leq T \wedge n$. Also, $A \cap \{S \leq n\} \in \mathcal{F}_{S\wedge n}$. Then apply the optional sampling theorem for bounded stopping times. By the monotone convergence theorem,

$$\lim_{n\to\infty} \mathbb{E}(X_T; A \cap \{T \leq n\})$$

and similarly for S. Thus

$$\mathbb{E}(X_T; A \cap \{T < \infty\}) \leq \mathbb{E}(X_S : A \cap \{S < \infty\})$$

so that

$$\mathbb{E}(X_T; A \cap \{T = \infty\}) = \mathbb{E}(X_S : A \cap \{S = \infty\}) = 0$$

and $\mathbb{E}(X_T; A) \leq \mathbb{E}(X_S; A)$. Since this holds for all $A \in \mathcal{F}_s$, $\mathbb{E}(X_T : \mathcal{F}_s) \leq X_s$. ■

2.7 Lemma. *Let $\{M_n\}_{n=0}^{\infty}$ be a martingale closed by M_∞. Let $T : \Omega \to \overline{\mathbb{Z}^+}$ be a stopping time. Then $M_T = \mathbb{E}(M_\infty | \mathcal{F}_T)$.*

Proof First assume $M_\infty \geq 0$. For any $A \in \mathcal{F}_T$,

$$\begin{aligned} E(M_T; A) &= \sum_{n=0}^{\infty} \mathbb{E}(M_n; A \cap \{T = n\}) \\ &= \sum_{n=0}^{\infty} \mathbb{E}(M_\infty, A \cap \{T = n\}) \\ &= \mathbb{E}(M_\infty; A). \end{aligned}$$

For the general case, decompose M_∞ into positive and negative parts. ■

2.8 Theorem. (Optional Sampling for Closed Supermartingales in Discrete Time) *Let $\{X_n\}_{n=0}^{\infty}$ be a $(\Omega, \mathcal{F}, \{\mathcal{F}_n\}_{n=0}^{\infty}, \mathbb{P})$ supermartingale closed by X_∞. Let $S, T : \Omega \to \overline{\mathbb{Z}^+}$ be two $\{\mathcal{F}_n\}_{n=0}^{\infty}$ stopping times with $S \leq T$. Then*

1. $\mathbb{E}(|X_T|) < \infty$
2. $\mathbb{E}(X_T | \mathcal{F}_s) \leq X_s$ *almost surely*

Proof

1. Define $M_n = \mathbb{E}(X_\infty | \mathcal{F}_n)$ and $A_n = X_n - M_n$ for $n = 0, 1, \ldots, \infty$. Since $\{X_n\}_{n=0}^{\infty}$ is a supermartingale closed by X_∞, $A_n \geq 0$, $A_\infty = 0$, and for $m \leq n$,

$$\begin{aligned} \mathbb{E}(A_n | \mathcal{F}_m) &= \mathbb{E}(X_n - \mathbb{E}(X_\infty | \mathcal{F}_n) | \mathcal{F}_m) \\ &\leq X_m - \mathbb{E}(X_\infty | \mathcal{F}_m) \\ &= A_m \end{aligned}$$

   Thus $\{A_n\}_{n=0}^{\infty}$ is a non-negative supermartingale with $A_\infty = 0$.
   One can prove that $\{M_n\}_{n=0}^{\infty}$ is a uniformly integrable martingale, i.e. $\lim_{c \to \infty} \sup_{n \in \mathbb{N}} \mathbb{E}(|X_n|; |X_n| > c) = 0$.
   By the previous proposition, $\mathbb{E}(A_T) < \infty$. On the other hand, $\mathbb{E}(|M_T|) < \infty$ by definition of the $\{M_n\}$. Thus $\mathbb{E}(|X_T|) < \infty$.
2. Apply the previous lemma so $\mathbb{E}(M_T | \mathcal{F}_s) = \mathbb{E}(\mathbb{E}(M_\infty | \mathcal{F}_T) \mathcal{F}_S) = \mathbb{E}(M_\infty | \mathcal{F}_s) = M_s$ by the optional sampling theorem for closed martingales. Meanwhile, as $\{A_n\}_{n=0}^{\infty}$ is a non-negative supermartingale with $A_\infty = 0$, $\mathbb{E}(A_T | \mathcal{F}_s) \leq A_s$ almost surely. Thus $\mathbb{E}(X_T | \mathcal{F}_s) \leq X_s$ almost surely. ■

To prepare for the optional sampling theorem for closed supermartingales in continuous time, we introduce one more discrete time notion:

Definition. Denote by $\mathbb{Z}^- = \{z \in \mathbb{Z} : z \leq 0\}$. A **negatively indexed supermartingale** $\{X_n\}_{n\in\mathbb{Z}^-}$ and $\{\mathcal{F}_n\}_{n\in\mathbb{Z}^-}$ where $\mathcal{F}_m \subseteq \mathcal{F}_n$ for $m \leq n$. Then $X_n \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, $\{X_n\}$ is adapted and $\mathbb{E}(X_n|\mathcal{F}_m) \leq X_m$.

2.9 Theorem. *Let $\{X_n\}_{n\in\mathbb{Z}^-}$ be a negatively indexed supermartingale such that $\sup_{n\in\mathbb{Z}^-} \mathbb{E}(X_n) < \infty$. Then $\{X_n\}_{n\in\mathbb{Z}^-}$ is uniformly L^1.*

Proof Fix $a = \sup_{n\in\mathbb{Z}^-} \mathbb{E}(X_n)$. For any $\epsilon > 0$, get $N(\epsilon) \in \mathbb{Z}^-$ such that $\mathbb{E}(X_{N(\epsilon)}) > a - \epsilon$ for $N \leq N(\epsilon)$. Since X is a supermartingale, $0 \leq \mathbb{E}(X_n) - \mathbb{E}(X_{N(\epsilon)}) \leq \epsilon$ for all $n \leq N(\epsilon)$. For $c > 0$, note that

$$|x|\chi_{|X|>c} = -x\chi_{X<-c} - X\chi_{X\leq x} + x.$$

Since $\{X_n\}$ is a supermartingale, for $n \leq N(\epsilon)$, $\mathbb{E}(X_{N(\epsilon)}|\mathcal{F}_n) \leq X_n$ and $\{X_n < -c\} \in \mathcal{F}_n$. Thus $\mathbb{E}(X_n; X_n < -c) \geq \mathbb{E}(X_{N(\epsilon)}; X_n < -c)$. Similarly, $\mathbb{E}(X_n; X_n \leq c) \geq \mathbb{E}(X_{N(\epsilon)}; X_n \leq c)$. Moreover, $\mathbb{E}(X_n) \leq \mathbb{E}(X_{N(\epsilon)}) + \epsilon$, so $\mathbb{E}(|X_n|; X_n \leq c) \leq \mathbb{E}(X_{N(\epsilon)}; X_n \leq c) + \epsilon$ for all $N \leq N(\epsilon)$. For this $N(\epsilon)$, there exists $\delta(\epsilon) > 0$ such that $\mathbb{E}(X_{N(\epsilon)}; A) < \epsilon$ for all $A \in \mathcal{F}$ and $\mathbb{P}(A) < \delta(\epsilon)$.

Since X is a supermartingale, $\{-2X_n^-\}_{n\in\mathbb{Z}^-}$ is also a supermartingale. To see this, define $f(x) = 2\min\{x, 0\}$, so f is increasing and concave. But then for $M \leq n$, by Jensen's inequality,

$$\mathbb{E}(f(X_n)|\mathcal{F}_m) \leq f(\mathbb{E}(X_n|\mathcal{F}_m)) \leq f(X_m)$$

and $-2X_n^- = f(X_n)$ is thus a supermartingale, so $\mathbb{E}(-2X_n^-) \geq \mathbb{E}(-2X_0^-)$. We also have $\mathbb{E}(X_n) \leq a$ where a is defined as above. Thus since $|X_n| = X_n + 2X_n^-$, we have

$$\mathbb{E}(|X_n|) \leq a - \mathbb{E}(-2X_0^-) = a + 2\,\mathbb{E}(X_0^-) < \infty.$$

Now, choose c so that $c \cdot \delta(\epsilon) > a + 2\,\mathbb{E}(X_0^-) \geq \mathbb{E}(|X_n|)$. Then by Markov's ineuality, $\mathbb{P}(|X_n| > c) \leq \delta(\epsilon)$. Thus $\mathbb{E}(|X_{N(\epsilon)}|; |X_n| > c) < \epsilon$ for $n \in \mathbb{Z}^-$ so that $\mathbb{E}(|X_m|; |X_n| > c) < 2\epsilon$ for all $n < N(\epsilon)$. But there are only finitely many terms with $n > N(\epsilon)$, so we are done. ■

2.10 Theorem. (Optional Sampling Theorem for Closed Supermartingales) *Let $\{X_t\}_{t\in\mathbb{R}^+}$ be an $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{R}^+}, \mathbb{P})$ supermartingale which is right continuous and closed. Let $S, T : \Omega \to \overline{\mathbb{R}^+}$ be two stopping times with $S \leq T$. Then*

(i) $\mathbb{E}(|X_T|) < \infty$

(ii) $\mathbb{E}(X_T|\mathcal{F}_s) \leq X_s$ *a.s.*

Proof Define random times $T_n := 2^{-n}(\lfloor 2^n T \rfloor + 1)$. Then the T_n are decreasing stopping times with $T = \lim T_n$ pointwise from above. For each n, $\{X_{m2^{-n}}\}_{n\in\mathbb{Z}^+}$ is a discrete-time closed supermartngale. Thus $\mathbb{E}(X_{T_n-1}|\mathcal{F}_{T_n}) \leq X_{T_n}$ a.s. (optional sampling theorem for closed supermartingale in discrete time). Set $Y_n = X_{T_n-1}$, so $\{Y_n\}_{n=-1,-2,\ldots}$ is a negatively indexed supermartingale and $\mathbb{E}(Y_n) = \mathbb{E}(X_{T-n-1}) \leq \mathbb{E}(X_0) < \infty$. Thus by the previous result, $\{X_{T_n}\}_{n\in\mathbb{Z}^+}$ is uniformly L^1. Since $T_n \to T$ from above and $X_{T_n} \to X_T$ a.s., uniform integrability gives that $X_{T_n} \to X_T$ in L^1. Thus $\mathbb{E}(|X_T|) = \lim_{n\to\infty} \mathbb{E}(|X_{T_n}|) < \infty$. Similarly, define $S_n = 2^{-n}(\lfloor 2^n S \rfloor + 1)$. Both S_n and T_n can be regarded as discrete-time stopping times. Thus by the discrete optional sampling theorem, $\mathbb{E}(X_{T_n}|\mathcal{F}_{S_n}) \leq X_{S_n}$ a.s.

Now for $A \in \mathcal{F}_S \subseteq \mathcal{F}_{S_n}$, $\mathbb{E}(X_{T_n}; A) \leq \mathbb{E}(X_{S_n}; A)$ so

$$|\,\mathbb{E}(X_{T_n}; A) - \mathbb{E}(X_T; A)|\; eq\, \mathbb{E}(|X_{T_n} - X_T|) \xrightarrow{n\to\infty} c$$

Also, $|\mathbb{E}(X_{S_n};A) - \mathbb{E}(X_S;A)| \to 0$. Therefore, as n goes to infinity,

$$\mathbb{E}(X_T;A) \le \mathbb{E}(X_S;A)$$

for all $A \in \mathcal{F}_S$ so $\mathbb{E}(X_T|\mathcal{F}_S) \le X_s$. ■

2.11 Corollary. *Let $\{X_t\}_{t\in\mathbb{R}^+}$ be a $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathbb{R}^+}$ martingale which is right continuous and closed. If $S, T : \Omega \to \overline{\mathbb{R}^+}$ are two stopping times with $S \le T$, then X_T is integrable and $|E(X_T|\mathcal{F}_S) = X_S$.*

2.12 Corollary. *Same results hold for bounded S, T and general right continuous martingale.*

2.13 Corollary. *A right continuous adapted process X is a martingale if and only if for every bounded stopping time T, X_T is L^1 and $\mathbb{E}(X_T) = \mathbb{E}(X_0)$.*

Proof The forward direction is contained in the above theorem. Conversely, for any $s < t$ and $A \in \mathcal{F}_s$, define $T = t\chi_{A^c} + s\chi_A$ which is a stopping time (exercise). Then $\mathbb{E}(X_0) = \mathbb{E}(X_T) = \mathbb{E}(X_t;A^c) + \mathbb{E}(X_s;A)$. On the other hand, $\mathbb{E}(X_0) = \mathbb{E}(X_t) = \mathbb{E}(X_t;A^c) + \mathbb{E}(X_t;A)$ so $\mathbb{E}(X_s;A) = \mathbb{E}(X_t;A)$. But $A \in \mathcal{F}_s$ is arbitrary, so $\mathbb{E}(X_t|\mathcal{F}_s) = \mathbb{E}(X_s)$. ■

2.14 Corollary. *Suppose $\{X_t\}_{t\in\mathcal{T}}$ is a martingale and T is a stopping time. Then the stopped process X^T where $X_t^T = X_{t\wedge T}$ for each $t \in \mathcal{T}$ is also a martingale.*

Proof The process X^T is clearly adapted. For any bounded stopping time S, $S \min T$ is also a bounded stopping time. Then

$$\mathbb{E}(X_s^T) = \mathbb{E}(X_{S\wedge T}) = \mathbb{E}(X_0) = \mathbb{E}(X_0^T)$$

so X^T is a martingale by the previous corollary. ■

2.4 Martingale Convergence

2.15 Theorem. *Let $\{X_t\}_{t\in\mathcal{T}}$ be a supermartingale. If $\sup_t \mathbb{E}(X_t^-) < \infty$, then $\lim_{t\to\infty} X_t$ exists almost surely.*

Proof Suppose $\lim_{t\to\infty} X_t$ does not exist a.s. Then there exist levels a and b such that X upcrosses $[a,b]$ infinitely many times with positive probability. In other words, $\mathbb{P}(\lim_{N\to\infty} U_N([a,b]) = \infty) > 0$, so $\mathbb{E}[U_N([a,b])] \to \infty$ as $N \to \infty$. However,

$$\mathbb{E}(U_N[a,b]) \le \frac{1}{b-a}\sup_t \mathbb{E}[(X_t - a)^-] \le \frac{1}{b-1}\sup_t(\mathbb{E}(X_t^-) + |a|) < \infty.$$ ■

2.16 Corollary. *A positive supermartingale converges a.s.*

2.17 Theorem. *Let $\{X_t\}_{t\in\mathcal{T}}$ be a martingale. The following conditions are equivalent:*
(i) The limit $\lim_{t\to\infty} X_t = X^*$ converges in L^1.
(ii) There is a random variable X_∞ in L^1 such that $X_t = \mathbb{E}(X_\infty|\mathcal{F}_t)$, i.e. X is closed by X_∞
(iii) X is uniformly L^1.

Moreover, in this case $X_\infty = X^*$.

PROOF First note that the upcrossing number $U_n^X([a,b],\omega)$ can be defined analgously for a continuous time process. Moreover, the upcrossing inequality sill holds in the continuous time case (right continuity + discretization).

$(ii \Rightarrow iii)$ Discussed earlier (find proof)

$(iii \Rightarrow i)$ Exercise: U.I. implies $\sup_t \mathbb{E}(X_t^-) < \infty$. Thus by the supermartingale convergence theorem, $\lim_{t\to\infty} X_t$ exists a.s. Then U.I. converges to X_∞ in L^1.

$(i \Rightarrow ii)$ $X + t - \mathbb{E}(X_{t+h}|\mathcal{F}_t)$ and take limit as $H \to \infty$ (??) ■

2.18 Corollary. *Let $\{X_t\}_{t\in\mathcal{T}}$ be a U.I. martingale. If $S, T : \Omega \to \overline{T}$ are two stopping times with $S \le T$. Then*

(i) $\mathbb{E}(|X_T|) < \infty$

(ii) $\mathbb{E}(X_T|\mathcal{F}_s) = X_s$

2.19 Corollary. *Let $\{X_t\}_{t\in\mathcal{T}}$ be a martingale and let T be a stopping time such that $|X_{t\wedge T}| \le 0$ for all $t \in \mathcal{T}$. Then $\mathbb{E}(X_T) = \mathbb{E}(X_0)$ where $X_T = \lim_{t\to\infty} X_{t\wedge T}$.*

PROOF The stopped process X^T is a bounded martingale and hence U.I. Thus $X_T = \lim_{t\to\infty} X_{t\wedge T}$ exists a.s. and $X_t^T \to X_T$ in L^1. Thus $\mathbb{E}(X_t^T) = \mathbb{E}(X_0)$ implies $E(X_T) = \mathbb{E}(X_0)$.■

2.5 BROWNIAN MOTION

Definition. An adapted process $B = \{B_t\}_{t\ge 0}$ is called a (standard, one-dimensional) **Brownian motion** if

1. $B_0 = 0$ a.s.
2. B has continuous sample paths
3. $B_t - B_s \perp \mathcal{F}_s$ (define independent), $B_t - B_s \sim \mathcal{N}(0, t-s)$.

Equivalently:

1' $\{B_t\}_{t\ge 0}$ is a Gaussian process

2' B has continuous paths

3' $\mathbb{E}(B_t) = 0$, $\mathbb{E}(B_sB_t) = s \wedge t$ for $s, t \ge 0$.

$(3 \Rightarrow 1')$ immediate

$(1 + 3 \Rightarrow 3')$ $\mathbb{E}(B_t) = 0$ for $t \ge 0$. Assume $s \le t$. Then $\mathbb{E}(B_sB_t) = \mathbb{E}(B_s^2) + \mathbb{E}(B_s(B_t - B_s))$, where $\mathbb{E}(B_s(B_t - B_s)) = 0$ by condition on $\mathcal{F}_s$, and $\mathbb{E}(B_s^2) \sim N(0, s)$.

$(3' \Rightarrow 1)$ $\mathbb{E}(B_0^2) = 0$ so $B_0 = 0$ a.s.

$(1' + 3' \Rightarrow 3)$ $(3')$ determines the variance-covariance structure for a Gaussian process, which coincides with the distribution given by (3).

2.20 Theorem. (Levy's Characterization) *Let X be a $\mathcal{F}_t$-adapted continuousprocess vanishing at 0. Then X is an $\mathcal{F}_t$-Brownian motion if and only if both X and $X_t^2 - t$ are martinales.*

PROOF "only if": Clearly X is a martingale. Then $X_t^2 - t$ is a martingale since

$$\begin{aligned}\mathbb{E}(X_t^2 - t|\mathcal{F}_s) &= \mathbb{E}[((X_t - X_s) + X_s)^2|\mathcal{F}_s] - t \\ &= \mathbb{E}[(X_t - X_s)^2|\mathcal{F}_s] + X_s^2 - t \\ &= t - s + X_s^2 = X + s^2 - s.\end{aligned}$$ ■

2.21 Proposition. *Brownian motion is a strong Markov process. Let T be a stopping time; then $B_T^* := \{B_{T+t} - B_t\}_{t\geq 0}$ is a Brownian motion, independent of $\mathcal{F}_T$.*

Proof Discretization and discussing all the possible values of T. ∎

Scaling property of BM: $\{B_{at}\}_{t\geq 0} = \{a^{1/2}B_t\}_{t\geq 0}$ in distribution.

2.22 Theorem. *Let $B = \{B_t\}_{t\geq 0}$ be a Brownian motion. Then the process $X_0 = 0$ and $X_t = tB(1/t)$ is also a Brownian motion.*

TODO: check.

2.23 Corollary. $\mathbb{P}(\inf\{t > 0 : B(t) = 0\} = 0) = 1$*, i.e. 0 is almost surely an accumulation point of zeros of $\{B_t\}_{t\geq 0}$.*

2.24 Proposition. (Reflection Principle) *Let $a > 0$, $\tau_a := \inf\{t : B(t) = a\}$, and $M_t := \sup_{s\in[0,t]} B_s$. Then $\mathbb{P}(M_t \geq a) = \mathbb{P}(\tau_a \leq t) = 2\,\mathbb{P}(B_t \geq a)$.*

Proof We have

$$\begin{aligned}\mathbb{P}(\tau_a \leq t) &= \mathbb{P}(\tau_a \leq t, B_t \geq a) + \mathbb{P}(\tau_a \leq t, B_t < a) \\ &= \mathbb{P}(\tau_a \leq t, B_t - B_{\tau_a} \geq 0) \\ &= \mathbb{E}(\mathbb{P}(B^*_{t-\tau_a} \geq 0 | \tau_a); \tau_a \leq t)\end{aligned}$$

Similarly,

$$\mathbb{P}(\tau_a \leq t, B_t < a) = \mathbb{E}(\mathbb{P}(B^*_{t-\tau_a} < 0 | \tau_a); \tau_a \leq t)$$

and by symmetry

$$\mathbb{P}(B^*_{t-\tau_a} \geq 0 | \tau_a) = \mathbb{P}(B^*_{t-\tau_a} < 0 | \tau_a) = \frac{1}{2}$$

so that

$$\mathbb{P}(\tau_a \leq t, B_t \geq a) = \mathbb{P}(\tau_a \leq t, B_t < a)$$ ∎

2.25 Theorem. *Almost surely the path of a Brownian motion is of unbounded variation on any compact interval.*

Proof Let $\{\pi_n\}_{n=1}^\infty$ be an increasing sequence of partitions of $[a,b]$ with part size converging uniformly to 0. Consider

$$\sum_{t_i\in\pi_n} (B_{t_{i+1}} - B_{t_i})^2 - (b-a) = \sum_{t_i\in\pi_n} \underbrace{[(B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i)]}_{\text{ind. rvs mean 0}}.$$

Moreover, note that

$$\frac{(B_{t_{i+1}})^2}{t_{i+1} - t_i} \sim Z^2$$

where $Z \sim N(0,1)$. Thus

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{t_i\in\pi_n}[(B_{t_{i+1}}-B_{t_i})^2-(t_{i+1}-t_i)]\right)^2\right] &= \sum_{t_i\in\pi_n}\mathbb{E}\left[\left((B_{t_{i+1}}-B_{t_i})^2-(t_{i+1}-t_i)^2\right)\right]\\ &= \sum_{t_i\in\pi_n}(t_{i+1}-t_i)^2\,\mathbb{E}((Z^2-1)^2)\\ &\leq \mathbb{E}((Z^2-1)^2)\sup_{t_i\in\pi_n}|t_{i+1}-t_i|\cdot(b-a)\end{aligned}$$

which converges to 0 as $n\to\infty$. Thus $\sum_{t_i\in\pi_n}(B_{t_{i+1}}-B_{t_i})^2\to b-a$ in L^2. Thus there exists a subsequence $(\pi_{n_k})_{k=1}^\infty$ such that $\sum_{t_i\in\pi_n}(B_{t_i+1}-B_{t_i})^2\to b-a$ almost surely. On the other hand,

$$\begin{aligned}\sum_{t_i\in\pi_{n_k}}(B_{t_i+1}-B_{t_i})^2 &\leq \sup_{t_i\in\pi_{n_k}}|B_{t_{i+1}}-B_{t_i}|\cdot\sum_{t_i\in\pi_{m_n}}|B_{t_{i+1}}-B_{t_i}|\\ &\leq \sup_{t_i\in\pi_{n_k}}|B_{t_{i+1}}-B_{t_i}|\cdot TV_{[a,b]}(B)\end{aligned}$$

Since each path B is uniformly continuous on $[a,b]$, since the part size of $\pi_{n_k}\to 0$, $\sup_{t_i\in\pi_{n_k}}|B_{t_{i+1}}-B_{t_i}|\to 0$. Then as $n\to\infty$ on both sides,

$$b-a\leq 0\cdot TV_{[a,b]}(B)$$

so that $TV_{[a,b]}(B)=\infty$. ∎

3 Stochastic Integration

Our goal is to define $\int f\,dB=\int f(t,w)dB_t(w)$. Since B does not have bounded variation, we cannot define this integral as a pathwise Lebesgue-Stieltjes integral. Let $I=[a,b]$ and define $V=V_I$ be the class of processes $f:[0,\infty)\times\Omega\to\mathbb{R}$ such that

(i) f is $\mathcal{B}\times\mathcal{F}$-measurable

(ii) f is $\mathcal{F}_t$-adapted.

(iii) $\mathbb{E}(\int_a^B f^2(t,\omega)dt)<\infty$, i.e. $f\in L^2([a,b]\times\Omega)$.

A process $\phi\in V$ is called **elementary** if ϕ has the form

$$\phi=\sum_j A_j\chi_{[t_j,t_{j+1})}(t)$$

where each A_j is a $\mathcal{F}_{t_j}$-measurable random variable for all j, and $\{t_j\}$ forms a partition of I. We then define

$$\int_I\phi\,dB=\sum_J A_j(B_{t_{j+1}}-B_{t_j}).$$

Next, we extend this definition to any process in V by approaching the process using elementtary process.

3.1 Lemma. *Let $g\in V$ be bounded and continuous. Then there exists $\{g_n\}_{n=1}^\infty$, $g_n\in V$ elementary, such that $g_n\to g$ in L^2.*

Proof Fix a partition $\pi_n = \{t_j\}$, we define

$$\phi_n = \sum g(t_j)\mathbf{1}[t_k, t_{j+1}).$$

Since g is $\mathcal{B} \otimes \mathcal{F}$-measurable and $\mathcal{F}_t$-adapted, so is ϕ_n. Moreover, since g is bounded, ϕ_n is also bounded, so $\phi_n \in V$ and ϕ_n is an elementary function. But $\int_a^b (g-\phi_n)^2 dt \to 0$ pointwise, so that $g \to \phi_n$ in L^2 by the dominated convergence theorem. ■

3.2 Lemma. *Let $h \in V$ be a bounded process. Then there exists $\{g_n\} \subseteq V$, g_n bounded and continuous, such that $g_n \to h$ in L^2.*

Proof We define mollifiers

$$\begin{aligned}
\rho(t) &= \begin{cases} c \cdot \exp\left(\frac{-1}{1-t^2}\right) & : |t| < 1 \\ 0 & : \text{otherwise} \end{cases} \\
\rho_\epsilon(t) &= \frac{1}{\epsilon}\rho(\epsilon^{-1} t)
\end{aligned}$$

Define $h(t) = 0$ for $t < a$ and let

$$\begin{aligned}
g_\epsilon(t) &= g_\epsilon * h(t) \\
&= \frac{1}{\epsilon}\int_{t-2\epsilon}^{t} \rho\left(\frac{t-s-\epsilon}{\epsilon}\right) h(s)\,\mathrm{d}s \\
&= \frac{1}{\epsilon}\int_{-\epsilon}^{\epsilon} \rho\left(\frac{z}{\epsilon}\right) h(t-z-\epsilon)\,\mathrm{d}z \\
&= \int_{-\epsilon}^{\epsilon} \rho_\epsilon(z) h(t-z-\epsilon)\,\mathrm{d}z
\end{aligned}$$

We first see that g_ϵ is adapted. By Cauchy-Schwarz, we have

$$\begin{aligned}
\int_a^b g_\epsilon^2(t)\,\mathrm{d}t &\le \int_a^b \left(\int_{-\epsilon}^{\epsilon} \rho_\epsilon(z)\,\mathrm{d}z \cdot \int_{-\epsilon}^{\epsilon} \rho_\epsilon(z) h^2(t-z-\epsilon)\,\mathrm{d}z\right)\mathrm{d}t \\
&\le \int_{-\epsilon}^{\epsilon} \rho_\epsilon(z)\left(\int_{a-1}^{b} h^2(t)\,\mathrm{d}t\right)\mathrm{d}z \\
&= \int_a^b h^2(t)\,\mathrm{d}t
\end{aligned}$$

and taking expectations, we have that $g_\epsilon \in L^2$.

It is clear that g_ϵ is bounded and continuous for any ϵ.

Next, we show convegence. Fix some $\omega \in \Omega$. Get $(u_n)_{n=1}^{\infty}$ such that $u_n(t) = 0$ and $u_n(t) \to h_n(t,\omega)$ in L^2. Since u_n is uniformly continuous, $\rho_\epsilon * u_n \to u_n$ uniformly on $[a,b]$. Thus

$$\begin{aligned}
\left\|\rho_\epsilon * h(\cdot,\omega) - h(\cdot,\omega)\right\|_2 &\le \left\|\rho_\epsilon * h(\cdot,\omega) - \rho_\epsilon * u_n(\cdot)\right\|_2 \\
&\quad + \left\|\rho_\epsilon * u_n - u_n\right\|_2 + \|u_n - h(\cdot,\omega)\|_2 \\
&\le 2\|u_n - h(\cdot,\omega)\|_2 + \left\|\rho_\epsilon * u_n - u_n\right\|_2
\end{aligned}$$

which both converge to 0 as $n \to \infty$. Since $g_\epsilon = \rho_\epsilon * h$ and h are uniformly bounded (by the bound of h), dominated convergence applies and we have $\mathbb{E}(\int_a^b (h - g_n)^2 \,\mathrm{d}t) \to 0$ as $n \to \infty$. ■

3.3 Lemma. *Let $f \in V$. Then there exists a sequence $\{h_n\} \subset V$ such that h_n is bounded for all n and $\mathbb{E}(\int_a^b |f - h_n|^2 \,\mathrm{d}t) \to 0$ as $n \to \infty$.*

Proof Take

$$h_n = \begin{cases} -n & : f < -n \\ f & : |f| \le n \\ & : f > n \end{cases}$$

and the result follows by the dominated convergence theorem. ■

Combing the proceding lemmas, we have:

For any process $f \in V$, there exists elementary processes $\phi_n \in V$ such that $\mathbb{E}(\int_a^b |f - \phi_n|^2 \,\mathrm{d}t) \to 0$ as $n \to \infty$. We want to define $\int_a^b f \,\mathrm{d}B$ as the limit of $\int_a^b \phi_n \,\mathrm{d}B$. To this end, we still need the existence and uniqueness of the limit.

3.4 Lemma. (Itô's Isometry for Elementary Processes) *Let ϕ be elementary. Then $\mathbb{E}[(\int_a^b \phi(t)\,\mathrm{d}b_t)^2] = \mathbb{E}(\int_a^b \phi^2(t)\,\mathrm{d}t)$.*

Proof Define $\Delta B_j = B_{j+1} - B_j$. Then

$$\mathbb{E}(A_i A_j \Delta B_i \Delta B_j) = \begin{cases} \mathbb{E}(A_j^2) \cdot (t_{j+1} - t_j) & : i = j \\ 0 & : i \neq j \end{cases}$$

Thus

$$\begin{aligned} \mathbb{E}[(\int_a^b \phi \,\mathrm{d}B)^2] &= \mathbb{E}[(\sum_j A_j \Delta B_j)^2] = \sum_{i,j} \mathbb{E}(A_i A_j \Delta B_i \Delta B_j) \\ &= \sum_j \mathbb{E}(A_j^2)(t_{j+1} - t_j) \\ &= \mathbb{E} \int_a^b \phi(t)^2 \,\mathrm{d}t \end{aligned}$$

■

By Lemma 4, since ϕ_n is a Cauchy sequence in $L^2(\mathbb{P} \times \lambda|_{[a,b]})$, $\int_a^b \phi_n(t)\,\mathrm{d}B_t$ is a Cauchy sequence in $L^2(\mathbb{P})$. Hence the limit exists. Moreover, if $\mathbb{E}(\int_a^b |f - \phi_n|^2 \,\mathrm{d}t) \to 0$ and $\mathbb{E}(\int_a^b |f - \phi_n'|^2 \,\mathrm{d}t) \to 0$, then $\mathbb{E}[(\int_a^b \phi_n(t)\,\mathrm{d}B_t) - \int_a^b \phi_n'(t)\,\mathrm{d}B_t)^2] \to 0$. The limit is unique, i.e. does not depend on the choice of ϕ_n. Thus we can define, for general $f \in V$,

$$\int_a^b f_t \,\mathrm{d}B_t := \lim_{n \to \infty} \int_a^b \phi_n(t)\,\mathrm{d}B_t$$

where $\{\phi_n\}$ is a sequence is an elementary process satisfying

$$\mathbb{E}(\int_a^b |f-\phi_n|^2 \,\mathrm{d}t) \to 0$$

in L^2.

3.5 Corollary. (Itô's Isometry) *Let $f \in V$, then*

$$\mathbb{E}[(\int_a^b f \,\mathrm{d}B)^2] = \mathbb{E}(\int_a^b f^2 \,\mathrm{d}t)$$

Proof This works for elementary processes; pass to the limit and use dominated convergence. ∎

3.1 Properties of the Stochastic Integral

The stochastic integral is linear.:

$$\int_a^b (cf+g)\,\mathrm{d}B_t = c\int_a^b f\,\mathrm{d}B_t + \int_a^b g\,\mathrm{d}B_t$$

In particular,

$$\int_a^b f\,\mathrm{d}B_t = \int_a^d f\,\mathrm{d}B_t + \int_d^b f\,\mathrm{d}B_t$$

for any $a < d < b$.

3.6 Proposition. *Let $\{M_n\}_{n=0}^N$ be a submartingale. Then for any $\lambda > 0$*

$$\lambda\,\mathbb{P}(\sup_n M_n \geq \lambda) \leq \mathbb{E}(|M_N|\chi_{\{\sup_n M_n \geq \lambda\}})$$

Proof Set $T :=\in \{n : M_n \geq \lambda\}$ if such an n exists, and N otherwise. Since T is a bounded stopping time, so is N. By the optional sampling theorem,

$$\begin{aligned}
\mathbb{E}(M_N) &\geq \mathbb{E}(M_T) \\
&= \mathbb{E}(M_T \mathbf{1}\sup_n M_n \geq \lambda) + \mathbb{E}(M_T \mathbf{1}\sup_n M_n < \lambda) \\
&\geq \lambda \cdot \mathbb{P}(\sup_n M_n \geq \lambda) + \mathbb{E}(M_N \mathbf{1}\sup_n M_n < \lambda)
\end{aligned}$$

so that

$$\begin{aligned}
\lambda\,\mathbb{P}(\sup_n M_n \geq \lambda) &\leq \mathbb{E}(M_n \mathbf{1}\sup_n M_n \geq \lambda) \\
&\leq \mathbb{E}(|M_N| \mathbf{1}\sup_n M_n \geq \lambda)
\end{aligned}$$

as required. ∎

3.7 Corollary. *Let $M = \{M_n\}_{n=0}^N$ be a martingale or a positive supermartingale. Then for any $P \geq 1$ and $\lambda > 0$,*

$$\lambda^= \mathbb{P}(\sup_N |M_n| \geq \lambda) \leq \mathbb{E}(|M_N|^p)$$

Proof By Jensen's inequality, $|M_n|^p$ is a submartingale. Apply the above proposition to $|M_n|^p$ and λ^p so that

$$\lambda^p\,\mathbb{P}(\sup_n |M_n|^p \geq \lambda^p) \leq \mathbb{E}(|M_N|^p \mathbf{1}\sup_n |M_n|^p \geq \lambda^p) \leq \mathbb{E}(|M_N|^p) \qquad \blacksquare$$

3.8 Theorem. (Doob's Martingale Inequality) *Let M_t be a continuous martingale. Then for $p \geq 1$, $T \geq 0$, and $\lambda > 0$,*

$$\mathbb{P}(\sup_{0\leq t\leq T} |M_t| \geq \lambda) \leq \frac{1}{\lambda^p}\,\mathbb{E}(|M_T|^p).$$

Proof Let D be a dense coutable subset of $[0,T]$ and D_n an increasing sequence of finite subsets of D such that $\bigcup_{n=1}^{\infty} D_n = D$. On each D_n, we can apply the discrete time result

$$\lambda^p\,\mathbb{P}(\sup_{t\in D_n} |M_t| \geq \lambda) \leq \mathbb{E}(|M_{D_n^*}|^p)$$

for $p \geq 1$, where $D_n^* = \max(D_n)$. Then as $n \to \infty$, for any $\epsilon > 0$,

$$\begin{aligned}
\mathbb{P}(\sup_{t\in D} |M_t| \geq \lambda) &\leq \mathbb{P}(\sup_{t\in D} |M_t|) > \lambda - \epsilon \\
&= \lim_{n\to\infty} \mathbb{P}(\sup_{t\in D_n} |M_t| > \lambda - \epsilon) \\
&\leq \liminf_{n\to\infty} \mathbb{P}(\sup_{t\in D_n} |M_t| \geq \lambda - \epsilon) \\
&= \liminf_{n\to\infty} \frac{1}{(\lambda-\epsilon)^p}\,\mathbb{E}(|M_{D_n^*}|^p).
\end{aligned}$$

Then as $\epsilon \to 0$,

$$\begin{aligned}
\mathbb{P}(\sup_{t\in D} |M_t| \geq \lambda) &\leq \lambda^{-p} \liminf_{n\to\infty} \mathbb{E}(|M_{D_n^*}|^p) \\
&\leq \lambda^{-p}\,\mathbb{E}(|M_T|^p).
\end{aligned} \qquad \blacksquare$$

3.9 Theorem. *Let $f \in V(0,T)$ for all $T \in \mathbb{R}^+$. Then $M_t = \int_0^t f_s \,\mathrm{d}B_s$ has a continuous version, which is a $\mathcal{F}_t$-martingale. Moreover,*

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

for $\lambda, T > 0$.

Proof Let $\{\phi_n\}$ be a sequence of elementary processes such that $\mathbb{E}(\int_0^t (f-\phi_n)^2 \,\mathrm{d}t) \to 0$ as $n \to \infty$. Define processes

$$I_n(t) = \int_0^t \phi_n(s)\,\mathrm{d}B_s.$$

Since $\phi_n = \sum_j A_j^{(n)} \mathbf{1}[t_j, t_{j+1})(t)$, we have

$$I_n(t) = \sum_{j\leq k-1} A_j^{(n)}(B_{t_{j+1}} - B_{t_j}) + A_k^{(n)}(B_t - B_{t_k})$$

for $t \in [t_k, t_{k+1}]$. It is easy to check that I_n is a continuous martingale for any n. Thus $I_n - I_m$ is also a continuous martingale for any m, n. By Doob's martingale inequality and Itô's isometry

$$\begin{aligned}\mathbb{P}(\sup_{0\le t\le T} |I_n(t) - I_m(t)| > \epsilon) &\le \frac{1}{\epsilon^2}\,\mathbb{E}(|I_n(T) - I_m(T)|^2) \\ &= \frac{1}{\epsilon^2}\,\mathbb{E}(\int_0^T (\phi_n - \phi_m)^2\,\mathrm{d}s)\end{aligned}$$

which converges to 0 as $m, n \to \infty$. Thus there exists a subsequence $(n_k)_{k=1}^\infty$ such that

$$\mathbb{P}(\sup_{0\le t\le T} |I_{n_{k+1}} - I_{n_k}| > 2^{-k}) < 2^{-k}.$$

Thus by Borel-Cantelli,

$$\mathbb{P}(\sup_{0\le t\le T} |I_{n_{k+1}}(t) - I_{n_k}(t)| > 2^{-k} \text{ infinitely often}) = 0.$$

Therefore, almost surely, there exists $k_1 = k_1(\omega)$ such that

$$\sup_{0\le t\le T} |I_{n_{k+1}}(t) - I_{n_k}(t)| \le 2^{-k}$$

for all $k \ge k_1$. Thus $I_{n_k}(t)$ converges uniformly for $t \in [0, T]$ to some $I(t)$. Then $I(t)$ is continuous. On the other hand, $I_{n_k}(t) \to M(t)$ in $L^2(\mathbb{P})$ for any t, hence $M(t) = I(t)$ almost surely for any $t \in [0, T]$. Thus $I(t)$ is a continuous version of $M(t)$.

Next, we always mean this continuous version when writing $M(t)$. $M(t)$ is certainly adapted (since it is the L^2 limit of $I_n(t)$). Moreover, for $t \ge 0$, since $I_n(t) \to M(t)$ in L^2, we also have

$$\mathbb{E}(I_n(t)|\mathcal{G}) \to \mathbb{E}(M(t)|\mathcal{G})$$

for any sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$. Thus

$$\begin{aligned}\mathbb{E}(M(t)|\mathcal{F}_s) &= \lim_{n\to\infty} \mathbb{E}(I_n(t)|\mathcal{F}_s) = \lim_{n\to\infty} I_n(s) \\ &= M(s)\end{aligned}$$

so that $M(t)$ is a $(\mathcal{F}_t)_{t\in\mathcal{I}}$-martingale. Then by Doob's martingale inequality,

$$\mathbb{P}(\sup_{0\le t\le T} |M_t| \ge \lambda) \le \frac{1}{\lambda^2}\,\mathbb{E}(M_T^2) = \frac{1}{\lambda^2}\,\mathbb{E}(\int_0^T f^2(s)\,\mathrm{d}s)$$ ■

3.2 Extensions of Itô's Integral

Stochastic integral with stopping time

3.10 Lemma. *Let $f \in V(0, T)$ and T a stopping time. Set $m(t) = \int_0^T f_s\,\mathrm{d}B_s$. Then $M(T) = \int_0^T f(s)\,\mathrm{d}B_s = \int_0^T f(s)\mathbf{1}s < T\,\mathrm{d}B_s$.*

Proof The proof clearly holds when f is an elementary process and τ is simple. For general f and τ, let f_n be an elementary process such that

$$\mathbb{E}(\int_0^T (f_n - f)^2 \, \mathrm{d}t) \to 0$$

as $n \to \infty$, and τ_n be simple stopping times such that $\tau_n \to \tau^+$ everywhere as $n \to \infty$. We have

$$\int_0^{\tau_n} f_n(t) \, \mathrm{d}B_t = \int_0^T f_n(t) \mathbf{1} t < \tau_n \, \mathrm{d}B_t. \tag{3.1}$$

Since $\mathbf{1} t < \tau_n \to \mathbf{1} t < \tau$ for any ω and $t \neq \tau$. Pathwisely, by the dominated convergence theorem,

$$\int_0^T |\mathbf{1} t < \tau_n - \mathbf{1} t < \tau|^2 f^2(t) \, \mathrm{d}t \to 0$$

almost surely as $n \to \infty$. We also have

$$\mathbb{E}\Big(\int_0^T \mathbf{1} t < \tau_n |f(t) - f_n(t)|^2 \, \mathrm{d}t\Big) \le \mathbb{E}\Big(\int_0^T |f(t) - f_n(t)|^2 \, \mathrm{d}t) \to 0$$

as $n \to \infty$. Thus

$$\begin{aligned}\Big[\mathbb{E}\Big(\int_0^T (f_n(t)\mathbf{1} t < \tau_n - f(t)\mathbf{1} t < \tau)^2 \, \mathrm{d}t\Big)\Big]^{1/2} &= \|f_n(t)\mathbf{1} t < \tau_n - f(t)\mathbf{1} t < \tau\|_{L^2(\mathbb{P}\times\lambda)} \\ &\le \|f_n(t)\mathbf{1} t < \tau_n - f(t)\mathbf{1} t < \tau_n\| + \|f(t)\mathbf{1} t < \tau_n - f(t)\mathbf{1} t < \tau\| \to 0\end{aligned}$$

so that

$$\int_0^T f_n(t)\mathbf{1} t < \tau_n \, \mathrm{d}B_t \to \int_0^T f(t)\mathbf{1} t < \tau \, \mathrm{d}B_t \tag{3.2}$$

in L^2. On the other hand, similarly as in the proof of the previous result,

$$\sup_{0 \le t \le T} |\int_0^t f_n(s) \, \mathrm{d}B_s - \int_0^t f(s) \, \mathrm{d}B_s| \to 0$$

in probability, so Doob's martingale inequality implies that

$$\int_0^{\tau_n} |f_n(s) - f(s)| \, \mathrm{d}B_s \to 0$$

in probability. Thus

$$\int_0^{\tau_n} f_n(s) \, \mathrm{d}B_s - \int_0^{\tau_n} f(s) \, \mathrm{d}B_s \to 0$$

as $n \to \infty$ in probability. By continuity of the stochastic integral, we also have

$$\int_0^{\tau_n} f(s) \, \mathrm{d}B_s - \int_0^{\tau} f(s) \, \mathrm{d}B_s \to 0$$

almost surely as $n \to \infty$. Thus

$$\int_0^{\tau_n} f_n(s) \, \mathrm{d}B_s \to \int_0^{\tau} f(s) \, \mathrm{d}B_s. \tag{3.3}$$

Combining (3.1), (3.2), and (3.3) completes the proof. ■

3.3 Multidimensional Itô's Integral

3.4 Variant's on Itô's Integral

- Multidimensional Integration Let $B = (B^1, \ldots, B^n)$ be a n-dimensional Brownian motion. Define $V_{\mathcal{H}}^{m\times n}(a,b)$ to be the set of $m \times n$ matrices of processes $V = (V_{ij}(t,\omega))$ such that each V_{ij} satisfies
  - V_{ij} is $\mathcal{B} \otimes \mathcal{H}$-measurable
  - V_{ij} is $\mathcal{H}_T$-adapted
  - $\mathbb{E}(\int_a^b V_{ij}^2(t)\,\mathrm{d}t) < \infty$.

  Then for $v \in V_{\mathcal{H}}^{m\times n}(a,b)$, define
  $$\int_a^b v\,\mathrm{d}B = \int_a^b \begin{pmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & & \vdots \\ v_{m1} & \cdots & v_{mn} \end{pmatrix} \cdot \begin{pmatrix} dB^1 \\ \vdots \\ db^n \end{pmatrix}.$$
  That is, $\int_a^b v\,\mathrm{d}B$ is a $m \times 1$ vector, and its ith component is given by $\sum_{j=1}^k \int_a^b v_{ij}\,\mathrm{d}B^j$.
- Weaken the integrability condition $\mathbb{E}(\int_a^b f^2(t)\,\mathrm{d}t) < \infty$. We can weaken this condition to $\mathbb{P}(\int_a^b f^2(t)\,\mathrm{d}t < \infty) = 1$ in two ways.
  1. Return to the construction of Itô's integral. Replace convergence in L^2 by convergence in probability when the former no longer goes through
  2. Define $\tau_n := \inf\{t : \int_a^t f^2(s,w)\,\mathrm{d}s \geq n\} \wedge b$. Then τ_n is an increasing sequence of stopping times and $\tau_n \to b$ almost surely. Define $f_n = f\mathbf{1}t < \tau_n$ and defne $\int_a^b f\,\mathrm{d}B$ as $\lim_{n\to\infty} \int_a^b f_n\,\mathrm{d}B$. The limit exists since $\tau_n \to b$ almost surely, and $\int_a^t f_m\,\mathrm{d}B$ and $\int_a^t f_n\,\mathrm{d}B$ agree on $\{t \leq \tau_n\}$ for $m > n$. (Exercise: see "stochastic integral with stopping time")
     Let $W_{\mathcal{H}}(a,b)$ be the class of pricesses satisfying the measurability condition, the adaptedness condition, and $\mathbb{P}(\int_a^b f^2(t)\,\mathrm{d}t < \infty) = 1$. In the matrix case, we write $W_{\mathcal{H}}^{m\times n}(a,b)$.

Remark. We have constructed $\int f\,\mathrm{d}B$ for $f \in W_{\mathcal{H}}(a,b)$. Most of the properties of stochastic integrals with $f \in V$ still hold. The only main exception is that $M_t = \int_0^t f(s)\,\mathrm{d}B_s$ is no longer necessarily a martingale since the proof requires convergence in L^2. However, as we have seen, by stopping, we know there exists an increasing sequence of stopping times $\{\tau_n\} \to \infty$ almost surely, and $M^{\tau_n} = \{M_{\tau_n \wedge t}\}_{t\geq 0}$ is a martingale for each n.

Definition. Let $\{X_t\}_{t\geq 0}$ be a stochastic processes. If there exists an increasing sequence of stopping times $\{\tau_n\}$ of $\{\mathcal{F}_t\}_{t\geq 0}$ such that $\tau_n \to \infty$ almost surely and $X^{\tau_n} = \{X_{t\wedge\tau_n}\}_{t\geq 0}$ is a martingale for each n, then $\{X_t\}_{t\geq 0}$ is called a **local martingale**.

3.5 Itô processes and Itô's formula for Brownian motion

Definition. Let $\{B_t\}_{t\geq 0}$ be a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. A Itô processes is a process $\{X_t\}_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ of the form
$$X_t = X_0 + \int_0^t u(s,\omega)\,\mathrm{d}s + \int_0^t v(s,\omega)\,\mathrm{d}B_s$$

where $v \in W_{\mathcal{H}} = \bigcap_{T>0} W_{\mathcal{H}}(0,T)$, u is $\mathcal{H}_t$-adapted, and

$$\int_0^T |u(s,\omega)|\,\mathrm{d}s < \infty$$

for all $t > 0$ almost surely. We write $dX_t = udt = vdB_t$ to denote the "dynamics of X"; this is notation for the expression above. Define the stochastic integral with respect to X:

$$\int_0^t f(s)\,\mathrm{d}X_s = \int_0^t f(s)u(s)\,\mathrm{d}s + \int_0^t f(s)v(s)\,\mathrm{d}B_s$$

whenever all the integrals are well-defined.

Recall that we write

$$dX_t = \underbrace{u\,\mathrm{d}t}_{\text{drift term}} + \underbrace{v\,\mathrm{d}B_t}_{\text{diffusion term}}$$

to denote the process

$$X_t = X_0 + \int_0^t u(s,\omega)\,\mathrm{d}s + \int_0^t v(s,\omega)\,\mathrm{d}B_s.$$

We then define $\int_0^t f(s,\omega)\,\mathrm{d}X_s = \int_0^t f(s)u(s)\,\mathrm{d}s + \int_0^t f(s)v(s)\,\mathrm{d}B_s$ whenever all the integrals are well-defined.

> ***3.11 Theorem. (1-dimensional Itô's Formula)*** *Let X be an Itô process given by $\mathrm{d}X_t = u\,\mathrm{d}t + v\,\mathrm{d}B_t$. Let $g(t,x) \in C^{1,2}([0,\infty)\times\mathbb{R})$, i.e. C^1 in t and C^2 in x. Then $Y = \{Y_t = g(t,X_t)\}_{t\geq 0}$ is an Itô process and*
>
> $$Y_t = Y_0 + \int_0^t \frac{\partial g}{\partial x}(s,X_s)v_s\,\mathrm{d}B_s + \int_0^t \left(\frac{\partial g}{\partial t}(s,X_s) + \frac{\partial g}{\partial x}(s,X_s)u_s + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(s,X_s)v_s^2\right)\mathrm{d}s.$$
>
> Proof Later ... ■

Remark. In differential notation, we have

$$\begin{aligned} dY_t &= \frac{\partial g}{\partial x}(t,X_t)v_t\,\mathrm{d}B_t + \left(\frac{\partial g}{\partial t}(t,X_t) + \frac{\partial g}{\partial x}(t,X_t)u_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)v_t^2\right)\mathrm{d}t \\ &= \frac{\partial g}{\partial t}(t,X_t)\,\mathrm{d}t + \frac{\partial g}{\partial x}(t,X_t)\,\mathrm{d}X_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)(\mathrm{d}X_t)^2 \end{aligned}$$

where $(dX_t)^2$ is calculated according to

$$\begin{aligned} (\mathrm{d}t)^2 &= \mathrm{d}t\,\mathrm{d}B_t = \mathrm{d}B_t\,\mathrm{d}t = 0 \\ (\mathrm{d}B_t)^2 &= \mathrm{d}t \end{aligned}$$

Say $Y_t = g(B_t)$ where $g \in C^2$. Then $\mathrm{d}Y_t = g'(B_t)\,\mathrm{d}B_t + \frac{1}{2}g''(B_t)\,\mathrm{d}t$.

Example. We will construct Geometric Brownian Motion. Let r and σ be constants, and let

$$X_t = X_0 \exp((r - \sigma^2/2)t + \sigma B_t).$$

Apply Itô's formula to get

$$\begin{aligned}\mathrm{d}\Big(\frac{X_t}{X_0}\Big) &= \sigma\exp(\ldots)\,\mathrm{d}B_t + \Big(r - \frac{1}{2}\sigma^2\Big)\exp(\ldots)\,\mathrm{d}t + \frac{1}{2}\sigma^2\exp(\ldots)\,\mathrm{d}t \\ &= \sigma\exp(\ldots)\,\mathrm{d}B_t + r\exp(\ldots)\,\mathrm{d}t\end{aligned}$$

so that

$$\frac{X_t}{X_0} = 1 + \int_0^t \sigma\exp(\ldots)\,\mathrm{d}B_s + \int_0^t r\exp(\ldots)\,\mathrm{d}s$$

and

$$X_t = X_0 + \int_0^t \sigma X_0\exp(\ldots)\,\mathrm{d}B_s + \int_0^t rX_0\exp(\ldots)\,\mathrm{d}s$$

which, in differential notation, is

$$\mathrm{d}X_T = \sigma X_t\,\mathrm{d}B_t + rX_t\,\mathrm{d}t$$

3.12 Theorem. (Integration by Parts) *Let X, Y be Itô proceses with*

$$\mathrm{d}X_t = u_1\,\mathrm{d}t + v_1\,\mathrm{d}B_t, \qquad \mathrm{d}Y_t = u_2\,\mathrm{d}t + v_2\,\mathrm{d}B_t.$$

Then

$$\begin{aligned}\mathrm{d}(XY)_t &= X_t\,\mathrm{d}Y_t + Y_t\,\mathrm{d}X_t + \mathrm{d}X_t\,\mathrm{d}Y_t \\ &= X_t\,\mathrm{d}Y_t + Y_t\,\mathrm{d}X_t + v_1v_2\,\mathrm{d}t.\end{aligned}$$

In other words,

$$\int_0^t X_s\,\mathrm{d}Y_s = X_tY_t - X_0Y_0 - \int_0^t Y_s\,\mathrm{d}X_s - \underbrace{\int_0^t v_1v_2\,\mathrm{d}t}_{=\int_0^t \mathrm{d}X_s\cdot\mathrm{d}Y_s}.$$

Proof By Itô's formula, we have

$$\begin{aligned}\mathrm{d}(X+Y)_t^2 &= 2(X+Y)_t\,\mathrm{d}(X+Y)_t + (v_1+v_2)^2\,\mathrm{d}t \\ &= 2X_t\,\mathrm{d}X_t + 2Y_t\,\mathrm{d}Y_t + 2X_t\,\mathrm{d}Y_t + 2Y_t\,\mathrm{d}X_t + (v_1+v_2)^2\,\mathrm{d}t\end{aligned}$$

and similarly

$$\mathrm{d}(X-Y)_t^2 = 2X_t\,\mathrm{d}X_t + 2Y_t\,\mathrm{d}Y_t - 2X_t\,\mathrm{d}Y_t - 2Y_t\,\mathrm{d}X_t + (v_1-v_2)^2\,\mathrm{d}t$$

so that

$$\begin{aligned}\mathrm{d}(XY)_t &= \frac{1}{4}(\mathrm{d}(X+Y)_t^2 - \mathrm{d}(X-Y)_t^2) \\ &= X_t\,\mathrm{d}Y_t + Y_t\,\mathrm{d}X_t + v_1v_2\,\mathrm{d}t\end{aligned}$$ ■

End of Midterm.

4 Continuous Semimartingales

Recall that for Brownian motion, we have $\sum_{t_i\in\pi_n}(B_{t_{i+1}}-B_{t_i})^2 \to b-a$ in L^2 if π_n is a partition of $[a,b]$ and the mesh goes to 0. In particular, take $a=0$ and t, then $\sum_{t_i\in\pi_n}(B_{t_{i+1}-t_i})^2 \to 0$. Meanwhile, we also know that B_t^2-t is a martingale. This is not a coincidense: intuitively, $(\mathrm{d}B)^2=\mathrm{d}t$ so that

$$\sum_{t_i}(B_{t_i+1}-B_{t_i})^2 = \int_0^t (\mathrm{d}B_s)^2\,\mathrm{d}t = t$$

In addition, $\mathrm{d}(B_t^2) = 2B_t\,\mathrm{d}B_t + (\mathrm{d}B_t)^2$ so B_t^2-t has no $\mathrm{d}t$ and is a local martingale (in fact, a martingale).

We can extend this to any continuous, L^2-bounded martingale.

4.1 Theorem. *Let M be a L^2-bounded continuous martingale, $M_0=0$ a.s. Then there exists a unique continuous increasing process $[M]$ vanishing at 0 such that $M^2-[M]$ is a uniformly integrable martingale.*

Proof First, we further assume that the martingale M is bounded. Let $\Delta=\{0=t_0<t_1<\cdots\}$ of $\mathbb{R}^+$ such that the nmber of parts in $[0,t]\cap\Delta$ is finite for any t. For a process X, define

$$T_t^\Delta(x) := \sum_{i=0}^{k-1}(X_{t_i+1}-X_{t_i})^2 + (X_t-X_{t_k})^2$$

for $r_n\le t<t_{k+1}$. We are going to show that the quadratic variation processes is indeed the limit of $T_t^\Delta(x)$ as the mesh of Δ goes to 0.

1. Show that the limit of $T_t^\Delta(M)$ exists as mesh Δ goes to 0. First note that

$$\mathbb{E}(T_t^\Delta(M)-T_s^\Delta(M)|\mathcal{F}_s) = \mathbb{E}[(M_t-M_{t_j})^2+(M_{t_j}-M_{t_{j-1}})^2+\cdots+(M_{t_{i+1}}-M_{t_i})^2]-(M_s-M_{t_i})^2|\mathcal{F}_s] \tag{4.1}$$

$$= \mathbb{E}((M_t-M_{t_j})^2+\cdots+(M_{t_{i+1}}-M_s)^2|\mathcal{F}_s) = \mathbb{E}(M_t^2-M_s^2|\mathcal{F}_s) \tag{4.2}$$

and this is independent of Δ. This implies that $M_t^2-M_t^\Delta(M)$ is a continuous martingale.

For $a>0$, let Δ and Δ' be two partitions of $\mathbb{R}^+$. Denote $\Delta\Delta'=\Delta\cup\Delta'$, the partition consisting of all the points of Δ and Δ', and set

$$X := T^\Delta(M)-T^{\Delta'}(M) = (M^2-T^\Delta(M))-(M^2-T^{\Delta'}(M))$$

is a martingale. If we can show that $\mathbb{E}(X_a^2)\to 0$ as $mesh(\Delta)\to 0$ and $\mathrm{mesh}(\Delta')\to 0$, then $T_a^\Delta(M)$ is a Cauchy sequence in L^2 as $\mathrm{mesh}(\Delta)\to 0$, and thus has a limit in L^2. Applying (4.1) to X, we have $\mathbb{E}(X_a^2)=\mathbb{E}(T_a^{\Delta\Delta'}(X))$. To show that $\mathbb{E}(T_a^{\Delta\Delta'}(X))\to 0$, note that

$$T_a^{\Delta\Delta'}(X) = \sum_{i=1}^{k-1}(X_{t_{i+1}}-X_{t_i})^2+(X_1-X_{t_k})^2$$

so that

$$X = T^\Delta(M)+(-T^{\Delta'}(M)) \le 2(T_a^{\Delta\Delta'}(T^\Delta(M))+T_a^{\Delta\Delta'}(T^{\Delta'}(M)).$$

Thus, it suffices to show that $\mathbb{E}[T_a^{\Delta\Delta'}(T^{\Delta}(M))] \to 0$ as $\mathrm{mesh}(Delta) + \mathrm{mesh}(Delta') \to 0$.

First note that

$$T^{\Delta}_{s_{k+1}}(M_T)^{\Delta}_{s_k}(M) = (M_{s_{k+1}} - M_{t_1})^2 - (M_{s_k} - M_{t_1})^2 = (M_{s_{k+1}} - M_{s_k}) \cdot (M_{s_{k+1}} + M_{s_k} - 2M_{t_i})$$

so that

$$\begin{aligned}\mathbb{E}[T_a^{\Delta\Delta'}(T^{\Delta}(M))] &= \mathbb{E}(\sum (M_{s_{k+1}} - M_{s_k})^2 \cdot (M_{s_{k+1}} + M_{s_k} - 2M_{t_1})^2) \\ &\leq \mathbb{E}[(\sup_k |M_{s_{k+1}} + M_{s_k} - 2M_{t_1}|^2) \cdot \sum (M_{s_{k+1}} - M_{s_k})^2)] \\ &= \textit{...missed}\end{aligned}$$

Note that $\sup_k |M_{s_{k+1}} - M_{s_k} - 2M_{t_1}|^4 \to 0$ a.s. since M is uniformly continuous over $[0,a]$. By dominated convergence, we have $\mathbb{E}((\sup_k |M_{s_{k+1}} + M_{s_k} - 2M_{t_1}|^4))^{1/2} \to 0$. Therefore, it suffices to show that $\mathbb{E}(T_a^{\Delta\Delta'}(M))^2$ is bounded by a constant independent of Δ and Δ'. Indeed, we have

$$\begin{aligned}(T_a^{\Delta}(M))^2 &= \left(\sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})^2 \right)^2 \\ &= \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}}) + 2 \sum_{k=1}^n (M_{t_k} - M_{t_{k-1}})^2 \cdot \sum_{j=k+1}^n (M_{t_j} - M_{t_{j-1}})^2.\end{aligned}$$

By (4.1), we have

$$\mathbb{E}(T_a^{\Delta}(M) - T_{t_k}^{\Delta}(M) | \mathcal{F}_{t_k}) = \mathbb{E}(M_a^2 - M_{t_k}^2 | \mathcal{F}_{t_k}) = \mathbb{E}((M_a - M_{t_k})^2 | \mathcal{F}_{t_k})$$

so that

$$\begin{aligned}\mathbb{E}(T_a^{\Delta}(M)^2) &= \sum_{k=1}^n \mathbb{E}(M_{t_k} - M_{t_{k-1}})^4 + 2 \sum_{k=1}^n \mathbb{E}((M_{t_k} - M_{t_{k-1}})^2 \cdot (M_a - M_{t_k})^2) \\ &\leq \mathbb{E}\left(\sup_k |M_{t_k} - M_{t_{k-1}}|^2 + 2 \sup_k |M_a - M_{t_k}|^2 \right) \cdot T_a^{\Delta}(M).\end{aligned}$$

Let c be such that $|M| \leq c$, so that the supremums are bounded above by $12c^2$. Thus $\mathbb{E}(T_a^{\Delta}(M)^2) \leq 12c^4$.

Thus we conclude that $T_a^{\Delta}(M)$ converges in L^2 to a limit, defined as $[M]_a$.

2. We now show that $[M]$ is continuous. Apply Doob's martingale inequality to the martingale $T^{\Delta_n} - T^{\Delta_m}$ and $p = 2$ so that

$$\mathbb{P}(\sup_{t \leq a} |T_t^{\Delta_n} - T_t^{\Delta_m}| \geq \lambda) \leq \lambda^{-2}\, \mathbb{E}((T_a^{\Delta_n} - T_a^{\Delta_m})^2).$$

Since $\mathbb{E}((T_a^{\Delta_n} - T_a^{\Delta_m})^2) \to 0$ almost surely as $\mathrm{mesh}(\Delta_n) \to 0$, for any k, there exists n_k such that for any $n, m \geq n_k$

$$\mathbb{P}(\sup_{t \leq a} |T_t^{\Delta_n} - T_t^{\Delta_m}| \geq 2^{-k}) \leq 2^{-k}.$$

Define $A_k := \sup_{t\le a}|T_t^{\Delta_{n_k}} - T_t^{\Delta_{n_{k+1}}}| \ge 2^{-k}$, so that the probability that A_k happens infinitely often is 0.

Thus almost surely, $\{T_t^{\Delta_{n_k}}\}$ is a Cauchy sequence with respect to the sup norm on $[0,a]$. Since $T_t^{\Delta_n}$ is continuous for any n by definition, its limit $[M]$ is also continuous.

3. We show that $[M]$ is increasing. Obviously, for any $s,t \in \Delta$, $s < t$ for some partition Δ, we have $T_s^\Delta \le T_t^\Delta$. The inequality holds passing to the limit.
4. $M^2 - [M]$ is a martingale. We have seen that $\mathbb{E}(M_t^2 - T_t^\Delta(M)|\mathcal{F}_s) = M_s^2 - T_s^\Delta(M)$. Pass to the limit. Checking integrability is easy.
5. $M^2 - [M]$ is uniformly integrable. Since $[M]$ is increasing, $\mathbb{E}([M]_t) = \mathbb{E}(M_t^2)$ is bounded.
6. $[M]$ is unique. Suppose Y_1, Y_2, both fanishing at 0, increasing continuous, and have $M^2 - Y_i$ a martingale for $i = 1,2$. Then $Z := Y_1 - Y_2$ is a continuous martingale vanishing at 0. Moreover, Z has finite variation. Apply the lemma below.

Finally, we need to generalize the results to L^2-bounded continuous martingales. Define $T_n := \inf\{t \ge 0 : |M_t| \ge n\}$. Then $T_n \to \infty$ almost surely. Define the stopped process $X_n := M^{T_n}$.

In particular, X_n is a bounded martinale so there exists some A_n continuous, vanishing at 0, such that $X_n^2 - A_n$ is a martingale. Similarly, there exists A_{n+1} such that $X_{n+1}^2 - A_{n+1}$ is a martingale. Thus $(X_{n+1}^2 - A_{n+1})^{T_n} = X_n^2 - A_{n+1}^{T_n}$ is a martingale, and $A_n = A_{n+1}^{T_n}$. This means we can consisdently define $[M] = A_n$ on $[0, T_n]$. Such a defined $[M]$ is clearly continuous, increasing, vanishing at 0, and unique.

We postpoine the proof that $M^2 - [M]$ is a U.I¿ martingale to after the next section. ■

4.2 Lemma. *Let Z be a continuous martingale, $Z_0 = 0$ almost surely, and Z has finite variation. Then $Z = 0$.*

Proof By similar arguments as used to derive the quadratic variation of Brownian motion, if Z has finite variation, then $[Z] = 0$. Then $Z^2 - [Z] = Z^2$ is a martingale, so $\mathbb{E}(Z_t^2) = \mathbb{E}(Z_0^2) = 0$ so $Z_t = 0$ almost surely. ■

5 Local Martingales

Recall:

Definition. An adapted process $\{X_t\}_{t\ge 0}$ is a $(\mathcal{F}_t, \mathbb{P})$-local martingale if there exists stopping times T_n such that T_n is increasing, $\lim T_n = \infty$ a.s., and for every n, the stopped process X^{τ_n} is a martingale.

When is a local martingale in fact a martingale? We have seen that $\int_0^t X_s \, \mathrm{d}B_s$ is a local martingale. If the integrand is in L^2, then it is a martingale.

Another criterion:

Definition. An adapted process X is said to be in class DL if for every $a > 0$, $\Big\{\{X_T\}_{0\le T\le a} : T \text{ stopping time}\Big\}$, is uniformly integrable.

5.1 Proposition. *A local martingale is a martingale if and only if it is in class DL.*

Proof ($\Longleftarrow$) Let $\{X_t\}_{t\geq 0}$ be a martingale. Then it is a local martingale. Moreover, we have for stopping time T, $0 \leq T \leq a$, $X_T = \mathbb{E}(X_a|\mathcal{F}_t)$ by the optional sampling with T and a. Thus by the lemma below, $\{X_T\}_{0\leq T\leq a}$ is uniformly integrable, and thus in class DL.

($\Longrightarrow$) Let $\{X_t\}$ be a local martingale in class DL with reducing sequence $\{X_t^{T_n}\}_{n=1,2,\ldots}$ is uniformly integrable. Thus $X_{T_n\wedge t} \to X_T$ implies that $X_{T_n\wedge t} \to X_t$ in L^1. Thus for all $A \in \mathcal{F}_s$ with $s \in [0,t]$,

$$\mathbb{E}(X_t;A) = \lim_{n\to\infty} \mathbb{E}(X_{T_n\wedge t};A) = \lim_{n\to\infty} \mathbb{E}(X_{T_n} \wedge s;A) = \mathbb{E}(X_s;A)$$

so that $\mathbb{E}(X_t|\mathcal{F}_s) = X_s$ and X is a martingale. ■

5.2 Lemma. *Let X be a $(\Omega,\mathcal{F},\mathbb{P})$-integrable random variable. Let $\{\mathcal{F}_t\}$ be a collection fo sub-σ-algebras of $\mathcal{F}$. Then $\{X_T := \mathbb{E}(X|\mathcal{F}_T)\}$ is uniformly integrable.*

Proof Exercise. ■

We can now show that $M^2 - [M]$ is a uniformly integrable martingale (when M is an L^2-bounded continuous martingale). Note that $\sup_t |M_t^2 - [M]_t| \leq (M_\infty^*)^2 + [M]_\infty$ where $M_\infty^* = \sup_{t\geq 0}|M_t|$. We use the following version of Doob's martingale inequality:

$$\mathbb{E}[(X_\infty^*)^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}(|X_\infty|^p)$$

so that $\mathbb{E}(M_\infty^*) < \infty$. Also,

$$\mathbb{E}([M]_\infty) = \mathbb{E}(\lim_{t\to\infty}[M]_t) = \lim_{t\to\infty}\mathbb{E}([M]_t)$$

and with T_n stopping times such that M^{T_n} are bounded, so that

$$\mathbb{E}([M]_t) = \mathbb{E}(\lim_{n\to\infty}[M]^{T_n}) = \lim_{n\to\infty}\mathbb{E}([M]_t^{T_n}) = \lim_{n\to\infty}\mathbb{E}[(M_t^{T_n})^2] \leq \mathbb{E}(M_\infty^*)^2 < \infty$$

so that $(M_\infty^*)^+[M]_\infty$ is integrable. Thus $M^2 - [M]$ is U.I.

We now show that $M^2 - [M]$ is a martingale. For any stopping time T, we have $|M_T - [M]_T| \leq (M_\infty^*)^2 + [M]_\infty$, which is integrable.. Thus $M^2 - [M]$ is in class DL, so $M^2 - [M]$ is a martingale.

Definition. Let M,N be two L^2-bounded continuous martingales vanishing at 0. Then the quadratic covariation process $[M,N]$ is defined by

$$[M,N] = \frac{1}{4}([M+N]-[M-N]).$$

Note that $[M] = [M,M]$.

5.3 Theorem. *The process $[M,N]$ is the unique finite variation process vanishing at 0 such that $MN - [M,N]$ is a uniformly integrable martingale.*

Proof Same as above. ■

Consider H of the form $H = Z\mathbf{1}(S,T)$ where $S \leq T$ are stopping times, Z is $\mathcal{F}_s$-measurable, and Z is bounded. Define

$$\int_0^t H(s)\,\mathrm{d}M_s = \int_0^t H\,\mathrm{d}M = Z(M_{T\wedge t} - M_{S\wedge t}).$$

5.4 Proposition. *The above integrale defines an adapted, L^2-bounded martingale.*

Proof Adaptedness is trivial. To see L^2-boundedness, we have

$$\|Z(M_{T\wedge t} - M_{S\wedge t})\|_2 \leq \|Z\|_\infty \cdot \|M_{T\wedge t} - M_{S\wedge t}\|_2.$$

By the optional sampling theorem, $\|M_{T\wedge t}\|_2 \leq \|M_t\|_2$, so that $Z(M_{T\wedge t} - M_{S\wedge t})$ is L^2-bounded.

We show the martingale property by using the converse of the optional sampling theorem: $\mathbb{E}(X_T) = \mathbb{E}(X_0)$ for any bounded stopping time T implies that X is a martingale. For any $A \in \mathcal{F}_s$, define $S_A = S$ on A and ∞ on $\Omega \setminus A$. Then S_A (and similarly T_A) is a stopping time. Let R be an arbitrary bounded stopping time. By the optional stopping theorem, we have $\mathbb{E}(M_{T_A\wedge R}) = \mathbb{E}(M_{S_A\wedge R})$, so that $\mathbb{E}(M_{T\wedge R}|\mathcal{F}_s) = M_{S\wedge R}$. Thus

$$\mathbb{E}(Z(M_{T\wedge R} - M_{S\wedge R})) = \mathbb{E}(\mathbb{E}(Z(M_{T\wedge R} - M_{S\wedge R})|\mathcal{F}_s)) = \mathbb{E}(Z\,\mathbb{E}(M_{T\wedge R} - M_{S\wedge R}|\mathcal{F}_s)) = 0$$

for any bounded stopping time R. ■

Use the linear combinations of $H = \sum_{i=1}^n Z_{i-1}\mathbf{1}T_{i-1}, T_i$, define $(H\cdot M)_t = \sum_{i=1}^n Z_{i-1}(M_{T_i\wedge t} - M_{T_{i-1}\wedge t})$. Then $H\cdot M$ is a L^2-bounded martingale.

We may identify L^2-bounded martingales with elements in $L^2(\mathcal{F}_\infty)$. Recall that by maringale convergence, if $\{M_t\}$ is L^2-bounded, then $M_t \to M_\infty \in L^2(\mathcal{F}_\infty)$, which implies that $M_t = \mathbb{E}(M_\infty|\mathcal{F}_t)$. In particular, we may define a topology on the space of L^2-bounded martingales by identification with $L^2(\mathcal{F}_\infty)$.

Now, note that

$$\mathbb{E}((H\cdot M)^2_\infty) = \mathbb{E}[(\sum Z_{i-1}(M_{T_i} - M_{T_{i-1}}))^2] = \mathbb{E}(\sum Z^2_{i-1}(M_{T_i} - M_{T_{i-}})^2)$$

where the optional sampling theorem implies that the cross terms are 0. Since $M_t^2 - [M]_t$ is a UI martingale for stopping times $S \leq T$,

$$\mathbb{E}((M_T - M_S)^2|\mathcal{F}_s) = \mathbb{E}(M_T^2 - M_S^t|\mathcal{F}_s) = \mathbb{E}([M]_T - [M]_S|\mathcal{F}_s)$$

so that

$$\mathbb{E}((H\cdot M)^2_\infty) = \mathbb{E}(\sum Z^2_{i-1}([M]_{T_i} - [M]_{T_{i-1}})) = \mathbb{E}\int_0^\infty H_s^2\,\mathrm{d}[M]_s$$

where the latter integral denotes the pathwise Lebesgue-Stieltjes integral. Define $\|H\|_M = (\mathbb{E}\int_0^\infty H_s^2\,\mathrm{d}[M]_s)^2$.

Let $b\mathcal{E}$ be the set of all processes H with the frm

$$H = \sum_{i=1}^n Z_{i-1}(T_{i-1}, T_i]$$

where $Z_i \in \mathcal{F}_{T_i}$ is bounded.

Definition. The σ-algebra $\mathcal{P}$ generated by $b\mathcal{E}$ is called the **preisible/predictable** σ-algebra. A process is called **previsible/predictable** if if is $\mathcal{P}$-measurable as a mapping from $(0,\infty)\times\Omega$ to $(\mathbb{R},\mathcal{B}(\mathbb{R}))$.

Remark. What process are previsible? By the monotone class theorem, all the cadlag processes are previsible.

Define $\|H\|_M = (\mathbb{E}[\int_0^\infty H_s^2\,\mathrm{d}[M]_s)^{1/2}$, and let $L^2(M)$ denote the set of presivible processes H such that $\|H\|_M < \infty$. We have defined $H\cdot M$ for $H \in b\mathcal{E}$. Our goal is to show that $\overline{b\mathcal{E}} = L^2(M)$, so we can define $H\cdot M$ for any $H\in L^2(M)$ as limits of $H\cdot M$ for $H\in b\mathcal{E}$. Note that

1. $\overline{b\mathcal{E}}$ contains all constant functions
2. If $H_n \in \overline{b\mathcal{E}}\cap L^2(M)$ for all n, the H_n converge uniformly on $(0,\infty)\times\Omega$ to H, then $H\in\overline{b\mathcal{E}}\cap L^2(M)$. Indeed, $\|H_n - H\|_M = (\mathbb{E}\int_0^\infty (H_n(s)-H(s))^2\,\mathrm{d}[M]_s)^{1/2}\to 0$.
3. H_n is non-negative and $H_n \to H$ increasing, then $H\in\overline{b\mathcal{E}}\cap L^2(M)$¿ Similar reason as above, since $\|H_n - H\|_M \to 0$, and apply monotone convergence.

5.5 Theorem. ***(Monotone Class)*** *Let X be a topological space. Suppose*

(i) $C_b(X)$ contains all constant functions,

(ii) $C_b(X)$ is closed under uniform convergence, and

(iii) If $f_n\to f$ is an increasing sequence of non-negative functions with $f_n\in C_b(X)$, and f is bounded, then $f\in C_b(X)$.

Then for any subset $C\subset C_b(X)$ that is closed under multiplication, $C_b(X)$ contains every bounded $\sigma(C)$-measurable function from X to $\mathbb{R}$.

Applying this theorem, we get that $\overline{b\mathcal{E}}\cap L^2(M)$ contains all the bounded previsible processes. For general elements, $H\in L^2(M)$, not necesarily bounded, take $H)n := H\mathbf{1}\{|H|\le n\}$, so

$$(\mathbb{E}\int_0^\infty H_s^2\mathbf{1}\{|H_s|>n\}\,\mathrm{d}[M]_s)^{1/2}\to 0$$

as $n\to\infty$. Thus, we conclude that $\overline{b\mathcal{E}}\cap L^2(M) = L^2(M)$ so $\overline{b\mathcal{E}} = L^2(M)$. Therefore $b\mathcal{E}$ is dense in $L^2(M)$.

We have seen that $\mathbb{E}((H\cdot M)_\infty^2) = \mathbb{E}\int_0^\infty H_s^2\,\mathrm{d}[M]_s$ for all $H\in b\mathcal{E}$, so that the mapping $I : b\mathcal{E}\cap L^2(M)\to L^2(\mathcal{F}_\infty)$ defined by $I(H) = (H\cdot M)_\infty$ for $H\in b\mathcal{E}$ is an isometry.

Since $b\mathcal{E}$ is dense in $L^2(M)$, this isometry extends uniqely to $\overline{b\mathcal{E}} = L^2(M)$. That is, one can define the stochastic integral of any previsible process H such that $\|H\|_M<\infty$.

5.6 Theorem. *Let M be a continuous bounded martingale vanishing at 0, and $H\in L^2(M)$. Then the stochastic integral $H\cdot M$, or $\int H\,\mathrm{d}M$, is the image of H under the extension of the isometry I to $L^2(M)$. In particular,*

$$\mathbb{E}((H\cdot M)_\infty^2) = \mathbb{E}(\int_0^\infty H_s^2\,\mathrm{d}[M]_s)$$

is the stochastic integral with respect to a continuous semimartingale.

We can extend the definition of the above integral by localication. Let $lb\mathcal{P}$ denote the set of locally bounded previsible processes, in other words the set of processes H for which for which there exists a sequence of stopping times increasing to infinity such that $H\mathbf{1}(0, T_n] \in b\mathcal{P}$ where $b\mathcal{P}$ is the space of bounded previsible processes. Let

$$M^2_{0,loc} := \{M : \exists (T_n)_{n=1}^{\infty} \to \infty \text{a.s., s.t.} M^{T_n} \in M^2_0\}$$

where M^2_0 is the set of L^2-bounded martingales vanishing at 0. One can choose $T_1, \dots, T_n$ to be the same for two processes (take the minimum). Define $H \cdot M$ up to T_n by $H\mathbf{1}(0, T_n] \cdot M^{T_n}$. This definition can be shown to be consistent. Then $H \cdot M^{T_n}(t) = H \cdot M^{T_m}(t)$ on $\{t \le T_m \wedge T_m\}$.

Definition. A stochastic processes X is called a **semimartingale** if $X = X_0 + M + A$ where $X_0 \in \mathcal{F}_0$, M is a local martingale vanishing at 0, A is an adapted process with paths of finite variation vanishing at 0.

Let $X = X_0 + M + A$ be a semimartingale, $H \in lb\mathcal{P}$. Then $H \cdot X = H \cdot M + H \cdot A$, where $H \cdot M$ is the stochastic integral and $H \cdot A$ is the pathwise Lebesgue-Stieltjes integral.

5.7 Lemma. *Let M be a bounded continuosu martingale, vanishing at* 0*, and V a continuous adapted process with finite variation, vanishing at* 0*. Then*

$$M_t^2 = \int_0^t 2M_s \, \mathrm{d}M_s + [M]_t$$

$$M_t V_t = \int_0^t M_s \, \mathrm{d}V_s + \int_0^t V_s \, \mathrm{d}M_s$$

Proof For $n = 1, 2, \dots$, define stopping times T_0^n, $T_{k+1}^n := \inf\{t : T_k^n : |M(t) - M(T_k^n)| > 2^{-n}\}$ and $t_k^n = t \wedge T_k^n$. Then

$$\begin{aligned} M_t^2 &= \sum_{k \ge 1} (M(t_k^n)^2 - M(t_{k-1}^n)^2) \\ &= 2 \sum_{k \ge 1} M(t_{k-1}^n)(M(t_k^n) - M(t_{k-1}^n)) + \sum_{k \ge 1} (M(t_k^n) - M(t_{k-1})^n)^2. \end{aligned}$$

Set

$$\begin{aligned} H_t^n &= \sum_{k \ge 1} M(T_{k-1}^n) \cdot \mathbf{1}(T_{k-1}^n, T_k^n] \\ A_t^n &= \sum_{k \ge 1} (M(t_k^n) - M(t_{k-1}^n))^2 \end{aligned}$$

so that $M_t^2 = 2(H^n \cdot M)_t + A_t^n$. Then as $n \to \infty$, $H^n M \to M \cdot M$ and $A^n \to [M]$. For the second equality, define $t_i^n = i2^{-n} \wedge t$. Then

$$M_t V_t = \sum_{i \ge 1} M(t_i^n)(V(t_i^n) - V(t_{i-1}^n)) + \sum_{i \ge 1} V(t_{i-1}^n)(M(t_i^n) - M(t_{i=1}^n))$$

and pass to the limit as $n \to \infty$. But the first term converges t $\int_0^t M_s \, \mathrm{d}V_s$ and the second term converges to $\int_0^t V_s \, \mathrm{d}M_s$. ■

5.8 Theorem. *Let X, Y be continuous semimartingales. Then*

$$X_t Y_t - X_0 Y_0 = \int_0^t X_s \, \mathrm{d}Y_s + \int_0^t Y_s \, \mathrm{d}X_s + [X, Y]_t$$

where $[X, Y]_t$.

Proof Without loss of generality, $X_0 = Y_0 = 0$. Let $X = M+A$, $Y = N+B$. By localization, we can further assume that M, N are bouded martingales. By the previous lemma,

$$(M+N)_t^2 = \int_0^t 2(M+N)_s \,\mathrm{d}(M+N)_s + [M+N]_t$$
$$(M-N)_t^2 = \int_0^t 2(M-N)_s \,\mathrm{d}(M-N)_s + [M-N]_T$$

and taking differences, we have

$$(MN)_t = \int_0^t M_s \,\mathrm{d}N_s + \int_0^t N_s \,\mathrm{d}M_s + [M,N]_t$$

We also have $A+B_t = \int_0^t A_s \,\mathrm{d}B_s + \int_0^t B_s \,\mathrm{d}A_s$, so that

$$\begin{aligned} X_tY_t &= M_tN_t + M_tB_t + N_tA_t + A_tB_t \\ &= \int_0^t (M+A)_s \,\mathrm{d}(N+B)_s + \int_0^t (N+B)_s \,\mathrm{d}(M+A)_s + [X,Y]_t \\ &= \int_0^t X_s \,\mathrm{d}Y_s + \int_0^t Y_s \,\mathrm{d}X_s + [X,Y]_t. \end{aligned}$$ ■

5.9 Proposition. *Let M,N be continuous local martingales and H,K be locally bounded previsible processes. Then*

$$[H\cdot M, K\cdot N]_t = \int_0^t A_sK_s \,\mathrm{d}[M,N]_s$$

or equivalently $\mathrm{d}[H\cdot M, K\cdot N] = HK\,\mathrm{d}[M,N]$.

5.10 Proposition. *1. $(H\cdot M)^T = H\mathbf{1}(0,T]\cdot M = H\cdot M^T$ for stopping time T.*
2. If $H,K \in lb\mathcal{P}$, then $H\cdot(K\cdot M) = (HK)\cdot M$.

Proof We only prove the results for M being L^2-bounded martingale and $H,K \in b\mathcal{P}$. The general case can be done by localization.

1. Given a stopping time T, we define

$$\begin{aligned} f_1 : L^2(\mathcal{F}_\infty) &\to L^1(\mathcal{F}_\infty) \\ Y &\mapsto \mathbb{E}(Y|\mathcal{F}_T) \\ f_2 : L^2(M) &\to L^2(M) \\ H &\mapsto H\mathbf{1}(0,T] \\ f_3 : L^2(M) &\to L^2(M^T) \\ H &\mapsto H \end{aligned} \qquad \begin{aligned} I : L^2(M) &\to L^2(\mathcal{F}_\infty) \\ H &\mapsto (H\cdot M)_\infty \\ I^{(T)} : L^2(M^T) &\to L^2(\mathcal{F}_\infty) \\ H &\mapsto (H\cdot M^T)_\infty \end{aligned}$$

   Then

$$\begin{aligned} (H\cdot M)_\infty^T &= \mathbb{E}((H\cdot M)_\infty|\mathcal{F}_T) = f_1\circ I(H) \\ H\mathbf{1}(0,T]\cdot M &= I\circ f_2(H) \\ H\cdot M^T &= I^{(T)}\circ f_3(H) \end{aligned}$$

where f_1, f_2, f_3 are all linear contractions, so that $f_1 \circ I$, $I \circ F_2$, and $I^{(T)} \circ f_3$ are continuous linear maps. It is easy to check that they agree on $b\mathcal{E}$, so they also agree on $L^2(M)$.

2. The identity clearly holds if $H, K \in b\mathcal{E}$. For $k \in b\mathcal{P}$, let K^n be a uniformly bounded sequence in $b\mathcal{E}$ such that $K^n \to K$ in $L^2(M)$. Then for $H \in b\mathcal{E}$, $(H \cdot M)^T = H\mathbf{1}(0,T] \cdot M = H \cdot M^T$, and if $H, K \in lb\mathcal{P}$, then $H \cdot (K \cdot M) = (HK) \cdot M$ and thus

$$\begin{aligned}
\mathbb{E}[(H \cdot (K \cdot M) - H \cdot (K^n - M))^2_\infty] &= \mathbb{E}[(H \cdot (K - K^n)) \cdot M)^2_\infty] \\
&= \mathbb{E}[\int_0^\infty H_s^2 \,\mathrm{d}[(K - K^n) \cdot M]_s \\
&= \mathbb{E} \int_0^\infty H_s^2 (K_s - K_s^n)^2 \,\mathrm{d}[M]_s \\
&= \mathbb{E}((HK - HK^n) \cdot M)^2_\infty
\end{aligned}$$

and as $n \to \infty$, $H \cdot (K^n \cdot M) \to H \cdot (K \cdot M)$ while $(HK^n) \cdot M \to (HK) \cdot M$. Thus the identity holds for $k \in b\mathcal{P}$. The argument for $H \in b\mathcal{P}$ is similar. ■

5.11 Theorem. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be C^2 and $X_t^1, \dots, X_t^n$ be continuous semimartingales. Then*

$$\begin{aligned}
f(X_t^1, \dots, X_t^n) - f(X_0^1, \dots, X_0^n) &= \sum_{i=1}^n \int_0^t \frac{\partial}{\partial i} f(X_s^1, \dots, X_s^n) \,\mathrm{d}X_s^i \\
&\quad + \frac{1}{2} \sum_{i,j=1}^n \int_0^t \frac{\partial^2}{\partial i \partial j} f(X_s^1, \dots, X_s^n) \,\mathrm{d}[X^i, X^j]_s
\end{aligned}$$

This is essentially the same as the Ito's formula for brownian motion, with the main difference that $\mathrm{d}[X^i, X^j]$ replaces $\mathrm{d}B_i \,\mathrm{d}B_j$, and is typically not $\delta_{ij}\,\mathrm{d}t$.

Proof ... ■