# 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 $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$. Suppose that at the $m$th 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 $b,\sigma$ 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 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 satisfying 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

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

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 $\sigma$-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 $\sigma$-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. $\square$

**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. $\square$

**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

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

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 $\Pi$ is a $\pi$-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 $|X| \le Y$ $\forall X \in \mathcal{X}$.

**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 $t \ge 0$ the family

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

is UI.

*Proof.* First suppose $X$ is a martingale. By the Optional Stopping Theorem, if $T \le t$ 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 $t \ge 0$. 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 $(T_n)_{n\ge 0}$ be a reducing sequence for $X$ and let $T \le t$ 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 $\{X_{T\wedge T_n} : n \ge 0\}$ 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 $|X_t| \le Y$ for all $t \ge 0$, 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 $\sup_{0 \leq s \leq t} |X_s(\omega)| < \infty$ so there exists $n(\omega, t) \in \mathbb{N}$ such that $n(\omega, t) \geq \sup_{0 \leq s \leq t} |X_s(\omega)|$. Then if $n \geq n(\omega, t)$ we have $T_n(\omega) \geq t$. 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.
$\square$

# 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 $X_\infty$ 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 $X_\infty$ 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

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

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

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

$\square$

## 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, $\mathcal{M}_c^2 = \mathcal{M}_c \cap \mathcal{M}^2$ 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 $(\mathcal{M}_c^2, \|\cdot\|)$ 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

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

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).

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

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

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

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

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

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

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 $1 \le k \le \lceil 2^n T \rceil$ 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 $\{k2^{-n} : 1 \le k \le \lceil 2^n T \rceil\}$. 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 $[M]_t^n = M_{2^{-n}\lceil 2^n t\rceil}^2 - 2X_{2^{-n}\lceil 2^n t\rceil}^n$. Hence

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

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 $(A_{t\wedge T_n}^{n+1})$, $(A_t^n)$ are indistinguishable. Let $A$ be the process such that

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

Then $M_{t\wedge T_n}^2 - A_{t\wedge T_n} \in \mathcal{M}_c$ 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 $\mathbb{P}[\sup_{t\in[0,T]} |[M^{T_k}]^n_t - [M^{T_k}]_t| > \varepsilon] \to 0$ for all $\varepsilon > 0, T > 0$. On the event $\{T_k > T\}$, $[M]^n_t = [M^{T_k}]^n_t$ and $[M]_t = [M^{T_k}]_t$ for all $t \leq T$. Thus

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

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 $\sigma$-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 $\mu$ 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 $\mu$ 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 $\nu$ 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 $\pi$-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 $L^2(\mathcal{P}, \nu)$. □

*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

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

□

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 $n \to \infty$ by the Itô isometry. Now we know $(H\mathbb{1}_{(0,T_{n,m}]}) \cdot M = (H \cdot M)^{T_{n,m}} \to (H \cdot M)^{T \wedge m}$ pointwise almost-surely by continuity of $H \cdot M$, so $(H\mathbb{1}_{(0,T\wedge m]}) \cdot M = (H \cdot M)^{T\wedge m}$. Repeating the same argument sending $m \to \infty$ we get $(H\mathbb{1}_{(0,T]}) \cdot M = (H \cdot M)^T$.

Now suppose $H \in L^2(M)$, $M \in \mathcal{M}_c^2$, $T$ a general stopping time. Let $(H^n)$ be a sequence in $S$ with $H^n \to H$ in $L^2(M)$. Then

$$\begin{aligned}
\|(H^n \cdot M)^T - (H \cdot M)^T\| &= \|(H^n \cdot M)_T - (H \cdot M)_T\|_{L^2} \\
&\leq \left\| \sup_{t \geq 0} |(H^n \cdot M)_t - (H \cdot M)_t| \right\|_{L^2} \\
&\leq 2\|(H^n \cdot M)_\infty - (H \cdot M)_\infty\|_{L^2} \\
&\qquad\qquad\qquad\qquad \text{(Doob's } L^2\text{-inequality)} \\
&= 2\|(H^n - H) \cdot M\| \\
&= 2\|H^n - H\|_M \qquad\qquad \text{(Itô isometry)} \\
&\xrightarrow{n \to \infty} 0.
\end{aligned}$$

Hence $(H^n \cdot M)^T \to (H \cdot M)^T$ in $\mathcal{M}_c^2$. On the other hand,

$$\begin{aligned}
\|H^n \mathbb{1}_{(0,T]} - H \mathbb{1}_{(0,T]}\|_M^2 &= \mathbb{E}\left[\int_0^T (H_t^n - H_t)^2 \mathrm{d}[M]_t\right] \\
&\leq \mathbb{E}\left[\int_0^\infty (H_t^n - H_t)^2 \mathrm{d}[M]_t\right] \\
&= \|H^n - H\|_M^2 \\
&\xrightarrow{n \to \infty} 0
\end{aligned}$$

so $H^n \mathbb{1}_{(0,T]} \cdot M \to H \mathbb{1}_{(0,T]} \cdot M$ in $\mathcal{M}_c^2$ by the Itô isometry. Since $(H^n \cdot M)^T = H^n \mathbb{1}_{(0,T]} \cdot M$ for all $n$ we have $(H \cdot M)^T = H \mathbb{1}_{(0,T]} \cdot M$. Next note

$$\begin{aligned}
\|H^n - H\|_{M^T}^2 &= \mathbb{E}\left[\int_0^\infty (H_s^n - H_s)^2 \mathrm{d}[M^T]_s\right] \\
&= \mathbb{E}\left[\int_0^T (H_s^n - H_s)^2 \mathrm{d}[M]_s\right] \\
&\leq \|H^n - H\|_M^2 \\
&\xrightarrow{n \to \infty} 0
\end{aligned}$$

and therefore $H^n \cdot M^T \to H \cdot M^T$ in $\mathcal{M}_c^2$ by the Itô isometry. Putting this together gives the result that $(H \cdot M)^T = H \cdot M^T$. $\square$

**Definition.** We say that a previsible process $H$ is *locally bounded* if there exists a sequence $(S_n)$ of stopping times where $S_n \uparrow \infty$ and $H\mathbb{1}_{(0,S_n]}$ is bounded for all $n$.

**Remark.** Every continuous adapted process is previsible and locally bounded.

**Definition.** Let $H$ be a locally bounded previsible process with $H\mathbb{1}_{(0,S_n]}$ bounded for all $n$ where $(S_n)$ is a sequence of stopping times with $S_n \uparrow \infty$. Let $M \in \mathcal{M}_{c,\text{loc}}$ have $M_0 = 0$ and let $S_n' = \inf\{t \geq 0 : |M_t| \geq n\}$ so that $M^{S_n'} \in \mathcal{M}_c^2$ for all $n$. Let $T_n = S_n \wedge S_n'$ and set

$$(H \cdot M)_t = ((H\mathbb{1}_{(0,T_n]}) \cdot M^{T_n})_t \ \forall t \in [0, T_n].$$

By the previous proposition, this definition is well-defined and consistent with the Itô integral with $M \in \mathcal{M}_c^2$, $H \in L^2(M)$.

**Proposition.** Let $M \in \mathcal{M}_{c,\text{loc}}$, $H$ locally bounded and previsible. Then $H \cdot M \in \mathcal{M}_{c,\text{loc}}$ where the sequence $(T_n)$ is a reducing sequence. Moreover, for any stopping time $T$ we have that

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

*Proof.* The fact $H \cdot M \in \mathcal{M}_{c,\text{loc}}$ with reducing sequence $(T_n)$ follows from the definition of $H \cdot M$. For any stopping time $T$,

$$\begin{aligned}
(H \cdot M)^T &= \lim_{n\to\infty} (H\mathbb{1}_{(0,T_n]} \cdot M^{T_n})^T \text{ pointwise} \\
&= \lim_{n\to\infty} (H\mathbb{1}_{(0,T_n]}\mathbb{1}_{(0,T]} \cdot M^{T_n}) && \text{(previous proposition)} \\
&= H\mathbb{1}_{(0,T]} \cdot M \\
&= \lim_{n\to\infty} (H \cdot (M^{T_n})^T) \\
&= \lim_{n\to\infty} (H \cdot M^{T_n \wedge T}) \\
&= H \cdot M^T.
\end{aligned}$$

□

**Proposition.** Let $M \in \mathcal{M}_{c,\text{loc}}$ and $H$ locally bounded and previsible. Then $[H \cdot M] = H^2 \cdot [M]$.

**Remark.** $H \cdot M$ is an Itô integral, while $H^2 \cdot [M]$ is a Lebesgue-Stieljes integral.

*Proof.* Suppose that $T$ is a bounded stopping time, $H, M$ are uniformly bounded. Then

$$\begin{aligned}
\mathbb{E}[(H \cdot M)_T^2] &= \mathbb{E}[(H\mathbb{1}_{(0,T]} \cdot M)_\infty^2] \\
&= \mathbb{E}[((H^2\mathbb{1}_{(0,T]}) \cdot [M])_\infty] && \text{(Itô isometry)} \\
&= \mathbb{E}[(H^2 \cdot [M])_T]
\end{aligned}$$

the OST then implies $(H \cdot M)^2 - H^2 \cdot [M]$ is a continuous martingale. Hence uniqueness of quadratic variation implies $[H \cdot M] = H^2 \cdot [M]$.

Now assume $H$ is locally bounded and previsible and that $M$ is a continuous local martingale. Let $(T_n)$ be a sequence of stopping times so that $H\mathbb{1}_{(0,T_n]}$ is bounded and $M^{T_n}$ is bounded and $T_n \uparrow \infty$. Then

$$\begin{aligned}
[H \cdot M] &= \lim_{n\to\infty} [H \cdot M]^{T_n} \\
&= \lim_{n\to\infty} [(H \cdot M)^{T_n}] && \text{(uniqueness of quadratic variation)} \\
&= \lim_{n\to\infty} [H\mathbb{1}_{(0,T_n]} \cdot M^{T_n}] \\
&= \lim_{n\to\infty} H^2\mathbb{1}_{(0,T_n]} \cdot [M^{T_n}] \\
&= H^2 \cdot [M]. && \text{(MCT)}
\end{aligned}$$

□

Since $H \cdot M \in \mathcal{M}_{c,\text{loc}}$ for $M \in \mathcal{M}_{c,\text{loc}}$, $H$ locally bounded and previsible, we may integrate against it.

**Proposition.** Let $M \in \mathcal{M}_{c,\text{loc}}$, $H, K$ locally bounded and previsible. Then

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

*Proof.* Elementary to check that this holds for $H, K \in S$. Now suppose that $H, K, M$ are uniformly bounded. We need to show $H \in L^2(K \cdot M)$ and $HK \in L^2(M)$. We have

$$\begin{aligned}
\|H\|_{L^2(K\cdot M)}^2 &= \mathbb{E}\left[(H^2 \cdot [K \cdot M])_\infty\right] \\
&= \mathbb{E}[(H^2 \cdot (K^2 \cdot [M]))_\infty] && \text{(previous proposition)} \\
&= \mathbb{E}[((HK)^2 \cdot [M])_\infty] && \text{(Lebesgue-Stieljes)} \\
&= \|HK\|_{L^2(M)}^2 && (*) \\
&\le \min\{\|H\|_{L^\infty}^2 \|K\|_{L^2(M)}^2, \|K\|_{L^\infty}^2 \|H\|_{L^2(M)}^2\} \\
&< \infty.
\end{aligned}$$

Let $(H^n), (K^n)$ be sequences in $S$ which converge to $H, K$ respectively in $L^2(M)$ and assume WLOG that $(H^n), (K^n)$ are uniformly bounded. Then

$$H^n \cdot (K^n \cdot M) = (H^n K^n) \cdot M.$$

Note

$$\begin{aligned}
&\|H^n \cdot (K^n \cdot M) - H \cdot (K \cdot M)\| \\
&\le \|(H^n - H) \cdot (K^n \cdot M)\| + \|H \cdot ((K^n - K) \cdot M)\| \\
&= \|H^n - H\|_{L^2(K^n \cdot M)} + \|H\|_{L^2((K^n - K)\cdot M)} && \text{(Itô isometry)} \\
&= \|(H^n - H)K^n\|_{L^2(M)} + \|H(K^n - K)\|_{L^2(M)} && \text{(by } (*)) \\
&\le \|K^n\|_{L^\infty} \|H^n - H\|_{L^2(M)} + \|H\|_{L^\infty} \|K^n - K\|_{L^2(M)} && \text{(by } (*) \text{ again)} \\
&\xrightarrow{n\to\infty} 0.
\end{aligned}$$

A similar argument shows that $(H^n K^n) \cdot M \to (HK) \cdot M$ in $\mathcal{M}_c^2$. Hence $H \cdot (K \cdot M) = (HK) \cdot M$.

Now suppose that $H, K$ are locally bounded, previsible and $M \in \mathcal{M}_{c,\text{loc}}$. Let $(T_n)$ be a sequence of stopping times so that $H\mathbb{1}_{(0,T_n]}, K\mathbb{1}_{(0,T_n]}, M^{T_n}$ are bounded and $T_n \uparrow \infty$. Then

$$H\mathbb{1}_{(0,T_n]} \cdot (K\mathbb{1}_{(0,T_n]} \cdot M^{T_n}) = HK\mathbb{1}_{(0,T_n]} \cdot M^{T_n}.$$

Also

$$K\mathbb{1}_{(0,T_n]} \cdot M^{T_n} = (K \cdot M)^{T_n}.$$

Hence

$$\begin{aligned}
H\mathbb{1}_{(0,T_n]} \cdot (K\mathbb{1}_{(0,T_n]} \cdot M^{T_n}) &= H\mathbb{1}_{(0,T_n]} \cdot (K \cdot M)^{T_n} && (1) \\
&= (H \cdot (K \cdot M))^{T_n} && (2) \\
&\xrightarrow{n\to\infty} H \cdot (K \cdot M). && (3)
\end{aligned}$$

Also

$$\begin{aligned}
HK\mathbb{1}_{(0,T_n]} \cdot M^{T_n} &= ((HK) \cdot M)^{T_n} \\
&\xrightarrow{n\to\infty} (HK) \cdot M.
\end{aligned}$$

Therefore $H \cdot (K \cdot M) = (HK) \cdot M$. □

**Remark.** We have repeatedly been using a "localisation" argument to reduce everything to the setting of a bounded integrand and bounded martingale. This is a standard procedure so details will me omitted from now on.

### Semimartingales

**Definition.** A continuous adapted process $X$ is called a *semimartingale* if it can be written as

$$X = X_0 + M + A$$

for $M \in \mathcal{M}_{c,\text{loc}}$, $A$ of finite variation and where $M_0 = A_0 = 0$. This decomposition is called the "Doob-Meyer decomposition" of the semimartingale.

For a continuous semimartingale $X = X_0 + M + A$, define its quadratic variation by $[X] = [M]$. This is justified since

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

UCP as $n \to \infty$ [see Example Sheet 1].

**Definition.** For $H$ locally bounded and previsible, $X = X_0 + M + A$ a continuous semimartingale, define

$$H \cdot X = \underbrace{H \cdot M}_{\text{Itô}} + \underbrace{H \cdot A}_{\text{Lebesgue-Stieltjes}} .$$

Then $H \cdot X$ is a semimartingale.

**Proposition.** Let $X$ be a continuous semimartingale and $H$ locally bounded, left-continuous and adapted. Then

$$\sum_{k=0}^{\lceil 2^n t \rceil - 1} H_{k2^{-n}}(X_{(k+1)2^{-n}} - X_{k2^{-n}}) \xrightarrow{n\to\infty} (H \cdot X)_t$$

UCP.

### Summary of the Stochastic Integral

Step 1: for $H \in S$, $H_t \in \sum_{k=0}^{n-1} Z_k \mathbb{1}_{(t_k,t_{k+1}]}(t)$, $Z_k$ bounded and $\mathcal{F}_{t_k}$-measurable, $M \in \mathcal{M}_c^2$ we set

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

Then $H \cdot M \in \mathcal{M}_c^2$.

Step 2: equip $\mathcal{M}_c^2$ with a Hilbert space structure with norm $\|M\| = \|M_\infty\|_{L^2}$.

Step 3 establish the existence of $[M]$, $M \in \mathcal{M}_{c,\text{loc}}$, where $[M]$ is the unique adapted non-decreasing continuous process with $[M]_0 = 0$ such that $M^2 - [M] \in \mathcal{M}_{c,\text{loc}}$.

Step 4: for $M \in \mathcal{M}_c^2$, used $[M]$ to define a Hilbert space $(L^2(M), \|\cdot\|_M)$ where $\|H\|_M = \left(\mathbb{E}\int_0^\infty H_s^2 \mathrm{d}[M]_s\right)^{1/2}$.

Step 5: extend the integral to $H \in L^2(M)$, $M \in \mathcal{M}_c^2$ using the Itô isometry: $\|H \cdot M\| = \|H\|_M$.

Step 6: extended to $H$ locally bounded and previsible, and $M \in \mathcal{M}_{c,\text{loc}}$ by setting $(H \cdot M)_t = (H\mathbb{1}_{(0,T_n]} \cdot M^{T_n})_t$ for all $t \leq T_n$. Then $H \cdot M \in \mathcal{M}_{c,\text{loc}}$.

Step 7: extend to $H$ locally bounded and previsible and $X = X_0 + M + A$ a continuous semimartingale by setting

$$H \cdot X = \underbrace{H \cdot M}_{\text{Itô}} + \underbrace{H \cdot A}_{Lebesgue-Stieltjes}.$$

Then $H \cdot X$ is a continuous semimartingale.

## Stochastic Calculus

**Definition.** For $M, N \in \mathcal{M}_{c,\text{loc}}$ define the *covariation* of $M, N$ by setting

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

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

**Theorem.** *Let $M, N \in \mathcal{M}_{c,loc}$. Then*

*(a) $[M,N]$ is the unique (up to indistinguishability) continuous, adapted, finite-variation process with $[M,N]_0 = 0$ so that $MN - [M,N] \in \mathcal{M}_{c,loc}$;*

*(b) For $n \geq 1$ set $[M,N]_t^n = \sum_{k=0}^{\lceil 2^n t\rceil -1} (M_{(k+1)2^{-n}} - M_{k2^{-n}})(N_{(k+1)2^{-n}} - N_{k2^{-n}})$. Then $[M,N]^n \to [M,N]$ UCP;*

*(c) If $M, N \in \mathcal{M}_c^2$, then $MN - [M,N]$ is a UI martingale;*

*(d) For $H$ locally bounded, previsible we have $[H \cdot M, N] + [M, H \cdot N] = 2H \cdot [M,N]$.*

*Proof.*

(a) Note $MN = \frac{1}{4}((M+N)^2 - (M-N)^2)$ so $MN - [M,N]$ is the sum of continuous local martingales, thus a local martingale. By definition $[M,N]$ is continuous, adapted and has finite variation. The same argument used for quadratic variation shows uniqueness;

(b) Note $[M,N]_t^n = \frac{1}{4}([M+N]_t^n - [M-N]_t^n)$;

(c) Follows from the corresponding result for quadratic variation;

(d) We have $[H \cdot (M \pm N)] = H^2 \cdot [M \pm N]$. Thus $[H \cdot M, H \cdot N] = H^2 \cdot [M,N]$. Moreover

$$\begin{aligned}(H+1)^2 \cdot [M,N] &= [(H+1) \cdot M, (H+1) \cdot N] \\ &= [H \cdot M + M, H \cdot N + N] \\ &= [H \cdot M, H \cdot N] + [M, H \cdot N] + [H \cdot M, N] + [M,N]\end{aligned}$$

since the covariation is bilinear (Example Sheet 2). Also $(H+1)^2 \cdot [M,N] = H^2 \cdot [M,N] + 2H \cdot [M,N] + [M,N]$, giving the result.

□

**Proposition** (Kunita-Watanabe identity)**.** Let $M, N \in \mathcal{M}_{c,\text{loc}}$, $H$ locally bounded and previsible. Then $[H \cdot M, N] = H \cdot [M,N]$.

*Proof.* We need to show $[H \cdot M, N] = [M, H \cdot N]$ as then we can apply part (d) of the previous theorem. We have that $(H \cdot M)N - [H \cdot M, N] \in \mathcal{M}_{c,\text{loc}}$. Also $M(H \cdot N) - [M, H \cdot N] \in \mathcal{M}_{c,\text{loc}}$. We will show that $(H \cdot M)N - M(H \cdot N) \in \mathcal{M}_{c,\text{loc}}$, which will give the theorem as this will have finite variation and starts from 0, meaning it must be identically 0.

By a localisation argument we may assume $M, N \in \mathcal{M}_c^2$ and $H$ is bounded. By the OST it suffices to show that for bounded stopping times $T$ we have

$$\mathbb{E}[(H \cdot M)_T N_T] = \mathbb{E}[M_T (H \cdot N)_T].$$

Indeed

$$\mathbb{E}[(H \cdot M)_T N_T] = \mathbb{E}[(H \cdot M)_\infty^T N_\infty^T] \text{ and } \mathbb{E}[M_T (H \cdot N)_T] = \mathbb{E}[M_\infty^T (H \cdot N)_\infty^T]$$

so it suffices to show

$$\mathbb{E}[(H \cdot M)_\infty N_\infty] = \mathbb{E}[M_\infty (H \cdot N)_\infty] \tag{$*$}$$

for all $M, N \in \mathcal{M}_c^2$ and $H$ bounded. Suppose that $H = Z\mathbb{1}_{(s,t]}$ for $Z$ a $\mathcal{F}_s$-measurable bounded random variable. Then

$$\begin{aligned}\mathbb{E}[(H \cdot M)_\infty N_\infty] = \mathbb{E}[Z(M_t - M_s)N_\infty] &= \mathbb{E}[ZM_t\mathbb{E}[N_\infty|\mathcal{F}_t] - ZM_s\mathbb{E}[N_\infty|\mathcal{F}_s]] \\ &= \mathbb{E}[Z(M_tN_t - M_sN_s)]\end{aligned}$$

which is symmetric, implying it is equal to $\mathbb{E}[M_\infty(H \cdot N)_\infty]$. Hence $(*)$ holds for $H = Z\mathbb{1}_{(s,t]}$. Also $(*)$ is linear in $H$ so is true for $H \in S$.

Suppose $H$ is bounded and previsible. Then we can find a sequence $(H^n)$ of simple processes such that $H^n \to H$ in $L^2(M)$ and $L^2(N)$ [in the lemma we showed $S$ is dense in $L^2(\mathcal{P}, \nu)$ for any finite $\nu$, so take $\nu$ given by $\nu(E) = \mathbb{E}[\int_0^\infty \mathbb{1}_E(\mathrm{d}[M]_s + \mathrm{d}[N]_s)]]$. Then $H^n \cdot M \to H \cdot M$ and $H^n \cdot N \to H \cdot N$ with respect to the $\|\cdot\|$ norm. Hence $(H^n \cdot M)_\infty \to (H \cdot M)_\infty$ and $(H^n \cdot N)_\infty \to (H \cdot N)_\infty$ in $L^2$.

Therefore

$$\|((H^n \cdot M)_\infty (H \cdot M)_\infty)N_\infty\|_{L^1}^2 \leq \|(H^n \cdot M)_\infty - (H \cdot M)_\infty\|_{L^2}\|N_\infty\|_{L^2} \to 0.$$

Hence $\mathbb{E}[(H^n \cdot M)_\infty N_\infty] \to \mathbb{E}[(H \cdot M)_\infty N_\infty]$ and similarly $\mathbb{E}[(H^n \cdot N)_\infty M_\infty] \to \mathbb{E}[(H \cdot N)_\infty M_\infty]$. $\square$

**Definition.** For continuous semimartingales $X, Y$, define $[X, Y]$ to be the co-variation of their martingale parts.

**Remark.**

- This definition is justified as

$$[X,Y]_t^n = \sum_{k=0}^{\lceil 2^n t\rceil -1} (X_{(k+1)2^{-n}} - X_{k2^{-n}})(Y_{(k+1)2^{-n}} - Y_{k2^{-n}}) \to [X,Y]_t$$

  UCP.

- The Kunita-Watanabe identity also holds for semimartingales.

**Proposition.** Let $X, Y$ be independent semimartingales. Then $[X, Y] = 0$.

*Proof.* Example Sheet 2. □

## Itô's formula

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

$$X_tY_t - X_0Y_0 = \int_0^t X_s \mathrm{d}Y_s + \int_0^t Y_s \mathrm{d}X_s + [X,Y]_t. \tag{$*$}$$

*Proof.* Note that the integrals are well-defined since any continuous adapted process is locally bounded and previsible.

Note that for $s \le t$, we have that

$$X_tY_t - X_sY_s = X_s(Y_t - Y_s) + Y_s(X_t - X_s) + (X_t - X_s)(Y_t - Y_s).$$

Since the LHS and RHS of $(*)$ are continuous, it suffices to show the result for $t$ of the form $t = m2^{-n}$, $m, n \in \mathbb{N}$. Then

$$\begin{aligned} X_tY_t - X_0Y_0 &= \sum_{k=0}^{m2^{j-n}-1} \Big[ (X_{k2^{-n}}(Y_{(k+1)2^{-n}} - Y_{k2^{-n}})) + (Y_{k2^{-n}}(X_{(k+1)2^{-n}} - X_{k2^{-n}})) \\ &\qquad + (X_{(k+1)2^{-n}} - X_{k2^{-n}})(Y_{(k+1)2^{-n}} - Y_{k2^{-n}}) \Big] \\ &\xrightarrow{UCP} (X \cdot Y)_t + (Y \cdot X)_t + [X,Y]_t. \end{aligned}$$

□

Note that the $[X, Y]$ term does not appear if either $X, Y$ or independent or one of $X, Y$ have no martingale part.

**Theorem** (Itô's formula)**.** *Let $X = (X^1, \dots, X^d)$ where each $X^i$ is a continuous semimartingale. Let $f : \mathbb{R}^d \to \mathbb{R}$ be $C^2$. Then*

$$f(X_t) = f(X_0) + \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(X_s)\mathrm{d}X_s^i + \frac{1}{2}\sum_{i,j=1}^{d} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)\mathrm{d}[X^i, X^j]_s.$$

**Remark.**

1. Integration by parts is a special case of Itô's formula with $f(x, y) = xy$;
2. For $d = 1$ Itô's formula is

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)\mathrm{d}[X]_s.$$

It is possible to derive Itô's formula using Taylor expansions since

$$\begin{aligned}
&f(X_t) \\
&= f(X_0) + \sum_{k=0}^{2^{-n}\lceil 2^n t\rceil - 1} \left(f(X_{(k+1)2^{-n}}) - f(X_{k2^{-n}})\right) + \left(f(X_t) - f(X_{2^{-n}\lceil 2^n t\rceil - 1})\right) \\
&= f(X_0) + \sum_{k=0}^{\lceil 2^n t\rceil - 1} f'(X_{k2^{-n}})(X_{(k+1)2^{-n}} - X_{k2^{-n}}) \\
&\qquad + \frac{1}{2}\sum_{k=0}^{\lceil 2^n t\rceil - 1} f''(X_{k2^{-n}})(X_{(k+1)2^{-n}} - X_{k2^{-n}})^2 + \text{error} \\
&\xrightarrow{UCP} f(X_0) + \int_0^t f'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)\mathrm{d}[X]_s.
\end{aligned}$$

We will prove Itô's formula a different way since the error term is inconvenient to deal with.

**Examples.**

1. $X = B$, $B$ a standard Brownian motion, $f(X) = X^2$. Then Itô's formula gives

$$f(B_t) = f(B_0) + \int_0^t f'(B_s)\mathrm{d}B_s + \frac{1}{2}\int_0^t f''(B_s)\mathrm{d}[B]_s \tag{4}$$

$$= \int_0^t 2B_s\mathrm{d}B_s + \frac{1}{2}\int_0^t 2\mathrm{d}s \tag{5}$$

$$= 2\int_0^t B_s\mathrm{d}B_s + t \tag{6}$$

which implies $B_t^2 - t = 2\int_0^t B_s\mathrm{d}B_s \in \mathcal{M}_{c,\mathrm{loc}}$.

2. Let $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ be $C^{1,2}$ and $X_t = (t, B_t^1, \dots, B_t^d)$ where $B^1, \dots, B^d$ are independent Brownian motions. By Itô's formula we have

$$f(t, B_t) - f(0, B_0) - \int_0^t (\frac{\partial}{\partial t} + \Delta) f(s, B_s) \mathrm{d}s = \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x_i} f(s, B_s) \mathrm{d}B_s^i \in \mathcal{M}_{c,\mathrm{loc}}.$$

If $f$ does not depend on $t$ and is harmonic in the spatial variables then $f(t, B_t) \in \mathcal{M}_{c,\mathrm{loc}}$. In furthermore $f$ is bounded then $f(B_t)$ is a martingale.

*Proof of Itô's formula.* We will prove the $d = 1$ case; the $d > 1$ case is an exercise. Let $X = X_0 + M + A$ and let $V$ be the total variation process of $A$. Let $T_r = \inf\{t \geq 0 : |X_t| + |V_t| + [M]_t > r\}$ for each $r > 0$. Then $(T_r)$ is a sequence of stopping times with $T_r \uparrow \infty$ as $r \to \infty$. It suffices to prove the formula in $[0, T_r]$ for each $r > 0$. Let $\mathcal{A}$ be the subset of $C^2(\mathbb{R})$ so that Itô's formula holds. We want to show $\mathcal{A} = C^2(\mathbb{R})$. We will do this by showing:

(a) $\mathcal{A}$ contains $f(x) = 1$ and $f(x) = x$;

(b) $\mathcal{A}$ is a vector space;

(c) $\mathcal{A}$ is an algebra, i.e if $f, g \in \mathcal{A}$ then $fg \in \mathcal{A}$;

(d) if $(f_n)$ is a sequence in $\mathcal{A}$ with $f_n \to f$ in $C^2(B(0,r))$ for each $r > 0$, then $f \in \mathcal{A}$. By $f_n \to f$ in $C^2(B(0,r))$ we mean

$$\Delta_{n,r} := \sup_{x \in B(0,r)} |f_n - f| + \sup_{x \in B(0,r)} |f_n' - f'| + \sup_{x \in B(0,r)} |f_n'' - f''| \xrightarrow{n \to \infty} 0.$$

From this, (a),(b),(c) imply all polynomials are in $\mathcal{A}$. Then by the Weierstrass approximation theorem, polynomials are dense in $C^2(B(0,r))$ for each $r > 0$, so (d) will give that $\mathcal{A} = C^2(\mathbb{R})$.

It is easy to see (a),(b) hold. So we prove (c). Suppose $f, g \in \mathcal{A}$ and define $F_t = f(X_t)$, $G_t = g(X_t)$. Then Itô's formula holds for $f, g$ so $F, G$ are continuous semimartingales. Integration by parts gives

$$F_t G_t - F_0 G_0 = \int_0^t F_s \mathrm{d}G_s + \int_0^t G_s \mathrm{d}F_s + \frac{1}{2}[F, G]_t. \tag{1}$$

Since Itô's formula holds for $g$,

$$\begin{aligned} \int_0^t F_s \mathrm{d}G_s &= \int_0^t F_s \mathrm{d}\left(\int_0^s g''(X_u)\mathrm{d}X_u + \frac{1}{2}\int_0^s g''(X_u)\mathrm{d}[X]_u\right) \\ &= \int_0^t f(X_s)g'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f(X_s)g''(X_s)\mathrm{d}[X]_s \end{aligned} \tag{2}$$

where we used $H \cdot (K \cdot M) = (HK) \cdot M$. Also

$$\int_0^t G_s \mathrm{d}F_s = \int_0^t f'(X_s)g(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)g(X_s)\mathrm{d}[X]_s \tag{3}$$

and

$$\begin{aligned} [F, G]_t &= [f(X), g(X)]_t \\ &= [f'(X) \cdot X, g'(X) \cdot X]_t && \text{(Itô's formula)} \\ &= \int_0^t f'(X_s)g'(X_s)\mathrm{d}[X]_s && (4) \end{aligned}$$

by Kunita-Watanabe. Plug (2)-(4) into (1) to get Itô's formula for $fg$, i.e $fg \in \mathcal{A}$.

Now we prove (d). Suppose $(f_n)$ is a sequence in $\mathcal{A}$ with $f_n \to f$ in $C^2(B(0,r))$ for each $r > 0$. Since $f_n \in \mathcal{A}$ we have

$$\begin{aligned} f_n(X_t) &= f_n(X_0) + \int_0^t f_n'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)\mathrm{d}[X]_s \\ &= f_n(X_0) + \underbrace{\left[\int_0^t f_n'(X_s)\mathrm{d}A_s + \frac{1}{2}\int_0^t f_n''(X_s)\mathrm{d}[X]_s\right]}_{\text{finite variation}} + \underbrace{\int_0^t f_n'(X_s)\mathrm{d}M_s}_{\text{cnts semimartingale}}. \end{aligned}$$

For the finite variation part

$$\begin{aligned} \int_0^{t\wedge T_r} |f_n'(X_s) - f'(X_s)|\mathrm{d}V_s + \frac{1}{2}\int_0^{t\wedge T_r} |f_n''(X_s) - f_n(X_s)|\mathrm{d}[M]_s &\leq \Delta_{n,r}(V_{t\wedge T_r} + \frac{1}{2}[M]_{t\wedge T_r}) \\ &\leq 2r\Delta_{n,r} \\ &\xrightarrow{n\to\infty} 0 \end{aligned}$$

and so

$$\int_0^{t\wedge T_r} f_n'(X_s)\mathrm{d}A_s + \frac{1}{2}\int_0^{t\wedge T_r} f_n''(X_s)\mathrm{d}[M]_s \xrightarrow{n\to\infty} \int_0^{t\wedge T_r} f'(X_s)\mathrm{d}A_s + \frac{1}{2}\int_0^{t\wedge T_r} f''(X_s)\mathrm{d}[M]_s$$

uniformly in $t$. For the semimartingale part, note $M^{T_r} \in \mathcal{M}_c^2$ since $[M^{T_r}] \leq r$. Then

$$\begin{aligned} \|(f_n'(X)\cdot M)^{T_r} - (f'(X)\cdot M)^{T_r}\|^2 &= \mathbb{E}\left[\int_0^{T_r} (f_n'(X_s) - f'(X_s))^2\mathrm{d}[M]_s\right] \\ &\leq \Delta_{n,r}^2\mathbb{E}[[M]_{T_r}] \\ &\leq r\Delta_{n,r} \\ &\xrightarrow{n\to\infty} 0 \end{aligned}$$

and so $(f_n'(X)\cdot M)^{T_r} \to (f'(X)\cdot M)^{T_r}$ in $\mathcal{M}_c^2$ as $n \to \infty$. Therefore

$$f(X_{t\wedge T_r}) = f(X_0) + \int_0^{t\wedge T_r} f'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^{t\wedge T_r} f''(X_s)\mathrm{d}[X]_s.$$

□

## Statonovich Integral

Let $X, Y$ be continuous semimartingales. The Statonovich integral of $X$ against $Y$ is defined as

$$\int_0^t X_s\partial Y_s = \underbrace{\int_0^t X_s\mathrm{d}Y_s}_{\text{Itô}} + \frac{1}{2}[X,Y]_t.$$

Then

$$\sum_{k=0}^{\lceil 2^n t\rceil - 1} \left( \frac{X_{(k+1)2^{-n}} + X_{k2^{-n}}}{2} \right) (Y_{(k+1)2^{-n}} - Y_{k2^{-n}}) \xrightarrow{n\to\infty} \int_0^t X_s \partial Y_s.$$

**Proposition.** Let $X^1, \ldots, X^d$ be continuous semimartingales and let $f : \mathbb{R}^d \to \mathbb{R}$ be $C^3(\mathbb{R}^3)$. Then

$$f(X_t) = f(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s)\partial X_s^i.$$

In particular, integration by parts is given by

$$X_t Y_t - X_0 Y_0 = \int_0^t X_s \partial Y_s + \int_0^t Y_s \partial X_s$$

so the Statonovich integral satisfies the "usual rules" of calculus. But the Statonovich integral against $M \in \mathcal{M}_{c,\mathrm{loc}}$ is not in general in $\mathcal{M}_{c,\mathrm{loc}}$.

For example,

$$\begin{aligned}\int_0^t B_s \partial B_s &= \int_0^t B_s \mathrm{d}B_s + \frac{1}{2}t \\ &= \frac{1}{2}B_t^2 \notin \mathcal{M}_{c,\mathrm{loc}}\end{aligned}$$

for $B$ a standard Brownian motion.

*Proof of proposition.* We consider the case $d = 1$, the $d > 1$ case is left as an exercise. Itô's formula gives

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)\mathrm{d}[X]_s \tag{1}$$

$$f'(X_t) = f'(X_0) + \int_0^t f''(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f'''(X_s)\mathrm{d}[X]_s. \tag{2}$$

Therefore

$$[f'(X), X] = [f''(X) \cdot X, X] = f''(X) \cdot [X]$$

by Kunita-Watanabe. Therefore

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t f'(X_s)\mathrm{d}X_s + \frac{1}{2}[f'(X), X] \\ &= f(X_0) + \int_0^t f'(X_s)\partial X_s. \end{aligned}$$

□

Some common shorthand notation is

$$\begin{aligned} Z_t = Z_0 + \int_0^t H_s \mathrm{d}X_s &\iff \mathrm{d}Z_t = H_t \mathrm{d}X_t \\ Z_t = Z_0 + \int_0^t H_s \mathrm{d}X_s &\iff \partial \mathrm{d}Z_t + H_t \partial X_t \\ Z_t = [X, Y]_t = \int_0^t \mathrm{d}[X, Y]_s &\iff \mathrm{d}Z_t = \mathrm{d}X_t \mathrm{d}Y_t. \end{aligned}$$

Then we have the following computational rules:

$$H_t \mathrm{d}(K_t \mathrm{d}X_t) = (H_t K_t)\mathrm{d}X_t \quad \text{(iterated integral)}$$

$$H_t(\mathrm{d}X_t \mathrm{d}Y_t) = \mathrm{d}(H_t \mathrm{d}X_t)\mathrm{d}Y_t \quad \text{(Kunita-Watanabe)}$$

$$\mathrm{d}(X_t Y_t) = X_t \mathrm{d}Y_t + Y_t \mathrm{d}X_t + \mathrm{d}X_t \mathrm{d}Y_t \quad \text{(integration by parts)}$$

$$\mathrm{d}f(X_t) = \sum_{i=1}^d \frac{\partial f}{\partial x_i}(X_t)\mathrm{d}X_t^i + \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t)\mathrm{d}X_t^i \mathrm{d}X_t^j.$$

Applications:

**Theorem** (Lévy characterisaton)**.** *Let $X^1, \ldots, X^d \in \mathcal{M}_{c,loc}$ and $X = (X^1, \ldots, X^d)$. Suppose $X_0 = 0$, $[X^i, X^j]_t = \delta_{ij} t$ for all $i, j$ and all $t \geq 0$. Then $X$ is a standard Brownian motion.*

*Proof.* We need to show for all $s \in [0,t]$, $X_t - X_s$ is independent of $\mathcal{F}_s$ and has the law of a $\mathcal{N}(0,(t-s)I_d)$. Equivalently,

$$\mathbb{E}[\exp[i\theta \cdot (X_t - X_s)]|\mathcal{F}_s] = \exp\left[-\frac{|\theta|^2}{2}(t-s)\right]$$

for all $\theta \in \mathbb{R}^d$. For $\theta \in \mathbb{R}^d$ set $Y_t = \theta \cdot X_t = \sum_{j=1}^d \theta^j X_t^j$. Then $Y \in \mathcal{M}_{c,\text{loc}}$ since $\mathcal{M}_{c,\text{loc}}$ is a vector space. Moreover

$$\begin{aligned}[Y]_t = [Y,Y]_t &= \left[\sum_{j=1}^d \theta^j X_t^j, \sum_{j=1}^d \theta^j X_t^j\right]_t \\ &= \sum_{j=1}^d \theta^j \theta^k [X^j, X^k]_t \\ &= |\theta|_t^2.\end{aligned}$$

Now let $Z_t = \exp[iY_t + \frac{1}{2}[Y]_t] = \exp[i(\theta \cdot X_t) + \frac{1}{2}|\theta|^2 t]$. By Itô's formula applied to $W = iY + \frac{1}{2}[Y]$, $f(w) = e^w$ we have that

$$\begin{aligned}\mathrm{d}Z_t &= Z_t(i\mathrm{d}Y_t + \frac{1}{2}[Y]_t) - \frac{1}{2}Z_t \mathrm{d}[Y]_t \\ &= iZ_t \mathrm{d}Y_t\end{aligned}$$

so $Z \in \mathcal{M}_{c,\text{loc}}$ since $Y \in \mathcal{M}_{c,\text{loc}}$. Since $Z$ is bounded on $[0,t]$ for all $t \geq 0$, $Z \in \mathcal{M}_c$. Hence

$$\mathbb{E}[Z_t|\mathcal{F}_s]Z_s \implies \mathbb{E}[\exp[i\theta \cdot (X_t - X_s)]|\mathcal{F}_s] = \exp[-\frac{1}{2}|\theta|^2(t-s)].$$

□

**Theorem** (Dubins-Schwarz)**.** *Let $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$, $[M]_\infty = \infty$. Set $\tau_s = \inf\{t \geq 0 : [M]_t > s\}$, $B_s = M_{\tau_s}$, $\mathcal{G}_s = \mathcal{F}_{\tau_s}$. Then $(\tau_s)$ is an $(\mathcal{F}_t)$-stopping time and $[M]_{\tau_s} = s$ for all $s \geq 0$. Moreover $B$ is a $(\mathcal{G}_s)$-Brownian motion with $M_t = B_{[M]_t}$.*

*Therefore every continuous local martingale starting from* 0 *is a time-change of a standard Brownian motion.*

*Proof.* Since $[M]$ is continuous and adapted, $\tau_s$ is a stopping time for each $s \geq 0$. Since $[M]_\infty = \infty$, $\tau_s < \infty$ for all $s \geq 0$. Moreover. $(\mathcal{G}_s)$ is a filtration since if $S, T$ are stopping times with $S \leq T$ then $\mathcal{F}_S \subseteq \mathcal{F}_T$. Hence $s \leq t$ implies $\mathcal{G}_s \subseteq \mathcal{G}_t$.

First we show $B$ is adapted to $(\mathcal{G}_s)$. So we need to show $M_{\tau_s}$ is $\mathcal{F}_{\tau_s}$-measurable for all $s \geq 0$. From Example Sheet 1 we know that if $X$ is cádlág, adapted and $T$ a stopping time, then $X_T \mathbb{1}\{T < \infty\}$ is $\mathcal{F}_T$-measurable. Applying this for

$X = M$, $T = \tau_s$ and since $\mathbb{1}\{\tau_s < \infty\} = 1$ we get the result.

Now we show $B$ is continuous. We know $s \mapsto \tau_s$ is non-decreasing and cádlág so it follows $B$ is cádlág. To show $B$ is continuous we need to show $B_{s^-} = B_s$ for all $s \geq 0$, i.e $M_{\tau_s^-} = M_{\tau_s}$ where $\tau_s^- = \inf\{t \geq 0 : [M]_t = s\}$. If $\tau_s = \tau_{s^-}$ we're done. If $\tau_s > \tau_{s^-}$ then $[M]$ is constant on $[\tau_{s^-}, \tau_s]$. So we show that whenever $[M]$ is constant on an interval, then $M$ is constant too. By localisation, WLOG $M \in \mathcal{M}_c^2$. Suppose $q \in \mathbb{Q} \cap (0, \infty)$ and let $S_q = \inf\{t > q : [M]_t > [M]_q\}$. It suffices to show $M$ is almost-surely constant on each $[q, S_q]$. We know that $M^2 - [M]$ is a UI martingale since $M \in calM_c^2$. By the OST we have that

$$\mathbb{E}[M_{S_q}^2 - [M]_{S_q} | \mathcal{F}_q] = M_q^2 - [M]_q. \tag{$*$}$$

Since $M \in \mathcal{M}_c^2$ we also have that

$$\begin{aligned}
\mathbb{E}[(M_{S_q} - M_q)^2 | \mathcal{F}_q] &= \mathbb{E}[M_{S_q}^2 - M_q^2 | \mathcal{F}_q] && \text{(MG orthogonality)} \\
&= \mathbb{E}[[M]_{S_q} - [M]_q | \mathcal{F}_q] && \text{(by $(*)$)} \\
&= 0.
\end{aligned}$$

Therefore $M_{S_q} - M_q = 0$ almost-surely so $M$ is almost-surely constant on $[q, S_q]$.

Next we show $B$ is a $(\mathcal{G}_s)$-Brownian motion. Fix $s > 0$. Then we know that $[M^{\tau_s}]_\infty = [M]_{\tau_s} = s$. Therefore $M^{\tau_s} \in \mathcal{M}_c^2$ since $\mathbb{E}[[M^{\tau_s}]_\infty] < \infty$. Therefore $(M^2 - [M])^{\tau_s}$ is a UI martingale. By the OST, for $0 \leq r < s < \infty$ we have

1. $\mathbb{E}[B_s | \mathcal{G}_r] = \mathbb{E}[M_\infty^{\tau_s} | \mathcal{F}_{\tau_r}] = M_{\tau_r} = B_r$;
2. $\mathbb{E}[B_s^2 - s | \mathcal{G}_r] = \mathbb{E}[(M^2 - [M])_\infty^{\tau_s} | \mathcal{F}_{\tau_r}] = M_{\tau_r}^2 - [M]_{\tau_r} = B_r^2 - r$.

(1) tells us $B \in \mathcal{M}_c$ and (2) tells us $[B]_t = t$. Hence $B$ is a $\mathcal{G}_s$-Brownian motion by the Lévy characterisaton. $\square$

Dubins-Schwarz requires $[M]_\infty = \infty$. There is an extension to the case $[M]_\infty < \infty$.

**Theorem.** *Let $M \in \mathcal{M}_{c,loc}$, $M_0 = 0$. Let $\beta$ be a Brownian motion which is independent of $M$. Set*

$$B_s = \begin{cases} M_{\tau_s} & s \le [M]_\infty \\ M_\infty + \beta_{s-[M]_\infty} & s > [M]_\infty \end{cases}.$$

*Then $B$ is a standard Brownian motion and $M_t = B_{[M]_t}$ for all $t \geq 0$.*

**Examples.**

(i) $B$ a standard Brownian motion, $h$ a (deterministic) $L^2([0,\infty))$ function. Let $M_t = \int_0^t h(s)\mathrm{d}B_s$. Then $M_0 = 0$, $M \in \mathcal{M}_{c,\mathrm{loc}}$ and $[M]_t = \int_0^t h(s)^2\mathrm{d}s$. Moreover, $M_\infty =^d B_{\int_0^\infty h(s)^2\mathrm{d}s} \sim \mathcal{N}(0, \|h\|_{L^2}^2)$.

(ii) Let $M \in \mathcal{M}_{c,\mathrm{loc}}$. Then

$$\begin{aligned} \{[M]_\infty < \infty\} &= \{\lim_{t\to\infty} M_t \text{ exists}\} \\ \{[M]_\infty < \infty\} &= \{\liminf_{t\to\infty} M_t = -\infty, \limsup_{t\to\infty} M_t = \infty\}. \end{aligned}$$

## Exponential Martingales

Let $M \in \mathcal{M}_{c,\mathrm{loc}}$, $M_0 = 0$. Set $Z_t = \exp[M_t - [M]_t/2]$. By Itô's formula

$$\mathrm{d}Z_t = Z_t(\mathrm{d}M_t - \frac{1}{2}\mathrm{d}[M]_t) + \frac{1}{2}Z_t\mathrm{d}[M]_t = Z_t\mathrm{d}[M]_t.$$

Hence $Z \in \mathcal{M}_{c,\mathrm{loc}}$, $Z_0 = 1$.

**Definition** (Exponential martingale)**.** In the above setting, this process

$$\mathcal{E}(M)_t := Z_t = \exp(M_t - [M]_t/2)$$

is the *stochastic exponential* or *exponential martingale* associated with $M$.

Note that $\mathcal{E}(M) \in \mathcal{M}_{c,\mathrm{loc}}$ and $\mathrm{d}\mathcal{E}(M)_t = \mathcal{E}(M)_t\mathrm{d}M_t$.

**Proposition.** Let $M \in \mathcal{M}_{c,\mathrm{loc}}$, $M_0 = 0$. For all $\varepsilon, \delta > 0$ we have that

$$\mathbb{P}[\sup_{t\geq 0} M_t \geq \varepsilon, [M]_\infty \leq \delta] \leq e^{-\frac{\varepsilon^2}{2\delta}}.$$

*Proof.* Fix $\varepsilon > 0$ and let $T = \inf\{t \geq 0 : M_t \geq \varepsilon\}$. Fix $\theta > 0$ and set $Z_t = \mathcal{E}(\theta M^T)_t = \exp[\theta M_t^T - \frac{1}{2}\theta^2[M]_t^T] \in \mathcal{M}_{c,\mathrm{loc}}$. Note that $|Z_t| \leq \exp[\theta\varepsilon]$ for all $t \geq 0$. Hence $Z$ is a bounded martingale, so $\mathbb{E}[Z_\infty] = Z_0 = 1$. For $\delta > 0$ we have that

$$\begin{aligned}
\mathbb{P}[\sup_{t\geq 0} M_t, [M]_\infty \leq \delta] &= \mathbb{P}[\sup_{t\geq 0} e^{\theta M_t^T} \geq e^{\theta\varepsilon}, [M]_\infty \leq \delta] \\
&\leq \mathbb{P}[\sup_{t\geq 0} Z_t \geq e^{\theta\varepsilon - \frac{\theta^2\delta}{2}}] \\
&\leq \exp\left[-\theta\varepsilon + \frac{\theta^2\delta}{2}\right] \qquad \text{(Doob)}
\end{aligned}$$

and optimising over $\theta$ gives the claim. □

**Proposition.** Let $M \in \mathcal{M}_{c,\mathrm{loc}}$, $M_0 = 0$. If $[M]$ is bounded, then $\mathcal{E}(M)$ is a UI martingale.

*Proof.* We will show $\mathcal{E}(M)$ is bounded by an integrable random variable. Note that

$$\sup_{t\geq 0} \mathcal{E}(M)_t \leq \exp[\sup_{t\geq 0} M_t]$$

since $[M]_t \geq 0$. We show that the RHS is integrable. Let $C > 0$ be so that $[M]_\infty \leq C$. Then

$$\begin{aligned}
\mathbb{P}[\sup_{t\geq 0} M_t \geq \varepsilon] &= \mathbb{P}[\sup_{t\geq 0} M_t \geq \varepsilon, [M]_\infty \leq C] \\
&\leq e^{-\frac{\varepsilon^2}{2C}}
\end{aligned}$$

and therefore

$$\begin{aligned}
\mathbb{E}[\exp[\sup_{t\geq 0} M_t]] &= \int_0^\infty \mathbb{P}(\exp[\sup_{t\geq 0} M_t] \geq \lambda)\mathrm{d}\lambda \\
&= \int_0^\infty \mathbb{P}[\sup_{t\geq 0} M_t \geq \log\lambda]\mathrm{d}\lambda \\
&\leq 1 + \int_1^\infty \exp\left(-\frac{(\log\lambda)^2}{2C}\right)\mathrm{d}\lambda \\
&< \infty
\end{aligned}$$

so $\mathcal{E}(M)$ is UI. □

Suppose that $\mathbb{P}, \mathbb{Q}$ are probability measures on $(\Omega, \mathcal{F})$. We say that $\mathbb{Q}$ is *absolutely continuous* with respect to $\mathbb{P}$ if $\mathbb{Q}(A) = 0$ whenever $\mathbb{P}(A) = 0$. We write $\mathbb{Q} \ll \mathbb{P}$ in this case.

The Radon-Nikodym theorem says that whenever $\mathbb{Q} \ll \mathbb{P}$, there exists a random variable $Z \geq 0$ such that $\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}[Z\mathbb{1}_A]$ for all $A \in \mathcal{F}$. We call $Z$ the *Radon-Nikodym derivative* of $\mathbb{Q}$ with respect to $\mathbb{P}$, written as $Z = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}$.

**Example.** Suppose that $X \sim \mathcal{N}(0,1)$, $\mu \in \mathbb{R}$. Let $Z = \exp(\mu X - \mu^2/2)$. Then $A \mapsto \mathbb{E}[\mathbb{1}_A Z]$ is a probability measure $\mathbb{Q}$ and under $\mathbb{Q}$, $X \sim \mathcal{N}(\mu, 1)$.

The Girsonov theorem generalises this idea to the setting of semimartingales, except instead of changing the mean of a random variable, we change the semimartingale decomposition.

**Theorem** (Girsonov theorem)**.** *Let $M \in \mathcal{M}_{c,loc}$, $M_0 = 0$ and assume that $Z = \mathcal{E}(M)$ is UI. Then we can construct a probability measure $\tilde{\mathbb{P}} \ll \mathbb{P}$ on $(\Omega, \mathcal{F})$ by setting $\tilde{\mathbb{P}}(A) = \mathbb{E}[Z_\infty \mathbb{1}_A]$ for $A \in \mathcal{F}$. If $X \in \mathcal{M}_{c,loc}(\mathbb{P})$, then $X - [X,M] \sim \mathcal{M}_{c,loc}(\tilde{\mathbb{P}})$.*

*Proof.* Since $Z$ is UI, we know that $Z_\infty$ exists and $Z_\infty \geq 0$ with $\mathbb{E}[Z_\infty] = \mathbb{E}[Z_0] = 1$. Therefore $\tilde{\mathbb{P}}$ indeed defines a probability measure with $\tilde{\mathbb{P}} \ll \mathbb{P}$.

Suppose that $X \in \mathcal{M}_{c,\mathrm{loc}}(\mathbb{P})$ and set $T_n = \inf\{t \geq 0 : |X_t - [X,M]_t| \geq n\}$. Since $X - [X,M]$ is continuous, we have that

$$\mathbb{P}(T_n \uparrow \infty) = 1 \implies \tilde{\mathbb{P}}(T_n \uparrow \infty) = 1.$$

To prove $Y := X - [X,M] \in \mathcal{M}_{c,\mathrm{loc}}(\tilde{\mathbb{P}})$, it suffices to show $Y^{T_n} \in \mathcal{M}_c(\tilde{\mathbb{P}})$ for all $n$.

From now we just write $X, Y$ in place of $X^{T_n}, Y^{T_n}$. Note

$$\begin{aligned}
\mathrm{d}(Z_tY_t) &= Y_t\mathrm{d}Z_t + Z_t\mathrm{d}Y_t + \mathrm{d}Y_t\mathrm{d}Z_t && \text{(IBP)} \\
&= Y_t\mathrm{d}Z_t + Z_t\mathrm{d}X_t - Z_t\mathrm{d}X_t\mathrm{d}M_t + \mathrm{d}X_t\mathrm{d}Z_t && (Y = X - [X,M]) \\
&= Y_t\mathrm{d}Z_t + Z_t\mathrm{d}X_t && (\mathrm{d}Z_t = Z_t\mathrm{d}M_t)
\end{aligned}$$

which implies $ZY \in \mathcal{M}_{c,\mathrm{loc}}(\mathbb{P})$. Moreover, $\{Z_T : T \leq t \text{ a stopping time}\}$ is UI for each $t \geq 0$. Since $Y$ is bounded, we also have that $\{Z_TY_T : T \leq t \text{ a stopping time}\}$ is UI, so in fact $ZY \in \mathcal{M}_c(\mathbb{P})$. Then for $s < t$ we have

$$\begin{aligned}
\tilde{\mathbb{E}}[Y_t - Y_s|\mathcal{F}_s] &= \frac{1}{Z_s}\mathbb{E}[Z_\infty Y_t - Z_\infty Y_s|\mathcal{F}_s] \\
&= \frac{1}{Z_s}\mathbb{E}[Z_tY_t - Z_sY_s|\mathcal{F}_s] && \text{(tower property)} \\
&= 0
\end{aligned}$$

so $Y \in \mathcal{M}_c(\tilde{\mathbb{P}})$. $\square$

**Remark.** The quadratic variation does not change when performing a change of measures (see Example Sheet 4).

**Corollary.** *Let $B$ be a standard Brownian motion under $\mathbb{P}$, $M \in \mathcal{M}_{c,loc}$, $M_0 = 0$. Suppose that $Z = \mathcal{E}(M)$ is UI and $\tilde{\mathbb{P}}(A) = \mathbb{E}[\mathbb{1}_A Z_\infty]$ for all $A \in \mathcal{F}$. Then $\tilde{B} = B - [B, M]$ is a $\tilde{\mathbb{P}}$-Brownian motion.*

*Proof.* Since $\tilde{B} \in \mathcal{M}_{c,\text{loc}}(\tilde{\mathbb{P}})$ by the Girsonov theorem, and $[\tilde{B}]_t = [B]_t = t$, the result follows from the Lévy characterisaton. □

**Example.** Suppose that $B$ is a $\mathbb{P}$-Brownian motion, $\mu \in \mathbb{R}$, $T > 0$ and let $M_t = \mu B_t^T$ so that $Z_t = \mathcal{E}(M)_t = \exp[\mu B_t^T - \mu^2 (t \wedge T)/2]$. Then $\tilde{\mathbb{P}}(A) = \mathbb{E}[Z_\infty \mathbb{1}_A] = \mathbb{E}[\exp[\mu B_T - \mu^2 T/2] \mathbb{1}_A]$ for all $A \in \mathcal{F}$. Then under $\tilde{\mathbb{P}}$, $B_t = \tilde{B}_t + \mu t$ for $t \in [0, T]$. and $\tilde{B}$ is a $\tilde{\mathbb{P}}$-Brownian motion.

# Stochastic Differential Equations

Let $M^{d \times m}(\mathbb{R})$ denote the space of $d \times m$ matrices with real entries. Suppose that $\sigma : \mathbb{R}^d \to M^{d \times m}(\mathbb{R})$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are measurable functions which are bounded on compact sets. Write $\sigma(x) = (\sigma_{ij}(x))$, $b(x) = (b_i(x))$. Consider

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

or equivalently

$$\mathrm{d}X_t^i = \sum_{j=1}^m \sigma_{ij}(X_t)\mathrm{d}B_t^j + b_i(X_t)\mathrm{d}t$$

for $1 \leq i \leq d$. A solution to $(*)$ consists of:

- A filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ where $(\mathcal{F}_t)_{t \geq 0}$ satisfies the usual conditions;
- An $(\mathcal{F}_t)_{t \geq 0}$-Brownian motion $B = (B^1, \ldots, B^m) \in \mathbb{R}^m$;
- An $(\mathcal{F}_t)_{t \geq 0}$-adapted continuous process $X = (X^1, \ldots, X^d)$ in $\mathbb{R}^d$ such that $X_t = X_0 + \int_0^t \sigma(X_s)\mathrm{d}B_s + \int_0^t b(X_s)\mathrm{d}s$.

When in addition $X_0 = x \in \mathbb{R}^d$, we say that $X$ is *started from* $x$. Some terminology:

- We say that an SDE has a *weak solution* if for all $x \in \mathbb{R}^d$, there is a solution starting from $x$;
- There is *uniqueness in law* if all solutions starting from each $x$ have the same distribution;
- There is *pathwise uniqueness* if when we fix $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and $B$, then any two solutions $X, X'$ with $X_0 = X'_0$ are indistinguishable, i.e $\mathbb{P}(X_t = X'_t \ \forall t) = 1$;

- We say that a solution started from $x$ is a *strong solution* if $X$ is adapted to the filtration generated by $B$.

**Example.** It is possible to have existence of weak solution and uniqueness in law without pathwise uniqueness. Suppose that $\beta$ is a standard Brownian motion on $\mathbb{R}$ with $\beta_0 = x$. Set $B_t = \int_0^t \operatorname{sign}(\beta_s)\mathrm{d}\beta_s$ where $\operatorname{sign}(x) = \mathbb{1}_{(0,\infty)}(x) - \mathbb{1}_{(-\infty,0]}$. Note that $\operatorname{sign}(\beta_s)$ is previsible, bounded so the integral is well-defined. Also

$$\begin{aligned} x + \int_0^t \operatorname{sign}(\beta_s)\mathrm{d}\beta_s &= x + \int_0^t (\operatorname{sign}(\beta_s))^2 \mathrm{d}\beta_s \\ &= x + \int_0^t \mathrm{d}\beta_s \\ &= x + \beta_t \end{aligned}$$

so $\beta$ solves the SDE $\mathrm{d}X_T = \operatorname{sign}(X_t)\mathrm{d}B_t$. This SDE has a weak solution and by the Lévy characterisaton any solution is a Brownian motion (continuous local martingale with quadratic variation equal to $t$). Hence we have uniqueness in law. However we do not have pathwise uniqueness. For example take $X_0 = x = 0$. Then we claim $\beta, -\beta$ are solutions. We know $\beta$ is a solution and

$$\begin{aligned} -\beta_t &= -\int_0^t \operatorname{sign}(\beta_s)\mathrm{d}B_s \\ &= \int_0^t (sign)(-\beta_s)\mathrm{d}B_s + 2\int_0^t \mathbb{1}\{\beta_s = 0\}\mathrm{d}B_s. \end{aligned}$$

The final term is in $\mathcal{M}_{c,\text{loc}}$, starts from 0 and has quadratic variation equal to $4\int_0^t \mathbb{1}\{B_s = 0\}\mathrm{d}s = 0$ (equal to zero since $\mathbb{P}(B_s = 0) = 0$ for all $s > 0$ and Fubini). Therefore $\beta, -\beta$ are both solutions on the same probability space with the same Brownian motion so we cannot have have pathwise uniqueness.

## Lipschitz Coefficients

For $d, m \geq 1$ we equip $M^{d\times m}(\mathbb{R})$ with the Frobenius norm, i.e if $A = (a_{ij})$ we have

$$|A| = \left(\sum_{i=1}^{d}\sum_{j=1}^{m} a_{ij}^2\right)^{1/2}.$$

If $f : U \to M^{d\times m}(\mathbb{R})$ say that $f$ is Lipschitz if there exists $K < \infty$ such that $|f(x) - f(y)| \leq K|x - y|$ for all $x, y \in U$.

**Theorem** (Existence and uniqueness)**.** *Suppose that $\sigma : \mathbb{R}^d \to M^{d\times m}(\mathbb{R})$, $b : \mathbb{R}^d \to \mathbb{R}^d$ are Lipschitz. Then there is pathwise uniqueness for the SDE*

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

*Moreover, for each filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying the usual conditions and each $(\mathcal{F}_t)$-Brownian motion $B$, $x \in \mathbb{R}^d$ there is a strong solution starting from $x$.*

The proof of this is analogous to existence/uniqueness for ODEs.

Recall the following result from undergraduate analysis.

**Theorem** (Contraction mapping theorem)**.** *Let $(X, d)$ be a complete metric space.*

*(a) Suppose that $F : X \to X$ is a contraction, i.e there exists $r < 1$ such that $d(F(x), F(y)) \leq rd(x, y)$ for all $x, y \in X$. Then $F$ has a unique fixed point.*

*(b) Suppose that $F : X \to X$ and there exists $n \in \mathbb{N}$ so that $F^n$ is a contraction. Then $F$ has a unique fixed point.*

**Lemma** (Gronwall)**.** *Let $T > 0$ and $f : [0, T] \to [0, \infty)$ be a bounded and measurable function. If there exists $a, b \geq 0$ such that*

$$f(t) \leq a + b \int_0^t f(s)\mathrm{d}s$$

*for all $t \in [0, T]$, then $f(t) \leq ae^{bt}$ for all $t \in [0, T]$.*

*Proof.* Example Sheet 3. □

Now we are ready to prove the main result.

*Proof of existence and uniqueness.* We will assume $d, m = 1$. Let $K$ be such that

$$\begin{aligned} |\sigma(x) - \sigma(y)| &\leq K|x - y| \\ |b(x) - b(y)| &\leq K|x - y| \end{aligned}$$

for all $x, y \in \mathbb{R}$.

First we show uniqueness. Suppose $X, X'$ are two solutions on the same probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ and Brownian motion $B$. We want to show $X_t = X'_t$ for all $t \geq 0$ almost-surely. Fix $M > 0$ and let $\tau = \inf\{t \geq 0 : |X_t| \vee |X'_t| \geq M\}$. Then

$$\begin{aligned} X_{t\wedge\tau} &= X_0 + \int_0^{t\wedge\tau} \sigma(X_s)\mathrm{d}B_s + \int_0^{t\wedge\tau} b(X_s)\mathrm{d}s \\ X'_{t\wedge\tau} &= X_0 + \int_0^{t\wedge\tau} \sigma(X'_s)\mathrm{d}B_s + \int_0^{t\wedge\tau} b(X'_s)\mathrm{d}s. \end{aligned}$$

Fix $T > 0$. If $t \in [0, T]$ we have

$$
\begin{aligned}
&\mathbb{E}[(X_{t\wedge\tau} - X'_{t\wedge\tau})^2] \\
&\leq 2\mathbb{E}\left[\left(\int_0^t (\sigma(X_s) - \sigma(X'_s))\mathrm{d}B_s\right)^2\right] + 2\mathbb{E}\left[\left(\int_0^t (b(X_s) - b(X'_s))\mathrm{d}s\right)^2\right] \\
&\leq 2\mathbb{E}\left[\left(\int_0^t (\sigma(X_s) - \sigma(X'_s))^2\mathrm{d}s\right)^2\right] + 2T\mathbb{E}\left[\left(\int_0^t (b(X_s) - b(X'_s))^2\mathrm{d}s\right)\right] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(Itô isometry \& C-S)} \\
&\leq 2K^2(1+T)\mathbb{E}\left[\int_0^{t\wedge T} (X_s - X'_s)^2\mathrm{d}s\right] \\
&\leq 2K^2(1+T)\int_0^t \mathbb{E}[(X_{s\wedge\tau} - X'_{s\wedge\tau})^2]\mathrm{d}s
\end{aligned}
$$

so if we let $f(t) = \mathbb{E}[(X_{t\wedge\tau} - X'_{t\wedge\tau})^2]$, then $0 \leq f \leq 4M^2$ and $f(t) \leq 2K^2(1 + T)\int_0^t f(s)\mathrm{d}s$ for all $t \in [0, T]$. Hence Gronwall implies $f(t) = 0$. Therefore $\mathbb{P}(X_{t\wedge\tau} = X'_{t\wedge\tau}\ \forall t \in [0, T]) = 1$. Since $M, T$ were arbitrary we have pathwise uniqueness.

Now we show existence. Suppose that $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ is a filtered probability space, $B$ is a $(\mathcal{F}_t)$-Brownian motion and $(\mathcal{F}_t^B)_{t\geq 0}$ the filtration generated by $B$ (so that $\mathcal{F}_t^B \subseteq \mathcal{F}_t$). We will use the contraction mapping theorem. To do so we need to specify a space and a map. For each $T > 0$ let $\mathcal{C}_T$ be the space of continuous adapted processes $X : [0, T] \to \mathbb{R}$. We equip $\mathcal{C}_T$ with the norm $|||X|||_T = \|\sup_{0\leq t\leq T} |X_t|\|_{L^2} < \infty$.

We have shown before that $\mathcal{C}_T$ is complete. Fix $x \in \mathbb{R}$. Using that $\sigma, b$ are Lipschitz, we have that

$$|\sigma(y)| \leq |\sigma(y) - \sigma(0)| + |\sigma(0)| \leq |\sigma(0)| + K|y| \text{ and} \tag{7}$$

$$|b(y)| \leq |b(0)| + K|y| \tag{8}$$

for all $y \in \mathbb{R}$. Fix $T > 0$ and $X \in \mathcal{C}_T$. Let $M_T = \int_0^t \sigma(X_s)\mathrm{d}B_s$ for $0 \leq t \leq T$. Then $[M]_T = \int_0^t \sigma(X_s)^2\mathrm{d}s$. Thus (1) implies $\mathbb{E}[[M]_T] \leq 2T(|\sigma(0)|^2 + K^2|||X|||_T) < \infty$. Therefore $M^T \in \mathcal{M}_c^2$ so Doob's inequality implies

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left(\int_0^t \sigma(X_s)\mathrm{d}B_s\right)^2\right] \leq 8T(|\sigma(0)|^2 + K^2|||X|||_T^2).$$

By (2), have

$$
\begin{aligned}
\mathbb{E}\left[\sup_{0\leq t\leq T}\left|\int_0^t b(X_s)\mathrm{d}s\right|^2\right] &\leq T\mathbb{E}\left[\int_0^T b(X_s)^2\mathrm{d}B_s\right] &\text{(C-S)} \\
&\leq 2T^2(|b(0)|^2 + K^2|||X|||_T^2) < \infty.
\end{aligned}
$$

Define $F : \mathcal{C}_T \to \mathcal{C}_T$ by $F(X)_t = x + \int_0^t \sigma(X_s)\mathrm{d}B_s + \int_0^t b(X_s)\mathrm{d}s$ (which is well-defined by the above).

Suppose that $X, Y \in \mathcal{C}_T$. Then by a similar argument to the above,

$$|||F(X) - F(Y)|||_t^2 \leq 2K^2(4+T)\int_0^t |||X-Y|||_s^2 \mathrm{d}s$$
$$:= C_T \int_0^t |||X-Y|||_S^2 \mathrm{d}s$$

for $0 \leq t \leq T$. Iterating $n$ times

$$\begin{aligned} |||F^n(X) - F^n(Y)|||_T^2 &\leq C_T^n \int_0^T \int_0^{t_1} \dots \int_0^{t_{n-1}} |||X-Y|||_{t_n}^2 \mathrm{d}t_n \dots \mathrm{d}t_1 \\ &\leq C_T^n \frac{T^n |||X-Y|||_T^2}{n!}. \end{aligned} \tag{(3)}$$

Take $n$ sufficiently large so that $C_T^n T^n / n! < 1$. Then the contraction mapping theorem implies there exists a unique fixed point $X^{(T)} \in \mathcal{C}_T$ of $F$. Pathwise uniqueness implies $X_t^{(T)} = X_t^{(T')}$ for all $t \leq T \wedge T'$ almost-surely. So define $X$ by setting $X_t = X_t^{(N)}$ whenever $t \leq N$, $N \in \mathbb{N}$. Hence $X$ is the pathwise unique solution to the SDE starting from $x$.

We need to show $X$ is a strong solution, i.e $X$ is adapted to $(\mathcal{F}_t^B)$. We will prove that for each fixed $T > 0$, $X^{(T)}$ is the limit of $(\mathcal{F}_t^B)$-adapted processes. Define $(Y^n)$ in $\mathcal{C}_T$ by setting $Y^0 = x$ and $Y^n = F(Y^{n-1})$ for each $n \geq 1$. Then $Y^n$ is adapted to $(\mathcal{F}_t^B)$ for each $n$. Since $F^{(n)}(X) = X$ for all $n \geq 1$, we have from (3) that

$$\begin{aligned} |||X - Y^n|||_T^2 &= |||F^n(X) - F^n(x)|||_T^2 \\ &\leq \frac{C_T^n T^n}{n!} |||X - x|||_T^2 \\ &\xrightarrow{n\to\infty} 0. \end{aligned}$$

This implies $Y^n \to X$ in $\mathcal{C}_T$. Therefore there exists a subsequence $(Y^{n_k})$ so that $Y^{n_k} \to X$ uniformly in $[0,T]$ almost-surely. Therefore $(X_t^T)$ is the almost-sure limit of $(\mathcal{F}_t^B)$-adapted processes and therefore is itself $(\mathcal{F}_t^B)$-adapted. Since $T > 0$ was arbitrary we are done. $\square$

**Remark.** The above proof implies the pathwise unique strong solution lies in $\mathcal{C}_T$ for all $T > 0$.

**Proposition.** Under the hypotheses of the theorem, there is uniqueness in law for the SDE

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

*Proof.* Example Sheet 3. □

**Example** (Ornstein-Uhlenbeck process)**.** Fix $\lambda \in \mathbb{R}$ and consider the SDE

$$\begin{aligned}\mathrm{d}V_t &= \mathrm{d}B_t - \lambda V_t \mathrm{d}t,\ V_0 = v_0\\ \mathrm{d}X_t &= V_t \mathrm{d}t.\end{aligned}$$

For $\lambda > 0$ this models the movement of a grain of pollen in liquid, $X$ being the position of the grain, $V$ the velocity. The term $-\lambda V$ dampens the system due to viscosity, since when $|V|$ is large the system moves to reduce $|V|$. The theorem implies the existence of a pathwise unique strong solution. We can explicitly solve it:

$$\begin{aligned}\mathrm{d}(e^{\lambda t}V_t) &= e^{\lambda t}\mathrm{d}V_t + \lambda e^{\lambda t}V_t\mathrm{d}t && \text{(integration-by-parts)}\\ &= e^{\lambda t}\mathrm{d}B_t\end{aligned}$$

and so $e^{\lambda t}V_t = v_0 + \int_0^t e^{\lambda s}\mathrm{d}B_s$, implying

$$V_t = e^{-\lambda t}v_0 + \int_0^t e^{\lambda(s-t)}\mathrm{d}B_s.$$

Therefore $V_t \sim \mathcal{N}\left(e^{-\lambda t}v_0, \frac{1-e^{-2\lambda t}}{2\lambda}\right)$. If $\lambda > 0$ then $V_t \xrightarrow{d} \mathcal{N}(0,(2\lambda)^{-2})$. Hence $\mathcal{N}(0,(2\lambda)^{-1})$ is the stationary distribution of $V$, i.e if $V_0 \sim \mathcal{N}(0,(2\lambda)^{-1})$ then $V_t \sim \mathcal{N}(0,(2\lambda)^{-1})$ for all $t \geq 0$.

## Local solutions

A *locally defined process* is a pair $(X, \mathcal{T})$ consisting of a stopping time $\mathcal{T}$ together with a map $X : \{(\omega, t) :\in \Omega \times [0, \infty) : t < \mathcal{T}(\omega)\} \to \mathbb{R}$. We say it is *cádlág* if the map $t \mapsto X_t(\omega)$ from $[0, \mathcal{T}(\omega)) \to \mathbb{R}$ is cádlág for all $\omega \in \Omega$. Let $\Omega_t = \{\omega \in \Omega : t < \mathcal{T}(\omega)\}$. Then $(X, \mathcal{T})$ is *adapted* if $X_t : \Omega_t \to \mathbb{R}$ is $\mathcal{F}_t$-measurable for all $t \geq 0$. We say that $(X, \mathcal{T})$ is a *locally defined martingale* if there exists stopping times $T_n \uparrow \mathcal{T}$ so that $X^{T_n}$ is a martingale for all $n$.

We say that $(H, \eta)$ is a *locally defined locally bounded previsible process* if there exist stopping times $S_n \uparrow \eta$ such that $H\mathbb{1}_{(0,S_n]}$ is bounded and previsible for all $n$. We define $(H \cdot X, \mathcal{T} \wedge \eta)$ by

$$(H \cdot X, \mathcal{T} \wedge \eta)_t^{T_n \wedge S_n} = (H\mathbb{1}_{(0,S_n \wedge T_n]} \cdot X^{T_n \wedge S_n})_t$$

for each $n$.

**Proposition** (Local Itô's formula)**.** Let $X^1, \dots, X^d$ be continuous semimartingales, let $U \subseteq \mathbb{R}^d$ be open and let $f : U \to \mathbb{R}$ be $C^2(U)$. Let $X = (X^1, \dots, X^d)$ and set $\mathcal{T} = \inf\{t \geq 0 : X_t \notin U\}$. Then for all $t < \mathcal{T}$ we have that

$$f(X_t) = f(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s)\mathrm{d}X_s^i + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s)\mathrm{d}[X^i, X^j]_s.$$

*Proof.* Apply Itô's formula to $X^{T_n}$ where $T_n = \inf\{t \geq 0 : \mathrm{dist}(X_t, U^c) \leq 1/n\}$ and note that $T_n \uparrow \mathcal{T}$. □

**Example.** Let $X = B$ where $B$ is a standard Brownian motion with $X_0 = B_0 = 1$, $U = (0, \infty)$, $f(x) = \sqrt{x}$. Then

$$\sqrt{B_t} = 1 + \frac{1}{2}\int_0^t B_s^{-1/2}\mathrm{d}B_s - \frac{1}{8}\int_0^t B_s^{-3/2}\mathrm{d}s$$

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

Let $U \subseteq \mathbb{R}^d$ be open, $\sigma : U \to M^{d\times m}(\mathbb{R})$, $b : U \to \mathbb{R}^d$ be measurable functions which are bounded on compact subsets of $U$.

A *local solution* to the SDE

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

consists of:

- A filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying the usual conditions;
- A $(\mathcal{F}_t)$-Brownian motion $B$ in $\mathbb{R}^m$;

- A continuous $(\mathcal{F}_t)$-adapted locally defined process $(X, \mathcal{T})$ with $X \in \mathbb{R}^d$ such that

$$X_t = X_0 + \int_0^t \sigma(X_s)\mathrm{d}B_s + \int_0^t b(X_s)\mathrm{d}s$$

  for all $t < \mathcal{T}$.

We say that a solution $(X, \mathcal{T})$ is a *maximal* local solution if for any other local solution $(Y, \eta)$ on the same filtered probability space which has $X_t = Y_t$ for all $t < \mathcal{T} \wedge \eta$ we have $\eta \leq \mathcal{T}$.

## Locally Lipschitz coefficients

Suppose that $U \subseteq \mathbb{R}^d$ is open. Then a function $f : U \to \mathbb{R}^d$ is *locally Lipschitz* if for each $C \subseteq U$ compact we have that $f|_C$ is Lipschitz.

**Theorem.** *Suppose $U \subseteq \mathbb{R}^d$ is open and $\sigma : U \to M^{d\times m}(\mathbb{R})$, $b : U \to \mathbb{R}^d$ are locally Lipschitz. Then for all $x \in U$ the SDE $\mathrm{d}X_t = \sigma(X_t)\mathrm{d}B_t + b(X_t)\mathrm{d}t$ has a pathwise unique maximal local solution $(X, \mathcal{T})$ starting from $x$. Moreover, for all compact $C \subseteq U$, on the event that $\{\mathcal{T} < \infty\}$ we have that*

$$\sup\{t < \mathcal{T} : X_t \in C\} < \mathcal{T}.$$

**Lemma.** *Let $U \subseteq \mathbb{R}^d$ be open, $C \subseteq U$ be compact. Then*

*(i) There exists $\varphi \in C^\infty(\mathbb{R}^d, \mathbb{R})$ such that $\varphi|_C = 1$, $\varphi|_{U^c} = 0$;*

*(ii) Given a locally Lipschitz function $f : U \to \mathbb{R}$, then there exists a globally Lipschitz function $g : \mathbb{R}^d \to \mathbb{R}$ such that $f|_C = g|_C$.*

*Proof.*

(i) Exercise (e.g see Part II Analysis of Functions);

(ii) Take $\varphi$ as in (i) and set $g = f\varphi$ (extending $f$ to 0 outside $U$ say).

□

*Proof of theorem.* Assume that $d = m = 1$. Fix $C \subseteq U$ compact. By the lemma, we can find Lipschitz functions $\tilde{\sigma}, \tilde{b}$ on $\mathbb{R}$ such that $\tilde{\sigma}|_C = \sigma|_C$, $\tilde{b}|_C = b|_C$. There exists a pathwise unique strong solution $\tilde{X}$ to

$$\begin{cases} \mathrm{d}\tilde{X}_t = \tilde{\sigma}(\tilde{X}_t)\mathrm{d}B_t + \tilde{b}(\tilde{X}_t)\mathrm{d}t \\ \tilde{X}_0 = x \end{cases}.$$

Let $T = \inf\{t \geq 0 : \tilde{X}_t \notin C\}$ and let $X = \tilde{X}|_{[0,T)}$. Then $(X, T)$ is a local solution in $C$. If $T < \infty$ then $X_T = \lim_{s\uparrow T} X_s$ exists and $X_T \in U \setminus \mathrm{int}(C)$. Suppose that $(X, T), (Y, S)$ are both local solutions in $C$. Let:

$$f(t) = \mathbb{E}[\sup_{0\leq s\leq t\wedge T\wedge S} |X_s - Y_s|^2].$$

As $b, \sigma$ are Lipschitz on $C$, we can use Gronwall as before in the basic uniqueness & existence theorem to see $f = 0$. Hence $X_t = Y_t$ for all $t < T \wedge S$ almost-surely.

Let $(C_n)$ be an increasing sequence of compact sets covering $U$. Let $(X^n, T^n)$ be the local solution as above in the case $C = C_n$. If $T_n < \infty$ then $X^n_{T_n^-} \in U \setminus \operatorname{int}(C_n)$. Since the $C_n$ are increasing, we have that $T_n \uparrow \mathcal{T} = \sup_n T_n$.

Define the local solution $(X, \mathcal{T})$ by setting $X_t = X^n_t$ for all $t < T_n$. This works by uniqueness above.

Suppose that $(Y, \eta)$ is another solution on the same probability space. For each $n$ set $S_n = \inf\{t \le \eta : Y_t \notin C_n\} \wedge \eta$. By the uniqueness of the solution in each $C_n$, we have $X_t = Y_t$ for all $t < S_n \wedge T_n$, so $S_n \le T_n$. As $n \to \infty$, $S_n \uparrow \eta$ and $T_n \uparrow \mathcal{T}$, which implies $\eta \le \mathcal{T}$ and so $X_t = Y_t$ for all $t \le \eta$. Therefore $(X, \mathcal{T})$ is maximal.

Suppose $C_1, C_2 \subseteq U$ are compact with $C_1 \subseteq \operatorname{int}(C_2) \subseteq C_2 \subseteq U$. Let $\varphi : U \to \mathbb{R}$ be a $C^\infty$ function with $\varphi|_{C_1} = 1$ and $\varphi|_{U \setminus \operatorname{int}(C_2)} = 0$. Let $R_0 = \inf\{t \ge 0 : X_t \notin C_2\}$ and

$$\begin{aligned} S_n &= \inf\{t \ge R_{n-1} : X_t \in C_1\} \wedge \mathcal{T} \\ R_n &= \inf\{t \ge S_n : X_t \notin C_2\} \wedge \mathcal{T}. \end{aligned}$$

Let $N$ be the number of crossings that $X$ makes from $C_2$ to $C_1$. On $\{\mathcal{T} \le t, N \ge n\}$ we have that

$$\begin{aligned} \sum_{k=1}^n (\underbrace{\varphi(X_{R_k})}_{0} - \underbrace{\varphi(X_{S_k})}_{1}) &= -n \\ &= \int_0^t \sum_{k=1}^n \mathbb{1}_{(S_k, R_k]}(s)(\varphi'(X_s)\mathrm{d}X_s + \frac{1}{2}\varphi''(X_s)\mathrm{d}[X]_s) \quad \text{(Itô)} \\ &= \int_0^t (H^n_s \mathrm{d}B_s + K^n_s \mathrm{d}s) \\ &=: Z^n_t \end{aligned}$$

where $H^n, K^n$ are previsible, bounded uniformly in $n$. Then

$$n^2 \mathbb{1}\{\mathcal{T} \le t, N \ge n\} \le (Z^n_t)^2 \implies \mathbb{P}[\mathcal{T} \le t, N \ge n] \le \frac{1}{n^2}\mathbb{E}[(Z^n_t)^2].$$

Since $H^n, K^n$ are uniformly bounded in $n$ and $Z^n_t$ is defined by integrating $H^n, K^n$ over a time interval which doesn't depend on $n$ we have that $\mathbb{E}[(Z^n_t)^2] \le C$ where $C < \infty$ does not depend on $n$. Thus

$$\mathbb{P}[\mathcal{T} \le t, N \ge n] \le \frac{C}{n^2}$$

which tells us $\mathbb{P}[Y \leq t, N = \infty] = 0$ and $\mathbb{P}[Y < \infty, N = \infty] = 0$. Therefore the number of crossings that $X$ makes from $C_2$ to $C_1$ is almost-surely finite on the event $\{\mathcal{T} < \infty\}$. Since each crossing $X$ makes from $C_1$, $C_2$ almost-surely takes a positive amount of time (since $C_1$ is contained in the interior of $C_2$ and continuity of $X$), $\sup\{t < \mathcal{T} : X_t \in C_1\} < \mathcal{T}$ on $\{\mathcal{T} < \infty\}$. $\square$

**Example** (Bessel processes)**.** Fix $\nu \in \mathbb{R}$ and consider the SDE in $U = (0, \infty)$ given by

$$\mathrm{d}X_t = \mathrm{d}B_t + \frac{\nu - 1}{2X_t}\mathrm{d}t, \quad X_0 = x_0 \in U.$$

Then there exists a unique maximal solution $(X, \mathcal{T})$ in $U$ and $\mathcal{T} = \inf\{t \geq 0 : X_t = 0\}$. We call $(X, \mathcal{T})$ a *Bessel process of dimension* $\nu$.

Suppose that $\nu \in \mathbb{N}$, $\beta$ is a Brownian motion in $\mathbb{R}^\nu$ with $|\beta_0| = x_0 > 0$. Set $Y_t = |\mathcal{B}_t|$ and $\eta = \inf\{t \geq 0 : B_t = 0\}$. By the local Itô formula, we have that

$$\mathrm{d}Y_t = \frac{(\beta_t, \mathrm{d}\beta_t)}{|\beta_t|} + \frac{\nu - 1}{2|\beta_t|}\mathrm{d}t, \quad t < \eta$$

where $(\beta_t, \mathrm{d}\beta_t) = \sum_{i=1}^{\nu} \beta_t^i \mathrm{d}\beta_t^i$. Then the process $W_t = \int_0^t \frac{(\beta_s, \mathrm{d}\beta_s)}{|\beta_s|}$ is in $\mathcal{M}_{c,\mathrm{loc}}$. Moreover,

$$\mathrm{d}[W]_t = \frac{1}{|\beta_t|^2} \sum_{i,j=1}^{\nu} \beta_t^i \beta_t^j \mathrm{d}[\beta^i, \beta^j]_t = \mathrm{d}t$$

and so the Lévy characterisaton shows $W$ is a standard Brownian motion. Thus

$$\mathrm{d}Y_t = \mathrm{d}W_t + \frac{\nu - 1}{2Y_t}\mathrm{d}t, \quad t < \eta.$$

A Bessel process of dimension $\nu$ describes the time evolution of the norm of a $\nu$-dimensional Brownian motion up to when it first hits 0.

## Diffusion processes

Suppose that $a : \mathbb{R}^d \to M^{d\times d}(\mathbb{R})$, $b : \mathbb{R}^d \to \mathbb{R}^d$ are bounded, measurable, $a$ is symmetric (i.e $a(x)$ is symmetric for all $x \in \mathbb{R}^d$). For $f \in C_b^2(\mathbb{R}^d)$ ($C^2$ with bounded derivatives) set

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

Let $X$ be a continuous, adapted process in $\mathbb{R}^d$. We say that $X$ is an *$L$-diffusion* if for all $f \in C_b^2(\mathbb{R}^d)$ we have that

$$M_t^f := f(X_t) - f(x_0) - \int_0^t Lf(X_s)\mathrm{d}s$$

is a martingale. We call the coefficient $a$ the *diffusion* and $b$ the *drift*.

**Example.** Suppose $\sigma, b$ are constant and $a = \sigma\sigma^T$. Let $B$ be a standard Brownian motion on $\mathbb{R}^d$. Then

$$X_t = \sigma B_t + b_t$$

is a $(a, b)$ diffusion. In the case that $\sigma = I$, $b = 0$, $X_t = B_t$ we have a $L$-diffusion where $L = \frac{1}{2}\Delta$.

**Proposition.** Suppose that $X$ solves

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

Let $f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ (bounded derivatives, $C^1$ in first variable and $C^2$ in the second variable). Then

$$M_t^f = f(t, X_t) - f(0, X_0) - \int_0^t \left(\frac{\partial}{\partial s} + L\right) f(s, X_s)\mathrm{d}s$$

is a continuous local martingale, for $a = \sigma\sigma^T$, $L$ as above. If in addition $a, b$ are bounded then $X$ is an $L$-diffusion.

*Proof.* Example Sheet 4. □

Question: which choices of $a$ can be written as $\sigma\sigma^T$ for such $\sigma$?

Suppose that $a, b$ are Lipschitz, bounded and there exists $\varepsilon > 0$ so that

$$(a(x)\xi, \xi) \geq \varepsilon^2 |\xi|^2 \quad \forall x, \xi \in \mathbb{R}^d.$$

If $a$ satisfies this we say it is *uniformly positive definite* (UPD). Then there exists $\sigma : \mathbb{R}^d \to M^{d\times d}(\mathbb{R})$ with $\sigma\sigma^T = a$.

For $d = 1$ we can just take $\sigma = \sqrt{a}$. For $d \geq 2$ we can write $a(x) = u(x)\lambda(x)u^T(x)$ where $\lambda(x)$ is the diagonal matrix of eigenvalues of $a$ and $u$ is the orthogonal matrix with columns the eigenvectors of $a$. Take $\sigma(x) = u(x)\sqrt{\lambda(x)}u^T(x)$. Then $\sigma$ is Lipschitz from differentiability of the square root map on the set of UPD matrices (and the fact $a$ is Lipschitz). For such $\sigma, b$ the SDE

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

has a unique strong solution which is an $(a, b)$-diffusion.

**Proposition.** Let $X$ be an $L$-diffusion and $T$ a finite stopping time. Set $\tilde{X}_t = X_{T+t}$ and $\tilde{\mathcal{F}}_t = \mathcal{F}_{T+t}$. Then $\tilde{X}$ is an $L$-diffusion with respect to the filtration $(\tilde{\mathcal{F}}_t)$.

*Proof.* Fix $f \in C_b^2(\mathbb{R}^d)$. Consider the process

$$\tilde{M}_t^f := f(\tilde{X}_t) - f(\tilde{X}_0) - \int_0^t Lf(\tilde{X}_s)\mathrm{d}s.$$

Then $\tilde{M}^f$ is adapted to $(\tilde{\mathcal{F}}_t)$ and is integrable. For $A \in \tilde{\mathcal{F}}_s$ and $n \geq 0$ we have

$$\begin{aligned}
\mathbb{E}[(\tilde{M}_t^f - \tilde{M}_s^f)\mathbb{1}_{A\cap\{T\leq n\}}] &= \mathbb{E}[(M_{t+T}^f - M_{s+T}^f)\mathbb{1}_{A\cap\{T\leq n\}}] \\
&= \mathbb{E}[(M_{t+T\wedge n}^f - M_{s+T\wedge n}^f)\mathbb{1}_{A\cap\{T\leq n\}}] \\
&= 0
\end{aligned}$$

by the OST. By sending $n \to \infty$ and applying the DCT we see $\mathbb{E}[\tilde{M}_t^f|\tilde{\mathcal{F}}_s] = \tilde{M}_s^f$. □

**Lemma.** *Let $X$ be an $L$-diffusion. Then for all $f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ the process*

$$M_t^f := f(t, X_t) - f(0, X_0) - \int_0^t \left(\frac{\partial}{\partial s} + L\right) f(s, X_s)\mathrm{d}s$$

*is a martingale.*

*Proof.* Fix $T > 0$ and consider

$$Z_n = \sup_{\substack{0\leq s\leq t\leq T \\ t-s\leq 1/n}} |\dot{f}(s, X_t) - \dot{f}(s, X_s)| + \sup_{\substack{0\leq s\leq t\leq T \\ t-s\leq 1/n}} |Lf(s, X_t) - Lf(t, X_t)|.$$

Then $Z_n$ is bounded ($f$ bounded with bounded derivatives), $Z_n \to 0$ as $n \to \infty$ (by continuity). The bounded convergence theorem tells us $\mathbb{E}Z_n \to 0$ as $n \to \infty$. We have

$$\begin{aligned} M_t^f - M_s^f = &\left( f(t, X_t) - f(s, X_t) - \int_s^t \dot{f}(r, X_t) \mathrm{d}t \right) \\ &+ \left( f(s, X_t) - f(s, X_s) - \int_s^t Lf(s, X_r) \mathrm{d}r \right) \\ &+ \left( \int_s^t (\dot{f}(r, X_t) - \dot{f}(r, X_r)) \mathrm{d}r \right) \\ &+ \int_s^t (Lf(s, X_r) - Lf(r, X_r)) \mathrm{d}r. \end{aligned}$$

Choose $s_0 < s_1 < \ldots < s_m$ with $s_0 = s$, $s_m = t$ and $s_{k+1} - s_k \leq 1/n$ for each $k$. The FTC shows the first term is 0. For the second line note that its $\mathcal{F}_s$-conditional expectation is 0 since $X$ is an $L$-diffusion.

For the last two lines, we have that

$$\mathbb{E}[|\mathbb{E}[M_{s_{k+1}}^f - M_{s_k}^f | \mathcal{F}_s]|] \leq (s_{k+1} - s_k) \mathbb{E}Z_n$$

which implies

$$\mathbb{E}[|\mathbb{E}[M_t^f - M_s^f | \mathcal{F}_s]|] \leq (t - s) \mathbb{E}[Z_n] \xrightarrow{n \to \infty} 0.$$

□

## The Dirichlet & Cauchy problems

### The Dirichlet Problem

Assume that $a, b$ are Lipschitz, $a$ UPD. Let $D \subseteq \mathbb{R}^d$ be a bounded open domain with smooth boundary. We will assume the following theorem.

**Theorem** (Dirichlet problem)**.** *For all $f \in C(\partial D)$, $\varphi \in C(\overline{D})$ there exists a unique function $u \in C(\overline{D}) \cap C^2(D)$ such that*

$$\begin{cases} Lu + \varphi = 0 & \text{in } D \\ u - f & \text{on } \partial D \end{cases}.$$

*Moreoever there exist continuous functions $m : D \times \partial D \to (0, \infty)$, $g : \{(x, y) \in D \times D : x \neq y\} \to (0, \infty)$ such that for all $f, \varphi$ as above, we have that*

$$u(x) = \int_D g(x, y) \varphi(y) \mathrm{d}y + \int_{\partial D} f(y) m(x, y) \lambda(\mathrm{d}y)$$

*where $\lambda$ denotes the natural boundary measure. We call $g$ the Green kernel and $m(x, y)\lambda(\mathrm{d}y)$ is called the harmonic measure on $\partial D$ as seen from $x$.*

We will now prove the following.

**Theorem.** *Suppose that $u \in C(\overline{D}) \cap C^2(D)$ satisfies*

$$\begin{cases} Lu + \varphi = 0 & \text{in } D \\ u - f & \text{on } \partial D \end{cases}$$

*where $f \in C(\partial D)$, $\varphi \in C(\overline{D})$. Then for any $L$-diffusion $X$ starting from $x \in D$ we have that*

$$u(x) = \mathbb{E}_x \left[ \int_0^T \varphi(X_s) \mathrm{d}s + f(X_T) \right]$$

*where $T = \inf\{t \geq 0 : X_t \notin D\}$. Moreover for all Borel sets $A \subseteq D$, $B \subseteq \partial D$,*

$$\mathbb{E}_x \left[ \int_0^T \mathbb{1}(X_s \in A) \mathrm{d}s \right] = \int_A g(x,y) \mathrm{d}y \text{ and}$$
$$\mathbb{P}_x[X_T \in B] = \int_B m(x,y) \lambda(\mathrm{d}y).$$

*Proof.* Fix $n \geq 1$ and let $T_n = \inf\{t \geq 0 : X_t \notin D_n\}$ where $D_n = \{x \in D : \operatorname{dist}(x, D^c) > 1/n\}$. Consider

$$M_t := u(X_{t \wedge T_n}) - u(X_0) + \int_0^{t \wedge T_n} \varphi(X_s) \mathrm{d}s.$$

There exists $\tilde{u} \in C_b^2(\mathbb{R}^d)$ with $u = \tilde{u}$ on $D_n$. Then $M = \tilde{M}^{T_n}$ where

$$\tilde{M}_t := \tilde{u}(X_t) - \tilde{u}(X_0) - \int_0^t L\tilde{u}(X_s) \mathrm{d}s.$$

Since $X$ is an $L$-diffusion, $\tilde{M}$ is a martingale. The OST implies $M$ is also a martingale so

$$u(X) = \mathbb{E}_x \left[ u(X_{t \wedge T_n}) + \int_0^{t \wedge T_n} \varphi(X_s) \mathrm{d}s \right]. \qquad (*)$$

We want to send $n \to \infty$. First we need to show $\mathbb{E}_x T < \infty$. Take $\varphi = 1$, $f = 0$ and $u^{1,0}$ the corresponding solution. Then $(*)$ holds for $u^{1,0}$ so

$$\mathbb{E}_x[T_n \wedge t] = u^{1,0}(x) - \mathbb{E}_x[u^{1,0}(X_{t \wedge T_n})].$$

Since $u^{1,0}$ is bounded, $T^n \uparrow T$ as $n \to \infty$ the MCT implies $\mathbb{E}_x[T] < \infty$ (by sending $n \to \infty$ and $t \to \infty$).

Now return to the general case in $(*)$. Have $T_n \wedge t \uparrow T$ as $n, t \to \infty$. Since $u$ is continuous on $\overline{D}$, $u(X_{t \wedge T_n}) \to f(X_T)$ as $n, t \to \infty$. Since $u$ is bounded on

$\overline{D}$ (by compactness) the bounded convergence theorem gives $\mathbb{E}_x[u(X_{t\wedge T_n})] \to \mathbb{E}_x[f(X_T)]$ as $t, n \to \infty$. Moreover

$$\mathbb{E}_x\left[\int_0^T |\varphi(X_s)|\mathrm{d}s\right] \leq \|\varphi\|_\infty \mathbb{E}_x T < \infty$$

so the DCT implies

$$\mathbb{E}_x\left[\int_0^{t\wedge T_n} \varphi(X_s)\mathrm{d}s\right] \to \mathbb{E}_x\left[\int_0^T \varphi(X_s)\mathrm{d}s\right].$$

Thus

$$u(x) = \mathbb{E}_x\left[f(X_T) + \int_0^T \varphi(X_s)\mathrm{d}s\right].$$

The final assertions follow by taking limits as $\varphi_n \to \mathbb{1}_A$, $f = 0$ and $f_n \to \mathbb{1}_B$, $\varphi = 0$. □

**The Cauchy Problem**

We assume the following.

**Theorem.** *For each $f \in C_b^2(\mathbb{R}^d)$ there exists a unique solution $u \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ such that*

$$\begin{cases} \frac{\partial u}{\partial t} = Lu & \text{on } \mathbb{R}_+ \times \mathbb{R}^d \\ u(0,\cdot) = f(\cdot) & \text{on } \mathbb{R}^d \end{cases}.$$

*Moreover there exists a continuous function $p : (0,\infty) \times \mathbb{R}^d \times \mathbb{R}^d \to (0,\infty)$ such that*

$$u(t,x) = \int_{\mathbb{R}^d} P(t,x,y) f(y)\mathrm{d}y \qquad \forall \mathbb{R}_+ \times \mathbb{R}^d.$$

*We call $p$ the "heat kernel".*

We will prove:

**Theorem.** *Assume that $f \in C_b^2(\mathbb{R}^d)$. Let $u$ satisfy*

$$\begin{cases} \frac{\partial u}{\partial t} = Lu & \text{on } \mathbb{R}_+ \times \mathbb{R}^d \\ u(0,\cdot) = f(\cdot) & \text{on } \mathbb{R}^d \end{cases}.$$

*Then for any $L$-diffusion $X$ starting from $x$, for all $t \in \mathbb{R}_+$, $0 \leq s \leq t$ we have that*

$$\mathbb{E}_x[f(X_t)|\mathcal{F}_s] = u(t-s, X_s) \text{ almost-surely.}$$

*In particular* $\mathbb{E}_x[f(X_t)] = u(t,x) = \int_{\mathbb{R}^d} p(t,x,y)f(y)\mathrm{d}y$. *Finally, under the measure* $\mathbb{P}_x$*, the finite dimensional distributions of* $X$ *are given by*

$$\mathbb{P}_x[X_{t_1} \in \mathrm{d}x_1, \dots, X_{t_n} \in \mathrm{d}x_n] = p(t_1, X_0, X_1) \dots p(t_n - t_{n-1}, X_{n-1}, X_n)\mathrm{d}x_1 \dots \mathrm{d}x_n$$

*for* $0 < t_1 < t_2 < \dots, t_n < \infty$*,* $x_1, \dots, x_n \in \mathbb{R}^d$*,* $x_0 = x$*.*

*Proof.* Fix $t \in (0, \infty)$. Consider $g(s,x) = u(t-s,x)$ for $s \le t$, $x \in \mathbb{R}^d$. Note that $\left(\frac{\partial}{\partial s} + L\right) g(s,x) = -\dot{u}(t-s,x) + Lu(t-s,x) = 0$. Therefore

$$\begin{aligned} M_s^g &:= g(s, X_t) - g(0, X_0) - \int_0^t \left(\frac{\partial}{\partial s} + L\right) g(s,x)\mathrm{d}s \\ &= g(s, X_t) - g(0, X_0) \end{aligned}$$

is a martingale. Therefore $\mathbb{E}[M_t^g|\mathcal{F}_s] = M_s^g$ so $\mathbb{E}[f(X_t)|\mathcal{F}_s] = u(t-s, X_s)$. Taking $s = 0, x_0 = x$ we see $u(t,x) = \mathbb{E}_x[f(X_t)]$.

Now to show the second part, set $P_t f(x) = \int_{\mathbb{R}^d} p(t,x,y)f(y)\mathrm{d}y = u(t,x)$. By the uniqueness of solutions to the Cauchy problem, $P_s(P_t f) = P_{s+t} f$. We claim (by induction) that

$$\begin{aligned} &\mathbb{E}_{x_0}\left[\prod_{i=1}^n f_i(X_{t_i})\right] \\ &= \int_{(\mathbb{R}^d)^n} p(t_1, x_0, x_1) f_1(x_1) \dots p(t_n - t_{n-1}, x_{n-1}, x_n) f_n(x_n) \mathrm{d}x_1 \dots \mathrm{d}x_n. \end{aligned}$$

We know this is true for $n = 1$ by the above. For the inductive step, use that

$$\begin{aligned} \mathbb{E}_{x_0}\left[\prod_{i=1}^n f_i(X_{t_i})|\mathcal{F}_{t_{n-1}}\right] &= \left(\prod_{i=1}^{n-1} f_i(X_{t_i})\right) \mathbb{E}[f_n(X_{t_n})|\mathcal{F}_{n-1}] \\ &= \left(\prod_{i=1}^{n-1} f_i(X_{t_i})\right) P_{t_n - t_{n-1}} f_n(X_{t_{n-1}}) \end{aligned}$$

and apply the case $n-1$. ☐