Introduction

Example. Suppose that we have a gambler who repeatedly tosses a fair coin, betting £1 on getting a heads for each toss. Let

$$\xi_k = \begin{cases} 1 & \text{heads on } k\text{th toss} \\ -1 & \text{otherwise} \end{cases}$$

so $(\xi_k)_{k\geq 1}$ is an iid Bern(1/2) sequence. Let $X_n = \sum_{k=1}^n \xi_k$ be the net winnings of the gambler and $X_0 = 0$. Note $(X_n)_{n\geq 0}$ is a simple random walk on $\mathbb{Z}$, hence is a martingale (MG) with respect to $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$. Suppose that at the mth toss, they bet £H_m on heads. Then the net winnings at time n are

$$(H \cdot X)_n = \sum_{k=1}^n H_k(X_k - X_{k-1}).$$

Assume $(H_m)_{m\geq 1}$ is deterministic. We claim $H \cdot X$ is an $\mathcal{F}_n$-MG. Indeed:

(a) Integrability: obvious;

(b) Adapted: obvious;

(c) $\mathbb{E}[(H \cdot X)_{n+1} - (H \cdot X)_n | \mathcal{F}_n] = H_{n+1}\mathbb{E}[X_{n+1} - X_n|\mathcal{F}_n] = 0.$

More generally, the same is true if H_{n+1} is integrable and $\mathcal{F}_n$ measurable for each n. This is called a *previsible process*. As before, $H \cdot X$ gives the winnings of the gambler. This is called a *martingale transform*.

The goal for the first part of the course: extend this reasoning to define

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

where H is previsible and X is a continuous martingale (e.g Brownian motion).

We cannot use the Lebesgue-Stieljes integral to define $(*)$ since this requires X to have finite variation, and the only continuous martingales with finite variation are constant (see later in course). Our strategy to define the Itô integral: set

$$(H \cdot X)_t \text{``}=\text{''} \lim_{\varepsilon\to 0} \sum_{k=1}^{\lfloor t/\varepsilon \rfloor} H_{k\varepsilon}(X_{(k+1)\varepsilon} - X_{k\varepsilon}).$$

However we need to be careful about the type of limit since X in general will be rough (not differentiable), like Brownian motion. To get convergence, we need to take advantage of cancellations. For example, if X is a Brownian motion and

H is a deterministic and continuous process we have

$$\begin{aligned}
&\mathbb{E}\left[\left[\sum_{k=0}^{\lfloor t/\varepsilon \rfloor} H_{k\varepsilon}(X_{(k+1)\varepsilon} - X_{k\varepsilon})\right]^2\right] \\
&= \mathbb{E}\left[\sum_{k=0}^{\lfloor t/\varepsilon \rfloor} H_{k\varepsilon}^2 (X_{(k+1)\varepsilon} - X_{k\varepsilon})^2 + \sum_{j \neq k} H_{k\varepsilon} H_{j\varepsilon} (X_{(k+1)\varepsilon} - X_{k\varepsilon})(X_{(j+1)\varepsilon} - X_{j\varepsilon})\right] \\
&= \mathbb{E}\left[\sum_{k=0}^{\lfloor t/\varepsilon \rfloor} H_{k\varepsilon}^2 (X_{(k+1)\varepsilon} - X_{k\varepsilon})^2\right] \\
&= \sum_{k=0}^{\lfloor t/\varepsilon \rfloor} H_{k\varepsilon}^2 \cdot \varepsilon \\
&\xrightarrow{\varepsilon \to 0} \int_0^t H_s^2 \mathrm{d}s.
\end{aligned}$$

The cancellations that make this work come from MG orthogonality and are what makes it possible to define the Itô integral.

After this we will learn about properties of the Itô integral:

- Stochastic analogue of the chain rule;
- Stochastic analogue of integration by parts.

The formulas will look like those in regular calculus, but with an extra term to reflect that X is rough (quadratic variation). We write

$$Y_t = \int_0^t H_s \mathrm{d}X_s \iff \mathrm{d}Y_t = H_t \mathrm{d}X_t.$$

Itô's formula tells use how to write $\mathrm{d}f(Y_t)$ in terms of $\mathrm{d}Y_t$ for $f \in C^2$. This has many applications, for example

Theorem (Dubins-Schwarz theorem)**.** *Any continuous martingale is a time-change of a Brownian motion.*

Then we will look at Stochastic Differential Equations (SDEs), i.e

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

where b, σ are "nice" and B is a Brownian motion. For $\sigma = 0$ this is just an ODE. For $\sigma \neq 0$ this corresponds to adding noise depending on the time and state of the the system.

Last part of the course: diffusion processes and how they are related to SDEs, as well as how they can be used to solve PDEs involving 2nd order elliptic operators.

0 Preliminaries

Recall that $a : [0, \infty) \to \mathbb{R}$ is *càdlàg* if it is right-continuous and has left limits. Let $a(x^-) = \lim_{y \to x^-} a(y)$ and $\Delta a(x) = a(x) - a(x^-)$. Suppose a is non-decreasing, càdlàg, $a(0) = 0$. Then there exists a unique Borel measure $\mathrm{d}a$ on $[0, \infty)$ such that $\mathrm{d}((s, t]) = a(t) - a(s)$ for all $0 \leq s < t$ (see Part II Probability & Measure).

For f measurable and integrable then the *Lebesgue-Stieljes* integral $f \cdot a$ is defined by

$$(f \cdot a)(t) = \int_{(0,t]} f(s) \mathrm{d}a(s) \ \forall t \geq 0.$$

Then $(f \cdot a)$ is right-continuous. Moreover if a is continuous then $(f \cdot a)$ is continuous and so we can write

$$\int_{(0,t]} f(s) \mathrm{d}a(s) = \int_0^t f(s) \mathrm{d}a(s).$$

We want to integrate against a wider class of functions. Suppose that a^+, a^- are functions satisifying the same conditions as from before (i.e non-decreasing and càdlàg) and set $a = a^+ - a^-$. Define

$$(f \cdot a)(t) = (f \cdot a^+)(t) - (f \cdot a^-)(t)$$

for all f measurable and such that both terms on the RHS are finite. The class of functions which are a difference of càdlàg non-decreasing functions coincides with the class of càdlàg functions of *finite variation.*

Definition. Let $a : [0, \infty) \to \mathbb{R}$ be càdlàg. For each $n \in \mathbb{N}$, $t \geq 0$, let

$$v^n(t) = \sum_{k=0}^{\lceil 2^n t \rceil - 1} |a((k+1)2^{-n}) - a(k2^{-n})|. \qquad (*)$$

Then the limit $v(t) := \lim_{n \to \infty} v^n(t)$ exists and is called the *total variation* of a on $(0, t]$. If $v(t) < \infty$ then we say that a has *finite variation* on $(0, t]$. If a has finite variation on $(0, t]$ for all $t \geq 0$, we say that a is of *finite variation.*

To see that $\lim_{n \to \infty} v^n(t)$ exists, fix $t > 0$ and let $t_n^+ = 2^{-n} \lceil 2^n t \rceil$, $t_n^- = 2^n(\lceil 2^n t \rceil - 1)$ so that $t_n^+ \geq t \geq t_n^-$ for all n and

$$v^n(t) = \sum_{k=0}^{2^n t_n^- - 1} |a((k+1)2^{-n}) - a(k2^{-n})| + |a(t_n^+) - a(t_n^-)|.$$

The triangle inequality implies that the sum is non-decreasing in n, so converges. The càdlàg property tells us that the second term on the RHS converges to $|\Delta a(t)|$, so $v^n(t)$ does indeed converge.

Lemma. *Let a be a càdlàg function of finite variation. Then v is càdlàg of finite variation with $\Delta v(t) = |\Delta a(t)|$ for all $t \geq 0$, and v is non-decreasing. In particular, if a is continuous then v is also continuous.*

Proof. See Example Sheet. □

Proposition. A càdlàg function can be written as a difference of two right-continuous non-decreasing if and only if it has finite variation.

Proof. First assume $a = a^+ - a^-$ for a^+, a^- càdlàg and non-decreasing. We show a has finite variation. Note

$$|a(t) - a(s)| \leq (a^+(t) - a^+(s)) + (a^-(t) - a^-(s)) \ \forall 0 \leq s < t.$$

Plugging this into $(*)$ and using the fact the sum telescopes for monotone functions to get

$$v^n(t) \leq (a^+(t_n^+) - a^+(0)) + (a^-(t_n^+) - a^-(0)).$$

Since a^+, a^- are right-continuous, the RHS converges to $(a^+(t) - a^+(0)) + (a^-(t) - a^-(0))$.

Now we show the reverse direction. Assume a has finite variation $v(t) < \infty$ for all $t > 0$. Set $a^+ = \frac{1}{2}(v + a)$ and $a^- = \frac{1}{2}(v - a)$. Then $a = a^+ - a^-$ and a^+, a^- are càdlàg since v, a are càdlàg (by the above lemma). We show a^+, a^- are non-decreasing. For $0 \leq s < t$ define t_n^+, t_n^- as before and s_n^+, s_n^- analogously. Then

$$\begin{aligned}
&a^+(t) - a^+(s) \\
&= \lim_{n\to\infty} \frac{1}{2}(v^n(t) - v^n(s) + a(t) - a(s)) \\
&= \lim_{n\to\infty} \frac{1}{2}\left[\sum_{k=2^n s_n^+}^{2^n t_n^- - 1} \big(|a((k+1)2^{-n}) - a(k2^{-n})| + a((k+1)2^{-n}) - a(k2^{-n})\big)\right. \\
&\qquad\qquad \left. + |a(t_n^+) - a(t_n^-)| + (a(t_n^+) - a(t_n^-))\right] \\
&\geq 0.
\end{aligned}$$

The same argument works for a^-. □

Random integrators: now we discuss integration against random fuctions of finite variations. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space. Recall a stochastic process $X : \Omega \times [0, \infty) \to \mathbb{R}$ is *adapted* if $X_t = X(\cdot, t)$ is $\mathcal{F}_t$-measurable for all $t \geq 0$. We also say X is *càdlàg* if $X(\omega, \cdot)$ is càdlàg for all $\omega \in \Omega$.

Definition. Given a càdlàg adapted process $A : \Omega \times [0, \infty) \to \mathbb{R}$, its *total variation process* $V : \Omega \times [0, \infty) \to \mathbb{R}$ is defined pathwise by setting $V(\omega, \cdot)$ to be the total variation of $A(\omega, \cdot)$.

Lemma. *If A is càdlàg, adapted and of finite variation, then V is càdlàg, adapted and non-decreasing.*

Proof. We just need to show V is adapted (the rest follows by previous results). For $t \geq 0$, set as before $t_n^- = 2^{-n}(\lceil 2^n t \rceil - 1)$. Then define

$$\tilde{V}_t^n = \sum_{k=0}^{2^n t_n^- - 1} |A_{(k+1)2^{-n}} - A_{k2^{-n}}|$$

so $\tilde{V}^n$ is adapted for all n as $t_n^- \leq t$. Then

$$V_t = \lim_{n \to \infty} \tilde{V}_t^n + |\Delta A(t)|$$

is $\mathcal{F}_t$-measurable as a limit/sum of $\mathcal{F}_t$-measurable functions. □

Recall that a discrete time process $(H_n)_{n\geq 0}$ is *previsible* with respect to $(\mathcal{F}_n)_{n\geq 0}$ if H_{n+1} is $\mathcal{F}_n$-measurable for all $n \geq 0$.

Definition. The *previsible σ-algebra* $\mathcal{P}$ on $\Omega \times (0,\infty)$ is generated by sets of the form $E \times (s,t]$ for $E \in \mathcal{F}_s$ and $s < t$. A process $H : \Omega \times (0,\infty) \to \mathbb{R}$ is *previsible* if it is $\mathcal{P}$-measurable.

Examples.

1. $H(\omega,t) = Z(\omega)\mathbb{1}_{(t_1,t_2]}(t)$ for $t_1 < t_2$ and Z being $\mathcal{F}_{t_1}$-measurable;

2. $H(\omega,t) = \sum_{k=0}^{n-1} Z_k(\omega)\mathbb{1}_{(t_k,t_{k+1}]}(t)$ for $0 = t_0 < \ldots < t_n$ and Z_k $\mathcal{F}_{t_k}$-measurable. H of this form is called a *simple process* and will be important for constructing the Itô integral.

Remark. Simple processes are left-continuous and adapted. It turns out that $\mathcal{P}$ is the smallest σ-algebra on $\Omega \times (0,\infty)$ such that all left-continuous adapted processes are measurable.

In general, a càdlàg process is <u>not</u> previsible, but their left-continuous modification is.

Proposition. Let X be a càdlàg adapted process and let $H_t = X_{t^-}$, $t \geq 0$. Then H is previsible.

Proof. Since X is càdlàg and adapted, it is clear that H is left-continuous and adapted. For each n set

$$H_t^n = \sum_{k=0}^{\infty} H_{k2^{-n}}\mathbb{1}_{(k2^{-n},(k+1)2^{-n}]}(t).$$

Then H_t^n is previsible for all n. By left continuity of H we have $\lim_{n\to\infty} H_t^n = H_t$ for all t. So H is previsible as the limit of previsible functions. □

Remark. The above proposition shows that continuous and adapted processes are previsible.

Proposition. If H is previsible then H_t is $\sigma(\mathcal{F}_s : s < t) = \mathcal{F}_{t^-}$-measurable for all t.

Proof. See Example Sheet. □

Remark. The Poisson process $(N_t)_{t\geq 0}$ is not previsible since N_t is not $\mathcal{F}_{t^-}$-measurable for $(\mathcal{F}_t)_{t\geq 0}$ the natural filtration for N.

We will not show that integrating a previsible process against a càdlàg process which is adapted and has finite variation yields an adapted càdlàg process of finite variation.

Theorem. *Let $A : \Omega \times (0, \infty) \to \mathbb{R}$ be a càdlàg process which is adapted and has finite variation V. Let H be a previsible process with*

$$\int_{(0,t]} |H(\omega, s)| \mathrm{d}V(s) < \infty \ \forall t > 0,\ \omega \in \Omega. \tag{1}$$

Then the process $H \cdot A : \Omega \times [0, \infty) \to \mathbb{R}$ given by

$$(H \cdot A)(\omega, t) = \int_{(0,t]} H(\omega, s) \mathrm{d}A(\omega, s),\ (H \cdot A)(\omega, 0) = 0 \tag{2}$$

is càdlàg adapted and of finite variation.

Proof. The integral in (2) is well-defined due to (1). Indeed, let H^+, H^- be the positive/negative parts of H respectively and let $A^\pm = \frac{1}{2}(V \pm A)$. Then $H = H^+ - H^-$, $A = A^+ - A^-$ and

$$(H \cdot A) = (H^+ - H^-) \cdot (A^+ - A^-) = H^+ \cdot A^+ - H^- \cdot A^+ - H^+ \cdot A^- + H^- \cdot A^-$$

and all terms on the RHS are finite by assumption (1).

We need to show $H \cdot A$ is (1) càdlàg, (2) adapted and (3) of finite variation.

Step 1: note $\mathbb{1}_{(0,s]} \to \mathbb{1}_{(0,t]}$ as $s \downarrow t$ and $\mathbb{1}_{(0,s]} \to \mathbb{1}_{(0,t)}$ as $s \uparrow t$. By definition $(H \cdot A)_t = \int H_s \mathbb{1}(s \in (0, t]) \mathrm{d}A_s$ so

$$\begin{aligned}
(H \cdot A)_t &= \int H_s \lim_{r \downarrow t} \mathbb{1}(s \in (0, r]) \mathrm{d}A_s \\
&= \lim_{r \downarrow t} \int H_s \mathbb{1}(s \in (0, r]) \mathrm{d}A_s && \text{(DCT)} \\
&= \lim_{r \downarrow t} (H \cdot A)_r
\end{aligned}$$

so $H \cdot A$ is right-continuous. An analogous argument shows $H \cdot A$ has left-limits, so is càdlàg. Also $\Delta(H \cdot A)_t = \int H_s \mathbb{1}(s = t) \mathrm{d}A_s = H_t \Delta A_s$.

Step 2: we'll use a "monotone class" style argument. Suppose $H = \mathbb{1}_{B \times (s,u]}$ where $B \in \mathcal{F}_s$ and $s < u$. Then $(H \cdot A)_t = \mathbb{1}_B (A_{t \wedge u} - A_{t \wedge s})$ which is $\mathcal{F}_t$-measurable. Let $\mathcal{A} = \{C \in \mathcal{P} : \mathbb{1}_C \cdot A \text{ is adapted}\}$. We want to show $\mathcal{A} = \mathcal{P}$. Let $\Pi = \{B \times (s, u] : B \in \mathcal{F}_s,\ s < u\}$ so $\Pi \subseteq \mathcal{A}$ and Π is a π-system generating $\mathcal{P}$ by definition. Not difficult to see that $\mathcal{A}$ is a d-system, implying $\mathcal{A} = \mathcal{P}$ by Dynkin's lemma.

Now suppose $H \geq 0$ is previsible. Set

$$\begin{aligned}
H_n &= (2^{-n} \lfloor 2^n H \rfloor) \wedge n \\
&= \sum_{k=0}^{2^n - 1} 2^{-nk} \mathbb{1}(\underbrace{H \in [2^{-n}k, 2^{-n}(k+1))}_{\in \mathcal{P}}) + \mathbb{1}(\underbrace{H \geq n}_{\in \mathcal{P}})
\end{aligned}$$

so H_n is a finite linea combination of functions of the form $\mathbb{1}_C$ for $C \in \mathcal{P}$. Thus $(H^n \cdot A)$ is adapted for all n. By the MCT $(H^n \cdot A)_t \to (H \cdot A)_t$ so $H \cdot A$ is itself adapted. For general previsible H we write $H = H^+ - H^-$ as usual.
Step 3: we have

$$H \cdot A = (H^+ \cdot A^+ + H^- \cdot A^-) - (H^- \cdot A^+ + H^+ \cdot A^-)$$

which is a difference of non-decreasing functions. $\square$

1 Local Martingales

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space.

Definition. Say that $(\mathcal{F}_t)_{t\geq 0}$ satisfies the *usual conditions* if

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

Throughout we assume that $(\mathcal{F}_t)$ satisfies the usual conditions.

For T a stopping time, set $\mathcal{F}_T = \{E \in \mathcal{F} : E \cap \{t \leq T\} \in \mathcal{F}_t \ \forall t \geq 0\}$. Then X_T is $\mathcal{F}_T$-measurable. If X is a martingale then $X^T = X_{T\wedge t}$ is also a martingale. Recall:

Theorem. *Optional stopping theorem Let X be an adapted, càdlàg, integrable process. Then the following are equivalent*

1. *X is a martingale;*
2. *X^T is a martingale for all stopping times T;*
3. *For all bounded stopping times $S \leq T$, we have*
$$\mathbb{E}[X_T|\mathcal{F}_S] = X_s \text{ almost-surely;}$$
4. *For all bounded stopping times T, we have that*
$$\mathbb{E}[X_T] = \mathbb{E}[X_0].$$

Definition. A càdlàg adapted process X is called a *local martingale* if there exists a sequence $(T_n)_{n\geq 1}$ of stopping times with $T_n \uparrow \infty$ almost-surely such that the stopped process X^{T_n} is a martingale for all $n \geq 1$. In this case, we say that $(T_n)_{n\geq 1}$ *reduces* X.

Note that a martingale is always a local martingale as any deterministic sequence $T_n \uparrow \infty$ will reduce it.

Example. Let B be a standard Brownian motion in $\mathbb{R}^3$ and let $M_t = \frac{1}{|B_t|}$. In Example Sheet 4 of Part III Advanced Probability we have seen that

(i) M is bounded in L^2;

(ii) $\mathbb{E}M_t \to 0$ as $t \to \infty$;

(iii) M is a supermartingale.

M cannot be a martingale as otherwise its expectation would vanish by (ii). Now we will show that M is a local martingale. For each $n \geq 1$ set $T_n = \inf\{t \geq 1 : |B_t| < 1/n\} = \inf\{t \geq 1 : |M_t| > n\}$. We want to show:

(i) $(M_t^{T_n})_{t\geq 1}$ is a martingale for all n;

(ii) $T_n \uparrow \infty$ as $n \to \infty$ almost-surely.

Note that $n \leq M_1$ implies $T_n = 1$ and $n > M_1$ implies $T_n > 1$. Since $|B_t|$ cannot hit $1/n$ before hitting $1/(n+1)$, we see T_n is non-decreasing.

In Advanced Probability we saw that for $f \in C_b^2(\mathbb{R}^3)$ (C^2 with bounded derivatives) we have

$$f(B_t) - f(B_0) - \frac{1}{2}\int_0^t \Delta f(B_s)\mathrm{d}s$$

is a martingale. Note that $f(x) = 1/|x|$ is harmonic in $\mathbb{R}^3 \setminus \{0\}$. Let $(f^n)_{n\geq 1}$ be a sequence of $C_b^2(\mathbb{R}^3)$ functions with $f^n(x) = 1/|x|$ on $\{|x| \geq 1/n\}$. If $0 < |B_1| < 1/n$ then $T_n = 1$ and $M_t^{T_n} = M_1$ is a martingale. Since $B_1 \neq 0$ almost-surely, we have $|B_1| > 1/n$ for all n sufficiently large in which case $f(B_{t\wedge T_n}) = f^n(B_{t\wedge T_n})$. Thus

$$\begin{aligned} M_{t\wedge T_n} &= f(B_{t\wedge T_n}) - f(B_1) + f(B_1) \\ &= \left(f(B_{t\wedge T_n}) - f(B_1) - \frac{1}{2}\int_1^{t\wedge T_n} \Delta f(B_s)\mathrm{d}s \right) + f(B_1) \\ &= \left(\underbrace{f^n(B_{t\wedge T_n}) - f^n(B_1) - \frac{1}{2}\int_1^{t\wedge T_n} \Delta f^n(B_s)\mathrm{d}s}_{\text{martingale}} \right) + f^n(B_1) \end{aligned}$$

so $M^{T_n} = (M_{t\wedge T_n})_{t\geq 1}$ is a martingale. Now we show $T_n \uparrow \infty$ almost-surely as $n \to \infty$. Since $T_n \leq T_{n+1}$ it suffices to show $T_n \to \infty$. For each R let $S_R = \inf\{t \geq 1 : |B_t| > R\} = \inf\{t \geq 1 : M_t < 1/R\}$. Then $S_R \to \infty$ as $R \to \infty$. We have

$$\begin{aligned} \mathbb{P}(\lim_n T_n < \infty) &\leq \mathbb{P}(\exists R : T_n < S_R \ \forall n) \\ &= \lim_{R\to\infty}\lim_{n\to\infty} \mathbb{P}(T_n < S_R). \end{aligned}$$

The OST says that $\mathbb{E}[M_{T_n \wedge S_R}] = \mathbb{E}[M_1] := \mu \in (0, \infty)$. Also

$$\begin{aligned} \mathbb{E}[M_{T_n\wedge S_R}] &= n\mathbb{P}(T_n < S_R) + \frac{1}{R}\mathbb{P}(S_R \leq T_n) \\ &= n\mathbb{P}(T_n < S_R) + \frac{1}{R}(1 - \mathbb{P}(T_n < S_R)) \\ &= \mu \end{aligned}$$

so $\mathbb{P}(T_n < S_R) = \frac{m-1/R}{n-1/R} \to 0$ as $n \to \infty$. Therefore M is a non-negative local martingale but not a martingale. It is also a super martingale and bounded in L^2.

We actually have:

Proposition. If X is a local martingale and X is non-negative then X is a supermartingale.

Proof. Let (T_n) be a reducing sequence for X. Then for any $s \leq t$ we have that

$$\begin{aligned}
\mathbb{E}[X_t|\mathcal{F}_s] &= \mathbb{E}[\lim_n X_{t\wedge T_n}|\mathcal{F}_s] \\
&\leq \liminf_{n\to\infty} \mathbb{E}[X_{t\wedge T_n}|\mathcal{F}s] && \text{(Fatou)} \\
&= \lim_{n\to\infty} X_{s\wedge T_n} \\
&= X_s \text{ almost-surely.}
\end{aligned}$$

□

We often work with local martingales instead of martingales because we want to avoid having to worry about integrability.

Definition. A collection $\mathcal{X}$ of random variables is *uniformly integrable* (UI) if

$$\sup_{X \in \mathcal{X}} \mathbb{E}[|X| \mathbb{1}(|X| > \lambda)] \to 0.$$

Some examples of UI families are:

1. Uniformly bounded random variables;
2. Uniformly L^p-bounded random variables for $p > 1$;
3. There exists Y integrable such that $|X| \leq Y \; \forall X \in \mathcal{X}$.

Lemma. *Suppose that $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Then*

$$\mathcal{X} = \{\mathbb{E}[X|\mathcal{G}] : \mathcal{G} \subseteq \mathcal{F} \text{ a sub-}\sigma\text{-algebra}\}$$

is a UI family.

Proof. Example Sheet 1. □

Proposition. The following are equivalent:

(i) X is a martingale;

(ii) X is a local martingale and for all $t \geq 0$ the family

$$\mathcal{X}_t = \{X_T : T \text{ is a stopping time with } T \leq t\}$$

is UI.

Proof. First suppose X is a martingale. By the Optional Stopping Theorem, if $T \leq t$ is a stopping time then $\mathbb{E}[X_t|\mathcal{F}_T] = X_T$ and so it follows by the previous lemma that $\mathcal{X}_t$ is UI.

Now for the converse, suppose X is a local martingale with $\mathcal{X}_t$ UI for all $t \geq 0$. To show X is a martingale, by the Optional Stopping Theorem it suffices to show that for all bounded stopping times T we have $\mathbb{E}X_T = \mathbb{E}X_0$. Let $(T_n)_{n \geq 0}$ be a reducing sequence for X and let $T \leq t$ be a stopping time. Then

$$\mathbb{E}X_0 = \mathbb{E}X_0^{T_n} = \mathbb{E}X_T^{T_n} = \mathbb{E}X_{T \wedge T_n}$$

by the OST applied to the martingale X^{T_n}. Since $\{X_{T \wedge T_n} : n \geq 0\}$ is UI and $X_{T \wedge T_n} \to X_T$ almost-surely as $n \to \infty$ we have $X_{T \wedge T_n} \to X_T$ in L^1. Hence $\mathbb{E}X_{T \wedge T_n} \to \mathbb{E}X_T$ implying $\mathbb{E}X_0 = \mathbb{E}X_T$. □

Corollary. *A bounded local martingale is a martingale. More generally, if X is a local martingale and there exists Y integrable such that $|X_t| \leq Y$ for all $t \geq 0$, then X is a martingale.*

Theorem. *Let X be a continuous local martingale with $X_0 = 0$. If X has finite variation then $X = 0$ almost-surely.*

Proof. Let V be the total variation process for X. Then $V_0 = 0$ and V is continuous, adapted and non-decreasing. Let $T_n = \inf\{t \geq 0 : V_t = n\}$ for $n \in \mathbb{N}$. Then $T_n \uparrow \infty$ as $n \to \infty$ since X has finite variation. Moreover $|X_t^{T_n}| = |X_{t\wedge T_n}| \leq V_{t\wedge T_n} \leq n$. Thus X^{T_n} is a bounded local martingale, so a martingale. To prove that $X = 0$ it suffices to show $X^{T_n} = 0$ for all n. Fix $n \geq 1$ and let $Y = X^{T_n}$. Y is a continuous bounded martingale with $Y_0 = 0$. To prove $Y = 0$ it suffices to show that $\mathbb{E}Y_t^2 = 0$ for all $t \geq 0$ [this implies $Y_t = 0$ for all $t \in \mathbb{Q}$ almost-surely, so by continuity $Y = 0$ almost-surely]. Fix $t \geq 0$ and $N \geq 1$ and let $t_k = \frac{k}{N}t$ for $0 \leq k \leq N$. Then

$$\begin{aligned}
\mathbb{E}Y_t^2 &= \mathbb{E}\left[\sum_{k=0}^{N-1}\left(Y_{t_{k+1}}^2 - Y_{t_k}^2\right)\right] \\
&= \mathbb{E}\left[\sum_{k=0}^{N-1}(Y_{t_{k+1}} - Y_{t_k})^2\right] && \text{(MG orthogonality)} \\
&\leq \mathbb{E}\left[\underbrace{\max_{0\leq k\leq N-1}|Y_{t_{k+1}} - Y_{t_k}|}_{\leq V_{t\wedge T_n}\leq n}\underbrace{\sum_{k=0}^{N-1}|Y_{t_{k+1}} - Y_{t_k}|}_{\leq V_{t\wedge T_n}\leq n}\right] \\
&\leq n^2.
\end{aligned}$$

Since Y is continuous, $\lim_{N\to\infty}\max_{0\leq k\leq N-1}|Y_{t_{k+1}} - Y_{t_k}| = 0$ almost-surely. Hence by the bounded convergence theorem, $\mathbb{E}Y_t^2 = 0$. □

Remark.

(i) The above proof requires continuity in an essential way; the theorem is not true otherwise.

(ii) The theorem implies Brownian motion has infinite variation, so cannot use Lebesgue-Stieljes integral to define the integral against a Brownian motion.

For a continuous local martingale, there is always an explicit way of choosing the reducing sequence.

Proposition. Let X be a continuous local martingale with $X_0 = 0$. Then $T_n = \inf\{t \geq 0 : |X_t| = n\}$ reduces X.

Proof. First we show T_n is a stopping time. Indeed

$$\begin{aligned}
\{T_n \leq t\} &= \{\sup_{0\leq s\leq t}|X_s| \geq n\} \\
&= \bigcap_{k=1}^{\infty}\bigcup_{\substack{s\leq t\\ s\in\mathbb{Q}}}\{|X_s| > n - 1/k\}.
\end{aligned}$$

Note that $\sup_{0\le s\le t}|X_s(\omega)| < \infty$ so there exists $n(\omega,t)\in\mathbb{N}$ such that $n(\omega,t)\ge \sup_{0\le s\le t}|X_s(\omega)|$. Then if $n\ge n(\omega,t)$ we have $T_n(\omega)\ge t$. Thus the T_n become arbitrarily large as $n\to\infty$, i.e $T_n\uparrow\infty$.

Now we show (T_n) reduces X. Let (T_n^*) denote a reducing sequence for X (exists since X is a local martingale). Then $X^{T_m^*}$ is a martingale for all m. The OST says $X^{T_n\wedge T_m^*}$ is a martingale for all m. Hence X^{T_n} is a local martingale with reducing sequence (T_m^*). Since X^{T_n} is also bounded it is therefore a martingale. □

2 The Stochastic Integral

Goal: to be able to integrate against a continuous local martingale. How does one construct an integral? We want a linear map $I : X \to Y$ for normed vector spaces X, Y. Steps:

1. Define I on some dense $\mathcal{D} \subseteq X$;
2. Show $I|_{\mathcal{D}} : \mathcal{D} \to Y$ is a continuous linear map. Then extend I to X by continuity.

So we need to specify $\mathcal{D}, X, Y$ and prove our integral is continuous (called the Itô isometry).

Theorem. *Let X be a càdlàg, L^2-bounded martingale. Then there exists X_∞ such that $X_t \to X_\infty$ almost-surely and in L^2. Furthermore $\mathbb{E}[X_\infty|\mathcal{F}_t] = X_t$ for all $t \geq 0$. We call X_∞ the "final value" of X.*

Proof. See Part III Advanced Probability. □

Proposition (Doob's L^2-inequality)**.** Let X be a càdlàg, L^2-bounded martingale. Then $\mathbb{E}[\sup_{t\geq 0} |X_t|^2] \leq 4\mathbb{E}[X_\infty^2]$.

We define

$$\begin{aligned}
\mathcal{M}^2 &= \{L^2\text{-bounded, càdlàg martingales}\}\\
\mathcal{M}^2_c &= \{L^2\text{-bounded, continuous martingales}\}\\
\mathcal{M}^2_{c,\text{loc}} &= \{L^2\text{-bounded, continuous local martingales}\}.
\end{aligned}$$

Definition. A process $H : \Omega \times (0, \infty) \to \mathbb{R}$ is called a *simple process* if it is of the form

$$H(\omega, t) = \sum_{k=0}^{n-1} Z_k(\omega)\mathbb{1}_{(t_k,t_{k+1}]}(t)$$

for $n \geq 1$, $0 = t_0 < \ldots, t_n$, Z_k bounded and $\mathcal{F}_{t_k}$-measurable random variables. Let S be the set of all simple processes.

We will define $(H \cdot M)_t$ for $H \in S$, $M \in \mathcal{M}^2$. Then we aim to extend the integral to more general integrands (e.g $M \in \mathcal{M}^2_C$).

Integrating a simple process

Suppose that $H_t = \sum_{k=0}^{n-1} Z_k\mathbb{1}_{(t_k,t_{k+1}]}(t)$ is a simple process, $M \in \mathcal{M}^2$. Set

$$(H \cdot M)_t = \sum_{k=0}^{n-1} Z_k(M_{t\wedge t_{k+1}} - M_{t\wedge t_k}).$$

Proposition. If $H \in S$, $M \in \mathcal{M}^2$, then $H \cdot M \in \mathcal{M}^2$. Moreover,

$$\mathbb{E}[(H \cdot M)^2_\infty] = \sum_{k=0}^{n-1} \mathbb{E}[Z_k^2(M_{t_{k+1}} - M_{t_k})^2] \leq 4\|H\|^2_\infty\mathbb{E}[(M_\infty - M_0)^2].$$

Proof. First we show $H \cdot M$ is a martingale. Suppose that $t_k \leq s < t \leq t_{k+1}$. Then we have

$$(H \cdot M)_t - (H \cdot M)_S = Z_k(M_t - M_s)$$

so that

$$\mathbb{E}[(H \cdot M)_t - (H \cdot M)_s | \mathcal{F}_s] = Z_k \mathbb{E}[M_t - M_s | \mathcal{F}_s] = 0$$

since Z_k is $\mathcal{F}_s$ measurable and $M \in \mathcal{M}^2$. Suppose that $0 \leq t_j \leq s \leq t_{j+1} \leq t_k \leq t \leq t_{k+1}$. Then

$$\begin{aligned}
&\mathbb{E}[(H \cdot M)_t - (H \cdot M)_s | \mathcal{F}_s] \\
&= \mathbb{E}\left[\sum_{i=0}^{k-1} Z_i(M_{t_{i+1}} - M_t) + Z_k(M_t - M_{t_k}) \right.\\
&\qquad \left. - \left(\sum_{i=0}^{j-1} Z_i(M_{t_{i+1}} - M_{t_i}) + Z_j(M_s - M_{t_j})\right) | \mathcal{F}_s \right] \\
&= \sum_{i=j+1}^{k-1} \mathbb{E}[Z_i(M_{t_{i+1}} - M_{t_i}) | \mathcal{F}_s] + \mathbb{E}[Z_j(M_{t_{j+1}} - M_s) | \mathcal{F}_s] \\
&\qquad + \mathbb{E}[Z_k(M_t - M_{t_k}) | \mathcal{F}_s] \\
&= 0
\end{aligned}$$

where we used

$$\mathbb{E}[Z_i(M_{t_{i+1}} - M_{t_i}) | \mathcal{F}_s] = \mathbb{E}[Z_i \mathbb{E}[M_{t_{i+1}} - M_{t_i} | \mathcal{F}_{t_i}] | \mathcal{F}_s] = 0$$

for all $j + 1 \leq i \leq k - 1$, as well as

$$\mathbb{E}[Z_j(M_{t_{j+1}} - M_s) | \mathcal{F}_s] = Z_j \mathbb{E}[M_{t_{j+1}} - M_s | \mathcal{F}_s] = 0$$

and

$$\mathbb{E}[Z_k(M_t - M_{t_k}) | \mathcal{F}_s] = \mathbb{E}[Z_k \mathbb{E}[M_t - M_{t_k} | \mathcal{F}_{t_k}] | \mathcal{F}_s] = 0.$$

Thus $H \cdot M$ is a martingale.

Now we show $H \cdot M$ is L^2-bounded. If $j < k$ we have

$$\begin{aligned}
&\mathbb{E}[Z_j(M_{t_{j+1}} - M_{t_j}) Z_k(M_{t_{k+1}} - M_{t_k})] \\
&= \mathbb{E}[Z_j(M_{t_{j+1}} - M_{t_j}) \mathbb{E}[Z_k(M_{t_{k+1}} - M_{t_k}) | \mathcal{F}_{t_k}]].
\end{aligned}$$

So

$$\begin{aligned}
\mathbb{E}[(H\cdot M)_t^2] &= \mathbb{E}\left[\left(\sum_{k=0}^{n-1} Z_k(M_{t_{k+1}\wedge t} - M_{t_k\wedge t})\right)^2\right] \\
&= \sum_{k=0}^{n-1} \mathbb{E}[Z_k^2(M_{t_{k+1}\wedge t} - M_{t_k\wedge t})^2] \\
&\le \|H\|_\infty^2 \sum_{k=0}^{n-1} \mathbb{E}[(M_{t_{k+1}\wedge t} - M_{t_k\wedge t})^2] \\
&\le 4\|H\|_\infty^2 \mathbb{E}[(M_\infty - M_0)^2]. && \text{(Doob's } L^2\text{-inequality)}
\end{aligned}$$

This bound is uniform in t so $H\cdot M$ is L^2-bounded, and $H\cdot M \in \mathcal{M}^2$.

Finally we have

$$\begin{aligned}
\mathbb{E}[(H\cdot M)_\infty^2] &\le \lim_{t\to\infty} \mathbb{E}[(H\cdot M)_t^2] && \text{(Fatou)} \\
&\le \sup_{t\ge 0} \mathbb{E}[(H\cdot M)_t^2] \\
&\le 4\|H\|_\infty^2 \mathbb{E}[(M_\infty - M_0)^2].
\end{aligned}$$

□

Space of integrators

We want a space of integrators. If X is càdlàg and adapted, define the norm $|||X||| = \|X^*\|_{L^2}$ where $X^* = \sup_{t\ge 0}|X_t|$. Let $\mathcal{C}^2$ be the set of càdlàg adapted processes X with $|||X||| < \infty$.

Define a norm on $\mathcal{M}^2$ by $\|X\| = \|X_\infty\|_{L^2}$ for $X \in \mathcal{M}^2$. This is clearly a semi-norm. To see that it's positive definite, suppose $\|X\| = \|X_\infty\|_{L^2} = 0$. Then $X_\infty = 0$ almost-surely. Hence $X_t = \mathbb{E}[X_\infty|\mathcal{F}_t] = 0$ almost-surely for all $t \ge 0$. The càdlàg property then implies $X = 0$ almost-surely.

Define $\mathcal{M}$ to be the space of càdlàg martingales, $\mathcal{M}_c$ to be the space of continuous martingales, and $\mathcal{M}_{c,\text{loc}}$ to be the space of continuous local martingales.

Proposition.

(a) $(\mathcal{C}^2, |||\cdot|||)$ is complete;

(b) $\mathcal{M}^2 = \mathcal{M} \cap \mathcal{C}^2$;

(c) $(\mathcal{M}^2, \|\cdot\|)$ is a Hilbert space, $\mathcal{M}^2_c = \mathcal{M}_c \cap \mathcal{M}^2$ is a closed subspace;

(d) For $\mathcal{F}_\infty = \sigma(\mathcal{F}_t : t \geq 0)$, the map $\mathcal{M}^2 \to L^2(\mathcal{F}_\infty)$ defined by $X \mapsto X_\infty$ is an isometry.

Remark. We can always identify an element of $\mathcal{M}^2$ with its final value so $(\mathcal{M}^2, \|\cdot\|)$ inherits the Hilbert space structure $(L^2(\mathcal{F}_\infty), \|\cdot\|_{L^2})$. Since $(\mathcal{M}^2_c, \|\cdot\|)$ is a closed linear subspace of $(\mathcal{M}^2, \|\cdot\|)$ by (c), it is also a Hilbert space. Thus is the collection of processes we will integrate against.

Proof.

(a) Suppose (X^n) is a Cauchy sequence in $\mathcal{C}^2$ with respect to $|||\cdot|||$. Then there exists a subsequence (X^{n_k}) of (X^n) such that $\sum_{k\geq 1} |||X^{n_k} - X^{n_{k+1}}||| < \infty$. Thus

$$\left\| \sum_{k\geq 1} \sup_{t\geq 0} |X^{n_k} - X^{n_{k+1}}| \right\|_{L^2} \leq \sum_{k\geq 1} |||X^{n_k} - X^{n_{k+1}}|||$$

and therefore $\sum_{k\geq 1} \sup_{t\geq 0} |X^{n_k} - X^{n_{k+1}}|$ is finite almost-surely. So (X^{n_k}) is uniformly Cauchy on $[0, \infty)$ almost-surely, hence converges uniformly to a cadlag limit X. Then

$$\begin{aligned}
|||X - X^n|||^2 &= \mathbb{E}[\sup_{t\geq 0} |X^n_t - X_t|^2] \\
&= \mathbb{E}[\lim_{k\to\infty} \sup_{t\geq 0} |X^n_t - X^{n_k}_t|^2] \\
&\leq \liminf_{k\to\infty} \mathbb{E}[\sup t \geq 0 |X^n_t - X^{n_k}_t|^2] && \text{(Fatou)} \\
&= \liminf_{k\to\infty} |||X^n_t - X^{n_k}_t|||^2 \\
&\xrightarrow{n\to\infty} 0 \text{ almost-surely}
\end{aligned}$$

since (X^n) is Cauchy.

(b) Suppose $X \in \mathcal{C}^2 \cap \mathcal{M}$. Then $|||X||| < \infty$ and so

$$\sup_{t\geq 0} \|X_t\|_{L^2} \leq \|\sup_{t\geq 0} |X_t|\|_{L^2} = |||X||| < \infty.$$

Hence $M \in \mathcal{M}^2$. Now suppose $X \in \mathcal{M}^2$. By Doob's L^2-inequality

$$|||X||| \leq 2\|X_\infty\|_{L^2} = 2\|X\| < \infty$$

so $X \in \mathcal{C}^2 \cap \mathcal{M}$.

(c) Note that $(X, Y) \mapsto \mathbb{E}[X_\infty Y_\infty]$ defines an inner product on $\mathcal{M}^2$. For $X \in \mathcal{M}^2$ we have

$$\|X\| \leq |||X||| \leq 2\|X\|$$

where the first inequality is obvious and the second follows by Doob's L^2-inequality. Hence $\|\cdot\|$ and $|||\cdot|||$ are equivalent on $\mathcal{M}^2$. So to show $(\mathcal{M}^2, \|\cdot\|)$ is complete it suffices to show $(\mathcal{M}^2, |||\cdot|||)$ is. So let (X^n) be a sequence in $\mathcal{M}^2$ such that $|||X^n - X||| \to 0$ as $n \to \infty$ for some $X \in \mathcal{C}^2$. We know X is cadlag, adapted and L^2-bounded since $X \in \mathcal{C}^2$. To prove its a martingale, fix $s < t$ we have that

$$\begin{aligned}
\|\mathbb{E}[X_t|\mathcal{F}_s] - X_s\|_{L^2} &= \|\mathbb{E}[X_t - X_t^n|\mathcal{F}_s] + X_s^n - X_s\|_{L^2} \\
&\leq \|\mathbb{E}[X_t - X_t^n|\mathcal{F}_s]\|_{L^2} + \|X_s^n - X_s\|_{L^2} \\
&\leq \|X_t^n - X_t\|_{L^2} + \|X_s^n - X_s\|_{L^2} \\
&\leq 2|||X^n - X||| \\
&\to 0 \text{ as } n \to \infty
\end{aligned}$$

and so $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ and X is a martingale.

(d) True by definition.

□

Space of integrands

Definition. Let (X^n) be a sequence of processes. We say that $X^n \to X$ *uniformly on compact sets in probability* (UCP) if for all $\varepsilon > 0$,

$$\mathbb{P}\left[\sup_{s \leq t} |X_s^n - X_s| > \varepsilon\right] \xrightarrow{n \to \infty} 0 \text{ almost-surely.}$$

Theorem. *Suppose that $M \in \mathcal{M}_{c,loc}$. Then there exists a unique (up to indistinguishability) continuous adapted non-decreasing process $[M]$ such that $[M]_0 = 0$, $M^2 - [M] \in \mathcal{M}_{c,loc}$. Moreover if we set*

$$[M]_t^n = \sum_{k=0}^{\lceil 2^n t \rceil - 1} (M_{(k+1)2^{-n}} - M_{k2^{-n}})^2$$

then $[M]^n \to [M]$ (UCP) as $n \to \infty$.

The process $[M]$ is called the quadratic variation of M.

Example. Let B be a standard Brownian motion. Then $(B_t^2 - t)_{t \geq 0}$ is a martingale. Therefore $[B]_t = t$. We will prove later that Brownian motion is characterised by this property, i.e $M \in \mathcal{M}_{\text{c,loc}}$ and $[M]_t = t$ for all $t \geq 0$ implies M is a standard Brownian motion (Levy characterisaton of Brownian motion).

Proof. Replace M with $M_t - M_0$ so WLOG $M_0 = 0$.

First we show uniqueness. If A, A' are two non-decreasing continuous adapted processes satisfying the conditions in the theorem, we have

$$A_t - A'_t = (M_t^2 - A'_t) - (M_t^2 - A_t).$$

Note the LHS is continuous of bounded variation. The RHS is a process in $\mathcal{M}_{c,\text{loc}}$. Together this implies $A - A'$ is constant and since $A_0 = A'_0 = 0$ we have $A = A'$.

□

First we need a lemma.

Lemma. *Suppose that $M \in \mathcal{M}$ is bounded. Then for any $N \in \mathbb{N}$ and $0 = t_0 < \ldots < t_n < \infty$ we have that*

$$\mathbb{E}\left[\left(\sum_{k=0}^{N-1}(M_{t_{k+1}} - M_{t_k})^2\right)\right] \leq 48\|M\|_{C^\infty}^4.$$

Proof. Define $\Delta_k = M_{t_{k+1}} - M_{t_k}$. We have

$$\mathbb{E}\left[\left(\sum_{k=0}^{N-1}\Delta_k^2\right)^2\right] = \sum_{k=0}^{N-1}\mathbb{E}[\Delta_k^4] + 2\sum_{k=0}^{N-1}\mathbb{E}[\Delta_k^2\sum_{j=k+1}^{N-1}\Delta_j^2]. \tag{$*$}$$

For each fixed k we have that

$$\begin{aligned}
\mathbb{E}\left[\Delta_k^2\sum_{j=k+1}^{N-1}\Delta_j^2\right] &= \mathbb{E}\left[\Delta_k^2\mathbb{E}\left[\sum_{j=k+1}^{N-1}\Delta_j^2|\mathcal{F}_{t_{k+1}}\right]\right] \\
&= \mathbb{E}\left[\Delta_k^2\mathbb{E}\left[\left(\sum_{j=k+1}^{N-1}\Delta_j\right)^2|\mathcal{F}_{t_{k+1}}\right]\right] \qquad \text{(MG orthogonality)} \\
&= \mathbb{E}\left[\Delta_k^2\mathbb{E}\left[(M_{t_N} - M_{t_{k+1}})^2|\mathcal{F}_{t_{k+1}}\right]\right] \\
&= \mathbb{E}[\Delta_k^2(M_{t_N} - M_{t_{k+1}})^2].
\end{aligned}$$

Thus

$$
\begin{aligned}
(*) &\leq \mathbb{E}\left[\left(\max_{0\leq j\leq N-1}|M_{t_{j+1}}-M_{t_j}|^2+\max_{0\leq j\leq N-1}|M_{t_N}-M_{t_j}|^2\right)\left(\sum_{k=0}^{N-1}\Delta_k^2\right)\right] \\
&\leq 2\|M\|_{C^\infty}^2\,\mathbb{E}\left[\sum_{k=0}^{N-1}\Delta_k^2\right] && ((a+b)^2\leq 2(a^2+b^2)) \\
&= 2\|M\|_{C^\infty}^2\,\mathbb{E}\left[\left(\sum_{k=0}^{N-1}\Delta_k\right)^2\right] && (\text{MG orthogonality}) \\
&= 2\|M\|_{C^\infty}^2\,\mathbb{E}[(M_{t_N}-M_{t_0})^2] \\
&= 48\|M\|_{C^\infty}^4.
\end{aligned}
$$

□

Now we prove the theorem.

Proof. Replace M with $M_t - M_0$ so WLOG $M_0 = 0$.

First we show uniqueness. If A, A' are two non-decreasing continuous adapted processes satisfying the conditions in the theorem, we have

$$A_t - A'_t = (M_t^2 - A'_t) - (M_t^2 - A_t).$$

Note the LHS is continuous of bounded variation. The RHS is a process in $\mathcal{M}_{c,\text{loc}}$. Together this implies $A - A'$ is constant and since $A_0 = A'_0 = 0$ we have $A = A'$.

Now we show existence. WLOG $M_0 = 0$ (by replacing M_t with $M_t - M_0$ if necessary).

Suppose $M \in \mathcal{M}_c$ is bounded (i.e $M \in \mathcal{M}_c^2$). Fix $T > 0$ and set

$$H_t^n = \sum_{K=0}^{\lceil 2^n T\rceil - 1} M_{k2^{-n}} \mathbb{1}_{(k2^{-n},(k+1)2^{-n}]}(t).$$

Then $H^n \in S$ for all n and set

$$H_t^n = (H^n \cdot M)_t = \sum_{k=0}^{\lceil 2^n T\rceil - 1} M_{k2^{-n}}(M_{(k+1)2^{-n}nt} - M_{k2^{-n}nt}).$$

Then $X^n \in \mathcal{M}_c$ is bounded so $X^n \in \mathcal{M}_c^2$. We will show (X^n) is Cauchy in $(\mathcal{M}_c^2, \|\cdot\|)$ and hence has a limit in $\mathcal{M}_c^2$. Fix $n \geq m \geq 1$ and write $H = H^n - H_m$

so that $X^n - X^m = (H^n - H^m) \cdot M = H \cdot M$. Then

$$\begin{aligned}
\|X^n - X^m\|^2 &= \mathbb{E}[(H \cdot M)_\infty^2] \\
&= \mathbb{E}[(H \cdot M)_T] \\
&= \mathbb{E}\left[\left(\sum_{k=0}^{\lceil 2^n T \rceil - 1} H_{k2^{-n}}(M_{(k+1)2^{-n}} - M_{k2^{-n}})\right)^2\right] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(MG orthogonality)} \\
&\leq \mathbb{E}\left[\sup_{t \in [0,T]} |H_t|^2 \sum_{k=0}^{\lceil 2^n T \rceil - 1} (M_{(k+1)2^{-n}} - M_{k2^{-n}})^2\right] \\
&\leq \left(\mathbb{E}\left[\sup_{t \in [0,T]} |H_t|^4\right]\right)^{1/2} \left(\mathbb{E}\left[\left(\sum_{k=0}^{\lceil 2^n T \rceil - 1} (M_{(k+1)2^{-n}} - M_{k2^{-n}})^2\right)^2\right]\right)^{1/2}.
\end{aligned}$$

Now the first term in the product is bounded as

$$\sup_{t \in [0,T]} |H_t|^4 = \sup_{t \in [0,T]} |H_t^n - H_t^m|^4 \leq 16\|M\|_{C^\infty}^4.$$

Also $\sup_{t \in [0,T]} |H_t^n - H_t^m|^4 \to 0$ as $n, m \to \infty$ since M is continuous. Bounded convergence hence shows the first term goes to 0. The second term is bounded by

$$(48\|M\|_{C^\infty}^4)^{1/2} < \infty$$

by the lemma, so $\|X^n - X_m\| \to 0$ as $n, m \to \infty$. Since $(\mathcal{M}_c^2, \|\cdot\|)$ is complete, there exists $Y \in \mathcal{M}_c^2$ such that $X_n \to Y$ almost-surely in $\mathcal{M}_c^2$.

For any n and $1 \leq k \leq \lceil 2^n T \rceil$ we have that

$$\begin{aligned}
M_{k2^{-n}}^2 - 2X_{k2^{-n}}^n &= \sum_{j=0}^{k-1} (M_{(j+1)2^{-n}} - M_{j2^{-n}})^2 \\
&= [M]_{k2^{-n}}^n.
\end{aligned}$$

Hence for all n, $M^2 - 2X^n$ is non-decreasing when restricted to times of the form $\{k2^{-n} : 1 \leq k \leq \lceil 2^n T \rceil\}$. To prove the same is also true for $M^2 - 2Y$, it suffices to show that $X^n \to Y$ almost-surely uniformly, at least along a subsequence. This follows from the equivalence of the norms $\|\cdot\|$ and $|||\cdot|||$.

Set $[M]_t = M_t^2 - 2Y_t$. Then $[M]$ is continuous adapted and non-decreasing and $M^2 - [M] = 2Y \in \mathcal{M}_c$. We can extend to all times by applying uniqueness to the above for $T = k$ for each $k \in \mathbb{N}$. Then the process obtained with $T = k, T = k+1$ restricted to $[0, k]$ is the same.

Next we show $[M]^n \to [M]$ UCP as $n \to \infty$. Since $X^n \to Y$ in $(\mathcal{M}_c^2, \|\cdot\|)$, we have $\sup_{0\le t\le T}|X_t^n - Y_t| \to 0$ in L^2 since $\|\cdot\|, |||\cdot|||$ are equivalent. Thus $\sup_{0\le t\le T}|X_t^n - Y_t| \xrightarrow{\mathbb{P}} 0$. Now $[M]_t^n = M^2_{2^{-n}\lceil 2^n t\rceil} - 2X^n_{2^{-n}\lceil 2^n t\rceil}$. Hence

$$\begin{aligned}
&\sup_{0\le t\le T}|[M]_t - [M]_t^n| \\
&\le \sup_{0\le t\le T}|M^2_{2^{-n}\lceil 2^n t\rceil} - M_t^2| + 2\sup_{0\le t\le T}|X^n_{2^{-n}\lceil 2^n t\rceil} - Y_{2^{-n}\lceil 2^n t\rceil}| \\
&\quad + \sup_{0\le t\le T}|Y_{2^{-n}\lceil 2^n t\rceil} - Y_t|
\end{aligned}$$

and each term on the RHS converges to 0 in probability.

Now suppose $M \in \mathcal{M}_{c,\text{loc}}$. For each $n \ge 1$ let $T_n = \inf\{t \ge 0 : |M_t| \ge n\}$. Then (T_n) is a reduces M and M^{T_n} is a bounded martingale for each n. Hence there is a unique continuous adapted non-decreasing process $[M^{T_n}]$ such that $[M^{T_n}]_0$, $(M^{T_n})^2 - [M^{T_n}] \in \mathcal{M}_{c,\text{loc}}$. Let $A^n = [M^{T_n}]$. By uniqueness $(A^{n+1}_{t\wedge T_n})$, (A_t^n) are indistinguishable. Let A be the process such that

$$A_{t\wedge T_n} = A_t^n \ \forall n.$$

Then $M^2_{t\wedge T_n} - A_{t\wedge T_n} \in \mathcal{M}_c$ for all n. Hence $M^2 - A \in \mathcal{M}_{c,\text{loc}}$ with reducing sequence (T_n). So we have $[M] = A$.

Now we show $[M]^n \to [M]$ UCP. We know $[M^{T_k}]^n \to [M^{T_k}]$ UCP for each k. Thus $\mathbb{P}[\sup_{t\in[0,T]}|[M^{T_k}]_t^n - [M^{T_k}]_t| > \varepsilon] \to 0$ for all $\varepsilon > 0, T > 0$. On the event $\{T_k > T\}$, $[M]_t^n = [M^{T_k}]_t^n$ and $[M]_t = [M^{T_k}]_t$ for all $t \le T$. Thus

$$\mathbb{P}[\sup_{t\in[0,T]}|[M]_t^n - [M]_t| > \varepsilon] = \mathbb{P}(T_k \le T) + \mathbb{P}[\sup_{t\in[0,T]}|[M^{T_k}]_t^n - [M^{T_k}]_t| > \varepsilon]$$

which converges to 0 by taking k large enough and then n large enough in the RHS. $\square$

Theorem. *Let $M \in \mathcal{M}_c^2$. Then $M^2 - [M]$ is a UI martingale.*

Proof. Let $T_n = \inf\{t \geq 0 : [M]_t \geq n\}$ for $n \in \mathbb{N}$. Then $T_n \uparrow \infty$ amd T_n is a stopping time with $[M]_{t\wedge T_n} \leq n$. Then

$$|M_{t\wedge T_n}^2 - [M]_{t\wedge T_n}| \leq n + \sup_{u\geq 0} M_u^2.$$

Doob's inequality implies the RHS is integrable and so $M_{t\wedge T_n}^2 - [M]_{t\wedge T_n} \in \mathcal{M}_c$.

The optional stopping theorem implies $\mathbb{E}[M_{t\wedge T_n}^2 - [M]_{t\wedge T_n}] = 0$ and so $\mathbb{E}[[M]_{t\wedge T_n}] = \mathbb{E}[M_{t\wedge T_n}^2]$. Taking $t \to \infty$ the MCT says $\mathbb{E}[[M]_{t\wedge T_n}] \to \mathbb{E}[[M]_{T_n}]$. The DCT says $\mathbb{E}[M_{t\wedge T_n}^2] \to \mathbb{E}[M_{T_n}^2]$. Hence $\mathbb{E}[[M]_{T_n}] = \mathbb{E}[M_{T_n}^2]$. Now take $n \to \infty$ so $\mathbb{E}[[M]_{T_n}] \to \mathbb{E}[[M]_\infty]$ by the MCT and $\mathbb{E}[M_{T_n}^2] \to \mathbb{E}[M_\infty^2]$ by the DCT. So we have $\mathbb{E}[[M]_\infty] = \mathbb{E}[M_\infty^2] < \infty$. Hence $[M]_\infty$ is integrable.

Moreover $|M_t^2 - [M]_t| \leq \sup_{u\geq 0} M_u^2 + [M]_\infty$. The RHS is integrable so $M_t^2 - [M]_t \in \mathcal{M}_c$ and is UI as it's dominated by an integrable random variable. □

The space $L^2(M)$, $M \in \mathcal{M}_c^2$

Recall that the previsible σ-algebra $\mathcal{P}$ is generated by sets of the form $E \times (s, t]$, $E \in \mathcal{F}_s$, $s < t$.

For $A \in \mathcal{P}$ define $\mu(A) = \mathbb{E}\left[\int_0^\infty \mathbb{1}_A(\omega, s)\mathrm{d}[M]_s\right]$. Then μ is a measure on $(\Omega \times (0, \infty), \mathcal{P})$. Moreover it is uniquely determined by

$$\mu(E \times (s, t]) = \mathbb{E}[\mathbb{1}_E([M]_t - [M]_s)]$$

for $s \leq t$¡ $E \in \mathcal{F}_s$. If $H \geq 0$ is previsible then

$$\int_{\Omega\times(0,\infty)} H\mathrm{d}\mu = \mathbb{E}\left[\int_0^\infty H_s \mathrm{d}[M]_s\right].$$

Definition. Let $L^2(M) = L^2(\Omega \times (0, \infty), \mathcal{P}, \mu)$.

Write $\|H\|_{L^2(M)} = \|H\|_M$.

Remark. $(L^2(M), \|\cdot\|_M)$ depends on M since μ depends on M, but simple processes are always in $L^2(M)$.

Itô integrals

Recall that for $H_t = \sum_{k=0}^{n-1} Z_k \mathbb{1}_{(t_k, t_k+1]} \in S$, $M \in \mathcal{M}_c^2$ we set $(H \cdot M)_t = \sum_{k=0}^{n-1} Z_k(M_{t_{k+1}\wedge t} - M_{t_k\wedge t}) \in \mathcal{M}_c^2$. This defines a map $L^2(M) \supseteq S \to \mathcal{M}_c^2$. We will prove that this map is an isometry between $(L^2(M), \|\cdot\|_M)$ and $(\mathcal{M}_c^2, \|\cdot\|)$ when restricted so S (Itô isometry).

Note

$$\|H \cdot M\|^2 = \|(H \cdot M)_\infty\|_{L^2}^2 = \sum_{k=0}^{n-1} \mathbb{E}[Z_k^2(M_{t_{k+1}} - M_{t_k})^2].$$

Since $M^2 - [M]$ is a martingale we have

$$\begin{aligned}
\mathbb{E}[Z_k^2(M_{t_{k+1}} - M_{t_k})^2] &= \mathbb{E}[Z_k^2\mathbb{E}[(M_{t_{k+1}} - M_{t_k})^2|\mathcal{F}_{t_k}]] \\
&= \mathbb{E}[Z_k^2\mathbb{E}[M_{t_{k+1}}^2 - M_{t_k}^2|\mathcal{F}_{t_k}]] \qquad \text{(MG orthogonality)} \\
&= \mathbb{E}[Z_k^2\mathbb{E}[[M]_{t_{k+1}} - [M]_{t_k}|\mathcal{F}_{t_k}]] \\
&= \mathbb{E}[Z_k^2([M]_{t_{k+1}} - [M]_{t_k})]
\end{aligned}$$

which implies

$$\begin{aligned}
\|H \cdot M\|^2 &= \mathbb{E}\left[\sum_{k=0}^{n-1} Z_k^2([M]_{t_{k+1}} - [M]_{t_k})\right] \\
&= \mathbb{E}\left[\int_0^\infty H_s^2 \mathrm{d}[M]_s\right] \\
&= \|H\|_M^2.
\end{aligned}$$

Theorem (Itô isometry)**.** *There exists a unique isometry $I : L^2(M) \to \mathcal{M}_c^2$ such that $I(H) = H \cdot M$ for all $H \in S$.*

Definition. For $M \in \mathcal{M}_c^2$, $H \in L^2(M)$, let $H \cdot M = I(H)$ where I is from the theorem.

To prove the theorem we first prove that the simple processes are dense in $L^2(M)$.

Lemma. *Let ν be a finite measure on $\mathcal{P}$. Then S is dense in $L^2(\mathcal{P}, \nu)$. In particular, if $M \in \mathcal{M}_c^2$ and we take $\nu = \mu$, we have that S is dense in $L^2(M)$.*

Proof. Since $H \in S$ we have $\|H\|_{C^\infty} < \infty$ and so $S \subseteq L^2(\mathcal{P}, \nu)$. Let $\overline{S}$ be the closure of S in $L^2(\mathcal{P}, \nu)$.

Let $\mathcal{A} = \{A \in \mathcal{P} : \mathbb{1}_A \in \overline{S}\}$. We claim $\mathcal{A} = \mathcal{P}$. Clearly $\mathcal{A} \subseteq \mathcal{P}$. For the other direction note that $\mathcal{A}$ contains the π-system $\{E \times (s,t] : E \in \mathcal{F}_s, s < t\}$ which generates $\mathcal{P}$. Also $\mathcal{A}$ is a d-system so $\mathcal{A} = \mathcal{P}$ by Dynkin's lemma. Therefore $\overline{S}$ contains all indicators of sets in $\mathcal{P}$.

The result now follows from the density of simple functions in $L^2(\mathcal{P}, \nu)$. $\square$

Proof of Itô isometry. Take $H \in L^2(M)$. The above lemma implies that there exists a sequence (H^n) in S with $\|H^n - H\|_M \to 0$. i.e

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

Thus (H^n) is Cauchy with respect to $\|\cdot\|_M$. We want to show $(I(H^n))$ is Cauchy with respect to $\|\cdot\|$. Indeed, we have

$$\begin{aligned}
\|I(H^n) - I(H^m)\| &= \|H^n \cdot M - H^m \cdot M\| \\
&= \|(H^n - H^m) \cdot M\| && \text{(linearity)} \\
&= \|H^n - H^m\|_M. && \text{(isometry property)}
\end{aligned}$$

Therefore $(I(H^n))$ converges with respect to $\|\cdot\|$ to an element of $\mathcal{M}_c^2$. We define $I(H)$ to be this limit. Now we just check that I is well-defined. Suppose (K^n) in S converges to H with respect to $\|\cdot\|_M$. Then $\|I(H^n) - I(K^n)\| = \|H^n - K^n\|_M$ as before so $I(H^n)$ and $I(K^n)$ have the same limit (up to indistinguishability).

Finally we show I is an isometry $L^2(M) \to \mathcal{M}_c^2$. For $H \in L^2(M)$ let (H^n) be a sequence in S converging to H in $L^2(M)$. Then

$$\|I(H)\| = \lim_{n\to\infty} \|H^n \cdot M\| = \lim_{n\to\infty} \|H^n\|_M = \|H\|_M.$$

$\square$

We write $I(H)_t = (H \cdot M)_t = \int_0^t H_s \mathrm{d}M_s$. This process $H \cdot M$ is the *Itô* (or *stochastic*) *integral of* H with respect to M.
Extensions: our goal now is to extend the definition of $H \cdot M$ to the setting that H is locally bounded and $M \in \mathcal{M}_{c,\text{loc}}$. We need to understand how this integral behaves under stopping.

Proposition. Let $H \in S$, $M \in \mathcal{M}$. Then for any stopping time T we have that $H \cdot (M^T) = (H \cdot M)^T$.

Proof. We have that

$$\begin{aligned}(H \cdot M^T)_t &= \sum_{k=0}^{n-1} Z_k (M^T_{t_{k+1}\wedge t} - M^T_{t_k \wedge t}) \\ &= \sum_{k=0}^{n-1} Z_k (M_{t_{k+1}\wedge t \wedge T} - M_{t_k \wedge t \wedge T}) \\ &= (H \cdot M)_{t\wedge T} = (H \cdot M)^T_t.\end{aligned}$$

□

Proposition. Let $H \in L^2(M)$, $M \in \mathcal{M}^2_c$, T a stopping time. Then $(H \cdot M)^T = (H\mathbb{1}_{(0,T]}) \cdot M = H \cdot (M^T)$.

Proof. First note that if $H \in L^2(M)$ then $H\mathbb{1}_{(0,T]} \in L^2(M)$ and $H \in L^2(M^T)$ so the integrals make sense.

Suppose $H \in S$, $M \in \mathcal{M}^2_c$ and T takes finitely many values. Then $H\mathbb{1}_{(0,T]} \in S$ and $(H \cdot M)^T = (H\mathbb{1}_{(0,T]}) \cdot M = H \cdot M^T$.

Now suppose $H \in S$, $M \in \mathcal{M}^2_c$ and T is a general stopping time. The previous proposition implies $(H \cdot M)^T = H \cdot M^T$. So we show $(H \cdot M)^T = (H\mathbb{1}_{(0,T]}) \cdot M$. We will prove this via an approximation argument. For $n, m \in \mathbb{N}$ let $T_{n,m} = (2^{-n}\lceil 2^n T \rceil) \wedge m$. Then $T_{n,m}$ takes on finitely many values and $T_{n,m} \downarrow T \wedge m$ as $n \to \infty$. Thus

$$\|H\mathbb{1}_{(0,T_{n,m}]} - H\mathbb{1}_{(0,T_{n,m}]}\|^2_M = \mathbb{E}\left[\int_0^\infty H_t^2 \mathbb{1}_{(T\wedge m, T_{n,m}]} \mathrm{d}[M]_t\right] \to 0$$

by the DCT with dominating function H_t^2. Hence $(H\mathbb{1}_{(0,T_{n,m}]})\cdot M \to (H\mathbb{1}_{(0,T\wedge m]})\cdot M$ in $\mathcal{M}^2_c$ as $n \to \infty$. Now we know the LHS is $(H \cdot M)^{T_{n,m}} \to (H \cdot M)^{T\wedge m}$ pointwise almost-surely by continuity of $H \cdot M$, so $(H\mathbb{1}_{(0,T\wedge m]})\cdot M = (H \cdot M)^{T\wedge m}$. □