Stochastic Calculus and Applications
Lent 2007

Dr. S. Großkinsky* and Prof. J. R. Norris†

Chris Almost‡

Contents

*stefan@statslab.cam.ac.uk, http://www.statslab.cam.ac.uk/~stefan/
†j.r.norris@statslab.cam.ac.uk, http://www.statslab.cam.ac.uk/~james/
‡cda23@cam.ac.uk

0 Introduction

The goal of this course is to develop tools to handle continuous-time Markov processes. Spatially homogenous examples of such processes in $\mathbb{R}^d$ include

(i) *scaled Brownian motion* $(W_t, t \geq 0)$ such that $W_t \sim \mathcal{N}(0, at)$, where $a = \mathrm{Var}(W_1)$ is the *diffusivity*;

(ii) *constant drift* $(bt, t \geq 0)$, where $b \in \mathbb{R}^d$; and

(iii) the *jump process* $(y_t, t \geq 0)$, which at the times of a Poisson process of rate $\lambda \geq 0$ makes a jump of distribution $\mu(dy)$ (on $\mathbb{R}^d$). We represent all the information in a jump process in a *Lévy measure* $K(dy) = \lambda \cdot \mu(dy)$. We may recover the information via $\lambda = K(\mathbb{R}^d)$ and $\mu(dy) = \frac{1}{\lambda} K(dy)$.

$(X_t, t \geq 0)$ defined as $X_t = W_t + bt + y_t$ is the most general spatially homogeneous process subject to the constraint of piecewise continuity (according to the Lévy-Khinchin Theorem).

In the inhomogeneous case, subject to reasonable regularity conditions, the general case corresponds to scaled Brownian motion with diffusivity $a(x)$, drift $b(x)$, and Lévy measure $K(x, dy)$ (the idea in the latter is to stay at x for a time $T \sim \mathcal{E}^{xp}(\lambda(x))$, where $\lambda(x) = K(x, \mathbb{R}^d)$, then make a jump with distribution $\mu(x, dy) = \frac{1}{\lambda(x)} K(x, dy)$).

The importance of martingales

Let us do a non-rigorous calculation in dimension $d = 1$ for illustration. Suppose $(X_t, t \geq 0)$ has parameters $a(x)$, $b(x)$, and $K(x, dy)$. Conditional on $X_s = x_s$ for all $s \leq t$,

$$X_{t+dt} = x_t + b(x_t)dt + G_t + y_t \mathbf{1}_{J_t}$$

where $G_t \sim \mathcal{N}(0, a(x_t)dt)$ (think $G_t = B_{t+dt} - B_t$), $y_t \sim \mu(x_t, dy)$, and $J_t \subset \Omega$ is the event that there is a jump at time t (so $\mathbb{P}(J_t) = \lambda(x_t)dt$). For $f \in C_b^2(\mathbb{R})$, by Taylor's Theorem,

$$\begin{aligned}
&\mathbb{E}[f(X_{t+dt}) - f(X_t) \mid X_s = x_s, s \leq t] \\
&\qquad = \mathbb{E}[f'(x_t)(b(x_t)dt + G_t) + \tfrac{1}{2} f''(x_t) G_t^2 \mid X_s = x_s, s \leq t] \\
&\qquad\qquad + \mathbb{P}(J_t)\mathbb{E}[f(x_t + y_t) - f(x_t) \mid X_s = x_s, s \leq t] = Lf(x_t)dt,
\end{aligned}$$

where $Lf(x)$ is the *generator*, defined by

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

Consider the process

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

We have "shown" that

$$\mathbb{E}[M_{t+dt}^f - M_t^f \mid X_s = x_s, s \leq t] = 0$$

for all t, so $(M_t^f, t \geq 0)$ is a martingale for all $f \in C_b^2(\mathbb{R})$. This property will serve to characterize the process $(X_t, t \geq 0)$.

For the first 18 lectures we take $K(x, dy) = 0$. Let $\sigma(x) = \sqrt{a(x)}$, so that $G_t \sim \sigma(x)(B_{t+dt} - B_t)$, where $(B_t, t \geq 0)$ is a standard Brownian motion. Whence we have the *stochastic differential equation*

$$X_{t+dt} - X_t = \sigma(X_t)(B_{t+dt} - B_t) + b(X_t)dt,$$

or more commonly,

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

or in integrated form,

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

where $\int_0^t \sigma(X_s)dB_s$ is an example of a *stochastic integral*. In general, stochastic integrals have the form $\int_0^t H(\omega, s)dA(\omega, s)$.

1 Preliminaries

1.1 Finite variation integrals

The first case we look at is the deterministic case. Recall that a function $f : \mathbb{R} \to \mathbb{R}$ is *càdlàg* if it is right-continuous and has left limits everywhere (also known as *r.c.l.l.* in English). We write $\Delta f(t) = f(t) - f(t^-)$. Suppose that $a : [0, \infty) \to \mathbb{R}$ is increasing and càdlàg. Then there exists a unique Borel measure da on $[0, \infty)$ such that $da((s, t]) = a(t) - a(s)$: the Lebesgue-Stieltjes measure with distribution function a. For a non-negative measurable function $h : \mathbb{R} \to [0, \infty)$ and $t \geq 0$, we define

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

We extend this to integrable $h : \mathbb{R} \to \mathbb{R}$ and to càdlàg functions of the form $a = a' - a''$ where a' and a'' are increasing and càdlàg. Subject to the finiteness of all terms involved, we define

$$h \cdot a := h \cdot a' - h \cdot a'' = h^+ \cdot a' - h^- \cdot a' - (h^+ \cdot a'' - h^- \cdot a''),$$

where $h^\pm = (\pm h) \vee 0$.

1.1.1 Lemma. *Let $a : [0, \infty) \to \mathbb{R}$ be càdlàg. Define for $t \geq 0$,*

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

Then $v(t) = \lim_{n\to\infty} v^n(t)$ exists for all $t \geq 0$ (but may be infinite) and $t \mapsto v(t)$ is increasing.

Notation. We write $t_n^+ := 2^{-n}\lceil 2^n t \rceil$, $t_n^- := 2^{-n}(\lceil 2^n t \rceil - 1)$, and

$$D_n := \{(k2^{-n}, (k+1)2^{-n}] : k \in \mathbb{N}\}.$$

With this notation $t \in (t_n^-, t_n^+] \in D_n$ for all $n \in \mathbb{N}$.

PROOF: We have

$$v^n(t) = \sum_{\substack{I \in D_n \\ \inf I < t}} |da(I)| = \overbrace{|a(t_n^+) - a(t_n^-)|}^{\text{converges to } |\Delta a(t)|} + \overbrace{\sum_{\substack{I \in D_n \\ \sup I < t}} |da(I)|}^{\text{increasing in } n},$$

so $\lim_{n\to\infty} v^n(t)$ exists (but is possibly infinite). Since $v^n(t)$ is increasing in t for each n, the same holds for $v(t)$. □

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

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

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

$$u^n(t) = \sum_{\substack{I \in D_n \\ t < \inf I < T}} |da(I)|$$

for $t \leq T$. Then $u^n(t)$ is decreasing in t for each n. For $I = (c,d] \in D_n$, u^n is constant on $[c,d)$, and hence is right-continuous. Thus $\{t \in [0,T] : u(t) \leq x\}$ is closed for all $x \geq 0$. On the other hand, $u^n(t)$ is increasing in n for each t, so has a limit $u(t)$ for all $t \leq T$. Therefore

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

is closed, and thus $u(t)$ is also right-continuous since it is decreasing. For $t < T$,

$$v^n(T) = v^n(t) + u^n(t) + \underbrace{|a(t + 2^{-n}) - a(t)|}_{\text{converges to 0 as } n \to \infty} \mathbf{1}_{\text{dyadic rationals}}(t),$$

so $v(t) = v(T) - u(t)$ is càdlàg and $T \in \mathbb{N}$ was arbitrary. Further, from the equation in the proof of 1.1.1, $v(t) = v(t^-) + |\Delta a(t)|$.

We have shown v is càdlàg, so $a^\pm = \frac{1}{2}(v \pm a)$ are also càdlàg. For each $m \in \mathbb{N}$, $dv^m(I) = |da(I)|$ for all $I \in D_m$ (by definition). By the triangle inequality, for each $n \geq m$, $dv^n(I) \geq |da(I)|$ for all $I \in D_m$. Thus

$$da^\pm(I) = \frac{1}{2} dv(I) \pm \frac{1}{2} da(I) \geq 0$$

by taking $n \to \infty$, for all $I \in \bigcup_{m \geq 1} D_m$. Whence $a^\pm$ are increasing, and $a = a^+ - a^-$ by definition.

Conversely, suppose that $a = a' - a''$. Without loss of generality, suppose $a(0) = a'(0) = a''(0) = 0$. Then $|da(I)| \leq da'(I) + da''(I)$ for all $I \in D_n$, for all $n \geq 0$. Summing over $I \in D_n$, the sum telescopes to show $v^n(t) \leq a'(t_n^+) + a''(t_n^+)$. Letting $n \to \infty$, $v(t) \leq a'(t) + a''(t) < \infty$, for all $t \geq 0$, so a is of finite variation. Finally, $a^\pm$ are the smallest such functions a' and a'' with this property because $v(t) \leq a'(t) + a''(t)$, whereas $a^\pm$ give equality, so any smaller functions would violate the inequality. □

Consider the filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t, t \geq 0), \mathbb{P})$. Recall that $X : \Omega \times [0, \infty) \to \mathbb{R}$ is *adapted* to the filtration $(\mathscr{F}_t, t \geq 0)$ if $X_t = X(\cdot, t)$ is $\mathscr{F}_t$-measurable for all $t \geq 0$, and X is *càdlàg* if $X(\omega, \cdot)$ is càdlàg for all $\omega \in \Omega$.

1.1.4 Definition. Let A be a càdlàg adapted process. Its *total variation process* V is defined pathwise (i.e. for each ω). We say that A is of *finite variation* if $A(\omega, \cdot)$ is of finite variation for every $\omega \in \Omega$.

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

PROOF: As in the proof of 1.1.1,

$$V_t = \lim_{n \to \infty} \sum_{k=0}^{2^n t_n^- - 1} |A_{(k+1)2^{-n}} - A_{k2^{-n}}| + |\Delta A_t|.$$

Each process on the right hand side is adapted since $t_n^- \leq t$ and since A is càdlàg. Thus V is adapted and is càdlàg and pathwise increasing for each $\omega \in \Omega$ by 1.1.3.□

1.2 Previsible processes

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

1.2.2 Proposition. *Let X be a càdlàg adapted process and set $H_t = X_{t^-}$ for $t > 0$. Then H is previsible.*

PROOF: Clearly H is left-continuous and adapted. Set $H_t^n = H_{t_n^-}$, so H^n is previsible for all n since $H_{t_n^-}$ is t_n^--measurable, as H is adapted and $t_n^- < t$. But $t_n^- \nearrow t$ as $n \to \infty$, so $H_t^n \to H_t$ by left-continuity, whence H is previsible. □

1.2.3 Proposition. *Let H be a previsible process. Then H_t is $\mathscr{F}_{t^-}$-measurable for all $t > 0$, where $\mathscr{F}_{t^-} = \sigma(\mathscr{F}_s : s < t)$.*

PROOF: See Example Sheet 1. □

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

1.2.4 Examples.

(i) Brownian motion is previsible because the paths are continuous (and hence left-continuous).

(ii) A Poisson process is *not* previsible, nor is any continuous-time Markov chain, since it has a discrete state space.

1.2.5 Proposition. *Let A be a càdlàg adapted process of finite variation, with total variation process V. Let H be previsible such that, for all $t \geq 0$, pathwise*

$$\int_{(0,t]} |H_s| dV_s < \infty.$$

Then the process defined pathwise by

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

is càdlàg, adapted, and of finite variation.

PROOF: $(H \cdot A)$ is càdlàg for each $\omega \in \Omega$ since

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

and by the Dominated Convergence Theorem. Further, $(H \cdot A)_{t^+} = (H \cdot A)_t$ since $\lim_{u \searrow t} \mathbf{1}_{(0,u]} = \mathbf{1}_{(0,t]}$ and $(H \cdot A)_{t^-} = \int_{(0,\infty)} H_s \mathbf{1}_{(0,t)}(s) dA_s$ exists, so $(H \cdot A)$ is càdlàg with $\Delta(H \cdot A)_t = H_t \Delta A_t$.

We show that $(H \cdot A)$ is adapted by a monotone class argument. Suppose that $H = \mathbf{1}_{B \times (s,u]}$, where $B \in \mathscr{F}_s$. Then $(H \cdot A)_t = \mathbf{1}_B(A_{u \wedge t} - A_{s \wedge t})$, which is clearly $\mathscr{F}_t$ measurable. Now let $\Pi = \{B \times (s,u] : B \in \mathscr{F}_s, s < u\}$, which is a π-system, and take $\mathscr{A} = \{C \in \mathscr{P} : (\mathbf{1}_C \cdot A)_t \text{ is } \mathscr{F}_t\text{-measurable}\}$. Then $\mathscr{A} \subseteq \mathscr{P} = \sigma(\Pi)$, and since $\mathscr{A}$ is a d-system and $\Pi \subseteq \mathscr{A}$, $\mathscr{A} = \mathscr{P}$ by Dynkin's Lemma.

Suppose now that H is non-negative and $\mathscr{P}$-measurable. For all $n \in \mathbb{N}$, set $H^n = (2^{-n} \lfloor 2^n H \rfloor) \wedge n$, so that $(H^n \cdot A)_t$ is $\mathscr{F}_t$-measurable since H^n is a linear combination of indicator functions $\mathbf{1}_P$ with $P \in \mathscr{P}$. Further, for fixed $t \geq 0$ and $\omega \in \Omega$, $(H^n \cdot A)_t \nearrow (H \cdot A)_t$ by the Monotone Convergence Theorem. Hence $(H \cdot A)_t$ is $\mathscr{F}_t$-measurable. This extends to $\mathscr{P}$-measurable $H = H^+ - H^-$ such that $|H| \cdot V = H^+ \cdot V + H^- \cdot V < \infty$ by linearity.

It remains to prove finite variation. Fix $\omega \in \Omega$, and let $H^\pm := (\pm H) \vee 0$ and $A^\pm := \frac{1}{2}(V \pm A)$. Analogous to the definition of $h \cdot a$, set

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

and from 1.1.3 it follows that $H \cdot A$ is of finite variation. □

Remark. The Wiener process $(W_t, t \geq 0)$ is *not* of finite variation, so we still can't make sense of $\int_{(0,t]} W_t dW_t$.

1.3 Martingales and local martingales

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t, t \geq 0), \mathbb{P})$ be a filtered probability space, where the filtration $(\mathscr{F}_t, t \geq 0)$ satisfies the *usual conditions*,

(i) $\mathscr{F}_0$ contains all $\mathbb{P}$-null sets; and

(ii) $\mathscr{F}_t = \mathscr{F}_{t^+} := \bigcap_{s>t} \mathscr{F}_s$ for all $t \geq 0$.

Recall that process X is an *integrable process* if $\mathbb{E}[|X_t|] < \infty$ for all $t \geq 0$, and an adapted integrable process is a *martingale* if

$$\mathbb{E}[X_t \mid \mathscr{F}_s] = X_s \text{ a.s.} \quad \text{for all } s \leq t.$$

We let $\mathscr{M}$ denote the set of all martingales. Finally, recall that $T : \Omega \to [0, \infty]$ is a *stopping time* if $\{T \leq t\} \in \mathscr{F}_t$ for all $t \geq 0$. We set

$$\mathscr{F}_T = \{A \in \mathscr{F} : A \cap \{T \leq t\} \in \mathscr{F}_t \text{ for all } t \geq 0\}.$$

If X is a càdlàg adapted process and T is a finite stopping time then X_T is $\mathscr{F}_T$-measurable.

1.3.1 Optional Stopping Theorem. *Let X be an adapted integrable process. The following are equivalent.*

(i) X is a martingale;

(ii) The stopped process $X^T := (X_{t\wedge T}, t \geq 0)$ is a martingale for all stopping times T;

(iii) For all stopping times S, T with $T < \infty$ a.s., $\mathbb{E}[X_T \mid \mathscr{F}_S] = X_{S\wedge T}$ a.s.; and

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

The OST implies that $\mathscr{M}$ is "stable under stopping." We will now define a localized version of this idea.

1.3.2 Definition. An adapted process is a *local martingale* if there exists a sequence $(T_n, n \in \mathbb{N})$ of stopping times with $T_n \nearrow \infty$ as $n \to \infty$ such that X^{T_n} is a martingale for all $n \in \mathbb{N}$. We say that $(T_n, n \in \mathbb{N})$ is a *reducing sequence* for X (or (T_n) *reduces* X). We write $\mathscr{M}_{loc}$ for the space of all local martingales.

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

$$\sup_{X\in\mathscr{X}} \mathbb{E}[|X|\mathbf{1}_{|X|\geq\lambda}] \to 0 \quad \text{as } \lambda \to \infty$$

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

$$\mathscr{X} = \{\mathbb{E}[X \mid \mathscr{G}] : \mathscr{G} \text{ is a sub-}\sigma\text{-algebra of } \mathscr{F}\}$$

is u.i.

PROOF: See exercise 9.1.7 in the Advanced Probability course notes. □

1.3.4 Proposition. *The following are equivalent.*

(i) $X \in \mathscr{M}$; and

(ii) $X \in \mathscr{M}_{loc}$ and for all $t \geq 0$

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

is u.i.

PROOF: If $X \in \mathscr{M} \subset \mathscr{M}_{loc}$ then by the OST, if T is a stopping time with $T \leq t$ then $X_T = \mathbb{E}[X_t \mid \mathscr{F}_T]$ a.s., so 1.3.3 gives the result.

Conversely, suppose (ii) holds and $(T_n, n \in \mathbb{N})$ reduces X. Then for every stopping time $T \leq t$ and $n \in \mathbb{N}$, X^{T_n} is a martingale so

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

But $\{X_{T\wedge T_n} : n \in \mathbb{N}\}$ is u.i., so $\mathbb{E}[X_{T\wedge T_n}] \to \mathbb{E}[X_T]$ since $T \wedge T_n \nearrow T$ a.s. as $n \to \infty$. Whence $\mathbb{E}[X_0] = \lim_{n\to\infty} \mathbb{E}[X_{T\wedge T_n}] = \mathbb{E}[X_T]$, so $X \in \mathscr{M}$ by the OST. □

Remark. 1.3.4 implies that a.s. bounded local martingales, or indeed any $M \in \mathscr{M}_{loc}$ such that $|M_t| \le Z$ a.s. for all $t \ge 0$ for some $Z \in L^1$, are in fact martingales.

1.3.5 Example. Let B be a standard Brownian motion in $\mathbb{R}^3$ with $|B_0| = 1$ and let $M_t = \frac{1}{|B_t|}$. The sequence $(S_n, n \in \mathbb{N})$ with $S_n = \inf\{t : |M_t| \ge n\}$ is reducing for M and $S_n \nearrow \infty$ as $n \to \infty$. For $t > s$,

$$\mathbb{E}[M_{t \wedge S_n} \mid \mathscr{F}_s] = M_{s \wedge S_n},$$

so $M^{S_n} \in \mathscr{M}$ and $M \in \mathscr{M}_{loc}$. But $\mathbb{E}[M_t \mid \mathscr{F}_s] < M_s$ since the variables $\{M_{t \wedge S_n} : n \in \mathbb{N}\}$ are not u.i. See Handout 1 for the details of this construction.

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

Notation. We let $\mathscr{M}_{c,loc}$ denote the collection of continuous local martingales.

PROOF: Note that for each $n \in \mathbb{N}$,

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

so S_n is a stopping time. For each $\omega \in \Omega$ and $t \ge 0$ we have by continuity $\sup_{s \le t} |M_s(\omega)| < \infty$ and $S_n(\omega) > t$ for all $n > \sup_{s \le t} |M_s(\omega)|$. Thus $S_n \nearrow \infty$ a.s. as $n \to \infty$.

Let $(T_k, k \in \mathbb{N})$ be a reducing sequence for M, so $M^{T_k} \in \mathscr{M}$ for each $k \in \mathbb{N}$. By the OST $M^{T_k \wedge S_n} \in M$ as well, and so $M^{S_n} \in \mathscr{M}_{loc}$ for each $n \in \mathbb{N}$ (with reducing sequence $(T_k, k \in \mathbb{N})$). But M^{S_n} is bounded by n, so $M^{S_n} \in \mathscr{M}$, for each $n \in \mathbb{N}$. □

1.3.7 Theorem. *Let M be a continuous local martingale of finite variation. If $M_0 = 0$ then $M \equiv 0$.*

Remark. In particular, Brownian motion is not of finite variation. The theory of finite variation integrals is not enough for our purposes.

PROOF: Let V be the total variation process of M. Then V is continuous and adapted with $V_0 = 0$. Set $R_n = \inf\{t \ge 0 : V_t \ge n\}$. Then R_n is a stopping time for all $n \in \mathbb{N}$ by 1.3.6, and $R_n \nearrow \infty$ a.s. as $n \to \infty$, since V_t is monotone and finite. It suffices to show $M^{R_n} \equiv 0$ for all $n \in \mathbb{N}$ since $R_n \nearrow \infty$ a.s. By the OST, $M^{R_n} \in \mathscr{M}_{loc}$. Also, $|M_t^{R_n}| \le |V_t^{R_n}| \le n$, so $M^{R_n} \in \mathscr{M}$ by 1.3.4.

Replace M by M^{R_n} in what follows to prove the result. Fix $t > 0$ and set $t_k := \frac{kt}{N}$. Then for all $k = 0, \dots, N$,

$$\mathbb{E}[M_{t_{k+1}} M_{t_k}] = \mathbb{E}[\mathbb{E}[M_{t_{k+1}} \mid \mathscr{F}_{t_k}] M_{t_k}] = \mathbb{E}[M_{t_k}^2],$$

so

$$\begin{aligned}\mathbb{E}[M_t^2] &= \mathbb{E}\left[\sum_{k=0}^{N-1}(M_{t_{k+1}}^2 - M_{t_k}^2)\right] \\ &= \mathbb{E}\left[\sum_{k=0}^{N-1}(M_{t_{k+1}} - M_{t_k})^2\right] \\ &\le \mathbb{E}\Bigg[\underbrace{\sup_{k<N}|M_{t_{k+1}} - M_{t_k}|}_{\text{converges to 0 as } N\to\infty} \underbrace{\sum_{k=0}^{N-1}|M_{t_{k+1}} - M_{t_k}|}_{\le V_t \le n}\Bigg]\end{aligned}$$

where the left term goes to zero by continuity. Therefore $\mathbb{E}[M_t^2] = 0$ for all $t \ge 0$, so $M \equiv 0$ a.s. □

1.3.8 Definition. A *semimartingale* X is an adapted càdlàg process which may be written as $X = X_0 + M + A$ with $M_0 = A_0 = 0$, where $M \in \mathscr{M}_{loc}$ and A is a finite variation process. (This is the *Doob-Meyer decomposition* of X.)

Remark. As a consequence of 1.3.7, the Doob-Meyer decomposition is unique for continuous semimartingales (where M is required to be continuous).

1.3.9 Example. Lévy processes are semimartingales. Let (a, q, Π) be a Lévy triple. The associated Lévy process is

$$X_t = at + \sqrt{q}B_t + y_t + Z_t,$$

where y_t is a compound Poisson process and Z_t is the compensated càdlàg L^2 martingale. $at + y_t$ is a process of finite variation and $\sqrt{q}B_t + Z_t$ is a martingale.

2 The Stochastic Integral

2.1 Simple integrands

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

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

where $n \in \mathbb{N}$, $0 = t_0 < \cdots < t_n < \infty$, and the Z_k are a bounded $\mathscr{F}_{t_k}$-measurable r.v.'s. Let $\mathscr{S}$ denote the set of all simple processes.

Remark. Notice that $\mathscr{S}$ is a vector space and by 1.2.2 every simple process is previsible.

Recall that X is bounded in L^2 if

$$\sup_{t\ge 0}\|X_t\|_2 = \sup_{t\ge 0}\mathbb{E}[X_t^2]^{\frac{1}{2}} < \infty.$$

Write $\mathscr{M}^2$ for the set of càdlàg L^2-bounded martingales. If $X \in \mathscr{M}^2$ then the L^2 Martingale Convergence Theorem ensures the existence of $X_\infty \in L^2(\mathscr{F}_\infty, \mathbb{P})$ such that $X_t \to X_\infty$ a.s. and in L^2 as $t \to \infty$. Moreover, $X_t = \mathbb{E}[X_\infty \mid \mathscr{F}_t]$ a.s. for all $t \geq 0$. Recall Doob's L^2 inequality, for $X \in \mathscr{M}^2$,

$$\|\sup_{t\geq 0}|X_t|\|_2 \leq 2\|X_\infty\|_2.$$

For $H \in \mathscr{S}$ (with the form above) and $M \in \mathscr{M}^2$ we define

$$\int_0^t H_s dM_s = (H\cdot M)_t := \sum_{k=0}^{n-1} Z_k(M_{t_{k+1}\wedge t} - M_{t_k\wedge t}).$$

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

(i) $H \cdot M^T = (H\cdot M)^T$;

(ii) $H \cdot M \in \mathscr{M}^2$;

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

PROOF:

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

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

(ii) For $t_k \leq s \leq t \leq t_{k+1}$ we have $(H\cdot T)_t - (H\cdot M)_s = Z_k(M_t - M_s)$ so that $\mathbb{E}[(H\cdot T)_t - (H\cdot M)_s \mid \mathscr{F}_s] = Z_k \mathbb{E}[M_t - M_s \mid \mathscr{F}_s] = 0$. This extends to general $s \leq t$ and hence $(H\cdot M)$ is a martingale. If $j < k$ we have

$$\begin{aligned}&\mathbb{E}[Z_j(M_{t_{j+1}} - M_{t_j})Z_k(M_{t_{k+1}} - M_{t_k})]\\ &\qquad = \mathbb{E}[Z_j(M_{t_{j+1}} - M_{t_j})Z_k \mathbb{E}[M_{t_{k+1}} - M_{t_k} \mid \mathscr{F}_{t_k}]] = 0\end{aligned}$$

so

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

as in the proof of 1.3.7. Applying Doob's inequality to $M_t - M_0$,

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

so $H \cdot M \in \mathscr{M}^2$.

(iii) Replace t by ∞ in the above argument. □

Notation. From now on, let $\mathscr{M}$ denote the set of càdlàg martingales and $\mathscr{M}_{loc}$ denote the set of càdlàg local martingales. $\mathscr{M}^2$ again denotes the set of L^2-bounded càdlàg martingales.

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

PROOF: Notice that $\mathscr{S} \subseteq L^2(\mathscr{P},\mu)$ since every simple process is bounded and of bounded support. Set $\mathscr{A} = \{A \in \mathscr{P} : \mathbf{1}_A \in \overline{\mathscr{S}}\}$, where $\overline{\mathscr{S}}$ is the closure of $\mathscr{S}$ in $L^2(\mathscr{P},\mu)$. Then $\mathscr{A}$ is a d-system and $\mathscr{A} \supseteq \Pi = \{B \times (s,t] : B \in \mathscr{F}_s, s < t\}$, a π-system generating $\mathscr{P}$. Therefore $\mathscr{A} = \mathscr{P}$ by Dynkin's Lemma. The result follows since linear combinations of indicator functions are dense in $L^2(\mathscr{P},\mu)$. □

2.2 L^2 properties

We work on a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t, t \geq 0), \mathbb{P})$, where $(\mathscr{F}_t, t \geq 0)$ satisfies the usual conditions.

2.2.1 Definition. For a cadlag adapted process X, define

$$|||X||| := \| \sup_{t \geq 0} |X_t| \|_2,$$

and let $\mathscr{C}^2$ denote the collection of all such processes X with $|||X||| < \infty$. On $\mathscr{M}^2$ define $\|X\| := \|X_\infty\|_2$.

2.2.2 Proposition. *We have*

(i) $(\mathscr{C}^2, |||\cdot|||)$ *is complete;*

(ii) $\mathscr{M}^2 = \mathscr{M} \cap \mathscr{C}^2$*;*

(iii) $(\mathscr{M}^2, \|\cdot\|)$ *is a Hilbert space and $\mathscr{M}^2_c$, the space of continuous L^2-bounded martingales, is a closed subspace; and*

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

Remark. We identify an element of $\mathscr{M}^2$ with its terminal value X_∞ and then $\mathscr{M}^2$ inherits the Hilbert space properties of L^2, with inner product

$$\mathscr{M}^2 \times \mathscr{M}^2 \to \mathbb{R} : (X,Y) \mapsto \mathbb{E}[X_\infty Y_\infty].$$

PROOF: (i) Suppose that $(X^n, n \in \mathbb{N})$ is a Cauchy sequence in $(\mathscr{C}^2, |||\cdot|||)$. 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 by the triangle inequality,

$$\left\| \sum_{k=1}^{\infty} \sup_{t \geq 0} |X_t^{n_{k+1}} - X_t^{n_k}| \right\|_2 \leq \sum_{k=1}^{\infty} |||X^{n_{k+1}} - X^{n_k}||| < \infty,$$

and so for a.e. $\omega \in \Omega$,

$$\sum_{k=1}^{\infty} \sup_{t \geq 0} |X_t^{n_{k+1}}(\omega) - X_t^{n_k}(\omega)| < \infty.$$

Thus $(X_t^{n_k}(\omega), k \in \mathbb{N})$ is a Cauchy sequence of real numbers, so it converges to some $X_t(\omega)$ by completeness of $\mathbb{R}$, uniformly in $t \geq 0$. $X(\omega)$ is cadlag as it is the uniform limit of cadlag functions (see Example Sheet 1, problem 10). Now

$$\begin{aligned} |||X^n - X|||^2 = \mathbb{E}[\sup_{t \geq 0} |X_t^n - X_t|^2] &\leq^{Fatou} \liminf_{k \to \infty} \mathbb{E}[\sup_{t \geq 0} |X_t^n - X_t^{n_k}|^2] \\ &= \liminf_{k \to \infty} |||X^n - X^{n_k}|||^2 \to 0 \end{aligned}$$

as $n \to \infty$ since $(X_n, n \in \mathbb{N})$ is Cauchy. Hence $(\mathscr{C}^2, ||| \cdot |||)$ is complete.

(ii) For $X \in \mathscr{C}^2 \cap \mathscr{M}$, by Fatou's Lemma

$$\sup_{t \geq 0} \|X_t\|_2 \leq \| \sup_{t \geq 0} |X_t| \|_2 = |||X||| < \infty,$$

so $X \in \mathscr{M}^2$. Conversely, for $X \in \mathscr{M}^2 \subseteq \mathscr{M}$, Doob's L^2 Inequality becomes

$$|||X||| \leq 2\|X\| < \infty$$

so $X \in \mathscr{C}^2$.

(iii) For $X \in \mathscr{M}^2$ we have $\|X\| \leq |||X||| \leq 2\|X\|$, so the norms are equivalent on $\mathscr{M}^2$. Thus $\mathscr{M}^2$ is complete for $\| \cdot \|$ if and only if it is complete for $||| \cdot |||$. By (i), it suffices to show that $\mathscr{M}^2$ is closed in $(\mathscr{C}^2, ||| \cdot |||)$. If $(X^n, n \in \mathbb{N}) \subset \mathscr{M}^2$ converges to some $X \in \mathscr{C}^2$ then X is certainly cadlag, adapted, and L^2-bounded. Further,

$$\begin{aligned} \| \mathbb{E}[X_t \mid \mathscr{F}_s] - X_s\|_2 &\leq \| \mathbb{E}[X_t - X_t^n \mid \mathscr{F}_s]\|_2 + \|X_s^n - X_s\|_2 \\ &\leq \|X_t - X_t^n\|_2 + \|X_s^n - X_s\|_2 \\ &\leq 2|||X^n - X||| \to 0, \end{aligned}$$

so $X \in \mathscr{M}^2$. By the same argument $\mathscr{M}_c^2$ is closed in $(\mathscr{M}^2, \| \cdot \|)$, where continuity follows from uniform convergence in $t \geq 0$ (since $t \mapsto X_t(\omega)$ is continuous).

(iv) This is by definition. □

2.3 Quadratic variation

2.3.1 Definition. For a sequence $(X^n, n \in \mathbb{N})$ we say that $X^n \to X$ *uniformly on compacts in probability* (or *u.c.p*) if for every $\varepsilon > 0$ and $t \geq 0$,

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

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

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

converges to $[M]$ u.c.p. as $n \to \infty$. We call $[M]$ the quadratic variation of M.

PROOF (UNIQUENESS): We consider the case when $M_0 = 0$ without loss of generality. If A and A' are two increasing processes that fulfill the conditions for $[M]$ then

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

and $A - A'$ has finite variation, so $A - A' \equiv 0$.

See handout for the proof of existence. The "spirit" of the proof is that

$$[M]_t = M_t^2 - 2\int_0^t M_s dM_s = M_t^2 - \int_0^t d(M_s^2) - \int_0^t (dM_s)^2. \qquad \square$$

2.3.3 Example. Let B be a standard Brownian Motion. Then $(B_t^2 - t, t \geq 0)$ is a martingale, so by the uniqueness statement in 2.3.2, $[B]_t = t$.

2.3.4 Theorem. *If $M \in \mathscr{M}_c^2$ then $M^2 - [M]$ is a u.i. martingale, i.e.*

$$\mathbb{E}[\sup_{t \geq 0} |M_t^2 - [M]_t|] < \infty.$$

PROOF: Let $R_n := \inf\{t \geq 0 : [M]_t \geq n\}$. Then R_n is a stopping time and $[M]_{t \wedge R_n} \leq n$ for all n. Thus

$$|M_{t\wedge R_n}^2 - [M]_{t \wedge R_n}| \leq \sup_{t \geq 0} M_t^2 + n$$

and the local martingale $(M^{R_n})^2 - [M]^{R_n}$ is bounded by an integrable r.v. and so is a martingale by 1.3.4. Thus $\mathbb{E}[[M]_{t\wedge R_n}] = \mathbb{E}[M_{t\wedge R_n}^2]$ for all $t \geq 0$. Taking $t \to \infty$ and then $R_n \to \infty$,

$$\begin{array}{ccccc} \mathbb{E}[[M]_{t\wedge R_n}] & = & \mathbb{E}[M_{t\wedge R_n}^2] & & \\ & \searrow^{MCT} & & \searrow^{DCT} & \\ & & \mathbb{E}[[M]_\infty] & = & \mathbb{E}[M_\infty^2] \end{array}$$

since $M \in \mathscr{M}_c^2$. Therefore

$$|M_t^2 - [M]_t| \leq |\sup_{t \geq 0} M_t^2 + [M]_\infty|,$$

which is integrable. Thus $M^2 - [M]$ is a martingale and is u.i. since, by Doob's L^2 Inequality,

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

Remark. Some books use $\langle M \rangle$ instead of $[M]$, which is slightly different since it is previsible. If M is continuous then they are equal, but in general they are different.

2.4 Itô integral

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

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

for all $s < t$ and $B \in \mathscr{F}_s$. Since events of this form generate $\mathscr{P}$, this uniquely determines the measure μ. Alternatively, we could define

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

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

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

2.4.1 Definition. Set $L^2(M) = L^2(\Omega \times (0,\infty), \mathscr{P}, \mu)$, and write

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

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

2.4.2 Theorem (Itô Isometry). *For $M \in \mathscr{M}_c^2$ there exists a unique isometry*

$$I : L^2(M) \to \mathscr{M}_c^2$$

such that $I(H) = H \cdot M$ for all $H \in \mathscr{S}$.

2.4.3 Definition. Itô's *stochastic integral* of H with respect to M is the process $H \cdot M := I(H)$, and we write

$$\int_0^t H_s dM_s = (H \cdot M)_t := I(H)_t.$$

PROOF: We must first check that $H \mapsto H \cdot M$ satisfies $\|H \cdot M\| = \|H\|_M$. Let $H = \sum_{k=0}^{n-1} Z_k \mathbf{1}_{(t_k, t_{k+1}]} \in \mathscr{S}$. By 2.1.2, $H \cdot M \in \mathscr{M}_c^2$ with

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

Now $M^2 - [M]$ is a martingale by 2.3.4, so for $k = 0, \dots, n-1$,

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

Thus

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

Now let $H \in L^2(M)$. Then by 2.1.3 there exist $H^n \in \mathscr{S}$ such that $H^n \to H$ in $L^2(M)$ as $n \to \infty$, i.e.

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

Following the first part,

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

so $(H^n \cdot M, n \in \mathbb{N})$ is Cauchy in $\mathscr{M}_c^2$. Thus, since $(\mathscr{M}_c^2, \|\cdot\|)$ is complete, $(H^n \cdot M, n \in \mathbb{N})$ converges to some $H \cdot M \in \mathscr{M}_c^2$. Then

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

so $H \mapsto H \cdot M$ is an isometry on $L^2(M)$. □

2.4.4 Proposition. *Let $M \in \mathscr{M}_c^2$ and $H \in L^2(M)$ and let T be a stopping time. Then $(H \cdot M)^T = (H\mathbf{1}_{(0,T]}) \cdot M = H \cdot M^T$.*

PROOF: Suppose first that $H \in \mathscr{S}$. If T takes only finitely many values then $H\mathbf{1}_{(0,T]} \in \mathscr{S}$ and $(H \cdot M)^T = (H\mathbf{1}_{(0,T]}) \cdot M$ can be checked elementarily. For general T, set $T^n = (2^{-n}\lfloor 2^n T\rfloor) \wedge n$, so T^n takes only finitely many values and $T^n \nearrow T$ as $n \to \infty$. Thus

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

by DCT. Thus $(H\mathbf{1}_{(0,T^n]}) \cdot M \to (H\mathbf{1}_{(0,T]}) \cdot M$ in $\mathscr{M}_c^2$ by the Itô Isometry. But $(H \cdot M)_t^{T^n} \to (H \cdot M)_t^T$ by continuity, and hence $(H\mathbf{1}_{(0,T]}) \cdot M = (H \cdot M)^T$. We know from 2.1.2 that $(H \cdot M)^T = H \cdot M^T$.

For $H \in L^2(M)$, choose $H^n \in \mathscr{S}$ such that $H^n \to H$ in $L^2(M)$ as $n \to \infty$. Then $H^n \cdot M \to H \cdot M$ in $\mathscr{M}_c^2$ by the Itô Isometry, so $(H^n \cdot M)^T \to (H \cdot M)^T$ in $\mathscr{M}_c^2$. Also, as $n \to \infty$,

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

thus $(H^n\mathbf{1}_{(0,T]}) \cdot M \to (H\mathbf{1}_{(0,T]}) \cdot M$ in $\mathscr{M}_c^2$. Therefore $(H \cdot M)^T = (H\mathbf{1}_{(0,T]}) \cdot M$. Moreover,

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

so $H^n \cdot M^T \to H \cdot M^T$, and thus $(H \cdot M)^T = H \cdot M^T$. □

2.4.5 Definition. (i) A previsible process H is *locally bounded* if there exist stopping times $R_n \nearrow \infty$ a.s. such that $H\mathbf{1}_{(0,R_n]}$ is bounded for all $n \in \mathbb{N}$.

(ii) Let $M \in \mathcal{M}_{c,loc}$ and H be a locally bounded previsible process. Let R_n be as in the previous definition and $S_n = \inf\{t \geq 0 : |M_t| \geq n\}$, so that $H\mathbf{1}_{(0,R_n]} \in L^2(M)$ and $M^{S_n} \in \mathcal{M}_c^2$ for all $n \in \mathbb{N}$. Set $T_n = R_n \wedge S_n$ and define the *stochastic integral* of H with respect to M by

$$(H \cdot M)_t := ((H\mathbf{1}_{(0,T_n]}) \cdot M^{T_n})_t$$

for all $t \leq T_n$

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

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

(ii) $H \cdot M \in \mathcal{M}_{c,loc}$*;*

(iii) $[H \cdot M] = H^2 \cdot [M]$*; and*

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

PROOF:

(i) With the notation from the definition of locally bounded processes, (i) follows directly from 2.4.4 applied to $H\mathbf{1}_{(0,T_n]} \in L^2(M)$ and $M^{T_n} \in \mathcal{M}_c^2$ for all $n \in \mathbb{N}$.

(ii) With the same notation, by (i), $(H \cdot M)^{T_n} = (H\mathbf{1}_{(0,T_n]}) \cdot M^{T_n} \in \mathcal{M}_c^2$ for all $n \in \mathbb{N}$, so $H \cdot M \in \mathcal{M}_{c,loc}$, i.e. $(T_n, n \in \mathbb{N})$ reduces $H \cdot M$.

Again using $(H \cdot M)^{T_n} = (H\mathbf{1}_{(0,T_n]}) \cdot M^{T_n}$, we can reduce (iii) and (iv) to the case where $M \in \mathcal{M}_c^2$ and H, K bounded uniformly in time. This type of argument is a *localization argument* and will be spelled out in (iii).

(iii) When H is uniformly bounded in t and $M \in \mathcal{M}_c^2$, for a stopping time T,

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

By the OST this implies that $(H \cdot M)^2 - H^2 \cdot [M]$ is a martingale. Therefore $[H \cdot M] = H^2 \cdot [M]$ by uniqueness in 2.3.2. But then we know

$$\begin{aligned}[H \cdot M]^{T_n} = [(H \cdot M)^{T_n}] &= [(H \cdot \mathbf{1}_{(0,T_n]}) \cdot M^{T_n}] \\ &= (H^2\mathbf{1}_{(0,T_n]}) \cdot [M^{T_n}] = (H^2 \cdot [M])^{T_n}\end{aligned}$$

for all $n \in \mathbb{N}$. Therefore $[H \cdot M]_t = (H^2 \cdot [M])_t$ for all $t \geq 0$ by letting $n \to \infty$.

(iv) The case $H, K \in \mathcal{S}$ can be checked elementarily. For H, K uniformly bounded in t there exist $H^n, K^n \in \mathcal{S}$, for $n \in \mathbb{N}$, such that $H^n \to H$ and $K^n \to K$ in $L^2(M)$. We have

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

for each $n \in \mathbb{N}$ and

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

because

$$\begin{aligned}
\|H\|^2_{L^2(K\cdot M)} &= \mathbb{E}[(H^2 \cdot [K \cdot M])_\infty] \\
&= \mathbb{E}[(H^2 \cdot (K^2 \cdot [M]))_\infty] \\
&= \mathbb{E}[(HK \cdot [M])_\infty] \\
&= \|HK\|^2_{L^2(M)} \\
&\le \min\{\|H\|^2_\infty \|K\|^2_{L^2(M)}, \|K\|^2_\infty \|H\|^2_{L^2(M)}\}.
\end{aligned}$$

Therefore $H^n \cdot (K^n \cdot M) \to H \cdot (K \cdot M)$ and $(H^n K^n) \cdot M \to (HK) \cdot M$ in $\mathscr{M}_c^2$. Together these imply the result. □

2.4.7 Definition. For a continuous semimartingale $X = X_0 + M + A$ and H a locally bounded previsible process, the *stochastic integral* of H with respect to X is $H \cdot X := H \cdot M + H \cdot A$, where $H \cdot M$ is the Itô stochastic integral and $H \cdot A$ is the finite variation integral defined in 1.2.5.

Remark.

(i) Notice that $H \cdot X$ is a semimartingale with Doob-Meyer decomposition $H \cdot M + H \cdot A$ since $H \cdot M \in \mathscr{M}_{c,loc}$ and $H \cdot A$ is of finite variation.

(ii) We set $[X] := [M]$ since, informally,

$$[A]_t = \int_0^t (dA)^2 \le |dA| \int_0^t |dA| \to 0.$$

This is mathematically justifiable, in that one can show

$$\sum_{k=0}^{\lfloor 2^n t \rfloor - 1} (X_{(k+1)2^{-n}} - X_{k2^{-n}})^2 \to [M]_t.$$

(iii) Under the additional assumption that H is left-continuous the Riemann sum approximation converges to the integral (see 2.4.8).

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

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

PROOF: Problem 6 on Example Sheet 1 deals with the finite variation part of X, so we show the result for $X = M \in \mathscr{M}_{c,loc}$. By localization it suffices to consider $M \in \mathscr{M}_c^2$ and H uniformly bounded in t. Let $H_t^n = H_{2^{-n}\lfloor 2^n t\rfloor}$. Then $H^n \to H$ as $n \to \infty$ by left-continuity. Now

$$(H^n \cdot M)_t = H_{2^{-n}\lfloor 2^n t\rfloor} \underbrace{(M_t - M_{2^{-n}\lfloor 2^n t\rfloor})}_{\substack{\text{converges to 0} \\ \text{as } n \to \infty}} + \sum_{k=0}^{\lfloor 2^n t\rfloor - 1} H_{k2^{-n}}(M_{(k+1)2^{-n}} - M_{k2^{-n}}).$$

By DCT,

$$\|H_n - H\|_M = \mathbb{E}\left[\int_0^\infty (H_t^n - H_t)^2 d[M]_t\right] \to 0$$

as $n \to \infty$, so by the Itô Isometry we get $H^n \cdot M \to H \cdot M$ in $\mathscr{M}_c^2$. By problem 14, convergence in $\mathscr{M}_c^2$ implies u.c.p. convergence. □

3 Stochastic Calculus

3.1 Covariation

3.1.1 Theorem (Covariation). *Let $M, N \in \mathscr{M}_{c,loc}$ and set*

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

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

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

(ii) $MN - [M,N] \in \mathscr{M}_{c,loc}$;

(iii) for $M, N \in \mathscr{M}_c^2$, $MN - [M,N]$ *is a u.i. martingale;*

(iv) $[H \cdot M, N] + [M, H \cdot N] = 2H \cdot [M,N]$ *for H locally bounded and previsible.*

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

Remark. Clearly $[M,M] = [M]$, and covariation is bilinear.

PROOF: Note that $MN = \frac{1}{4}((M+N)^2 - (M-N)^2)$ so we take

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

This is the *polarization identity*. Compare $[M,N]^n$ to $[M]^n$ in 2.3.2 to get

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

Statements (i)–(iii) follow directly from 2.3.2 on quadratic variation.

For statement (iv), from 2.4.6, $[H\cdot(M\pm N)] = H^2\cdot[M\pm N]$, so the polarization identity gives $[H\cdot M, H\cdot N] = H^2\cdot[M,N]$. By bilinearity,

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

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

Proof: By 3.1.1, part (iv), it suffices to show that

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

By part (ii) of that theorem

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

We will show that

$$(H\cdot M)N - M(H\cdot N) \in \mathscr{M}_{c,loc},$$

which will imply that

$$[H\cdot M,N] - [M,H\cdot N] \in \mathscr{M}_{c,loc}.$$

Since H is of finite variation, 1.3.7 implies $[H\cdot M,N] \equiv [M,H\cdot N]$. By localization we may assume that $M,N \in \mathscr{M}_c^2$ and H is bounded uniformly in time. By the OST it suffices to show that

$$\mathbb{E}[(H\cdot M)^T N^T] = \mathbb{E}[M^T (H\cdot N)^T]$$

for all stopping times T. By 2.4.4 we may replace M by M^T and N by N^T, so it suffices to show that

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

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

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

By linearity this extends to all $H \in \mathscr{S}$. We can always find a sequence $H^n \in \mathscr{S}$ such that $H^n \to H$ simultaneously in $L^2(M)$ and $L^2(N)$, e.g. by taking $\mu = d[M] + d[N]$ in 2.1.3. Then

$$(H^n\cdot M)_\infty \to (H\cdot M)_\infty \quad \text{and} \quad (H^n\cdot N)_\infty \to (H\cdot N)_\infty$$

in L^2. This finishes the proof. □

3.2 Itô's Formula

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

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

PROOF: Since both sides are continuous in time we consider only times of the form $t = M2^{-N}$, $M, N \geq 1$. Note that

$$X_t Y_t - X_s Y_s = X_s(Y_t - Y_s) + Y_s(X_t - X_s) + (X_t - X_s)(Y_t - Y_s)$$

so for $n \geq N$,

$$\begin{aligned} X_t Y_t - X_0 Y_0 \\ = \sum_{k=0}^{M2^{n-N}-1} X_{k2^{-n}}(Y_{(k+1)2^{-n}} - Y_{k2^{-n}}) + Y_{k2^{-n}}(X_{(k+1)2^{-n}} - X_{k2^{-n}}) \\ + (X_{(k+1)2^{-n}} - X_{k2^{-n}})(Y_{(k+1)2^{-n}} - Y_{k2^{-n}}). \end{aligned}$$

This converges u.c.p. to

$$(X \cdot Y)_t + (Y \cdot X)_t + [X,Y]_t$$

as $n \to \infty$ by 2.4.8 and 3.1.1. □

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

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

Remark. In particular, $f(X)$ is a continuous semimartingale in its Doob-Meyer decomposition. Recall that the covariation of X^i and X^j is simply $[M^i, M^j]$.

$$\underbrace{\sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) dM_s^i}_{\text{in } \mathscr{M}_{c,loc}} + \underbrace{\sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) dA_s^i + \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) d[X^i, X^j]_s}_{\text{of finite variation}}.$$

Remark. The intuitive proof for the case $d = 1$ is given by the Taylor expansion of f,

$$\begin{aligned} f(X_{\lfloor 2^n t \rfloor}) &= f(X_0) + \sum_{k=0}^{\lfloor 2^n t \rfloor - 1} (f(X_{(k+1)2^{-n}}) - f(X_{k2^{-n}})) \\ &= f(X_0) + \sum_{k=0}^{\lfloor 2^n t \rfloor - 1} f'(X_{k2^{-n}})(X_{(k+1)2^{-n}} - X_{k2^{-n}}) \\ &\quad + \frac{1}{2} f''(X_{k2^{-n}})(X_{(k+1)2^{-n}} - X_{k2^{-n}})^2 + \text{error} \\ &\xrightarrow{u.c.p.} f(x_0) + \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) d[X_s] \end{aligned}$$

In practise, detailed estimation and control of the error term is more difficult than the proof given below.

PROOF: We prove the case $d = 1$, namely that

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

Write $X = X_0 + M + A$, where A has total variation process V. Let

$$T_r = \inf\{t \geq 0 : |X_t| + V_t + [M]_t > r\},$$

so $(T_r, r \geq 0)$ is a family of stopping times with $T_r \nearrow \infty$ a.s. as $r \to \infty$. It is sufficient to show (1) for all $t \leq T_r$. Let $\mathscr{A} := \{f \in C^2(\mathbb{R}) : (1) \text{ holds}\}$. Then

(i) $\mathscr{A}$ contains $x \mapsto 1$ and $x \mapsto x$; and

(ii) $\mathscr{A}$ is a vector space.

We will show

(iii) $f, g \in \mathscr{A}$ implies $fg \in \mathscr{A}$; and

(iv) if $f_n \in \mathscr{A}$ and $f_n \to f$ in $C^2(\mathbb{R})$ then $f \in \mathscr{A}$ (i.e. $\mathscr{A}$ is closed).

(Recall that $f_n \to f$ in $C^2(\mathbb{R})$ if

$$\Delta_{n,r} = \max\{\sup_{x \in B_r} |f_n^{(i)}(x) - f^{(i)}(x)| : i = 0, 1, 2\}$$

What is the precise statement of the Weierstrass theorem that is being applied here?

goes to zero as $n \to \infty$.) Once we do this, (i)–(iii) imply that $\mathscr{A}$ contains all polynomials. The Weierstrass Approximation Theorem shows that the polynomials are dense in $C^2(\mathbb{R})$, so (iv) implies $\mathscr{A} = C^2(\mathbb{R})$.

To prove (iii), take $f, g \in \mathscr{A}$ and set $F_t = f(X_t)$ and $G_t = g(X_t)$, which are continuous semimartingales. By integration by parts,

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

By 2.4.6, part (iv), $H \cdot (K \cdot M) = (HK) \cdot M$, so

$$\begin{aligned}\int_0^t F_s dG_s = (F \cdot G)_t &= (F \cdot (1 \cdot G))_t \\ &= \int_0^t F_s d(g(X_s) - g(X_0)) \\ &= \int_0^t f(X_s)g'(X_s)dX_s + \frac{1}{2}\int_0^t f(X_s)g''(X_s)d[X]_s\end{aligned}$$

by Itô's formula applied to g. By Kunita-Watanabe,

$$[F, G]_t = [f'(X) \cdot X, g'(X) \cdot X] = \int_0^t f'(X_s)g'(X_s)d[X]_s,$$

since $\int_0^t f'(X_s)dX_s = \int_0^t dF_s$. Substituting in the integration by parts formula yields Itô's formula for fg.

To prove (iv), let $f_n \in \mathscr{A}$ be such that $f_n \to f$ in $C^2(\mathbb{R})$. Then

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

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

$$\|(f_n' \cdot M)^{T_r} - (f' \cdot M)^{T_r}\|_2 = \mathbb{E}\left[\int_0^{T_r} (f_n'(X_s) - f'(X_s))^2 d[M]_s\right]$$
$$\le \Delta_{n,r}^2\, \mathbb{E}[[M]_{T_r}] \le r\Delta_{n,r}^2 \to 0 \text{ as } n \to \infty.$$

By taking the limit $n \to \infty$ in Itô's formula for the f_n we get

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

for all $r \ge 0$, which implies the result. □

Remark. For the case $d > 1$, (i) becomes "$\mathscr{A}$ contains $x \mapsto 1$ and $x \mapsto x^i$ for $i = 1, \ldots, d$." The argument follows as before.

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

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

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

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

3.3 Stratonovich integral

3.3.1 Definition. Let X, Y be continuous semimartingales. The *Stratonovich integral* of Y with respect to X is

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

Remark. We have discrete approximations for the right hand side converging u.c.p. to the left hand side,

$$\sum_{k=0}^{\lfloor 2^n t\rfloor - 1} \frac{Y_{(k+1)2^{-n}} + Y_{k2^{-n}}}{2}(X_{(k+1)2^{-n}} - X_{k2^{-n}}).$$

Informally, the Stratonovich integral is the limit of Riemann sums approximating via the midpoint, whereas the Itô stochastic integral is the limit of Riemann sums approximating via the lefthand endpoint The Itô integral can be approximated by Riemann sums if the integrand is left-continuous, by the Stratonovich integral requires that the integrand is a continuous semi-martingale.

3.3.2 Proposition. *Let $X^1,\dots,X^d$ be continuous semi-martingales and $f \in C^3(\mathbb{R}^d,\mathbb{R})$. Then*

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

In particular, the integration by parts formula becomes

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

PROOF: Again we prove only the case $d = 1$. By Itô's formula

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

and

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

By Kunita-Watanabe and the second formula, $[f'(X),X] = f''(X)\cdot[X]$. then by the first formula and the definition of the Stratonovich integral,

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

□

3.3.3 Example. Let B be standard Brownian motion. Then

$$\int_0^t B_s \partial B_s = \int_0^t B_s dB_s + \frac{1}{2}[B_t] = \frac{1}{2}B_t^2 \notin \mathcal{M}_{loc}$$

by Itô's formula.

3.4 Notation and summary

3.4.1 Conventions. *The following shorthand conventions will be used for the remainder of the course.*

(i) $dZ_t = H_t dX_t$ *means* $Z_t - Z_0 = \int_0^t H_s dX_s$. *Keep in mind that in this notation* $H_t dX_t = d(H\cdot X)_t$;

(i') $\partial Z_t = H_t \partial X_t$ *means* $Z_t - Z_0 = \int_0^t H_s \partial X_s$;

(ii) $dZ_t = dX_t dY_t$ *means* $Z_t - Z_0 = [X,Y]_t$. *Keep in mind that the latter is equal to* $\int_0^t dX_s dY_s$.

We have shown in 2.4.6, part (iv), that

$$H_t(K_t dX_t) = (H_t K_t) dX_t.$$

The Kunita-Watanabe idenity is

$$H_t(dX_t dY_t) = (H_t dX_t) dY_t,$$

integration by parts is

$$d(X_t Y_t) = Y_t dX_t + X_t dY_t + dX_t dY_t,$$

and Itô's formula is

$$\begin{aligned} d(f(X_t)) &= \frac{\partial f}{\partial x^i}(X_t) dX_t^i + \frac{1}{2}\frac{\partial^2 f}{\partial x^i \partial x^j}(X_t) dX_t^i dX_t^j \\ &= \langle Df(X_t), dX_t\rangle + \langle \tfrac{1}{2} D^2 f(X_t) dX_t, dX_t\rangle \end{aligned}$$

or in the case $d = 1$,

$$df(X_t) = f'(X_t) dX_t + \tfrac{1}{2} f''(X_t) d[X]_t.$$

4 Applications of Stochastic Calculus

4.1 Brownian motion

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t, t \geq 0), \mathbb{P})$ be a filtered probability space, where $(\mathscr{F}_t, t \geq 0)$ fulfills the usual conditions.

4.1.1 Lévy's Characterization of Brownian Motion. *If $X^1, \dots, X^d \in \mathscr{M}_{c,loc}$ are such that $[X^i, X^j]_t = \delta_{i,j} t$ then $X = (X^1, \dots, X^d)$ is a Brownian motion in $\mathbb{R}^d$.*

PROOF: We have to show that for $0 \leq s \leq t$, $X_t - X_s \sim \mathscr{N}(0, (t-s) I_d)$ and the increments are independent of $\mathscr{F}_s$. This happens if and only if

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

for all $\theta \in \mathbb{R}^d$ and $s \leq t$. Fix $\theta \in \mathbb{R}^d$ and set $Y_t = \langle\theta, X_t\rangle$. Then Y is an $\mathbb{R}$-valued continuous local martingale and $[Y]_t = |\theta|^2 t$. Let

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

By Itô's formula applied to the function $f(x, y) = \exp(ix + \frac{1}{2}[Y]_t)$,

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

Therefore Z_t is a local martingale. Moreover, Z is bounded on $[0, t]$ for all $t \geq 0$, so it is a martingale by 1.3.4. Hence $\mathbb{E}[Z_t \mid \mathscr{F}_s] = Z_s$, which implies the result. □

4.1.2 Proposition. *Let B be a standard Brownian motion and let $h : \mathbb{R}^+ \to \mathbb{R}$ be a square-integrable with respect to Lebesgue measure. If $X = \int_0^\infty h_s dB_s$ then $X \sim N(0, \int_0^\infty h^2(s) ds)$.*

PROOF: Let $M_t = \int_0^t h_s dB_s$, so that $M \in \mathscr{M}_{c,loc}$ and, since $[h \cdot B] = h^2[B]$, $[M]_t = \int_0^t h_s^2 ds$. Now let

$$Z_t = \exp(iuM_t + \tfrac{1}{2}u^2[M]_t),$$

which again is a local martingale and uniformly bounded by

$$\exp\left(\frac{1}{2}u^2 \int_0^\infty h^2(s)ds\right) < \infty.$$

Therefore Z is a martingale and

$$1 = \mathbb{E}[Z_0] = \mathbb{E}[Z_\infty] = \mathbb{E}[\exp(iuX)] \exp\left(\frac{1}{2}u^2 \int_0^\infty h^2(s)ds\right).$$

□

4.1.3 Theorem (Dubins-Schwarz). *Let $M \in \mathscr{M}_{c,loc}$ be such that $M_0 = 0$ and $[M]_\infty = \infty$ a.s., and set $\tau_s = \inf\{t \geq 0 : [M]_t > s\}$. Then τ_s is a stopping time and $[M]_{\tau_s} = s$ for all $s \geq 0$. Moreover, if $B_s := M_{\tau_s}$ and $\mathscr{G}_s := \mathscr{F}_{\tau_s}$ then $(B_s, s \geq 0)$ is a $(\mathscr{G}_s, s \geq 0)$ Brownian motion and $M_t = B_{[M]_t}$.*

Remark. It follows that any continuous local martingale is a continuous time change of a Brownian motion.

PROOF: Since $[M]$ is continuous and adapted, τ_s is a stopping time, and since $[M]_\infty = \infty$, $\tau_s < \infty$ a.s. for all $s \geq 0$. B is adapted to $(\mathscr{G}_t)$ by Proposition 4.1.1 in Advanced Probability.

We first show that B is continuous. Notice that $s \mapsto \tau_s$ is cadlag and increasing, and M is continuous, so $B_s = M_{\tau_s}$ is also right continuous. We need to show that $B_{s^-} = B_s$, i.e. that $M_{\tau_{s^-}} = M_{\tau_s}$, where

$$\tau_{s^-} = \inf\{t \geq 0 : [M]_t = s\}.$$

By localization we may assume that $M \in \mathscr{M}_c^2$. Then

$$\mathbb{E}[M_{\tau_s}^2 - [M]_{\tau_s} \mid \mathscr{F}_{\tau_s}] = M_{\tau_{s^-}}^2 - [M]_{\tau_{s^-}},$$

and since

$$\mathbb{E}[M_{\tau_s}^2 - M_{\tau_{s^-}}^2 \mid \mathscr{F}_{\tau_{s^-}}] = \mathbb{E}[(M_{\tau_s} - M_{\tau_{s^-}})^2 \mid \mathscr{F}_{\tau_{s^-}}] = 0,$$

we have $M_{\tau_{s^-}} = M_{\tau_s}$ a.s. and B is left-continuous as well.

Fix $s > 0$ and consider $[M^{\tau_s}]_\infty = [M]_{\tau_s} = s$. Therefore by Example Sheet 1, problem 11, $M^{\tau_s} \in \mathscr{M}_c^2$ since $\mathbb{E}[[M^{\tau_s}]_\infty] < \infty$. By 2.3.4, $(M^2 - [M])^{\tau_s}$ is a u.i. martingale. By the OST, for $r \leq s$,

$$\mathbb{E}[B_s \mid \mathscr{G}_r] = \mathbb{E}[M_\infty^{\tau_s} \mid \mathscr{F}_{\tau_r}] = M_{\tau_r} = B_r.$$

Further,

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

Therefore $B \in \mathscr{M}_c$ with $[B]_s = s$, so B is a Brownian motion adapted to $(\mathscr{G}_s, s \geq 0)$.

□

4.2 Exponential martingales

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

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

so $Z \in \mathscr{M}_{c,loc}$. Z is the *stochastic exponential* of M, denoted by $\mathscr{E}(M)$.

4.2.2 Proposition. *Let $M \in \mathscr{M}_{c,loc}$ with $M_0 = 0$. Then for all $\varepsilon, \delta > 0$,*

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

This is the exponential martingale inequality.

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

$$Z_t = \exp(\theta M_t^T - \tfrac{1}{2}\theta^2[M]_t^T) = \mathscr{E}(\theta M^T),$$

a continuous local martingale. Further, $|Z| \leq e^{\theta\varepsilon}$ for all $t \geq 0$, so $Z \in \mathscr{M}_c^2$. By the OST, $\mathbb{E}[Z_\infty] = \mathbb{E}[Z_0] = 1$, so for $\delta > 0$ we get

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

Optimizing with respect to θ, take $\theta = \frac{\varepsilon}{\delta}$ to give the result. □

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

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

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

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

$$\mathbb{E}[\exp(\sup_{t\geq 0} M_t)] = \int_0^\infty \mathbb{P}(\sup_{t\geq 0} M_t \geq \log u)du \leq 1 + \int_1^\infty e^{-\frac{(\log u)^2}{2C}} du < \infty$$

so $\mathscr{E}(M)$ is u.i., and by 1.3.4, $\mathscr{E}(M)$ is a martingale. □

Recall that for two probability measures $\mathbb{P}_1$ and $\mathbb{P}_2$ on $(\Omega, \mathscr{F})$, $\mathbb{P}_1 \ll \mathbb{P}_2$ (i.e. $\mathbb{P}_1$ is *absolutely continuous* with respect to $\mathbb{P}_2$) if $\mathbb{P}_2(A) = 0$ implies $\mathbb{P}_1(A) = 0$ for all $A \in \mathscr{F}$. In this case there is $f : \Omega \to [0, \infty)$ measurable (i.e. the *Radon-Nikodym derivative* of $\mathbb{P}_1$ with respect to $\mathbb{P}_2$) such that $\mathbb{P}_1 = f \cdot \mathbb{P}_2$.

4.2.4 Theorem (Girsanov). *Let $M \in \mathscr{M}_{c,loc}$ with $M_0 = 0$ and suppose that $Z = \mathscr{E}(M)$ is a u.i. martingale. Then $\tilde{\mathbb{P}}(A) := \mathbb{E}[Z_\infty \mathbf{1}_A]$ is a probability measure on $(\Omega, \mathscr{F})$ and $\tilde{\mathbb{P}} \ll \mathbb{P}$ (with density Z_∞). Moreover, if $X \in \mathscr{M}_{c,loc}(\mathbb{P})$ then $X - [X, M] \in \mathscr{M}_{c,loc}(\tilde{\mathbb{P}})$.*

PROOF: Z is a u.i. martingale, so Z_∞ exists and $\mathbb{E}[Z_\infty] = \mathbb{E}[Z_0] = 1$, and $Z_t \geq 0$ for all $t \geq 0$. Therefore $\tilde{\mathbb{P}}(\Omega) = 1$, $\tilde{\mathbb{P}}(\emptyset) = 0$, and countable additivity follows by linearity of $\mathbb{E}$ and bounded convergence. Hence $\tilde{\mathbb{P}}$ is a probability measure, and $\tilde{\mathbb{P}} \ll \mathbb{P}$ since $\tilde{\mathbb{P}}(A) = \int_A Z_\infty d\,\mathbb{P} = 0$ if $\mathbb{P}(A) = 0$.

Let $X \in \mathscr{M}_{c,loc}$ and set $T_n := \inf\{t \geq 0 : |X_t - [M,X]_t| \geq n\}$. Since $X - [M,X]$ is continuous, $\mathbb{P}(T_n \nearrow \infty) = 1$, so $\tilde{\mathbb{P}}(T_n \nearrow \infty) = 1$. Therefore it is enough to show that $Y^{T_n} := X^{T_n} - [X^{T_n}, M] \in \mathscr{M}_c(\tilde{\mathbb{P}})$ for all $n \in \mathbb{N}$. Replace Y by Y^{T_n} and X by X^{T_n} in what follows. By the integration by parts formula,

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

so $ZY \in \mathscr{M}_{c,loc}(\mathbb{P})$. Also, $\{Z_T : T \text{ is a stopping time}\}$ is u.i., so since Y is bounded (by n) $\{Z_T Y_T : T \text{ is a stopping time}\}$ is u.i. Hence $ZY \in \mathscr{M}_c(\mathbb{P})$, so for $s \leq t$,

$$\tilde{\mathbb{E}}[Y_t - Y - s \mid \mathscr{F}_s] = \mathbb{E}[Z_\infty (Y_t - Y_s) \mid \mathscr{F}_s] = \mathbb{E}[Z_t Y_t - Z_s Y_s \mid \mathscr{F}_s] = 0. \qquad \square$$

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

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

PROOF: $\tilde{B} \in \mathscr{M}_{c,loc}(\tilde{\mathbb{P}})$ by Girsanov's Theorem, and $[\tilde{B}]_t = [B] = t$, so by Lévy's characterization $\tilde{B}$ is a Brownian motion. $\square$

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

4.2.6 Definition. Define the *Cameron-Martin space*

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

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

4.2.7 Theorem. *Fix $h \in H$ and set $\mu^h(A) := \mu(\{w \in W : w + h \in A\})$ for all $A \in \mathscr{W}$. Then μ^h is a probability measure on $(W, \mathscr{W})$ and $\mu^h \ll \mu$ with density*

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

for μ-a.a. $w \in W$.

Remark. So if we take Brownian motion and shift it by a deterministic function then its law is absolutely continuous with respect to the original law of the Brownian motion.

PROOF: Set $\mathscr{W}_t = \sigma(X_s : s \leq t)$ and $M = \int_0^t \dot{h}(s) dX_s$. Then $M \in \mathscr{M}_c^2$ on $(W, \mathscr{W}, (\mathscr{W}_t, t \geq 0), \mu)$, and

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

By 4.2.3, $\mathscr{E}(M)$ is u.i., so we can define a probability measure $\tilde{\mu} \ll \mu$ on $(W, \mathscr{W})$ by

$$\begin{aligned} \frac{d\tilde{\mu}}{d\mu}(w) &= \exp(M_\infty(w) - \tfrac{1}{2}[M]_\infty(w)) \\ &= \exp\left(\int_0^\infty \dot{h}(s) dw(s) - \frac{1}{2} \int_0^\infty |\dot{h}(s)|^2 ds \right). \end{aligned}$$

$\tilde{X} := X - [X, M]$ by Girsanov's Theorem is a continuous local martingale with respect to $\tilde{\mu}$ since X is a Brownian motion, by 4.2.5, $\tilde{X}$ is a $\tilde{\mu}$ Brownian motion. But

$$[X, M] = \int_0^t \tilde{h}(s) ds = h(t)$$

(since $dM_t dX_t = \dot{h} dX_t dX_t$?) and so $\tilde{X}(w) = X(w) - h = w - h$. Hence for $A \in \mathscr{W}$,

$$\mu^h(A) = \mu(\{w : X(w) + h \in A\}) = \tilde{\mu}(\{w : \underbrace{\tilde{X}(w) + h}_{X(w)=w} \in A\}) = \tilde{\mu}(A).$$

Therefore $\mu^h = \tilde{\mu}$ and we are done. □

5 Stochastic Differential Equations

Equations of the form

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

(or in differential form $dX_t = b(X_t)dt + \sigma(X_t)dB_t$) arise in many areas, and they are the topic of this chapter.

5.1 General definitions

Suppose that $\sigma : \mathbb{R}^d \to \mathbb{R}^{d \times m}$ and $b : \mathbb{R}^d \to R^d$ are measurable functions that are bounded on compact sets, with $\sigma(x) = (\sigma_{ij}(x))_{i=1,\ldots,d; j=1,\ldots,m}$ and $b(x) = (b_i(x))_{i=1,\ldots,d}$. The SDE

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

written in components is

$$dX_t^i = b_i(X_t)dt + \sum_{j=1}^m \sigma_{ij}(X_t) dB_t^j \tag{2}$$

for $i = 1, \ldots, d$. A *solution to the SDE* (2) is

(i) a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t, t \geq 0), \mathbb{P})$ satisfying the usual conditions;

(ii) an $(\mathcal{F}_t, t \geq 0)$ Brownian motion $B = (B^1, \ldots, B^m)$; and

(iii) an $(\mathcal{F}_t, t \geq 0)$ adapted continuous process $(X^1, \ldots, X^d)$ such that

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

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

5.1.1 Definition.

(i) We say that an SDE has a *weak solution* if for all $x \in \mathbb{R}^d$, there exists a solution to the SDE started from x.

(ii) We say that a solution X of an SDE started from x is a *strong solution* if X is adapted to the natural filtration generated by B.

(iii) There is *uniqueness in law* for an SDE if all solutions to the SDE started from x have the same distribution.

(iv) There is *pathwise uniqueness* for an SDE if for a given $(\Omega, \mathcal{F}, (\mathcal{F}_t, t \geq 0), \mathbb{P})$ and a given Brownian motion B, any two solutions X and X' satisfying $X_0 = X'_0$ a.s. are such that $\mathbb{P}(X_t = X'_t \text{ for all } t) = 1$.

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

5.1.2 Example. Suppose that β is a Brownian motion with $\beta_0 = x$. Then $\text{sgn}\beta_s$ is previsible by Example Sheet 1, problem 3, so

$$B_t := \int_0^t \text{sgn}\beta_s d\beta_s \in \mathcal{M}_{c,loc}.$$

Further, $[B]_t = t$ so B is a Brownian motion started from 0. On the other hand,

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

so β is a solution started from x of the SDE $dX_t = \text{sgn}(X_t) dB_t$. In general, $X \in \mathcal{M}_{c,loc}$ and $[X] = t$, so we have uniqueness in law. But in general there is no pathwise uniqueness, e.g. β and $-\beta$ are both solutions at $x = 0$ but $\mathbb{P}(\beta_t = -\beta_t \text{ for all } t) = 0$. It turns out that β is not a strong solution to this SDE (see Example Sheet 3).

5.2 Lipschitz coefficients

For $U \subseteq \mathbb{R}^d$ and $f : U \to \mathbb{R}^d$, we say that f is a *Lipschitz function* with *L-constant* K if for all $x, y \in U$, $|f(x) - f(y)| \le K|x - y|$, where $|\cdot|$ denotes the Euclidean norm. If $f : U \to \mathbb{R}^{d\times m}$ we use

$$|f| := \left(\sum_{i=1}^{d}\sum_{j=1}^{m} f_{ij}^2\right)^{\frac{1}{2}}.$$

We write

$$\mathscr{C}_T := \{X : [0,T] \to \mathbb{R} \text{ continuous, adapted, } |||X|||_T = \|\sup_{t\le T}|X_t|\|_2 < \infty\}$$

and $\mathscr{C} = \mathscr{C}_\infty$. Recall from 2.2.2 that $(\mathscr{C}_T, |||\cdot|||_T)$ is complete.

5.2.1 Theorem (Contractive Mapping Theorem). *Let (X, d) be a complete metric space and $F : X \to X$. Suppose that F^n is a contraction for some $n \in \mathbb{N}$, (i.e. F^n is Lipshitz with L-constant strictly less than one). Then F has a unique fixed point.*

5.2.2 Lemma (Gronwall). *Let $T > 0$ and let f be a non-negative, bounded, measurable function on $[0,T]$. Suppose that there are $a, b \ge 0$ such that for all $t \in [0,T]$,*

$$f(t) \le a + b\int_0^t f(s)ds.$$

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

5.2.3 Theorem. *Suppose that $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are Lipschitz. Then the SDE*

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

has pathwise uniqueness. Moreover, for each $(\Omega, \mathscr{F}, (\mathscr{F}_T, t \ge 0), \mathbb{P})$ and each $(\mathscr{F}_t)$-Brownian motion B there exists a strong solution started from x, for all $x \in \mathbb{R}^d$.

Proof: We prove the case $d = m = 1$. Fix $(\Omega, \mathscr{F}, (\mathscr{F}_T, t \ge 0), \mathbb{P})$ and a Brownian motion B with natural filtration $(\mathscr{F}_t^B, t \ge 0)$ such that $\mathscr{F}_t^B \subseteq \mathscr{F}_t$. Suppose that K is an L-constant for β and σ.

Uniqueness. Suppose that X and X' are two solutions on Ω such that $X_0 = X_0'$ a.s. Fix $M > 0$ and let $\tau = \inf\{t \ge 0 : |X_t| \vee |X_t'| \ge n\}$. Then

$$X_{t\wedge\tau} = X_0 + \int_0^{t\wedge\tau} \sigma(X_s)dB_s + \int_0^{t\wedge\tau} b(X_s)ds,$$

and similarily for X'. Let $T > 0$. Recall the inequality $(a+b)^2 \le 2a^2 + 2b^2$, which follows from the AM-GM inequality. For $0 \le t \le T$, we have

$$
\begin{aligned}
f(t) &:= \mathbb{E}[(X_{t\wedge\tau} - X'_{t\wedge\tau})^2] \\
&\le 2\,\mathbb{E}\left[\left(\int_0^{t\wedge\tau} \sigma(X_s) - \sigma(X'_s) dB_s\right)^2\right] + 2\,\mathbb{E}\left[\left(\int_0^{t\wedge\tau} b(X_s) - b(X'_s) ds\right)^2\right] \\
&\le \underbrace{2\,\mathbb{E}\left[\int_0^{t\wedge\tau} (\sigma(X_s) - \sigma(X'_s))^2 ds\right]}_{\text{by the Itô isometry}} + \underbrace{2T\,\mathbb{E}\left[\int_0^{t\wedge\tau} (b(X_s) - b(X'_s))^2 ds\right]}_{\text{by Cauchy-Schwarz}} \\
&\le 2K^2(1+T)\,\mathbb{E}\left[\int_0^{t\wedge\tau} (X_s - X'_s)^2 ds\right] \\
&\le 2K^2(1+T)\int_0^t \mathbb{E}[(X_{s\wedge\tau} - X'_{s\wedge\tau})^2] ds \\
&= 0 + 2K^2(1+T)\int_0^t f(s) ds
\end{aligned}
$$

Since $f(t)$ is bounded by $4M^2$, by the magic of Gronwall's Lemma, $f(t) = 0$ for all $t \in [0,T]$. Whence $X_{t\wedge\tau} = X'_{t\wedge\tau}$ a.s., and letting $M, T \to \infty$, we obtain $X = X'$ a.s.

Existence of a strong solution. Note that by Lipschitz-ness,

$$|\sigma(y)| \le |\sigma(0)| + K|y| \qquad \text{and} \qquad |b(y)| \le |b(0)| + K|y|.$$

Suppose that $X \in \mathscr{C}_T$ for some T. Let

$$M_t = \int_0^t \sigma(X_s) dB_s$$

for $t \in [0,T]$. Then $[M]_T = \int_0^T \sigma(X_s)^2 ds$ and so

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

By Example Sheet 1, problem 11, $(M_t)_{0\le t\le T}$ is a martingale bounded in L^2. By Doob's L^2 inequality,

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

By the Cauchy-Schwarz inequality,

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

Define $F : \mathscr{C}_T \to \mathscr{C}_T$ by

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

For $X, Y \in \mathscr{C}_T$, we have

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

Then for all $n \in \mathbb{N}$,

$$\begin{aligned} |||F^n(X) - F^n(Y)|||_T &\le C_T \int_0^T |||F^{n-1}(X) - F^{n-1}(Y)|||_{t_{n-1}}^2 \, dt_{n-1} \\ &\le C_T^n \int_0^T \int_0^{t_{n-1}} \cdots \int_0^{t_2} \underbrace{\int_0^{t_1} |||X - Y|||_{t_0}^2 \, dt_0}_{\le t_0 |||X-Y|||_T^2} \, dt_1 \cdots dt_{n-1} dt_n \\ &\le |||X - Y|||_T^2 C_T^n \frac{T^n}{n!} \to 0 \qquad \text{as } n \to \infty, \end{aligned}$$

so F^n is a contraction for sufficiently large n. Hence, since $(\mathscr{C}_T, |||\cdot|||_T)$ is complete, F has a unique fixed point $X^{(T)} \in \mathscr{C}_T$. By uniqueness, $X_t^{(T)} = X_t^{(T')}$ for all $t \le T \wedge T'$ a.s. and so we can consistently define $X_t = X_t^{(N)}$ for $t \le N$ ($N \in \mathbb{N}$). It remains to show that X is $(\mathscr{F}_t^B, t \ge 0)$-adapted. Define $(Y^n, n \ge 0)$ in $\mathscr{C}_T$ by $Y^0 \equiv x$ and $Y^n := F(Y^{n-1})$ for $n \ge 1$. Then Y^n is $(\mathscr{F}_t^B, t \ge 0)$-adapted by construction, for each $n \ge 0$. Since $X = F^n(X)$ for all $n \in \mathbb{N}$ we have

$$|||X - Y^n|||_T^2 \le C_T^n \frac{T^n}{n!} |||X - x|||_T^2.$$

Hence

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

which implies that

$$\sum_{n=0}^{\infty} \sup_{t \le T} |X_t - Y_t^n| < \infty \text{ a.s.},$$

and $Y^n \to X$ a.s. as $n \to \infty$ uniformly on $[0, T]$. Hence X is also $(\mathscr{F}_t^B, t \ge 0)$-adapted. □

5.2.4 Proposition. *Under the hypotheses of of 5.2.3, the SDE*

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

has uniqueness in law.

PROOF: See Example Sheet 3, problem 3. □

5.2.5 Example. Fix $\lambda \in \mathbb{R}$ and consider the SDE in $\mathbb{R}^2$

$$\begin{aligned} dV_t &= dB_t - \lambda V_t dt, & V_0 &= v_0 \\ dX_t &= V_t dt, & X_0 &= x_0 \end{aligned}$$

V is called the *Ornstein-Uhlenbeck process* (or *velocity process*). By Itô's formula,

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

and so

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

This is the pathwise unique strong solution of 5.2.3. By 4.1.2,

$$V_t \sim \mathcal{N}(v_0 e^{-\lambda t}, \frac{1}{2\lambda}(1 - e^{-2\lambda t})).$$

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

5.3 Local solutions

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

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

It is *cadlag* if $[0, \zeta(\omega)) \to \mathbb{R} : t \mapsto X_t(\omega)$ is cadlag for all $\omega \in \Omega$, and *adapted* if $X_t : \Omega_t \to \mathbb{R}$ is $\mathcal{F}_t$-measurable for all $t \geq 0$, where $\Omega_t = \{\omega : t < \zeta(\omega)\}$.

We say that (X, ζ) is a *locally defined continuous local martingale* if there is a sequence of stopping times $T_n \nearrow \zeta$ a.s. such that X^{T_n} is a continuous martingale for all $n \in \mathbb{N}$. Say that (H, η) is a *locally defined locally bounded previsible process* if there is a sequence of stopping times $S_n \to \eta$ a.s. such that $H\mathbf{1}_{(0,S_n]}$ is bounded and previsible for all $n \in \mathbb{N}$. Finally, define $((H \cdot X), \zeta \wedge \eta)$ via

$$(H \cdot X)^{S_n \wedge T_n} = ((H\mathbf{1}_{(0,S_n]}) \cdot X^{S_n \wedge T_n})$$

for all $n \in \mathbb{N}$.

5.3.2 Proposition (Local Itô formula). *Let $X = (X^1, \dots, X^d)$ be a vector of continuous semimartingales. Let $U \subseteq \mathbb{R}^d$ be open and let $f \in C^2(U, \mathbb{R})$. Set $\zeta = \inf\{t \geq 0 : X \notin U\}$. Then for all $t < \zeta$,*

$$f(X_t) = f(X_0) + \langle Df(X_t), dX_t\rangle + \langle \tfrac{1}{2}D^2 f(X_t)dX_t, dX_t\rangle.$$

PROOF: Apply Itô's formula to X^{T_n}, where

$$T_n = \inf\{t \geq 0 : |X_t - y| \leq \frac{1}{n} \text{ for some } y \notin U\}$$

and use $T_n \nearrow \zeta$ a.s. □

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

$$\sqrt{B_t} = 1 + \int_0^t \tfrac{1}{2}B_s^{-\frac{1}{2}}dB_s + \int_0^t (-\tfrac{1}{4}B_s^{-\frac{3}{2}})ds.$$

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

5.3.4 Definition. Let $U \subseteq \mathbb{R}^d$ be open and let $\sigma : U \to \mathbb{R}^{d\times m}$, and $b : U \to \mathbb{R}^d$ be measurable and bounded on compact subsets of U. A *local solution* to the SDE $dX_t = \sigma(X_t)dB_t + b(X_t)dt$ is

(i) a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t, t \geq 0), \mathbb{P})$ satisfying the usual conditions;

(ii) an $(\mathscr{F}_t)$-Brownian motion $B = (B^1, \dots, B^m)$; and

(iii) an $(\mathscr{F}_t)$-adapted, continuous, locally defined process (X, ζ), where X has values in U and

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

for all $t < \zeta$, a.s.

We say that (X, ζ) is a *maximal solution* if (X', ζ') satisfies (iii) and $X_t = X'_t$ for all $t < \zeta \wedge \zeta'$ then $\zeta' \leq \zeta$.

5.3.5 Definition. We say that b is *locally Lipschitz* in U if for all compact sets $C \subseteq U$ there exists $K_C < \infty$ such that $|f(x) - f(y)| \leq K_C|x - y|$ for all $x, y \in C$.

Notice in particular that if b is continuously differentiable then b is locally Lipschitz.

5.3.6 Theorem. *Let $(\Omega, \mathscr{F}, (\mathscr{F}_t, t \geq 0), \mathbb{P})$ and B be given, $X_0 \in \mathbb{R}^d$, and assume that $\sigma_1, \dots, \sigma_m, b$ are locally Lipschitz. There exists a unique maximal local solution (X, ζ) starting from X_0. Moreover, for all compacts $C \subseteq U$,* $\sup\{t < \zeta : X_t \in C\} < \zeta$ *a.s. on $\{\zeta < \infty\}$.*

The last assertion can be interpreted as saying that if the solution blows up in a finite time then at least it will not return to any compact set infinitely often as it approaches that time.

5.3.7 Lemma. *Given $C \subseteq U$ compact, set $C_n = \{x \in \mathbb{R}^d : d(x, C) \leq \frac{1}{n}\}$. By compactness, $C_n \subseteq U$ for some n. There exists a C^∞ p.d.f. ψ with $\psi(x) = 0$ for all $|x| \geq \frac{1}{2n}$. Set*

$$\phi(x) = \int_{\mathbb{R}^d} \psi(x - y)\mathbf{1}_{C_{2n}}(y)dy = \varphi * \mathbf{1}_{C_{2n}}(x).$$

Then ϕ is C^∞, $\phi \equiv 1$ on C and $\phi \equiv 0$ off C_n.

Hence if b is locally Lipschitz then $\tilde{b} = \phi b$ is Lipschitz and $\tilde{b} = b$ on C.

Proof (of 5.3.6): We prove the case $d = m = 1$. Fix $C \subseteq U$ compact and choose $\tilde{\sigma}, \tilde{b}$ Lipschitz with $\tilde{\sigma} = \sigma$ and $\tilde{b} = b$ on C. By 5.2.3, there is a solution $\tilde{X}$ to

$$d\tilde{X}_t = \tilde{\sigma}(\tilde{X}_t)dB_t + \tilde{b}(\tilde{X}_t)dt, \quad \tilde{X}_0 = X_0.$$

Set $T = \inf\{t \geq 0 : \tilde{X}_t \notin C\}$ and let X be the restriction of $\tilde{X}$ to $[0, T)$. Then (X, T) is a local solution with values in C and $X_{T^-} = \tilde{X}_T$ exists in ∂C if $T < \infty$.

Suppose that (X', T') is another local solution in C. Consider for $t \geq 0$

$$f(t) = \mathbb{E}[\sup_{s<t\wedge T\wedge T'} |X'_s - X_s|^2].$$

Then $f(t) < \infty$ for all $t_0 \geq 1$, and as in the proof of 5.2.3,

$$f(t) \leq 10K_C^2 t_0 \int_0^t f(s)ds.$$

Therefore by Gronwall's lemma $f \equiv 0$. Hence $X_t = X'_t$ for all $t < T \wedge T'$ a.s. If $X'_{T'-}$ exists in ∂C on $\{T' < \infty\}$ then this forces $T = T'$ a.s.

Take compacts $C_n \nearrow U$ and a sequence of local solutions (X^n, T_n), with X_n taking values in C_n and $X^n_{T_n^-} \in \partial C_n$ a.s. on $\{T_n < \infty\}$. Then $T_n \leq T_{n+1}$ and $X^n_t = X^{n+1}_t$ for $t < T_n$ a.s. (exercise). Therefore, set $\zeta = \sup T_n$ and $X_t = X^n_t$ for $t < T_n$. Then (X, ζ) is a local solution and $X_t \to \partial U \cup \{\infty\}$ as $t \nearrow \zeta$ a.s. on $\{\zeta < \infty\}$. Further, (X, ζ) is maximal (exercise, an easy consequence of uniqueness on compacts).

Given $C \subseteq U$ compact, choose $C' \subseteq U$ also compact with the property that $C \subseteq C'^{\circ}$, and $\varphi : U \to \mathbb{R}$ twice continuously differentiable with $\varphi \equiv 1$ on C and $\varphi \equiv 0$ on $U \setminus C'$. Set $R_1 = \inf\{t < \zeta : X_t \notin C'\}$ and recursively define, for $n \geq 1$, $S_n = \inf\{t \in [R_n, \zeta) : X_t \in C\}$ and $R_{n+1} = \inf\{t \in [S_n, \zeta) : X_t \notin C'\}$. Let N be the number of crossings by X from C' to C. On $\{\zeta \leq t, N \geq n\}$, we have

$$\begin{aligned} n &= \sum_{k=1}^n (\varphi(X_{S_k}) - \varphi(X_{R_k})) \\ &= \int_0^t \sum_{k=1}^n \mathbf{1}_{(R_k, S_k]}(s)(\varphi'(X_s)dX_s + \tfrac{1}{2}\varphi''(X_s)dX_s dX_s) \\ &=: \int_0^t (\xi^n_s dB_s + \eta^n_s ds) =: Z^n_t \end{aligned}$$

where ξ^n, η^n are previsible and bounded uniformly in n. Now

$$n^2 \mathbf{1}_{\{\zeta \leq t, N \geq n\}} \leq (Z^n_t)^2$$
$$\text{so } n^2\, \mathbb{P}(\zeta \leq t, N \geq n) \leq \mathbb{E}(Z^n_t)^2 \leq C < \infty$$

so $\mathbb{P}(\zeta \leq t, N \geq n)$ decreases at least as fast as $\frac{1}{n^2}$. Therefore

$$\sum_{n=1}^\infty \mathbb{P}(\zeta \leq t, N \geq n) < \infty \qquad \text{so} \qquad \mathbb{P}(\zeta \leq t, N = \infty) = 0$$

by the Borel-Cantelli lemma. Whence $\mathbb{P}(\zeta < \infty, N = \infty) = 0$ and $\sup\{t < \zeta : X_t \in C\} < \zeta$ a.s. on $\{\zeta < \infty\}$. □

6 Diffusion Processes

6.1 L-diffusions

6.1.1 Definition. For $f \in C^2(\mathbb{R}^d)$ define

$$Lf(x) := \frac{1}{2} \sum_{i,j=1}^d a^{ij}(x) \frac{\partial^2}{\partial x^i \partial x^j} f(x) + \sum_{i=1}^d b^i(x) \frac{\partial}{\partial x^i} f(x),$$

where $a : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are bounded measurable functions. We say a is the *diffusivity* and b is the *drift*.

We always assume that a is a non-negative-definite symmetric matrix. We will work in the context of a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t, t \geq 0), \mathbb{P})$. Let $X = (X_t, t \geq 0)$ be a continuous random process. We say that X is an *L-diffusion* if, for all $f \in C_b^2(\mathbb{R}^d)$, the following process is a martingale.

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

Assume that exists an m-dimensional $(\mathscr{F}_t)$-Brownian motion B on the f.p.s. above.

6.1.2 Proposition. *Let X be a solution to the SDE*

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

with bounded measurable coefficients. Let $f \in C_b^{1,2}(\mathbb{R}^+ \times \mathbb{R}^d)$. Then the following process is a martingale.

$$M_t^f = f(t,X_t) - f(0,X_0) - \int_0^t (\dot{f} + Lf)(s,X_s)ds,$$

where we take $a = \sigma\sigma^T$. In particular, X is an L-diffusion.

Remark. Take $m = d$ and suppose that $a : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ are Lipschitz and a is *uniformly positive-definite*, i.e. $\xi^T a(x)\xi \geq \varepsilon|\xi|^2$ for all $\xi, x \in \mathbb{R}^d$, for some $\varepsilon > 0$. Then the positive definite square root $\sigma = \sqrt{a}$ is also Lipschitz. So the SDE above has a solution with this σ and an L-diffusion exists, for any starting point X_0.

PROOF: By Itô's formula, (using the sum convention)

$$\begin{aligned} df(t,X_t) &= \dot{f}(t,X_t)dt + \frac{\partial}{\partial x_i} f(t,X_t)dX_t^i + \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_j} f(t,X_t)dX_t^i dX_t^j \\ &= \dot{f}(t,X_t)dt + \frac{\partial}{\partial x_i} f(t,X_t)\sigma_{ij}(X_t)dB_t^j \\ &\qquad + \underbrace{\frac{\partial}{\partial x_i} f(t,X_t)b_i(X_t)dt + \frac{1}{2}\frac{\partial^2}{\partial x_i \partial x_j} f(t,X_t)\overbrace{\sigma_{ik}dB_t^k \sigma_{j\ell}dB_t^\ell}^{\sigma_{ik}\sigma_{j\ell}\delta_{k\ell}dt}}_{Lf(t,X_t)dt} \end{aligned}$$

so

$$df(t,X_t) - (\dot{f} + Lf)(t,X_t)dt = \frac{\partial}{\partial x_i} f(t,X_t)\sigma_{ij}(X_t)dB_t^j.$$

But $\sup_{s\leq t}|M_s^f| \leq Ct$ for all $t \geq 0$ since f, b, and σ (and hence a) are bounded, so M^f is a martingale. □

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

PROOF: Consider, for $f \in C_b^2(\mathbb{R}^d)$,

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

Clearly $\tilde{M}^f$ is $(\tilde{\mathscr{F}}_t)$ adapted and integrable. For $A \in \tilde{\mathscr{F}}_s$ and $n \geq 0$, we have $A \cap \{T \leq n\} \in \mathscr{F}_{(T\wedge n)+s}$ and

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

by the OST. Let $n \to \infty$ and use dominated convergence to see $\tilde{M}^f$ is a martingale. □

6.1.4 Lemma. *Let X be an L-diffusion and let $f \in C_b^{1,2}(\mathbb{R}^+ \times \mathbb{R}^d)$. Then the following process is a martingale.*

$$M_t^f := f(t, X_t) - f(0, X_0) - \int_0^t (\dot{f} + Lf)(s, X_s)ds,$$

PROOF: Fix $T > 0$ and consider

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

Then Z_n is uniformly bounded and $Z_n \to 0$ as $n \to \infty$ (using uniform continuity of continuous functions on compacts). So $\mathbb{E}[Z_n] \to 0$ as $n \to \infty$. For $0 \leq s \leq t \leq T$,

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

Choose $s_0 \leq s_1 \leq \cdots \leq s_m$ with $s_0 = s$, $s_m = t$, and $s_{k+1} - s_k \leq \frac{1}{n}$. Then

$$\mathbb{E}[\mathbb{E}[M^f_{s_{k+1}} - M^f_{s_k} \mid \mathscr{F}_s]] \leq (s_{k+1} - s_k)\mathbb{E}[Z_n]$$

for all $0 \leq k < m$, so

$$\mathbb{E}[\mathbb{E}[M_t^f - M_s^f \mid \mathscr{F}_s]] \leq (t - s)\mathbb{E}[Z_n]$$

for all n. Therefore $\mathbb{E}[M_t^f - M_s^f \mid \mathscr{F}_s] = 0$ a.s. as required. □

Please note that $f \in C_b^{1,2}$ means that f is bounded, $\dot{f}$ is bounded and continuous, and first and second partials in the space coordinates are bounded and continuous.

6.2 Dirichlet and Cauchy problems

For this section we assume that a and b (in the definition of L) are Lipschitz and a is uniformly positive-definite. Take D to be a bounded domain (i.e. an open, connected region) in $\mathbb{R}^d$ with smooth boundary. We accept the following result from the theory of PDEs.

6.2.1 Theorem. *For all $f \in C(\partial D)$ and $\phi \in C(\overline{D})$ there exists a unique $u \in C(\overline{D}) \cap C^2(D)$ such that*

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

Moreover, there exists a continuous function $g : D \times D \setminus diag(D) \to (0, \infty)$ and there exists a continuous kernel $m : D \times \mathscr{B}(\partial D) \to [0, \infty)$ such that

$$u(x) = \int_D g(x,y)\phi(y)dy + \int_{\partial D} m(x,dy)f(y) \qquad \text{for all } x, f, \phi.$$

Remark.

(i) Recall that m is a *kernel* if $m(\cdot, A)$ is a measurable function for all $A \in \mathscr{B}(\partial D)$ and $m(x, \cdot)$ is a measure for all $x \in D$, and it is continuous if $m(\cdot, A)$ is continuous for all A.

(ii) g is called the *Green function* and m is the *harmonic measure*. They depend only on the domain D (and not on f or ϕ).

PROOF: See Bass, *Diffusions and elliptic operators*. □

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

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

with $\phi \in C(\overline{D})$. Then for any L-diffusion X starting from $x \in D$,

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

where $T = \inf\{t \geq 0 : X_t \notin D\}$. In particular, for all $A \in \mathscr{B}(D)$, the expected total time in A is

$$\mathbb{E}_x \int_0^T \mathbf{1}_{X_s \in A} ds = \int_A g(x,y)dy,$$

and for all $B \in \mathscr{B}(\partial D)$, the hitting distribution is

$$\mathbb{P}_x(X_T \in B) = m(x, B).$$

PROOF: Fix $n \in \mathbb{N}$. Set $D_n := \{x \in D : d(x, D^C) > \frac{1}{n}\}$ and $T_n := \inf\{t \geq 0 : X_t \notin D_n\}$. Define

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

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

$$M^{\tilde{u}} = \tilde{u}(X_t) - \tilde{u}(X_0) - \int_0^t L\tilde{u}(X_s)ds.$$

So M is a martingale since X is an L-diffusion and by the OST. Therefore

$$0 = \mathbb{E}_x(M_0) = \mathbb{E}_x(M_t) = \mathbb{E}_x(u(X_{t\wedge T_n})) - u(x) + \mathbb{E}_x \int_0^{t\wedge T_n} \phi(X_s)ds. \qquad (3)$$

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

$$\mathbb{E}_x(T_n \wedge t) = u(x) - \mathbb{E}_x(u(X_{t\wedge T_n})).$$

By 6.2.1, u is bounded (indeed, it is a continuous function on the compact set $\overline{D}$), and $T_n \nearrow T$ a.s., so we deduce by monotone convergence (letting $t = n \to \infty$) that $\mathbb{E}_x(T) < \infty$. Returning to the general case, let $t \to \infty$ and $n \to \infty$ in (3). Since u is continuous on $\overline{D}$, $u(X_{t\wedge T_n}) \to f(X_T)$. Also, $t \wedge T_n \nearrow T$ and

$$\mathbb{E}_x \int_0^T |\phi(X_s)|ds \le \|\phi\|_\infty \mathbb{E}_x(T) < \infty,$$

so by multiple applications of dominated convergence,

$$\mathbb{E}_x \int_0^{t\wedge T_n} \phi(X_s)ds \to \mathbb{E}_x \int_0^T \phi(X_s)ds.$$

We have shown, for all $f \in C(\partial D)$ and $\phi \in C(\overline{D})$,

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

This identity extends to all bounded measurable ϕ and f by a monotone class argument (fill this in). In particular, it extends to the cases with $\phi = \mathbf{1}_A$ and $f = 0$ and with $\phi = 0$ and $f = \mathbf{1}_B$. □

We accept the following from the theory of PDEs.

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

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

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

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

Remark. The problem above is the *Cauchy problem*, and p is called the *heat kernel*.

PROOF: See Bass, *Diffusions and elliptic operators*. □

6.2.4 Theorem. *Suppose that $u \in C_b^{1,2}(\mathbb{R}^+ \times \mathbb{R}^d)$ satifies*

$$\begin{cases} \dot{u} = Lu & \text{on } \mathbb{R}^+ \times \mathbb{R}^d \\ u(0,\cdot) = g & \text{on } \mathbb{R}^d \end{cases}$$

Then for any L-diffusion X, we have $u(t,x) = \mathbb{E}_x(g(X_t))$. Moreover, for $T > t \geq 0$ and all $x \in \mathbb{R}^d$, a.s.

$$\mathbb{E}_x(g(X_T) \mid \mathscr{F}_t) = u(T-t, X_t) = \int_{\mathbb{R}^d} p(T-t, X_t, y) g(y) dy.$$

Hence X is an $(\mathscr{F}_t)$-Markov process with transition density function p.

Proof: Fix T and take $f(t,x) = u(T-t,x)$ in 6.1.4 to see that $M_t = u(T-t, X_t)$, $0 \leq t \leq T$, is a martingale. Hence

$$\mathbb{E}_x(g(X_T) \mid \mathscr{F}_t) = \mathbb{E}(M_T \mid \mathscr{F}_t) =^{a.s.} M_t = u(T-t, X_t). \qquad \square$$

6.2.5 Theorem (Feynman-Kac formula).
Let $f \in C_b^2(\mathbb{R}^d)$ and $V \in L^\infty(\mathbb{R}^d)$. Suppose $u \in C_b^{1,2}(\mathbb{R}^+ \times \mathbb{R}^d)$ satisfies

$$\begin{cases} \dot{u} = \frac{1}{2}\Delta u + Vu & \text{on } \mathbb{R}^+ \times \mathbb{R}^d \\ u(0,\cdot) = f & \text{on } \mathbb{R}^d \end{cases}$$

where Δ is the Laplacian. Let B be an $\mathbb{R}^d$-valued Brownian motion. Then for all t, x,

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

Proof: Fix $T > 0$ and consider $M_t = u(T-t, B_t) \exp \int_0^t V(B_s) ds$. Then by Itô's formula

$$\begin{aligned} dM_t &= \nabla u(T-t, B_t) E_t dB_t + (-\dot{u} + \tfrac{1}{2}\Delta u + Vu)(T-t, B_t) E_t dt \\ &= \nabla u(T-t, B_t) E_t dB_t, \end{aligned}$$

where $E_t := \exp \int_0^t V(B_s) ds$, so M is a local martingale. It is uniformly bounded on $[0,T]$, so M is a martingale. Whence

$$u(T,x) = \mathbb{E}_x(M_0) = \mathbb{E}_x(M_T) = \mathbb{E}_x(f(B_T) E_T). \qquad \square$$

7 Markov Jump Processes

7.1 Definitions and basic example

Let $(E, \mathscr{E})$ be a measurable space. Measurable kernels encode all the data we will need when dealing with jump processes.

7.1.1 Definition. A *measurable kernel* is a function $q : E \times \mathscr{E} \to [0, \infty)$ such that

(i) for all $A \in \mathscr{E}$, $x \mapsto q(x,A) : E \to [0,\infty)$ is measurable; and

(ii) for all $x \in E$, $A \mapsto q(x,A) : \mathscr{E} \to [0,\infty)$ is a measure.

7.1.2 Example. The most important example is the countable case, as follows. Take E to be countable and $\mathscr{E} = 2^E$. A *Q-matrix* is $Q = (q_{xy} : x, y \in E)$, where

(i) $q_{xy} \geq 0$ for all distinct $x, y \in E$; and

(ii) $q_x = -q_{xx} = \sum_{y \neq x} q_{xy} < \infty$ for all $x \in E$.

Define the measurable kernel $q(x,A) = \sum_{y \in A \setminus \{x\}} q_{xy}$ with respect to the Q-matrix Q.

For $f : E \to \mathbb{R}$ bounded and measurable, write

$$Qf(x) := \int_E (f(y) - f(x)) q(x, dy).$$

Let $q(x) := q(x,E)$ and write

$$\pi(x,A) = \begin{cases} \frac{q(x,A)}{q(x)} & \text{if } q(x) > 0 \\ \delta_x(A) & \text{if } q(x) = 0 \end{cases}$$

We always assume that $q(x,\{x\}) = 0$. In the countable case,

$$(Qf)_x = \sum_{y \in E} q_{xy} f_y$$

(This formula works because of how the diagonal terms of a Q-matrix are defined.)

7.1.3 Definition. Say $X = (X_t, t \geq 0)$ is a *(minimal) jump process* if, for some random times $0 = J_0 \leq J_1 \leq \cdots \leq J_n \nearrow \zeta$ with $J_n > J_{n-1}$ if $J_{n-1} < \infty$, and for some E-valued process $(Y_n, n \in \mathbb{N})$,

$$X_t = \begin{cases} Y_n & \text{if } J_n \leq t < J_{n+1} \\ \partial & \text{if } t \geq \zeta \end{cases}$$

Remark.

(i) A *random time* is a random variable in $[0,\infty]$.

(ii) The "minimality" refers to the fact that X is sent to the *cemetary state* ∂ from its *explosion time* ζ onward.

(iii) In the case where $J_n = \infty$, we insist $Y_n = Y_{n-1}$. Then $(Y_n, n \in \mathbb{N})$ and $(J_n, n \in \mathbb{N})$ are determined uniquely by $(X_t, t \geq 0)$ and *visa versa*.

Remark. There are three distinct cases which may arise from this definition.

(i) If $J_n = \infty$ for some $n \in \mathbb{N}$ then X takes finitely many jumps and then gets "stuck." In this case it is reasonable to define $X_\infty = X_{J_n}$, the final value taken.

(ii) If $\zeta < \infty$ then X takes infinitely many jumps in a finite amount time.

(iii) If $\zeta = \infty$ and $J_n < \infty$ for all n then X takes infinitely many jumps in infinite time.

7.1.4 Definition. Fix $(\Omega, \mathscr{F}, (\mathscr{F}_t, t \geq 0), \mathbb{P})$. We say that a jump process X is a *Markov jump process* with generator Q (or *Markov with generator* Q) if X is adapted and for all $s, t \geq 0$ and $B \in \mathscr{E}$,

$$\mathbb{P}(J_1(t) > t + s, Y_1(t) \in B \mid \mathscr{F}_t) = \pi(X_t, B)e^{-q(X_t)s} \text{ a.s.},$$

where $J_1(t) = \inf\{s > t : X_s \neq X_t\}$ is the time of the first jump after t and $Y_1(t) = X_{J_1(t)}$ is the value of that jump. Define $q(\partial) := 0$

7.1.5 Definition. We say that $(Y_n, n \in \mathbb{N})$ is *Markov with transition kernel* π if for all $n \geq 0$ and $B \in \mathscr{E}$,

$$\mathbb{P}(Y_{n+1} \in B \mid Y_0, \dots, Y_n) = \pi(Y_n, B) \text{ a.s.}$$

Given a probability measure λ on $(E, \mathscr{E})$, there is a unique probability measure μ on $(E^{\mathbb{N}}, \mathscr{E}^{\otimes \mathbb{N}})$ such that

(i) $\mu(Y_0 \in B) = \lambda(B)$; and

(ii) $(Y_n, n \in \mathbb{N})$ is Markov with transition kernel π,

where $(Y_n : E^{\mathbb{N}} \to E)$ are the coordinate functions (i.e. $Y_n(y) := y_n$ for $y = (y_0, y_1, \dots)$). We can define μ_n on $\sigma(Y_m : m \leq n)$ by

$$\mu_n((Y_0, \dots, Y_n) \in B) = \int_B \lambda(dy_0) \prod_{i=0}^{n-1} \pi(y_i, dy_{i+1}).$$

The properties (i) and (ii) force any such $\mu = \mu_n$ on $\sigma(Y_m : m \leq n)$, so we get uniqueness of μ by a π-system argument. Existence comes via Daniell's Extension Theorem. (If the space is countable then such sophistication is not necessary.)

7.1.6 Proposition. *Let $(Y_n, n \geq 0)$ be Markov with transition kernel π, and let $(T_n, n \geq 1)$ be a sequence of independent $\mathscr{E}xp(1)$ r.v.'s, independent of $(Y_n, n \geq 0)$. Set*

$$S_n := \frac{1}{q(Y_{n-1})} T_n, \quad J_n := S_1 + \cdots + S_n, \quad J_0 := 0,$$

and write $(X_t, t \geq 0)$ for the associated jump process, and $(\mathscr{F}_t, t \geq 0)$ for its natural filtration. Then $(X_t, t \geq 0)$ is $(\mathscr{F}_t)$-Markov with generator Q.

7.1.7 Lemma. *Set $\mathscr{G}_n := \sigma(Y_m, J_m : m \leq n)$, $n \geq 0$. For all $A \in \mathscr{F}_t$ and $n \geq 0$ there exists $\tilde{A}_n \in \mathscr{G}_n$ with the property that*

$$A \cap \{J_n \leq t < J_{n+1}\} = \tilde{A}_n \cap \{t < J_{n+1}\}.$$

PROOF: Let $\mathscr{A}_t$ denote the set of all $A \in \mathscr{F}_t$ for which the conclusion of the Lemma holds. Then $\mathscr{A}_t$ is a σ-algebra (exercise) and X_s is $\mathscr{A}_t$-measurable for all $s \leq t$ (exercise), hence $\mathscr{A}_t = \mathscr{F}_t$. □

Proof (of 7.1.6): For $s,t\geq 0$, $B\in\mathscr{E}$, $A\in\mathscr{F}_t$, and $n\geq 0$,

$$\begin{aligned}\mathbb{P}(J_1(t)>t+s,Y_1(t)\in B,A,J_n\leq t<J_{n+1})\\ &=\mathbb{P}(J_{n+1}>t+s,Y_{n+1}\in B,\tilde{A}_n)\\ &=\mathbb{E}[\pi(Y_n,B)e^{-q(Y_n)(t+s-J_n)}\mathbf{1}_{\tilde{A}_n}] \qquad \text{since } S_{n+1}\sim\mathscr{E}^{xp}(q(Y_n))\\ &=\mathbb{E}[\pi(Y_n,B)e^{-q(X_n)s}\mathbf{1}_{\tilde{A}_n\cap\{J_{n+1}>t\}}]\\ &=\mathbb{E}[\pi(X_t,B)e^{-q(X_t)s}\mathbf{1}_{A\cap\{J_n\leq t<J_{n+1}\}}]\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}(J_1(t)>t+s,Y_1(t)\in B,A,t\geq\zeta) \text{ “}=\delta_\partial(B)\text{”}\\ &=0=\mathbb{E}[\pi(X_t,B)e^{-q(X_t)s}\mathbf{1}_{A\cap\{t\geq\zeta\}}].\end{aligned}$$

On summing the above,

$$\mathbb{P}(J_1(t)>t+s,Y_1(t)\in B\mid\mathscr{F}_t)=\pi(X_t,B)e^{-q(X_t)s}\text{ a.s.}$$ □

7.1.8 Proposition. *Let $(X_t,t\geq 0)$ be an $(\mathscr{F}_t,t\geq 0)$-Markov jump process with generator Q and let T be a stopping time. Then for all $s\geq 0$ and $B\in\mathscr{E}$, on $\{T<\infty\}$,*

$$\mathbb{P}(J_1(T)>T+s,Y_1(T)\in B\mid\mathscr{F}_T)=\pi(X_T,B)e^{-q(X_T)s}\text{ a.s.}$$

Proof: Consider the sequence of stopping times $(T_m:=2^{-m}\lceil 2^mT\rceil,m\in\mathbb{N})$. Note that $T_m\searrow T$ as $m\to\infty$, so since X is right-continuous,

$$X_{T_m}=X_T,\quad J_1(T_m)=J_1(T),\quad Y_1(T_m)=Y_1(T)$$

eventually as $m\to\infty$, a.s. Let $A\in\mathscr{F}_T$ with $A\subseteq\{T<\infty\}$. For all $k\in\mathbb{Z}^+$, $A\cap\{T_m=k2^{-m}\}\in\mathscr{F}_{k2^{-m}}$, so

$$\begin{aligned}\mathbb{P}(J_1(T_m)>T_m+s,Y_1(T_m)\in B,A,T_m=k2^{-m})\\ &=\mathbb{E}[\pi(X_{T_m},B)e^{-q(X_{T_m})s}\mathbf{1}_{A\cap\{T_m=k2^{-m}\}}]\end{aligned}$$

and summing over k,

$$\mathbb{P}(J_1(T_m)>T_m+s,Y_1(T_m)\in B,A)=\mathbb{E}[\pi(X_{T_m},B)e^{-q(X_{T_m})s}\mathbf{1}_A]$$

Letting $m\to\infty$, we can replace T_m by T (bounded convergence), proving the result. □

7.2 Some martingales

Define random measures μ,ν on $(0,\infty)\times E$ by

$$\mu:=\sum_{t:X_t\neq X_{t^-}}\delta_{(t,X_t)}=\sum_{n=1}^{\infty}\delta_{(J_n,Y_n)}\qquad\text{and}\qquad\nu(dt,dy):=q(X_{t^-},dy)dt.$$

7.2.1 Proposition. *Fix $B \in \mathscr{E}$ and set $M_t = (\mu - \nu)((0,t] \times B)$. Then M is a local martingale on $[0, \zeta)$.*

PROOF: It will suffice to show that M^{J_n} is a martingale for all n. Note that

$$M_t^{J_{n+1}} - M_t^{J_n} = \mathbf{1}_{\{J_{n+1} \le t, Y_{n+1} \in B\}} - \int_0^t \mathbf{1}_{\{J_n \le r < J_{n+1}\}} q(X_{r^-}, B) dr.$$

Take $s \le t$, $A \in \mathscr{F}_s$, and set $\tilde{A} := A \cap \{s < J_{n+1}\}$ and $T := J_n \vee s$. On $\tilde{A}$ we have $J_{n+1} = J_1(T)$ and $Y_{n+1} = Y_1(T)$, so

$$\begin{aligned}\mathbb{P}(s < J_{n+1} \le t, Y_{n+1} \in B, A) &= \mathbb{P}(J_1(T) \le t, Y_1(T) \in B, \tilde{A}) \\ &= \mathbb{E}[\pi(X_T, B) \mathbf{1}_{\tilde{A}} \int_0^{(t-T)^+} q(X_T) e^{-q(X_T) r} dr].\end{aligned}$$

On the other hand,

$$\begin{aligned}\mathbb{E}\left[\mathbf{1}_A \int_s^t \mathbf{1}_{J_n \le r < J_{n+1}} q(X_r, B) dr\right] &= \mathbb{E}\left[\mathbf{1}_{\tilde{A}} q(X_T, B) \int_T^{t \vee T} \mathbf{1}_{J_1(T) > r} dr\right] \\ &= \mathbb{E}\left[\mathbf{1}_{\tilde{A}} q(X_T, B) \int_0^{(t-T)^+} \mathbf{1}_{J_1(T) > T + r} dr\right]\end{aligned}$$

and $\mathbb{P}(J_1(T) > T + r \mid \mathscr{F}_T) = e^{-q(X_T) r}$ a.s. We have shown that

$$\mathbb{E}[\{(M_t^{J_{n+1}} - M_t^{J_n}) - (M_s^{J_{n+1}} - M_s^{J_n})\} \mathbf{1}_A] = 0$$

so $M^{J_{n+1}} - M^{J_n}$ is a martingale. But $M^{J_0} = 0$, so this implies that M^{J_n} is a martingale for all n. □

Recap:

(i) $q(x, dy) = q(x) \pi(x, dy)$, where $q(x)$ is the rate of jumping from x and $\pi(x, dy)$ is the probability of jumping into dy from x.

(ii) For a Markov jump process $(X_t, t \ge 0)$, $X_t = y_n$ for $J_n \le t < J_{n+1}$, $\mu = \sum_{t : X_t \ne X_{t^-}} \delta_{(t, X_t)} = \sum_{n=1}^{\infty} \delta_{(J_n, Y_n)}$ and $\nu(dy, dt) = q(X_{t^-}, dy) dt$ can be thought of as random measures on $(0, \infty) \times E$. Then for $B \in \mathscr{E}$, $(M_t = (\mu - \nu)((0,t] \times B), t \ge 0)$ is a local martingale on $[0, \zeta)$.

What about $\mathbb{P}(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n)$?

7.2.2 Definition. We say that $H : \Omega \times (0, \infty) \times E \to \mathbb{R}$ is *previsible* if it is $\mathscr{P} \otimes \mathscr{E}$ measurable, where $\mathscr{P}$ is the previsible σ-algebra on $\Omega \times (0, \infty)$ for $(\mathscr{F}_t, t \ge 0)$.

7.2.3 Theorem. *Let H be previsible and assume that*

$$\mathbb{E}\left[\int_0^t \int_E |H(s, y)| \nu(ds, dy)\right] < \infty$$

for all $t \ge 0$. Then the following process is a martingale

$$M_t = \int_{(0,t] \times E} H(s, y)(\mu - \nu)(ds, dy).$$

Proof: Define for $C \in \mathscr{P} \otimes \mathscr{E}$, $\overline{\mu}(C) = \mathbb{E}\,\mu(C)$ and $\overline{\nu}(C) = \mathbb{E}\,\nu(C)$. Note that

$$\overline{\mu}(C) = \overline{\mu}(C \cap (0,\zeta)) = \lim_{n\to\infty} \overline{\mu}(C \cap (0, J_n]),$$

where $(0,\zeta) = \{(\omega, t, y) : t < \zeta(\omega), y \in E\})$ and similarly for $\overline{\nu}$. Now $\mathscr{P} \otimes \mathscr{C}$ is generated by the π-system of sets $\{A \times (s,t] \times B : s < t, A \in F_s, B \in \mathscr{E}\}$. Set $M_t := (\mu - \nu)((0,t] \times B)$. By 7.2.1,

$$\overline{\mu}(D \cap (0,J_n]) - \overline{\nu}(D \cap (0,J_n]) = \mathbb{E}[\mathbf{1}_A(M_t^{J_n} - M_s^{J_n})] = 0$$

for such D. So $\overline{\mu}(D \cap (0,J_n]) = \overline{\nu}(D \cap (0,J_n])$ for all $D \in \mathscr{P} \otimes \mathscr{E}$ by uniqueness of extension (by considering the restricted (finite) measures $\overline{\mu}(\cdot \cap (0,J_n])$ and $\overline{\nu}(\cdot \cap (0,J_n]))$. Hence $\overline{\mu} = \overline{\nu}$ on $\mathscr{P} \otimes \mathscr{E}$. For $H \geq 0$ previsible, for $s \leq t$ and $A \in \mathscr{F}_s$,

$$\begin{aligned}
\mathbb{E}\left[\mathbf{1}_A \int_{(s,t]\times E} H(r,y)\mu(dr,dy)\right] &= \int_{A\times(s,t]\times E} H\,d\overline{\mu} \\
&= \int_{A\times(s,t]\times E} H\,d\overline{\nu} \\
&= \mathbb{E}\left[\mathbf{1}_A \int_{(s,t]\times E} H(r,y)\nu(dr,dy)\right],
\end{aligned}$$

hence, taking $s = 0$, if H satisfies the integrability property in the statement of the theorem, then M_t is integrable and, now with general s, $\mathbb{E}[(M_t - M_s)\mathbf{1}_A] = 0$. The result extends to general H by taking differences. □

Let $f : E \to \mathbb{R}$ be bounded and measurable. For any $t < \zeta$ there is n such that $J_n \leq t < J_{n+1}$. Then

$$\begin{aligned}
f(X_t) = f(Y_n) &= f(Y_0) + \sum_{m=0}^{n-1}\{f(Y_{m+1}) - f(Y_m)\} \\
&= f(Y_0) + \int_{(0,t]\times E} \{f(y) - f(X_{s^-})\}\mu(ds,dy).
\end{aligned}$$

But

$$\begin{aligned}
\int_{(0,t]\times E} \{f(y) - f(X_{s^-})\}\nu(ds,dy) &= \int_0^t \int_E \{f(y) - f(X_{s^-})\}q(X_{s^-},dy)ds \\
&= \int_0^t Qf(X_s)ds,
\end{aligned}$$

so

$$f(X_t) = f(X_0) + M_t^f + \int_0^t Qf(X_s)ds$$

where M^f is a martingale given by

$$M_t^f := \int_{(0,t]\times E} \{f(y) - f(X_{s^-})\}(\mu - \nu)(ds,dy).$$

Also,

$$M_t^2 = 2\int_{(0,t]\times E} M_{s^-}\{f(y) - f(X_{s^-})\}(\mu - \nu)(ds, dy)$$
$$+ \int_{(0,t]} \{f(y) - f(X_{s^-})\}^2 \mu(ds, dy),$$

(This formula can be checked at each jump time, and follows from calculus at the other times. Recall $F_t = F_0 + \int_0^t F_s' ds + \sum_{s\le t} \Delta F_s$.) Therefore

$$\mathbb{E}[M_t^2] = \int_0^t \alpha(X_s) ds, \qquad \text{where} \qquad \alpha(x) = \int_E \{f(y) - f(s)\}^2 q(x, dy).$$

Recap: For $(X_t, t \ge 0)$ a Markov jump process with generator Q and $f : E \to \mathbb{R}$ measurable (but not now not necessarily bounded),

$$f(X_t) = f(X_0) + M_t + \int_0^t Qf(X_s) ds$$

where $M_t^2 = N_t + \int_0^t \alpha(X_s)ds$, with $\alpha(x) = \alpha^f(x) := \int_E \{f(y) - f(x)\}^2 q(x, dy)$, and M and N are *local* martingales on $[0, \zeta)$. Recall $Qf(x) = \int_E \{f(y) - f(x)\} q(x, dy)$ and we need $\gamma(x) := \int_E |f(y)| q(x, dy) < \infty$ for all x. Then

$$T_n = \inf\{t \ge 0 : q(X_t) \ge n, \gamma(X_t) \ge n, \alpha(X_n) \ge n\} \nearrow \zeta,$$

a sequence of stopping times reducing M and N.

We seek $\phi = \phi^f : E \to \mathbb{R}$ such that

$$Z_t := Z_t^f := \exp\left[M_t - \int_0^t \phi(M_s) ds\right]$$

is a local martingale on $[0, \zeta)$. Let

$$\begin{aligned} Z_t &= Z_0 + \sum_{s\le t}(Z_s - Z_{s^-}) - \int_0^t Z_s(Qf + \phi)(X_s) ds \\ &= Z_0 + \int_{(0,t]\times E} Z_{s^-}(e^{f(y) - f(X_{s^-})} - 1)\mu(ds, dy) - \int_0^t Z_s(Qf + \phi)(X_s) ds. \end{aligned}$$

Then

$$\phi(x) = \int_E \{e^{f(y)-f(x)} - 1 - (f(y) - f(x))\} q(x, dy)$$

has the desired property, assuming that $\phi(x) < \infty$ for all x.

Uniqueness in law

Suppose that X is Markov with generator Q. Apply 7.1.8 with $T = J_n$ to obtain

$$\begin{aligned} \mathbb{P}(J_{n+1} - J_n > t, Y_{n+1} \in B \mid \mathscr{F}_{J_n}) &= \mathbb{P}(J_1(J_n) > J_n + t, Y_1(J_n) \in B \mid \mathscr{F}_{J_n}) \\ &= \pi(X_{J_n}, B) e^{-q(X_{J_n})t} \\ &= \pi(Y_n, B) e^{-q(Y_n)t} \end{aligned}$$

a.s. on $\{J_n < \infty\}$. Define

$$T_{n+1} = \begin{cases} q(Y_n)(J_{n+1} - J_n) & \text{if } J_n < \infty \\ \tilde{T}_{n+1} & \text{otherwise} \end{cases}$$

where $(\tilde{T}_n, n \geq 1)$ are independent $\mathscr{E}^{xp}(1)$ r.v.'s. Then $(Y_n, n \geq 0)$ is Markov with transition kernel $\pi(x, dy)$ and $(T_n, n \geq 1)$ are $\mathscr{E}^{xp}(1)$ r.v.'s, all independent. Since $(X_t, t \geq 0)$ is a measurable function of these sequences, its law is determined.

Finite-dimensional distributions—transition semigroups

Define $P_t(x, B) = \mathbb{P}_x(X_t \in B)$, $B \in \mathscr{E}$, a measurable kernel. Then by the uniqueness argument $\mathbb{P}_x(X_t \in B \mid \mathscr{F}_0) = P_t(X_0, B)$ a.s. Then

$$\mathbb{P}_x(X_{t+s} \in B \mid \mathscr{F}_s) = P_s(X_t, B) \text{ a.s.}$$

since $(X_{t+s}, s \geq 0)$ is $(\mathscr{F}_{t+s}, s \geq 0)$-Markov, starting from X_t. Iterating, we obtain for $0 = t_0 \leq t_1 \leq \cdots \leq t_n$, $B_1, \ldots, B_n \in \mathscr{E}$,

$$\mathbb{P}_{x_0}(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n) = \int_{x_1 \in B_1, \ldots, x_n \in B_n} \prod_{i=0}^{n-1} P_{t_{i+1} - t_i}(x_i, dx_{i+1}).$$

The kernel $P_t(x, dy)$ defines a semigroup

$$P_t f(x) = \int_E f(y) P_t(x, dy)$$

acting on bounded measurable functions f, called the *transition semigroup* of $(X_t, t \geq 0)$. (Exercise: $P_t(P_s f) = P_{s+t} f$.) By conditioning on the first jump,

$$P_t(x, B) = \mathbb{P}(X_t \in B) = e^{-q(x)t} \delta_x(B) + \int_0^t \int_E q(x) e^{-q(x)s} \pi(x, dy) P_{t-s}(y, B) ds$$

so

$$e^{q(x)t} P_t(x, B) = \delta_x(B) + \int_0^t \underbrace{\int_E e^{q(x)u} q(x, dy) P_u(y, B)\, du}_{\text{continuous in } u \text{ by DCT}}$$

It follows that $P_t(x, B)$ is differentiable in t with $\dot{P}_t = Q P_t$ (fill this in), where $Q P_t(x, B) = (Q P_t(\cdot, B))(x)$. When E is finite, this shows that $P_t = e^{tQ} = \sum_{n=0}^{\infty} \frac{1}{n!} (tQ)^n$ (and this holds in more generality as well).

Martingale inequalities

Let $(X_t, t \geq 0)$ be Markov with generator Q, and $f : E \to \mathbb{R}$ be measurable. Then

$$f(X_t) = f(X_0) + M_t + \int_0^t Q f(X_s) ds,$$

$$M_t^2 = N_t + \int_0^t \alpha(X_s) ds,$$

and

$$Z_t = \exp(M_t - \int_0^t \phi(X_s)ds),$$

where these local martingales M, N, and Z require progressively stronger assumptions about f. There exist stopping times $T_n \nearrow \zeta$ such that all three stopped at T_n are martingales. Letting $n \to \infty$, since $Z \geq 0$, Z is a supermartingale, (assuming $\zeta = \infty$). Also,

$$\mathbb{E}[(M_t^{T_n})^2] = \mathbb{E}\left[\int_0^{T_n \wedge t} \alpha(X_s)ds\right],$$

so by Doob's L^2 inequality

$$\mathbb{E}[\sup_{s\leq t}(M_s^{T_n})^2] \leq 4\,\mathbb{E}\left[\int_0^{T_n \wedge t} \alpha(X_s)ds\right]$$

and by the monotone convergence theorem (since $M_t = M_s$ for $t \geq \zeta$),

$$\mathbb{E}[\sup_{s\leq t} M_s^2] \leq 4\int_0^t \alpha(X_s)ds. \tag{4}$$

For next time, let $Qf(x) =: \beta(x)$ with $f(x) = x$. Then $X_t = X_0 + M_t + \int_0^t \beta(X_s)ds$ so X is close to the solution to the deterministic equation $x_t = x_0 + \int_0^t \beta(x_s)ds$ when M is small.

7.2.4 Proposition (Exponential martingale inequality).
Assume for all $x \in E$ that

(i) $q(x) \leq \Lambda$ (i.e. bounded rates); and

(ii) $q(x, \{y \in E : f(y) - f(x) \geq J\}) = 0$ (i.e. $f(x)$ makes jumps up by at most J).

Fix $\delta > 0$ and $t \geq 0$ and choose C so that $C \geq J^2\Lambda e^{\frac{\delta J}{Ct}}$. Then

$$\mathbb{P}(\sup_{s\leq t} M_s > \delta) \leq e^{-\frac{\delta^2}{2Ct}}.$$

Remark. We use this inequality when $\Lambda \approx N$ and $J \approx \frac{1}{N}$.

PROOF: Set $\theta = \frac{\delta}{Ct}$ and apply the OST to $Z = Z^{\theta f}$ at the stopping time $T \wedge t$, where $T := \inf\{t \geq 0 : M_t > \delta\}$. Note

$$\phi^{\theta f}(x) \leq (e^{\theta J} - 1 - \theta J)\Lambda \leq \frac{1}{2}\theta^2 J^2 e^{\theta J}\Lambda \leq \frac{1}{2}\theta^2 C$$

so, on $\{T \leq t\}$,

$$Z_T^{\theta f} = \exp\{\theta M_T - \int_0^T \phi^{\theta f}(X_s)ds\} \geq e^{\theta\delta - \frac{1}{2}\theta^2 Ct} = e^{-\frac{\delta^2}{2Ct}}.$$

Hence by Chebyshev's inequality,

$$\mathbb{P}(T \leq t) \leq e^{-\frac{\delta^2}{2Ct}}\,\mathbb{E}[Z_{T\wedge t}^{\theta f}] \leq e^{-\frac{\delta^2}{2Ct}}. \qquad \square$$

7.3 Fluid limit for Markov jump processes

We can use the exponential martingale inequality to give conditions for when the jump process can be approximated closely (and how closely) by a solution to a deterministic DE.

Let $x^i : E \to \mathbb{R}$, $i = 1, \dots, d$, $\vec{x} = (x^1, \dots, x^d) : E \to \mathbb{R}^d$, and $\vec{X}_t = \vec{x}(X_t)$. Let $\beta^i = \beta^{x^i} = Qx^i$ and $\alpha^i = \alpha^{x^i}$. Then

$$\beta(\xi) = (\beta^1(\xi), \dots, \beta^d(\xi)) = \int_E \{\vec{x}(\eta) - \vec{x}(\xi)\} q(\xi, d\eta)$$

is the *drift vector*, and

$$\alpha(\xi) = \sum_{i=1}^d \alpha^i(\xi) = \int_E |\vec{x}(\eta) - \vec{x}(\xi)|^2 q(\xi, d\eta).$$

Assume that $q(\xi) \le \Lambda$ and $q(\xi, \{\eta \in E : \|\vec{x}(\eta) - \vec{x}(\xi)\| \ge J\}) = 0$ (where $\|\vec{x}\| = \sup_i |x_i|$). Then $\vec{X}_t = \vec{X}_0 + M_t + \int_0^t \beta(X_s) ds$, where $M_t = (M_t^1, \dots, M_t^d)$, $M_t^i = M_t^{x_i}$. Assume further that for some $x_0 \in \mathbb{R}^d$ and some Lipschitz vector field b on $\mathbb{R}^d$ with L-constant K, $\vec{X}_0 = x_0$, $\beta(\xi) = b(\vec{x}(\xi))$ for all $\xi \in E$. We compare $(\vec{X}_t, t \ge 0)$ with the solution to $\dot{x}_t = b(x_t)$ starting from x_0.

7.3.1 Theorem (Kurtz). *Suppose that $\alpha(\xi) \le A$ for all $\xi \in E$. Then given $\varepsilon > 0$ and $t_0 \ge 0$, set $\delta = \varepsilon e^{-Kt_0}$. Then*

$$\mathbb{P}(\sup_{t \le t_0} |\vec{X}_t - \vec{x}_t| > \varepsilon) \le \frac{4At_0}{\delta^2}.$$

PROOF: Write $\vec{X}_t - \vec{x}_t = M_t + \int_0^t (b(\vec{X}_s) - b(\vec{x}_s)) ds$. By the Lipschitz property of b, on the set $\{\sup_{t \le t_0} |M_t| \le \delta\}$ we have

$$f(t) := \sup_{s \le t} |\vec{X}_s - x_s| \le \delta + K \int_0^t f(s) ds,$$

so by Gronwall's Lemma, $f(t_0) \le \varepsilon$. Whence by (4),

$$\mathbb{P}(\sup_{t \le t_0} |\vec{X}_t - x_t| > \varepsilon) \le \mathbb{P}(\sup_{t \le t_0} |M_t| > \delta\}) \le \frac{1}{\delta^2} \mathbb{E}[\sup_{t \le t_0} |M_t|^2] \le \frac{4At_0}{\delta^2}. \qquad \square$$

Suppose now that $\|b(x) - b(y)\| \le L\|x - y\|$.

7.3.2 Theorem. *Given $\varepsilon > 0$ and $t_0 \ge 0$, set $\delta = \varepsilon e^{-Lt_0}$. Choose C so that $C \ge J^2 \Lambda e^{\frac{\delta J}{Ct_0}}$. Then*

$$\mathbb{P}(\sup_{t \le t_0} \|\vec{X}_t - x_t\| > \varepsilon) \le 2d e^{-\frac{\delta^2}{2Ct_0}}.$$

PROOF: By the exponential martingale inequality,

$$\mathbb{P}(\sup_{t \le t_0} M_t^{x_i} > \delta) \le e^{-\frac{\delta^2}{2Ct_0}}$$

for $i = 1, \dots, d$, so

$$\mathbb{P}(\sup_{t \le t_0} \|M_t\| > \delta) \le 2de^{-\frac{\delta^2}{2Ct_0}}.$$

But outside of $\{\sup_{t \le t_0} \|M_t\| > \delta\}$, $g(t) := \sup_{s \le t} \|\vec{X}_s - x_s\| \le \delta + L \int_0^t g(s)ds$, so again by Gronwall's Lemma, $g(t_0) \le \varepsilon$. □

7.3.3 Example (Stochastic epidemic). Take

$$E = \{\xi = (\xi^{(1)}, \xi^{(2)}) \in (\mathbb{Z}^+)^2 : \xi^{(1)} + \xi^{(2)} \le N\}.$$

Let

$$q(\xi, \xi') = \begin{cases} \lambda \xi^{(1)} \xi^{(2)}/N & \xi' = \xi + (-1, 1) \\ \mu \xi^{(2)} & \xi' = \xi + (0, -1) \end{cases}$$

Interpret $X_t = (\xi_t^{(1)}, \xi_t^{(2)})$ to be (susceptibles, infectives). Let $\vec{x}(\xi) = \xi/N$. Assume that $\vec{X}_0 = x_0 = X_0/N = (1 - p, p)$ with an initial proportion p of infectives. Define $(x_t, t \ge 0)$ by

$$\dot{x}_t^{(1)} = -\lambda x_t^{(1)} x_t^{(2)}$$
$$\dot{x}_t^2 = \lambda x_t^{(1)} x_t^{(2)} - \mu x_t^{(2)}.$$

Exercise: find $C = C(\lambda, \mu, t_0) < \infty$ such that

$$\mathbb{P}(\sup_{t \le t_0} \|\vec{X}_t - x_t\| > \varepsilon) \le 4e^{-\frac{N\varepsilon^2}{C}}.$$

Index