Stochastic Calculus

Stefan Grosskinsky and James Norris

Cambridge, Lent 2007

These notes and other information about the course are available on
`www.statslab.cam.ac.uk/˜stefan/teaching/stochcalc.html`

The text is copied from notes by Christina Goldschmidt and James Norris, who deserve all the credit for the contents. Responsible for errors is Stefan Grosskinsky. Comments and corrections are most welcome and should be sent to `stefan@statslab.cam.ac.uk`.

Contents

Introduction **2**

1 Preliminaries **4**
- 1.1 Finite variation integrals 4
- 1.2 Previsible processes 6
- 1.3 Local martingales 8

2 The stochastic integral **11**
- 2.1 Simple integrands 11
- 2.2 L^2 properties 12
- 2.3 Quadratic variation 14
- 2.4 Itô integrals 17

3 Stochastic calculus **22**
- 3.1 Covariation 22
- 3.2 Itô's formula 23
- 3.3 Stratonovich integral 26
- 3.4 Notation and Summary 27

4 Applications of Stochastic Calculus **29**
- 4.1 Brownian motion 29
- 4.2 Exponential martingales 30

5 Stochastic Differential Equations **33**
- 5.1 General definitions 33
- 5.2 Lipschitz coefficients 34
- 5.3 Local solutions 38

Introduction

In this course we will develop the tools needed to handle continuous-time Markov processes in $\mathbb{R}^d$. Three basic examples of spatially homogeneous processes are:

(i) scaled Brownian motion: $(W_t)_{t\geq 0}$ where $W_t \sim N(0, at)$ and $a = \mathrm{var}\big(W(1)\big)$ is the diffusivity

(ii) constant drift: bt where $b \in \mathbb{R}^d$ determines the speed $|b|$ and the direction $b/|b|$

(iii) jump process: $(Y_t)_{t\geq 0}$ which at times of a Poisson process with rate λ makes jumps of distribution $\mu(dy)$. $(Y_t)_{t\geq 0}$ can be completely characterized by the *Lévy measure* $K(dy) = \lambda\mu(dy)$, via $\lambda = K(\mathbb{R}^d)$ and $\mu(dy) = K(dy)/\lambda$.

By the *Lévy-Khinchin Theorem* (see Advanced Probability, Chapter 8)

$$(X_t)_{t\geq 0} \quad \text{defined by} \quad X_t = W_t + bt + Y_t$$

is the most general process with piecewise continuous paths $t \mapsto X_t(\omega)$ in $\mathbb{R}^d$.
Subject to reasonable regularity conditions, the most general inhomogeneous continuous-time Markov process can be characterized by the diffusivity $a(x)$, the drift $b(x)$ and the Lévy measure $K(x, dy)$ depending on $x \in \mathbb{R}^d$. The interpretation of $K(x, dy)$ is to stay at x for a random time $T \sim \mathrm{Exp}\big(\lambda(x)\big)$ where $\lambda(x) := K(x, \mathbb{R}^d)$, then make a jump with distribution $\mu(x, dy) = K(x, dy)/\lambda(x)$.

Martingales play an important role in the description of the time evolution of such processes, which is demonstrated in the following non-rigorous calculation. Suppose that $d = 1$ and $(X_t)_{t\geq 0}$ is characterized by $a(x)$, $b(x)$ and $K(x, dy)$. Conditional on $X_s = x_s$ for all $s \leq t$

$$X_{t+dt} = x_t + b(x_t)\,dt + G_t + Y_t\,\mathbb{1}_{J_t}\ ,$$

where $G_t \sim N\big(0, a(x_t)\,dt\big)$, $Y_t \sim \mu(x_t, dy)$ and $J_t = \{\omega : \text{jump at } t\}$ is the event of having a jump in the time interval $(t, t+dt]$, i.e. $\mathbb{P}(J_t) = \lambda(x_t)\,dt$. Take $f \in C_b^2(\mathbb{R})$ (bounded and two times continuously differentiable). Then by Taylor expansion

$$\mathbb{E}\big(f(X_{t+dt}) - f(X_t)\,\big|\,X_s = x_s,\ s\leq t\big) = \\ = \mathbb{E}\Big(f'(x_t)\Big(b(x_t)\,dt + G_t + \frac{1}{2}f''(x_t)\,G_t^2\Big)\Big) + \mathbb{P}(J_t)\,\mathbb{E}\big(f(x_t + Y_t) - f(x_t)\big)\ .$$

Up to order dt this includes the second derivative f'', since $G_t \sim \sqrt{a(x_t)\,dt}\,N(0,1) = O(\sqrt{dt})$. This fact will be revisited later in a rigorous way as one of the main results of stochastic calculus (*Itô's formula*). Calculating the expectations we get

$$\mathbb{E}\big(f(X_{t+dt}) - f(X_t)\,\big|\,X_s = x_s,\ s\leq t\big) = L\,f(x_t)\,dt\ ,$$

where

$$L\,f(x) = \frac{1}{2}a(x)\,f''(x) + b(x)\,f'(x) + \int_{\mathbb{R}^d}\big(f(x+y) - f(x)\big)\,K(x, dy)\ ,$$

which is also called the *generator* of the process $(X_t)_{t\geq 0}$. It describes the infinitesimal expected time evolution of the process in the following sense. Consider

$$M_t^f = f(X_t) - f(X_0) - \int_0^t L\,f(X_s)\,ds\ .$$

We have calculated that $\mathbb{E}\big(M^f_{t+dt} - M^f_t \,\big|\, X_s = x_s,\ s \le t\big) = 0$ for all t, and so the process $(M^f_t)_{t\ge 0}$ is a martingale for all f. This property will often serve to characterize $(X_t)_{t\ge 0}$.

In the first part of this course, we will consider processes without jumps, i.e. $K \equiv 0$. Writing $\sigma(x) = \sqrt{a(x)}$ we have $G_t \sim \sigma(x_t)(B_{t+dt} - B_t)$ where $(B_t)_{t\ge 0}$ is a standard Brownian motion. In this case we get the *stochastic differential equation (SDE)*

$$X_{t+dt} - X_t = b(X_t)\,dt + \sigma(X_t)(B_{t+dt} - B_t) \quad \text{or} \quad dX_t = b(X_t)\,dt + \sigma(X_t)dB_t\ ,$$

with the integral form

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

The third term on the right is a stochastic integral, a sort of path integral in time along a Brownian motion, which we will define precisely in the next sections. After constructing the stochastic integral we will discuss various applications, including stochastic differential equations. The last part of the course will deal with pure Markov jump processes.

Stochastic calculus has very important application in sciences (biology or physics) as well as mathematical finance. For example, we will develop all the necessary tools to rigorously prove results like the Black-Scholes formula. But it is also a very beautiful part of modern probability and has let to a considerable enrichment of the subject itself.

1 Preliminaries

To make sense of a stochastic integral $\int H(s)\,dA(s)$, it turns out that the most general integrands H are previsible processes and the most general integrators are semimartingales, i.e. processes of the form $H = H_0 + M + A$ where M is a local martingale and A is a finite variation process. In this section we will define these processes and study some basic properties.

1.1 Finite variation integrals

Since finite variation is a pathwise property, we will first establish integrals with respect to deterministic integrants and lift it to stochastic processes in the last part of this section.
Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is *càdlàg* or *rcll* if it is right-continuous and has left limits. For such functions we write $\Delta f(t) := f(t) - f(t-)$ where $f(t-) = \lim_{s\uparrow t} f(s)$. Suppose $a : [0,\infty) \to \mathbb{R}$ is an increasing càdlàg function. Then there exists a unique Borel measure da on $(0,\infty)$ such that $da\big((s,t]\big) = a(t) - a(s)$, the *Lebesgue-Stieltjes measure* with distribution function a. For a non-negative measurable function h and $t \geq 0$ define

$$(h \cdot a)(t) = \int_{(0,t]} h(s)\,da(s)\ . \tag{1.1}$$

We may extend this definition to a càdlàg function $a = a' - a''$, where a' and a'' are both increasing càdlàg, and to integrable $h : [0,\infty) \to \mathbb{R}$. Subject to the finiteness of all the terms on the right we define

$$h \cdot a = h^+ \cdot a' - h^+ \cdot a'' - h^- \cdot a' + h^- \cdot a''\ . \tag{1.2}$$

where $h^\pm := \max\{\pm h, 0\}$ are the positive and negative part of h.
Which functions a may be expressed as the difference of two increasing càdlàg functions?

Lemma 1.1 *Let $a : [0,\infty) \to \mathbb{R}$ be càdlàg and define $v^n(0) = 0$, and for all $t > 0$*

$$v^n(t) = \sum_{k=0}^{\lceil 2^n t\rceil - 1} \big|a((k+1)2^{-n}) - a(k2^{-n})\big|\ . \tag{1.3}$$

Then $v(t) := \lim_{n\to\infty} v^n(t)$ exists for all $t \geq 0$ and is increasing in t.

Proof. Let $t_n^+ = 2^{-n}\lceil 2^n t\rceil$ and $t_n^- = 2^{-n}\big(\lceil 2^n t\rceil - 1\big)$ and write

$$v^n(t) = \sum_{\substack{I\in D_n\\ \inf I<t}} \big|da(I)\big| = \sum_{\substack{I\in D_n\\ \sup I<t}} \big|da(I)\big| + \big|a(t_n^+) - a(t_n^-)\big|\ . \tag{1.4}$$

where $D_n = \big\{(k2^{-n}, (k+1)2^{-n}]\ :\ k \in \mathbb{N}\big\}$. The first term is increasing in n by the triangle inequality, and so has a limit as $n \to \infty$. The second converges to $\big|\Delta a(t)\big| = \big|a(t) - a(t-)\big|$ as a is càdlàg, and so $v(t)$ exists for all $t \geq 0$.
Since $v^n(t)$ is increasing in t for all n, the same holds for $v(t)$. □

Definition 1.1 $v(t)$ is called the *total variation of a over* $(0,t]$ and a is said to be *of finite variation* if $v(t) < \infty$ for all $t \geq 0$.

Proposition 1.2 *A càdlàg function $a : [0,\infty) \to \mathbb{R}$ can be expressed as $a = a' - a''$, with a', a'' increasing and càdlàg, iff a is of finite variation. In this case, $t \mapsto v(t)$ is càdlàg with $\Delta v(t) = \big|\Delta a(t)\big|$ and $a^{\pm} := \frac{1}{2}(v \pm a)$ are the smallest functions a' and a'' with that property.*

Proof. Suppose $v(t) < \infty$ for all $t \geq 0$. Fix $T \in \mathbb{N}$ and consider

$$u^n(t) = \sum_{\substack{I\in D_n\\ t<\inf I<T}} \big|da(I)\big| \quad \text{for } t \leq T\ . \tag{1.5}$$

$u^n(t)$ is clearly decreasing in t. For $I = (c,d] \in D_n$, u^n is constant on $[c,d)$ and hence right-continuous. Therefore $\big\{t \in [0,T] \,:\, u^n(t) \leq x\big\}$ is closed for all $x \geq 0$.
$u^n(t)$ is increasing in n and so has a limit $u(t)$ for all $t \leq T$. We have that

$$\big\{t \in [0,T] \,:\, u(t) \leq x\big\} = \bigcap_{n\in\mathbb{N}} \big\{t \in [0,T] \,:\, u^n(t) \leq x\big\} \tag{1.6}$$

is closed as a countable intersection of closed sets. Thus, since $u(t)$ is decreasing in t it is right-continuous. For all $t < T$

$$v^n(T) = v^n(t) + u^n(t) + \big|a(t + 2^{-n}) - a(t)\big|\, 1_{t\in D_n}\ , \tag{1.7}$$

where the final term on the right converges to 0 as $n \to \infty$ because a is right-continuous. Hence for all $t < T$ we have $v(t) = v(T) - u(t)$ and since T was arbitrary v is càdlàg. The left limits exist since v is increasing, and taking the limit $n \to \infty$ in (1.4) we get $v(t) = v(t-) + \big|\Delta a(t)\big|$. With v also

$$a^+ = \frac{1}{2}(v + a) \quad \text{and} \quad a^- = \frac{1}{2}(v - a) \quad \text{are càdlàg.} \tag{1.8}$$

Then for each $m \in \mathbb{Z}^+$,

$$\begin{aligned} & dv^m(I) = \big|da(I)\big| && \text{for all } I \in D_m \\ \text{and} \quad & dv^n(I) \geq \big|da(I)\big| && \text{for all } I \in D_m \text{ if } n \geq m\ . \end{aligned} \tag{1.9}$$

Thus $da^{\pm}(I) = \frac{1}{2}dv(I) \pm \frac{1}{2}da(I) \geq 0$ for all $I \in \bigcup_{m\geq 1} D_m$, and so $a^{\pm}$ are increasing. Suppose now $a = a' - a''$ where a', a'' are increasing with $a(0) = a'(0) = a''(0) = 0$ wlog. Then for any $I \in D_n$, $n \geq 0$,

$$\big|da(I)\big| \leq da'(I) + da''(I) \tag{1.10}$$

and the sum (1.4) telescopes to give

$$v(t) \leftarrow v^n(t) \leq a'(t_n^+) + a''(t_n^+) \to a'(t) + a''(t)\ , \tag{1.11}$$

since a' and a'' are right-continuous. Thus $v(t) \leq a'(t) + a''(t) < \infty$ for all $t \geq 0$.
The last inequality also shows that a^+ and a^- as defined in (1.8) are the smallest increasing functions a', a'', such that $a = a' - a''$. □

Suppose now that we have a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. Recall that a process $X : \Omega \times [0,\infty) \to \mathbb{R}$ is *adapted* to $(\mathcal{F}_t)_{t\geq 0}$ if $X_t = X(.,t)$ is $\mathcal{F}_t$-measurable for all $t \geq 0$, and X is *càdlàg* if $X(\omega,.)$ is càdlàg for all $\omega \in \Omega$.

Definition 1.2 Let A be a càdlàg adapted process. Its *total variation process* V is defined pathwise (for each $\omega \in \Omega$) as the total variation of $A(\omega, .)$. We say that A is of *finite variation* if $A(\omega, .)$ is of finite variation for all $\omega \in \Omega$.

Lemma 1.3 *Let A be a càdlàg adapted process with finite total variation V. Then V is càdlàg adapted and pathwise increasing.*

Proof. Using the same partition as in (1.4) we get

$$V_t = \lim_{n\to\infty} \tilde{V}_t^n + \big|\Delta A_t\big| \tag{1.12}$$

where $\tilde{V}_t^n := \sum_{k=0}^{2^n t_n^- - 1} \big|A_{(k+1)2^{-n}} - A_{k2^{-n}}\big|$ is adapted for all $n \in \mathbb{N}$ since $t_n^- < t$ and ΔA_t is $\mathcal{F}_t$-measurable since A is càdlàg adapted. Thus V is adapted and it is càdlàg and increasing because $V(\omega, .)$ is càdlàg and increasing for all $\omega \in \Omega$. □

In the next section we will introduce a suitable class of integrands H for a pathwise definition of the stochastic integral

$$(H \cdot A)(\omega, t) = \int_{(0,t]} H(\omega, s)\, dA(\omega, s)\ . \tag{1.13}$$

1.2 Previsible processes

Definition 1.3 The *previsible σ-algebra $\mathcal{P}$* on $\Omega \times (0, \infty)$ is the σ-algebra generated by sets of the form $E \times (s, t]$ where $E \in \mathcal{F}_s$ and $s < t$. A *previsible process* H is a $\mathcal{P}$-measurable map $H : \Omega \times (0, \infty) \to \mathbb{R}$.

Proposition 1.4 *Let X be càdlàg adapted and $H_t = X_{t-}$, $t > 0$. Then H is previsible.*

Proof. $H : \Omega \times (0, \infty) \to \mathbb{R}$ is left-continuous and adapted.
Set $t_n^- = k2^{-n}$ when $k2^{-n} < t \le (k + 1)2^{-n}$ and

$$H_t^n = H_{t_n^-} = \sum_{k=0}^{\infty} H_{k2^{-n}}\, 1_{(k2^{-n},(k+1)2^{-n}]}(t)\ . \tag{1.14}$$

So H^n is previsible for all $n \in \mathbb{N}$ since $H_{t_n^-}$ is $F_{t_n^-}$-measurable as H is adapted and $t_n^- < t$. But $t_n^- \nearrow t$ and so $H_t^n \to H_t$ as $n \to \infty$ by left-continuity and H is also previsible. □

Proposition 1.5 *Let H be 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 first example sheet.

Remark. $\mathcal{P}$ is the smallest σ-algebra s.th. all adapted left-cont. processes are measurable.

Examples. (i) Brownian motion is previsible by Proposition 1.4, since it is continuous.

(ii) A Poisson process $(N_t)_{t\ge 0}$ or, indeed, any other continuous-time Markov chain with discrete state space is not previsible, since N_t is not $\mathcal{F}_{t-}$-measurable.

Proposition 1.6 *Let A be a càdlàg adapted finite variation process with total variation V. Let H be previsible such that for all $t \geq 0$ and all $\omega \in \Omega$*

$$\int_{(0,t]} \big|H(\omega,s)\big|\, dV(\omega,s) < \infty\ . \tag{1.15}$$

Then the process defined pathwise by

$$(H\cdot A)_t = \int_{(0,t]} H_s\, dA_s \tag{1.16}$$

is càdlàg, adapted and of finite variation.

Proof. We show that $(H\cdot A)$ is càdlàg for each fixed $\omega\in\Omega$.
We have $\mathbb{1}_{(0,s]} \to \mathbb{1}_{(0,t]}$ as $s \searrow t$, $\mathbb{1}_{(0,s]} \to \mathbb{1}_{(0,t)}$ as $s \nearrow t$ and

$$(H\cdot A)_t = \int_{(0,\infty)} H_s\, \mathbb{1}_{(0,t]}(s)\, dA_s\ . \tag{1.17}$$

Hence, by dominated convergence, the following limits exist

$$(H\cdot A)_{t+} = (H\cdot A)_t \quad \text{and} \quad (H\cdot A)_{t-} = \int_{(0,\infty)} H_s\, \mathbb{1}_{(0,t)}(s)\, dA_s\ , \tag{1.18}$$

and $H\cdot A$ is càdlàg with $\Delta(H\cdot A)_t = \int_{(0,\infty)} H_s\, \mathbb{1}_{\{t\}}(s)\, dA_s = H_t \Delta A_t$.

Next, we show that $H\cdot A$ is adapted via a monotone class argument.
Suppose first $H = \mathbb{1}_{B\times(s,u]}$ where $B\in\mathcal{F}_s$. Then $(H\cdot A)_t = \mathbb{1}_B(A_{t\wedge u} - A_{s\wedge t})$ which is clearly $\mathcal{F}_t$-measurable. Now let

$$\Pi = \big\{B\times(s,u]\ :\ B\in\mathcal{F}_s, s<u\big\} \quad \text{and} \tag{1.19}$$

$$\mathcal{A} = \big\{C\in\mathcal{P}\ :\ (\mathbb{1}_C\cdot A)_t \text{ is } \mathcal{F}_t\text{-measurable}\big\} \tag{1.20}$$

so that Π is a π-system and $\Pi\subseteq\mathcal{A}$. But $\mathcal{A}\subseteq\mathcal{P}=\sigma(\Pi)$ and $\mathcal{A}$ is a d-system. $\big[\,\Omega\times(0,\infty)\in\mathcal{A}$; if $C\subseteq D\in\mathcal{A}$ then $\big((\mathbb{1}_D - \mathbb{1}_C)\cdot A\big)_t$ is $\mathcal{F}_t$-measurable, which gives $D\setminus C\in\mathcal{A}$; if $C_n\in\mathcal{A}$ with $C_n\nearrow C$ then $C\in\mathcal{A}$ since a limit of measurable functions is measurable.$\big]$
Hence, by Dynkin's lemma, $\mathcal{A}=\mathcal{P}$. Suppose now that H is non-negative, $\mathcal{P}$-measurable. For all $n\in\mathbb{N}$ set

$$H^n := \big(2^{-n}\lfloor 2^n H\rfloor\big)\wedge n = \sum_{k=1}^{2^n n} 2^{-n}k\, \mathbb{1}\underbrace{\big\{H\in[2^{-n}k, 2^{-n}(k+1))\big\}}_{\in\mathcal{P}}\ , \tag{1.21}$$

so that $(H^n\cdot A)_t$ is $\mathcal{F}_t$-measurable. We have $(H^n\cdot A)_t \nearrow (H\cdot A)_t$ by monotone convergence (applied for each ω). Hence, $(H\cdot A)_t$ is $\mathcal{F}_t$-measurable. This extends in the usual way to $\mathcal{P}$-measurable $H = H^+ - H^-$ such that $|H|\cdot V = (H^+\cdot V) + (H^-\cdot V) < \infty$.

We show finite variation for each fixed $\omega\in\Omega$. If $H^\pm = \max\{\pm H, 0\}$, $A^\pm = \frac{1}{2}(V\pm A)$ then analogous to (1.2)

$$H\cdot A = (H^+\cdot A^+ + H^-\cdot A^-) - (H^+\cdot A^- + H^-\cdot A^+)\ . \tag{1.22}$$

This is the difference of two increasing functions and thus $H\cdot A$ is of finite variation. □

Example. Suppose that H is a previsible process, such as Brownian motion, and that

$$\int_{(0,t]} |H_s|\, ds < \infty \quad \text{for all } \omega\in\Omega \text{ and } t\geq 0\ . \tag{1.23}$$

Then $\int_{(0,t]} H_s ds$ is càdlàg, adapted and of finite variation.

1.3 Local martingales

We 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, i.e. $\mathcal{F}$ is $\mathbb{P}$-complete, $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets in $\mathcal{F}$ and $(\mathcal{F}_t)_{t\geq 0}$ is right-continuous in that

$$\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s \quad \text{for all } t \geq 0\ . \tag{1.24}$$

Recall that an adapted integrable $\big(\mathbb{E}(|X_t|) < \infty$ for all $t \geq 0\big)$ process X is a *martingale* if

$$\mathbb{E}(X_t|\mathcal{F}_s) = X_s\ a.s. \quad \text{for all } s \leq t\ . \tag{1.25}$$

We write $\mathcal{M}$ for the set of all càdlàg martingales. $T : \Omega \to [0, \infty]$ is a *stopping time* if $\{T \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$ and we set $\mathcal{F}_T = \big\{A \in \mathcal{F} \,:\, A \cap \{T \leq t\} \in \mathcal{F}_t,\, t \geq 0\big\}$. If X is a càdlàg adapted process and T is a finite stopping time then X_T is $\mathcal{F}_T$-measurable (see Prop. 4.1.1 of Advanced Probability).

Optional stopping theorem (OST)
Let X be an adapted integrable process. Then the following are equivalent:

(i) X is a martingale (ii) $X^T = (X_{t\wedge T})_t$ is a martingale for all stopping times T.

(iii) For all stopping times S, T with $T < \infty$, $\mathbb{E}(X_T|\mathcal{F}_S) = X_{S\wedge T}\ a.s.$.

(iv) $\mathbb{E}(X_T) = \mathbb{E}(X_0)$ for all stopping times $T < \infty$.

(i) implies that $\mathcal{M}$ is *stable under stopping*. For any such set there exists a localized version.

Definition 1.4 A càdlàg adapted process X is a *local martingale*, $X \in \mathcal{M}_{loc}$, if there exists a sequence $(T_n)_{n\in\mathbb{N}}$ of stopping times with $T_n \nearrow \infty$ such that $(X_t^{T_n})_{t\geq 0} \in \mathcal{M}$ for all $n \in \mathbb{N}$. We say that the sequence $(T_n)_{n\in\mathbb{N}}$ *reduces* X.

In particular $\mathcal{M} \subseteq \mathcal{M}_{loc}$ since any sequence $(T_n)_{n\in\mathbb{N}}$ of stopping times reduces X by OST(i). Recall that a set $\mathcal{X}$ of random variables is *uniformly integrable (UI)* if

$$\sup_{X\in\mathcal{X}} \mathbb{E}\big(|X|\, 1_{\{|X|\geq\lambda\}}\big) \to 0 \quad \text{as } \lambda \to \infty\ . \tag{1.26}$$

Lemma 1.7 *If $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ then the set*

$$\mathcal{X} = \big\{\mathbb{E}(X|\mathcal{G}) \,:\, \mathcal{G} \text{ is a sub-}\sigma\text{-algebra of } \mathcal{F}\big\} \quad \text{is UI}\ . \tag{1.27}$$

Proof. See exercise 9.1.7 of Advanced Probability

We can now give necessary and sufficient conditions for a local martingale to be a martingale.

Proposition 1.8 *The following statements are equivalent:*

(i) X is a martingale

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

$$\mathcal{X}_t = \big\{X_T \,:\, T \text{ is a stopping time},\ T \leq t\big\} \quad \text{is UI}\ . \tag{1.28}$$

Proof. Suppose (i) holds. By OST, if T is a stopping time with $T \leq t$, then $X_T = \mathbb{E}(X_t|\mathcal{F}_T)\ a.s.$.
But then by Lemma 1.7 $\mathcal{X}_t$ is UI.
If (ii) holds, suppose $(T_n)_{n\geq 0}$ reduces X. But then for any bounded stopping time, $T \leq t$, say,

$$\mathbb{E}(X_0) = \mathbb{E}(X_0^{T_n}) \overset{OST}{=} \mathbb{E}(X_T^{T_n}) = \mathbb{E}(X_{T\wedge T_n}) \ . \tag{1.29}$$

$\{X_{T\wedge T_n} \ : \ n \in \mathbb{N}\}$ is UI and so $\mathbb{E}(X_{T\wedge T_n}) \to \mathbb{E}(X_T)$ as $n \to \infty$. But then by OST again, X must be a martingale. □

Remark. This implies that a bounded local martingale, or indeed a local martingale M such that for all $t \geq 0$, $|M_t| \leq Z$ for some $Z \in L^1$, is a true martingale.

Example. Let B be a standard Brownian motion in $\mathbb{R}^3$ with $|B_0| = 1$, and let $M_t := 1/|B_t|$. The sequence $(T_n)_{n\in\mathbb{N}}$ with $T_n := \inf\{t : |M_t| > n\}$ is reducing for M, and for $t > s$

$$\mathbb{E}(M_{t\wedge T_n}|\mathcal{F}_s) = M_{s\wedge T_n} \quad \text{but} \quad \mathbb{E}(M_t|\mathcal{F}_s) < M_s \ , \tag{1.30}$$

since the variables $\{M_{t\wedge T_n} \ : \ n \in \mathbb{N}\}$ are not UI.

Proposition 1.9 *Let M be a continuous local martingale ($M \in \mathcal{M}_{c,loc}$) starting from 0. Set $S_n = \inf\big\{t \geq 0 \ : \ |M_t| = n\big\}$. Then S_n is a stopping time for all $n \in \mathbb{N}$, $S_n \nearrow \infty\ a.s.$ as $n \to \infty$ and M^{S_n} is a martingale for all n, i.e. $(S_n)_{n\geq 0}$ reduces M.*

Proof. Note that $$\{S_n \leq t\} = \bigcap_{k\in\mathbb{N}} \bigcup_{\substack{s\in\mathbb{Q}\\ s\leq t}} \big\{|M_s| > n - 1/k\big\} \in \mathcal{F}_t \ ,$$

and so S_n is a stopping time. For each $\omega \in \Omega$ and $t \geq 0$ we have by continuity

$$\sup_{s\leq t} \big|M_s(\omega)\big| < \infty \quad \text{and} \quad S_n(\omega) > t \text{ for all } n > \sup_{s\leq t} \big|M_s(\omega)\big| \ . \tag{1.31}$$

Hence, $S_n \nearrow \infty\ a.s.$. Let $(T_k)_{k\in\mathbb{N}}$ be a reducing sequence for M, i.e. $M^{T_k} \in \mathcal{M}$. By OST, also $M^{S_n\wedge T_k} \in \mathcal{M}$ and so $M^{S_n} \in \mathcal{M}_{loc}$ for each $n \in \mathbb{N}$. But M^{S_n} is bounded and so also a martingale. □

Theorem 1.10 *Let $M \in \mathcal{M}_{c,loc}$ be of finite variation, started at 0. Then $M \equiv 0\ a.s.$.*

Remarks. (i) In particular Brownian motion is not of finite variation.

(ii) This makes it clear that the theory of finite variation integrals we have developed is useless for integrating w.r.t. continuous local martingales.

Proof. Let V denote the total variation process of M. Then V is continuous and adapted with $V_0 = 0$. Set $S_n = \inf\{t \geq 0 \ : \ V_t = n\}$. Then S_n is a stopping time for all $n \in \mathbb{N}$ since V is adapted, and $S_n \nearrow \infty$ as $n \to \infty$ since V_t is monotone increasing and finite for all t
$geq0$.
It suffices to show $M^{S_n} \equiv 0$ for all $n \in \mathbb{N}$. By OST, $M^{S_n} \in \mathcal{M}_{loc}$. Also

$$|M_t^{S_n}| \leq |V_t^{S_n}| \leq n \ , \tag{1.32}$$

and so, by Proposition 1.8, $M^{S_n} \in \mathcal{M}$.
Replacing M by M^{S_n} we can reduce to the case where M is a bounded martingale of bounded variation, i.e. V is bounded. Fix $t > 0$ and set $t_k = kt/N$. Then

$$\mathbb{E}(M_{t_k} M_{t_{k+1}}) = \mathbb{E}\big(M_{t_k}\, \mathbb{E}(M_{t_{k+1}} | \mathcal{F}_{t_k})\big) = \mathbb{E}(M_{t_k}^2) \;, \tag{1.33}$$

which leads to

$$\begin{aligned}\mathbb{E}(M_t^2) &= \mathbb{E}\bigg(\sum_{k=0}^{N-1} (M_{t_{k+1}}^2 - M_{t_k}^2)\bigg) = \mathbb{E}\bigg(\sum_{k=0}^{N-1} (M_{t_{k+1}} - M_{t_k})^2\bigg) \leq \\ &\leq \mathbb{E}\bigg(\underbrace{\sup_{k<N} |M_{t_{k+1}} - M_{t_k}|}_{\leq V_t \leq n}\; \underbrace{\sum_{k=0}^{N-1} |M_{t_{k+1}} - M_{t_k}|}_{\leq V_t \leq n}\bigg) \;.\end{aligned} \tag{1.34}$$

As M is bounded and continuous

$$\sup_{k<N} |M_{t_{k+1}} - M_{t_k}| \to 0 \quad \text{as } N \to \infty \;, \tag{1.35}$$

and so, by bounded convergence,

$$\mathbb{E}\bigg(\sup_{k<N} |M_{t_{k+1}} - M_{t_k}| \sum_{k=0}^{N-1} |M_{t_{k+1}} - M_{t_k}|\bigg) \to 0 \quad \text{as } N \to \infty \;. \tag{1.36}$$

Hence, $\mathbb{E}(M_t^2) = 0$ for all $t \geq 0$ which implies that $M \equiv 0$ $a.s.$. □

Definition 1.5 A *semimartingale* X is an adapted càdlàg process which may be written as

$$X = X_0 + M + A \quad \text{with} \quad M_0 = A_0 = 0 \;, \tag{1.37}$$

where $M \in \mathcal{M}_{loc}$ and A is a finite variation process. This is sometimes called the *Doob-meyer decomposition*.

As a consequence of Theorem 1.10, the decomposition is unique for continuous semimartingales, where M is required to be continuous.

Example. Lévy processes
Let (a, q, Π) be a Lévy triple. By the Lévy-Îto theorem, the associated Lévy process in $\mathbb{R}$ may be written as

$$X_t = at + \sqrt{q}\, B_t + Y_t + Z_t \tag{1.38}$$

where Y is a compound Poisson process and Z a compensated càdlàg martingale with

$$Y_t = \int_0^t \int_{|y|\geq 1} y\, M(ds, dy) \quad \text{and} \quad Z_t = \int_0^t \int_{|y|<1} y\, \tilde{M}(ds, dy) \;. \tag{1.39}$$

$M \sim PRM\big(dt \otimes \Pi(dx)\big)$ is a Poisson random measure and $\tilde{M}$ is the compensated version of M. $at + Y_t$ is a process of finite variation and $\sqrt{q}\, B_t + Z_t$ is a (square-integrable) martingale. So X is a semimartingale.

2 The stochastic integral

In this section we establish the stochastic integral with respect to continuous semimartingales. In places, we develop parts of the theory also for càdlàg semimartingales, where this involves no extra work. However, parts of the construction will use crucially the assumption of continuity. A more general theory exists, but it is beyond the scope of this course. Similar to the construction to the Lebesgue integral in measure theory, we start with various (weak) forms of a stochastic integrals for special integrands, before working our way up to the most general form.

2.1 Simple integrands

Definition 2.1 A *simple process* is any map $H : \Omega \times (0, \infty) \to \mathbb{R}$ of the form

$$H(\omega, t) = \sum_{k=0}^{n-1} Z_k(\omega)\, 1_{(t_k,t_{k+1}]}(t) \tag{2.1}$$

where $n \in \mathbb{N}$, $0 = t_0 < \ldots < t_n < \infty$ and Z_k is a bounded $\mathcal{F}_{t_k}$-measurable random variable for all k. We denote the set of simple processes by $\mathcal{S}$.

Note that $\mathcal{S}$ is a vector space and that by Proposition 1.4 every simple process is previsible. Recall that we say a process X is *bounded in* L^2 if

$$\sup_{t\geq 0} \|X_t\|_2 = \sup_{t\geq 0} \mathbb{E}\big(|X_t|^2\big)^{1/2} < \infty\ . \tag{2.2}$$

Write $\mathcal{M}^2$ for the set of all càdlàg L^2-bounded martingales. If $X \in \mathcal{M}^2$ the L^2 **martingale convergence theorem** ensures the existence of $X_\infty \in L^2$ such that

$$X_t \to X_\infty \quad a.s. \text{ and in } L^2, \text{ as } t \to \infty\ . \tag{2.3}$$

Moreover, $X_t = \mathbb{E}(X_\infty|\mathcal{F}_t)$ $a.s.$ for all $t \geq 0$.
Recall **Doob's** L^2 **inequality** for $X \in \mathcal{M}^2$, $\quad \big\|\sup_{t\geq 0} |X_t|\big\|_2 \leq 2\|X_\infty\|_2\ .$

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

$$(H \cdot M)_t = \sum_{k=0}^{n-1} Z_k\big(M_{t_{k+1}\wedge t} - M_{t_k\wedge t}\big)\ . \tag{2.4}$$

Proposition 2.1 *Let $H \in \mathcal{S}$ and $M \in \mathcal{M}^2$. Let T be a stopping time. Then*

(i) $H \cdot M^T = (H \cdot M)^T$. *(ii)* $H \cdot M \in \mathcal{M}^2$.

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

Proof. (i) For all $t \geq 0$ we have

$$\begin{aligned}(H \cdot M^T)_t &= \sum_{k=0}^{n-1} Z_k\big(M^T_{t_{k+1}\wedge t} - M^T_{t_k\wedge t}\big) = \\ &= \sum_{k=0}^{n-1} Z_k\big(M_{t_{k+1}\wedge t\wedge T} - M_{t_k\wedge t\wedge T}\big) = (H \cdot M)_{t\wedge T} = (H \cdot M)^T_t\ .\end{aligned} \tag{2.5}$$

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

$$\mathbb{E}\big((H \cdot M)_t - (H \cdot M)_s \,\big|\, \mathcal{F}_s\big) = Z_k \,\mathbb{E}(M_t - M_s|\mathcal{F}_s) = 0 \;. \tag{2.6}$$

This extends easily to general $s \leq t$ and hence $H \cdot M$ is a martingale. To show it is bounded in L^2, note that if $j < k$ we have

$$\begin{aligned} &\mathbb{E}\Big(Z_j(M_{t_{j+1}} - M_{t_j})\, Z_k(M_{t_{k+1}} - M_{t_k})\Big) = \\ &\qquad = \mathbb{E}\Big(Z_j(M_{t_{j+1}} - M_{t_j})\, Z_k \,\mathbb{E}\big(M_{t_{k+1}} - M_{t_k} \,\big|\, \mathcal{F}_{t_k}\big)\Big) = 0 \;. \end{aligned} \tag{2.7}$$

Thus

$$\begin{aligned} \mathbb{E}\big((H \cdot M)_t^2\big) &= \mathbb{E}\bigg(\Big(\sum_{k=0}^{n-1} Z_k(M_{t_{k+1}} - M_{t_k})\Big)^2\bigg) = \sum_{k=0}^{n-1} \mathbb{E}\big(Z_k^2(M_{t_{k+1}\wedge t} - M_{t_k\wedge t})^2\big) \leq \\ &\leq \|H\|_\infty^2 \sum_{k=0}^{n-1} \mathbb{E}\big((M_{t_{k+1}\wedge t} - M_{t_k\wedge t})^2\big) = \|H\|_\infty^2 \mathbb{E}\big((M_t - M_0)^2\big) \;, \end{aligned} \tag{2.8}$$

and so, by Doob's L^2 inequality applied to $M - M_0$,

$$\sup_{t\geq 0} \mathbb{E}\big((H \cdot M)_t^2\big) \leq 4\|H\|_\infty^2 \mathbb{E}\big((M_\infty - M_0)^2\big) \;. \tag{2.9}$$

Hence, $H \cdot M \in \mathcal{M}^2$. (iii) (2.8) holds in particular for $t = \infty$ to give the result. □

Proposition 2.2 *Let μ be a finite measure on the previsible σ-algebra $\mathcal{P}$. Then $\mathcal{S}$ is a dense subspace of $L^2(\mathcal{P}, \mu)$.*

Proof. As the Z_k's are bounded, it is certainly true that $\mathcal{S} \subseteq L^2(\mathcal{P}, \mu)$. Denote by $\bar{\mathcal{S}}$ the closure of $\mathcal{S}$ in $L^2(\mathcal{P}, \mu)$. Set

$$\mathcal{A} = \big\{A \in \mathcal{P} \,:\, \mathbb{1}_A \in \bar{\mathcal{S}}\big\} \;. \tag{2.10}$$

Then $\mathcal{A}$ is a d-system. $\big[$ Check: $\mathbb{1}_{\Omega\times(0,n)} \in \mathcal{S}$ so $\mathbb{1}_{\Omega\times(0,\infty)} \in \bar{\mathcal{S}}$ and $\Omega \times (0, \infty) \in \mathcal{A}$; if $C \subseteq D \in \mathcal{A}$ then $D \setminus C \in \mathcal{A}$; if $C_n \in \mathcal{A}$ and $C_n \nearrow C$ then $C \in \mathcal{A}$ since $\bar{\mathcal{S}}$ is the closure of $\mathcal{S}$ in $L^2(\mathcal{P}, \mu)$ $\big]$. $\mathcal{A}$ contains the π-system $\{B \times (s, t] \,:\, B \in \mathcal{F}_s,\ s \leq t\}$, which generates $\mathcal{P}$. Hence, by Dynkin's lemma, $\mathcal{A} = \mathcal{P}$. The result now follows since linear combinations of measurable indicator functions are dense in L^2. □

2.2 L^2 properties

To extend the simple integral defined in the last section, we will need some Hilbert space properties of the set of integrators we are considering. As before, we 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.

Definition 2.2 For all càdlàg adapted processes X define the norm[1] $|||X||| := \|X^*\|_2$ where $X^* = \sup_{t\geq 0} |X_t|$ and write $\mathcal{C}^2$ for the set of all càdlàg adapted processes X with $|||X||| < \infty$. On $\mathcal{M}^2 = \big\{X \in \mathcal{M} \,:\, \sup_{t\geq 0} \|X_t\|_2\big\}$ define the norm $\|X\| := \|X_\infty\|_2$.

[1] Usual conventions about versions apply, dealing with equivalence classes analogous to L^p-spaces

Proposition 2.3 *We have* *(i)* $(\mathcal{C}^2, |||.|||)$ *is complete* *(ii)* $\mathcal{M}^2 = \mathcal{M} \cap \mathcal{C}^2$

(iii) $(\mathcal{M}^2, \|.\|)$ *is a Hilbert space with* $\mathcal{M}_c^2 = \mathcal{M}_c \cap \mathcal{M}^2$ *as a closed subspace*

(iv) $X \mapsto X_\infty : \mathcal{M}^2 \to L^2(\mathcal{F}_\infty)$ *is an isometry*

Remark. We identify an element of $\mathcal{M}^2$ with its terminal value and then $\mathcal{M}^2$ inherits the Hilbert space structure of $L^2(\mathcal{F}_\infty)$ with inner product $(X, Y) \mapsto \mathbb{E}(X_\infty Y_\infty)$.

Proof. (i) Suppose $(X^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $(\mathcal{C}^2, |||.|||)$. Then we can find a subsequence $(n_k)_{k\in\mathbb{N}}$ such that $\sum_{k=1}^\infty |||X^{n_{k+1}} - X^{n_k}||| < \infty$. Then

$$\Big\| \sum_{k=1}^\infty \sup_{t\geq 0} \big|X_t^{n_{k+1}} - X_t^{n_k}\big| \Big\|_2 \leq \sum_{k=1}^\infty |||X^{n_{k+1}} - X^{n_k}||| < \infty \tag{2.11}$$

and so for a.e. $\omega \in \Omega$, $\displaystyle\sum_{k=1}^\infty \sup_{t\geq 0} \big|X_t^{n_{k+1}}(\omega) - X_t^{n_k}(\omega)\big| < \infty$,

and so $\big(X_t^{n_k}(\omega)\big)_{k\in\mathbb{N}} \to X(\omega)$ as $k \to \infty$ uniformly in $t \geq 0$. $X(\omega)$ is càdlàg as it is the uniform limit of càdlàg functions (see example sheet 1.10). Now

$$\begin{aligned} |||X^n - X|||^2 &= \mathbb{E}\big(\sup_{t\geq 0} |X_t^n - X_t|^2\big) \leq \liminf_{k\to\infty} \mathbb{E}\big(\sup_{t\geq 0} |X_t^n - X_t^{n_k}|^2\big) = \qquad \text{(Fatou)} \\ &= \liminf_{k\to\infty} |||X^n - X^{n_k}|||^2 \to 0 \quad as n \to \infty \end{aligned} \tag{2.12}$$

because $(X^n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Hence $(\mathcal{C}^2, |||.|||)$ is complete.
(ii) For $X \in \mathcal{C}^2 \cap \mathcal{M}$ we have

$$\sup_{t\geq 0} \|X_t\|_2 \leq \big\| \sup_{t\geq 0} |X_t| \big\|_2 = |||X||| < \infty \ , \quad \text{and so } X \in \mathcal{M}^2 \ . \tag{2.13}$$

On the other hand, if $X \in \mathcal{M}^2$, by Doob's inequality,

$$|||X||| \leq 2\|X\| < \infty \ , \quad \text{and so } X \in \mathcal{C}^2 \cap \mathcal{M} \ . \tag{2.14}$$

(iii) $(X, Y) \mapsto \mathbb{E}(X_\infty Y_\infty)$ defines an inner product on $\mathcal{M}^2$. Moreover, for $X \in \mathcal{M}^2$, we have shown in (ii) that

$$\|X\| \leq |||X||| \leq 2\|X\| \ , \quad \text{i.e. } \|.\| \text{ and } |||.||| \text{ are equivalent} \ . \tag{2.15}$$

Thus $\mathcal{M}^2$ is complete for $\|.\|$ iff it is complete for $|||.|||$, and by (i) it is thus sufficient to show that $\mathcal{M}^2$ is closed in $(\mathcal{C}^2, |||.|||)$. If $X^n \in \mathcal{M}^2$ and $|||X^n - X||| \to 0$ as $n \to \infty$ for some X, then X is certainly càdlàg adapted and L^2-bounded. Furthermore,

$$\begin{aligned} \big\|\mathbb{E}(X_t|\mathcal{F}_s) - X_s\big\|_2 &\leq \big\|\mathbb{E}(X_t - X_t^n|\mathcal{F}_s)\big\|_2 + \|X_s^n - X_s\|_2 \leq \\ &\leq \|X_t - X_t^n\|_2 - \|X_s^n - X_s\|_2 \leq 2|||X^n - X||| \to 0 \end{aligned} \tag{2.16}$$

as $n \to \infty$ and so $X \in \mathcal{M}^2$. By the same argument $\mathcal{M}_c^2$ is closed in $(\mathcal{M}^2, \|.\|)$, where continuity of $t \mapsto X_t(\omega)$ follows by uniform convergence in t.
(iv) For $X, Y \in \mathcal{M}^2$, $\|X - Y\| = \|X_\infty - Y_\infty\|_2$ by definition. □

2.3 Quadratic variation

Definition 2.3 For a sequence $(X^n)_{n\in\mathbb{N}}$ we say that $X^n \to X$ *uniformly on compacts in probability* (*u.c.p.*) if

$$\forall \epsilon > 0\, \forall t \geq 0\ :\ \mathbb{P}\big(\sup_{s\leq t} |X_s^n - X_s| > \epsilon\big) \to 0 \quad \text{as } n \to \infty\ . \tag{2.17}$$

Theorem 2.4 Quadratic variation
For each $M \in \mathcal{M}_{c,loc}$ there exists a (up to versions) unique continuous adapted increasing process $[M]$ such that $M^2 - [M] \in \mathcal{M}_{c,loc}$. Moreover, for

$$[M]_t^n := \sum_{k=0}^{\lfloor 2^n t\rfloor -1} \big(M_{(k+1)2^{-n}} - M_{k2^{-n}}\big)^2 \tag{2.18}$$

we have $[M]^n \to [M]$ $u.c.p.$ *as $n \to \infty$. We call $[M]$ the quadratic variation process of M.*

Example. Let B be a standard Brownian motion. Then we know that $(B_t^2 - t)_{t\geq 0} \in \mathcal{M}_c$. But then by Theorem 2.4, $t = [B]_t$.

Lemma. Let $M \in \mathcal{M}$ be bounded. Suppose that $l \in \mathbb{N}$ and $0 = t_0 < t_1 < \ldots < t_l < \infty$. Then $\mathbb{E}\Big(\Big(\sum_{k=0}^{l-1}(M_{t_{k+1}} - M_{t_k})^2\Big)^2\Big)$ is bounded.

Proof of the Lemma. First note that

$$\mathbb{E}\Big(\Big(\sum_{k=0}^{l-1}(M_{t_{k+1}} - M_{t_k})^2\Big)^2\Big) = \sum_{k=0}^{l-1}\mathbb{E}\big((M_{t_{k+1}} - M_{t_k})^4\big) + \\ +2\sum_{k=0}^{l-1}\mathbb{E}\Big((M_{t_{k+1}} - M_{t_k})^2 \sum_{j=k+1}^{l-1}(M_{t_{j+1}} - M_{t_j})^2\Big)\ . \tag{2.19}$$

For each fixed k we have

$$\mathbb{E}\Big((M_{t_{k+1}} - M_{t_k})^2 \sum_{j=k+1}^{l-1}(M_{t_{j+1}} - M_{t_j})^2\Big) = \\ = \mathbb{E}\Big((M_{t_{k+1}} - M_{t_k})^2 \mathbb{E}\Big(\sum_{j=k+1}^{l-1}(M_{t_{j+1}} - M_{t_j})^2\Big|\mathcal{F}_{t_{k+1}}\Big)\Big) = \\ = \mathbb{E}\Big((M_{t_{k+1}} - M_{t_k})^2 \mathbb{E}\Big(\sum_{j=k+1}^{l-1}(M_{t_{j+1}}^2 - M_{t_j}^2)\Big|\mathcal{F}_{t_{k+1}}\Big)\Big) = \\ = \mathbb{E}\Big((M_{t_{k+1}} - M_{t_k})^2 \mathbb{E}\big(M_{t_l}^2 - M_{t_{k+1}}^2\big|\mathcal{F}_{t_{k+1}}\big)\Big) = \\ = \mathbb{E}\big((M_{t_{k+1}} - M_{t_k})^2 (M_{t_l}^2 - M_{t_{k+1}}^2)^2\big)\ . \tag{2.20}$$

After inserting this in (2.19) we get the estimate

$$\mathbb{E}\Big(\Big(\sum_{k=0}^{l-1}(M_{t_{k+1}} - M_{t_k})^2\Big)^2\Big) \leq \\ \leq \mathbb{E}\Big(\Big(\sup_j |M_{t_{j+1}} - M_{t_j}|^2 + 2\sup_j |M_{t_l} - M_{t_j}|^2\Big)\sum_{k=0}^{l-1}(M_{t_{k+1}} - M_{t_k})^2\Big)\ . \tag{2.21}$$

Now, M is uniformly bounded by C, say. So using the inequality $(x-y)^2 \le 2(x^2+y^2)$, we obtain

$$\mathbb{E}\Big(\Big(\sum_{k=0}^{l-1}(M_{t_{k+1}}-M_{t_k})^2\Big)^2\Big) \le 12C^2\mathbb{E}\Big(\sum_{k=0}^{l-1}(M_{t_{k+1}}-M_{t_k})^2\Big) = \\ = 12C^2\mathbb{E}\big((M_{t_l}-M_{t_0})^2\big) \le 48C^4 \;. \tag{2.22}$$

Proof of Theorem 2.4. Wlog we will consider the case $M_0 = 0$.
Uniqueness: If A and A' are two increasing processes satisfying the conditions for $[M]$ then

$$A_t - A'_t = (M_t^2 - A'_t) - (M_t^2 - A_t) \in \mathcal{M}_{c,loc} \tag{2.23}$$

is of finite variation and thus $A \equiv A'$ $a.s.$ by Theorem 1.10.
Existence: First we assume that M is bounded, which implies $M \in \mathcal{M}_c^2$. Fix $T > 0$ deterministic. Let

$$H_t^n = M^T_{2^{-n}\lfloor 2^n t\rfloor} = \sum_{k=0}^{\lfloor 2^n T\rfloor -1} M_{k2^{-n}}\, \mathbf{1}_{(k2^{-n},(k+1)2^{-n}]}(t) \;. \tag{2.24}$$

Then $H^n \in \mathcal{S}$ for all $n \in \mathbb{N}$. Hence X^n defined by

$$X_t^n = (H^n \cdot M)_t = \sum_{k=0}^{\lfloor 2^n T\rfloor -1} M_{k2^{-n}}\big(M_{(k+1)2^{-n}\wedge t} - M_{k2^{-n}\wedge t}\big) \;, \tag{2.25}$$

is in $\mathcal{M}_c^2$ by Proposition 2.1 and by continuity of M. Recall that $\|X^n\| = \|X_\infty^n\|_2 = \|X_T\|_2$ since X_t^n is constant for $t \ge T$. For $n \ge m$ we have

$$\begin{aligned}\|X^n - X^m\|^2 &= \mathbb{E}\big(((H^n - H^m)\cdot M)_T^2\big) \le \\ &\le \mathbb{E}\Big(\sup_{0\le t\le T}|H_t^n - H_t^m|^2 \sum_{k=0}^{\lfloor 2^n T\rfloor -1}(M_{(k+1)2^{-n}} - M_{k2^{-n}})^2\Big) \le \\ &\le \mathbb{E}\big(\sup_{0\le t\le T}|H_t^n - H_t^m|^4\big)^{1/2}\mathbb{E}\Big(\Big(\sum_{k=0}^{\lfloor 2^n T\rfloor -1}(M_{(k+1)2^{-n}} - M_{k2^{-n}})^2\Big)^2\Big)^{1/2}\end{aligned} \tag{2.26}$$

by Hölder's inequality. Since M is bounded, the second term is bounded by the lemma, and

$$\mathbb{E}\Big(\sup_{0\le t\le T}|H_t^n - H_t^m|^4\Big) = \mathbb{E}\Big(\sup_{0\le t\le T}|M_{2^{-n}\lfloor 2^n t\rfloor} - M_{2^{-m}\lfloor 2^m t\rfloor}|^4\Big) \to 0 \tag{2.27}$$

as $n, m \to \infty$ by uniform continuity of M on $[0,T]$ and bounded convergence. Hence X^n is a Cauchy sequence in $\big(\mathcal{M}_c^2, \|.\|\big)$ and so, by Proposition 2.3, converges to a limit $Y = (Y_t, 0 \le t \le T) \in \mathcal{M}_c^2$. Now for any n and $1 \le k \le \lfloor 2^n T\rfloor$,

$$M_{k2^{-n}}^2 - 2X_{k2^{-n}}^n = \sum_{j=0}^{k-1}(M_{(j+1)2^{-n}} - M_{j2^{-n}})^2 = [M]_{k2^{-n}}^n \;. \tag{2.28}$$

Hence, $M_t^2 - 2X_t^n$ is increasing along the sequence of times $\big(k2^{-n}, 1 \le k \le \lfloor 2^n T\rfloor\big)$. Passing to the limit $n \to \infty$, $M_t^2 - 2X_t^n$ must be $a.s.$ increasing. Set $[M]_t := M_t^2 - 2Y_t$ for $t \in [0,T]$

on the set where $M^2 - 2Y$ is increasing and $[M] \equiv 0$ otherwise. Hence, $[M]$ is a continuous increasing process and $M^2 - [M] = 2Y$ is a martingale on $[0,T]$.
We extend the definition of $[M]_t$ to $t \in [0,\infty)$ by applying the foregoing for all $T \in \mathbb{N}$. Note that the process $[M]$ obtained with T is the restriction to $[0,T]$ of $[M]$ defined with $T+1$.
Now, $\quad [M]^n_t = M^2_{2^{-n}\lfloor 2^n t\rfloor} - 2\,X^n_{2^{-n}\lfloor 2^n t\rfloor} \quad$ and so

$$\sup_{0\le t\le T}\big|[M]_t - [M]^n_t\big| \le \sup_{0\le t\le T}\big|M^2_{2^{-n}\lfloor 2^n t\rfloor} - M^2_t\big| + \\ +2\sup_{0\le t\le T}\big|X^n_{2^{-n}\lfloor 2^n t\rfloor} - Y_{2^{-n}\lfloor 2^n t\rfloor}\big| + 2\sup_{0\le t\le T}\big|Y_{2^{-n}\lfloor 2^n t\rfloor} - Y_t\big|\,. \tag{2.29}$$

Each term on the right converges to 0 in probability as $n\to\infty$, the first and third by continuity and the second because $X^n \to Y$ in $\big(\mathcal{M}^2_c, \|.\|\big)$ implies $\sup\limits_{0\le t\le T}|X^n_t - Y_t| \to 0$ in L^2.
Now we turn to the general case $M \in \mathcal{M}_{c,loc}$. Define $T_n := \inf\big\{t\ge 0 : |M_t|\ge n\big\}$. Then $(T_n)_{n\in\mathbb{N}}$ reduces M and we can apply the bounded case to M^{T_n}, writing $A^n = [M^{T_n}]$. By uniqueness, $A^{n+1}_{t\wedge T_n}$ and A^n_t are indistinguishable, i.e. $\mathbb{P}\big(A^{n+1}_{t\wedge T_n} = A^n_t \text{ for all } t\ge 0\big) = 1$. Thus there exists an increasing process A such that for all $n\in\mathbb{N}$, $A_{t\wedge T_n}$ and A^n_t are indistinguishable. By construction, $(M^2_{t\wedge T_n} - A_{t\wedge T_n})_{t\ge 0} \in \mathcal{M}_c$ and so $(M^2_t - A_t)_{t\ge 0} \in \mathcal{M}_{c,loc}$ and we take $[M]_t = A_t$.
Finally, we have $\quad [M^{T_n}]^m \to [M^{T_n}]$ $u.c.p.$ as $m\to\infty$ for each $n\in\mathbb{N}$. Note that for each $t\ge 0$, $\mathbb{P}(t\le T_n)\to 1$ as $n\to\infty$ to obtain that $\quad [M]^m \to [M] \quad u.c.p. \quad$ as $m\to\infty$. □

DANGER. The following formula is nonsense in a mathematical sense for several reasons, but captures the main idea of the quadratic variation:

$$[M]_t := M^2_t - 2\int_0^t M_s\,dM_s = \underbrace{M^2_t - \int_0^t d(M^2_s)}_{=M_0=0} + \int_0^t (dM_s)^2\,, \tag{2.30}$$

so $[M]_t \approx \int_0^t (dM_s)^2$ which vanishes if M is of finite variation, since then $(dM_s)^2 \approx o(ds)$. But for M being e.g. a standard Brownian motion, $(dM_s)^2 = ds$ and we get $[M]_t = t$ as expected.

Theorem 2.5 *If $M \in \mathcal{M}^2_c$, $M^2 - [M]$ is a UI martingale.*

Proof. Let $S_n = \inf\big\{t\ge 0 : [M]_t \ge n\big\}$. S_n is a stopping time and $[M]_{t\wedge S_n}\le n$. Thus, the local martingale

$$M^2_{t\wedge S_n} - [M]_{t\wedge S_n} \le n + \sup_{t\ge 0} M^2_t \tag{2.31}$$

is bounded by an integrable random variable and thus a true martingale (see remark after Proposition 1.8). Thus

$$\mathbb{E}\big([M]_{t\wedge S_n}\big) = \mathbb{E}(M^2_{t\wedge S_n}) \quad \text{for all } t\ge 0\,. \tag{2.32}$$

We take the limit $t\to\infty$, using monotone convergence on the left and dominated convergence on the right, and then $n\to\infty$ by the same arguments to get

$$\mathbb{E}\big([M]_\infty\big) = \mathbb{E}(M^2_\infty) < \infty\,. \tag{2.33}$$

Hence, $M_t^2 - [M]_t$ is dominated by $\sup_{t\geq 0} M_t^2 + [M]_\infty$ which is integrable. Thus $M^2 - [M]$ is a true martingale and is UI since by Doob's inequality

$$\mathbb{E}\big(\sup_{t\geq 0} |M_t^2 - [M]_t|\big) \leq \mathbb{E}\big((\sup_{t\geq 0} M_t)^2 + [M]_\infty\big) \leq 5\mathbb{E}(M_\infty^2) < \infty \,. \qquad \square \quad (2.34)$$

Remark. Some textbooks use the notation $\langle M \rangle$ rather than $[M]$ for the quadratic variation. In general, $\langle M \rangle$ should be previsible and means something slightly different (beyond the scope of this course), but it coincides with $[M]$ when M is continuous.

2.4 Itô integrals

Given $M \in \mathcal{M}_c^2$, define a measure μ on $\mathcal{P}$ by

$$\mu\big(A \times (s,t]\big) = \mathbb{E}\big(\mathbb{1}_A([M]_t - [M]_s)\big) \quad \text{for all } s < t,\ A \in \mathcal{F}_s \,. \tag{2.35}$$

Since $\mathcal{P}$ is generated by the π-system of events of this form, this uniquely specifies μ. Alternatively, write

$$\mu(d\omega \otimes dt) = \lambda(\omega, dt)\, \mathbb{P}(d\omega) \,, \tag{2.36}$$

where $\lambda(\omega, .)$ is the Lebesgue-Stieltjes measure associated to $[M](\omega)$. Thus, for a previsible process $H \geq 0$,

$$\int_{\Omega\times(0,\infty)} H\, d\mu = \mathbb{E}\Big(\int_0^\infty H_s\, d[M]_s\Big) \,. \tag{2.37}$$

Definition 2.4 Set $L^2(M) = L^2\big(\Omega \times (0,\infty), \mathcal{P}, \mu\big)$ and write

$$\|H\|_M^2 = \|H\|_{L^2(M)}^2 = \mathbb{E}\Big(\int_0^\infty H_s^2\, d[M]_s\Big) \,, \tag{2.38}$$

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

Note that the simple processes $\mathcal{S} \subseteq L^2(M)$ for all $M \in \mathcal{M}_c^2$.

Theorem 2.6 Itô isometry
For every $m \in \mathcal{M}_c^2$ there exists a unique isometry $I : L^2(M) \to \mathcal{M}_c^2$ such that $I(H) = H \cdot M$ for all $H \in \mathcal{S}$.

Proof. Let $H = \sum_{k=0}^{n-1} Z_k\, \mathbb{1}_{(t_k,t_{k+1}]} \in \mathcal{S}$. By Proposition 2.1, $H \cdot M \in \mathcal{M}_c^2$ with

$$\|H \cdot M\|^2 = \sum_{k=0}^{n-1} \mathbb{E}\big(Z_k^2 (M_{t_{k+1}} - M_{t_k})^2\big) \,. \tag{2.39}$$

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

$$\begin{aligned} \mathbb{E}\big(Z_k^2 (M_{t_{k+1}} - M_{t_k})^2\big) &= \mathbb{E}\Big(Z_k^2 \mathbb{E}\big((M_{t_{k+1}} - M_{t_k})^2 \,\big|\, \mathcal{F}_{t_k}\big)\Big) = \\ &= \mathbb{E}\Big(Z_k^2 \mathbb{E}\big(M_{t_{k+1}}^2 - M_{t_k}^2 \,\big|\, \mathcal{F}_{t_k}\big)\Big) = \mathbb{E}\big(Z_k^2 ([M]_{t_{k+1}} - [M]_{t_k})\big) \,, \end{aligned} \tag{2.40}$$

and so $\quad \|H\cdot M\|^2 = \mathbb{E}\Big(\int_0^\infty H_s^2\, d[M]_s\Big) = \|H\|_M^2$.

Now let $H \in L^2(M)$. Then by Proposition 2.2 there exist $H^n \in \mathcal{S}$ for all $n \in \mathbb{N}$ and $H^n \stackrel{L^2(M)}{\longrightarrow} H$ as $n \to \infty$, i.e.

$$\mathbb{E}\Big(\int_0^\infty (H_s^n - H_s)^2\, d[M]_s\Big) \to 0 \quad \text{as } n \to \infty\ . \tag{2.41}$$

Now following the first part

$$\big\|(H^n - H^m)\cdot M\big\| = \|H^n - H^m\|_M \to 0 \quad \text{as } n, m \to \infty\ , \tag{2.42}$$

so $H^n \cdot M$ is a Cauchy sequence in $\mathcal{M}_c^2$. Thus, since $\big(\mathcal{M}_c^2, \|.\|\big)$ is complete this sequence has a limit in $\mathcal{M}_c^2$, which we define to be $I(H)$. Then

$$\|I(H)\| = \lim_{n\to\infty} \|H^n \cdot M\| = \lim_{n\to\infty} \|H^n\|_M = \|H\|_M\ , \tag{2.43}$$

and so $I(H)$ is, indeed, an isometry on $L^2(M)$. For $H \in \mathcal{S}$ we have consistently $I(H) = H\cdot M$ by choosing $H^n \equiv H$. □

Definition 2.5 We write $\quad I(H)_t = (H\cdot M)_t = \int_0^t H_s\, dM_s \quad$ for all $H \in L^2(M)$. The process $H \cdot M$ is *Itô's stochastic integral* of H with respect to M.

Remark. By Theorem 2.6, this is consistent with our previous definition of $H \cdot M$ for $H \in \mathcal{S}$.

Proposition 2.7 *Let $M \in \mathcal{M}_c^2$ and $H \in L^2(M)$. Let T be a stopping time. Then*

$$(H\cdot M)^T = (H\,\mathbb{1}_{(0,T]})\cdot M = H\cdot(M^T)\ . \tag{2.44}$$

Proof. Take $M \in \mathcal{M}_c^2$ and suppose first that $H \in \mathcal{S}$. If T takes only finitely many values, $H\,\mathbb{1}_{(0,T]} \in \mathcal{S}$ and $(H\cdot M)^T = (H\,\mathbb{1}_{(0,T]})\cdot M$ is easily checked. For general T, set $T_n = (2^{-n}\lfloor 2^n T\rfloor)\wedge n$ which takes only finitely many values. Then $T_n \nearrow T$ as $n\to\infty$ and so

$$\big\|H\,\mathbb{1}_{(0,T_n]} - H\,\mathbb{1}_{(0,T]}\big\|_M^2 = \mathbb{E}\bigg(\int_0^\infty H_t^2\,\mathbb{1}_{(T_n,T]}(t)\, d[M]_t\bigg) \to 0 \quad \text{as } n\to\infty \tag{2.45}$$

by dominated convergence, and so $H\,\mathbb{1}_{(0,T_n]}\cdot M \to H\,\mathbb{1}_{(0,T]}\cdot M$ in $\mathcal{M}_c^2$ by Theorem 2.6. But $(H\cdot M)_t^{T_n} \to (H\cdot M)_t^T$ *a.s.* by continuity and hence, $(H\cdot M)^T = (H\,\mathbb{1}_{(0,T]})\cdot M$ since $(H\cdot M)^{T_n} = (H\,\mathbb{1}_{(0,T_n]})\cdot M$ for all $n\in\mathbb{N}$ by the first part. Note that we already know $(H\cdot M)^T = H\dot{(}M^T)$ by Proposition 2.1.

Now for $H \in L^2(M)$ choose $H^n \in \mathcal{S}$ 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,

$$\big\|H^n\,\mathbb{1}_{(0,T]} - H\,\mathbb{1}_{(0,T]}\big\|_M^2 = \mathbb{E}\bigg(\int_0^T (H^n - H)_s^2\, d[M]_s\bigg) \le \|H^n - H\|_M^2 \to 0 \tag{2.46}$$

as $n\to\infty$, so $\quad (H^n\mathbb{1}_{(0,T]})\cdot M \to (H\,\mathbb{1}_{(0,T]})\cdot M$ in $\mathcal{M}_c^2$. Again, by equating the limits of both sequences we get $(H\cdot M)^T = (H\,\mathbb{1}_{(0,T]})\cdot M$. Moreover,

$$\begin{aligned}\|H^n - H\|_{M^T}^2 &= \mathbb{E}\bigg(\int_0^\infty (H^n - H)_s^2\, d[M^T]_s\bigg) = \\ &= \mathbb{E}\bigg(\int_0^T (H^n - H)_s^2\, d[M]_s\bigg) \le \|H^n - H\|_M \to 0\ ,\end{aligned} \tag{2.47}$$

so $H^n \cdot (M^T) \to H \cdot (M^T)$ in $\mathcal{M}_c^2$. Hence, $(H \cdot M)^T = H \cdot (M^T)$. $\square$

Proposition 2.7 allows us to make a final extension of Itô's integral to locally bounded, previsible integrands.

Definition 2.6 Let H be previsible. Say that H is *locally bounded* if there exist stopping times $S_n \nearrow \infty$ $a.s.$ such that $H\,\mathbb{1}_{(0,S_n]}$ is bounded for all $n \in \mathbb{N}$, i.e. $\sup_{t\geq 0} \big|H_t\,\mathbb{1}_{(0,S_n]}(t)\big| < \infty$ $a.s.$. For $M \in \mathcal{M}_{c,loc}$ choose stopping times $S_n' = \inf\big\{t \geq 0 \,:\, |M_t| \geq n\big\} \nearrow \infty$ $a.s.$ such that $M^{S_n'} \in \mathcal{M}_c^2$ for all $n \in \mathbb{N}$. Set $T_n = S_n \wedge S_n'$ and define

$$(H \cdot M)_t := \big((H\,\mathbb{1}_{(0,T_n]} \cdot M^{T_n}\big)_t \quad \text{for all } t \leq T_n\,. \tag{2.48}$$

Note that every bounded previsible process is in $L^2(M)$ whenever $M \in \mathcal{M}_c^2$. Moreover, with $H\,\mathbb{1}_{(0,S_n]}$ also $H\,\mathbb{1}_{(0,T_n]}$ is bounded and $M^{T_n} = (M^{S_n'})^{T_n} \in \mathcal{M}_c^2$. Hence, by Proposition 2.7, (2.48) is well-defined and consistent.

Proposition 2.8 *Let $M \in \mathcal{M}_{c,loc}$ and H, K be locally bounded previsible processes. Let T be a stopping time. Then*

(i) $(H \cdot M)^T = (H\,\mathbb{1}_{(0,T]}) \cdot M = H \cdot (M^T)\,,$ *(ii)* $H \cdot M \in \mathcal{M}_{c,loc}\,,$

(iii) $[H \cdot M] = H^2 \cdot [M]\,,$ *(iv)* $H \cdot (K \cdot M) = (H\,K) \cdot M\,.$

Proof. For each fixed time $t > 0$, (i) follows directly from applying Proposition 2.7 to $H\,\mathbb{1}_{(0,T_n]}$ and M^{T_n} for large enough $n \in \mathbb{N}$. Then

$$(H \cdot M)^{T_n} = (H\,\mathbb{1}_{(0,T_n]}) \cdot M^{T_n} \in \mathcal{M}_c^2 \tag{2.49}$$

which implies (ii).
(2.49) holds for all stopping times T, and we can reduce (iii) and (iv) to the case where $M \in \mathcal{M}_c^2$, and H, K are bounded uniformly in time. The general case then follows via

$$[H \cdot M]^{T_n} = \big[H\,\mathbb{1}_{(0,T_n]} \cdot M^{T_n}\big] = H^2\mathbb{1}_{(0,T_n]} \cdot [M^{T_n}] = \big(H^2 \cdot [M]\big)^{T_n} \tag{2.50}$$

for all times $t \leq T_n$ and thus for all $t > 0$ as $n \to \infty$. This is called a **localization argument**.
(iii) For any stopping time T, we have

$$\begin{aligned}\mathbb{E}\big((H \cdot M)_T^2\big) &= \mathbb{E}\big((H\,\mathbb{1}_{(0,T]} \cdot M)_\infty^2\big) = \qquad \text{(by Itô isometry we get)}\\ &= \mathbb{E}\big((H^2\,\mathbb{1}_{(0,T]} \cdot [M])_\infty\big) = \mathbb{E}\big((H^2 \cdot [M])_T\big)\end{aligned} \tag{2.51}$$

By OST this implies that $(H \cdot M)^2 - H^2 \cdot [M]$ is a martingale and so, by Theorem 2.4, we have $[H \cdot M] \equiv H^2 \cdot [M]$.
(iv) The case $H, K \in \mathcal{S}$ can be checked elementary. For H, K uniformly bounded there exist $H^n, K^n \in \mathcal{S}$, $n \in \mathbb{N}$ such that $H^n \to H$ and $K^n \to K$ in $L^2(M)$. We first prove an upper bound on $\|H\|_{L^2(K\cdot M)}$:

$$\begin{aligned}\|H\|^2_{L^2(K\cdot M)} &= \mathbb{E}\big((H^2 \cdot [K \cdot M])_\infty\big) \overset{\text{by}(iii)}{=} \mathbb{E}\big((H^2 \cdot (K^2 \cdot [M]))_\infty\big) =\\ &\overset{*}{=} \mathbb{E}\big(((H\,K)^2 \cdot [M])_\infty\big) = \|H\,K\|^2_{L^2(M)} \leq\\ &\leq \min\big\{\|H\|^2_\infty \|K\|^2_{L^2(M)}, \|H\|^2_{L^2(M)} \|K\|^2_\infty\big\}\,,\end{aligned} \tag{2.52}$$

where $*$ holds by problem 5 on example sheet 1 since $[M]$ is monotone and thus of finite variation. We have $H^n \cdot (K^n \cdot M) = (H^n K^n) \cdot M$ and using (2.52)

$$\begin{aligned} \big\|H^n \cdot (K^n \cdot M) - H \cdot (K \cdot M)\big\| \leq \big\|(H^n - H) \cdot (K^n \cdot M)\big\| + \\ + \big\|H \cdot ((K^n - K) \cdot M)\big\| = \|H^n - H\|_{L^2(K^n \cdot M)} + \|H\|_{L^2((K^n - K)\cdot M)} \leq \\ \leq \|H^n - H\|_{L^2(M)} \|K^n\|_\infty + \|H\|_\infty \|K^n - K\|_{L^2(M)} \to 0 \quad \text{as } n \to \infty \,. \end{aligned} \tag{2.53}$$

So $H^n \cdot (K^n \cdot M) \to H \cdot (K \cdot M)$ in $\mathcal{M}_c^2$. Similarly, $(H^n K^n) \cdot M \to (H K) \cdot M$ in $\mathcal{M}_c^2$, which implies the result. □

Given a continuous semimartingale $X = X_0 + M + A$ with $M \in \mathcal{M}_{c,loc}$, A a finite variation process and $M_0 = A_0 = 0$. Inspired by the dangerous formula (2.30), we set $[X] := [M]$. This definition is mathematically justified, since then ($\to$ check)

$$\sum_{k=0}^{\lfloor 2^n t \rfloor - 1} (X_{(k+1)2^{-n}} - X_{k2^{-n}})^2 \to [X]_t \quad u.c.p. \text{ as } n \to \infty \,. \tag{2.54}$$

Definition 2.7 For a continuous semimartingale X and H locally bounded and previsible, we define the *stochastic integral*

$$H \cdot X = H \cdot M + H \cdot A \,, \quad \text{writing also} \quad (H \cdot X)_t = \int_0^t H_s \, dX_s \,, \tag{2.55}$$

where $H \cdot M$ is Itô's integral from Definition 2.6 and $H \cdot A$ is the finite variation integral defined in Proposition 1.6.

Note that $H \cdot X$ is already given in Doob-Meyer decomposition and is thus obviously a continuous semimartingale. Under the additional assumption that H is left-continuous, one can show that the Riemann sum approximation to the integral converges.

Proposition 2.9 *Let X be a continuous semimartingale and H be a locally bounded* ***left-continuous*** *adapted process. Then*

$$\sum_{k=0}^{\lfloor 2^n t \rfloor - 1} H_{k2^{-n}} (X_{(k+1)2^{-n}} - X_{k2^{-n}}) \to (H \cdot X)_t \quad u.c.p. \textit{ as } n \to \infty \,. \tag{2.56}$$

Proof. We can treat the finite variation part $X_0 + A$ and the local martingale part M separately. The first is proved in problem 6 on example sheet 1 (in fact, uniformly on compacts everywhere). So it suffices to show that

$$\sum_{k=0}^{\lfloor 2^n t \rfloor - 1} H_{k2^{-n}} (M_{(k+1)2^{-n}} - M_{k2^{-n}}) \to (H \cdot M)_t \quad u.c.p. \text{ as } n \to \infty \tag{2.57}$$

when $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$. By localization, we can reduce to the case where $M \in \mathcal{M}_c^2$ and H_t is bounded uniformly for $t > 0$. Let $H_t^n = H_{2^{-n}\lfloor 2^n t \rfloor}$. Then $H_t^n \to H_t$ as $n \to \infty$ by left continuity. Now,

$$(H^n \cdot M)_t = \sum_{k=0}^{\lfloor 2^n t \rfloor - 1} H_{k2^{-n}} (M_{(k+1)2^{-n}} - M_{k2^{-n}}) + H_{2^{-n}\lfloor 2^n t \rfloor} (M_t - M_{2^{-n}\lfloor 2^n t \rfloor}) \tag{2.58}$$

where $M_t - M_{2^{-n}\lfloor 2^n t \rfloor} \to 0$ as $n \to \infty$ by continuity, so we can ignore the second term on the right. Now

$$\|H^n - H\|_M = \mathbb{E}\bigg(\int_0^\infty (H_t^n - H_t)^2 \, d[M]_t \bigg) \to 0 \quad \text{as } n \to \infty \tag{2.59}$$

by bounded convergence and so $H^n \cdot M \to H \cdot M$ in $\mathcal{M}_c^2$. See problem 14 on example sheet 1 to check that this implies $u.c.p.$ convergence. □

3 Stochastic calculus

In practice we do not calculate integrals from first principles, but rather use tools of calculus such as integration by parts or the chain rule. In this section we derive these tools for stochastic integrals, which differ from ordinary calculus in certain correction terms.

3.1 Covariation

An important process used in this section is the covariation of two local martingales.

Theorem 3.1 Covariation
Let $M, N \in \mathcal{M}_{c,loc}$ and set

$$[M,N]^n_t := \sum_{k=0}^{\lfloor 2^n t \rfloor -1} (M_{(k+1)2^{-n}} - M_{k2^{-n}})(N_{(k+1)2^{-n}} - N_{k2^{-n}}) \,. \tag{3.1}$$

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

(i) $[M,N]^n \to [M,N]$ $u.c.p.$ *as* $n \to \infty$ *,*

(ii) $M\,N - [M,N] \in \mathcal{M}_{c,loc}$ *and* $[M,N]$ *is the unique process with that property ,*

(iii) for $M, N \in \mathcal{M}^2_c$*,* $M\,N - [M,N]$ *is a UI martingale*

(iv) for H *locally bounded and previsible,*

$$[H \cdot M, N] + [M, H \cdot N] = 2H \cdot [M,N] \,. \tag{3.2}$$

We call $[M,N]$ *the covariation of* M *and* N*.*

Remark. Of course, $[M,M] = [M]$. Moreover, $[M,N]$ is a symmetric bilinear form (see problem 12 on example sheet 1), so for example

$$[M_1 + M_2, N_1 + N_2] = [M_1, N_1] + [M_2, N_1] + [M_1, N_2] + [M_2, N_2] \,. \tag{3.3}$$

Proof. Note that $M\,N = \frac{1}{4}\big((M+N)^2 - (M-N)^2\big)$. So we take

$$[M,N] := \frac{1}{4}\big([M+N] - [M-N]\big) \,, \qquad \text{(polarization identity)} \tag{3.4}$$

and show that it has the desired properties. Comparing (2.18) and (3.1) we get

$$[M,N]^n_t = \frac{1}{4}\big([M+N]^n_t - [M-N]^n_t\big) \,, \tag{3.5}$$

so (i)-(iii) follow from Theorems 2.4 and 2.5.
For (iv), $\big[H \cdot (M \pm N)\big] = H^2[M \pm N]$ by Proposition 2.8, thus using (3.4) we get $[H \cdot M, H \cdot N] = H^2 \cdot [M,N]$. Then by bilinearity

$$\begin{aligned}(H+1)^2 \cdot [M,N] &= \big[(H+1) \cdot M, (H+1) \cdot N\big] = \big[H \cdot M + M, H \cdot N + N\big] = \\ &= [H \cdot M, H \cdot N] + [M, H \cdot N] + [H \cdot M, N] + [M,N] \,.\end{aligned} \tag{3.6}$$

Also $(H+1)^2 \cdot [M,N] = H^2 \cdot [M,N] + 2H \cdot [M,N] + [M,N]$. Matching up terms,

$$[M, H \cdot N] + [H \cdot M, N] = 2H \cdot [M,N] \,. \qquad \square \tag{3.7}$$

Proposition 3.2 Kunita-Watanabe Identity
Let $M, N \in \mathcal{M}_{c,loc}$ and H be a locally bounded previsible process. Then

$$[H \cdot M, N] = H \cdot [M, N] \, . \tag{3.8}$$

Proof. By Theorem 3.1(iv) it suffices to show $[H \cdot M, N] = [M, H \cdot N]$. By Theorem 3.1(ii)

$$(H \cdot M)\, N - [H \cdot M, N] \in \mathcal{M}_{c,loc} \quad \text{and} \quad M\,(H \cdot N) - [M, H \cdot N] \in \mathcal{M}_{c,loc} \, . \tag{3.9}$$

Below we show that $(H \cdot M)\, N - M\,(H \cdot N) \in \mathcal{M}_{c,loc}$ which implies $[H \cdot M, N] - [M, H \cdot N] \in \mathcal{M}_{c,loc}$. But this is also of finite variation, since it is a linear combination of quadratic variations, which are increasing. Therefore, by Theorem 1.10, $[H \cdot M, N] \equiv [M, H \cdot N]\ a.s.$ finishing the proof.
By localization, we may assume that $M, N \in \mathcal{M}_c^2$ and H is uniformly bounded in time. By OST, it suffices to show $\mathbb{E}\big((H \cdot M)_T N_T\big) = \mathbb{E}\big(M_T (H \cdot N)_T\big)$ for all bounded stopping times T. By Proposition 2.7 we can replace M by M^T and N by N^T. So it suffices to show

$$\mathbb{E}\big((H \cdot M)_\infty N_\infty\big) = \mathbb{E}\big(M_\infty (H \cdot N)_\infty\big) \quad \text{for all } N, M \in \mathcal{M}_c^2,\ H \text{ bounded} \, . \tag{3.10}$$

Consider $H = Z\, \mathbb{1}_{(s,t]}$ where Z is bounded and $\mathcal{F}_s$-measurable. Then

$$\begin{aligned} \mathbb{E}\big((H \cdot M)_\infty N_\infty\big) &= \mathbb{E}\big(Z\,(M_t - M_s)\, N_\infty\big) = \mathbb{E}\big(Z\,(M_t N_t - M_s N_s)\big) = \\ &= \mathbb{E}\big(M_\infty\, Z\,(N_t - N_s)\big) = \mathbb{E}\big(M_\infty (H \cdot N)_\infty\big) \, . \end{aligned} \tag{3.11}$$

This extends by linearity to all $H \in \mathcal{S}$. We can always find $H^n \in \mathcal{S}$, $n \in \mathbb{N}$ such that $H^n \to H$ simultaneously in $L^2(M)$ and $L^2(N)$, by taking e.g. $\mu = d[M] + d[N]$ in Proposition 2.2. Then

$$(H^n \cdot M)_\infty \to (H \cdot M)_\infty \quad \text{and} \quad (H^n \cdot N)_\infty \to (H \cdot N)_\infty \quad \text{in } L^2 \, , \tag{3.12}$$

so (3.11) extends to all uniformly bounded, previsible H. □

3.2 Itô's formula

Theorem 3.3 Integration by parts
Let X, Y be continuous semimartingales. Then

$$X_t\, Y_t - X_0\, Y_0 = \int_0^t X_s\, dY_s + \int_0^t Y_s\, dX_s + [X, Y]_t \, . \tag{3.13}$$

Proof. Since both sides are continuous in t, it suffices to consider $t = M\, 2^{-N}$ for $M, N \geq 1$. Note that

$$X_t\, Y_t - X_s\, Y_s = X_s(Y_t - Y_s) + Y_s(X_t - X_s) + (X_t - X_s)(Y_t - Y_s) \tag{3.14}$$

so for $n \geq N$

$$\begin{aligned} X_t\, Y_t - X_0\, Y_0 = \sum_{k=0}^{M 2^{n-N} - 1} \Big(& X_{k2^{-n}} (Y_{(k+1)2^{-n}} - Y_{k2^{-n}}) + Y_{k2^{-n}} (X_{(k+1)2^{-n}} - X_{k2^{-n}}) + \\ & + (X_{(k+1)2^{-n}} - X_{k2^{-n}})(Y_{(k+1)2^{-n}} - Y_{k2^{-n}}) \Big) \\ \stackrel{u.c.p.}{\longrightarrow}\ & (X \cdot Y)_t + (Y \cdot X)_t + [X, Y]_t \quad \text{as } n \to \infty \end{aligned} \tag{3.15}$$

by Proposition 2.9 and Theorem 3.1. □

Note the extra covariation term which we do not get in the deterministic case. The next result, Itô's formula, tells us that a smooth function of a continuous semimartingale is again a continuous semimartingale and gives us its precise decomposition in a sort of chain rule.

Theorem 3.4 Itô's formula
Let $X^1, X^2, \ldots, X^d$ be continuous semimartingales and set $X = (X^1, \ldots, X^d)$. Let $f \in C^2(\mathbb{R}^d, \mathbb{R})$. Then

$$f(X_t) = f(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x^i}(X_s)\, dX_s^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s)\, d[X^i, X^j]_s \,. \quad (3.16)$$

Remarks. (i) In particular, $f(X)$ is a continuous semimartingale with decomposition

$$f(X_t) = f(X_0) + \underbrace{\sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x^i}(X_s)\, dM_s^i}_{\in \mathcal{M}_{c,loc}} +$$

$$+ \underbrace{\sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x^i}(X_s)\, dA_s^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s)\, d[M^i, M^j]_s}_{\text{finite variation}} \,. \quad (3.17)$$

where the covariation of the $\mathbb{R}^d$-valued semimartingale $X = X_0 + A + M$ is $[X^i, X^j] = [M^i, M^j]$, due to quadratic variation and the polarization identity (3.4).

(ii) Intuitive proof by Taylor expansion for $d = 1$:

$$\begin{aligned} f(X_t) &= f(X_0) + \sum_{k=0}^{\lfloor 2^n t \rfloor - 1} \big(f(X_{(k+1)2^{-n}}) - f(X_{k2^{-n}}) \big) + \big(f(X_t) - f(X_{\lfloor 2^n t \rfloor 2^{-n}}) \big) = \\ &= f(X_0) + \sum_{k=0}^{\lfloor 2^n t \rfloor - 1} f'(X_{k2^{-n}}) \big(X_{(k+1)2^{-n}} - X_{k2^{-n}} \big) + \\ &\qquad + \frac{1}{2} \sum_{k=0}^{\lfloor 2^n t \rfloor - 1} f''(X_{k2^{-n}}) \big(X_{(k+1)2^{-n}} - X_{k2^{-n}} \big)^2 + \text{ error terms} \\ &\xrightarrow{u.c.p.} f(X_0) + \int_0^t f'(X_s)\, dX_s + \frac{1}{2} \int_0^t f''(X_s)\, d[X]_s \,. \end{aligned} \quad (3.18)$$

We will not follow this method of proof, because the error terms are hard to deal with.

Proof. (for $d = 1$)
Write $X = X_0 + M + A$, where A has total variation process V. Let

$$T_r = \inf \big\{ t \geq 0 \,:\, |X_t| + V_t + [M]_t > r \big\} \,. \quad (3.19)$$

Then $(T_r)_{r \geq 0}$ is a family of stopping times with $T_r \nearrow \infty$. It is sufficient to prove (3.16) on the time intervals $[0, T_r]$. Let $\mathcal{A} \subseteq C^2(\mathbb{R}, \mathbb{R})$ denote the subset of functions f for which the formula holds. Then

(i) $\mathcal{A}$ contains the functions $f(x) \equiv 1$ and $f(x) = x$. (ii) $\mathcal{A}$ is a vector space.

Below we will show that $\mathcal{A}$ is, in fact, an algebra, i.e. in addition

(iii) $f, g \in \mathcal{A} \quad \Rightarrow \quad f\, g \in \mathcal{A}$.

Finally we will show that

(iv) if $f_n \in \mathcal{A}$ and $f_n \to f$ in $C^2(B_r, \mathbb{R})$ for all $r > 0$ then $f \in \mathcal{A}$, where $f_n \to f$ in $C^2(B_r, R)$ means that $\Delta_{n,r} \to 0$ as $n \to \infty$ with $B_r = \{x : |x| \leq r\}$ and

$$\Delta_{n,r} := \max\Big\{ \sup_{x\in B_r} \big|f_n(x)-f(x)\big|, \sup_{x\in B_r} \big|f_n'(x)-f'(x)\big|, \sup_{x\in B_r} \big|f_n''(x)-f''(x)\big| \Big\} . \tag{3.20}$$

(i) -(iii) imply that $\mathcal{A}$ contains all polynomials. By Weierstrass' approximation theorem, these are dense in $C^2(B_r, \mathbb{R})$ and so (iv) implies $\mathcal{A} = C^2(B_r, \mathbb{R})$.
Proof of (iii): Suppose $f, g \in \mathcal{A}$ and set $F_t = f(X_t)$, $G_t = g(X_t)$. Since the formula holds for f and g, F and G are continuous semimartingales. Integration by parts (Theorem 3.3) yields

$$F_t\, G_t - F_0\, G_0 = \int_0^t F_s\, dG_x + \int_0^t G_x\, dF_s + [F, G]_t \, . \tag{3.21}$$

By Proposition 2.8(iv) we have $F \cdot G = F \cdot (1 \cdot G)$, and using Itô's formula for $(1 \cdot G)_s = g(X_s) - g(X_0)$ we get again by Proposition 2.8(iv)

$$\int_0^t F_s\, dG_s = \int_0^t f(X_s)\, g'(X_s)\, dX_s + \frac{1}{2} \int_0^t f(X_s)\, g''(X_s)\, d[X]_s \, . \tag{3.22}$$

By the Kunita-Watanabe identity (Proposition 3.2) we have $[f' \cdot X, G] = f' \cdot [X, G]$. Applying this a second time for G leads to

$$[F, G]_t = \big[f'(X) \cdot X, g'(X) \cdot X\big]_t = \int_0^t f'(X_s)\, g'(X_s)\, d[X]_s \, . \tag{3.23}$$

Substituting these into (3.21), we obtain Itô's formula for $f\, g$.
Proof of (iv): Let $f_n \in \mathcal{A}$ sucht that $f_n \to f$ in $C^2(B_r, \mathbb{R})$. Then

$$\begin{aligned} &\int_0^{t\wedge T_r} \big|f_n'(X_s) - f'(X_s)\big|\, dA_s + \frac{1}{2} \int_0^{t\wedge T_r} \big|f_n''(X_s) - f''(X_s)\big|\, d[M]_s \leq \\ &\quad \leq \Delta_{n,r}\big(V_{t\wedge T_r} + \frac{1}{2}[M]_{t\wedge T_r}\big) \leq r\Delta_{n,r} \to 0 \quad \text{as } n \to \infty \, . \end{aligned} \tag{3.24}$$

and so

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

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

$$\begin{aligned} &\Big\|\big(f_n'(X) \cdot M\big)^{T_r} - \big(f'(X) \cdot M\big)^{T_r}\Big\| = \mathbb{E}\Big(\int_0^{T_r} \big(f_n'(X_s) - f'(X_s)\big)^2\, d[M]_s \Big) \leq \\ &\quad \leq \Delta_{n,r}^2 \mathbb{E}\big([M]_{T_r}\big) \leq r\Delta_{n,r}^2 \to 0 \quad \text{as } n \to \infty \, , \end{aligned} \tag{3.25}$$

and so $\big(f_n'(X)\cdot M\big)^{T_r} \to \big(f'(X)\cdot M\big)^{T_r}$ in $\mathcal{M}_c^2$. For any fixed r, $X^{T_r}\in B_r$ and taking the limit $n\to\infty$ in Itô's formula for f_n we obtain

$$f(X_{t\wedge T_r}) = f(X_0) + \int_0^{t\wedge T_r} f'(X_s)\,dX_s + \frac{1}{2}\int_0^{t\wedge T_r} f''(X_s)\,d[X]_s\ . \tag{3.26}$$

□

Remark. For $d>1$, (i) becomes '$\mathcal{A}$ contains the constant 1 and the coordinate functions $f_1(x)=x^1,\ldots,f_d(x)=x^d$.' Check that you can then follow the same argument, dealing with all the different components $X^i, M^i, [M^i,M^j]$ etc.

Example. Local martingales associated to Brownian motion.
Let $X=B$ be a standard Brownian motion and $f(x)=x^2$. Then by Itô's formula

$$B_t^2 - t = 2\int_0^t B_s\,dB_s \in \mathcal{M}_{c,loc}\ . \tag{3.27}$$

Let $f\in C^{1,2}\big([0,\infty)\times\mathbb{R}^d,\mathbb{R}\big)$ and $X_t=(t,B^1,\ldots,B^d)$ where $B^1,\ldots,B^d$ are independent standard Brownian motions, $B=(B^1,\ldots,B^d)$. Then by Itô's formula,

$$f(t,B_t) - f(t,B_0) - \int_0^t \Big(\frac{1}{2}\Delta + \frac{\partial}{\partial t}\Big) f(s,B_s)\,ds = \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x^i} f(s,B_s)\,dB_s^i \tag{3.28}$$

is a local martingale. See also Proposition 6.4.3 of Advanced Probability.

3.3 Stratonovich integral

Definition 3.1 Let X,Y be continuous semimartingales. Define the *Stratonovich integral*

$$\int_0^t Y_s\,\partial X_s := \int_0^t Y_s\,dX_s + \frac{1}{2}[X,Y]_t\ . \tag{3.29}$$

Remark. As we have discrete approximations for the terms on the right, we have

$$\sum_{k=0}^{\lfloor 2^n t\rfloor -1} \frac{(Y_{(k+1)2^{-n}} + Y_{k2^{-n}})}{2}(X_{(k+1)2^{-n}} - X_{k2^{-n}}) \overset{u.c.p.}{\longrightarrow} \int_0^t Y_s\,\partial X_s\ . \tag{3.30}$$

Proposition 3.5 *Let $X^1,X^2,\ldots,X^d$ be continuous semimartingales, $X=(X^1,\ldots,X^d)$. Let $f\in C^3(\mathbb{R}^d,\mathbb{R})$. Then*

$$f(X_t) = f(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x^i}(X_s)\,\partial X_s^i\ . \tag{3.31}$$

In particular, the integration by parts formula becomes

$$X_t\,Y_t - X_0\,Y_0 = \int_0^t X_s\,\partial Y_s + \int_0^t Y_s\,\partial X_s\ . \tag{3.32}$$

Proof. We treat the $d = 1$ case, the same argument works for $d \geq 2$. By Itô's formula,

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\, dX_s + \frac{1}{2}\int_0^t f''(X_s)\, d[X]_s \tag{3.33}$$

$$f'(X_t) = f'(X_0) + \int_0^t f''(X_s)\, dX_s + \frac{1}{2}\int_0^t f'''(X_s)\, d[X]_s \tag{3.34}$$

By Kunita-Watanabe and (3.34) $\big[f'(X), X\big] = f''(X) \cdot [X]$. But then by (3.33) and the definition of the Stratonovich integral $f(X_t) = f(X_0) + \int_0^t f'(X_s)\, \partial X_s$. □

So the Stratonovich integral has the advantage of being subject to the rules of ordinary calculus. However, the integrand must be a continuous semimartingale (not the case for the Itô integral), and it requires more regularity in the function f. Moreover a Stratonovich integral with respect to a local martingale is not necessarily itself a local martingale.

Example. Let B be a standard Brownian motion and define the discrete approximation

$$B^n = \sum_{k\geq 0} B_{t_k^*} \mathbb{1}_{(t_k, t_{k+1}]} \quad \text{where} \quad t_k = k2^{-n} . \tag{3.35}$$

For $t_k^* = t_k$, $B^n \in \mathcal{S}$ and $\mathbb{E}\big((B^n \cdot B)_t\big) = \sum_{k=0}^{\lfloor 2^n t\rfloor - 1} \mathbb{E}\big(B_{t_k}(B_{t_{k+1}} - B_{t_k})\big) = 0$,
corresponding to the fact that the Itô integral $(B^n \cdot B) \in \mathcal{M}_c^2$.
For $t_k^* > t_k$, $B^n \notin \mathcal{S}$ and in particular for $t_k^* = (t_{k+1} + t_k)/2$ we have

$$\sum_{k=0}^{\lfloor 2^n t\rfloor - 1} \mathbb{E}\Big(\frac{B_{t_{k+1}} + B_{t_k}}{2}(B_{t_{k+1}} - B_{t_k})\Big) = \mathbb{E}\big((B_{t_{k+1}} - B_{t_k})^2/2\big) = t/2 , \tag{3.36}$$

corresponding to the fact that the Stratonovich integral is (by Itô's formula)

$$\int_0^t B_s\, \partial B_s = \int_0^t B_s\, dB_s + \frac{1}{2}t = \frac{1}{2}B_t^2 \notin \mathcal{M}_{loc} . \tag{3.37}$$

The Itô and Stratonovich integrals are both useful in different situations (see pp. 35-37 in Øksendal for a nice discussion). Since we know how to convert back and forth between them, we will deal only with the Itô integral for the rest of the course.

3.4 Notation and Summary

We agree that

$$\begin{aligned} dZ_t = H_t\, dX_t &\quad \text{means} \quad Z_t - Z_0 = \int_0^t H_s\, dX_s \\ \partial Z_t = H_t\, dX_t &\quad \text{means} \quad Z_t - Z_0 = \int_0^t H_s\, \partial X_s \\ dZ_t = dX_t\, dY_t &\quad \text{means} \quad Z_t - Z_0 = [X, Y]_t . \end{aligned} \tag{3.38}$$

Note also that $H_t\,dX_t = d(H\cdot X)_t$ and $dX_t\,dY_t = d[X,Y]_t$.
We have shown:

$$\begin{aligned}
H_t(K_t\,dX_t) &= (H_t\,K_t)\,dX_t && \text{(Prop. 2.8(iv))}\\
H_t(dX_t\,dY_t) &= (H_t\,dX_t)\,dY_t && \text{(Kunita-Watanabe)}\\
d(X_t\,Y_t) &= Y_t\,dX_t + X_t\,dY_t + dX_t\,dY_t && \text{(integration by parts)}\\
d\big(f(X_t)\big) &= \frac{\partial f}{\partial x^i}(X_t)\,dX_t^i + \frac{1}{2}\frac{\partial^2 f}{\partial x^i \partial x^j}(X_t)\,dX_t^i\,dX_t^j && \text{(Itô's formula)}\\
&= Df(X_t)\,dX_t + \frac{1}{2}D^2 f(X_t)\,dX_t \otimes dX_t\ , && (3.39)
\end{aligned}$$

where we use the sum convention in Itô's formula. The relationship between the Itô and Stratonovich integrals is expressed as

$$\partial Z_t = Y_t\,\partial X_t \quad \Leftrightarrow \quad dZ_t = Y_t\,dX_t + \frac{1}{2}dX_t\,dY_t\ . \tag{3.40}$$

4 Applications of Stochastic Calculus

4.1 Brownian motion

We 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 4.1 Lévy's characterization of Brownian motion
Let $X^1, \ldots, X^d \in \mathcal{M}_{c,loc}$ with $[X^i, X^j]_t = \delta_{ij} t$. Then $X = (X^1, \ldots, X^d)$ is a Brownian motion in $\mathbb{R}^d$.

Proof. It suffices to show that, for $0 \leq s \leq t$, $X_t - X_s \sim N\big(0, (t-s)I\big)$ and the increment is independent of $\mathcal{F}_s$. By uniqueness of characteristic functions, this is equivalent to showing that for all $s \leq t$

$$\mathbb{E}\big(\exp(i\langle\theta, X_t - X_s\rangle)\,\big|\,\mathcal{F}_s\big) = \exp\big(-\tfrac{1}{2}|\theta|^2(t-s)\big) \quad \text{for all } \theta \in \mathbb{R}^d, s \leq t\,. \tag{4.1}$$

Fix θ and set $Y_t = \langle\theta, X_t\rangle = \theta_1 X_t^1 + \ldots + \theta_n X_t^d$.
Then $Y \in \mathcal{M}_{c,loc}$ and $[Y]_t = |\theta|^2 t$. Consider

$$Z_t = \exp\big(iY_t + \tfrac{1}{2}[Y]_t\big) = \exp\big(i\langle\theta, X_t\rangle + \tfrac{1}{2}|\theta|^2 t\big)\,. \tag{4.2}$$

By Itô's formula,

$$dZ_t = Z_t\big(i\,dY_t + \tfrac{1}{2}d[Y]_t\big) + \tfrac{1}{2}Z_t i^2\,d[Y]_t = i\,Z_t\,dY_t\,. \tag{4.3}$$

So Z is a local martingale. Moreover, Z is bounded on $[0,t]$ for all $t \geq 0$ and so is a true martingale by Proposition 1.8. Hence, $\mathbb{E}(Z_t|\mathcal{F}_s) = Z_s$, which implies (4.1). □

Proposition 4.2 *Let B be a Brownian motion and let h be a deterministic function in $L^2(\mathbb{R}^+, \mathcal{B}, \lambda)$ with Lebesgue measure λ. Set $X = \int_0^\infty h_s\,dB_s$. Then $X \sim N\big(0, \|h\|_2^2\big)$.*

Proof. Let $M_t = \int_0^t h_s\,dB_s$. Then $M \in \mathcal{M}_{c,loc}$ and (since $[h \cdot B] = h^2 \cdot [B]$) we have $[M]_t = \int_0^t h_s^2\,ds$. Now let $Z_t = \exp\big(iuM_t + \frac{1}{2}u^2[M]_t\big)$. This is in $\mathcal{M}_{loc}$ as in the proof of Lévy's characterization and, as it is uniformly bounded by $\exp\big(\frac{1}{2}u^2\|h\|_2^2\big)$, is in fact in $\mathcal{M}^2$. Hence,

$$1 = \mathbb{E}(Z_0) = \mathbb{E}(Z_\infty) = \mathbb{E}\big(\exp(iuX)\big)\,\exp\big(\tfrac{1}{2}u^2\|h\|_2^2\big)\,. \tag{4.4}$$

□

Theorem 4.3 Dubins-Schwarz Theorem
Let $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$ and $[M]_\infty = \infty$ a.s.. Set $\tau_s = \inf\big\{t \geq 0 : [M]_t > s\big\}$, $B_s = M_{\tau_s}$ and $\mathcal{G}_s = \mathcal{F}_{\tau_s}$. Then τ_s is an $(\mathcal{F}_t)_{t\geq 0}$-stopping time and $[M]_{\tau_s} = s$ for all $s > 0$. Moreover, B is a $(\mathcal{G}_t)_{t\geq 0}$-Brownian motion and $M_t = B_{[M]_t}$.

Remark. So any continuous local martingale is a (stochastic) time-change of Brownian motion. In this sense, Brownian motion is the most general continuous local martingale.

Proof. Since $[M]$ is continuous and adapted, τ_s is a stopping time, and $[M]_\infty = \infty$ ensures

that $\tau_s < \infty$ $a.s.$ for all $s \geq 0$. Then B is adapted to $(\mathcal{G}_t)_{t\geq 0}$ (by Proposition 4.1.1 of Advanced Probability).
We show first that B is continuous: $s \mapsto \tau_s$ is càdlàg and increasing and thus B is right-continuous. So it remains to show $B_{s-} = B_s$, i.e. $M_{\tau_{s-}} = M_{\tau_s}$, where

$$\tau_{s-} = \inf\big\{t \geq 0 \,:\, [M]_t = s\big\} \tag{4.5}$$

So we need to show that M is constant between τ_{s-} and τ_s whenever $\tau_{s-} < \tau_s$, i.e. whenever $[M]$ is constant. By localization, we may assume $M \in \mathcal{M}_c^2$.
But $\mathbb{E}\big(M_{\tau_s}^2 - [M]_{\tau_s} \,\big|\, \mathcal{F}_{\tau_{s-}}\big) = M_{\tau_{s-}}^2 - [M]_{\tau_{s-}}$. Since $[M]_{\tau_s} = [M]_{\tau_{s-}}$ and M is a martingale, we obtain $\mathbb{E}\big(M_{\tau_s}^2 - M_{\tau_{s-}}^2 \,\big|\, \mathcal{F}_{\tau_{s-}}\big) = \mathbb{E}\big((M_{\tau_s} - M_{\tau_{s-}})^2 \,\big|\, \mathcal{F}_{\tau_{s-}}\big) = 0$, and so M is $a.s.$ constant between τ_{s-} and τ_s.
Now we show that B is a Brownian motion: Fix $s > 0$. Then $[M^{\tau_s}]_\infty = [M]_{\tau_s} = s$, and so by example sheet 1, problem 11, $M^{\tau_s} \in \mathcal{M}_c^2$ since $\mathbb{E}\big([M^{\tau_s}]_\infty\big) < \infty$. But then by Theorem 2.5, $\big(M^2 - [M]\big)^{\tau_s}$ is a UI martingale. By OST, for $r \leq s$,

$$\begin{aligned} \mathbb{E}(B_s|\mathcal{G}_r) &= \mathbb{E}(M_\infty^{\tau_s}|\mathcal{F}_{\tau_r}) = M_{\tau_r} = B_r \\ \mathbb{E}(B_s^2 - s|\mathcal{G}_r) &= \mathbb{E}\big((M^2 - [M])_\infty^{\tau_s} \,\big|\, \mathcal{F}_{\tau_r}\big) = M_{\tau_r}^2 - [M]_{\tau_r} = B_r^2 - r\,. \end{aligned} \tag{4.6}$$

Hence, $B \in \mathcal{M}_c$ with $[B]_s = s$ and so, by Lévy's characterization, B is a $(\mathcal{G}_t)_{t\geq 0}$-Brownian motion. □

4.2 Exponential martingales

Definition 4.1 Let $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$. Set $Z_t = \exp\big(M_t - \frac{1}{2}[M]_t\big)$. By Itô's formula

$$dZ_t = Z_t\big(dM_t - \tfrac{1}{2}d[M]_t\big) + \tfrac{1}{2}Z_t\, d[M]_t = Z_t\, dM_t\,, \tag{4.7}$$

and so $Z \in \mathcal{M}_{c,loc}$. We call Z the *stochastic exponential* of M and write $Z = \mathcal{E}(M)$.

Proposition 4.4 Exponential martingale inequality
Let $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$. Then for all $\epsilon, \delta > 0$,

$$\mathbb{P}\Big(\sup_{t\geq 0} M_t \geq \epsilon\,,\ [M]_\infty \leq \delta\Big) \leq e^{-\epsilon^2/(2\delta)}\,. \tag{4.8}$$

Proof. Fix $\epsilon > 0$ and set $T = \inf\{t \geq 0 : M_t \geq \epsilon\}$. Fix $\theta \in \mathbb{R}$ and set

$$Z_t = \exp\big(\theta M_t^T - \tfrac{1}{2}\theta^2[M]_t^T\big)\,. \tag{4.9}$$

Then $Z \in \mathcal{M}_{c,loc}$ and $|Z| \leq e^{\theta\epsilon}$. Hence, $Z \in \mathcal{M}_c^2$ and, by OST, $\mathbb{E}(Z_\infty) = \mathbb{E}(Z_0) = 1$. For $\delta > 0$ we get by the Markov inequality

$$\mathbb{P}\big(\sup_{t\geq 0} M_t \geq \epsilon\,,\ [M]_\infty \leq \delta\big) \leq \mathbb{P}\big(Z_\infty \geq e^{\theta\epsilon - \frac{1}{2}\theta^2\delta}\big) \leq e^{-\theta\epsilon + \frac{1}{2}\theta^2\delta}\,. \tag{4.10}$$

Optimizing in θ gives $\theta = \epsilon/\delta$ and the result follows. □

Proposition 4.5 *Let $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$ and suppose that $[M]$ is $a.s.$ uniformly bounded. Then $\mathcal{E}(M)$ is a UI martingale.*

Proof. Let C be such that $[M]_\infty \leq C\ a.s.$. By the exponential martingale inequality

$$\mathbb{P}\Big(\sup_{t\geq 0} M_t \geq \epsilon\Big) = \mathbb{P}\Big(\sup_{t\geq 0} M_t,\ [M]_\infty \leq C\Big) \leq e^{-\epsilon^2/(2C)}\ . \tag{4.11}$$

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

$$\mathbb{E}\Big(\exp\big(\sup_{t\geq 0} M_t\big)\Big) = \int_0^\infty \mathbb{P}\Big(\sup_{t\geq 0} M_t \geq \log\lambda\Big)\, d\lambda \leq 1 + \int_1^\infty e^{-(\log\lambda)^2/(2C)} < \infty\ . \tag{4.12}$$

Hence, $\mathcal{E}(M)$ is UI and, by Proposition 1.8, $\mathcal{E}(M)$ is a martingale. □

The next result is of fundamental importance in mathematical finance. Remember that for two probability measures $\mathbb{P}_1, \mathbb{P}_2$ on a measurable space $(\Omega, \mathcal{F})$, $\mathbb{P}_1$ is absolutely continuous with respect to $\mathbb{P}_2$, $\mathbb{P}_1 \ll \mathbb{P}_2$, if

$$\mathbb{P}_2(A) = 0 \quad \Rightarrow \quad \mathbb{P}_1(A) = 0 \quad \text{for all } A \in \mathcal{F}\ . \tag{4.13}$$

In this case there exists a $\mathbb{P}_2$-unique measurable density $f : \Omega \to [0, \infty)$ such that $\mathbb{P}_1 = f\,\mathbb{P}_2$. f is also called the Radon-Nikodym derivative.

Theorem 4.6 Girsanov's theorem
Let $M \in \mathcal{M}_{c,loc}$ with $M_0 = 0$. Suppose that $Z = \mathcal{E}(M)$ is a UI martingale. Then we can define a new probability measure $\tilde{\mathbb{P}} \ll \mathbb{P}$ on $(\Omega, \mathcal{F})$ by

$$\tilde{\mathcal{P}}(A) = \mathbb{E}(Z_\infty 1_A)\ , \quad A \in \mathcal{F}\ . \tag{4.14}$$

Moreover, if $X \in \mathcal{M}_{c,loc}(\mathbb{P})$ then $X - [X, M] \in \mathcal{M}_{c,loc}(\tilde{\mathbb{P}})$.

Proof. Since Z is UI, Z_∞ exists $a.s.$, $Z_\infty \geq 0$ and $\mathbb{E}(Z_\infty) = \mathbb{E}(Z_0) = 1$. Thus $\tilde{\mathbb{P}}(\Omega) = 1$, $\tilde{\mathbb{P}}(\emptyset) = 0$ and countable additivity follows by linearity of the expected value and bounded convergence. $\tilde{\mathbb{P}}(A) = \int_A Z_\infty d\mathbb{P} = 0$ if $\mathbb{P}(A) = 0$ so $\tilde{\mathbb{P}} \ll \mathbb{P}$.
Let $X \in \mathcal{M}_{c,loc}$ and set

$$T_n = \inf\big\{t \geq 0\ :\ \big|X_t - [X, M]_t\big| \geq n\big\}\ . \tag{4.15}$$

Since $X - [X, M]$ is continuous, $\mathbb{P}(T_n \nearrow \infty) = 1$ with implies $\tilde{\mathbb{P}}(T_n \nearrow \infty) = 1$. So to show that $Y = X - [X, M] \in \mathcal{M}_{c,loc}(\tilde{\mathbb{P}})$, it suffices to show that

$$Y^{T_n} = X^{T_n} - [X^{T_n}, M] \in \mathcal{M}_c(\tilde{\mathbb{P}}) \quad \text{for all } n \in \mathbb{N}\ . \tag{4.16}$$

Replacing X by X^{T_n}, we reduce to the case where Y is uniformly bounded. By the integration by parts formula

$$\begin{aligned} d(Z_t\, Y_t) &= dZ_t\, Y_t + Z_t\, dY_t + dZ_t\, dY_t = \\ &= Z_t\, dM_t\big(X_t - [X, M]_t\big) + Z_t(dX_t - dX_t\, dM_t) + Z_t\, dM_t\, dX_t \end{aligned} \tag{4.17}$$

and so $Z\,Y \in \mathcal{M}_{c,loc}(\mathbb{P})$. Also $\big\{Z_T\ :\ T \text{ is a stopping time}\big\}$ is UI.
So since Y is bounded, $\big\{Z_T\, Y_T\ :\ T \text{ is a stopping time}\big\}$ is UI. Hence, $Z\,Y \in \mathcal{M}_c(\mathbb{P})$. But then for $s \leq t$,

$$\tilde{\mathbb{E}}(Y_t - Y_s \,|\, \mathcal{F}_s) = \mathbb{E}\big(Z_\infty(Y_t - Y_s)\,\big|\,\mathcal{F}_s\big) = \mathbb{E}\big(Z_t\, Y_t - Z_s\, Y_s\,\big|\,\mathcal{F}_s\big) = 0 \tag{4.18}$$

since $Z\,Y \in \mathcal{M}_c(\mathbb{P})$. So $Y \in \mathcal{M}_c(\tilde{\mathbb{P}})$ as required. □

Remark. The quadratic variation $[Y]$ is the same under $\tilde{\mathbb{P}}$ as it is under $\mathbb{P}$ (see example sheet 2, problem 1).

Corollary 4.7 *Let B be a standard Brownian motion under $\mathbb{P}$ and $M \in \mathcal{M}_{c,loc}$ such that $M_0 = 0$. Suppose $Z = \mathcal{E}(M)$ is a UI martingale and $\tilde{\mathbb{P}}(A) = \mathbb{E}(Z_\infty 1_A)$ for all $A \in \mathcal{F}$. Then $\tilde{B} := B - [B, M]$ is a $\tilde{P}$-Brownian motion.*

Proof. Since $\tilde{B} \in \mathcal{M}_{c,loc}(\tilde{\mathbb{P}})$ by Theorem 4.6 and has $[\tilde{B}]_t = [B]_t = t$, by Lévy's characterization, it is a Brownian motion. □

Let $(W, \mathcal{W}, \mu)$ be **Wiener space**. Recall that $W = C\big([0,\infty), \mathbb{R}\big)$, $\mathcal{W} = \sigma(X_t : t \geq 0)$ where $X_t : W \to \mathbb{R}$ with $X_t(w) = w(t)$. The Wiener measure μ is the unique probability measure on $(W, \mathcal{W})$ such that $(X_t)_{t\geq 0}$ is a Brownian motion started from 0.

Definition 4.2 Define the *Cameron-Martin space*

$$H = \Big\{ h \in W \,:\, h(t) = \int_0^t \phi(s)\, ds \text{ for some } \phi \in L^2\big([0,\infty)\big) \Big\} . \tag{4.19}$$

For $h \in H$, write $\dot{h}$ for a version of the weak derivative ϕ.

Theorem 4.8 Cameron-Martin theorem
Fix $h \in H$ and set

$$\mu^h(A) = \mu\big(\{w \in W \,:\, w + h \in A\}\big) \quad \textit{for all } A \in \mathcal{W} . \tag{4.20}$$

Then μ^h is a probability measure on $(W, \mathcal{W})$ and $\mu^h \ll \mu$ with density

$$\frac{d\mu^h}{d\mu}(w) = \exp\bigg(\int_0^\infty \dot{h}(s)\, dw(s) - \frac{1}{2} \int_0^\infty |\dot{h}(s)|^2\, ds \bigg) \quad \textit{for } \mu - a.a.\ w . \tag{4.21}$$

Remark. So 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 with respect to that of the original Brownian motion.

Proof. Set $\mathcal{W}_t = \sigma(X_s, s \leq t)$ and $M_t = \int_0^t \dot{h}(s)\, dX_s$. Then $M \in \mathcal{M}_c^2\big(W, \mathcal{W}, (\mathcal{W}_t)_{t\geq 0}, \mu\big)$ and

$$[M]_\infty = \int_0^\infty |\dot{h}(s)|^2\, ds =: C < \infty . \tag{4.22}$$

By Proposition 4.5, $\mathcal{E}(M)$ is a UI martingale, so we can define a new probability measure $\tilde{\mu} \ll \mu$ on $(W, \mathcal{W})$ by

$$\frac{d\tilde{\mu}}{d\mu}(w) = \exp\Big(M_\infty(w) - \frac{1}{2}[M]_\infty(w)\Big) = \exp\bigg(\int_0^\infty \dot{h}(s) dw(s) - \frac{1}{2}\int_0^\infty |\dot{h}(s)|^2 ds\bigg) \tag{4.23}$$

and $\tilde{X} = X - [X, M] \in \mathcal{M}_{c,loc}(\tilde{\mu})$ by Girsanov's theorem. Since X is a μ-Brownian motion, by Corollary 4.7, $\tilde{X}$ is a $\tilde{\mu}$-Brownian motion. Thus

$$\mu^h(A) = \mu\big(\{w : X(w) + h \in A\}\big) = \tilde{\mu}\big(\{w : \tilde{X}(w) + h \in A\}\big) . \tag{4.24}$$

But $[X, M]_t = \int_0^t \dot{h}(s)\, ds = h(t)$ and so $\tilde{X}(w) = X(w) - h = w - h$. Hence, for $A \in \mathcal{W}$,

$$\mu^h(A) = \tilde{\mu}\big(\{w : w \in A\}\big) = \tilde{\mu}(A) . \tag{4.25}$$

Hence, $\mu^h = \tilde{\mu}$ as required. □.

5 Stochastic Differential Equations

Suppose we have a differential equation, say $\frac{dx(t)}{dt} = b\big(x(t)\big)$, or, in integral form,

$$x(t) = x(0) + \int_0^t b\big(x(s)\big)\, ds \,, \tag{5.1}$$

which describes a system evolving in time, be it the growth of a population, the trajectory of a moving object or the price of an asset. Taking into account random perturbations, it may be more realistic to add a noise term:

$$X_t = X_0 + \int_0^t b(X_s)\, ds + \sigma B_t \,, \tag{5.2}$$

where B is a Brownian motion and σ is a constant controlling the intensity of the noise. It may be that this intensity depends on the state of the system, in which case we have to consider an equation of the form

$$X_t = X_0 + \int_0^t b(X_s)\, ds + \int_0^t \sigma(X_s)\, dB_s \,, \tag{5.3}$$

where the last term is, of course, an Itô integral. (5.3) is a *stochastic differential equation* and may also be written

$$dX_t = b(X_t)\, dt + \sigma(X_t)\, dB_t \,. \tag{5.4}$$

5.1 General definitions

Suppose $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are measurable functions, bounded on compact sets with $\sigma(x) = \big(\sigma_{ij}(x)\big)_{\substack{1\le i\le d\\ 1\le j\le m}}$ and $b(x) = \big(b_i(x)\big)_{1\le i\le d}$. Consider the equation

$$dX_t = \sigma(X_t)\, dB_t + b(X_t)\, dt \,, \tag{5.5}$$

which may be written componentwise as

$$dX_t^i = \sum_{j=1}^m \sigma_{ij}(X_t)\, dB_t^j + b_i(X_t)\, dt \,, \quad 1 \le i \le d \,. \tag{5.6}$$

A *solution* to (5.5) consists of

- a filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}, \mathbb{P}\big)$ satisfying the usual conditions;
- an $(\mathcal{F}_t)_{t\ge0}$-Brownian motion $B = (B^1, \dots, B^m)$ taking values in $\mathbb{R}^m$;
- an $(\mathcal{F}_t)_{t\ge0}$-adapted continuous process $X = (X^1, \dots, X^d) \in \mathbb{R}^d$ such that

$$X_t = X_0 + \int_0^t \sigma(X_s)\, dB_s + \int_0^t b(X_s)\, ds \,. \tag{5.7}$$

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 SDE's.

Definition 5.1 We say that a SDE has a *weak solution* if for all $x \in \mathbb{R}^d$, there exists a solution to the SDE 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 $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\big)$ and B then any two solutions X and X' satisfying $X_0 = X_0'$ $a.s.$ are such that $\mathbb{P}(X_t = X_t' \text{ for all } t) = 1$. We say that a solution X of a SDE started from x is a *strong solution* if X is adapted to the natural filtration of B.

Remark. In general, $\sigma(B_s, s \leq t) \subseteq \mathcal{F}_t$ and a weak solution might not be measurable with respect to the Brownian motion B. A strong solution depends only on $x \in \mathbb{R}^d$ and the Brownian motion B.

Example. It is possible to have existence of a weak solution and uniqueness in law without pathwise uniqueness. Suppose β is a Brownian motion in $\mathbb{R}$ with $\beta_0 = x$. Set

$$B_t = \int_0^t \operatorname{sgn}(\beta_s)\, d\beta_s \quad \text{where} \quad \operatorname{sgn}(x) = \begin{cases} -1 & \text{if } x \leq 0 \\ 1 & \text{if } x > 0 \end{cases} . \tag{5.8}$$

In problem 3 on example sheet 1, we showed that $\big(\operatorname{sgn}(\beta_t)\big)_{t\geq 0}$ is previsible, so that the Itô integral is well defined and $B \in \mathcal{M}_{c,loc}$. By Lévy's characterization, B is a Brownian motion started from 0, since $[B]_t = [\beta]_t = t$. It is also true that

$$\beta_t = x + \int_0^t \operatorname{sgn}(\beta_s)\, dB_s\ . \tag{5.9}$$

Hence, β is a solution to the SDE

$$dX_t = \operatorname{sgn}(X_t)\, dB_t\ , \quad X_0 = x\ . \tag{5.10}$$

Thus (5.10) has a weak solution. Applying Lévy's characterization again, it is clear that any solution must be a Brownian motion and so there is uniqueness in law. On the other hand, pathwise uniqueness does not hold: Suppose that $\beta_0 = 0$. Then both β and $-\beta$ are solutions to (5.10) started from 0.
It also turns out that β is not a strong solution to (5.10) – see example sheet 3 for a proof.

5.2 Lipschitz coefficients

For $U \subseteq \mathbb{R}^d$ and $f : U \to \mathbb{R}^d$, say f is *Lipschitz* with *Lipschitz constant* $K < \infty$ if

$$\big|f(x) - f(y)\big| \leq K|x - y| \quad \text{for all } x, y \in U\ , \tag{5.11}$$

where $|.|$ denotes the Euclidean norm on $\mathbb{R}^d$. For $f : U \to \mathbb{R}^{d\times m}$ we use the norm $|f| = \Big(\sum_{i=1}^d \sum_{j=1}^m f_{ij}^2\Big)^{1/2}$ in the definition of Lipschitz. We will consider the existence and uniqueness of solutions to SDEs with Lipschitz coefficients.

We write $\mathcal{C}_T$ for the set of continuous adapted processes $X : [0, T] \to \mathbb{R}$ such that

$$|||\, X|||_T := \big\| \sup_{t\leq T} |X_t| \big\|_2 < \infty \tag{5.12}$$

and $\mathcal{C}$ for the set of continuous adapted processes $X : [0, \infty) \to \mathbb{R}$ such that

$$|||\, X|||_T < \infty \quad \text{for all } T > 0\ . \tag{5.13}$$

Recall from Proposition 2.3 that $\big(\mathcal{C}_T, |||.\, |||_T\big)$ is complete. We will make use of two theorems:

Theorem 5.1 Contraction Mapping Theorem
Let (X,d) be a complete metric space and $F: X \to X$. Suppose that the iterated function F^n is a contraction for some $n \in \mathbb{N}$, i.e.

$$\exists r < 1 \,\forall x, y \in X \;:\; d\big(F^n(x), F^n(y)\big) \le r\, d(x,y)\ . \tag{5.14}$$

Then F has a unique fixed point.

Lemma 5.2 Gronwall's Lemma
Let $T > 0$ and let f be a non-negative bounded measurable function on $[0,T]$. Suppose that

$$\exists a, b \ge 0 \,\forall t \in [0,T] \;:\; f(t) \le a + b \int_0^t f(s)\, ds\ . \tag{5.15}$$

Then $f(t) \le a \exp(b\,t)$ for all $t \in [0,T]$.

Theorem 5.3 *Suppose that $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are Lipschitz. Then there is pathwise uniqueness for the SDE*

$$dX_t = \sigma(X_t)\, dB_t + b(X_t)\, dt\ . \tag{5.16}$$

Moreover, for each filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P}\big)$ and each $(\mathcal{F}_t)_{t\ge 0}$-Brownian motion B, there exists a strong solution to the SDE started from x, for any $x \in \mathbb{R}^d$.

Proof. (for $d = m = 1$)
Fix $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P}\big)$ and B. Let $(\mathcal{F}_t^B)_{t\ge 0}$ be the natural filtration generated by B so that $\mathcal{F}_t^B \subseteq \mathcal{F}_t$ for all $t \ge 0$. Suppose that K is the Lipschitz constant for σ and b.
Uniqueness:
Suppose X and X' are two solutions on $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P}\big)$ with $X_0 = X'_0$ $a.s.$. Fix M and let

$$\tau = \inf\big\{t \ge 0 \;:\; |X_t| \vee |X'_t| \ge M\big\}\ . \tag{5.17}$$

Then $\quad X_{t\wedge\tau} = X_0 + \int_0^{t\wedge\tau} \sigma(X_s)\, dB_s + \int_0^{t\wedge\tau} b(X_s)\, ds\ ,$ and similarly for X'.
Let $T > 0$. If $0 \le t \le T$ then, since $(x+y)^2 \le 2(x^2+y^2)$,

$$\begin{aligned}
&\mathbb{E}\big((X_{t\wedge\tau} - X'_{t\wedge\tau})^2\big) \le \\
&\quad \le 2\mathbb{E}\bigg(\Big(\int_0^{t\wedge\tau} \big(\sigma(X_s) - \sigma(X'_s)\big)\, dB_s\Big)^2\bigg) + 2\mathbb{E}\bigg(\Big(\int_0^{t\wedge\tau} \big(b(X_s) - b(X'_s)\big)\, ds\Big)^2\bigg) \\
&\quad \le 2\mathbb{E}\bigg(\int_0^{t\wedge\tau} \big(\sigma(X_s) - \sigma(X'_s)\big)^2 ds\bigg) + 2T\mathbb{E}\bigg(\int_0^{t\wedge\tau} \big(b(X_s) - b(X'_s)\big)^2 ds\bigg) \\
&\qquad\quad \text{(by the Itô isometry and the Cauchy-Schwarz inequality)} \\
&\quad \le 2K^2(1+T)\,\mathbb{E}\bigg(\int_0^{t\wedge\tau} (X_s - X'_s)^2\, ds\bigg) \qquad \text{(by the Lipschitz property)} \\
&\quad \le 2K^2(1+T) \int_0^t \mathbb{E}\big((X_{s\wedge\tau} - X'_{s\wedge\tau})^2\big)\, ds\ .
\end{aligned} \tag{5.18}$$

Let $f(t) = \mathbb{E}\big((X_{t\wedge\tau} - X'_{t\wedge\tau})^2\big)$. Then $f(t)$ is bounded by $4M^2$ and

$$f(t) \le 2K^2(1+T) \int_0^t f(s)\, ds\ . \tag{5.19}$$

Hence, by Gronwall's lemma, $f(t) = 0$ for all $t \in [0,T]$. So $X_{t\wedge\tau} = X'_{t\wedge\tau}$ $a.s.$ and, letting $M, T \to \infty$, we obtain $X = X'$ $a.s.$.
Existence of a strong solution:
Note that for all $y \in \mathbb{R}$,

$$\big|\sigma(y)\big| \le \big|\sigma(0)\big| + k|y| \; , \quad \big|b(y)\big| \le \big|b(0)\big| + k|y| \; . \tag{5.20}$$

Suppose $X \in \mathcal{C}_T$ for some T. Let $M_t = \int_0^t \sigma(X_s)\, dB_s$, $0 \le t \le T$. Then $[M]_T = \int_0^T \sigma(X_s)^2\, ds$ and so by (5.20)

$$\mathbb{E}\big([M]_T\big) \le 2T\big(|\sigma(0)|^2 + K^2 \;|||\; X \;|||_T\;\big) < \infty \; . \tag{5.21}$$

Hence, by example sheet 1, problem 11, $(M_t)_{0\le t\le T}$ is a martingale bounded in L^2. So by Doob's L^2 inequality and (5.20)

$$\mathbb{E}\Big(\sup_{t\le T}\Big|\int_0^t \sigma(X_s)\, dB_s\Big|^2\Big) \le 8T\Big(\big|\sigma(0)\big|^2 + K^2 \;|||\; X \;|||_T^2\;\Big) < \infty \; . \tag{5.22}$$

By (5.20) and the Cauchy-Schwarz inequality,

$$\mathbb{E}\Big(\sup_{t\le T}\Big|\int_0^t b(X_s)\, ds\Big|^2\Big) \le T\mathbb{E}\Big(\int_0^T \big|b(X_s)|^2\, ds\Big) \le$$
$$\le 2T\Big(\big|b(0)\big|^2 + K^2 \;|||\; X \;|||_T^2\;\Big) < \infty \; . \tag{5.23}$$

So we can define $F : \mathcal{C}_T \to \mathcal{C}_T$ by

$$F(X)_t = x + \int_0^t \sigma(X_s)\, dB_s + \int_0^t b(X_s)\, ds \quad \text{for all } t \le T \; . \tag{5.24}$$

Note that for $X, Y \in \mathcal{C}_T$, by similar arguments to those above

$$|||\; F(X) - F(Y) |||_T^2 \le \underbrace{2K^2(4+T)}_{C_T} \int_0^T |||X - Y\;|||_t^2\; dt \; . \tag{5.25}$$

By induction using (5.25), we have for all $n \ge 0$ that

$$|||\; F^n(X) - F^n(Y) |||_T^2 \le C_T^n \;|||\; X - Y\;|||_T^2 \int_0^T \int_0^{t_{n-1}} \dots \int_0^{t_2} t_1\, dt_1 \dots dt_{n-1} =$$
$$= \frac{C_T^n T^n}{n!} \;|||\; X - Y\;|||_T^2 \; . \tag{5.26}$$

For n sufficiently large, F^n is a contraction on the complete metric space $\big(\mathcal{C}_T, |||.|||_T\big)$. Hence, by the Contraction Mapping Theorem, F has a (unique) fixed point $X^{(T)} \in \mathcal{C}_T$. By uniqueness, $X_t^{(T)} = X_t^{(T')}$ for all $t \le T \wedge T'$ $a.s.$ and so we may consistently define $X \in \mathcal{C}$ by

$$X_t = X_t^{(N)} \quad \text{for } t \le N,\; N \in \mathbb{N} \; . \tag{5.27}$$

This is the pathwise unique solution to the SDE started from x. It remains to prove that it is $(\mathcal{F}_t^B)_{t\ge 0}$-adapted. Define a sequence $(Y^n)_{n\ge 0}$ in $\mathcal{C}_T$ by

$$Y^0 \equiv x \; , \quad Y^n = F(Y^{n-1}) \quad \text{for } n \ge 1 \; . \tag{5.28}$$

Then Y^n is $(\mathcal{F}_t^B)_{t\geq 0}$-adapted for each $n \geq 0$. Since $X = F^n(X)$ for all $n \geq 0$ by (5.26) we have

$$|||\, X - Y^n |||_T^2 \leq \frac{C_T^n T^n}{n!}\, |||\, X - x \,|||_T^2 \ . \tag{5.29}$$

Hence,

$$\mathbb{E}\bigg(\sum_{n=0}^{\infty} \sup_{0\leq t\leq T} |X_t - Y_t^n| \bigg) \leq \sum_{n=0}^{\infty} |||X - Y^n|||_T < \infty \tag{5.30}$$

which gives that $\displaystyle\sum_{n=0}^{\infty} \sup_{0\leq t\leq T} |X_t - Y_t^n| < \infty \ a.s.$.

So $Y^n \to X$ $a.s.$ as $n \to \infty$, uniformly in $[0,T]$. Hence, X is also $(\mathcal{F}_t^B)_{t\geq 0}$-adapted. □

Proposition 5.4 *Under the hypothesis of Theorem 5.3, there is uniqueness in law for the SDE*

$$dX_t = \sigma(X_t)\, dB_t + b(X_t)\, dt\ . \tag{5.31}$$

Proof. see example sheet 3, problem 3.

Thus all strong solutions started from x have the same law, even if defined on different probability spaces. We now give an important example of the Lipschitz setting.

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

$$dV_t = dB_t - \lambda V_t\, dt\ , \quad V_0 = v_0\ , \qquad dX_t = V_t\, dt, \quad X_0 = x_0\ . \tag{5.32}$$

When $\lambda > 0$ this models the motion of a pollen grain on the surface of a liquid. X represents the x-coordinate of the grain's position and V represents its velocity in the x-direction. $-\lambda V$ is the friction force due to viscosity. Whenever $|V|$ becomes large, the system acts to reduce it. V is called the *Ornstein-Uhlenbeck (velocity) process*. This is a rare example of a SDE we can solve explicitly. By Itô's formula,

$$d(e^{\lambda t}\, V_t) = e^{\lambda t}\, dV_t + \lambda\, e^{\lambda t}\, V_t\, dt = e^{\lambda t}\, dB_t \tag{5.33}$$

and so $\quad e^{\lambda t}\, V_t = v_0 + \displaystyle\int_0^t e^{\lambda s}\, dB_s \quad$ i.e.

$$V_t = v_0\, e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)}\, dB_s\ . \tag{5.34}$$

This is the pathwise unique strong solution guaranteed by Theorem 5.3. By Proposition 4.2,

$$V_t \sim N\big(v_0\, e^{-\lambda t}, \tfrac{1}{2\lambda}(1 - e^{-2\lambda t})\big)\ . \tag{5.35}$$

If $\lambda > 0$, $V_t \xrightarrow{d} N(0, \frac{1}{2\lambda})$ as $t \to \infty$. Indeed, $N(0, \frac{1}{2\lambda})$ is a stationary distribution for V in the sense that if $V_0 \sim N(0, \frac{1}{2\lambda})$ then $V_t \sim N(0, \frac{1}{2\lambda})$ for all $t \geq 0$.

5.3 Local solutions

Definition 5.2 A *locally defined* process (X, ζ) is a stopping time ζ together with a map

$$X : \big\{(\omega, t) \in \Omega \times [0, \infty) \ : \ t < \zeta(\omega)\big\} \to \mathbb{R} \,. \tag{5.36}$$

It is *càdlàg* if $t \mapsto X_t(\omega) : [0, \zeta(\omega)) \to \mathbb{R}$ is càdlàg for all $\omega \in \Omega$. Let $\Omega_t = \big\{\omega \ : \ t < \zeta(\omega)\big\}$. Then (X, ζ) is *adapted* if $X_t : \Omega_t \to \mathbb{R}$ is $\mathcal{F}_t$-measurable for all $t \geq 0$.
We say (X, ζ) is a *locally defined continuous local martingale* if there exist stopping times $T_n \nearrow \zeta \ a.s.$ such that X^{T_n} is a martingale for all $n \in \mathbb{N}$. Say (H, η) is a *locally defined locally bounded previsible process* if there exist stopping times $S_n \nearrow \eta \ a.s.$ such that $H\,\mathbb{1}_{(0,S_n]}$ is bounded and previsible for all $n \in \mathbb{N}$. Then we can define $(H \cdot X, \zeta \wedge \eta)$ via

$$(H \cdot X)_t^{S_n \wedge T_n} = \big((H\,\mathbb{1}_{(0,S_n \wedge T_n]}) \cdot X^{S_n \wedge T_n}\big)_t \quad \text{for all } n \in \mathbb{N} \,. \tag{5.37}$$

Proposition 5.5 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 $f \in C^2(U, \mathbb{R})$. Set $\zeta = \inf\{t \geq 0 \ : \ X_t \notin U\}$. Then for all $t < \zeta$,

$$f(X_t) = f(X_0) + \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x^i}(X_s)\, dX_s^i + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t \frac{\partial^2 f}{\partial x^i \partial x^j}(X_s)\, d[X^i, X^j]_s \,. \tag{5.38}$$

Proof. Apply Itô's formula (Theorem 3.4) to X^{T_n}, where

$$T_n = \inf\Big\{t \geq 0 \ : \ |X_t - y| \leq \frac{1}{n} \text{ for some } y \in U^c\Big\} \,, \tag{5.39}$$

and use $T_n \nearrow \zeta \ a.s.$. □

Example. Let $X = B$ be a Brownian motion started from 1, $d = 1$, $U = (0, \infty)$ and $f(x) = \sqrt{x}$. Then

$$\sqrt{B_t} = 1 + \int_0^t \frac{1}{2} B_s^{-1/2}\, dB_s + \frac{1}{2} \int_0^t (-\frac{1}{4}\, B_s^{-3/2})\, ds \tag{5.40}$$

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

Suppose now that $U \subseteq \mathbb{R}^d$ is open and $\sigma : U \to \mathbb{R}^{d \times m}$ and $b : U \to \mathbb{R}^d$ are measurable functions, bounded on compact sets. By a *local solution* to the SDE

$$dX_t = \sigma(X_t)\, dB_t + b(X_t)\, dt \,, \tag{5.41}$$

we mean

- a filtered probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}\big)$ satisfying the usual conditions;
- an $(\mathcal{F}_t)_{t \geq 0}$-Brownian motion $B = (B^1, \ldots, B^m)$ taking values in $\mathbb{R}^m$;
- an $(\mathcal{F}_t)_{t \geq 0}$-adapted continuous locally defined process (X, ζ) such that $X \in U$ and

$$X_t = X_0 + \int_0^t \sigma(X_s)\, dB_s + \int_0^t b(X_s)\, ds \quad \text{for all } t < \zeta \,. \tag{5.42}$$

We say that (X, ζ) is *maximal* if, for any other local solution (X, η) on the same set-up,

$$(X_t = Y_t \text{ for all } t < \zeta \wedge \eta) \quad \Rightarrow \quad \eta \leq \zeta \,. \tag{5.43}$$