STOCHASTIC CALCULUS

JASON MILLER

Contents

Preface

These lecture notes are for the University of Cambridge Part III course Stochastic Calculus, given Lent 2016. The contents are very closely based on a set of lecture notes for this course due to Christina Goldschmidt. Please notify jpmiller@statslab.cam.ac.uk for corrections.

1. Introduction

1.1. **Overview.** In ordinary calculus, one learns how to integrate, differentiate, and solve ordinary differential equations. In this course, we will develop the theory for the stochastic analogs of these constructions: the Itô integral, the Itô derivative, and stochastic differential equations (SDEs).

1.2. **Motivating example.** Fix $\epsilon > 0$. Suppose that:

- S_t is the value of an asset at time $t \geq 0$
- X_t^ϵ is the total amount of the asset held at time $t \geq 0$
- V_t^ϵ is the total value of the portfolio, assuming always invested in the asset at time $t \geq 0$

Assume that X_t^ϵ only changes every $\epsilon > 0$ units of time. That is, $X^\epsilon|_{[k\epsilon,(k+1)\epsilon]}$ is constant for each $k \in \{0, 1, \ldots\}$.

Then we have that

$$V_t^\epsilon - V_0^\epsilon = \sum_{k=0}^{\lfloor t\epsilon \rfloor} X_{k\epsilon}^\epsilon (S_{(k+1)\epsilon} - S_{k\epsilon}).$$

Assume for the moment that S_t is a smooth function of t. By Taylor's formula, we have that $S_{(k+1)\epsilon} = S_{k\epsilon} + \epsilon S'_{k\epsilon} + o(\epsilon)$ as $\epsilon \to 0$. Therefore

$$V_t^\epsilon - V_0^\epsilon = \sum_{k=0}^{\lfloor t/\epsilon \rfloor} X_{k\epsilon}^\epsilon \cdot \epsilon S'_{k\epsilon} + o(1). \tag{1.1}$$

Assume that $X^\epsilon \to X$ as $\epsilon \to 0$. Then (1.1) converges to

$$V_t - V_0 = \int_0^t X_u S'_u du = \int_0^t X_u dS_u.$$

Remark 1.1. *This construction does not require that S is smooth, only that it has finite variation (we will review this precisely later) and $\int_0^t X_u dS_u$ is a Lebesgue-Stieljes integral.*

Goal: Make sense of the same type of integral when S is a random process, with the main example being $S = B$ where B is a standard Brownian motion.

Recall from Advanced probability that B does not have smooth sample paths, so it is not possible to construct $\int_0^t X_u dB_u$ as above. (In fact, B is only α-Hölder continuous for each $\alpha \in (0, 1/2)$: $|B_s - B_t| \leq C|s-t|^\alpha$.)

As we will see later in the course, the reason that the approximations to the integral $\int_0^t X_u dB_u$ do converge even though B is not smooth (or of finite variation) is due to cancellation. In particular, assuming that X^ϵ is deterministic, we have that

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{k=0}^{\lfloor t/\epsilon \rfloor} X_{k\epsilon}^\epsilon (B_{(k+1)\epsilon} - B_{k\epsilon})\right)^2\right] \\
=&\mathbb{E}\left[\sum_{k=0}^{\lfloor t/\epsilon \rfloor} (X_{k\epsilon}^\epsilon)^2 \left(B_{(k+1)\epsilon} - B_{k\epsilon}\right)^2 + \sum_{j \neq k}^{\lfloor t/\epsilon \rfloor} X_{j\epsilon} X_{k\epsilon} (B_{(k+1)\epsilon} - B_{k\epsilon})(B_{(j+1)\epsilon} - B_{j\epsilon})\right] \\
=&\sum_{k=0}^{\lfloor t/\epsilon \rfloor} (X_{k\epsilon}^\epsilon)^2 \cdot \epsilon \to \int_0^t X_s^2 ds \quad \text{as} \quad \epsilon \to 0.
\end{aligned}$$

So, even though B is not smooth, the approximations to the integral do not blow up. We will construct $\int_0^t X_u dB_u$ more carefully later in the course.

The tools developed in this course will lead us to:

(I) A general theory of stochastic integration
(II) Study the properties of Brownian motion and continuous time martingales
(III) Stochastic differential equations
(IV) Convergence of discrete processes to solutions of SDEs

We will also see that Brownian motion is closely related to PDEs involving Δ. More generally, SDEs are closely related to PDEs involving second order elliptic operators.

Remark 1.2. *The motivating example was about finance. Stochastic calculus, however, is a very important tool in many other areas of probability.*

2. Preliminaries

2.1. **Finite variation integrals.** Recall that a function $a\colon [0,\infty) \to \mathbb{R}$ is said to be *cádlág* if it is right-continuous and has left hand limits. In other words, for all $x \in (0,\infty)$ we have both

$$\lim_{y\to x^-} a(y) \quad \text{exists} \quad \text{and} \quad \lim_{y\to x^+} a(y) = a(x).$$

Assume that a is non-decreasing and cadlag with $a(0)=0$. Then there exists a unique Borel measure da on $[0,\infty)$ such that for all $s<t$, we have that $da((s,t]) = t-s$. The Lebesgue-Stieljes integral $f \cdot a$ is defined as

$$f \cdot a(t) = \int_0^t f(s)da(s). \tag{2.1}$$

If a is the difference of two non-decreasing functions a_1, a_2, then the definition (2.1) extends by setting

$$f \cdot a(t) = f \cdot a_1(t) - f \cdot a_2(t)$$

provided both terms on the right hand side are finite.

We will now characterize when a function a can be written as the difference of two non-decreasing functions.

For each $n \in \mathbb{N}$ and $t \geq 0$, we let

$$V^n(t) = \sum_{k=0}^{\lceil 2^n t\rceil -1} \left| a((k+1)2^{-n}) - a(k2^{-n}) \right|.$$

Proposition 2.1. *The function V^n has a limit V as $n \to \infty$. Moreover, a can be expressed as the difference of two non-decreasing functions if and only if $V(t) < \infty$ for all $t \geq 0$.*

$V(t)$ is the *total variation* of a on $(0,t]$. If $V(t)<\infty$, then a is said to have *finite variation* on $(0,t]$.

Proof of Proposition 2.1. Let $t_n^+ = 2^{-n}\lceil 2^n t\rceil$ and $t_n^- = 2^{-n}(\lceil 2^n t\rceil - 1)$. Then we have that

$$V^n(t) = \sum_{k=0}^{2^n t_n^- -1} \left| a((k+1)2^{-n}) - a(k2^{-n}) \right| + |a(t_n^+) - a(t_n^-)|. \tag{2.2}$$

By the triangle inequatilty, the first summand in the definition of $V^n(t)$ is non-decreasing in n. Consequently, it has a limit as $n \to \infty$. By the cádlág property, the second summand converges to $|a(t) - a(t^-)| = |\Delta a(t)|$ as $n \to \infty$. Therefore $V^n(t)$ has a limit $V(t)$ as $n\to\infty$. Moreover, it is not difficult to see that V is both non-decreasing and cádlág.

Let

$$a^+ = \frac{1}{2}(V+a) \quad \text{and} \quad a^- = \frac{1}{2}(V-a).$$

Then a^+, a^- are cádlág as V, a are cádlág and $a = a^+ - a^-$. We will now argue that a^+, a^- are non-decreasing. We will just give the prove in the case of a^+ as the argument in the case of a^- is analogous. Fix $s < t$. Then we have that

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

Thus if $V(t) < \infty$, then a can be written as a difference of non-decreasing functions. The reverse implication is obvious. □

Suppose that $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is a filtered probability space. A *cádlág adapted process* X is a map $X \colon \Omega \times [0,\infty) \to \mathbb{R}$ such that:

(1) $X_t = X(\cdot, t)$ is $\mathcal{F}_t$-measurable for all $t \geq 0$
(2) $X(\omega, \cdot) \colon [0,\infty) \to \mathbb{R}$ is cádlág for all $\omega \in \Omega$

If A is a cádlág, adapted process then we can define its total variation process V for each fixed ω. In this case, V is cádlág and non-decreasing. It is also adapted. Indeed, if we let

$$\widetilde{V}_t^n = \sum_{k=0}^{2^n t_n^- - 1} |A_{(k+1)2^{-n}} - A_{k2^{-n}}|, \quad \text{where} \quad t_n^- = 2^{-n}(\lceil 2^n t \rceil - 1)$$

then

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

is $\mathcal{F}_t$-measuarble as it is a limit of $\mathcal{F}_t$-measurable functions.

We can define the stochastic integral against a process of finite variation by

$$(H \cdot A)(\omega, t) = \int_0^t H(\omega, s) dA(\omega, s).$$

This integral is defined for each ω. For our later purposes, we will want $H \cdot A$ to be both cádlág and adapted. This constrains the possibilities for h.

2.2. Previsible processes.

Definition 2.2. *The previsible σ-algebra $\mathcal{P}$ on $\Omega \times [0,\infty)$ is the σ-algebra generated by sets of the form $A \times (s,t]$ for $A \in \mathcal{F}_s$ and $s < t$. A process $H \colon \Omega \times [0,\infty) \to \mathbb{R}$ is called previsible if it is $\mathcal{P}$-measurable.*

Proposition 2.3. *Let X be a cádlág, adapted process. Then $H_t = X_{t^-} = \lim_{s\to t^-} X_s$ is a previsible process.*

Proof. Since X is cádlág and adapted, we have that $H \colon \Omega \times [0,\infty) \to \mathbb{R}$ is left-continuous and adapted. Set $t_n^- = k2^{-n}$ for $2^{-n}k < t \leq (k+1)2^{-n}$ and let

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

As $H_{k2^{-n}}$ is $\mathcal{F}_{k2^{-n}}$-measurable, it follows that H_t^n is previsible.

Since $t_n^- \uparrow t$ as $n \to \infty$, it follows that $H_t^n \to H_t$ as $n \to \infty$ by left-continuity. Therefore, H is previsible. □

Proposition 2.4. *Suppose that H is a previsible process. Then H_t is $\mathcal{F}_{t^-}$-measurable for all $t > 0$ where $\mathcal{F}_{t^-} = \sigma(\mathcal{F}_s : s < t)$.*

Proof. See Example Sheet 1. □

Note that Brownian motion is previsible as it is a continuous process. On the other hand, a Poisson process (N_t) is *not* previsible since N_t is not $\mathcal{F}_{t^-}$-measurable.

Proposition 2.5. *Let A be a cádlág, adapted, finite variation process with associated total variation process V. Let H be a previsible process such that for all $t \geq 0$, $\omega \in \Omega$, we have that*

$$\int_0^t |H(\omega, s)| dV(\omega, s) < \infty.$$

Then the process $(H \cdot A)_t = \int_0^t H_s dA_s$ is cádlág, adapted, and of finite variation.

Proof. Step 1: cádlág. Note that $\mathbf{1}_{(0,s]} \to \mathbf{1}_{(0,t]}$ as $s \downarrow t$ and $\mathbf{1}_{(0,s]} \to \mathbf{1}_{(0,t)}$ as $s \uparrow t$. Recall that $(H \cdot A)_t = \int H_s \mathbf{1}(s \in (0,t]) dA_s$. By the dominated convergence theorem (with ω fixed), we have that

$$\begin{aligned}
(H \cdot A)_t &= \int H_s \lim_{r \to t^+} \mathbf{1}(s \in (0,r]) dA_s \\
&= \lim_{r \to t^+} \int H_s \mathbf{1}(s \in (0,r]) dA_s \\
&= \lim_{r \to t^+} (H \cdot A)_r.
\end{aligned}$$

Therefore $H \cdot A$ is right-continuous. A similar argument implies that A has left-hand limits.

Moreover, we have that

$$\Delta (H \cdot A)_t = \int H_s \mathbf{1}(s = t) dA_s = H_t \Delta A_t.$$

Step 2: adapted. Suppose that $H = \mathbf{1}_{B \times (s,u]}$ where $B \in \mathcal{F}_s$ and $u > s$. Then we have that

$$(H \cdot A)_t = \mathbf{1}_B \left(A_{t \wedge u} - A_{t \wedge s} \right),$$

which is $\mathcal{F}_t$-measurable. Let

$$\mathcal{A} = \{ C \in \mathcal{P} : \mathbf{1}_C \cdot A \quad \text{is adapted to} \quad (\mathcal{F}_t) \}$$

and let

$$\Pi = \{ B \times (s,u] : B \in \mathcal{F}_s,\ s < u \}.$$

Then Π is a π-system and $\Pi \subseteq \mathcal{A}$. We have shown so far that $\Pi \subseteq \mathcal{A}$.

It is easy to see that $\mathcal{A}$ is a d-system, so by Dynkin's lemma we have that $\mathcal{P} = \sigma(\Pi) \subseteq \mathcal{A} \subseteq \mathcal{P}$. Therefroe $\mathcal{A} = \mathcal{P}$.

Suppose that $H \geq 0$ is previsible. Set

$$\begin{aligned} H^n &= (2^{-n} \lfloor 2^n H \rfloor) \wedge n \\ &= \sum_{k=1}^{2^n n - 1} 2^{-n} k \mathbf{1}(H \in [2^{-n}k, 2^{-n}(k+1))) + n \mathbf{1}(H > n). \end{aligned}$$

Thus $(H^n \cdot A)_t$ is $\mathcal{F}_t$-measurable for all t. By the monotone convergence theorem, we have that $(H^n \cdot A)_t \to (H \cdot A)_t$ as $n \to \infty$. Therefore $(H \cdot A)_t$ is $\mathcal{F}_t$-measurable. Extending in the usual way to the case of general H is straightforward, and therefore $H \cdot A$ is adapted.

Step 3: finite variation. Let $H^+ = \max(H, 0)$, $H^- = \max(-H, 0)$, $A^+ = \frac{1}{2}(V + A)$, $A^- = \frac{1}{2}(V - A)$ so that $H = H^+ - H^-$ and $A = A^+ - A^-$. Then we have that

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

is a difference of non-decreasing processes, hence of finite variation. □

2.3. **Local martingales.** Suppose that $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ is a filtered probability space and that $(\mathcal{F}_t)$ satisfies the *usual conditions*. In other words,

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

Recall that an integrable, adapted process X_t is called a martingale if $\mathbb{E}[X_t \,|\, \mathcal{F}_s] = X_s$ for all $s \leq t$. Similarly, X_t is called a submartingale (resp. supermartingale) if $\mathbb{E}[X_t \,|\, \mathcal{F}_s] \geq X_s$ (resp. $\mathbb{E}[X_t \,|\, \mathcal{F}_s] \leq X_s$) for all $s \leq t$.

A map $T \colon \Omega \to [0, \infty]$ is called a *stopping time* if $\{T \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$. If X is cádlág and adapted and T is a finite stopping time, then X_T is $\mathcal{F}_T$-measurable where

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t \quad \forall t \geq 0\}.$$

Theorem 2.6 (Optional stopping, OST)**.** *Let X be a cádlág, adapted process. TFAE:*

(a) X is a martingale.
(b) $X_t^T := X_{t \wedge T}$ is a martingale for all stopping times T.
(c) For all stopping times S, T with T bounded, $\mathbb{E}[X_T \,|\, \mathcal{F}_S] = X_{S \wedge T}$.
(d) $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ for all bounded stopping times T.

Definition 2.7. *A cádlág, adapted process X is called a* local martingale *if there exists a sequence (T_n) of stopping times such that $T_n \uparrow \infty$ as $n \to \infty$ and X^{T_n} is a martingale for every n. In this case, we say that (T_n)* reduces *X.*

Remark 2.8. *Every martingale is a local martingale by the OST.*

We will now give an important example which will be important to keep in mind. Let B be a standard Brownian motion in $\mathbb{R}^3$. Let $M_t = 1/|B_t|$. Recall from Example Sheet 3, Exercise 7.7 of Advanced Probability that:

- $(M_t)_{t \geq 1}$ is bounded in L^2. That is, $\sup_{t \geq 1} \mathbb{E}[M_t^2] < \infty$.
- $\mathbb{E}[M_t] \to 0$ as $t \to \infty$.
- M is a supermartingale.

For each n, let $T_n = \inf\{t \geq 1 : 1/|B_t| < 1/n]\} = \inf\{t \geq 1 : |M_t| > n\}$. We will now argue that $(M_t^{T_n})_{t\geq 1}$ is a martingale for all n and $T_n \uparrow \infty$ a.s. That is, $(M_t)_{t\geq 1}$ is a local martingale.

We begin by noting that if $n \leq M_1(\omega)$ then $T_n(\omega) = 1$ and if $n > M_1(\omega)$ then $T_n(\omega) > 1$. Moreover, since $|B_t|$ cannot hit $1/(n+1)$ without first having hit $1/n$, it follows that $T_n(\omega)$ is non-decreasing in n.

Recall from Advanced Probability that if $f \in C^2$ with bounded, continuous derivatives, then

$$f(B_t) - f(B_0) - \frac{1}{2}\int_0^t \Delta f(B_s) ds, \quad t \geq 0,$$

is a martingale. Noting that $f(x) = 1/|x|$ is harmonic in $\mathbb{R}^3$ for $x \neq 0$, we therefore see that $(M_t^{T_n})_{t\geq 1}$ is a martingale. To finish showing that $(M_t)_{t\geq 1}$ is a local martingale, we need to show that $T_n \uparrow \infty$ as $n \to \infty$. To see this, for each $R > 0$ we let $S_R = \inf\{t \geq 1 : |B_t| > R\} = \inf\{t \geq 1 : M_t < 1/R\}$. By the OST, we have that

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

We can also rewrite the left hand side as

$$n\mathbb{P}[T_n < S_R] + \frac{1}{R}\mathbb{P}[S_R \leq T_n].$$

Using that $\mathbb{P}[S_R \leq T_n] = 1 - \mathbb{P}[T_n < S_R]$, we can solve for $\mathbb{P}[T_n < S_R]$ in the above to get that

$$\mathbb{P}[T_n < S_R] = \frac{\mu - 1/R}{n - 1/R} \to \frac{\mu}{n} \quad \text{as} \quad R \to \infty.$$

Let $\underline{B} = \inf_{t\geq 1} |B_t|$. Then it follows that $\mathbb{P}[\underline{B} \leq 1/n] \leq \mu/n$. Therefore $T_n \to \infty$ a.s. as $n \to \infty$.

Since $\mathbb{E}[M_t] \to 0$ as $t \to \infty$, it follows that $(M_t)_{t\geq 1}$ is not a martingale.

In summary, $(M_t)_{t\geq 1}$ is:

- a local martingale but not a martingale,
- a supermartinagle, and
- L^2 bounded.

We need a stronger condition than L^2-boundedness for a local martingale to be a martingale, which we will come back to later.

Recall that a set $\mathcal{X}$ of random variables is said to be uniformly integrable (UI) if

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

Lemma 2.9. *If $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, then the set*

$$\mathcal{X} = \{\mathbb{E}[X \mid \mathcal{G}] : \mathcal{G} \subseteq \mathcal{F} \quad \textit{is a } \sigma\textit{-algebra}\}$$

is UI.

Proof. See Example Sheet 1. □

Proposition 2.10. *TFAE:*

(a) X is a martingale
(b) X is a local martingale and, for all $t \geq 0$, the set

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

is UI.

Proof. Suppose that (a) holds. By the OST, if T is a stopping time with $T \le t$, then we have that $X_T = \mathbb{E}[X_t \,|\, \mathcal{F}_T]$. Lemma 2.9 then implies that $\mathcal{X}_t$ is UI.

Suppose that (b) holds and let (T_n) be a reducing sequence for X. Then, for all bounded stopping times $T \le t$,

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

As $\{X_{T\wedge T_n} : n \ge 0\}$ is UI, and $X_{T\wedge T_n} \to X_T$ as $n \to \infty$, it follows that $\mathbb{E}[X_{T\wedge T_n}] \to \mathbb{E}[X_T]$. That is, $\mathbb{E}[X_0] = \mathbb{E}[X_T]$. Therefore X is a martingale by the OST. $\square$

Remark 2.11. *(i) Every bounded, local martingale is a martingale.*
(ii) If there exists $Z \in L^1$ such that $|X_t| \le Z$ for all t, then X is a martingale

Proposition 2.12. *Suppose that X is a local martingale such that $X_t \ge 0$ for all t. Then X is a supermartingale.*

Proof. Let (T_n) be a reducing sequence. Then, for all $s \le t$ and for all $n \in \mathbb{N}$, we have that $X_{s\wedge T_n} = \mathbb{E}[X_{t\wedge T_n} \,|\, \mathcal{F}_s]$. Therefore,

$$X_s = \liminf_{n\to\infty} X_{s\wedge T_n} = \liminf_{n\to\infty} \mathbb{E}[X_{t\wedge T_n} \,|\, \mathcal{F}_s] \ge \mathbb{E}[\liminf_{n\to\infty} X_{t\wedge T_n} \,|\, \mathcal{F}_s] = \mathbb{E}[X_t \,|\, \mathcal{F}_s].$$

$\square$

Proposition 2.13. *Let M be a continuous, local martingale with $M_0 = 0$. Set $S_n = \inf\{t \ge 0 : |M_t| = n\}$. Then S_n is a stopping time for all $n \in \mathbb{N}$, $S_n \uparrow \infty$ as $n \to \infty$, and M^{S_n} is a martingale for all n. That is, (S_n) reduces M.*

Proof. Note that

$$\{S_n \le t\} = \bigcap_{k\in\mathbb{N}} \bigcup_{\substack{s\in\mathbb{Q} \\ s\le t}} \{|M_s| > n - 1/k\} \in \mathcal{F}_t.$$

Therefore S_n is a stopping time. By the continuity of M, we have that $\sup_{s\le t} |M_s(\omega)| < \infty$ for all $t \ge 0$ and $\omega \in \Omega$. Whenever $n > \sup_{s\le t} |M_s(\omega)|$ we have that $S_n(\omega) > t$, hence $S_n \uparrow \infty$ a.s.

By assumption, we know that there exists a sequence of stopping times (T_k) with $T_k \uparrow \infty$ such that M^{T_k} is a martingale for every k. By the OST, we have that $M^{T_k\wedge S_n}$ is also a martingale. Therefore M^{S_n} is a local martingale. Since M^{S_n} is also bounded, it is in fact a martingale. That is, (S_n) reduces M. $\square$

Theorem 2.14. *Let M be a continuous, local martingale of finite variation. Suppose that $M_0 = 0$. Then $M \equiv 0$ a.s.*

Proof. Let V denote the total variation process associated with M. Then V is continuous and adapted with $V_0 = 0$. Let $S_n = \inf\{t \ge 0 : V_t = n\}$. Then S_n is a stopping time for all n and $S_n \uparrow \infty$ as $n \to \infty$. It suffices to show that $M^{S_n} \equiv 0$ a.s. for all $n \in \mathbb{N}$. By the OST, M^{S_n} is a local martingale. Moreover, $|M_t^{S_n}| \le |V_t^{S_n}| \le n$. Thus M^{S_n} is bounded, hence it is a martingale. Therefore, for the rest of the proof, we may assume without loss of generality that M is a bounded martingale of bounded variation.

Fix $t > 0$ and let $t_k = tk/N$. Using that M is a bounded martingale, we have that

$$\mathbb{E}[(M_{t_{k+1}} - M_{t_k})^2] = \mathbb{E}[M_{t_{k+1}}^2] - \mathbb{E}[M_{t_k}^2].$$

Thus,

$$\begin{aligned}\mathbb{E}[M_t^2] &= \mathbb{E}\left[\sum_{k=0}^{N-1}(M_{t_{k+1}} - M_{t_k})^2\right] \\ &\le \mathbb{E}\left[\left(\max_{k<N}|M_{t_{k+1}} - M_{t_k}|\right)\sum_{k=0}^{N-1}|M_{t_{k+1}} - M_{t_k}|\right].\end{aligned}$$

Both factors inside of the expectation are bounded by V_t hence by n. Since M is continuous, we moreover have that $\max_{k<N}|M_{t_{k+1}} - M_{t_k}| \to 0$ as $N \to \infty$. Combining, the bounded convergence theorem implies that

$$\mathbb{E}\left[\left(\max_{k<N}|M_{t_{k+1}} - M_{t_k}|\right)\sum_{k=0}^{N-1}|M_{t_{k+1}} - M_{t_k}|\right] \to 0 \quad \text{as} \quad N \to \infty.$$

Therefore $\mathbb{E}[M_t^2] = 0$ which implies that $M_t = 0$ a.s. As M is continuous, we conclude that $M \equiv 0$ a.s. □

Remark 2.15. *(i) Brownian motion is not of finite variation.*

(ii) The theorem implies that the theory of finite variation integrals cannot be used to define the integral against a continuous local martingale.

Definition 2.16. *A cádlág, adapted process X is called a semimartingale if it can be written in the form*

$$X = X_0 + M + A$$

where M is a local martingale with $M_0 = 0$ and A is a process of finite variation starting from 0. This is the Doob-Meyer decomposition of X.

3. The stochastic integral

We are now going to construct the stochastic integral with respect to a continuous semimartingale X. Some of the theory that we will develop will also apply to cádlág semimartingales, however parts of the theory will depend on the assumption that X is continuous.

Throughout, our integrals will be over the interval $(0, t]$ and will take the value 0 at 0.

3.1. **Simple integrands.**

Definition 3.1. *A simple process is a map $H \colon \Omega \times (0,\infty) \to \mathbb{R}$ of the form*

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

where $n \in \mathbb{N}$, $0 = t_0 < \cdots < t_n < \infty$, and Z_k is bounded and $\mathcal{F}_{t_k}$-measurable for all k.

We will denote the set of simple processes by $\mathcal{S}$. Note that $\mathcal{S}$ is a vector space and that any simple process is previsible. Recall that we say that a process X is L^2-bounded if

$$\sup_{t\ge 0}\|X_t\|_{L^2} < \infty.$$

Let $\mathcal{M}^2$ be the set of cádlág, L^2-bounded martingales. If $X \in \mathcal{M}^2$, then the L^2 martingale convergence theorem says that there exists $X_\infty \in L^2$ such that $X_t \to X_\infty$ in L^2 as $t \to \infty$. Moreover, $X_t = \mathbb{E}[X_\infty \,|\, \mathcal{F}_t]$. We also recall Doob's inequality: if $X^* = \sup_{t \geq 0} |X_t|$, then $\|X^*\|_{L^2} \leq 2\|X_\infty\|_{L^2}$.

For $H = \sum_{k=0}^{n-1} Z_k \mathbf{1}_{(t_k, t_{k+1}]} \in \mathcal{S}$ and $M \in \mathcal{M}^2$, we set

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

This is a continuous-time version of the so-called martingale transform.

Proposition 3.2. *Let $H \in \mathcal{S}$ and $M \in \mathcal{M}^2$. Let T be a stopping time. Then*

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

Proof. We have that

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

This proves (a).

For $t_k \leq s \leq t \leq t_{k+1}$, we have that $(H \cdot M)_t - (H \cdot M)_s = Z_k (M_t - M_s)$, so that

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

as M is a martingale. This extends easily to general $s \leq t$ by the tower property. Therefore $(H \cdot M)$ is a martingale. Note that if $j < k$ then

$$\mathbb{E}[Z_j (M_{t_{j+1}} - M_{t_j}) Z_k (M_{t_{k+1}} - M_{t_k})] = \mathbb{E}[Z_j (M_{t_{j+1}} - M_{t_j}) Z_k \mathbb{E}[M_{t_{k+1}} - M_{t_k} \,|\, \mathcal{F}_{t_k}]] = 0.$$

Therefore

$$\begin{aligned}\mathbb{E}[(H \cdot M)_t^2] &= \mathbb{E}\left[\left(\sum_{k=0}^{n-1} Z_k (M_{t_{k+1} \wedge t} - M_{t_k \wedge t})\right)^2\right] \\ &= \sum_{k=0}^{n-1} \mathbb{E}\left[Z_k^2 (M_{t_{k+1} \wedge t} - M_{t_k \wedge t})^2\right] \\ &\leq \|H\|_{L^\infty}^2 \sum_{k=0}^{n-1} \mathbb{E}[(M_{t_{k+1} \wedge t} - M_{t_k \wedge t})^2] \\ &= \|H\|_{L^\infty}^2 \mathbb{E}[(M_t - M_0)^2]. \qquad (3.1)\end{aligned}$$

By Doob's L^2-inequality applied to $M_t - M_0$, we have that

$$\sup_{t \geq 0} \mathbb{E}[(H \cdot M)_t^2] \leq 4\|H\|_{L^\infty}^2 \mathbb{E}[(M_\infty - M_0)^2].$$

Therefore $H \cdot M \in \mathcal{M}^2$. This proves (b).

Part (c) follows because (3.1) holds for $t = \infty$. $\square$

Proposition 3.3. *Let μ be a finite measure on $\mathcal{P}$, the previsible σ-algebra. Then $\mathcal{S}$ is a dense subspace of $L^2(\mathcal{P}, \mu)$.*

Proof. As the Z_k's are bounded, it is certainly the case that $\mathcal{S} \subseteq L^2(\mathcal{P}, \mu)$. Denote by $\overline{\mathcal{S}}$ the closure of $\mathcal{S}$ in $L^2(\mathcal{P}, \mu)$. Let $\mathcal{A} = \{A \in \mathcal{P} : \mathbf{1}_A \in \overline{\mathcal{S}}\}$. Observe that $\mathcal{A}$ is a d-system, which contains the π-system $\{B \times (s,t] : B \in \mathcal{F}_s,\ s \leq t\}$ which generates $\mathcal{P}$. So, by Dynkin's lemma, we have that $\mathcal{A} = \mathcal{P}$. The result then follows because finite linear combinations of measurable indicators are dense in L^2. $\square$

3.2. L^2 **properties.** In order to proceed we need to define some Hilbert space properties of the spaces of integration that we will consider. This will then enable us to extend the simple integral that we have already defined.

We let X be a cádlág, adapted process. We then define a norm by

$$|||X||| = \|X^*\|_{L^2} \quad \text{where} \quad X^* = \sup_{t \geq 0} |X_t|.$$

We let $\mathcal{C}^2$ be the set of cádlág, adapted processes X such that $|||X||| < \infty$. We also let $\mathcal{M}$ be the set of cádlág martingales, $\mathcal{M}_c$ be the set of continuous martingales, $\mathcal{M}_{c,loc}$ be the set of continuous, local martingales, and $\mathcal{M}^2$ be the set of cádlág L^2-bounded martingales. We also define the norm $\|X\| = \|X_\infty\|_{L^2}$ on $\mathcal{M}^2$.

Proposition 3.4. *(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 and $\mathcal{M}_c^2 = \mathcal{M}_c \cap \mathcal{M}^2$ is a closed subspace.
(d) The map $\mathcal{M}^2 \to L^2(\mathcal{F}_\infty)$ given by $X \mapsto X_\infty$ is an isometry.

We can identify an element of $\mathcal{M}^2$ with its terminal value and then $\mathcal{M}^2$ inherits the Hilbert space structure of $L^2(\mathcal{F}_\infty)$.

Proof of Proposition 3.4. Suppose that $(X^n)_{n \geq 1}$ is a Cauchy sequence in $(\mathcal{C}^2, |||\cdot|||)$. Then we can find a subsequence $(X^{n_k})_{k \geq 1}$ such that $\sum_k |||X^{n_{k+1}} - X^{n_k}||| < \infty$. Then we have that

$$\left\| \sum_k \sup_{t \geq 0} |X_t^{n_{k+1}} - X_t^{n_k}| \right\|_{L^2} \leq \sum_k |||X^{n_{k+1}} - X^{n_k}||| < \infty.$$

Therefore for a.e. ω, $\sum_k \sup_{t \geq 0} |X_t^{n_{k+1}} - X_t^{n_k}| < \infty$. Therefore $(X^{n_k}(\omega))_{k \geq 1}$ converges as $k \to \infty$, uniformly in $t \geq 0$. The limit $X(\omega)$ is a cádlág process because it is a uniform limit of cádlág functions. Moreover,

$$\begin{aligned}
|||X^n - X||| &= \mathbb{E}\left[\sup_{t \geq 0} |X_t^n - X_t|^2\right] \\
&\leq \liminf_{k \to \infty} \mathbb{E}\left[\sup_{t \geq 0} |X_t^n - X_t^{n_k}|^2\right] \quad \text{(Fatou's lemma)} \\
&= \liminf_{k \to \infty} |||X^n - X^{n_k}|||^2 \to 0 \quad \text{as} \quad n \to \infty
\end{aligned}$$

as (X^n) is Cauchy. This proves (a).

Suppose that $X \in \mathcal{C}^2 \cap \mathcal{M}$. Then $|||X||| < \infty$. But then

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

Therefore $X \in \mathcal{M}^2$. On the other hand, if $X \in \mathcal{M}^2$, by Doob's L^2-inequality $|||X||| \leq 2\|X\|_{L^2} < \infty$, so $X \in \mathcal{C}^2 \cap \mathcal{M}$. Hence $\mathcal{M}^2 = \mathcal{M} \cap \mathcal{C}^2$. This proves (b).

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 just shown that

$$\|X\|_{L^2} \leq |||X||| \leq 2\|X\|_{L^2}.$$

Therefore $\|\cdot\|$ and $|||X|||$ are equivalent norms. So, showing that $(\mathcal{M}^2, |||\cdot|||)$ is complete is equivalent to showing that $(\mathcal{M}^2, \|\cdot\|_{L^2})$ is complete. By (a), it suffices to show that $\mathcal{M}^2$ is closed in $(\mathcal{C}^2, |||\cdot|||)$. But if $X^n \in \mathcal{M}^2$ and $|||X^n - X||| \to 0$ as $n \to \infty$ for some X, then X is cádlág, adapted, and L^2-bounded. Furthermore,

$$\begin{aligned}
\|\mathbb{E}[X_t \,|\, \mathcal{F}_s] - X_s\|_{L^2} &\leq \|\mathbb{E}[X_t - X_t^n \,|\, \mathcal{F}_s] + X_s^n - X_s\|_{L^2} \\
&\leq \|\mathbb{E}[X_t - X_t^n \,|\, \mathcal{F}_s\|_{L^2} + \|X_s^n - X_s\|_{L^2} \quad \text{(Minkowski's inequality)} \\
&\leq \|X_t - X_t^n\|_{L^2} + \|X_s^n - X_s\|_{L^2} \\
&\leq 2|||X_s^n - X||| \to 0 \quad \text{as} \quad n \to \infty.
\end{aligned}$$

Therefore $X \in \mathcal{M}^2$. Exactly the same argument applies to show that $\mathcal{M}_c^2$ is a closed, subspace of $(\mathcal{M}^2, \|\cdot\|)$. This proves (c).

Part (d) follows from the definition. □

3.3. **Quadratic variation.** The tool which will allow us to construct the full stochastic integral is the so-called quadratic variation of a local martingale.

Let $(X^n)_{n \geq 1}$ be a sequence of processes. We say that $X^n \to X$ uniformly on compacts in probability (ucp) if for every $\epsilon, t > 0$ we have that

$$\mathbb{P}[\sup_{s \leq t} |X_s^n - X_s| > \epsilon) \to 0 \quad \text{as} \quad n \to \infty.$$

Theorem 3.5 (Quadratic variation)**.** *Let M be a continuous local martingale. Then there exists a unique continuous, adapted, and increasing process $[M]$ such that $M^2 - [M]$ is a continuous, local martingale. Moreover, if we define*

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

then $[M]^n \to [M]$ ucp as $n \to \infty$.

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

Example 3.6. *Let B be a standard Brownian motion. Then $(B_t^2 - t)_{t \geq 0}$ is a martingale. Therefore $[B]_t = t$. As we will see later, that $[B]_t = t$ singles out the standard Brownian motion among continuous local martingales. This is the so-called Lévy characterization of Brownian motion.*

Proof of Theorem 3.5. By replacing M with $M_t - M_0$ if necessary, we may assume without loss of generality that $M_0 = 0$.

Step 1: uniqueness. If A and A' are two increasing processes satisfying the conditions of the theorem, then we can write

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

The left hand side is a continuous process of bounded variation as it is given by the difference of two non-decreasing functions while the right hand side is a continuous local martingale as it is given by the difference of two continuous local martingales. Since a continuous local martingale starting from 0 with bounded variation is almost surely equal to 0, it follows that $A - A' \equiv 0$. That is, $A \equiv A'$, as desired.

Step 2: existence, M bounded. We shall first consider the case that M is a bounded, continuous martingale. (We will later reduce the general case of a continuous local martingale to this case using a localization argument.) Then we have that $M \in \mathcal{M}_c^2$. Fix $T > 0$ deterministic and let

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

Then H^n is a simple process. That is, $H^n \in \mathcal{S}$ for all n. Consequently, it follows that

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

is an L^2-bounded martingale. Moreover, X^n is continuous. Then for $n, m \geq 1$ we have that

$$\begin{aligned}
\|X^n - X^m\|^2 &= \mathbb{E}[((H^n - H^m) \cdot M)_T^2] \\
&\leq \mathbb{E}[\sup_{0 \leq t \leq T} |H_t^n - H_t^m|^2] \|M^T\|^2 \\
&= \mathbb{E}[\sup_{0 \leq t \leq T} |M_{2^{-n}\lfloor 2^n t \rfloor} - M_{2^{-m}\lfloor 2^m t \rfloor}|^2] \|M^T\|^2 \to 0 \quad \text{as} \quad n, m \to \infty
\end{aligned}$$

by the uniform continuity of M on $[0, T]$ and the bounded convergence theorem.

We have thus shown that (X^n) is a Cauchy sequence in $(\mathcal{M}_c^2, \|\cdot\|)$, hence there exists $Y \in (\mathcal{M}_c^2, \|\cdot\|)$ such that $X^n \to Y$ as $n \to \infty$.

Now, for any n and $1 \leq k \leq \lceil 2^n T \rceil$, we can write

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

Consequently, $M_t^2 - 2X_t^n$ is non-decreasing when restricted to the set of times $(k2^{-n} : 1 \leq k \leq \lceil 2^n T \rceil)$. Taking a limit as $n \to \infty$, it thus follows that $M_t^2 - 2Y_t$ is non-decreasing. Set

$$[M]_t = M_t^2 - 2Y_t.$$

Then $[M]$ is a continuous, non-decreasing process and $M^2 - [M] = 2Y$ is a martingale on $[0, T]$.

We can extend this definition to $[0, \infty)$ by applying the above for each $T = k \in \mathbb{N}$. By uniqueness, we note that the process obtained with $T = k$ must be the restriction to $[0, k]$ of the process obtained with $T = k + 1$.

Note that the convergence $X^n \to Y$ in $(\mathcal{M}_c^2, \|\cdot\|)$ implies that $\sup_{0 \leq t \leq T} |X_t^n - Y_t| \to 0$ as $n \to \infty$ in L^2. Now,

$$[M]_t^n = M_{2^{-n}\lfloor 2^n t \rfloor}^n - 2X_{2^{-n}\lfloor 2^n t \rfloor}^n$$

and so

$$\sup_{0\le t\le T}|[M]_t-[M]^n_t|\le \sup_{0\le t\le T}|M^2_{2^{-n}\lfloor 2^n t\rfloor}-M^2_t|+2\sup_{0\le t\le T}|X^n_{2^{-n}\lfloor 2^n t\rfloor}-Y_{2^{-n}\lfloor 2^n t\rfloor}|+2\sup_{0\le t\le T}|Y^n_{2^{-n}\lfloor 2^n t\rfloor}-Y_t|.$$

Each term in the right hand side converges to 0 in probability. Indeed, this follows for the first and third terms by continuity and for the second term because $\sup_{0\le t\le T}|X^n_t-Y_t|\to 0$ in L^2 as $n\to\infty$.

Step 3: the general case. We now suppose that $M\in\mathcal{M}_{c,loc}$. For each n, we let $T_n=\inf\{t\ge 0: |M_t|\ge n\}$. Then (T_n) reduces M and we can apply the case of bounded continuous martingales to M^{T_n}. Write $A^n=[M^{T_n}]$. By uniqueness, $A^{n+1}_{t\wedge T_n}$ and A^n_t must be indistinguishable. We thus take A to be the non-decreasing process such that $A_{t\wedge T_n}=A^n_t$ for all n. By construction, $M^2_{t\wedge T_n}-A_{t\wedge T_n}$ is a martingale for each n. Therefore M^2-A is a local martingale. Thus we can take $[M]_t=A_t$.

Note that $[M^{T_n}]^m\to[M^{T_n}]$ as $m\to\infty$ ucp. Thus as $\mathbb{P}[T_n\le T]\to 1$ as $n\to\infty$ for each $T>0$, it follows that $[M]^m\to[M]$ as $m\to\infty$ ucp. □

Theorem 3.7. *If $M\in\mathcal{M}^2_c$, then $M^2-[M]$ is a uniformly integrable martingale.*

Proof. Let $S_n=\inf\{t\ge 0:[M]_t\ge n\}$. Then S_n is a stopping time and $[M]_{t\wedge S_n}\le n$. Therefore the local martingale $M^2_{t\wedge S_n}-[M]_{t\wedge S_n}$ is dominated by the integrable random variable (by Doob's inequality) $n+\sup_{t\ge 0}M^2_t$. Therefore $M^2_{t\wedge S_n}-[M]_{t\wedge S_n}$ must be a true martingale. Therefore

$$\mathbb{E}[[M]_{t\wedge S_n}]=\mathbb{E}[M^2_{t\wedge S_n}].$$

Letting $t\to\infty$ and using the monotone convergence theorem on the left and side and the dominated convergence theorem on the right hand side, we see that

$$\mathbb{E}[[M]_{S_n}]=\mathbb{E}[M^2_{S_n}].$$

Sending $n\to\infty$ and using the same argument, we thus have that

$$\mathbb{E}[[M]_\infty]=\mathbb{E}[M^2_\infty]<\infty.$$

So, $M^2_t-[M]_t$ is dominated by the integrable random variable $[M]_\infty+\sup_{t\ge 0}M^2_t$. Thus $M^2_t-[M]_t$ is a true martingale and is UI. □

3.4. **Ito integrals.** Given $M\in\mathcal{M}^2_c$, we define a measure μ on $\mathcal{P}$ by setting

$$\mu(A\times(s,t])=\mathbb{E}[\mathbf{1}(A)([M]_t-[M]_s)]\quad\text{for}\quad s\le t.$$

Then for a previsible process $H\ge 0$, we have that

$$\int H d\mu=\mathbb{E}[\int_0^\infty H_s d[M]_s].$$

The integral inside of the expectation is with respect to the Lebesgue-Stieljes integral associated with the non-decreasing process $[M]$. Let $L^2(M)=L^2(\Omega\times(0,\infty),\mathcal{P},\mu)$ and write

$$\|H\|_{L^2(M)}=\|H\|_M=\left(\mathbb{E}[\int_0^\infty H^2_s d[M]_s]\right)^{1/2}.$$

Then $L^2(M)$ is the set of previsible processes such that $\|H\|_M<\infty$.

Suppose that

$$H=\sum_{k=0}^{n-1}Z_k\mathbf{1}_{(t_k,t_{k+1}]}(t)\in\mathcal{S}.$$

Then we know that $H\cdot M\in\mathcal{M}_c^2$ and

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

But $M^2-[M]$ is a martingale, so that

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

Therefore

$$\|H\cdot M\|^2=\mathbb{E}[\int_0^\infty H_s^2d[M]_s]=\|H\|_M^2.$$

Theorem 3.8 (Ito isometry)**.** *There exists a unique isometry $I\colon L^2(M)\to\mathcal{M}_c^2$ such that $I(H)=H\cdot M$ for all $H\in\mathcal{S}$.*

Proof. Suppose that $H\in L^2(M)$ and let (H^n) be a sequence in $\mathcal{S}\cap L^2(M)$ such that $H^n\to H$ as $n\to\infty$. That is,

$$\mathbb{E}[\int_0^\infty(H_s^n-H_s)^2d[M]_s]\to 0\quad\text{as}\quad n\to\infty.$$

Now, $\|H^n\cdot M\|=\|H^n\|_M$ for all n and

$$\|(H^n-H^m)\cdot M\|=\|H^n-H^m\|_M\to 0\quad\text{as}\quad n,m\to\infty.$$

So, $H^n\cdot M$ is a Cauchy sequence in $\mathcal{M}_c^2$. As $(\mathcal{M}_c^2,\|\cdot\|)$ is complete, the sequence $H^n\cdot M$ has a limit $H\cdot M\in\mathcal{M}_c^2$.

Note that

$$\|H\cdot M\|=\lim_{n\to\infty}\|H^n\cdot M\|=\lim_{n\to\infty}\|H^n\|_M=\|H\|_M$$

and so $I(H)=H\cdot M$ is an isometry for any H obtainable as a limit of simple processes. But $\mathcal{S}$ is dense in $L^2(M)$, so the integral extends to all of $L^2(M)$. □

We write $I(H)_t=(H\cdot M)_t=\int_0^t H_sdM_s$. The process $H\cdot M$ is Ito's stochastic integral of H with respect to M.

Example 3.9. *Suppose that B is a standard Brownian motion. Then B_t^2-t is a martingale, hence $[B]_t=t$. We will show later in this course that:*

(1) If M is any continuous local martingale with $[M]_t=t$, then M is a standard Brownian motion (Lévy characterization of Brownian motion).
(2) Every continuous local martingale can be realized as a time-change of a standard Brownian motion (time-change the process so that its quadratic variation is equal to t, apply Lévy characterization). (Dubins-Schwarz theorem.)

Proposition 3.10. *Let $M\in\mathcal{M}_c^2$ and $H\in L^2(M)$. Let T be a stopping time. Then*

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

Proof. Case 1. Fix $H \in \mathcal{S}$, $M \in \mathcal{M}_c^2$, T taking only finitely many values. Then one can check that $H\mathbf{1}_{(0,T]} \in \mathcal{S}$ and $(H \cdot M)^T = (H\mathbf{1}_{(0,T]}) \cdot M$.

Case 2. Fix $H \in \mathcal{S}$, $M \in \mathcal{M}_c^2$, T a general stopping time. Then we have already shown that $(H \cdot M)^T = H \cdot (M^T)$.

For n, m, we let $T_{n,m} = (2^{-n}\lceil 2^n t \rceil) \wedge m$. Then $T_{n,m}$ takes on only finitely many values and $T_{n,m} \downarrow T \wedge m$ as $n \to \infty$. Thus,

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

by the dominated convergence theorem. Therefore $(H\mathbf{1}_{(0,T_{n,m}]}) \cdot M \to (H \cdot \mathbf{1}_{(0,T\wedge m]}) \cdot M$ in $\mathcal{M}_c^2$ as $n \to \infty$. But $(H \cdot M)_t^{T_{n,m}} \to (H \cdot M)_t^{T\wedge m}$ as $n \to \infty$ by the continuity of $H \cdot M$. Therefore $(H \cdot M)^{T\wedge m} = (H\mathbf{1}_{(0,T\wedge m]}) \cdot M$. Sending $m \to \infty$ and applying a similar argument implies that $(H \cdot M)^T = (H\mathbf{1}_{(0,T]}) \cdot M$.

Case 3. $H \in L^2(M)$, $M \in \mathcal{M}_c^2$, T a general stopping time. Chose a sequence (H^n) in $\mathcal{S}$ such that $H^n \to H$ in $L^2(M)$. Then $H^n \cdot M \to H \cdot M$ in $\mathcal{M}_c^2$, so $(H^n \cdot M)^T \to (H \cdot M)^T$ in $\mathcal{M}_c^2$. Also,

$$\|H^n\mathbf{1}_{(0,T]} - H\mathbf{1}_{(0,T]}\|_M^2 = \mathbb{E}[\int_0^T (H_t^n - H_t) d[M]_t] \leq \|H^n - H\|_M^2 \to 0 \quad \text{as} \quad n \to \infty.$$

So we also have that $(H^n\mathbf{1}_{(0,T]}) \cdot M \to (H\mathbf{1}_{(0,T]}) \cdot M$ in $\mathcal{M}_c^2$. Hence, $(H \cdot M)^T = (H\mathbf{1}_{(0,T]}) \cdot M$. Moreover,

$$\|H^n - H\|_{M^T}^2 = \mathbb{E}\left[\int_0^\infty (H_s^n - H_s)^2 d[M^T]\right] = \mathbb{E}[\int_0^T (H_s^n - H_s)^2 d[M]_s] \leq \|H^n - H\|_M^2 \to 0 \quad \text{as} \quad n \to \infty.$$

So, we also have that $(H \cdot M)^T = H \cdot (M^T)$. □

The previous proposition allows us to make an extension to the Ito integral. Let H be a previsible process. We say that H is *locally bounded* if there exists a sequence (S_n) of stopping times such that $S_n \uparrow \infty$ as $n \to \infty$ and such that $H\mathbf{1}_{(0,S_n]}$ is bounded for all n. Let M be a continuous local martingale and let (S_n') be a reducing sequence for M. Let $T_n = S_n \wedge S_n'$ and define $(H \cdot M)_t = ((H\mathbf{1}_{(0,T_n]}) \cdot M^{T_n})_t$ whenever $t \leq T_n$. Then the previous proposition implies that $H \cdot M$ is well-defined and is a continuous local martingale.

Summary of the stochastic integral

Step 1. $H \in \mathcal{S}$, $M \in \mathcal{M}_c^2$.

$H_t = \sum_{k=0}^{n-1} Z_k \mathbf{1}_{(t_k, t_{k+1}]}(t)$ where Z_k is a bounded, $\mathcal{F}_{t_k}$-measurable random variable for each k. We set

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

and showed that $H \cdot M \in \mathcal{M}_c^2$.

Step 2. Establish the existence of $[M]$

For $M \in \mathcal{M}_{c,loc}$, there exists a unique, adapted, continuous, non-decreasing process $[M]$ such that $M^2 - [M] \in \mathcal{M}_{c,loc}$.

Step 3. Extend the integral to $H \in L^2(M)$, $M \in \mathcal{M}_c^2$

Define the norms $\|M\| = (\mathbb{E}[M_\infty^2])^{1/2}$ on $\mathcal{M}_c^2$ and $\|H\|_M = \|H\|_{L^2(M)} = (\mathbb{E}[(H^2 \cdot [M])_\infty])^{1/2}$. Here, $L^2(M)$ is the space of previsible processes H such that $\|H\|_M < \infty$. For $H \in \mathcal{S}$, $M \in \mathcal{M}_c^2$, $\|H\|_M = \|H \cdot M\|$. Let $I \colon \mathcal{S} \to \mathcal{M}_c^2$ be given by $I(H) = H \cdot M$. As $\mathcal{S}$ is dense in $L^2(M)$ and $(\mathcal{M}_c^2, \|\cdot\|)$ is complete, I extends uniquely to all of $L^2(M)$. In other words, to each $H \in L^2(M)$, I associates a unique process $X \in \mathcal{M}_c^2$ such that $\|H\|_M = \|X\|$. We define $H \cdot M = X$.

Step 4. Extend to H locally bounded and $M \in \mathcal{M}_{c,loc}$

Suppose that there exists a sequence (S_n) of stopping times, $S_n \uparrow \infty$ as $n \to \infty$, such that $H\mathbf{1}_{(0,S_n]}$ is bounded for all n. In other words, we assume that H is locally bounded. There also exists a sequence (S_n') of stopping times, $S_n' \uparrow \infty$ as $n \to \infty$, such that $M^{S_n'} \in \mathcal{M}_c^2$ for all n. Set $T_n = S_n \wedge S_n'$. We know that for $M \in \mathcal{M}_c^2$ and $H \in L^2(M)$ that for any stopping time, the integral is consistent:

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

So, for H locally bounded and $M \in \mathcal{M}_{c,loc}$, we can unambigously set

$$(H \cdot M)_t = ((H\mathbf{1}_{(0,T_n]}) \cdot M^{T_n})_t \quad \text{for} \quad t \le T_n.$$

Proposition 3.11. *Let $M \in \mathcal{M}_{c,loc}$ and let H be a locally bounded previsible process. Let T be a stopping time. Then:*

(a) $(H \cdot M)^T = (H\mathbf{1}_{(0,T]}) \cdot M = H \cdot M^T$
(b) $H \cdot M \in \mathcal{M}_{c,loc}$
(c) $[H \cdot M] = H^2 \cdot [M]$
(d) $H \cdot (K \cdot M) = (HK) \cdot M$

Proof. Parts (a) and (b) follow from Proposition 3.10. By (a), we can reduce (c) and (d) to the case where M, H, and K are all uniformly bounded. (This is an example of a so-called localization argument.) Then we have that

$$\mathbb{E}[(H \cdot M)_T^2] = \mathbb{E}[((H\mathbf{1}_{(0,T]} \cdot M)_\infty^2] = \mathbb{E}[((H^2\mathbf{1}_{(0,T]} \cdot [M])_\infty] = \mathbb{E}[(H^2 \cdot [M])_T].$$

By the OST, we thus have that $(H \cdot M)^2 - H^2 \cdot [M]$ is a martingale. By Theorem 3.5, the unique non-decreasing, adapted, continuous process X such that $(H \cdot M)^2 - X$ is a martingale is $[H \cdot M]$. Therefore $[H \cdot M] = H^2[M]$. This proves (c).

We will now prove (d). The case that $H, K \in \mathcal{S}$ is an elementary exercise. For H, K uniformly bounded, there exist sequences (H^n), (K^n) in $\mathcal{S}$ such that $H^n \to H$ in $L^2(M)$ and $K^n \to K$ in $L^2(M)$. We will first prove an upper bound on $\|H\|_{L^2(K \cdot M)}$. We have that

$$\|H\|_{L^2(K\cdot M)}^2 = \mathbb{E}[(H^2 \cdot [K \cdot M])_\infty] = \mathbb{E}[(H^2 \cdot (K^2 \cdot [M]))_\infty] = \mathbb{E}[((HK)^2 \cdot [M])_\infty]$$

by the corresponding property for the Stieljes integral. This, in turn, is equal to

$$\|HK\|_{L^2(M)}^2 \le \min\left(\|H\|_\infty^2 \|K\|_{L^2(M)}^2, \|K\|_\infty^2 \|H\|_{L^2(M)}^2\right).$$

We have already that $H^n \cdot (K^n \cdot M) = (H^n K^n) \cdot M$. Also,

$$\begin{aligned}
&\|H^n \cdot (K^n \cdot M) - H \cdot (K \cdot M)\| \\
\le &\|(H^n - H) \cdot (K^n \cdot M)\| + \|H \cdot ((K^n - K) \cdot M)\| \\
= &\|H^n - H\|_{L^2(K^n \cdot M)} + \|H\|_{L^2((K^n - K)\cdot M)} \\
\le &\|H^n - H\|_{L^2(M)} \|K^n\|_\infty + \|H\|_\infty \|K^n - K\|_{L^2(M)} \to 0 \quad \text{as} \quad n \to \infty.
\end{aligned}$$

A similar argument implies that $(H^n K^n) \cdot M \to (HK) \cdot M$ in $\mathcal{M}_c^2$, which implies the result. $\square$

Proposition 3.12. *Let X be a continuous semimartingale and H be a locally bounded left-continuous process which is adapted. Then*

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

Proof. We can treat the local martingale and finite variation parts separately. The latter is Exercise 6 on Example Sheet 2, so this leaves just the martingale part. By applying a localization argument, we can reduce to the case in which $M\in\mathcal{M}_c^2$ and H is bounded. Let $H_t^n = H_{2^{-n}\lfloor 2^n t\rfloor}$. Then, by the left-continuity of H, we have that $H_t^n \to H_t$ as $n\to\infty$. Moreover, we have that

$$(H^n\cdot M)_t = \sum_{k=0}^{\lfloor 2^n t\rfloor -1} H_{k2^{-n}}\big(M_{(k+1)2^{-n}} - M_{k2^{-n}}\big) + H_{2^{-n}\lfloor r^n\rfloor}(M_t - M_{2^{-n}\lfloor 2^n t\rfloor})$$

where $|M_t - M_{2^{-n}\lfloor 2^n t\rfloor}|\to 0$ as $n\to\infty$ by the left-continuity of M. Consequently, we can ignore the second term on the right hand side.

Now,

$$\mathbb{E}\left[\int_0^\infty (H_t^n - H_t)^2 d[M]_t\right]\to 0 \quad \text{as} \quad n\to\infty$$

by the bounded convergence theorem. Therefore $H^n\cdot M\to H\cdot M$ in $\mathcal{M}_c^2$ by the Ito isometry. But, convergence in $(\mathcal{M}_c^2, \|\cdot\|)$ implies ucp convergence, which completes the proof. □

Theorem 3.13. *Let $M,N\in\mathcal{M}_{c,loc}$. Let*

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

Then there exists a unique continuous, adapted, finite variation process $[M,N]$ such that

(a) $[M,N]_t^n\to[M,N]_t$ ucp as $n\to\infty$.
(b) $MN-[M,N]\in\mathcal{M}_{c,loc}$.
(c) For $M,N\in\mathcal{M}_c^2$, $MN-[M,N]$ is a UI martingale.
(d) For H locally bounded and previsible,

$$[H\cdot M,N]+[M,H\cdot N] = 2H\cdot[M,N].$$

$[M,N]$ is called the covariation of M and N. Note that $[M,M]=[M]$. Moreover, $[M,N]$ is a symmetric bilinear form.

Proof of Theorem 3.13. Note that

$$MN = \frac{1}{4}\left((M+N)^2 - (M-N)^2\right).$$

Thus we take

$$[M,N] = \frac{1}{4}\left([M+N]-[M-N]\right) \tag{3.2}$$

and we will show that it satisifes the desired properties. (The identity (3.2) is the so-called polarization identity.) We also note that

$$[M,N]_t^n = \frac{1}{4}\left([M+N]_t^n - [M-N]_t^n\right).$$

Then (a)– (c) follow from the equivalent properties for quadratic variation.

For (d), we note that

$$[H \cdot (M \pm N)] = H^2[M \pm N].$$

Hence, by (3.2), we have that

$$[H \cdot M, H \cdot N] = H^2[M, N].$$

We then have that

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

We also have that

$$(H+1)^2[M,N] = (H^2 + 2H + 1)[M,N] = H^2[M,N] + 2H[M,N] + [M,N].$$

By matching up terms, we see that

$$2H[M,N] = [M, H \cdot N] + [H \cdot M, N],$$

as required. □

Proposition 3.14 (Kunita-Watanabe identity)**.** *Let $M, N \in \mathcal{M}_{c,loc}$ and H a locally bounded, previsible process. Then $[H \cdot M, N] = H \cdot [M, N]$.*

Proof. It suffices to show that $[H \cdot M, N] = [M, H \cdot N]$. We have that $(H \cdot M)N - [H \cdot M, N] \in \mathcal{M}_{c,loc}$ and $M(H \cdot N) - [M, H \cdot N] \in \mathcal{M}_{c,loc}$. Hence if we can show that $(H \cdot M)N - M(H \cdot N) \in \mathcal{M}_{c,loc}$, then we can deduce that $[H \cdot M, N] - [M, H \cdot N] \in \mathcal{M}_{c,loc}$. But this last process can be expressed as sums and differences of quadratic variations and, as quadratic variations are non-decreasing, this will imply that $[H \cdot M, N] - [M, H \cdot N]$ is of finite variation. Hence we will be able to deduce that $[H \cdot M, N] = [M, H \cdot N]$.

By localization, we may assume that $M, N \in \mathcal{M}_c^2$ and that H is bounded. By OST, it suffices to show that

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

for all bounded stopping times T. We know that we can replace M by M^T and N by N^T. Therefore it suffices to show that

$$\mathbb{E}[(H \cdot M)_\infty \cdot N_\infty] = \mathbb{E}[M_\infty \cdot (H \cdot N)_\infty] \tag{3.3}$$

for all $M, N \in \mathcal{M}_c^2$ and H bounded. Consider first the case that $H = Z\mathbf{1}_{(s,t]}$ where Z is bounded and $\mathcal{F}_s$-measurable. Then

$$\begin{aligned}\mathbb{E}[(H \cdot M)_\infty N_\infty] &= \mathbb{E}[Z(M_t - M_s)N_\infty] = \mathbb{E}[Z M_t \mathbb{E}[N_\infty \,|\, \mathcal{F}_t] - Z M_s \mathbb{E}[N_\infty \,|\, \mathcal{F}_s]] \quad \text{(tower property)} \\ &= \mathbb{E}[Z(M_t N_t - M_s N_s)].\end{aligned}$$

The same argument implies that

$$\mathbb{E}[M_\infty (H \cdot N)_\infty] = \mathbb{E}[Z(M_t N_t - M_s N_s)]$$

which proves (3.3) for $H = Z\mathbf{1}_{(s,t]}$.

By linearity, this extends to all $H \in \mathcal{S}$. If H is bounded, then we can find a sequence (H^n) such that $H^n \to H$ in both $L^2(M)$ and $L^2(N)$. Therefore $(H^n \cdot M)_\infty \to (H \cdot M)_\infty$ and $(H^n \cdot N)_\infty \to (H \cdot N)_\infty$ as $n \to \infty$ in L^2. This proves (3.3) for all H bounded, which completes the proof. □

4. Stochastic calculus

4.1. Ito's formula.

Theorem 4.1 (Integration by parts). *Let X, Y be continuous semimartingales. Then*

$$X_tY_t - X_0Y_0 = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + [X,Y]_t.$$

Proof. Since both sides are continuous, it suffices to prove the result for t of the form $t = M2^{-N}$ for $M, N \in \mathbb{N}$. Note that for $s \le t$ we have that

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

Applying this for $n \ge N$, we have that

$$\begin{aligned} X_tY_t - X_0Y_0 = \sum_{k=0}^{M2^{n-N}} \Big(& X_{k2^{-n}}(Y_{(k+1)2^{-n}} - Y_{k2^{-n}}) + Y_{k2^{-n}}(X_{(k+1)2^{-n}} - X_{k2^{-n}}) \\ & + (X_{(k+1)2^{-n}} - X_{k2^{-n}})(Y_{(k+1)2^{-n}} - Y_{k2^{-n}})\Big) \\ & \to (X \cdot Y)_t + (Y \cdot X)_t + [X,Y] \quad \text{ucp as} \quad n \to \infty. \end{aligned}$$

□

Note that the covariation term does not appear in the usual integration by parts formula for Lebesgue-Stieljes integrals. Note also that if either X or Y is of finite variation then the covariation term also does not appear.

Theorem 4.2 (Itô's formula). *Let $X^1, \ldots, X^d$ be continuous semimartingales and let $X = (X^1, \ldots, X^d)$. Let $f \colon \mathbb{R}^d \to \mathbb{R}$ be C^2. Then*

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

Remark 4.3. *(1) Note that Theorem 4.2 implies that $f(X_t)$ is a semimartingale since each of the summands in Itô's formula is a semimartingale.*

(2) One can derive Itô's formula using Taylor expansion. Indeed, we have that

$$\begin{aligned} f(X_t) &= f(X_0) + \sum_{k=0}^{\lfloor 2^n t \rfloor} \big(f(X_{(k+1)2^{-n}}) - f(X_{k2^{-n}})\big) + \big(f(X_t) - f(X_{\lfloor 2^n t \rfloor})\big) \\ &= f(X_0) + \sum_{k=0}^{\lfloor 2^n t \rfloor} f'(X_{k2^{-n}})(X_{(k+1)2^{-n}} - X_{k2^{-n}}) + \frac{1}{2} \sum_{k=0}^{\lfloor 2^n t \rfloor} f''(X_{k2^{-n}})(X_{(k+1)2^{-n}} - X_{k2^{-n}})^2 + \quad \text{error} \\ &\to f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d[X]_s \quad \textit{ucp as} \quad n \to \infty. \end{aligned}$$

This is Itô's formula for $d = 1$.

Proof of Theorem 4.2. We are going to give the proof of Itô's formula for $d = 1$. We leave the proof for $d > 1$ as an exercise. Write $X = X_0 + M + A$ where A has corresponding total variation process V. Let $T_r = \inf\{t \ge 0 : |X_t| + V_t + [M]_t > r\}$. Then (T_r) is a sequence of stopping times with

$T_r \uparrow \infty$ as $r \to \infty$. It suffices to prove the formula for arbitrary $f \in C^2(\mathbb{R})$ in the time-interval $[0, T_r]$. Let $\mathcal{A}$ denote the subset of $C^2(\mathbb{R})$ where the formula holds. Then:

(a) $\mathcal{A}$ contains the functions $f(x) \equiv 1$ and $f(x) = x$
(b) $\mathcal{A}$ is a vector space
(c) We will show that $\mathcal{A}$ is in fact an algebra. That is, if $f, g \in \mathcal{A}$ then $fg \in \mathcal{A}$.
(d) Finally, we will show that if (f_n) is a sequence in $\mathcal{A}$ with $f_n \to f$ in $C^2(B_r)$ for each $r > 0$, then $f \in \mathcal{A}$. Here, $B_r = \{x : |x| < r\}$ and $f_n \to f$ in $C^2(B_r)$ means that with

$$\Delta_{n,r} = \max\{\sup_{x \in B_r} |f_n(x) - f(x)|, \sup_{x \in B_r} |f_n'(x) - f'(x)|, \sup_{x \in B_r} |f_n''(x) - f'(x)|\}$$

we have that $\Delta_{n,r} \to 0$ as $n \to \infty$ with r fixed.

We note that (a), (b), (c) together imply that all polynomials are in $\mathcal{A}$. By Weierstrass' approximation theorem, polynomials are dense in $C^2(B_r)$ for each $r > 0$, so (d) implies that $\mathcal{A} = C^2(\mathbb{R})$.

Proof of (c) Suppose that $f, g \in \mathcal{A}$. Set $F_t = f(X_t)$ and $G_t = g(X_t)$. Since Itô's formula holds, F and G must be continuous semimartingales. By integration by parts,

$$F_t G_t - F_0 G_0 = \int_0^t F_s dG_s + \int_0^t G_s dF_s + [F, G]_t. \tag{4.1}$$

Using that $H \cdot (K \cdot M) = (HK) \cdot M$ and Itô's formula for g, we also have that

$$\int_0^t F_s dG_s = \int_0^t f(X_s) g'(X_s) dX_s + \frac{1}{2} \int_0^t f(X_s) g''(X_s) d[X]_s. \tag{4.2}$$

By the Kunita-Watanabe identity, we have that $[H \cdot M, N] = H \cdot [M, N]$, hence

$$[F, G]_t = [f(X), g(X)]_t = \int_0^t f'(X_s) g(X_s) d[X]_s. \tag{4.3}$$

Inserting (4.2) and (4.3) into (4.1) implies that Itô's formula holds for fg. That is, $fg \in \mathcal{A}$.

Proof of (d) Suppose that (f_n) is a sequence in $\mathcal{A}$ and we have that $f_n \to f$ in $C^2(B_r)$ for all $r > 0$. Then

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

Consequently,

$$\int_0^{t \wedge T_r} f_n'(X_s) dA_s + \frac{1}{2} \int_0^{t \wedge T_r} f_n''(X_s) d[M]_s \to \int_0^{t \wedge T_r} f'(X_s) dA_s + \frac{1}{2} \int_0^{t \wedge T_r} f''(X_s) d[M]_s.$$

Moreover, $M^{T_r} \in \mathcal{M}_c^2$, so

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

Therefore $(f_n(X) \cdot M)^{T_r} \to (f(X) \cdot M)^{T_r}$ in $\mathcal{M}_c^2$. Hence, for any r, we can pass to the limit in Itô's formula to obtain

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

□

Example 4.4. *Let $X = B$ where B is a standard Brownian motion in $\mathbb{R}$ and let $f(x) = x^2$. Applying Itô's formula, we then have that*

$$B_t^2 = 2\int_0^t B_s dB_s + t.$$

Rearranging, we see that

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

Example 4.5. *Let $f\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ be $C^{1,2}$ and let $X_t = (t, B_t^1, \ldots, B_t^d)$ where $B^1, \ldots, B^d$ are independent standard Brownian motions. Then, by Itô's formula, we have that*

$$f(t, B_t) - f(0, B_0) - \int_0^t \left(\frac{1}{2}\Delta + \frac{\partial}{\partial t}\right) f(s, B_s) ds = \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x_i} f(s, B_s) dB_s^i \in \mathcal{M}_{c,loc}.$$

In particular, if f does not depend on t and f is harmonic in x, then $f(B_t) \in \mathcal{M}_{c,loc}$.

4.2. **Stratonovich integral.** Let X, Y be continuous semimartingales. We define the *Stratonovich integral* as

$$\int_0^t X_s \partial Y_s = \int_0^t X_s dY_s + \frac{1}{2}[X, Y]_t.$$

Here, the integral on the right hand side is an Itô integral. Using the Riemann sum approximations for the terms on the right hand side, we have that

$$\sum_{k=0}^{\lfloor 2^n t \rfloor - 1} \left(\frac{X_{(k+1)2^{-n}} - X_{k2^{-n}}}{2}\right)(Y_{(k+1)2^{-n}} - Y_{k2^{-n}}) \to \int_0^t X_s \partial Y_s \quad \text{ucp as} \quad n \to \infty.$$

Proposition 4.6. *Let $X^1, \ldots, X^d$ be continuous semimartingales and let $f\colon \mathbb{R}^d \to \mathbb{R}$ be C^3. Then*

$$f(X_t) = f(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) \partial X_s^i. \tag{4.4}$$

In particular, the integration by parts formula is given by

$$X_t Y_t - X_0 Y_0 = \int_0^t X_s \partial Y_s + \int_0^t Y_s \partial X_s. \tag{4.5}$$

Note that Proposition 4.6 implies that the Stratonovich integral satisfies the usual rules of calculus.

Proof of Proposition 4.6. We will treat the case that $d = 1$ and leave the case $d \geq 2$ as an exercise. By Itô's formula, we have that

$$f(X_t) = f(X_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2}\int_0^t f''(X_s) d[X]_s \tag{4.6}$$

$$f'(X_t) = f'(X_0) + \int_0^t f''(X_s) dX_s + \frac{1}{2}\int_0^t f'''(X_s) d[X]_s. \tag{4.7}$$

By the Kunita-Watanabe identity and (4.7), we have that $[f'(X_s), X] = [\int_0^t f''(X_s)dX_s, X] = f''(X) \cdot [X]$. Combining this with (4.6) and the definition of the Stratonovich integral, we then have that

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\partial X_s.$$

□

Remark 4.7. *Note that in Proposition 4.6, we require more regularity of the function f in order to obtain* (4.4) *in comparison to Itô's formula. We also emphasize that for the Stratonovich integral, the integrand must be a semimartingale. This is also in contrast to the case of Itô's formula, in which case the integrand need only be locally bounded and previsible. Finally, the Stratonovich integral with respect to a local martingale is not necessarily a local martingale. For example,*

$$\int_0^t B_s \partial B_s = \int_0^t B_s dB_s + \frac{1}{2}t = \frac{1}{2}B_t^2,$$

which is not a local martingale.

4.3. **Shorthand notation.** We will now record some shorthand notation which is common for stochastic calculus.

- $Z_t - Z_0 = \int_0^t H_s dX_s$ is equivalent to $dZ_t = H_t dX_t$.
- $Z_t - Z_0 = \int_0^t H_s \partial X_s$ is equivalent to $\partial Z_t = H_t \partial X_t$.
- $Z_t = [X, Y]_t = \int_0^t d[X, Y]_s$ is equivalent to $dZ_t = dX_t dY_t$.

We then have the following computational rules:

- $H_t(K_t dX_t) = (H_t K_t)dX_t$
- $H_t(dX_t dY_t) = (H_t dX_t)dYt$
- $d(X_t Y_t) = X_t dY_t + Y_t dX_t + dX_t dY_t$
- $df(X_t) = \sum_{i=1}^d \frac{\partial f}{\partial x_i}(X_t)dX_t^i + \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t)dX_t^i dX_t^j$.

The relationship between the Itô and Stratonovich integrals is that $\partial Z_t = Y_t \partial X_t$ if and only if $dZ_t = Y_t dX_t + \frac{1}{2}dX_t dY_t$.

5. Applications

5.1. **Brownian motion.** Throughout, we will work on 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.

Theorem 5.1 (Lévy's characterization of Brownian motion)**.** *Let $X^1, \ldots, X^d$ be continuous local martingales and set $X = (X^1, \ldots, X^d)$. Suppose that $[X^i, X^j]_t = \delta_{ij}t$ for all i, j and $t \geq 0$. Then X is a standard Brownian motion on $\mathbb{R}^d$.*

Proof. It suffices to show that, for all $0 \leq s \leq t < \infty$, we have that $X_t - X_s$ is distributed as a $N(0, (t-s)I)$ random variable and is independent of $\mathcal{F}_s$. This is equivalent to show that

$$\mathbb{E}[\exp(i(\theta, X_t - X_s)) \mid \mathcal{F}_s] = \exp\left(-\frac{1}{2}\theta^2(t-s)\right), \quad \text{for all} \quad \theta \in \mathbb{R}^d.$$

Fix $\theta \in \mathbb{R}^d$ and set $Y_t = (\theta, X_t) = \sum_{j=1}^d \theta^j X_t^j$. Then $Y \in \mathcal{M}_{c,loc}$ as $\mathcal{M}_{c,loc}$ is a vector space. By assumption, we have that

$$[Y]_t = [Y,Y]_t = [\sum_{j=1}^d \theta^j X^j, \sum_{j=1}^d \theta^j X^j] = \sum_{j,k=1}^d \theta^j \theta^k [X^j, X^k] = |\theta|^2 t.$$

Consider

$$Z_t = \exp(iY_t + \tfrac{1}{2}[Y]_t) = \exp(i(\theta, X_t) + \tfrac{1}{2}|\theta|^2 t).$$

By Itô's formula with $X = iY_t + \frac{1}{2}[Y]_t$ and $f(x) = \exp(x)$, we have that

$$dZ_t = Z_t(idY_t + \tfrac{1}{2}d[Y]_t) - \tfrac{1}{2}Z_t d[Y]_t = iZ_t dY_t.$$

Consequently, $Z \in \mathcal{M}_{c,loc}$. Moreover, Z is bounded on $[0,t]$ for each $t \geq 0$, hence $Z \in \mathcal{M}_c$. Therefore $\mathbb{E}[Z_t \,|\, \mathcal{F}_s] = Z_s$, which implies that

$$\mathbb{E}[\exp(i(\theta, X_t - X_s)) \,|\, \mathcal{F}_s] = \exp(-\tfrac{1}{2}\theta^2(t-s)).$$

□

Theorem 5.2 (Dubins-Schwartz theorem)**.** *Let $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$ and $[M]_\infty = \infty$. Set $\tau_s = \inf\{t \geq 0 : [M]_t > s\}$, $B_s = M_{\tau_s}$, and $\mathcal{G}_s = \mathcal{F}_{\tau_s}$. Then τ_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 and $M_t = B_{[M]_t}$.*

Remark 5.3. *Theorem 5.2 states that any continuous local martingale starting from* 0 *can be represented as a stochastic time-change of Brownian motion. Thus, in some sense, Brownian motion is the "most general" continuous local martingale.*

Proof of Theorem 5.2. Since $[M]$ is continuous and adapted, τ_s is a stopping time for each $s \geq 0$. The hypothesis $[M]_\infty = \infty$ implies that $\tau_s < \infty$ for all $s \geq 0$. Note that B is adapted to $(\mathcal{G}_s)_{s \geq 0}$. To finish the proof, we will show that B is continuous and that it is a local martingale with $[B]_t = t$. The result will then follow from Theorem 5.1.

Step 1: B is continuous. We first note that $s \mapsto \tau_s$ is non-decreasing and cádlág. Therefore B is cádlág Moreover, $\tau_{s^-} = \inf\{t \geq 0 : [M]_t = s\}$. To prove the continuity of B, we must therefore show that $B_{s^-} = B_s$ for all $s \geq 0$. Equivalently, $M_{\tau_{s^-}} = M_{\tau_s}$ for all $s \geq 0$. In order to show that this is the case, we need to show that if $[M]$ is constant on a given interval then M is also constant on the same interval. By applying localization, we may assume that $M \in \mathcal{M}_c^2$. It suffices to show that for all $q \in \mathbb{Q}$, $q > 0$, with $S_q = \inf\{t > q : [M]_t > [M]_q\}$ we have that M is constant $[q, S_q]$.. We know that $M^2 - [M]$ is a UI martingale as in $M \in \mathcal{M}_c^2$. By the OST, we have that

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

Since $[M]_{S_q} = M_q$ and M is a martingale, we also have that

$$\mathbb{E}[(M_{S_q} - M_q)^2 \,|\, \mathcal{F}_q] = \mathbb{E}[[M]_{S_q} - M_q \,|\, \mathcal{F}_q] = 0.$$

Therefore M is almost surely constant on $[q, S_q]$, hence B is continuous.

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

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

Combining, we have that $B \in \mathcal{M}_c$ with $[B]_t = t$ for all $t \geq 0$. Thus by Theorem 5.1, we have that B is a $(\mathcal{G}_t)_{t\geq 0}$-Brownian motion. □

5.2. **Exponential martingales.**

Example 5.4. *Let $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$. Set $Z_t = \exp(M_t - \frac{1}{2}[M]_t)$. By Itô's formula applied to $X = M - \frac{1}{2}[M]$ and $f(x) = \exp(x)$, we have that*

$$dZ_t = Z_t(dM_t - \tfrac{1}{2}d[M]_t) + \tfrac{1}{2}Z_t d[M]_t = Z_t dM_t.$$

Consequently, $Z_t \in \mathcal{M}_{c,loc}$ as the integral of Z with respect to an element of $\mathcal{M}_{c,loc}$ is an element of $\mathcal{M}_{c,loc}$. We call Z the stochastic exponential (or exponential martingale associated with) M and write $Z = \mathcal{E}(M)$.

Proposition 5.5. *Let $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$. Then for all $\epsilon, \delta > 0$ we have that*

$$\mathbb{P}[\sup_{t\geq 0} M_t \geq \epsilon,\ [M]_\infty \leq \delta] \leq \exp\left(-\frac{\epsilon^2}{2\delta}\right).$$

Proof. Fix $\epsilon > 0$ and set $T = \inf\{t \geq 0 : M_t \geq \epsilon\}$. Fix $\theta \in \mathbb{R}$ and set $Z_t = \exp(\theta M_t^T - \frac{1}{2}\theta^2[M]_t^T)$. Then $Z_t \in \mathcal{M}_{c,loc}$ with $|Z_t| \leq e^{\theta\epsilon}$ for all $t \geq 0$. Consequently, Z is in fact a bounded martingale. Therefore $\mathbb{E}[Z_\infty] = Z_0 = 1$. For each $\delta > 0$, we have that

$$\begin{aligned}\mathbb{P}[\sup_{t\geq 0} M_t \geq \epsilon,\ [M]_\infty \leq \delta] &\leq \mathbb{P}[Z_\infty \geq e^{\theta\epsilon - \theta^2\delta/2}] \\ &\leq \exp(-\theta\epsilon + \theta^2\delta/2) \quad \text{(Markov's inequality)}.\end{aligned}$$

The result follows by optimizing this bound over θ. □

Proposition 5.6. *Let $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$. Suppose that $[M]$ is uniformly bounded. Then $\mathcal{E}(M)$ is a UI martingale.*

Proof. Let c be such that $[M]_\infty \leq c$. By Proposition 5.5, we have that

$$\mathbb{P}[\sup_{t\geq 0} M_t \geq \epsilon] = \mathbb{P}[\sup_{t\geq 0} M_t \geq \epsilon,\ [M]_\infty \leq c] \leq \exp\left(-\frac{\epsilon^2}{2c}\right).$$

Note that

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

as $[M]_t \geq 0$ for all $t \geq 0$. Moreover,

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

Therefore $\mathcal{E}(M)$ is UI, as desired. □

Suppose that $\mathbb{P}$ and $\mathbb{Q}$ are two probability measures on $(\Omega, \mathcal{F})$. Then we say that $\mathbb{Q}$ is *absolutely continuous* with respect to $\mathbb{P}$ and write $\mathbb{Q} \ll \mathbb{P}$ if for any $A \in \mathcal{F}$, we have that $\mathbb{P}[A] = 0$ implies $\mathbb{Q}[A] = 0$.

We define the *stochastic exponential* (or *exponential martingale*) of $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$ by $\mathcal{E}(M)_t = \exp(M_t - \frac{1}{2}[M]_t)$.

We showed that if $[M]$ is uniformly bounded, then $\mathcal{E}(M)$ is a UI martingale. This provides one circumstance in which the hypotheses of the following theorem are satisfied.

Theorem 5.7 (Girsanov's theorem)**.** *Let $M \in \mathcal{M}_{c,loc}$ be such that $M_0 = 0$. Suppose that $Z = \mathcal{E}(M)$ is a UI martingale. Then we can define a new probability measure $\widetilde{\mathbb{P}}$ on $(\Omega, \mathcal{F})$ by $\widetilde{\mathbb{P}}[A] = \mathbb{E}[Z_\infty \mathbf{1}(A)]$ for $A \in \mathcal{F}$. If $X \in \mathcal{M}_{c,loc}(\mathbb{P})$, then $X - [X, M] \in \mathcal{M}_{c,loc}(\widetilde{\mathbb{P}})$.*

Proof. Since Z is UI, we know that Z_∞ exists with $Z_\infty \geq 0$ and $\mathbb{E}[Z_\infty] = \mathbb{E}[Z_0] = 1$. It is not difficult to see that $\widetilde{\mathbb{P}}$ is a probability measure with $\widetilde{\mathbb{P}} \ll \mathbb{P}$. Let $X \in \mathcal{M}_{c,loc}(\mathbb{P})$ and set $T_n = \inf\{t \geq 0 : |X_t - [X, M]_t] \geq n\}$. As $X - [X, M]$ is continuous, we have that $\mathbb{P}[T_n \uparrow \infty] = 1$. By absolute continuity, we therefore have that $\widetilde{\mathbb{P}}[T_n \uparrow \infty] = 1$. So, to show that $Y = X - [X, M] \in \mathcal{M}_{c,loc}(\widetilde{\mathbb{P}})$, it suffices to show that $Y^{T_n} = X^{T_n} - [X^{T_n}, M] \in \mathcal{M}_{c,loc}(\widetilde{\mathbb{P}})$ for all n. Replacing X by X^{T_n}, we can reduce to the case where Y is uniformly bounded.

By integration by parts, we have that

$$\begin{aligned} d(Z_t Y_t) = Y_t dZ_t + Z_t dY_t &= (X_t - [X, M]_t) Z_t dM_t + Z_t(dX_t - dX_t dM_t) + Z_t dM_t dX_t \\ &= Z_t X_t dM_t - Z_t [X, M]_t dM_t + Z_t dX_t. \end{aligned}$$

Consequently, $Y \in \mathcal{M}_{c,loc}$. Also, $\{Z_T : T \quad \text{is a stopping time}\}$ is UI. Since Y is bounded, we also have that $\{Z_T Y_T : T \quad \text{is a stopping time}\}$ is UI as well. Hence, $ZY \in \mathcal{M}_c(\mathbb{P})$. But, for $s \leq t$, we have that

$$\widetilde{\mathbb{E}}[Y_t - Y_s \,|\, \mathcal{F}_s] = \mathbb{E}[Z_\infty Y_t - Z_\infty Y_s \,|\, \mathcal{F}_s] = \mathbb{E}[Z_t Y_t - Z_s Y_s \,|\, \mathcal{F}_s] = 0,$$

as $ZY \in \mathcal{M}_c(\mathbb{P})$. Therefore $Y \in \mathcal{M}_c(\widetilde{\mathbb{P}})$, as required. □

Remark 5.8. *Note that the quadratic variation does not change when performing a change of measures from $\mathbb{P}$ to $\widetilde{\mathbb{P}}$.*

Corollary 5.9. *Let B be a standard Brownian motion under $\mathbb{P}$ and let $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$. Suppose that $Z = \mathcal{E}(M)$ is a UI martingale and $\widetilde{\mathbb{P}}[A] = \mathbb{E}[Z_\infty \mathbf{1}_A]$ for $A \in \mathcal{F}$. Then $\widetilde{B} = B - [B, M]$ is a $\widetilde{\mathbb{P}}$-Brownian motion.*

Proof. Since $\widetilde{B}$ is a continuous $\widetilde{\mathbb{P}}$-local martingale and has $[\widetilde{B}]_t = [B]_t = t$, it follows from the Lévy characterization that $\widetilde{B}$ must be a $\widetilde{\mathbb{P}}$-Brownian motion. □

Let $(W, \mathcal{W}, \mu)$ be the Wiener space, i.e., $W = C(\mathbb{R}_+, \mathbb{R})$, $\mathcal{W} = \sigma(X_t : t \geq 0)$, where $X_t \colon W \to \mathbb{R}$ is given by $X_t(w) = w(t)$, μ is the Wiener measure, the unique probability measure on $(W, \mathcal{W})$ such that $(X_t)_{t \geq 0}$ is a standard Brownian motion starting from 0.

Let

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

be the *Cameron-Martin space*. For $h \in H$, we will write $\dot{h}$ for φ.

Theorem 5.10 (Cameron-Martin)**.** *Fix $h \in H$ and set $\mu^h(A) = \mu(\{w \in W : w + h \in A\})$ for $A \in \mathcal{W}$. Then μ^h is a probability measure on $(W, \mathcal{W})$ which is absolutely continuous with respect to μ and with density*

$$\frac{d\mu^h}{d\mu}(w) = \exp\left(\int_0^\infty \dot{h}(s)dw(s) - \frac{1}{2}\int_0^\infty |\dot{h}(s)|^2 ds\right).$$

Remark 5.11. *In other words, if we take a Brownian motion and shift it by a deterministic function $h \in H$, then the resulting process has a law which is absolutely continuous wrt that of the original Brownian motion.*

Proof of Theorem 5.10. Set $\mathcal{W}_t = \sigma(X_s : s \leq t)$ and $M_t = \int_0^t \dot{h}(s)dX_s$. Then $M \in \mathcal{M}_c^2(W, \mathcal{W}, (\mathcal{W}_t)_{t\geq 0}, \mu)$ and $[M]_\infty = \int_0^\infty |\dot{h}(s)|^s ds =: C < \infty$.

We know that $\mathcal{E}(M)$ is a UI martingale, so we can define a measure $\widetilde{\mu}$ on $(W, \mathcal{W})$ by

$$\frac{d\widetilde{\mu}}{d\mu}(\omega) = \exp(M_\infty(\omega) - \tfrac{1}{2}[M]_\infty(\omega))$$

and $\widetilde{X} = X - [X, M] \in \mathcal{M}_{c,loc}(\widetilde{\mu})$ by Girsanov's theorem. Since X is a μ-Brownian motion, it follows that $\widetilde{X}$ is a $\widetilde{\mu}$-Brownian motion. But $[X, M]_t = \int_0^t \dot{h}(s)ds = h(t)$. So, $\widetilde{X}(w) = X(w) - h = w - h$. Hence, for $A \in \mathcal{W}$, $\mu^h(A) = \mu(\{w : X(w) + h \in A\}) = \widetilde{\mu}(\{w : \widetilde{X}(w) + h \in A\}) = \widetilde{\mu}(A)$, Hence, $\widetilde{\mu} = \mu^h$, as required. ☐

6. Stochastic differential equations

6.1. **Motivation.** Suppose that we have a differential equation, for example $dx(t)/dt = b(x(t))$. Equivalently, $x(t) = x(0) + \int_0^t b(x(s))ds$. In certain situations, it can be natural to take into account random perturbations and add a "noise term" so that we have the equation $X_t = X_0 + \int_0^t b(X_s)ds + \sigma B_t$ where B is a Brownian motion and σ is a parameter which controls the intensity of the noise. If the intensity depends on the state of the system, it may make more sense to consider an equation of the form

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

where the last term is an Ito integral. This is an example of a stochastic differential equation (SDE) which is written in integral form. The corresponding differential form of the SDE is

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

6.2. **General definitions.** Let $M^{d\times m}(\mathbb{R})$ denote the set of $d\times m$ matrices with real entries. Suppose that $\sigma\colon \mathbb{R}^d \to M^{d\times m}(\mathbb{R})$ and $b\colon \mathbb{R}^d \to \mathbb{R}^d$ are measurable functions which are bounded on compact sets. We write $\sigma(x) = (\sigma_{ij}(x))$ and $b(x) = (b_i(x))$. Consider the equation

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt. \tag{6.1}$$

We can write this equation component-wise as

$$dX_t^i = \sum_{j=1}^m \sigma_{ij}(X_t)dB_t^j + b_i(X_t)dt.$$

A solution to (6.1) 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 $(\mathcal{F}_t)$-Brownian motion $B = (B^1, \ldots, B^m)$ in $\mathbb{R}^m$.
- An $(\mathcal{F}_t)$-adapted continuous process $X = (X^1, \ldots, X^d)$ in $\mathbb{R}^d$ such that

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

When, in addition, $X_0 = x \in \mathbb{R}^d$, we say that X is a solution started from x.

There are several different notions of existence and uniqueness for solutions to SDEs.

We say that an SDE has a *weak solution* if for all $x \in \mathbb{R}^d$, there exists a solution to the equation started from x. There is *uniqueness in law* if all solutions to the SDE started from 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 and X' such that $X_0 = X_0'$ are such that $\mathbb{P}[X_t = X_t' \quad \text{for all} \quad t \geq 0] = 1$. We say that a solution X of a SDE started from x is a *strong solution* if X is adapted to the filtration generated by B.

In the case of a weak solution, the filtration $(\mathcal{F}_t)_{t\geq 0}$ is not necessarily the filtration generated by B. In other words, it is not necessarily true that $X_t(\omega)$ is a measurable function of $(B_s(\omega) : s \leq t)$ and x.

Example 6.1. *It is possible to have the existence of a weak solution and uniqueness in law without pathwise uniqueness.*

Suppose that β is a Brownian motion in $\mathbb{R}$ with $\beta_0 = x$. Set $B_t = \int_0^t \operatorname{sgn}(B_s) d\beta_s$. We take the convention that $\operatorname{sgn}(x) = \mathbf{1}_{(0,\infty)}(x) - \mathbf{1}_{(-\infty,0]}(x)$. Note that $(\operatorname{sgn}(\beta_t))_{t\geq 0}$ is previsible, so that the integral is well-defined. Note that

$$x + \int_0^t \operatorname{sgn}(\beta_s) dB_s = x + \int_0^t (\operatorname{sgn}(\beta_s))^2 d\beta_s = x + \int_0^t d\beta_s = \beta_t.$$

Consequently, β solves the SDE

$$dX_t = \operatorname{sgn}(X_t) dB_t, \quad X_0 = x. \tag{6.2}$$

Therefore (6.2) *has a weak solution. By the Lévy characterization of Brownian motion, any solution to this SDE is a Brownian motion. Therefore we have uniqueness in law.*

On the other hand, there is not pathwise uniqueness. Indeed, to see this we first observe that $\int_0^t \mathbf{1}(\beta_s = 0) ds = 0$ as the zero set of Brownian motion has zero Lebesgue measure. Thus, by the Itô isometry, we have that $\int_0^t \mathbf{1}(\beta_s = 0) dB_s = 0$.

Suppose that $\beta_0 = 0$. Then both β and $-\beta$ are solutions to (6.2)*. So, pathwise uniqueness does not hold. It turns out that β is not a strong solution either.*

6.3. **Lipschitz coefficients.** For $U \subseteq \mathbb{R}^d$ and $f \colon U \to \mathbb{R}^n$, we say that f is *Lipschitz* if there exists $K < \infty$ such that $|f(x) - f(y)| \leq K|x - y|$ for all $x, y \in U$.

Suppose that $A \in M^{d\times m}(\mathbb{R})$ with $A = (a_{ij})$. Then we will use the norm given by

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

So, if $f\colon U \to M^{d\times m}(\mathbb{R})$ instead, then we say that f Lipschitz if there exists $K < \infty$ such that $|f(x) - f(y)| \le K|x-y|$ for all $x, y \in U$.

We will establish the existence and uniqueness of solutions to SDEs when the coefficients σ and b are Lipschitz. We write $\mathcal{C}_T$ for the set of continuous, adapted processes $X\colon [0,T] \to \mathbb{R}$ such that $|||X|||_T = \|\sup_{t\le T} |X_t|\|_2 < \infty$. We write $\mathcal{C}$ for the set of continuous, adapted processes $X\colon [0,T] \to \mathbb{R}$ such that $|||X|||_T < \infty$ for all $T > 0$.

Recall that $\mathcal{C}_T$ is complete.

In our proof the existence of solutions to SDEs with Lipschitz coefficients, we will make use of the following theorem which we state without proof.

Theorem 6.2 (Contraction mapping theorem)**.** *Let (X, d) be a complete metric space.*

(a) Suppose that $F\colon X \to X$ is a contraction. That is, there exists $r < 1$ such that $d(F(x), F(y)) \le rd(x,y)$ for all $x, y \in X$. Then F has a unique fixed point.
(b) Suppose that $F\colon X \to X$ is such that there exists $r < 1$ and $n \in \mathbb{N}$ with $d(F^n(x), F^n(y)) < rd(x,y)$ for all $x, y \in X$, then F has a unique fixed point.

Theorem 6.3. *Suppose that $\sigma\colon \mathbb{R}^d \to M^{d\times m}(\mathbb{R})$ and $b\colon \mathbb{R}^d \to \mathbb{R}^d$ are Lipschitz. Then there is pathwise uniqueness for the SDE $dX_t = \sigma(X_t)dB_t + b(X_t)dt$. Moreover, for each filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{t\ge 0}, \mathbb{P})$ which satisfies the usual conditions and each $(\mathcal{F}_t)$-Brownian motion B, there exists for each $x \in \mathbb{R}^d$ a strong solution starting from x.*

Lemma 6.4 (Gronwall's lemma)**.** *Let $T > 0$ and let f be a non-negative, bounded, measurable function on $[0,T]$. Suppose that there exists $a, b \ge 0$ such that*

$$f(t) \le a + b\int_0^t f(s)ds \quad \textit{for all} \quad t \in [0,T].$$

Then

$$f(t) \le ae^{bt} \quad \textit{for all} \quad t \in [0,T].$$

Proof. Let $C = \sup_{0\le t\le T} f(t) < \infty$. By induction on n, we have that

$$f(t) \le a\left(1 + bt + \cdots + \frac{(bt)^n}{n!}\right) + b^{n+1}\int_0^t \int_0^{s_1} \cdots \int_0^{s_n} f(s_{n+1})ds_{n+1}\cdots ds_1.$$

The last term on the right hand side is bounded from above by $C(bt)^{n+1}/(n+1)! \to 0$ as $n \to \infty$. $\square$

Proof of Theorem 6.3. We will give the proof in the case that $d, m = 1$ and leave the general case as an exercise.

We assume that we have fixed a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$ and a standard $(\mathcal{F}_t)$-Brownian motion B. We also let $(\mathcal{F}_t^B)$ be the filtration generated by B; note that $\mathcal{F}_t^B \subseteq \mathcal{F}_t$ for all $t \ge 0$.

Suppose that $K > 0$ is such that

$$|\sigma(x) - \sigma(y)| \le K|x-y| \quad \text{and} \quad |b(x) - b(y)| \le K|x-y| \quad \text{for all} \quad x, y \in \mathbb{R}.$$

Uniqueness. Suppose that X, X' are solutions to the SDE with $X_0 = X_0'$. We will show that $X_t = X_t'$ for all $t \geq 0$. Fix $M > 0$ and let $\tau = \inf\{t \geq 0 : |X_t| \geq M \quad \text{or} \quad |X_t'| \geq M\}$. Then we have that

$$\begin{aligned} X_{t\wedge\tau} &= X_0 + \int_0^{t\wedge\tau} \sigma(X_s)dB_s + \int_0^{t\wedge\tau} b(X_s)ds \\ X'_{t\wedge\tau} &= X'_0 + \int_0^{t\wedge\tau} \sigma(X'_s)dB_s + \int_0^{t\wedge\tau} b(X'_s)ds. \end{aligned}$$

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

$$\begin{aligned} \mathbb{E}[(X_{t\wedge\tau} - X'_{t\wedge\tau})^2] &\leq 2\mathbb{E}\left[\left(\int_0^{t\wedge\tau}(\sigma(X_s) - \sigma(X'_s))dB_s\right)^2\right] + 2\mathbb{E}\left[\left(\int_0^{t\wedge\tau}(b(X_s) - b(X'_s))ds\right)^2\right] \\ &\leq 2\mathbb{E}\left[\int_0^{t\wedge\tau}(\sigma(X_s) - \sigma(X'_s))^2 ds\right] + 2T\mathbb{E}\left[\int_0^{t\wedge\tau}(b(X_s) - b(X'_s))^2 ds\right] \\ &\leq 2K^2(1+T)\mathbb{E}\left[\int_0^{t\wedge\tau}(X_s - X'_s)^2 ds\right] \\ &\leq 2K^2(1+T)\int_0^t \mathbb{E}[(X_{s\wedge\tau} - X'_{s\wedge\tau})^2]ds. \end{aligned}$$

Let $f(t) = \mathbb{E}[(X_{t\wedge\tau} - X'_{t\wedge\tau})^2]$. Then we have shown that

$$f(t) \leq 2K^2(1+T)\int_0^t f(s)ds \quad \text{for all} \quad t \in [0,T].$$

Hence by Gronwall's lemma, we have that $f(t) = 0$ for all $t \in [0,T]$. This implies that $X_{t\wedge\tau} = X'_{t\wedge\tau}$ for all $t \in [0,T]$. Since $M > 0$ was arbitrary, we therefore have that $X_t = X'_t$ for all $t \geq 0$.

Existence of a strong solution.

Fix $x \in \mathbb{R}$. Recall that $|||X|||_T = \|\sup_{0\leq t\leq T}|X_t|\|_2$ and that $\mathcal{C}_T$ (resp. $\mathcal{C}$) denotes the set of continuous, $(\mathcal{F}_t)$-adapted processes X such that $|||X|||_T < \infty$ (resp. $|||X|||_T < \infty$ for all $T > 0$). Using the Lipschitz property of σ and b, we note that for all $y \in \mathbb{R}$ we have that

$$|\sigma(y)| = |\sigma(y) - \sigma(0) + \sigma(0)| \leq |\sigma(0)| + |\sigma(y) - \sigma(0)| \leq |\sigma(0)| + K|y|. \tag{6.3}$$

We similarly have that

$$|b(y)| \leq |b(0)| + K|y|. \tag{6.4}$$

Fix $T > 0$ and $X \in \mathcal{C}_T$. Let $M_t = \int_0^t \sigma(X_s)dB_s$, $0 \leq t \leq T$. Then $[M]_T = \int_0^T \sigma^2(X_s)ds$. Consequently, by (6.3) we have that

$$\mathbb{E}[[M]_T] \leq 2T(|\sigma(0)|^2 + K^2|||X|||_T^2) < \infty.$$

It thus follows that $(M_t)_{t\in[0,T]}$ is an L^2-bounded martingale. Doob's L^2 inequality then implies that

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

Using (6.4) and the Cauchy-Schwarz inequality, we also have that

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|\int_0^t b(X_s)ds\right|^2\right] \leq T\mathbb{E}\left[\int_0^T b^2(X_s)ds\right] \leq 2T^2\left(|b(0)|^2 + K|||X|||_T^2\right) < \infty.$$

Combining, we see that the map F defined on $\mathcal{C}_T$ by

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

takes values in $\mathcal{C}_T$. Note that for $X, Y \in \mathcal{C}_T$, similar arguments to those above imply that

$$|||F(X) - F(Y)|||_t \le 2K^2(4+T)\int_0^t |||X - Y|||_s^2 ds.$$

In particular, there exists a constant $C > 0$ such that

$$|||F(X) - F(Y)|||_t^2 \le C|||X - Y|||_t^2, \quad \forall t \le T.$$

Thus by iterating, we have that

$$|||F^{(n)}(X) - F^{(n)}(Y)|||_T^2 \le \frac{C^n T^n}{n!}|||X - Y|||_T^2.$$

Note that we can make n sufficiently large so that $C^nT^n/n! < 1$. The contraction mapping theorem therefore implies that there exists a unique fixed point $X^{(T)} \in \mathcal{C}_T$ of F.

By uniqueness, we have that $X_t^{(T)} = X_t^{(T')}$ for all $t \le T \wedge T'$. We thus define $X_t = X_t^{(N)}$ whenever $t \le N$ and $N \in \mathbb{N}$. This is the pathwise unique solution to the SDE started from x. It remains to prove that X is $(\mathcal{F}_t^B)$-adapted.

We define a sequence (Y^n) in $\mathcal{C}_T$ by setting $Y_0 = x$ and $Y^n = F(Y^{n-1})$ for each $n \in \mathbb{N}$. Then Y^n is $(\mathcal{F}_t^B)_{t\ge 0}$ adapted for all n. Since $X = F^{(n)}(X)$ for all $n \ge 0$, it follows that

$$|||X - Y^n||| = |||F^{(n)}(X) - F^{(n)}(x)||| \le \frac{C^nT^n}{n!}|||X - x|||_T^2.$$

Consequently, it follows that

$$\mathbb{E}\left[\sum_{n=0}^{\infty} \sup_{0\le t\le T} |X_t - Y_t^n|^2\right] \le \sum_{n=0}^{\infty} |||X - Y^n|||_T^2 < \infty.$$

This implies that $\sum_{n=0}^{\infty} \sup_{0\le t\le T} |X_t - Y_t^n| < \infty$ almost surely, which implies that $Y^n \to X$ almost surely as $n \to \infty$ uniformly on $[0,T]$. Therefore X is also $(\mathcal{F}_t^B)$-adapted. That is, X is a strong solution. $\square$

Proposition 6.5. *Under the hypotheses of Theorem 6.3, there is uniqueness in law for the SDE $dX_t = \sigma(X_t)dB_t + b(X_t)dt$.*

Proof. See example sheet 3. $\square$

Example 6.6 (Ornstein-Uhlenbeck process)**.** *Fix $\lambda \in \mathbb{R}$ and consider the SDE in $\mathbb{R}^2$*

$$dV_t = dB_t - \lambda V_t dt, \quad V_0 = v_0$$
$$dX_t = V_t dt, \quad X_0 = x_0.$$

When $\lambda > 0$, this SDE models the movement of a pollen grain on the surface of a liquid. X represents the x-coordinate of the position and V is the velocity in the x-direction. The term $-\lambda V$ is the "damping" due to the viscosity of the liquid. Whenever $|V|$ gets large, the system acts to reduce $|V|$.

Theorem 6.3 implies that there exists a unique strong solution to this SDE started at each x_0, v_0. This is a rare example of an SDE that we can solve explicitly. By Itô's formula, we have that

$$d(e^{\lambda t}V_t) = e^{\lambda t}dV_t + \lambda e^{\lambda t}V_t dt = e^{\lambda t}dB_t$$

and so

$$e^{\lambda t}V_t = V_0 + \int_0^t e^{\lambda s} dB_s.$$

Rearranging, we have that

$$V_t = e^{-\lambda t}V_0 + \int_0^t e^{\lambda(s-t)} dB_s.$$

Note that V_t has the normal distribution with mean $V_0 e^{-\lambda t}$ and variance $(1-e^{-2\lambda t})/2\lambda$. If $\lambda > 0$, then $V_t \xrightarrow{d} N(0,(2\lambda)^{-1})$ as $t \to \infty$. In particular, $N(0,(2\lambda)^{-1})$ is the stationary distribution for V in the sense that if $V_0 \sim N(0,(2\lambda)^{-1})$ then $V_t \sim N(0,(2\lambda)^{-1})$ for all $t \geq 0$.

6.4. **Local solutions.** A *locally defined process* $(X, \mathcal{T})$ is a stopping time $\mathcal{T}$ together with a map $X\colon \{(\omega,t) \in \Omega \times [0,\infty) : t < \mathcal{T}(\omega)\} \to \mathbb{R}$. It is cádlág if $t \mapsto X_t(\omega) : [0, \mathcal{T}(\omega)) \to \mathbb{R}$ is cádlág for all $\omega \in \Omega$. Let $\Omega_t = \{\omega \in \Omega : t < \mathcal{T}(\omega)\}$. Then $(X,\mathcal{T})$ is *adapted* if $X_t\colon \Omega_t \to \mathbb{R}$ is $\mathcal{F}_t$-measurable for all t. We say that $(X,\mathcal{T})$ is a *locally defined local martingale* if there exist stopping times $T_n \uparrow \mathcal{T}$ such that X^{T_n} is a martingale for all n. We say that (H,η) is a *locally defined locally bounded previsible process* if there exist stopping times $S_n \uparrow \eta$ such that $H\mathbf{1}_{(0,S_n]}$ is bounded and previsible for all n. Then we can define $(H \cdot X, \mathcal{T} \wedge \eta)$ by

$$(X \cdot X)_t^{S_n \wedge T_n} = \big((H\mathbf{1}_{(0,S_n\wedge T_n]}) \cdot X^{X_n \wedge T_n}\big)_t, \quad \text{for all} \quad n.$$

Proposition 6.7 (Local Itô formula)**.** *Let $X^1, \ldots, X^d$ be continuous semimartingales and let $X = (X^1,\ldots,X^d)$. Let $U \subseteq \mathbb{R}^d$ be open and let $f\colon U \to \mathbb{R}^d$ be C^2. Set $\mathcal{T} = \inf\{t \geq 0 : X_t \notin U\}$. Then for all $t < \mathcal{T}$, we have that*

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

Proof. This follows by applying Itô's formula to X^{T_n} where $T_n = \inf\{t \geq 0 : \mathrm{dist}(X_t, U^c) \leq 1/n\}$. Then it is clearly the case that $T_n \uparrow \mathcal{T}$ as $n \to \infty$. ∎

Example 6.8. *Let $X = B$ where B is a standard Brownian motion starting from 1 with $d = 1$, $U = (0,\infty)$, and $f(x) = \sqrt{x}$. Then Proposition 6.7 implies that*

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

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

Suppose that $U \subseteq \mathbb{R}^d$ is open and $\sigma\colon U \to M^{d\times m}(\mathbb{R})$ and $b\colon U \to \mathbb{R}^d$ are measurable functions which are bounded on compact sets. A *local solution* to the SDE consists of:

- A filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying the usual conditions.
- An $(\mathcal{F}_t)$-Brownian motion B
- A continuous $(\mathcal{F}_t)$-adapted locally defined process $(X,\mathcal{T})$ such that $X_t = X_0 + \int_0^t \sigma(X_s) dB_s + \int_0^t b(X_s) ds$ for all $t < \mathcal{T}$.

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

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

Theorem 6.9. *Suppose that $U \subseteq \mathbb{R}^d$ is open and $\sigma\colon U \to M^{d\times m}(\mathbb{R})$ and $b\colon U \to \mathbb{R}^d$ are locally Lipschitz. Then the SDE $dX_t = \sigma(X_t)dB_t + b(X_t)dt$ has a pathwise unique maximal local solution $(X, \mathcal{T})$ started from $x \in U$. Moreover, for all compact sets $C \subseteq U$, on the event $\{\mathcal{T} < \infty\}$ we have that* $\sup\{t < \mathcal{T} : X_t \in C\} < \mathcal{T}$.

Lemma 6.10. *Let $U \subseteq \mathbb{R}^d$ be open, $C \subseteq U$ compact. Then:*

(1) there exists a C^∞ function $\varphi\colon \mathbb{R}^d \to [0,1]$ such that $\varphi|_C \equiv 1$ and $\varphi|_{U^c} \equiv 0$.
(2) given a locally Lipschitz function $f\colon U \to \mathbb{R}$, there exists a Lipschitz function $g\colon \mathbb{R}^d \to \mathbb{R}$ such that $f|_C = g|_C$.

Proof. We leave the first part as an exercise. For the second part, we take φ as in the second part and then set $g = f\varphi$. □

Proof of Theorem 6.9. We will assume for simplicity that $d = m = 1$. Fix $C \subseteq U$ compact. By Lemma 6.10 we can find Lipschitz functions $\widetilde{\sigma}$, $\widetilde{b}$ on $\mathbb{R}$ such that $\widetilde{\sigma}|_C = \sigma|_C$ and $\widetilde{b}|_C = b|_C$. By Theorem 6.3, there exists a pathwise unique strong solution $\widetilde{X}$ to $d\widetilde{X}_t = \widetilde{\sigma}(\widetilde{X}_t)dB_t + \widetilde{b}(\widetilde{X}_t)dt$, $\widetilde{X}_0 = x$.

Set $T = \inf\{t \geq 0 : \widetilde{X}_t \notin C\}$ and denote by X the restriction of $\widetilde{X}$ to $[0,T)$. Then (X,T) is a local solution in C and X_{T-} exists in $U \setminus C^\circ$ if $T < \infty$. Suppose that (X,T) and (Y,S) are both local solutions in C. Consider $f(t) = \mathbb{E}[\sup_{s \leq S\wedge T \wedge t} |X_s - Y_s|^2]$. Then $f(t) < \infty$. Let K_C be the Lipschitz constant of σ and b in C and fix $t_0 < \infty$. Then for $t \leq t_0$, as in the proof of Theorem 6.3, we have that

$$f(t) \leq 2K_C^2(4+t_0)\int_0^t f(s)ds, \quad t \leq t_0.$$

Therefore by Gronwall's lemma we have that $f \equiv 0$ and it thus follows that $X_t = Y_t$ for all $t \leq S \wedge T$.

Take compact sets C_n which increase to U and construct for each n a local solution (X_n, T^n) in C_n by the above procedure. Then if $T_n < \infty$, we have that $X^n_{T_n-} \in U \setminus C_n^\circ$. Then $T_n \uparrow \mathcal{T}$ for some $\mathcal{T} > 0$ and $X^n_t = X^{n+1}_t$ for all $t < T_n$. Thus by taking $X = X^n_t$ for $t < T_n$, we have defined a local solution $(X, \mathcal{T})$.

We will now show that $(X, \mathcal{T})$ is a unique and maximal. Suppose that (Y, η) is another local solution on the same space and, for each n, set $S_n = \inf\{t \geq 0 : Y_t \notin C_n\}$. Then by the uniqueness of the solution on each C_n, we have that $X_t = Y_t$ for all $t < T_n \wedge S_n$. Therefore $S_n \leq T_n$. As $n \to \infty$, we have that $S_n \uparrow \eta$ and $T_n \uparrow \mathcal{T}$, hence $\eta \leq \mathcal{T}$ and $X_t = Y_t$ for all $t \leq \eta$.

We now prove the final assertion of the theorem. Suppose that C_1, C_2 are compact sets in U with $C_1 \subseteq C_2^\circ \subseteq C_2 \subseteq U$ and also a C^∞ function $\varphi\colon U \to \mathbb{R}$ with $\varphi|_{C_1} \equiv 1$ and $\varphi|_{(C_2^\circ)^c} \equiv 0$. We then set

$$\begin{aligned} R_1 &= \inf\{t \geq 0 : X_t \notin C_2\} \\ S_n &= \inf\{t \geq R_n : X_t \in C_1\} \wedge \mathcal{T} \\ R_n &= \inf\{t \geq S_{n-1} : X_t \notin C_2\} \wedge \mathcal{T}. \end{aligned}$$

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

$$\begin{aligned} n &= \sum_{k=1}^{n} \big(\varphi(X_{S_k}) - \varphi(X_{R_k})\big) \\ &= \int_0^t \sum_{k=1}^{n} \mathbf{1}_{(R_k,S_k]}(s)\big(\varphi'(X_s)dX_s + \frac{1}{2}\varphi''(X_s)d[X]_s\big) \\ &= \int_0^t \big(H_s^n dB_s + K_s^n ds\big) := Z_t^n \end{aligned}$$

where H^n and K^n are previsible and bounded uniformly in n. This implies that

$$n^2 \mathbf{1}_{\{\mathcal{T} \le t, N \ge n\}} \le (Z_t^n)^2$$

hence

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

Note that $\mathbb{E}[(Z_t^n)^2] \le C < \infty$ where C is independent of n as H^n and K^n are bounded uniformly in n and Z_t^n is defined by integrating H^n, K^n over a time-interval which does not depend on n. Thus by letting $n \to \infty$, we see that $\mathbb{P}[\mathcal{T} \le t, N = \infty] = 0$. Therefore, on $\{\mathcal{T} < \infty\}$, X eventually stops coming back to C_1, so $\sup\{t < \mathcal{T} : X_t \in C_1\} < \mathcal{T}$. $\square$

Example 6.11 (Bessel processes)**.** *Fix $\nu \in \mathbb{R}$. Consider the SDE in $U = (0,\infty)$ given by*

$$dX_t = dB_t + \left(\frac{\nu - 1}{2X_t}\right) dt, \quad X_0 = x_0 \in U.$$

Theorem 6.9 implies that there exists a unique maximal local solution $(X, \mathcal{T})$ in U with $\liminf_{t \uparrow \mathcal{T}} X_t = 0$ if $\mathcal{T} < \infty$. We call $(X,\mathcal{T})$ a Bessel process of dimension ν.

Suppose that $\nu \in \mathbb{N}$ and let β be a Brownian motion in $\mathbb{R}^\nu$ with $|\beta_0| = x_0$. Set $|Y_t| = |\beta_t|$ and $\eta = \inf\{t \ge 0 : \beta_t = 0\}$. By the local Itô formula, we have that

$$dY_t = \frac{(\beta_t, d\beta_t)}{|\beta_t|} + \left(\frac{\nu - 1}{2|\beta_t|}\right) dt, \quad t < \eta.$$

Consider $W_t = \int_0^t \frac{(\beta_s, d\beta_s)}{|\beta|_s} ds, \quad t \ge 0$. Then $W \in \mathcal{M}_{c,loc}$ and

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

Therefore, by Lévy's characterization of Brownian motion, we have that W is a standard Brownian motion in $\mathbb{R}$. Therefore

$$dY_t = dW_t + \left(\frac{\nu - 1}{2Y_t}\right) dt, \quad t < \eta.$$

7. Diffusion processes

Suppose that we are given bounded, measurable functions $a \colon \mathbb{R}^d \to M^{d\times d}(\mathbb{R})$ and $b \colon \mathbb{R}^d \to \mathbb{R}^d$ with a symmetric (i.e., each for each $x \in \mathbb{R}^d$ the matrix $a(x)$ satisfies $a(x) = (a(x))^T$). For $f \in C_b^2(\mathbb{R}^d)$ (i.e., the space of $C^2(\mathbb{R}^d)$ functions with bounded derivatives), we set

$$Lf(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{i=1}^{d} b_i(x) \frac{\partial f}{\partial x_i}. \tag{7.1}$$

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

$$M_t^f := f(X_t) - f(X_0) - \int_0^t Lf(x_s)ds$$

is a martingale. We call a the *diffusivity* and b the *drift.* If a and b are clear from the context, we will sometimes refer to an L-diffusion as an (a, b)-diffusion.

Example 7.1. *Suppose that σ and b are constant and $a = \sigma\sigma^T$. Let B be a standard Brownian motion in $\mathbb{R}^d$. Then the process $X_t = \sigma B_t + bt$ is an (a, b)-diffusion.*

One special case of this is when $\sigma = I$ and $b = 0$. Then $X_t = B_t$ is an L-diffusion with $L = \frac{1}{2}\Delta$.

Proposition 7.2. *Suppose that X is a solution to the SDE $dX_t = \sigma(X_t)dB_t + b(X_t)dt$. Let $f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ (i.e., the space of functions which are C^1 in the first variable and C^2 in the remaining d-variables, with bounded derivatives). Then the process*

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

is a continuous local martingale where $a = \sigma\sigma^T$ and L is as in (7.1)*. In particular, if σ, b are bounded, then X is an L-diffusion.*

Proof. See Example Sheet 3. □

Proposition 7.2 gives us a way to construct L-diffusions using solutions of SDEs. In particular, we suppose that a, b are Lipschitz and bounded and that there exists $\epsilon > 0$ such that

$$(\xi, a(x)\xi) \geq \epsilon^2 |\xi|^2 \quad \text{for all} \quad x. \tag{7.2}$$

(When (7.2) holds, we say that a is *uniformly positive definite* or UPD.) Then there exists a Lipschitz map $\sigma\colon \mathbb{R}^d \to \mathbb{R}^d$ with $\sigma\sigma^T = a$. Indeed, if $d = 1$ then we can take $\sigma = \sqrt{a}$. For $d \geq 2$, we can write $a(x) = u(x)\lambda(x)(u(x))^T$ where $\lambda(x)$ is the diagonal matrix of eigenvalues of $a(x)$ and $u(x)$ consists of the corresponding eigenvectors. So, in this case, we can take $\sigma(x) = u(x)\sqrt{\lambda(x)}(u(x))^T$. For such u and σ, Theorem 6.3 implies that the SDE $dX_T = \sigma(X_t)dB_t + b(X_t)dt$ has a solution.

Proposition 7.3. *Let X be an L-diffusion and T a finite stopping time. Set $\widetilde{X}_t = X_{T+t}$ and $\widetilde{\mathcal{F}}_t = \mathcal{F}_{T+t}$. Then $\widetilde{X}$ is an L-diffusion with respect to $(\widetilde{\mathcal{F}}_t)_{t \geq 0}$.*

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

$$\widetilde{M}_t^f = f(\widetilde{X}_t) - f(\widetilde{X}_0) - \int_0^t Lf(\widetilde{X}_s)ds.$$

Then $\widetilde{M}_t^f$ is a $\widetilde{\mathcal{F}}_t$-adapted and integrable. For $A \in \widetilde{\mathcal{F}}_s$ and $n \geq 0$, we note that

$$\mathbb{E}[(\widetilde{M}_t^f - \widetilde{M}_s^f)\mathbf{1}_{A \cap \{T \leq n\}}] = \mathbb{E}[(M_{T+t}^f - M_{T+s}^f)\mathbf{1}_{A \cap \{T \leq n\}}] = 0$$

by OST. Letting $n \to \infty$, we see that $\widetilde{M}^f$ is a martingale. □

Lemma 7.4. *Let X be an L-diffusion. Then for all $f \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$, the process*

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

is a martingale.

Proof. We 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 and $Z_n \to 0$ as $n \to \infty$. Consequently, $\mathbb{E}[Z_n] \to 0$ as $n \to \infty$. We note that:

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

Choose $s_0 \leq s_1 \leq \cdots \leq s_m$ with $s_0 = s$ and $s_m = t$ and $s_{k+1} - s_k \leq 1/n$. As the first line of our formula for $M_t^f - M_s^f$ is zero and the second line has zero conditional expectation given $\mathcal{F}_s$ (as X is an L-diffusion), it thus follows that

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

Combining this with the triangle inequality, we thus see that

$$\mathbb{E}\left[\left| \mathbb{E}[M_t^f - M_s^f \,|\, \mathcal{F}_s] \right| \right] \leq (t - s)\mathbb{E}[Z_n].$$

As the right hand side goes to zero as $n \to \infty$, it follows that $\mathbb{E}[M_t^f \,|\, \mathcal{F}_s] = M_s^f$, as desired. $\square$

Assume that a, b are Lipschitz and that a is UPD: $(\xi, a(x)\xi) \geq \epsilon |\xi|^2$ for all x, ξ. Let D be an bounded, open subset of $\mathbb{R}^d$ with smooth boundary $\partial \mathbb{D}$. For each $\epsilon > 0$, we let $A^\epsilon = \{x \in \mathbb{R}^d : |x - A| < \epsilon\}$. Then we have that $\mathrm{Leb}(A^\epsilon) = 2\epsilon\lambda(A) + o(\epsilon)$ as $\epsilon \to 0$ where λ is the surface area measure on ∂D. We shall assume the following result from PDE.

Theorem 7.5 (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*

$$Lu + \varphi = 0 \quad \text{on} \quad D \quad \text{and} \quad u = f \quad \text{on} \quad \partial D.$$

Moreover, there exist continuous functions $m \colon D \times \partial D \to (0, \infty)$ and $g \colon \{(x, y) \in D \times D : x \neq y\} \to (0, \infty)$ such that for all such f and φ we have that

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

We call g the *Green kernel* and $m(x, y)\lambda(dy)$ the *harmonic measure* on ∂D starting from x.

Theorem 7.6. *Suppose that $u \in C(\overline{D}) \cap C^2(D)$ satisfies*

$$Lu + \varphi = 0 \quad \text{on} \quad D \quad \text{and} \quad u = f \quad \text{on} \quad \partial D$$

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

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

where $T = \inf\{t \geq 0 : X_t \notin D\}$. In particular, for all Borel sets $A \subseteq D$ and $B \subseteq \partial D$, we have that

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

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

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

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

$$M_t^{\widetilde{u}} = \widetilde{u}(X_t) - \widetilde{u}(X_0) - \int_0^t L\widetilde{u}(X_s)ds$$

so M is a martingale by OST. Hence,

$$u(x) = \mathbb{E}_x\left[u(X_{t\wedge T_n}) + \int_0^{t\wedge T_n} \varphi(X_s)ds\right]. \tag{7.3}$$

Consider the case $\varphi = 1$ and $f = 0$. Then

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

Since $u^{1,0}$ is bounded and $T_n \uparrow T$, we deduce by the monotone convergence theorem that $\mathbb{E}_x[T] < \infty$. Returning to the general case, we can let $t \to \infty$ and $n \to \infty$ in (7.3). Since u is continuous on $\overline{D}$, we have that $u(X_{t\wedge T_n}) \to f(X_T)$. Also, $t \wedge T_n \uparrow T$ and

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

We obtain by dominated convergence that

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

Consequently,

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

Finally, we have for all $\varphi \in C(\overline{D})$, $f \in C(\partial D)$ that

$$\mathbb{E}_x\left[\int_0^T \varphi(X_s)ds + f(X_T)\right] = \int_{\partial D} m(x,y)f(y)\lambda(dy) + \int_D g(x,y)\varphi(y)dy.$$

We let $\varphi_n \to \mathbf{1}_A$ and $f_n \to \mathbf{1}_B$ to obtain the final result. □

Theorem 7.7 (Cauchy problem)**.** *For each $f \in C_b^2(\mathbb{R}^d)$, there exists a unique $u \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ such that*

$$\frac{\partial u}{\partial t} = Lu \quad \textit{on} \quad \mathbb{R}_+ \times \mathbb{R}^d \quad \textit{and} \quad u(0,\ldots) = f \quad \textit{on} \quad \mathbb{R}^d.$$

Moreover, there exists a continuous function $p\colon (0,\infty) \times \mathbb{R}^d \times \mathbb{R}^d \to (0,\infty)$ such that for all $f \in C_b^2(\mathbb{R}^d)$, u is given by

$$u(t,x) = \int_{\mathbb{R}^d} p(t,x,y)f(y)dy.$$

In the setting of Theorem 7.7, the function p is called the *heat kernel*.

Theorem 7.8. *Assume that $f \in C_b^2(\mathbb{R}^d)$. Let $u \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ satisfy*

$$\frac{\partial u}{\partial t} = Lu \quad on \quad \mathbb{R}_+ \times \mathbb{R}^d \quad and \quad u(0,\ldots) = f \quad on \quad \mathbb{R}^d.$$

Then, for any L-diffusion X, for all $t \in \mathbb{R}_+$, $x \in \mathbb{R}^d$, $s \leq t$, we have that $\mathbb{E}[f(X_t) \,|\, \mathcal{F}_s] = u(t-s, X_s)$. Hence, $u(t,x) = \mathbb{E}_x[f(X_t)]$ and under $\mathbb{P}_x$ the finite dimensional distributions of X are given by

$$\mathbb{P}_x[X_{t_1} \in dx_1, \ldots, X_{t_n} \in dx_n] = p(t_1, x_0, x_1) \cdots p(t_n - t_{n-1}, X_{t_{n-1}}, X_{t_n}) dx_1 \cdots dx_n$$

for $0 < t_1 < \cdots < t_n$ and $x_1, \ldots, x_n \in \mathbb{R}^d$.

Proof. Fix $t \in (0,\infty)$ and consider the function $g(s,x) = u(t-s,x)$. Note that

$$\left(\frac{\partial f}{\partial s} + L\right) g(s,x) = -\dot{u}(t-s,x) + Lu(t-s,x) = 0.$$

Consequently, $M_s^g = g(s,X_s) - g(0,x_0)$ is a martingale. This implies that $\mathbb{E}[f(X_t) \,|\, \mathcal{F}_s] = u(t-s,X_s)$. Taking $s = 0$ and $X_0 = x$ implies that $\mathbb{E}[f(X_t)] = u(t,x)$. This proves the first part of the theorem.

For the second part of the theorem, we set

$$P_t f(x) = \int_{\mathbb{R}^d} p(t,x,y) f(y) dy = u(t,x).$$

By the uniqueness of solutions to the Cauchy problem, we have that $P_s(P_t f) = P_{s+t} f$. We claim by induction on n that

$$\mathbb{E}_{x_0}\left[\prod_{i=1}^n f_i(X_{t_i})\right] = \int_{(\mathbb{R}^d)^n} p(t_1, x_0, x_1) f_1(x_1) \cdots p(t_n - t_{n-1}, x_{n-1}, x_n) f_n(x_n) dx_1 \cdots dx_n.$$

For the induction step, we use that

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

and then apply the $n-1$ case. □

Theorem 7.9 (Feynman-Kac)**.** *Let $f \in C_b^2(\mathbb{R}^d)$ and $V \in C_b(\mathbb{R}^d)$. Suppose that $u \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ satisfies the PDE*

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u + Vu \quad on \quad \mathbb{R}_+ \times \mathbb{R}^d \quad and \quad u(0,\cdot) = f \quad on \quad \mathbb{R}^d.$$

Then for all $xt \in \mathbb{R}_+$, $x \in \mathbb{R}^d$ we have that

$$u(t,x) = \mathbb{E}_x\left[f(B_t) \exp\left(\int_0^t V(B_s) ds\right)\right]$$

where B is a standard Brownian motion.

Proof. Let

$$E_t = \exp\left(\int_0^t V(B_s) ds\right).$$

Fix $T \in (0,\infty)$ and set $M_t = u(T-t, B_t)E_t$. By Itô's formula, we have that

$$dM_t = \left(\sum_{i=1}^d \frac{\partial u}{\partial x_i}(T-t,B_t)dB_t^i + \frac{1}{2}\Delta u(T-t,B_t)dt - \dot{u}(T-t,B_t)dt\right)E_t + E_t u(T-t,B_t)V(B_t)dt.$$

So, M is a local martingale which is uniformly bounded on $[0,T]$ as $u \in C_b^{1,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ and V is bounded on $[0,T]$. Therefore M is a martingale. Hence,

$$u(T,x) = M_0 = \mathbb{E}_x[M_T] = \mathbb{E}_x[f(B_T)E_T],$$

as desired. □

8. Complementary material

8.1. Existence of solutions to SDEs with continuous coefficients.

Theorem 8.1. *Consider the SDE $dX_t = b(X_t)dt + \sigma(X_t)dB_t$ where $b\colon \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma\colon \mathbb{R}^d \to M^{d\times d}(\mathbb{R})$ are bounded and continuous functions. Then for each $x \in \mathbb{R}^d$, there exists a weak solution started from x.*

Suppose that $b\colon \mathbb{R}^d \to \mathbb{R}^d$, $\sigma\colon \mathbb{R}^d \to M^{d\times d}(\mathbb{R})$ are bounded and measurable. Let $a = \sigma\sigma^T$ and

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

A solution to the so-called *local martingale problem* associated with L is a continuous, adapted process X such that

$$M_t^f = f(X_t) - f(X_0) - \int_0^t Lf(X_s)ds$$

is a continuous local martingale for every $f \in C_b^2(\mathbb{R}^d)$.

Theorem 8.2. *The existence of a solution to the local martingale problem associated with L is equivalent to the existence of a weak solution to the SDE $dX_t = \sigma(X_t)dB_t + b(X_t)dt$.*

Proof. See Example Sheet 3, exercise 9. □

Proof of Theorem 8.1. For each $j \geq 0$ and $n \in \mathbb{N}$, we let $t_j^n = 2^{-n}j$ and let $\psi_n(t) = t_j^n$ for $t \in [t_j^n, t_{j+1}^n)$. We recursively define

$$X_t^n = X_{t_j^n}^n + b(X_{t_j^n}^n)(t - t_j^n) + \sigma(X_{t_j^n}^n)(B_t - B_{t_j^n}) \quad \text{for} \quad j \geq 0, \quad t_j^n \leq t < t_{j+1}^n$$

with $X_0 = x$. Then X_t^n solves the integral equation

$$X_t^n = x + \int_0^t b(X_{\psi_n(s)}^n)ds + \int_0^t \sigma(X_{\psi_n(s)}^n)dB_s, \quad 0 \leq t < \infty.$$

By Example Sheet 4, exercise 9, we have that for each $m \in \mathbb{N}$ and $T > 0$ there exists a constant $C > 0$ such that

$$\sup_{n\geq 1}\mathbb{E}\|X_t^n - X_s^n\|^{2m} \leq C(t-s)^m, \quad \text{for all} \quad 0 \leq s < t \leq T.$$

This implies that, for each $T > 0$, the family of laws $(X^n|_{[0,T]})$ is tight with respect to the uniform topology (i.e., the topology of $C([0,T],\mathbb{R}^d)$). Therefore we can find a subsequence (X^{n_k}) such that for each $T > 0$, the law of $X^{n_k}|_{[0,T]}$ converges to the law of a limiting process X restricted to $[0,T]$. By the Skorohod representation theorem for weak convergence, we may assume that (X^{n_k}) and X are defined on a common probability space so that for each $T > 0$ we have that $X^{n_k}|_{[0,T]}$ converges uniformly to $X|_{[0,T]}$.

To show that X satisfies $dX_t = b(X_t)dt + \sigma(X_t)dB_t$, it suffices to show that $f(X_t) - f(X_0) - \int_0^t Lf(X_s)ds$ is a continuous martingale for each $f \in C_b^2(\mathbb{R}^d)$ which is bounded. To show this, it suffices to show that for all $0 \leq s < t$ and bounded, continuous functions $g\colon C([0,\infty),\mathbb{R}^d) \to \mathbb{R}$ such that $g(y)$ only depends on $y|_{[0,s]}$, we have that

$$\mathbb{E}\left[\left(f(X_t) - f(X_s) - \int_s^t Lf(X_u)du\right) g(X)\right] = 0. \tag{8.1}$$

By the construction of X^{n_k}, we know that

$$\mathbb{E}\left[\left(f(X_t^{n_k}) - f(X_s^{n_k}) - \int_s^t L_u^n f(X_u^{n_k})du\right) g(X)\right] = 0 \tag{8.2}$$

for all k where

$$L_u^{n_k} f(x) = \frac{1}{2}\sum_{i,j=1}^d a_{ij}(X^{n_k}_{\psi_{n_k}(u)})\frac{\partial^2 f}{\partial x_i \partial x_j}(x) + \sum_{i=1}^d b_i(X^{n_k}_{\psi_{n_k}(u)})\frac{\partial f}{\partial x_i}(x).$$

To finish the proof, it therefore suffices to show that

$$F_{n_k} = f(X_t^{n_k}) - f(X_s^{n_k}) - \int_s^t L_u^{n_k}(X_u^{n_k})du \to F = f(X_t) - f(X_s) - \int_s^t Lf(X_u)du \quad \text{as} \quad k \to \infty.$$

Indeed, then we have that (8.1) by (8.2) and the bounded convergence theorem. Since we know that $X^{n_k} \to X$ as $k \to \infty$ uniformly on each compact subset of $[0,\infty)$, we in particular have that $f(X_t^{n_k}) \to f(X_t)$ and $f(X_s^{n_k}) \to f(X_s)$ as $k \to \infty$.

We can write,

$$\begin{aligned}
&\int_s^t L_u^{n_k} f(X_u^{n_k})du - \int_s^t Lf(X_u)du \\
=& \int_s^t (L_u^{n_k} - L)f(X_u^{n_k})du + \int_s^t Lf(X_u^{n_k}) - Lf(X_u)du.
\end{aligned}$$

Note that the second term on the right hand side tends to 0 as $k \to \infty$ by the convergence of X^{n_k} to X. As for the first term, note that we can write

$$\begin{aligned}
&(L_u^{n_k} - L)f(X_u^{n_k}) \\
=&\frac{1}{2}\sum_{i,j=1}^d \big(a_{ij}(X^{n_k}_{\psi_{n_k}(u)}) - a_{ij}(X_u)\big)\frac{\partial^2 f}{\partial x_i \partial x_j}(X_u^{n_k}) + \sum_{i=1}^d \big(b_i(X^{n_k}_{\psi_{n_k}(u)}) - b_i(X_u)\big)\frac{\partial f}{\partial x_i}(X_u^{n_k}).
\end{aligned}$$

Hence, $(L_u^{n_k} - L)f(X_u^{n_k})$ is bounded and tends to 0 on $[s,t]$ as $k \to \infty$ by the uniform convergence of X^{n_k} to X on $[s,t]$. $\square$

8.2. **Yamada-Watanabe criterion for pathwise uniqueness and its consequences.**

Proposition 8.3 (Yamada and Watanabe)**.** *Consider the SDE $dX_t = b(X_t)dt + \sigma(X_t)dB_t$. Assume that the coefficients b, σ satisfy the conditions*

$$|b(x) - b(y)| \le K|x-y| \quad \textit{and} \quad |\sigma(x) - \sigma(y)| \le h(|x-y|)$$

where $K > 0$ is a constant and $h\colon [0,\infty) \to [0,\infty)$ is strictly increasing with $h(0) = 0$. Assume further that

$$\int_0^\epsilon \frac{1}{h^2(u)} du = \infty \quad \textit{for all} \quad \epsilon > 0.$$

Then pathwise uniqueness holds for the SDE.

Proof. Due to the conditions imposed on h, there exists a strictly decreasing sequence (a_n) in $(0,1]$ with $a_0 = 1$, $\lim_{n\to\infty} a_n = 0$, and $\int_{a_n}^{a_{n-1}} h^{-2}(u)du = n$ for all n. For each n, we can find a continuous function ρ_n supported in (a_n, a_{n-1}) so that $0 \le \rho_n(x) \le 2/(nh^2(x))$ for all x and $\int_{a_n}^{a_{n-1}} \rho_n(x)dx = 1$. Let

$$\psi_n(x) = \int_0^{|x|} \int_0^y \rho_n(u)dudy, \quad x \in \mathbb{R}.$$

Then ψ_n is even, C^2, $|\psi'| \le 1$, and $\lim_{n\to\infty} \psi_n(x) = |x|$ for $x \in \mathbb{R}$.

Suppose that X^1, X^2 are two solutions to the SDE with $X_0^1 = X_0^2$. By localization, we may assume that

$$\mathbb{E}\int_0^t |\sigma(X_s^i)|^2 ds < \infty \quad \text{for} \quad i = 1,2 \quad \text{and all} \quad t \ge 0.$$

Let

$$\Delta_t = X_t^1 - X_t^2 = \int_0^t b(X_s^1) - b(X_s^2)ds + \int_0^t \sigma(X_s^1) - \sigma(X_s^2)dW_s.$$

By Ito's formula, we have that

$$\begin{aligned}\psi_n(\Delta_t) =& \int_0^t \psi_n'(\Delta_s)\big(b(X_s^1) - b(X_s^2)\big)ds + \\ & \frac{1}{2}\int_0^t \psi_n'(\Delta_s)\big(\sigma(X_s^1) - \sigma(X_s^2)\big)^2 ds + \\ & \int_0^t \psi_n'(\Delta_s)\big(\sigma(X_s^1) - \sigma(X_s^2)\big)dB_s \\ =& I_t^1 + I_t^2 + I_t^3.\end{aligned}$$

By the martingale property, we have that $\mathbb{E}[I_t^3] = 0$. We also have that

$$\begin{aligned}\mathbb{E}[I_t^2] =& \frac{1}{2}\mathbb{E}\int_0^t \psi_n''(\Delta_s)\big(\sigma(X_s^1) - \sigma(X_s^2)\big)^2 ds \\ \le& \frac{1}{2}\mathbb{E}\int_0^t \psi_n''(\Delta_s)h^2(|\Delta_s|)ds \\ \le& \frac{t}{n}.\end{aligned}$$

Consequently, we have that

$$\begin{aligned}\mathbb{E}[\psi_n(\Delta_t)] \leq& \mathbb{E}\int_0^t \psi_n'(\Delta_s)\big(b(X_s^1) - b(X_s^2)\big)ds + \frac{t}{n}\\ \leq& K\int_0^t |\Delta_s|ds + \frac{t}{n}.\end{aligned}$$

Sending $n \to \infty$, we thus have that

$$\mathbb{E}|\Delta_t| \leq K\int_0^t \mathbb{E}|\Delta_s|ds \quad \text{for all} \quad t \geq 0.$$

The Gronwall inequality therefore implies that $\Delta_t = 0$ for all t, which completes the proof. □

Example 8.4. *Fix $\delta \in \mathbb{R}$ and consider the SDE*

$$dZ_t = \sqrt{|Z_t|}dB_t + \delta dt.$$

This is the so-called square Bessel equation *of dimension δ. Its importance is that if Z solves this SDE, then $\sqrt{Z}$ solves the Bessel SDE of dimension δ. This is how solutions to the Bessel SDE are constructed, even after they hit* 0*. Theorem 8.1 implies that this SDE has a weak solution started at x for each $x \in \mathbb{R}$ and Proposition 8.3 implies that the SDE has pathwise uniqueness. We are now going to show that weak existence and pathwise uniqueness imply that there is a unique strong solution.*

Proposition 8.5. *Pathwise uniqueness implies that we have uniqueness in law.*

Proof. See Example Sheet 3, problem 5. □

Theorem 8.6. *Consider the SDE $dX_t = \sigma(X_t)dB_t + b(X_t)dt$. If it has a weak solution and pathwise uniqueness, there is a unique strong solution.*

Proof. Suppose that (X^1, B^1) and (X^2, B^2) are two solutions defined on possibly different probability spaces. Our first goal is to put them on the same space so that they are driven by the same Brownian motion instance. In order to do so, we:

- Sample a standard Brownian motion B.
- Sample X^1 from its conditional law given $B^1 = B$.
- Sample X^2 from its conditional law given $B^2 = B$.

In this procedure, we take X^1 and X^2 to be conditionally independent given B. In other words, for all events A^1 and A^2, we have that

$$\mathbb{P}[X^1 \in A^1, X^2 \in A^2 \,|\, \sigma(B)] = \mathbb{P}[X^1 \in A^1 \,|\, \sigma(B)]\mathbb{P}[X^2 \in A^2 \,|\, \sigma(B)].$$

We leave as an exercise the following useful fact. Suppose that X, Y are random variables which are conditionally independent given a σ-algebra $\mathcal{G}$ such that $X = Y$ almost surely. Then $X = Y$ is $\mathcal{G}$-measurable.

Pathwise uniqueness implies that $X^1 = X^2$ almost surely. Therefore the exercise implies that $X^1 = X^2$ is $\sigma(B)$-measurable. The same argument applied to $X^1|_{[0,t]}$, $X^2|_{[0,t]}$, and $B|_{[0,t]}$ implies that $X^1|_{[0,t]} = X^2|_{[0,t]}$ is $\sigma(B|_{[0,t]})$-measurable. This says that $X^1 = X^2$ is adapted to the filtration generated by B, hence we have a strong solution. □

8.3. **The Kazamaki and Novikov integrability conditions for exponential martingales.** Recall that if $M \in \mathcal{M}_{c,loc}$, we defined $\mathcal{E}(M)_t = \exp(M_t - [M]_t/2)$. This is the exponential local martingale associated with M. By construction, $\mathcal{E}(M)_t$ is always a continuous local martingale. Moreover, it is not difficult to see using Fatou's lemma that $\mathcal{E}(M)_t$ is a supermartingale. In order to apply the Girsanov theorem, it is important that $\mathcal{E}(M)_t$ is a genuine martingale. To show that this is the case, we just need to show that $\mathbb{E}[\mathcal{E}(M)_t] = 1$ for all $t \geq 0$. Previously, we showed that a sufficient condition is that $[M]_\infty \leq C$ for some constant C. We will now relax this condition.

Theorem 8.7 (Kazamaki condition)**.** *Suppose that $M \in \mathcal{M}_{c,loc}$. If*

$$\mathbb{E}\exp([M]_t/2) < \infty \quad \textit{for all} \quad t \geq 0$$

then

$$\mathbb{E}[Z_t] = 1 \quad \textit{for all} \quad t \geq 0 \quad \textit{where} \quad Z_t = \mathcal{E}(M_t).$$

In particular, Z_t is a martingale.

Proof. For each $s \geq 0$, we let $\tau_s = \inf\{t \geq 0 : [M]_t > s\}$. By the Dubins-Schwartz theorem, the process $B_s = M_{\tau_s}$ is a Brownian motion for the filtration $\mathcal{G}_s = \mathcal{F}_{\tau_s}$. For each $b < 0$, we let $S_b = \inf\{t \geq 0 : B_s - s = b\}$. Then we have that

$$\begin{aligned}1 &= \mathbb{E}[\exp(B_{S_b} - S_b/2)] \\ &= \mathbb{E}[\exp(B_{S_b} - S_b + S_b - S_b/2)] \\ &= \mathbb{E}[\exp(b + S_b/2)].\end{aligned}$$

Rearranging, we thus see that

$$\mathbb{E}[\exp(S_b/2)] = \exp(-b).$$

Conisder the martingale $Y_s = \exp(B_s - s/2)$ and let $N_s = Y_{s \wedge S_b}$. As N is a martingale and $\mathbb{P}[S_b < \infty] = 1$, we have that

$$N_\infty = \lim_{t \to \infty} N_s = \exp(B_{S_b} - S_b/2).$$

As $\mathbb{E}[N_\infty] = 1$, it follows that for any stopping time R we have that

$$\mathbb{E}[\exp(B_{R \wedge S_b} - R \wedge S_b/2)] = \mathbb{E}[N_R] = 1.$$

As $[M]_t$ is a stopping time for $(\mathcal{G}_s)$, it follows that for any stopping time $b < 0$ and $t \geq 0$ we have that

$$1 = \mathbb{E}[\mathbf{1}_{S_b \leq [M]_t} \exp(b + S_b/2)] + \mathbb{E}[\mathbf{1}_{[M]_t < S_b} \exp(M_t - [M]_t/2)].$$

The first expectation is at most $e^b \mathbb{E}[\exp([M]_t/2)] \to 0$ as $b \to -\infty$. The second expectation converges to $\mathbb{E}[Z_t]$ by the monotone convergence theorem. Therefore $\mathbb{E}[Z_t] = 1$ for all $t \geq 0$. $\square$

Corollary 8.8 (Novikov)**.** *Let B be a d-dimensional Brownian motion and $X^1, \ldots, X^d$ locally bounded and previsible. If*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t \|X_s\|^2 ds\right)\right] < \infty \quad \textit{for all} \quad t \geq 0 \tag{8.3}$$

then $Z_t = \exp\left(\int_0^t X_s dB_s - \frac{1}{2}\int_0^t \|X_s\|^2 ds\right)$ is a martingale with $\mathbb{E}Z_t = 1$ for all $t \geq 0$.

Corollary 8.9. *The previous corollary holds if* (8.3) *is replaced by the following assumption. There exists a sequence of real numbers* (t_n) *with* $0 = t_0 < t_1 < \cdots < t_n \uparrow \infty$ *such that*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_{t_{n-1}}^{t_n} \|X_s\|^2 ds\right)\right] < \infty \quad \textit{for all} \quad n \in \mathbb{N}.$$

Proof. For each n, we let $X_t^n = (X_t^1 \mathbf{1}_{[t_{n-1},t_n)}(t), \ldots, X_t^d \mathbf{1}_{[t_{n-1},t_n)})$. Then

$$Z_t^n = \mathcal{E}\left(\int X_s^n dB_s\right)_t$$

is a maritngale by the previous corollary. Thus,

$$\mathbb{E}[Z_{t_n}^n \,|\, \mathcal{F}_{t_{n-1}}] = Z_{t_{n-1}}^n = 1 \quad \text{for all} \quad n \in \mathbb{N}.$$

Consequently,

$$\mathbb{E}[Z_{t_n}] = \mathbb{E}[Z_{t_{n-1}} \mathbb{E}[Z_{t_n}^n \,|\, \mathcal{F}_{t_{n-1}}]] = \mathbb{E}[Z_{t_{n-1}}] \quad \text{for all} \quad n \in \mathbb{N}.$$

By induction on n, we thus have that $\mathbb{E}[Z_{t_n}] = 1$ for all n. A similar argument implis that $\mathbb{E}[Z_t] = 1$ for all n, and therefore Z is a martingale. □

Statslab, Center of Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK

E-mail address: jpmiller@statslab.cam.ac.uk