Stochastic Calculus

Lecture notes based on the book by J.-F. Le Gall

Russell Lyons

May 9, 2025

Contents

Preface

I gave these lectures at Indiana University during the academic year 2017–18. Initially, one of the students, ChunHsien Lu, typed the notes during class. Later, another student who was not in the course, Zhifeng Wei, used my handwritten notes to correct and complete the typed notes. I am very grateful to both of them for all their work. Zhifeng deserves special thanks for figuring out how to add reasons beautifully to displayed equations, as well as for being attentive in general to all my typesetting requests. I then did some further editing and added some illustrations and a bit more material. I would be grateful to learn of any errors or improvements; please email me at rdlyons@iu.edu.

The course was based on the book, *Brownian Motion, Martingales, and Stochastic Calculus*, by Jean-François Le Gall. The same theorem and exercise numbers are used here, although I have not reproduced the exercises. I also added a large number of exercises, especially in order to have some that were useful for learning new concepts and definitions. I assigned homework once per week, and have included the dates those assignments were due in order that others may gauge the pace. A few new problems were added after the course ended; these do not have due dates. Furthermore, the last homework exercises also do not have dates due because they were given at the end of the term. I spent a substantial amount of time in class going over solutions to the homework, but no solutions are presented here. I am grateful to Jean-François for his advice on teaching this course. This turned out to be one of my most enjoyable teaching experiences ever. I had never taught this material before, and always promptly forgot it whenever I had learned some of it in the past. This time, however, teaching it and working hard on the exercises led to actually learning it.

Other differences from Le Gall's book arise from using somewhat different proofs and sometimes giving more general results. A couple of proofs are substantially different. In addition, I covered Chapter 8 on SDEs before Chapter 7 on PDEs. I did not have time to cover Chapter 9 on local times, nor Sections 5.4–5.6. I later made up for this in part by including appendices on the Cameron–Martin theorem and Girsanov's theorem. A couple of appendices provide material I gave to the students from other sources. Occasionally I refer to Le Gall's book for details not given in lecture.

The format of the typed notes tries to reproduce the format of my handwritten notes and most of what went on the board.

The First Day

We begin with some

Motivation (A special case of Itô's formula)**.** *If* $(B_t)_{t\geqslant 0}$ *is a standard Brownian motion and* $f \in C^2(\mathbb{R})$*, then*

$$\mathrm{d}f(B_t) = f'(B_t)\,\mathrm{d}B_t + \frac{1}{2}f''(B_t)\,\mathrm{d}t.$$

This is like calculus, but there is a second term on the right-hand side: $|\mathrm{d}B_t| \approx \sqrt{\mathrm{d}t}$. So $(\mathrm{d}B_t)^2 \approx \mathrm{d}t$. This shows partly why $L^2(\mathbf{P})$ is a key.

SDEs (semester 2) are defined via stochastic integration (semester 1). Other relations to PDEs and harmonic functions are in semester 2, including conformal invariance of complex Brownian motion.

We will start with preparatory material: Gaussian processes, construction of Brownian motion and its basic properties, and a quick review of discrete-time martingales. Then we will study new material on continuous-time martingales and continuous semimartingales.

Before that, recall that a class $\mathcal{U}$ of random variables on $(\Omega, \mathscr{F}, \mathbf{P})$ is ***uniformly integrable*** if

$$\lim_{t\to\infty} \sup_{X\in\mathcal{U}} \mathbf{E}\big[|X|\mathbf{1}_{[|X|>t]}\big] = 0.$$

This holds if (and, it turns out, only if) $\sup_{X\in\mathcal{U}} \mathbf{E}\big[\varphi(|X|)\big] < \infty$ for some function $\varphi\colon [0,\infty) \to [0,\infty)$ with $\lim_{t\to\infty} \frac{\varphi(t)}{t} = \infty$. If X_n and X are integrable and $X_n \xrightarrow{\mathbf{P}} X$, then the following are equivalent:

1. $\{X_n\}$ is uniformly integrable;
2. $\mathbf{E}\big[|X - X_n|\big] \to 0$;
3. $\mathbf{E}\big[|X_n|\big] \to \mathbf{E}\big[|X|\big]$.

Chapter 1

Gaussian Variables and Gaussian Processes

1.1. Gaussian Random Variables

The ***standard Gaussian*** (or ***normal***) ***density*** is

$$p_X : x \mapsto \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \quad (x \in \mathbb{R}).$$

The complex Laplace transform of such a random variable, X, is

$$z \mapsto \mathbf{E}\big[e^{zX}\big] = \int_{-\infty}^{\infty} e^{zx} p_X(x)\, dx = e^{z^2/2} \quad (z \in \mathbb{C}).$$

One sees this by first calculating the integral for $z \in \mathbb{R}$ and then using analytic continuation (see page 2 of Le Gall's book). In particular, the characteristic function (Fourier transform) is

$$\xi \mapsto \mathbf{E}\big[e^{i\xi X}\big] = e^{-\xi^2/2} \quad (\xi \in \mathbb{R}).$$

Recall that the Fourier transform determines the law of X uniquely. By expanding in a Taylor series, one gets the moments of X, such as $\mathbf{E}[X] = 0$ and $\mathbf{E}\big[X^2\big] = 1$.

We say $Y \sim \mathcal{N}(m, \sigma^2)$ for $m \in \mathbb{R}$ and $\sigma > 0$ if $(Y - m)/\sigma$ is standard normal. This is equivalent to:

$$Y \text{ has density } y \mapsto \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(y-m)^2}{2\sigma^2}\right\}$$

and to

$$Y \text{ has Fourier transform } \xi \mapsto e^{im\xi - \sigma^2\xi^2/2}.$$

Note that then $\mathbf{E}[Y] = m$ and $\mathrm{Var}(Y) = \sigma^2$. If $Y = m$ a.s., we will also say $Y \sim \mathcal{N}(m, 0)$.

Using the Fourier transform, one sees that a sum of two independent normal random variables is also normal.

One proves properties of stochastic processes with a continuous parameter by taking limits in various senses from finite or countable subsets of parameters. This is how we will use the following:

Proposition 1.1. *Suppose $X_n \sim \mathcal{N}(m_n, \sigma_n^2)$ converges to X in L^2 (i.e., $\mathbf{E}\big[|X_n - X|^2\big] \to 0$ as $n \to \infty$). Then*

(i) *$X \sim \mathcal{N}(m, \sigma^2)$ with $m := \lim m_n$ and $\sigma := \lim \sigma_n$;*

(ii) *$X_n \to X$ in L^p for every $p \in (0, \infty)$.*

Proof. (i) That $\lim m_n = \mathbf{E}[X]$ and $\lim \sigma_n^2 = \mathrm{Var}(X)$ does not use that $(X_n)_{n\geqslant 1}$ are Gaussian. The fact that X is Gaussian then follows from using the Fourier transform.

(ii) Because $X_n \overset{\mathscr{D}}{=} \sigma_n N + m_n$ with $N \sim \mathcal{N}(0,1)$, we see that

$$\forall q > 0 \qquad \sup_n \mathbf{E}\big[|X_n|^q\big] < \infty,$$

whence

$$\sup_n \mathbf{E}\big[|X_n - X|^q\big] < \infty.$$

(Recall that $\|\cdot\|_q$ satisfies the triangle inequality for $q \geqslant 1$ and $\|\cdot\|_q^q$ does for $q < 1$.) Given $p \in (0,\infty)$, we get that $\big\{|X_n - X|^p,\ n \geqslant 1\big\}$ is bounded in L^2 (use $q := 2p$) and tends to 0 in probability because $X_n \xrightarrow{\mathbf{P}} X$, whence is uniformly integrable. Therefore, $\mathbf{E}\big[|X_n - X|^p\big] \to 0$. ◂

1.2. Gaussian Vectors

Let E be a Euclidean space, i.e., a finite-dimensional inner-product space. Let X be an E-valued random variable with $\mathbf{E}\big[\|X\|^2\big] < \infty$. We claim that there exist some $m_X \in E$ and a non-negative quadratic form q_X on E such that

$$\forall u \in E \qquad \mathbf{E}\big[\langle u, X\rangle\big] = \langle u, m_X\rangle \text{ and } \mathrm{Var}\big(\langle u, X\rangle\big) = q_X(u).$$

We will then write $\mathbf{E}[X] := m_X$. To see our claim, take an orthonormal basis $(e_1, \dots, e_d)$ of E, write $X = \sum X_j e_j$, and define

$$\begin{aligned} m_X &:= \sum \mathbf{E}[X_j]\, e_j, \\ q_X(u) &:= \sum u_j u_k \operatorname{Cov}(X_j, X_k) = \mathrm{Var}\Big(\sum u_j X_j\Big) \geqslant 0. \end{aligned}$$

Calculation shows this works.

We also write $\gamma_X \colon E \to E$ for the symmetric linear mapping such that

$$\forall u \in E \qquad q_X(u) = \big\langle u, \gamma_X(u)\big\rangle;$$

its matrix is $\big(\operatorname{Cov}(X_j, X_k)\big)_{j,k\leqslant d}$. The eigenvalues of γ_X are non-negative.

We call X ***Gaussian*** if $\forall u \in E$ $\langle u, X\rangle$ is Gaussian; we also call the components of X ***jointly Gaussian***.

Example. If $X_1, X_2, \dots, X_d$ are independent Gaussian, then $\sum X_j e_j$ is a Gaussian vector.

If X is Gaussian, then $\langle u, X\rangle \sim \mathcal{N}\big(\langle u, m_X\rangle, q_X(u)\big)$, so

$$\mathbf{E}\, \mathrm{e}^{\mathrm{i}\langle u,X\rangle} = \mathrm{e}^{\mathrm{i}\langle u,m_X\rangle - q_X(u)/2}. \tag{1.1}$$

We write $X \sim \mathcal{N}(m_X, q_X)$.

Proposition 1.2. *If X is Gaussian, $(e_1, \dots, e_d)$ is an orthonormal basis of E, and $X = \sum X_j e_j$, then $\big(\mathrm{Cov}(X_j, X_k)\big)_{j,k\leqslant d}$ is diagonal if and only if $X_1, X_2, \dots, X_d$ are (mutually) independent.*

Proof. $\Leftarrow$: Independence implies pairwise independence. Thus, $\mathrm{Cov}(X_j, X_k) = 0$ for distinct j and k.

$\Rightarrow$: Conversely, when the covariance matrix is diagonal, the right-hand side of Eq. (1.1) factors as a product over j, and independence follows. ◀

In particular, for jointly Gaussian random variables, pairwise independence implies mutual independence.

For simplicity, we now consider ***centered*** Gaussian vectors, i.e., ones with mean 0. We will not use the following:

Theorem 1.3. (i) *If γ is a positive semi-definite linear map on E, then there exists a Gaussian vector X on E such that $\gamma_X = \gamma$.*

(ii) *Let $X \sim \mathcal{N}(0, \gamma_X)$. Let $(\varepsilon_1, \dots, \varepsilon_d)$ be an orthonormal basis of eigenvectors of γ_X with eigenvalues $\lambda_1 \geqslant \cdots \geqslant \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_d$. Then there exist independent $Y_j \sim \mathcal{N}(0, \lambda_j)$ such that*

$$X = \sum_{j=1}^{r} Y_j\, \varepsilon_j.$$

The support of the law P_X of X is the linear span of $\{\varepsilon_1, \dots, \varepsilon_r\}$. Also, P_X is absolutely continuous with respect to Lebesgue measure if and only if $r = d$, in which case the density of X is

$$p_X \colon x \mapsto \frac{1}{(2\pi)^{d/2}\sqrt{\det \gamma_X}} \mathrm{e}^{-\langle x, \gamma_X^{-1}(x)\rangle/2}.$$

1.3. Gaussian Processes and Gaussian Spaces

We will often omit the word "centered".

Another way to say that $(X_1, X_2, \dots, X_d) \in \mathbb{R}^d$ is a Gaussian vector is to say that the linear span of $\{X_1, X_2, \dots, X_d\}$ in $L^2(\Omega, \mathbf{P})$ consists only of Gaussian random variables.

Definition 1.4. *A **(centered) Gaussian space** is a closed linear subspace of $L^2(\Omega, \mathbf{P})$ that contains only centered Gaussian variables.*

Definition 1.5. *Let T be a set and $(E, \mathscr{E})$ be a measurable space. A **stochastic process** (or **random process**) indexed by T with values in E is a collection $(X_t)_{t\in T}$ of E-valued random variables. If $(E, \mathscr{E})$ is not specified, then we assume that $E = \mathbb{R}$ and $\mathscr{E} = \mathscr{B}(\mathbb{R})$ is its Borel σ-field. Usually, $T = \mathbb{R}_+ := [0, \infty)$.*

Definition 1.6. *A stochastic process* $(X_t)_{t\in T} \in \mathbb{R}^T$ *is a* ***Gaussian process*** *if for every finite subset* T' *of* T, $(X_t)_{t\in T'} \in \mathbb{R}^{T'}$ *is a Gaussian vector.*

By Proposition 1.1, we get

Proposition 1.7. *If* $(X_t)_{t\in T}$ *is a Gaussian process, then the closed linear span of* $(X_t)_{t\in T}$ *in* $L^2(\Omega, \mathbf{P})$ *is a Gaussian space, called the* ***Gaussian space generated by*** $(X_t)_{t\in T}$. ◀

Exercise (due 8/31). Exercise 1.15 (4 parts).

The undergrad notion that jointly normal, centered random variables (X, Y) are independent if and only if they are orthogonal in L^2 (i.e., $\mathbf{E}(XY) = 0$), which we proved and extended in Proposition 1.2, has the following further extension:

Theorem 1.9. *Let* H *be a centered Gaussian space and* $\mathcal{K}$ *be a collection of linear subspaces of* H. *Then the subspaces of* $\mathcal{K}$ *are (pairwise) orthogonal* ($\perp$) *in* L^2 *if and only the* σ*-fields* $\sigma(K)$ $(K \in \mathcal{K})$ *are independent* ($\perp\!\!\!\perp$).

Proof. Independence implies orthogonality trivially.

For the converse, it suffices to show that if $K_1, K_2, \ldots, K_p \in \mathcal{K}$ are distinct, then $\sigma(K_1), \sigma(K_2), \ldots, \sigma(K_p)$ are independent, because this is the definition of independence for infinitely many σ-fields. In turn, this follows if we show that $(\xi_1^1, \xi_2^1, \ldots, \xi_{n_1}^1), \ldots, (\xi_1^p, \xi_2^p, \ldots, \xi_{n_p}^p)$ are independent for $\xi_i^j \in K_j$. (This is a standard fact and follows from Dynkin's π-λ theorem, which is called in the book "the monotone class lemma"; see Appendix 1 for that and this application. Halmos' monotone class lemma is given on page 89 of the book.) Now let $(\eta_1^j, \eta_2^j, \ldots, \eta_{m_j}^j)$ be an orthonormal basis of the span of $(\xi_1^j, \xi_2^j, \ldots, \xi_{n_j}^j)$. Orthogonality gives that the vector

$$(\eta_1^1, \eta_2^1, \ldots, \eta_{m_1}^1, \eta_1^2, \eta_2^2, \ldots, \eta_{m_2}^2, \ldots, \eta_1^p, \eta_2^p, \ldots, \eta_{m_p}^p)$$

has covariance matrix the identity. This is a Gaussian vector since its components are in H. Proposition 1.2 then yields that all η_i^j are independent, whence

$$(\eta_1^1, \eta_2^1, \ldots, \eta_{m_1}^1), \ldots, (\eta_1^p, \eta_2^p, \ldots, \eta_{m_p}^p)$$

are independent. This gives the result. ◀

If $X\colon (\Omega, \mathscr{F}, \mathbf{P}) \to (E, \mathscr{E})$ and $\mathscr{G}$ is a sub-σ-field of $\mathscr{F}$, then a ***regular conditional distribution*** for X given $\mathscr{G}$ is a function $\mu\colon \Omega \times \mathscr{E} \to [0, 1]$ such that

(1) $\forall \omega \in \Omega \;\; \mu(\omega, \cdot)$ is a probability measure

and

(2) $\forall A \in \mathscr{E} \;\; \mu(\cdot, A)$ is a version of $\mathbf{P}[X \in A \mid \mathscr{G}]$.

This exists if $(E, \mathscr{E})$ is a standard Borel space (Borel isomorphic to a Borel subset of $\mathbb{R}$), such as a Borel subset of a Polish space (complete, separable, metrizable space); see Durrett's book, *Probability: Theory and Examples*.

Corollary 1.10. *Let H be a (centered) Gaussian space and K be a closed linear subspace of H. Let $p_K\colon H \to K$ be the orthogonal projection. If $X_1, X_2, \ldots, X_d \in H$, then the $\sigma(K)$-conditional distribution of $(X_1, X_2, \ldots, X_d)$ is*

$$\mathcal{N}\Big(\big(p_K(X_i)\big)_{i=1}^d,\ q_{(p_K^{\perp}(X_i))_{i=1}^d}\Big).$$

Proof. We have

$$X_i = \underbrace{p_K^{\perp}(X_i)}_{\perp\!\!\!\perp\ \sigma(K)} + \underbrace{p_K(X_i)}_{\in\ \sigma(K)}, \quad 1 \leqslant i \leqslant d. \qquad ◀$$

See the book for more details when $d = 1$. Note that here $\mathbf{E}[X \mid \sigma(K)] = p_K(X)$, whereas in general (outside the context of Gaussian random variables), it is $p_{L^2(\Omega,\sigma(K),\mathbf{P})}(X)$.

Exercise (due 9/7). Exercise 1.17.

1.4. Gaussian White Noise

White noise is an engineering term that refers to a signal with constant Fourier transform. In the case of a stationary stochastic process, we look at the spectral measure (page 11 in the book), whose Fourier transform is the covariance function; it should be $c \cdot \delta_0$. That is, the process should have no correlations; in the Gaussian case, this is equivalent to independence. This makes most sense if the index set is $\mathbb{Z}$. But we are interested in $\mathbb{R}$. However, the index set for us will not be $\mathbb{R}$, but $\mathscr{B}(\mathbb{R})$. Motivations include increments of Brownian motion and the Poisson process in $\mathbb{R}$ or $\mathbb{R}^2$ . . . . Thus, each Borel set $A \in \mathscr{B}(\mathbb{R})$ gives a Gaussian random variable, $G(A)$. If $A_1 \cap A_2 = \varnothing$, then we want $G(A_1) \perp\!\!\!\perp G(A_2)$.

Definition 1.12. *Let $(E, \mathscr{E})$ be a measurable space and μ be a measure on $(E, \mathscr{E})$. A **Gaussian white noise with intensity** μ is an isometry G from $L^2(E, \mathscr{E}, \mu)$ into a (centered) Gaussian space.*

Thus, for $f, g \in L^2(E)$, we have

$$\operatorname{Cov}\big(G(f), G(g)\big) = \langle f, g\rangle_{L^2},$$
$$\operatorname{Var}\big(G(f)\big) = \|f\|_{L^2}^2.$$

If $A \in \mathscr{E}$ with $\mu(A) < \infty$, we set $G(A) := G(\mathbf{1}_A) \sim \mathcal{N}\big(0, \mu(A)\big)$. If $A_1, A_2, \ldots, A_n \in \mathscr{E}$ with $\mu(A_j) < \infty$, then $\big(G(A_1), G(A_2), \ldots, G(A_n)\big)$ is a Gaussian vector with covariance

$$\operatorname{Cov}\big(G(A_i), G(A_j)\big) = \mu(A_i \cap A_j).$$

In particular, if $A_1, A_2, \ldots, A_n$ are disjoint, then the covariance matrix is diagonal, so by Proposition 1.2, the variables $G(A_1), G(A_2), \ldots, G(A_n)$ are independent.

If $A \in \mathscr{E}$, $\mu(A) < \infty$, is partitioned into $A_1, A_2, \ldots \in \mathscr{E}$, then $\mathbf{1}_A = \sum_j \mathbf{1}_{A_j}$ in L^2, so by isometry,

$$G(A) = \sum_j G(A_i) \quad \text{in } L^2.$$

Kolmogorov's theorem shows that we also have almost sure convergence. However, in general, it is not possible to make $A \mapsto G(A)$ a signed measure almost surely, even when $(E, \mathscr{E}) = \big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$, as Corollary 2.17 will show.

Proposition 1.13. *Let $(E, \mathscr{E})$ be a measurable space and μ be a measure on $(E, \mathscr{E})$. There exists a probability space $(\Omega, \mathscr{F}, \mathbf{P})$ and a Gaussian white noise on $L^2(\Omega, \mathscr{F}, \mathbf{P})$ with intensity μ.*

Proof. Let $(f_i)_{i\in I}$ be an orthonormal basis for $L^2(E, \mathscr{E}, \mu)$. Choose a probability space on which there exist i.i.d. random variables $X_i \sim \mathcal{N}(0,1)$ $(i \in I)$. Define $G\colon L^2(\mu) \to L^2(\mathbf{P})$ by $G(f_i) := X_i$. That is, for $f \in L^2(\mu)$, we define

$$G(f) := \sum_{i\in I} \langle f, f_i \rangle X_i.$$

The fact that G is an isometry uses only that the variables $(X_i)_{i\in I}$ are orthonormal. The fact that G takes values in a Gaussian space uses that $(X_i)_{i\in I}$ is standard normal and Proposition 1.1(i). ◀

Exercise. Let $(E, \mathscr{E})$ be a measurable space and μ be a measure on $(E, \mathscr{E})$. Let $f_1, f_2 \in L^2(\mu)$. Let G be a Gaussian white noise on $L^2(\Omega, \mathscr{F}, \mathbf{P})$ with intensity μ. Calculate the joint distribution of $G(f_1)$ and $G(f_2)$ and the conditional distribution of $G(f_2)$ given $G(f_1)$.

Exercise (due 9/7). Exercise 1.18.

Proposition 1.14. *Let G be a Gaussian white noise on $(E, \mathscr{E})$ with intensity μ and $A \in \mathscr{E}$ have $\mu(A) < \infty$. If for each $n \in \mathbb{N}$, A is partitioned as $A = \bigcup_{j=1}^{k_n} A_j^n$ with*

$$\lim_{n\to\infty} \max_{1\leqslant j\leqslant k_n} \mu(A_j^n) = 0,$$

then

$$\lim_{n\to\infty} \sum_{j=1}^{k_n} G(A_j^n)^2 = \mu(A) \quad \text{in } L^2(\mathbf{P}).$$

Proof. We have $G(A_j^n) \sim \mathcal{N}\big(0, \mu(A_j^n)\big)$ are independent. From page 2 of the book, we know $G(A_j^n)^2$ have variance $2\,\mu(A_j^n)^2$. Therefore,

$$\operatorname{Var}\left(\sum_{j=1}^{k_n} G(A_j^n)^2\right) = 2\sum_{j=1}^{k_n} \mu(A_j^n)^2 \leqslant 2\underbrace{\big(\max_{1\leqslant j\leqslant k_n} \mu(A_j^n)\big)}_{\to 0} \cdot \underbrace{\sum_{j=1}^{k_n} \mu(A_j^n)}_{\mu(A)} \to 0.$$

But this is precisely $\left\|\sum_{j=1}^{k_n} G(A_j^n)^2 - \mu(A)\right\|_2^2$. ◀

If the partitions are successive refinements, then we have almost sure convergence by Doob's martingale convergence theorem.

Chapter 2

Brownian Motion

Although Exercise 1.18 constructed Brownian motion (on $[0,1]$), we will give another construction that yields more information, via a lemma of Kolmogorov that will also be used in later chapters.

2.1. Pre-Brownian Motion

The following is a natural extension of Exercise 1.18(3).

Definition 2.1. *A **pre-Brownian motion** $(B_t)_{t\geqslant 0}$ is a stochastic process such that*

$$B_t = G\big(\mathbf{1}_{[0,t]}\big) \qquad \text{“c.d.f. of } G\text{”}$$

for some Gaussian white noise G on $\mathbb{R}_+$ whose intensity is Lebesgue measure.

Proposition 2.2. *Every pre-Brownian motion is a centered Gaussian process with covariance $K(s,t) = \min\{s,t\} =: s \wedge t$.*

Proof. $\operatorname{Cov}(B_s, B_t) = \operatorname{Leb}\big([0,s] \cap [0,t]\big) = s \wedge t$. ◂

Proposition 2.3. *Let $(X_t)_{t\geqslant 0}$ be a (real-valued) stochastic process. The following are equivalent:*

(i) *$(X_t)_{t\geqslant 0}$ is a pre-Brownian motion;*

(ii) *$(X_t)_{t\geqslant 0}$ is a centered Gaussian process with covariance $K(s,t) = s \wedge t$;*

(iii) $X_0 = 0$ a.s. *and* $\forall 0 \leqslant s < t$ $X_t - X_s \sim \mathcal{N}(0, t-s)$ *is independent of* $\sigma(X_r, r \leqslant s)$*;*

(iv) $X_0 = 0$ a.s. *and* $\forall 0 = t_0 < t_1 < \cdots < t_p$ $X_{t_i} - X_{t_{i-1}} \sim \mathcal{N}(0, t_i - t_{i-1})$ *are independent for* $1 \leqslant i \leqslant p$.

Proof. (i) $\Rightarrow$ (ii): Proposition 2.2.

(ii) $\Rightarrow$ (iii): $X_0 \sim \mathcal{N}(0,0)$; $X_t - X_s = G\big([s,t]\big) \sim \mathcal{N}(0, t-s)$; if H_s is the closed linear span of $(X_r)_{0\leqslant r\leqslant s}$ and $\widetilde{H}_s$ the closed linear span of $(X_t - X_s)_{t\geqslant s}$, then $H_s \perp \widetilde{H}_s$ since $X_r \perp (X_t - X_s)$ ($\mathbf{E}\big[X_r(X_t - X_s)\big] = r \wedge t - r \wedge s = r - r = 0$) for $r \leqslant s \leqslant t$, whence by Theorem 1.9, $\sigma(H_s) \perp\!\!\!\perp \sigma(\widetilde{H}_s)$.

(iii) $\Rightarrow$ (iv): By (iii), $X_{t_i} - X_{t_{i-1}} \perp\!\!\!\perp \sigma(X_{t_j} - X_{t_{j-1}}\,;\ 0 \leqslant j < i)$ for each $i \in [1,p]$.

(iv) $\Rightarrow$ (i): We need to define $G(f)$ for $f \in L^2(\mathbb{R}_+)$. We start with step functions $f = \sum_{i=1}^n \lambda_i \mathbf{1}_{(t_{i-1},t_i]}$, where $0 = t_0 < t_1 < \cdots < t_n$. For such f, we define

$$G(f) \coloneqq \sum_{i=1}^n \lambda_i (X_{t_i} - X_{t_{i-1}}).$$

This does not depend on the representation of f: to see this, use a common refinement. Similarly, to see that $\mathbf{E}\big[G(f)G(g)\big] = \int_{\mathbb{R}_+} fg$ for each f and g, use a common refinement. Thus, G is an isometry from step functions on $\mathbb{R}_+$ into the Gaussian space generated by X. Since step functions are dense in $L^2(\mathbb{R}_+)$, we may extend G to an isometry on $L^2(\mathbb{R}_+)$. By construction, $G\big((0,t]\big) = X_t - X_0 = X_t$. ◂

Exercise. Show that $(X_t)_{t\geqslant 0}$ is a pre-Brownian motion iff $X_0 = 0$ a.s. and $(X_t)_{t\geqslant 0}$ is a centered Gaussian process with $\forall 0 \leqslant s < t \;\; \operatorname{Var}(X_t - X_s) = t - s$.

The ***finite-dimensional distributions*** of pre-Brownian motion—the laws of $(B_{t_1}, B_{t_2}, \ldots, B_{t_n})$ for $0 < t_1 < \cdots < t_n$—are unique by the equivalence of (i) and (iv) in Proposition 2.3. To be explicit:

Corollary 2.4. *Let $(B_t)_{t\geqslant 0}$ be a pre-Brownian motion and $0 = t_0 < t_1 < \cdots < t_n$. Then $(B_{t_1}, B_{t_2}, \cdots, B_{t_n})$ has density on $\mathbb{R}^n$*

$$(x_1, \ldots, x_n) \mapsto \prod_{i=1}^n \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\Big\{-\sum_{i=1}^n \frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})}\Big\},$$

where $x_0 \coloneqq 0$.

Proof. Independence of increments gives the joint density of the increments. Then we change variables $(y_1, \ldots, y_n) \mapsto (x_1, \ldots, x_n)$ via $x_i \coloneqq \sum_{j=1}^i y_j$, which has Jacobian determinant 1. ◂

Some simple properties of pre-Brownian motion:

Proposition 2.5. *Let $(B_t)_{t\geqslant 0}$ be a pre-Brownian motion.*

(i) *$(-B_t)_{t\geqslant 0}$ is a pre-Brownian motion.*

(ii) *$\forall \lambda > 0$ $\big(B_t^\lambda\big)_{t\geqslant 0}$ defined by $B_t^\lambda \coloneqq \frac{1}{\lambda} B_{\lambda^2 t}$ is a pre-Brownian motion.*

(iii) *$\forall s \geqslant 0$ $\big(B_t^{(s)}\big)_{t\geqslant 0}$ defined by $B_t^{(s)} \coloneqq B_{s+t} - B_s$ is a pre-Brownian motion and is independent of $\sigma(B_r, r \leqslant s)$.*

Proof. (i) and (ii) follow from (say) Proposition 2.3(ii).

In the notation of the proof of Proposition 2.3, we have $\sigma\big(B_t^{(s)}, t \geqslant 0\big) = \sigma(\widetilde{H}_s)$, which we saw is independent of $\sigma(H_s) = \sigma(B_r, r \leqslant s)$. The finite-dimensional distributions are correct as a special case of those for B itself. ◂

We defined B in terms of G, but G is also determined by B: we did this in Proposition 2.3 (iv)$\Rightarrow$(i), using step functions and limits. One sometimes writes

$$G(f) = \int_0^\infty f(s)\,\mathrm{d}B_s \quad \big(f \in L^2(\mathbb{R}_+)\big).$$

This is called the ***Wiener integral***. However, $G(\cdot)$ is not an almost sure measure and this integral makes no sense pointwise. We will extend integration to random f in Chapter 5.

Exercise. Nevertheless, one can integrate by parts in the Wiener integral: Suppose that μ is a finite, signed measure on $(0,t]$ for some $t > 0$ and that $f(s) = \mu(0,s]$ for $s \leqslant t$. Set $f(s) := 0$ for $s > t$. Assume that B is continuous a.s. Show that

$$G(f) = f(t)B_t - \int_{(0,t]} B_s \, \mu(\mathrm{d}s) \quad \text{a.s.}$$

2.2. The Continuity of Sample Paths

In Exercise 1.18, we defined Brownian motion B' by taking a limit of continuous functions based on B, thus getting almost surely a continuous function $t \mapsto B'_t(\omega)$. In our redevelopment, we haven't done that yet. Finite-dimensional distributions cannot guarantee that, since we could always change the process at an independent $U[0,1]$ random time to be 0, say, which would not change the f.d.d.s, yet would make the process discontinuous. We now discuss such modifications more generally.

Definition 2.6. *Let $(X_t)_{t\in T}$ be a stochastic process with values in E. The **sample paths** of X are the maps $t \mapsto X_t(\omega)$ for each $\omega \in \Omega$.*

Definition 2.7. *Let $(X_t)_{t\in T}$ and $(\widetilde{X}_t)_{t\in T}$ be stochastic processes indexed by the same T and taking values in the same E. We say $\widetilde{X}$ is a **modification** of X if*

$$\forall t \in T \qquad \mathbf{P}[\widetilde{X}_t = X_t] = 1.$$

This gives the same finite-dimensional distributions, but that is not enough for us.

Definition 2.8. *With the same notation, we say $\widetilde{X}$ is **indistinguishable** from X if*

$$\mathbf{P}[\forall t \in T \quad \widetilde{X}_t = X_t] = 1.$$

To be more precise, we use the completion of $\mathbf{P}$ here, or, alternatively, the condition is that there exists a subset $N \subseteq \Omega$ with $\mathbf{P}(N) = 0$ such that

$$\forall \omega \in N^{\mathrm{c}} \;\; \forall t \in T \qquad \widetilde{X}_t(\omega) = X_t(\omega).$$

Notice that if T is a separable metric space and $X, \widetilde{X}$ both have continuous sample paths almost surely, then $\widetilde{X}$ is a modification of X if and only if $\widetilde{X}$ is indistinguishable from X. In case $T \subseteq \mathbb{R}$, the same assertion holds with "continuous" replaced with "right-continuous" or "left-continuous".

We are going to prove more than continuity, namely, Hölder continuity. In the context of metric spaces, a function $f\colon (E_1, d_1) \to (E_2, d_2)$ is ***Hölder continuous of order*** α if

$$\exists C < \infty \;\; \forall s, t \in E_1 \qquad d_2\big(f(s), f(t)\big) \leqslant C \cdot d_1(s,t)^{\alpha}.$$

Kolmogorov showed that when f is replaced by a stochastic process on a domain in $\mathbb{R}^k$ that satisfies the above inequality when the left-hand side is replaced by the expectation of a power of the distance and $\alpha > k$ on the right-hand side, then the process has almost sure Hölder continuity of order higher than 0:

Theorem 2.9 (Kolmogorov's lemma, or Kolmogorov's continuity theorem)**.** *Consider a stochastic process $X = (X_t)_{t \in I}$ on a bounded rectangle $I \subseteq \mathbb{R}^k$ that takes values in a complete metric space (E, d). If there exist positive q, ε, C such that*

$$\forall s, t \in I \qquad \mathbf{E}\Big[d(X_s, X_t)^q\Big] \leqslant C|s-t|^{k+\varepsilon},$$

then there exists a modification $\widetilde{X}$ of X whose sample paths are Hölder continuous of order α for all $\alpha \in (0, \frac{\varepsilon}{q})$. Indeed, $\widetilde{X}$ can be chosen to satisfy

$$\forall \alpha < \frac{\varepsilon}{q} \qquad \mathbf{E}\left[\sup_{\substack{s,t \in I\\ s \neq t}} \left(\frac{d(\widetilde{X}_s, \widetilde{X}_t)}{|s-t|^\alpha}\right)^q\right] < \infty. \tag{$*$}$$

Note that for unbounded I, this gives locally Hölder sample paths. Recall that continuous sample path modifications are unique up to indistinguishability.

Proof. We do only $k = 1$. We also take $I = [0, 1]$ for simplicity; the presence of endpoints would not matter. Note that Eq. ($*$) implies that for each $\alpha \in (0, \frac{\varepsilon}{q})$, there is a Hölder-$\alpha$ modification. Using a sequence $\alpha_j \uparrow \frac{\varepsilon}{q}$, we get that there is a modification that is Hölder-α_j for all j (by uniqueness up to indistinguishability). This gives Hölder-α for all $\alpha \in (0, \frac{\varepsilon}{q})$.

Now for $s \neq t$, the hypothesis yields

$$\mathbf{E}\left[\frac{d(X_s, X_t)^q}{|s-t|^{\alpha q}}\right] \leqslant \frac{C|s-t|^{1+\varepsilon}}{|s-t|^{\alpha q}} = C|s-t|^{1+\varepsilon-\alpha q}.$$

Hence,

$$\begin{aligned}
\mathbf{E}\left[\overbrace{\sup_{n \geqslant 1} \sup_{1 \leqslant i \leqslant 2^n} \left(\frac{d(X_{(i-1)2^{-n}}, X_{i2^{-n}})}{(2^{-n})^\alpha}\right)^q}^{K(\omega) :=}\right] &\leqslant \sum_{n \geqslant 1} \sum_{1 \leqslant i \leqslant 2^n} \mathbf{E}\left[\left(\frac{d(X_{(i-1)2^{-n}}, X_{i2^{-n}})}{2^{-n\alpha}}\right)^q\right] \\
&\leqslant \sum_{n \geqslant 1} \sum_{1 \leqslant i \leqslant 2^n} C 2^{-n(1+\varepsilon-\alpha q)} = \sum_n C 2^{-n(\varepsilon - \alpha q)} < \infty.
\end{aligned}$$

We now use:

Lemma 2.10. *Let $D := \{i2^{-n}\,;\ n \geqslant 1,\ 0 \leqslant i \leqslant 2^n\}$, $f\colon D \to (E, d)$, $\alpha > 0$. Then*

$$\sup_{\substack{s,t \in D\\ s \neq t}} \frac{d\big(f(s), f(t)\big)}{|s-t|^\alpha} \leqslant \frac{2}{1-2^{-\alpha}} \sup_{n \geqslant 1} \sup_{1 \leqslant i \leqslant 2^n} \frac{d\Big(f\big((i-1)2^{-n}\big), f(i2^{-n})\Big)}{(2^{-n})^\alpha}.$$

Proof. Take a "chain" from s to t that uses at most two hops of order ℓ for every $\ell \geqslant p$, where $2^{-p} \leqslant |s-t| < 2^{-p+1}$. See page 26 of the book for details. ◂

This gives Eq. ($*$) restricted to $s, t \in D$ for $\widetilde{X}_r := X_r$ $(r \in D)$. In particular, X is almost surely Hölder-α continuous on D, hence continuous on D. Define

$$\widetilde{X}_t(\omega) := \lim_{D \ni s \to t} X_s(\omega) \quad \text{when } K(\omega) < \infty$$

and $\widetilde{X}_t(\omega) := x_0$ for some fixed $x_0 \in E$ when $K(\omega) = \infty$. Then

$$\forall \omega \in \Omega \qquad \sup_{\substack{s,t \in I \\ s \neq t}} \frac{d(\widetilde{X}_s, \widetilde{X}_t)}{|s-t|^\alpha} \leqslant \frac{2}{1-2^{-\alpha}} K(\omega).$$ ◄

Corollary 2.11. *Pre-Brownian motion has a modification with continuous sample paths. Every such modification is indistinguishable from one (all of) whose sample paths are locally Hölder continuous of order α for all $\alpha < \frac{1}{2}$.*

Proof. Recall that a standard normal random variable has a finite qth moment for each $q < \infty$. Thus, for $s < t$, there exists a standard normal U such that

$$B_t - B_s = \sqrt{t-s} \cdot U \in L^q$$

with

$$\mathbf{E}\big[|B_t - B_s|^q\big] = (t-s)^{q/2} \cdot \mathbf{E}\big[|U|^q\big].$$

If $q > 2$, we can apply Theorem 2.9 with $\varepsilon := \frac{q}{2} - 1$ to get Hölder continuity with $\alpha < \frac{\varepsilon}{q} = \frac{1}{2} - \frac{1}{q}$. We may take q arbitrarily large. ◄

Remark. The optimal result is known as "Lévy's modulus of continuity":

$$\lim_{\varepsilon \downarrow 0} \sup_{t \geqslant 0} \frac{|B_{t+\varepsilon} - B_t|}{\sqrt{2\varepsilon \log \frac{1}{\varepsilon}}} = 1 \quad \text{a.s.}$$

Definition 2.12. *A* ***Brownian motion*** *is a pre-Brownian motion with continuous sample paths.*

We have proved Brownian motion exists. Since $-B$, B^λ, $B^{(s)}$ have continuous sample paths when B does, the statements of Proposition 2.5 holds when "pre" is removed everywhere.

In order to discuss the law of the sample paths, we use the space $C(\mathbb{R}_+, \mathbb{R})$ of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}$ equipped with the topology τ of uniform convergence on every compact set. This topology is locally compact. The corresponding Borel σ-field is generated by the coordinate maps $w \mapsto w(t)$ $(t \in \mathbb{R}_+)$.

Exercise. Check that τ is locally compact and its Borel σ-field is generated as claimed.

Then $\omega \mapsto \big(t \mapsto B_t(\omega)\big)$ is measurable since composing it with each coordinate map $w \mapsto w(s)$ gives the measurable B_s. The pushforward of $\mathbf{P}$ is the ***Wiener measure*** W, the law of sample paths: $W(A) = \mathbf{P}[B_\cdot \in A]$ for measurable $A \subseteq C(\mathbb{R}_+, \mathbb{R})$. Corollary 2.4, the finite-dimensional distributions of pre-Brownian motion, gives the finite-dimensional distributions of W, i.e., the collection of laws of $\big(w(t_0), w(t_1), \ldots, w(t_n)\big)$ for $n \geqslant 0$, $0 = t_0 < t_1 < \cdots < t_n$. The ***cylinder sets*** are the sets

$$\big\{w \in C(\mathbb{R}_+, \mathbb{R})\,;\ w(t_0) \in A_0, \ldots, w(t_n) \in A_n\big\}$$

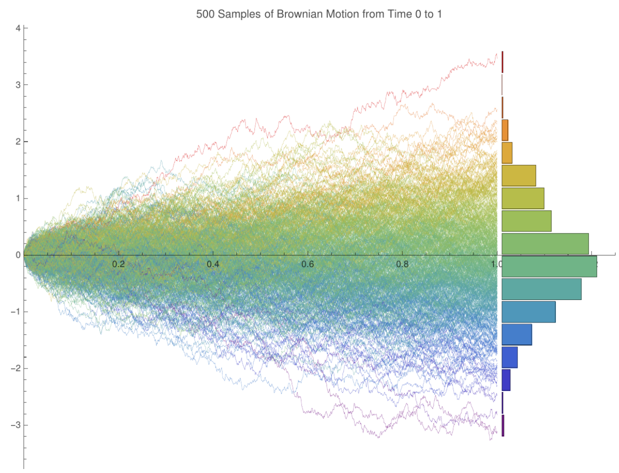


Figure 2.1: Simulation of Brownian motion

for $A_0, \ldots, A_n \in \mathscr{B}(\mathbb{R})$. The class of cylinder sets is obviously closed under finite intersections; by definition, this class generates the σ-field, whence by the π-λ theorem (number 1 on page 262 of the book), the finite-dimensional distributions of W determine W. Thus, there is only one Wiener measure.

Exercise (due 9/14). Exercise 2.25 (time inversion).

Exercise. Suppose that $f \in L^2_{\mathrm{loc}}(\mathbb{R}_+)$, i.e., $f \in L^2\bigl([0,t]\bigr)$ for all $t > 0$. Define the stochastic process $X \colon t \mapsto G\bigl(f\mathbf{1}_{[0,t]}\bigr)$ for a Gaussian white noise G on $\mathbb{R}_+$ with intensity Lebesgue measure. Define $A_t := \int_0^t |f(s)|^2 \, \mathrm{d}s$ and $\tau_t := \inf\{s \geq 0\,;\ A_s > t\}$. Let $\beta_t := X_{\tau_t}$. Show that $(\beta_t)_{0 \leqslant t < A_\infty}$ is a pre-Brownian motion restricted to $[0, A_\infty)$. Let B be a modification of β that is continuous. Write $d(s,t) := \bigl(\int_s^t |f(u)|^2 \, \mathrm{d}u\bigr)^{1/2}$. Show that $Y \colon t \mapsto B_{A_t}$ is a modification of X that is locally Hölder continuous of order α for all $\alpha < 1$ with respect to the pseudometric d on $\mathbb{R}_+$. We call such a process Y a ***Wiener integral process***.

2.3. Properties of Brownian Sample Paths

Lemma. *Let $X_1, X_2, \ldots, X_n$ be random variables and A be an event. Then A is independent of $\sigma(X_1, X_2, \ldots, X_n)$ if and only if*

$$\forall g \in C_{\mathrm{c}}(\mathbb{R}^n, \mathbb{R}) \qquad \mathbf{E}\bigl[\mathbf{1}_A g(X_1, X_2, \ldots, X_n)\bigr] = \mathbf{P}(A)\, \mathbf{E}\bigl[g(X_1, X_2, \ldots, X_n)\bigr].$$

Proof. The law of $(X_1, X_2, \ldots, X_n)$, as a Borel probability measure on $\mathbb{R}^n$, is determined by its integral of $g \in C_{\mathrm{c}}(\mathbb{R}^n)$, as is the law conditional on A. ◂

Let $(B_t)_{t \geqslant 0}$ be a Brownian motion. Write $\mathscr{F}_t := \sigma(B_s, s \leqslant t)$ and $\mathscr{F}_{0^+} := \bigcap_{s>0} \mathscr{F}_s$. This latter describes how Brownian motion "starts". Are there ways it can start that have non-trivial probability? No:

Theorem 2.13 (Blumenthal's 0-1 Law)**.** *$\mathscr{F}_{0^+}$ is trivial in the sense that all its sets have probability* 0 *or* 1.

Proof. We want to show that $\mathscr{F}_{0^+}$ is independent of $\mathscr{F}_{0^+}$, for which it suffices to show that $\mathscr{F}_{0^+}$ is independent of $\mathscr{F}_s$ for some $s > 0$, since $\mathscr{F}_{0^+} \subseteq \mathscr{F}_s$. Take any $s > 0$; since $\mathscr{F}_s = \sigma(B_t, 0 < t \leqslant s)$ (as $B_0 = 0$ a.s.), it suffices to show that $\mathscr{F}_{0^+}$ is independent of $\sigma(B_{t_1}, B_{t_2}, \ldots, B_{t_n})$ for any $0 < t_1 < t_2 < \cdots < t_n \leqslant s$. In light of the above lemma, we calculate, for $g \in C_c(\mathbb{R}^n, \mathbb{R})$ and $A \in \mathscr{F}_{0^+}$,

$$\begin{aligned}
\mathbf{E}\big[\mathbf{1}_A g(B_{t_1}, B_{t_2}, \ldots, B_{t_n})\big] &= \lim_{\varepsilon \downarrow 0} \mathbf{E}\big[\mathbf{1}_A g(B_{t_1} - B_\varepsilon, B_{t_2} - B_\varepsilon, \ldots, B_{t_n} - B_\varepsilon)\big] \\
&\qquad \text{[bounded convergence theorem]} \\
&= \lim_{\varepsilon \downarrow 0} \mathbf{P}(A)\, \mathbf{E}\big[g(B_{t_1} - B_\varepsilon, B_{t_2} - B_\varepsilon, \ldots, B_{t_n} - B_\varepsilon)\big] \\
&\qquad [\mathscr{F}_{0^+} \subseteq \mathscr{F}_\varepsilon \perp\!\!\!\perp \sigma(B_{t_1} - B_\varepsilon, B_{t_2} - B_\varepsilon, \ldots, B_{t_n} - B_\varepsilon) \\
&\qquad \text{for } \varepsilon < t_1 \text{ by Proposition 2.5}] \\
&= \mathbf{P}(A)\, \mathbf{E}\big[g(B_{t_1}, B_{t_2}, \ldots, B_{t_n})\big]. \qquad \blacktriangleleft
\end{aligned}$$

As a corollary, we deduce the following:

Proposition 2.14. (i) *Almost surely,*

$$\forall \varepsilon > 0 \qquad \sup_{0 \leqslant s \leqslant \varepsilon} B_s > 0 \textit{ and } \inf_{0 \leqslant s \leqslant \varepsilon} B_s < 0.$$

(ii) *Almost surely,* $\overline{\lim}_{t \to \infty} B_t = \infty$ *and* $\underline{\lim}_{t \to \infty} B_t = -\infty$.

Note that these *are* random variables since we may restrict to rational times.

Proof. (i) We have

$$\begin{aligned}
\mathbf{P}\Big[\forall \varepsilon > 0 \quad \sup_{0 \leqslant s \leqslant \varepsilon} B_s > 0\Big] &= \mathbf{P}\Big[\bigcap_{\varepsilon > 0} \Big[\sup_{0 \leqslant s \leqslant \varepsilon} B_s > 0\Big]\Big] \\
&= \lim_{\varepsilon \downarrow 0} \mathbf{P}\Big[\sup_{0 \leqslant s \leqslant \varepsilon} B_s > 0\Big] \geqslant \lim_{\varepsilon \downarrow 0} \mathbf{P}[B_\varepsilon > 0] = \frac{1}{2},
\end{aligned}$$

whence the above probability equals one by Theorem 2.13. Symmetry gives the other result.

(ii) Let $Z := \sup_t B_t$. Recall that $B_t^\lambda := \frac{1}{\lambda} B_{\lambda^2 t}$ gives a Brownian motion by Proposition 2.5(ii). Thus, the law of Z is the same as the law of Z/λ for all $\lambda > 0$, which means it is concentrated on $\{0, \infty\}$. By part (i), $Z > 0$ a.s., whence $Z = \infty$ a.s. Therefore, $\overline{\lim}_{t \to \infty} B_t = \infty$ a.s. as well. Symmetry gives the other assertion. ◀

Exercise (due 9/21). Exercise 2.29. In fact, show that for any sequence $(t_k)_{k \geqslant 1} \subset (0, \infty)$ with $t_k \to 0$, we have $\overline{\lim}_{k \to \infty} B_{t_k}/\sqrt{t_k} = \infty$ and $\underline{\lim}_{k \to \infty} B_{t_k}/\sqrt{t_k} = -\infty$ almost surely.

Exercise. Show that for any sequence $(t_k)_{k \geqslant 1} \subset (0, \infty)$ with $t_k \to \infty$, we have $\overline{\lim}_{k \to \infty} B_{t_k}/\sqrt{t_k} = \infty$ and $\underline{\lim}_{k \to \infty} B_{t_k}/\sqrt{t_k} = -\infty$ almost surely.

Exercise. Show that the tail σ-field $\bigcap_{t > 0} \sigma(B_s, s \geqslant t)$ is trivial.

Another corollary:

Corollary 2.15. *Almost surely, Brownian motion is not monotone on any nontrivial interval.*

Proof. By Proposition 2.5(iii),

$$\forall t \geqslant 0 \qquad \mathbf{P}\Big[\forall \varepsilon > 0 \quad \sup_{t \leqslant s \leqslant t+\varepsilon} B_s > B_t \text{ and } \inf_{t \leqslant s \leqslant t+\varepsilon} B_s < B_t\Big] = 1.$$

Apply this to $t \in \mathbb{Q}_+$. ◂

Of course, Brownian motion does have local maxima and minima.

We give two last properties that do not depend on Theorem 2.13:

Proposition 2.16. *Fix $t > 0$. If $0 = t_0^n < t_1^n < \cdots < t_{t_{p_n}}^n = t$ satisfies*

$$\lim_{n\to\infty} \max_{1\leqslant i\leqslant p_n} (t_i^n - t_{i-1}^n) = 0,$$

then

$$\lim_{n\to\infty} \sum_{i=1}^{p_n} (B_{t_i^n} - B_{t_{i-1}^n})^2 = t \quad \text{in } L^2(\mathbf{P}).$$

Proof. Immediate from Proposition 1.14. ◂

Corollary 2.17. *Almost surely, Brownian motion has infinite variation on every nontrivial interval.*

Proof. As in the proof of Corollary 2.15, it suffices to prove this for each interval $[0, t]$, $t > 0$. By taking a subsequence, we may assume almost sure convergence in Proposition 2.16. Since

$$\sum_{i=1}^{p_n} (B_{t_i^n} - B_{t_{i-1}^n})^2 \leqslant \max_{1\leqslant i\leqslant p_n} \big|B_{t_i^n} - B_{t_{i-1}^n}\big| \cdot \sum_{i=1}^{p_n} \big|B_{t_i^n} - B_{t_{i-1}^n}\big|,$$

the left-hand side tends to t almost surely, and $\lim_{n\to\infty} \max_{1\leqslant i\leqslant p_n} |B_{t_i^n} - B_{t_{i-1}^n}| = 0$ by continuity, the result follows. ◂

Thus, the Wiener integral cannot be defined as an ordinary integral.

2.4. The Strong Markov Property of Brownian Motion

We want to extend the Markov property, that what happens at times before t is independent of the increments after time t, by replacing t with a suitable class of random times, T. Clearly, such T should not "depend on the future".

Remark. This T is not the index set of t.

Define $\mathscr{F}_\infty := \sigma(B_s, s \geqslant 0)$.

Definition 2.18. *A $[0, \infty]$-valued random variable T is a* ***stopping time*** *if*

$$\forall t \geqslant 0 \qquad [T \leqslant t] \in \mathscr{F}_t.$$

Examples. Two examples of stopping times:

(1) $T \equiv s$;

(2) $T = T_a := \inf\{s \geqslant 0\,;\ B_s = a\}$ is a stopping time, since $T_a \leqslant t$ if and only if

$$\inf_{s \in \mathbb{Q} \cap [0,t]} |B_s - a| = 0.$$

If T is a stopping time, then

$$\forall t > 0 \quad [T < t] = \bigcup_{s \in \mathbb{Q} \cap [0,t)} [T \leqslant s] \in \mathscr{F}_t.$$

What is the σ-field of events "determined up to time T"? We might guess A is such an event if for each $t \geqslant 0$, $A \cap [T = t] \in \mathscr{F}_t$. But we know it might be problematic to make such a fine disintegration of A. Perhaps it would be better to require $A \cap [T \leqslant t] \in \mathscr{F}_t$. Moreover, this is enough at the intuitive level since then $A \cap [T < t] \in \mathscr{F}_t$ and so $A \cap [T = t] \in \mathscr{F}_t$.

Definition 2.19. *If T is a stopping time, the* ***σ-field of the past before*** *T is*

$$\mathscr{F}_T := \left\{A \in \mathscr{F}_\infty\,;\ \forall t \geqslant 0 \quad A \cap [T \leqslant t] \in \mathscr{F}_t\right\}.$$

It is easy to check that

(1) $\mathscr{F}_T$ is a σ-field, and

(2) T is $\mathscr{F}_T$-measurable.

What is Brownian motion at time T? When $T = \infty$, this makes no sense, so define

$$\widetilde{B}_T(\omega) := \begin{cases} B_{T(\omega)}(\omega) & \text{if } T(\omega) < \infty, \\ 0 & \text{if } T(\omega) = \infty. \end{cases}$$

We claim that $\widetilde{B}_T$ is $\mathscr{F}_T$-measurable. We use the left-continuity of B to write

$$\widetilde{B}_T = \lim_{n \to \infty} \sum_{i \geqslant 0} \mathbf{1}_{[\frac{i}{n} \leqslant T < \frac{i+1}{n}]} B_{\frac{i}{n}} = \lim_{n \to \infty} \sum_{i \geqslant 0} \mathbf{1}_{[T < \frac{i+1}{n}]} \mathbf{1}_{[\frac{i}{n} \leqslant T]} B_{\frac{i}{n}}.$$

Thus, we see it suffices to show that

$$\forall s \geqslant 0 \qquad \mathbf{1}_{[s \leqslant T]} B_s \in \mathscr{F}_T.$$

Indeed, if $A \in \mathscr{B}(\mathbb{R})$ and $0 \notin A$, then for each $t \geqslant 0$,

$$\left[\mathbf{1}_{[s \leqslant T]} B_s \in A\right] \cap [T \leqslant t] = \begin{cases} \varnothing & \text{if } t < s, \\ [B_s \in A] \cap [T < s]^{\mathrm{c}} \cap [T \leqslant t] & \text{if } t \geqslant s \end{cases}$$
$$\in \mathscr{F}_t.$$

In case $0 \in A$, just use A^{c} in what we just established. This gives our claim.

Theorem 2.20 (Strong Markov Property)**.** *Let T be a stopping time with $\mathbf{P}[T < \infty] > 0$. Define*

$$\widetilde{B}_t^{(T)} := \widetilde{B}_{T+t} - \widetilde{B}_T \quad (t \geqslant 0).$$

Then under $\mathbf{P}[\,\cdot \mid T < \infty]$, the process $\big(\widetilde{B}_t^{(T)}\big)_{t\geqslant 0}$ is a Brownian motion independent of $\mathscr{F}_T$.

Proof. Suppose first $T < \infty$ a.s. The assertions will follow from

$$\begin{aligned} &\forall A \in \mathscr{F}_T \ \ \forall 0 \leqslant t_1 < \cdots < t_p \ \ \forall F \in C_c(\mathbb{R}^p, \mathbb{R}) \\ &\qquad \mathbf{E}\Big[\mathbf{1}_A F\big(\widetilde{B}_{t_1}^{(T)}, \widetilde{B}_{t_2}^{(T)}, \ldots, \widetilde{B}_{t_p}^{(T)}\big)\Big] = \mathbf{P}(A) \cdot \mathbf{E}\big[F(B_{t_1}, B_{t_2}, \ldots, B_{t_p})\big]. \end{aligned} \tag{2.1}$$

For then $A := \Omega$ shows that $\big(\widetilde{B}_t^{(T)}\big)_{t\geqslant 0}$ has the same finite-dimensional distributions as Brownian motion, so by Proposition 2.3 is a pre-Brownian motion. Sample paths are continuous, so it is a Brownian motion. Also, Eq. (2.1) shows that $(\widetilde{B}_{t_1}^{(T)}, \widetilde{B}_{t_2}^{(T)}, \ldots, \widetilde{B}_{t_p}^{(T)})$ is independent of $\mathscr{F}_T$. So by the π-λ theorem, $\widetilde{B}^{(T)}$ is independent of $\mathscr{F}_T$.

To show Eq. (2.1), we use the following notation: $\lceil t \rceil_n := \frac{\lceil nt \rceil}{n} \geqslant t$. Now the bounded convergence theorem yields

$$\begin{aligned} \mathbf{E}\Big[\mathbf{1}_A F\big(\widetilde{B}_{t_1}^{(T)}, \widetilde{B}_{t_2}^{(T)}, \ldots, \widetilde{B}_{t_p}^{(T)}\big)\Big] &= \lim_{n\to\infty} \mathbf{E}\Big[\mathbf{1}_A F\big(\widetilde{B}_{t_1}^{(\lceil T\rceil_n)}, \widetilde{B}_{t_2}^{(\lceil T\rceil_n)}, \ldots, \widetilde{B}_{t_p}^{(\lceil T\rceil_n)}\big)\Big] \\ &= \lim_{n\to\infty} \sum_{k=0}^{\infty} \mathbf{E}\Big[\mathbf{1}_A \mathbf{1}_{\left[\frac{k-1}{n} < T \leqslant \frac{k}{n}\right]} F\big(B_{\frac{k}{n}+t_1} - B_{\frac{k}{n}}, B_{\frac{k}{n}+t_2} - B_{\frac{k}{n}}, \ldots, B_{\frac{k}{n}+t_p} - B_{\frac{k}{n}}\big)\Big]. \end{aligned}$$

Note that

$$A \cap \left[\tfrac{k-1}{n} < T \leqslant \tfrac{k}{n}\right] = \underbrace{A \cap \left[T \leqslant \tfrac{k}{n}\right]}_{\in \mathscr{F}_{k/n} \text{ since } A \in \mathscr{F}_T} \cap \underbrace{\left[T \leqslant \tfrac{k-1}{n}\right]^{c}}_{\in \mathscr{F}_{\frac{k-1}{n}} \subseteq \mathscr{F}_{\frac{k}{n}}} \in \mathscr{F}_{\frac{k}{n}}.$$

Thus, the kth term in the sum equals

$$\mathbf{P}\Big[A \cap \left[\tfrac{k-1}{n} < T \leqslant \tfrac{k}{n}\right]\Big]\, \mathbf{E}\Big[F\big(B_{t_1}, B_{t_2}, \ldots, B_{t_p}\big)\Big].$$

Summing over k gives Eq. (2.1).

In case $\mathbf{P}[T = \infty] > 0$, the same arguments work with $A \cap [T < \infty]$ in place of A, yielding Eq. (2.1) for such sets in $\mathscr{F}_T$, and this gives the result similarly. ◂

When $T < \infty$ a.s., we will omit the tildes in $\widetilde{B}_T$ and $\widetilde{B}^{(T)}$.

A very nice application of the strong Markov property is the reflection principle.

Theorem 2.21. *For $t > 0$, write $S_t := \max_{0\leqslant s\leqslant t} B_s \geqslant 0$. Then*

$$\forall a \geqslant 0 \ \ \forall b \in (-\infty, a] \qquad \mathbf{P}[S_t \geqslant a, B_t \leqslant b] = \mathbf{P}[B_t \geqslant 2a - b].$$

Moreover, $S_t \overset{\mathscr{D}}{=} |B_t|$.

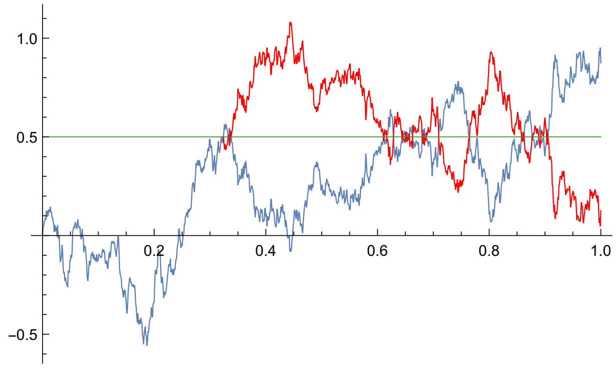


Figure 2.2: Illustration of the reflection principle

Proof. We use the stopping time $T_a := \inf\{t \geqslant 0\,;\ B_t = a\}$. By Proposition 2.14, $T_a < \infty$ almost surely. We have by Theorem 2.20,

$$\begin{aligned}\mathbf{P}[S_t \geqslant a, B_t \leqslant b] = \mathbf{P}[T_a \leqslant t, B_t \leqslant b] &= \mathbf{P}\big[T_a \leqslant t, B^{(T_a)}_{t-T_a} \leqslant b-a\big] \\ &= \mathbf{P}[T_a \leqslant t] \cdot \mathbf{P}\big[B^{(T_a)}_{t-T_a} \leqslant b-a \,\big|\, T_a \leqslant t\big] \\ &= \mathbf{P}[T_a \leqslant t] \cdot \mathbf{P}\big[-B^{(T_a)}_{t-T_a} \leqslant b-a \,\big|\, T_a \leqslant t\big] \\ &= \mathbf{P}[T_a \leqslant t] \cdot \mathbf{P}\big[B^{(T_a)}_{t-T_a} \geqslant a-b \,\big|\, T_a \leqslant t\big] = \mathbf{P}\big[T_a \leqslant t, B^{(T_a)}_{t-T_a} \geqslant a-b\big] \\ &= \mathbf{P}[T_a \leqslant t, B_t \geqslant 2a-b] = \mathbf{P}[B_t \geqslant 2a-b]\end{aligned}$$

since $2a-b \geqslant a$. The crucial fourth equality uses the independence of $B^{(T_a)}$ and $\mathscr{F}_{T_a}$. It follows that

$$\mathbf{P}[S_t \geqslant a] = \mathbf{P}[S_t \geqslant a, B_t \geqslant a] + \mathbf{P}[S_t \geqslant a, B_t \leqslant a] = 2\,\mathbf{P}[B_t \geqslant a] = \mathbf{P}\big[|B_t| \geqslant a\big]. \qquad ◀$$

Exercise (due 9/28). (1) Exercise 2.28.
(2) Prove (2.2) in the book.

Corollary 2.22. *$\forall a \neq 0$ $T_a \overset{\mathscr{D}}{=} \dfrac{a^2}{B_1^2}$ and $\mathbf{E}\big[\sqrt{T_a}\,\big] = \infty$.*

Proof. We may assume by symmetry that $a > 0$. For each $t \geqslant 0$,

$$\begin{aligned}\mathbf{P}[T_a \leqslant t] = \mathbf{P}[S_t \geqslant a] &= \underbrace{\mathbf{P}\big[|B_t| \geqslant a\big]}_{\text{[Theorem 2.21]}} \\ &= \mathbf{P}\big[(B_t)^2 \geqslant a^2\big] = \mathbf{P}[t(B_1)^2 \geqslant a^2] = \mathbf{P}\Big[\frac{a^2}{(B_1)^2} \leqslant t\Big].\end{aligned}$$

Therefore,

$$\mathbf{E}\big[\sqrt{T_a}\,\big] = \mathbf{E}\big[a/|B_1|\big] = a\int_{-\infty}^{\infty} \frac{p_X(x)}{|x|}\,\mathrm{d}x = \infty,$$

where p_X is the standard normal density. ◀

Exercise (due 9/28). Verify the density in Corollary 2.22 in the book.

An amusing and immediate consequence of Corollary 2.22 is that $\mathbf{E}[T_a^{-1}] = a^{-2}$.

Exercise. Show that for $a \neq 0$, if $S_a := \sup\{s\,;\ B_s = as\}$, then $S_a \overset{\mathscr{D}}{=} B_1^2/a^2$. *Hint:* use the result of Exercise 2.25.

Exercise. Let $X_t := \int_0^t B_s \, \mathrm{d}s$ be ***integrated Brownian motion***. Show that almost surely, $\overline{\lim}_{t\to\infty} X_t = \infty$ and $\underline{\lim}_{t\to\infty} X_t = -\infty$. *Hint:* For a finite stopping time T, write $X_t = X_{t\wedge T} + (t-T)^+ B_T + Y_{(t-T)^+}$ with Y a copy of X that is independent of $\mathscr{F}_T$. Use $T_n := \inf\{t \geq n\,;\ B_t = 0\}$ to show that $\mathbf{P}[\sup_t X_t = \infty] \in \{0,1\}$. Use $T := \inf\{t \geq 1\,;\ B_t = -1\}$ to show that $\mathbf{P}[\sup_t X_t = \infty] = 1$.

We now extend Brownian motion to initial values other than 0 and to finite dimensions.

Definition 2.23. *If Z is an $\mathbb{R}$-valued random variable and B is a Brownian motion independent of Z, then we call $(Z + B_t)_{t\geqslant 0}$ a **real Brownian motion started from** Z.*

Definition 2.24. *If $B^1, \ldots, B^d$ are independent real Brownian motions started from 0, then we call $\big((B_t^1, \ldots, B_t^d)\big)_{t\geqslant 0}$ a d**-dimensional Brownian motion started from** 0. If we add an independent starting vector, Z, then we get d**-dimensional Brownian motion started from** Z.*

Note that by Corollary 2.4, if B is a d-dimensional Brownian motion (from 0), then for $0 = t_0 < t_1 < \cdots < t_n$ and $x^1 = (x_1^1, \ldots, x_d^1), \ldots, x^n = (x_1^n, \ldots, x_d^n)$, the density of $(B_{t_1}, \ldots, B_{t_n})$ at $(x^1, \ldots, x^n)$ is

$$\prod_{k=1}^d \prod_{i=1}^n \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \cdot \exp\Big\{-\sum_{i=1}^n \frac{(x_k^i - x_k^{i-1})^2}{2(t_i - t_{i-1})}\Big\} = \prod_{i=1}^n \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}^d} \cdot \exp\Big\{-\sum_{i=1}^n \frac{|x^i - x^{i-1}|^2}{2(t_i - t_{i-1})}\Big\}.$$

This is invariant under isometries of $\mathbb{R}^d$. Therefore, the law of d-dimensional Brownian motion (started at 0) is invariant under isometries of $\mathbb{R}^d$ that fix 0. Thus, we really have E-valued Brownian motion for finite-dimensional inner-product spaces, E.

It is easy to check that Blumenthal's 0-1 law and the strong Markov property hold for d-dimensional Brownian motion, where now a stopping time is defined with respect to the collection of σ-fields

$$\mathscr{F}_t := \sigma\big((B_s^1, \ldots, B_s^d), s \leqslant t\big).$$

The proofs are the same but with more notation.

Appendix: The Cameron–Martin Theorem

Let B be a Brownian motion. How does adding drift to B change its law? Let G be the corresponding Gaussian white noise. Let $f \in L^2(\mathbb{R}_+)$, and denote $F_t := \int_0^t f(s)\,\mathrm{d}s$. We will consider the process $X := B + F$. We claim that the law of X is absolutely continuous with respect to the law of B; in fact, the law of X is equal to the law of B with respect to $\mathrm{e}^{G(f) - \|f\|_{L^2}^2/2}\,\mathbf{P}$; this result is due to Cameron and Martin. To be even more explicit, let W be Wiener measure. Recall that $\int f(s)\,\mathrm{d}w(s)$ is defined for W-a.e. w as in Proposition 2.3 (iv)⇒(i), using step functions and limits.

Proposition 5.24. *For $f \in L^2(\mathbb{R}_+)$ and $F_t := \int_0^t f(s)\,\mathrm{d}s$, we have for all measurable $A \subseteq C(\mathbb{R}_+, \mathbb{R})$,*

$$\int \mathrm{d}W(w)\, \mathbf{1}_{[w+F\in A]} = \int \mathrm{d}W(w)\, \mathbf{1}_{[w\in A]}\, \mathrm{e}^{\int f\,\mathrm{d}w - \|f\|^2/2}.$$

That is, the $\mathbf{P}$*-law of* X *is absolutely continuous with respect to* W*, having Radon–Nikodym derivative* $w \mapsto e^{\int f\, dw - \|f\|^2/2}$.

This allows us to conclude, for example, that every $\mathbf{P}$-a.s. property of B also holds for X.

The class of F that are absolutely continuous with derivative in $L^2(\mathbb{R}_+)$ and with $F_0 = 0$ is known as the ***Cameron–Martin*** space, $\mathscr{H}$. It is easy to see that $\mathscr{H}$ is dense in $C_*(\mathbb{R}_+, \mathbb{R}) := \{F \in C(\mathbb{R}_+, \mathbb{R})\,;\ F_0 = 0\}$. It follows that the support of W is all of $C_*(\mathbb{R}_+, \mathbb{R})$: Indeed, Proposition 5.24 tells us that the support of W is unchanged by addition of any function in $\mathscr{H}$, because the Radon–Nikodym derivative is nonzero W-a.s. Thus, if w_0 is one point of the support of W, then $w_0 + \mathscr{H}$ also lies in the support. Since the closure of $w_0 + \mathscr{H}$ is $C_*(\mathbb{R}_+, \mathbb{R})$, our claim follows.

Proposition 5.24 is actually a very simple consequence of basic manipulations with Gaussian random variables. Consider any (centered) Gaussian space, $H \subset L^2(\mathbf{P})$, and any nonzero $Y \in H$. Define $\phi\colon H \to L^2(\mathbf{P})$ by $\phi(Z) := Z + \langle Z, Y\rangle_{L^2(\mathbf{P})}$. Obviously $\phi(Z) = Z$ whenever $Z \perp Y$, which is the same as $Z \perp\!\!\!\perp Y$, whereas $\phi(Y) = Y + \|Y\|^2_{L^2(\mathbf{P})}$. The law of $\phi(Y)$, i.e., $\mathscr{N}\big(\|Y\|^2, \|Y\|^2\big)$, is absolutely continuous with respect to that of Y with Radon–Nikodym derivative $y \mapsto e^{y - \|Y\|^2/2}$. Therefore, if $(Z_1, \ldots, Z_n) \in Y^\perp$, then the law of $\big(\phi(Z_1), \ldots, \phi(Z_n), \phi(Y)\big)$ also has Radon–Nikodym derivative $(z_1, \ldots, z_n, y) \mapsto e^{y - \|Y\|^2/2}$ with respect to that of $(Z_1, \ldots, Z_n, Y)$. In other words, the $\mathbf{P}$-law of $\big(\phi(Z_1), \ldots, \phi(Z_n), \phi(Y)\big)$ is equal to the $e^{Y - \|Y\|^2/2}\,\mathbf{P}$-law of $(Z_1, \ldots, Z_n, Y)$. Since this determines the finite-dimensional distributions of all of H, we conclude that the $\mathbf{P}$-law of $\big(\phi(Z)\big)_{Z \in H}$ is equal to the $e^{Y - \|Y\|^2/2}\,\mathbf{P}$-law of $(Z)_{Z \in H}$.

Coming back to Brownian motion, let us apply this general result to H being the image of the Gaussian white noise, G. Note that $F_t = \langle \mathbf{1}_{[0,t]}, f\rangle_{L^2} = \langle B_t, G(f)\rangle_{L^2(\mathbf{P})}$. Thus, we are exactly in the situation just analyzed: $X_t = \phi(B_t)$. Therefore, the $\mathbf{P}$-law of X is the $e^{G(f) - \|f\|^2/2}\,\mathbf{P}$-law of B, as claimed.

Exercise. Deduce that X is a Brownian motion with respect to $\mathbf{Q} := e^{-G(f) - \|f\|^2_{L^2}/2}\,\mathbf{P}$. Alternatively, give a direct proof of this property by showing that X is a pre-Brownian motion for the Gaussian white noise $\widetilde{G}\colon L^2(\mathbb{R}_+) \to L^2(\mathbf{Q})$ defined by $\widetilde{G}(h) := G(h) + \langle h, f\rangle_{L^2}$.

If F is not in the Cameron–Martin space, then the laws of $X = B + F$ and B are mutually singular, a result of Segal. This is obvious if $F_0 \neq 0$. When $F_0 = 0$, note that $B_t \mapsto F_t$ extends uniquely to a linear functional, γ, on the linear span V of $\{B_t\,;\ t > 0\}$, because the random variables B_t are linearly independent. This map γ is bounded, i.e., $\exists C < \infty$ such that $|\gamma(Z)| \leqslant C\|Z\|$ for all $Z \in V$, iff γ extends continuously to the closure of V, which is equivalent to $\gamma(Z) = \langle Z, Y\rangle$ for some $Y \in H$, i.e., $F \in \mathscr{H}$. Thus, if $F \notin \mathscr{H}$, then there exist $Z \in V$ of norm 1 with arbitrarily large $\gamma(Z)$. Let Φ be the c.d.f. of the standard normal distribution. For $\|Z\| = 1$, we have $\mathbf{P}\big[Z > \gamma(Z)/2\big] = 1 - \Phi\big(\gamma(Z)/2\big) = \mathbf{P}\big[Z + \gamma(Z) \leqslant \gamma(Z)/2\big]$. This leads us to choose Z_n such that $\|Z_n\| = 1$ and $\alpha_n := \gamma(Z_n)$ satisfies $\sum_n \big[1 - \Phi(\alpha_n/2)\big] < \infty$. Let $\xi_n := Z_n + \gamma(Z_n)$. Then $Z_n > \alpha_n/2$ for only finitely many n a.s., whereas $\xi_n \leqslant \alpha_n/2$ for only finitely many n a.s. The explicit forms of Z_n and ξ_n are $Z_n = \sum_{i=1}^{k_n} a_{n,i} B_{t_{n,i}}$ and $\xi_n = \sum_{i=1}^{k_n} a_{n,i} X_{t_{n,i}}$ for some constants $a_{n,i}$ and times $t_{n,i}$. Thus the laws of X and B are mutually singular.

Exercise. Let $F \in \mathscr{H}$ with F' having bounded variation on $[0, t]$ for some $t > 0$. Show that

$$\lim_{\varepsilon \downarrow 0} \frac{\mathbf{P}\big[\|B - F\|_{L^\infty[0,t]} \leqslant \varepsilon\big]}{\mathbf{P}\big[\|B\|_{L^\infty[0,t]} \leqslant \varepsilon\big]} = \exp\{-\tfrac{1}{2}\|F'\|^2_{L^2[0,t]}\}.$$

Note that the denominator here is positive, because W has full support. See the discussion of the exercise on page 79 for the value of the denominator; it is asymptotic to $\frac{4}{\pi} \exp\big\{-\frac{\pi^2 t}{8\varepsilon^2}\big\}$ as $\varepsilon \downarrow 0$.

Exercise. For a function $F : \mathbb{R}_+ \to \mathbb{R}$, define

$$M(F) := \sup \sum_i \frac{\big[F(t_{i+1}) - F(t_i)\big]^2}{t_{i+1} - t_i},$$

where the supremum is over all sequences $(t_i)_i$ with $0 \leqslant t_1 < t_2 < \cdots$.

(1) Show that if $F \in \mathscr{H}$ with derivative F', then $M(F) \leqslant \|F'\|^2$.

(2) Show that if $M(F) < \infty$ and $(s_i, t_i]$ are disjoint intervals in $\mathbb{R}_+$, then

$$\sum_i |F(t_i) - F(s_i)| \leq \Big(M(F) \sum_i (t_i - s_i)\Big)^{1/2},$$

and deduce that F is absolutely continuous.

(3) Show that if $M(F) < \infty$, then $F \in \mathscr{H}$ with $\|F'\|^2 \leqslant M(F)$.

We conclude that $M(F) < \infty$ iff $F \in \mathscr{H}$, in which case $M(F) = \|F'\|^2$.

Exercise. For $F : \mathbb{R}_+ \to \mathbb{R}$, let $\mathcal{T}F$ be the function $t \mapsto tF(1/t)$ for $t > 0$ and $0 \mapsto 0$. Note that if $F(0) = 0$, then $\mathcal{T}\mathcal{T}F = F$. Let $\langle F, K\rangle_{\mathscr{H}} := \int_0^\infty F'(t)K'(t)\,\mathrm{d}t$ be the natural inner product on $\mathscr{H}$, making $\mathscr{H}$ a Hilbert space.

(1) Show that if $F, K \in \mathscr{H}$ are continuously differentiable with compact support in $(0, \infty)$, then $\langle F, \mathcal{T}K\rangle_{\mathscr{H}} = \langle \mathcal{T}F, K\rangle_{\mathscr{H}}$.

(2) Show that if $F \in \mathscr{H}$ is continuously differentiable with compact support in $(0, \infty)$, then $\|\mathcal{T}F\|_{\mathscr{H}} = \|F\|_{\mathscr{H}}$.

(3) Show that if $F \in \mathscr{H}$, then $\mathcal{T}F \in \mathscr{H}$ with $\|\mathcal{T}F\|_{\mathscr{H}} = \|F\|_{\mathscr{H}}$.

(4) Give another proof of (3) by using the fact that $\mathcal{T}B$ is a Brownian motion, together with Proposition 5.24.

Chapter 3

Filtrations and Martingales

Please review martingales in discrete time, Appendix A2 of the book.

Here, we select just a few things to review. Time is usually $\mathbb{N} = \{0, 1, \dots\}$. We are given an increasing sequence $(\mathscr{G}_n)_{n\in\mathbb{N}}$ of sub-σ-fields. For a sequence $(Y_n)_{n\in\mathbb{N}}$ of integrable random variables with $Y_n \in \mathscr{G}_n$, e.g., $\mathscr{G}_n := \sigma(Y_0, Y_1, \dots, Y_n)$, we call $(Y_n)_{n\in\mathbb{N}}$

(1) a ***martingale*** if $\mathbf{E}[Y_n \mid \mathscr{G}_m] = Y_m$ when $0 \leqslant m \leqslant n$;

(2) a ***submartingale*** if $\mathbf{E}[Y_n \mid \mathscr{G}_m] \geqslant Y_m$ when $0 \leqslant m \leqslant n$;

(3) a ***supermartingale*** if $\mathbf{E}[Y_n \mid \mathscr{G}_m] \leqslant Y_m$ when $0 \leqslant m \leqslant n$.

If $M^Y_{ab}(n)$ denotes the number of upcrossings by $(Y_0, Y_1, \dots, Y_n)$ of an interval $[a, b]$ $(a < b)$, then one version of Doob's upcrossing inequality is that for a supermartingale, $(Y_n)_{n\in\mathbb{N}}$,

$$\forall n \in \mathbb{N}\ \ \forall a < b \qquad \mathbf{E}\big[M^Y_{ab}(n)\big] \leqslant \frac{\mathbf{E}\big[(Y_n - a)^-\big]}{b - a}.$$

There is a corresponding version for submartingales.

The maximal inequality given in Appendix A2 can be hard to find, so here is a proof. The inequality states that for a submartingale or supermartingale $(Y_n)_{n\in\mathbb{N}}$,

$$\forall k \in \mathbb{N}\ \ \forall \lambda > 0 \qquad \lambda\, \mathbf{P}\Big[\max_{0\leqslant n\leqslant k} |Y_n| \geqslant \lambda\Big] \leqslant \mathbf{E}\big[|Y_0|\big] + 2\,\mathbf{E}\big[|Y_k|\big].$$

We may assume that Y is a supermartingale, because if not, then $-Y$ is a supermartingale and $|-Y| = |Y|$. The desired inequality will follow from adding the following two inequalities:

$$\forall k \in \mathbb{N}\ \ \forall \lambda > 0 \qquad \lambda\, \mathbf{P}\Big[\max_{0\leqslant n\leqslant k} Y_n \geqslant \lambda\Big] \leqslant \mathbf{E}[Y_0] + \mathbf{E}\big[|Y_k|\big]$$

and

$$\forall k \in \mathbb{N}\ \ \forall \lambda > 0 \qquad \lambda\, \mathbf{P}\Big[\min_{0\leqslant n\leqslant k} Y_n \leqslant -\lambda\Big] \leqslant \mathbf{E}\big[|Y_k|\big].$$

Fix $k \in \mathbb{N}$ and $\lambda > 0$. Let $T := \inf\{n\,;\ Y_n \geqslant \lambda\} \wedge k$. By the optional stopping theorem, we have

$$\mathbf{E}[Y_0] \geqslant \mathbf{E}[Y_T] = \mathbf{E}\big[Y_T \mathbf{1}_{[\max_{0\leqslant n\leqslant k} Y_n \geqslant \lambda]}\big] + \mathbf{E}\big[Y_T \mathbf{1}_{[\max_{0\leqslant n\leqslant k} Y_n < \lambda]}\big] \geqslant \lambda\, \mathbf{P}\Big[\max_{0\leqslant n\leqslant k} Y_n \geqslant \lambda\Big] - \mathbf{E}\big[|Y_k|\big],$$

which gives the first inequality. For the second, define $T := \inf\{n\,;\ Y_n \leqslant -\lambda\} \wedge k$. By the optional stopping theorem, we have

$$\mathbf{E}[Y_k] \leqslant \mathbf{E}[Y_T] = \mathbf{E}\big[Y_T \mathbf{1}_{[\min_{0\leqslant n\leqslant k} Y_n \leqslant -\lambda]}\big] + \mathbf{E}\big[Y_T \mathbf{1}_{[\min_{0\leqslant n\leqslant k} Y_n > -\lambda]}\big] \leqslant -\lambda\, \mathbf{P}\Big[\min_{0\leqslant n\leqslant k} Y_n \leqslant -\lambda\Big] + \mathbf{E}\big[|Y_k|\big],$$

which gives the second inequality.

3.1. Filtrations and Processes

Let $(\Omega, \mathscr{F}, \mathbf{P})$ be a probability space.

Definition 3.1. *A* ***filtration*** *on* $(\Omega, \mathscr{F}, \mathbf{P})$ *is a collection* $(\mathscr{F}_t)_{0\leqslant t\leqslant\infty}$ *of sub-σ-fields of* $\mathscr{F}$ *such that* $\mathscr{F}_s \subseteq \mathscr{F}_t$ *for* $0 \leqslant s \leqslant t \leqslant \infty$.

We also call $\big(\Omega, \mathscr{F}, (\mathscr{F}_t)_{0\leqslant t\leqslant\infty}, \mathbf{P}\big)$ a ***filtered probability space***.

Example. In Chapter 2, we used the filtration associated to Brownian motion

$$\mathscr{F}_t = \sigma(B_s, 0 \leqslant s \leqslant t), \quad \mathscr{F}_\infty = \sigma(B_s, s \geqslant 0).$$

Example. More generally, if $(X_t)_{t\geqslant 0}$ is any stochastic process, then its ***canonical filtration*** is

$$\mathscr{F}_t^X := \sigma(X_s, 0 \leqslant s \leqslant t), \quad \mathscr{F}_\infty^X := \sigma(X_s, s \geqslant 0).$$

These are not the only filtrations of interest, since there may be other stochastic processes we want to include, or other randomness.

Similar to $\mathscr{F}_{0^+}$ that we considered in Chapter 2, define

$$\mathscr{F}_{t^+} := \bigcap_{s>t} \mathscr{F}_s, \quad \mathscr{F}_{\infty^+} := \mathscr{F}_\infty.$$

Clearly, $(\mathscr{F}_{t^+})_{0\leqslant t\leqslant\infty}$ is a filtration and $\mathscr{F}_t \subseteq \mathscr{F}_{t^+}$. If $\mathscr{F}_t = \mathscr{F}_{t^+}$ for each $t \geqslant 0$, then we say that $(\mathscr{F}_t)_{0\leqslant t\leqslant\infty}$ is ***right-continuous***.

Example. Let $(\mathscr{F}_t)_{t\geqslant 0}$ be the canonical filtration of a Poisson process, where the process is modified so as to be left-continuous. Then $\mathscr{F}_t \neq \mathscr{F}_{t^+}$ for every $t > 0$.

A filtration $(\mathscr{F}_t)_{0\leqslant t\leqslant\infty}$ is ***complete*** if $\mathscr{F}_0$ contains every subset of each $\mathbf{P}$-negligible set of $\mathscr{F}_\infty$. Every filtration can be ***completed*** to a filtration $(\mathscr{F}_t')_{0\leqslant t\leqslant\infty}$, where $\mathscr{F}_t' := \sigma(\mathscr{F}_t, \mathcal{N})$ and $\mathcal{N}$ is the collection of $(\mathscr{F}_\infty, \mathbf{P})$-negligible sets (those $A \subseteq B \in \mathscr{F}_\infty$ with $\mathbf{P}(B) = 0$).

In discrete time, there are no pesky issues of measurability, other than $X_n \in \mathscr{G}_n$. Now, however, there are additional issues. We say $(X_t)_{t\geqslant 0}$ is ***adapted*** to $(\mathscr{F}_t)_{0\leqslant t\leqslant\infty}$ if $\forall t \geqslant 0\ \ X_t \in \mathscr{F}_t$. We will want, e.g., to integrate a stochastic process and get a random variable. This requires some joint measurability. We will also want the result to be an adapted process. These properties will hold automatically when (X_t) has continuous sample paths.

Definition 3.2. *A process* $(X_t)_{t\geqslant 0}$ *with values in a measurable space* $(E, \mathscr{E})$ *is* ***(jointly) measurable*** *if*

$$\begin{gathered}(\omega, t) \mapsto X_t(\omega) \\ \big(\Omega \times \mathbb{R}_+, \mathscr{F} \otimes \mathscr{B}(\mathbb{R}_+)\big) \to (E, \mathscr{E})\end{gathered}$$

is measurable.

Fix a filtered probability space.

Definition 3.3. *A set $A \subseteq \Omega \times \mathbb{R}_+$ is called* ***progressively measurable****, written $A \in \mathscr{P}$, if*

$$\forall t \geqslant 0 \qquad A \cap \big(\Omega \times [0,t]\big) \in \mathscr{F}_t \otimes \mathscr{B}\big([0,t]\big).$$

The set $\mathscr{P}$ is a σ-field, called the ***progressive σ-field****. We call $(X_t)_{t\geqslant 0}$* ***progressive*** *if*

$$\begin{aligned} (\omega,t) &\mapsto X_t(\omega) \\ (\Omega\times\mathbb{R}_+,\mathscr{P}) &\to (E,\mathscr{E}) \end{aligned}$$

is measurable. Equivalently, $(X_t)_{t\geqslant 0}$ is progressive if for all $t > 0$,

$$\begin{aligned} (\omega,s) &\mapsto X_s(\omega) \\ \big(\Omega\times[0,t],\mathscr{F}_t\otimes\mathscr{B}\big([0,t]\big)\big) &\to (E,\mathscr{E}) \end{aligned}$$

is measurable. Note: every progressive process is measurable and adapted.

Exercise (due 10/5). Let $(\Omega,\mathscr{F},\mathbf{P}) := \big([0,1],\mathscr{L},\mu\big)$, where $\mathscr{L}$ is the collection of Lebesgue-measurable sets and μ is Lebesgue measure. Let $\mathscr{L}_0 := \big\{B\in\mathscr{L}\,;\ \mu(B)\in\{0,1\}\big\}$. Let $\mathscr{F}_t := \mathscr{L}_0$ for each $t\in[0,\infty]$. Define

$$A := \big\{(x,x)\,;\ 0\leqslant x\leqslant \tfrac{1}{2}\big\} \subseteq \Omega\times\mathbb{R}_+.$$

Write $X_t(\omega) := \mathbf{1}_A(\omega,t)$ for $t\geqslant 0$. Show that $(X_t)_{t\geqslant 0}$ is a measurable and adapted process, but is not progressive. *Hint:* show that for each $C\in\mathscr{P}$,

$$\int_{[0,1]} \mathbf{1}_C(x,x)\,\mu(\mathrm{d}x) = \int_{[0,1]^2} \mathbf{1}_C(x,y)\,\mu^{(2)}(\mathrm{d}x,\ \mathrm{d}y).$$

Proposition 3.4. *Let E a metric space. Suppose that $(X_t)_{t\geqslant 0}$ is a stochastic process with values in $\big(E,\mathscr{B}(E)\big)$ that is adapted and has right-continuous sample paths. Then X is progressive. The same holds if "right-continuous" is replaced by "left-continuous".*

Proof. The case of right-continuous is in the book, so we do left-continuous. We approximate $(X_t)_{t\geqslant 0}$ by processes that are easily seen to be progressive and use that the class of progressive processes is closed under limits (this uses that E is a metric space).

For $n\geqslant 1$, define $X_t^n := X_{\frac{\lfloor nt\rfloor}{n}}$; then $\lim_{n\to\infty} X_t^n(\omega) = X_t(\omega)$ for all (ω,t). Also, given $t\geqslant 0$ and $B\in\mathscr{B}(E)$,

$$\begin{aligned} &\big\{(\omega,s)\in\Omega\times[0,t]\,;\ X_s^n(\omega)\in B\big\} \\ &\quad= \bigcup_{0\leqslant k\leqslant nt} \Big(\big\{\omega\,;\ X_{\frac{k}{n}}(\omega)\in B\big\}\times\big([\tfrac{k}{n},\tfrac{k+1}{n})\cap[0,t]\big)\Big) \in \mathscr{F}_t\otimes\mathscr{B}\big([0,t]\big), \end{aligned}$$

whence $(X_t^n)_{t\geqslant 0}$ is progressive. Since $X^n\to X$, so is X. ◀

3.2. Stopping Times and Associated σ-Fields

Definition 3.5. *A random variable $T\colon \Omega \to [0,\infty]$ is a **stopping time** of $(\mathscr{F}_t)_{0\leqslant t\leqslant\infty}$ if $\forall t \geqslant 0$ $[T \leqslant t] \in \mathscr{F}_t$. We write*

$$\mathscr{F}_T := \left\{A \in \mathscr{F}_\infty\,;\ \forall t \geqslant 0\ \ A \cap [T \leqslant t] \in \mathscr{F}_t\right\}$$

*for the **σ-field of the past before** T.*

As we saw for Brownian motion, a stopping time T for $(\mathscr{F}_t)$ also satisfies

$$\forall t \geqslant 0 \quad [T < t] \in \mathscr{F}_t,$$

but this is not sufficient for T to be a stopping time (example: use the canonical filtration for a *left*-continuous Poisson process and let T be the time of the first jump).

Since $\mathscr{F}_t \subseteq \mathscr{F}_{t^+}$, an $(\mathscr{F}_t)$-stopping time is also an $(\mathscr{F}_{t^+})$-stopping time, but not conversely (same example).

Proposition 3.6. *Write $\mathscr{G}_t := \mathscr{F}_{t^+}$ for $t \in [0,\infty]$.*

(i) *The following are equivalent:*

  (a) $\forall t > 0\ \ [T < t] \in \mathscr{F}_t$;

  (b) *T is a $(\mathscr{G}_t)$-stopping time;*

  (c) $\forall t > 0\ \ T \wedge t \in \mathscr{F}_t$.

(ii) *If T is a $(\mathscr{G}_t)$-stopping time, then*

$$\mathscr{G}_T = \left\{A \in \mathscr{F}_\infty\,;\ \forall t > 0\ \ A \cap [T < t] \in \mathscr{F}_t\right\}.$$

We write $\mathscr{F}_{T^+} := \mathscr{G}_T$.

Proof. (i) (a) $\Rightarrow$ (b): $\forall 0 \leqslant t < s$,

$$[T \leqslant t] = \bigcap_{q\in(t,s)\cap\mathbb{Q}} \overbrace{[T < q]}^{\in\mathscr{F}_q\subseteq\mathscr{F}_s} \in \mathscr{F}_s,$$

so $[T \leqslant t] \in \mathscr{G}_t$.

(b) $\Rightarrow$ (c): $\forall 0 < s < t\ \ [T \wedge t \leqslant s] = [T \leqslant s] \in \mathscr{G}_s \subseteq \mathscr{F}_t$, so $T \wedge t \in \mathscr{F}_t$.

(c) $\Rightarrow$ (a): $\forall t > 0$

$$[T < t] = \bigcup_{q\in(0,t)\cap\mathbb{Q}} \overbrace{[T \leqslant q]}^{=[T\wedge t\leqslant q]\in\mathscr{F}_t} \in \mathscr{F}_t.$$

(ii) Similar; see the book. ◂

Here are some easy properties (see the book for proofs):

(a) If T is a stopping time, then $\mathscr{F}_T \subseteq \mathscr{F}_{T^+}$, with equality when $(\mathscr{F}_t)$ is right-continuous.

(b) If $T = t$ is constant, then $\mathscr{F}_T = \mathscr{F}_t$ and $\mathscr{F}_{T^+} = \mathscr{F}_{t^+}$.

(c) If T is a stopping time, then $T \in \mathscr{F}_T$.

(d) Let T be a stopping time, $A \in \mathscr{F}_\infty$ and

$$T^A(\omega) := \begin{cases} T(\omega) & \text{if } \omega \in A, \\ \infty & \text{if } \omega \notin A. \end{cases}$$

Then $A \in \mathscr{F}_T$ if and only if T^A is a stopping time.

(e) If $S \leqslant T$ are stopping times, then $\mathscr{F}_S \subseteq \mathscr{F}_T$ and $\mathscr{F}_{S^+} \subseteq \mathscr{F}_{T^+}$.

(f) If S and T are stopping times, then $S \vee T$ and $S \wedge T$ are stopping times, $\mathscr{F}_{S\wedge T} = \mathscr{F}_S \cap \mathscr{F}_T$, $\mathscr{F}_{S\vee T} = \sigma(\mathscr{F}_S, \mathscr{F}_T)$, $[S \leqslant T] \in \mathscr{F}_{S\wedge T}$, and $[S = T] \in \mathscr{F}_{S\wedge T}$.

(g) If $(S_n)_n$ is a monotone increasing sequence of stopping times, then $\lim_{n\to\infty} S_n$ is a stopping time.

(h) If (S_n) is a monotone decreasing sequence of stopping times, then $S := \lim_{n\to\infty} S_n$ is an $(\mathscr{F}_{t^+})$-stopping time and

$$\mathscr{F}_{S^+} = \bigcap_n \mathscr{F}_{S_n^+}.$$

(i) If (S_n) is a monotone decreasing of stopping times that is eventually constant (stabilizes), then $S := \lim_{n\to\infty} S_n$ is a stopping time and

$$\mathscr{F}_S = \bigcap_n \mathscr{F}_{S_n}.$$

(j) Let T be a stopping time and $Y\colon [T < \infty] \to E$. Then $Y \in \mathscr{F}_T$ if and only if $\forall t \geqslant 0$ $\big(Y \restriction [T \leqslant t]\big) \in \mathscr{F}_t$. (Here, we use implicitly the fact that for any measurable space $(\Omega, \mathscr{F})$ and any $A \subseteq \Omega$, there is an induced σ-field $\{A \cap F\,;\ F \in \mathscr{F}\}$ on A.)

Exercise (due 10/5). Show that $\mathscr{F}_{S\vee T} = \sigma(\mathscr{F}_S, \mathscr{F}_T)$. *Hint:* one may use the fact that

$$A = \big(A \cap [S \leqslant T]\big) \cup \big(A \cap [T \leqslant S]\big).$$

Note that the graph of a measurable function is measurable: if $Y\colon (\Omega, \mathscr{F}) \to (E, \mathscr{E})$ is measurable, then $\mathrm{id} \otimes Y\colon (\Omega, \mathscr{F}) \to (\Omega \times E, \mathscr{F} \otimes \mathscr{E})$, defined by $\omega \mapsto \big(\omega, Y(\omega)\big)$, is measurable.

Here is our first use of progressive measurability:

Theorem 3.7. *Let $(X_t)_{t\geqslant 0}$ be a progressive $(E, \mathscr{E})$-valued process and T be a stopping time. Then $\omega \mapsto X_{T(\omega)}(\omega) =: X_T(\omega)$, defined on $[T < \infty]$, is $\mathscr{F}_T$-measurable.*

Proof. By (j) above, it suffices to verify that $\forall t \geqslant 0$ $\big(X_T \restriction [T \leqslant t]\big) \in \mathscr{F}_t$. Now $X_T \restriction [T \leqslant t]$ is a composition:

$$\omega \mapsto (\omega, T(\omega) \wedge t)$$
$$\big([T \leqslant t], \mathscr{F}_t\big) \to \Big([T \leqslant t] \times [0, t], \mathscr{F}_t \otimes \mathscr{B}\big([0, t]\big)\Big)$$

with

$$(\omega, s) \mapsto X_s(\omega)$$
$$\big(\Omega \times [0,t], \mathscr{F}_t \otimes \mathscr{B}([0,t])\big) \to (E, \mathscr{E}).$$

Both of these are measurable: the first by our observation about graphs and the measurability of $T \wedge t$ from Proposition 3.6(i); the second by definition of progressive measurability. ◀

Note how $\mathscr{F}_T$-measurability dovetails well with progressive measurability. (Actually, it suffices that T be an $(\mathscr{F}_{t^+})$-stopping time.)

We will need to approximate a stopping time by a stopping time that takes discrete values.

If T is a stopping time, $S \leqslant T$, $S \in \mathscr{F}_T$, then S need not be a stopping time. However, $S \geqslant T$ works:

Proposition 3.8. *If T is a stopping time, $S \geqslant T$, and $S \in \mathscr{F}_T$, then S is a stopping time. If T is a stopping time and*

$$T_n := \frac{\lceil 2^n T \rceil}{2^n},$$

then T_n are stopping times with $T_n \downarrow T$.

Proof. $\forall t \geqslant 0$ $[S \leqslant t] = [S \leqslant t] \cap [T \leqslant t] \in \mathscr{F}_t$ since $[S \leqslant t] \in \mathscr{F}_T$; that is, S is a stopping time. The remainder follows since $T \in \mathscr{F}_T$, so $\sigma(T) \subseteq \mathscr{F}_T$. ◀

Our stopping times will be of the following types:

Proposition 3.9. *Let $(X_t)_{t \geqslant 0}$ be an adapted process with values in a metric space (E, d). For $A \subseteq E$, write*

$$T_A := \inf\{t \geqslant 0 \,;\ X_t \in A\}.$$

(i) *If the sample paths of X are, at each time, either left-continuous or right-continuous and A is open, then T_A is an $(\mathscr{F}_{t^+})$-stopping time.*

(ii) *If the sample paths of X are continuous and A is closed, then T_A is a stopping time.*

Proof. (i) $\forall t > 0$ $[T_A < t] = \bigcup_{s \in [0,t) \cap \mathbb{Q}} [X_s \in A] \in \mathscr{F}_t$, so the result follows immediately from Proposition 3.6(i).

(ii) $\forall t > 0$ $[T_A \leqslant t] = \big[\inf_{s \in [0,t] \cap \mathbb{Q}} d(X_s, A) = 0\big] \in \mathscr{F}_t$. ◀

Exercise (due 10/12). Give an example of an adapted process for each of the following:

(a) X is left-continuous and A is open, but T_A is not a stopping time;

(b) X is right-continuous and A is open, but T_A is not a stopping time;

(c) X is left-continuous and A is closed, but T_A is not a stopping time.

Much more general sets and processes give stopping times under some common restrictions on the filtration. Namely, suppose that $(\mathscr{F}_t)_t$ is right-continuous and complete and that E is a topological space. Let X be E-valued and progressive and $A \subset E$ be Borel. Then both the following are stopping times: $\inf\{t \geqslant 0 \,;\ X_t \in A\}$ and $\inf\{t > 0 \,;\ X_t \in A\}$. This is a consequence of the ***debut theorem*** that if $B \subseteq \Omega \times \mathbb{R}_+$ is progressive, then $\omega \mapsto \inf\{t \geqslant 0 \,;\ (t, \omega) \in B\}$ is a stopping time. The usual proofs of this use analytic sets and capacities; for a more elementary proof, see Richard F. Bass, “The measurability of hitting times,” *Electron. Commun. Probab.* **15** (2010), 99–105, with a correction at *Electron. Commun. Probab.* **16** (2011), 189–191. We will not use these extensions.

3.3. Continuous-Time Martingales and Supermartingales

Fix a filtered probability space $\big(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, \mathbf{P}\big)$. Unless otherwise stated, processes in the remainder of the chapter will be $\mathbb{R}$-valued.

Definition 3.10. *Let $(X_t)_{t\geqslant 0}$ be an adapted process with $X_t \in L^1$ for all t. We say X is a* ***martingale*** *if*

$$0 \leqslant s \leqslant t \quad \Longrightarrow \quad \mathbf{E}[X_t \mid \mathscr{F}_s] = X_s \quad a.s.$$

If “$= X_s$” is replaced by “$\leqslant X_s$” [“$\geqslant X_s$”], we say X is a ***supermartingale [submartingale]****.*

Many examples are from processes, like Brownian motion, that have independent increments, where an $\mathbb{R}^d$-valued process $(Z_t)_{t\geqslant 0}$ has ***independent increments with respect to*** $(\mathscr{F}_t)$ if Z is adapted and $0 \leqslant s \leqslant t \Rightarrow Z_t - Z_s \perp\!\!\!\perp \mathscr{F}_s$. If Z is $\mathbb{R}$-valued and has this property, then the following hold:

(i) if $\forall t \geqslant 0\ \ Z_t \in L^1$, then $\widetilde{Z}_t := Z_t - \mathbf{E}[Z_t]$ is a martingale;

(ii) if $\forall t \geqslant 0\ \ Z_t \in L^2$, then $Y_t := \widetilde{Z}_t^{\,2} - \mathbf{E}[\widetilde{Z}_t^{\,2}]$ is a martingale;

(iii) if $\theta \in \mathbb{R}$ and $\forall t \geqslant 0\ \ \mathbf{E}[\exp(\theta Z_t)] < \infty$, then $X_t := \mathrm{e}^{\theta Z_t}/\mathbf{E}[\mathrm{e}^{\theta Z_t}]$ is a martingale.

Proof. These are easy to prove. For example, for (ii), when $0 \leqslant s \leqslant t$,

$$\begin{aligned}
\mathbf{E}\big[\,\widetilde{Z}_t^{\,2} \bigm| \mathscr{F}_s\big] &= \mathbf{E}\big[(\widetilde{Z}_t - \widetilde{Z}_s + \widetilde{Z}_s)^2 \bigm| \mathscr{F}_s\big] \\
&= \widetilde{Z}_s^{\,2} + 2\widetilde{Z}_s\,\mathbf{E}\big[\widetilde{Z}_t - \widetilde{Z}_s \bigm| \mathscr{F}_s\big] + \mathbf{E}\big[(\widetilde{Z}_t - \widetilde{Z}_s)^2 \bigm| \mathscr{F}_s\big] \\
&= \widetilde{Z}_s^{\,2} + \mathbf{E}\big[(\widetilde{Z}_t - \widetilde{Z}_s)^2\big] \\
&= \widetilde{Z}_s^{\,2} + \mathbf{E}\big[\widetilde{Z}_t^{\,2}\big] - 2\,\mathbf{E}\big[\widetilde{Z}_t\widetilde{Z}_s\big] + \mathbf{E}\big[\widetilde{Z}_s^{\,2}\big] \\
&\qquad\qquad\qquad\qquad = \mathbf{E}\big[\mathbf{E}[\widetilde{Z}_t\widetilde{Z}_s \mid \mathscr{F}_s]\big] = \mathbf{E}\big[\widetilde{Z}_s^{\,2}\big] \\
&= \widetilde{Z}_s^{\,2} + \mathbf{E}\big[\widetilde{Z}_t^{\,2}\big] - \mathbf{E}\big[\widetilde{Z}_s^{\,2}\big].
\end{aligned}$$

For (iii), see the book—it is even shorter. ◂

Exercise. On the probability space of $[0,1]$ with Lebesgue measure, define $X_t := (t+1)\mathbf{1}_{[0,1/(t+1)]}$. Show that $(X_t)_{t\geqslant 0}$ is a martingale (with respect to some filtration).

We will now derive even more martingales from Brownian motion.

Definition 3.11. *A (d-dimensional) Brownian motion that has independent increments with respect to $(\mathscr{F}_t)$ is called a* ***(d-dimensional) $(\mathscr{F}_t)$-Brownian motion****.*

From the above, if B is an $(\mathscr{F}_t)$-Brownian motion, started from a fixed real number, then

$$B_t; \quad B_t^2 - t; \quad \mathrm{e}^{\theta B_t - \frac{\theta^2}{2}t}\ (\theta \in \mathbb{R})$$

are martingales with continuous sample paths. These last are called ***exponential martingales of Brownian motion***.

Here are some more: Suppose $f \in L^2(\mathbb{R}_+, \mathrm{Leb})$. Let G be a Gaussian white noise on $L^2(\mathbb{R}_+)$ and define, as we did earlier,

$$Z_t := \int_0^t f(s)\,\mathrm{d}B_s := G\big(f\mathbf{1}_{[0,t]}\big),$$

where $B_s = G\big(\mathbf{1}_{[0,s]}\big)$. Then Z has independent increments with respect to the canonical filtration of Brownian motion. In fact, first, $Z_t \in \mathscr{F}_t$ because one may approximate $f\mathbf{1}_{[0,t]}$ in L^2 by step functions, and second, when $0 \leqslant s \leqslant t$,

$$Z_t - Z_s = G(f\mathbf{1}_{[s,t]}) \perp G(h\mathbf{1}_{[0,s]})$$

for all $h \in L^2(\mathbb{R}_+)$, whence $Z_t - Z_s \perp\!\!\!\perp \mathscr{F}_s$ by Theorem 1.9. This yields the martingales

$$\int_0^t f(s)\,\mathrm{d}B_s; \quad \Big(\int_0^t f(s)\,\mathrm{d}B_s\Big)^2 - \underbrace{\int_0^t f(s)^2\,\mathrm{d}s}_{\|f\mathbf{1}_{[0,t]}\|^2_{L^2(\mathbb{R}_+)}};$$

since $Z_t \sim \mathcal{N}\big(0, \int_0^t f(s)^2\,\mathrm{d}s\big)$

$$\exp\Big\{\theta\int_0^t f(s)\,\mathrm{d}B_s - \frac{\theta^2}{2}\int_0^t f(s)^2\,\mathrm{d}s\Big\} \quad (\theta \in \mathbb{R}).$$

The first of these is a Wiener integral process, so it has a modification with continuous sample paths by the exercise on page 12. Therefore, so do all the rest. This also follows from Theorem 5.6.

If N is a Poisson process with parameter λ and $(\mathscr{F}_t)$ is its canonical filtration, then we get the martingales

$$N_t - \lambda t; \quad (N_t - \lambda t)^2 - \lambda t; \quad \exp\big\{\theta N_t - \lambda t(\mathrm{e}^\theta - 1)\big\} \quad (\theta \in \mathbb{R}).$$

Of course, these cannot be modified to have continuous sample paths.

Exercise (due 10/12). Let B be an $(\mathscr{F}_t)$-Brownian motion. Write $M_t(\theta) := \mathrm{e}^{\theta B_t - \frac{\theta^2}{2}t}$. Show that for each $\theta \in \mathbb{R}$ and $n \in \mathbb{N}$, $\big((\frac{\mathrm{d}}{\mathrm{d}\theta})^n M_t(\theta)\big)_{t \geqslant 0}$ is an $(\mathscr{F}_t)$-martingale. By using $n = 3$ and $\theta = 0$, deduce that $\big(B_t^3 - 3tB_t\big)_{t\geqslant 0}$ is an $(\mathscr{F}_t)$-martingale.

We now give some properties of (sub)(super)martingales. The first is proved exactly as in the discrete case:

Proposition 3.12. *Let $(X_t)_{t\geqslant 0}$ be adapted and $f\colon \mathbb{R} \to \mathbb{R}_+$ be convex. Suppose that $\forall t \geqslant 0$ $\mathbf{E}[f(X_t)] < \infty$.*

(i) *If X_t is a martingale, then $\big(f(X_t)\big)_{t\geqslant 0}$ is a submartingale.*
(ii) *If (X_t) is a submartingale and f is increasing, then $\big(f(X_t)\big)_{t\geqslant 0}$ is a submartingale.*

Our next result is trivial in the discrete case:

Proposition 3.13. *Let $(X_t)_{t\geqslant 0}$ be a submartingale or supermartingale. Then*

$$\forall t > 0 \qquad \sup_{0\leqslant s\leqslant t} \mathbf{E}\big[|X_s|\big] < \infty.$$

Proof. By symmetry, it is enough to prove this when (X_t) is a submartingale. We use $|X_s| = 2X_s^+ - X_s$. By Proposition 3.12, (X_t^+) is a submartingale, so

$$0 \leqslant s \leqslant t \quad \Longrightarrow \quad \mathbf{E}[X_s^+] \leqslant \mathbf{E}[X_t^+].$$

Also,

$$\mathbf{E}[X_s] \geqslant \mathbf{E}[X_0].$$

Hence,

$$\mathbf{E}\big[|X_s|\big] \leqslant 2\,\mathbf{E}[X_t^+] - \mathbf{E}[X_0].$$ ◀

Our next proofs will use the fact that if (X_t) is a (sub)(super)martingale and $t_1 < t_2 < \cdots < t_p$, then $\big((X_{t_i}, \mathscr{F}_{t_i})\big)_{1\leqslant i\leqslant p}$ is a discrete time (sub)(super)martingale.

The following points towards quadratic variation. We call a process (X_t) ***square-integrable*** if $\forall t \;\; X_t \in L^2$.

Proposition 3.14. *Let $(X_t)_{t\geqslant 0}$ be a square-integrable martingale and $0 \leqslant t_0 < \cdots < t_p$. Then*

$$\mathbf{E}\Big[\sum_{i=1}^{p} (X_{t_i} - X_{t_{i-1}})^2 \;\Big|\; \mathscr{F}_{t_0}\Big] \overset{1}{=} \mathbf{E}\big[X_{t_p}^2 - X_{t_0}^2 \;\big|\; \mathscr{F}_{t_0}\big] \overset{2}{=} \mathbf{E}\big[(X_{t_p} - X_{t_0})^2 \;\big|\; \mathscr{F}_{t_0}\big].$$

Hence, the same holds unconditionally.

Proof. This is a type of Pythagorean theorem and depends on orthogonality. We have $\forall i \in [1, p]$

$$\begin{aligned}
\mathbf{E}[(X_{t_i} - X_{t_{i-1}})^2 \mid \mathscr{F}_{t_0}] &= \mathbf{E}\Big[\mathbf{E}\big[(X_{t_i} - X_{t_{i-1}})^2 \;\big|\; \mathscr{F}_{t_{i-1}}\big] \;\Big|\; \mathscr{F}_{t_0}\Big] \\
&= \mathbf{E}\Big[\mathbf{E}\big[X_{t_i}^2 \mid \mathscr{F}_{t_{i-1}}\big] - 2X_{t_{i-1}}\,\mathbf{E}\big[X_{t_i} \mid \mathscr{F}_{t_{i-1}}\big] + X_{t_{i-1}}^2 \;\Big|\; \mathscr{F}_{t_0}\Big] \\
&= \mathbf{E}\Big[\mathbf{E}\big[X_{t_i}^2 \mid \mathscr{F}_{t_{i-1}}\big] - X_{t_{i-1}}^2 \;\Big|\; \mathscr{F}_{t_0}\Big] \\
&= \mathbf{E}\big[X_{t_i}^2 - X_{t_{i-1}}^2 \;\big|\; \mathscr{F}_{t_0}\big].
\end{aligned}$$

Now, sum on i to get 1. If we take $p = 1$, we get 2. ◀

Next, we have analogues of discrete-time inequalities. Note that if $f\colon [0,t] \to \mathbb{R}$ is right-continuous, then $\sup_{0\leqslant s\leqslant t} f(s) = \sup_{s\in[0,t]\cap(\mathbb{Q}\cup\{t\})} f(s)$, whence we obtain measurability of the supremum for a stochastic process with right-continuous sample paths.

Proposition 3.15. *Let $(X_t)_{t\geqslant 0}$ be a submartingale or supermartingale with right-continuous sample paths.*

(i) *(Maximal inequality)*

$$\forall t > 0 \;\; \forall \lambda > 0 \qquad \lambda\,\mathbf{P}\Big[\sup_{0\leqslant s\leqslant t} |X_s| \geqslant \lambda\Big] \leqslant \mathbf{E}\big[|X_0|\big] + 2\,\mathbf{E}\big[|X_t|\big].$$

(ii) *(Doob's L^p inequality) If X is a martingale, then*

$$\forall t > 0 \;\; \forall p > 0 \qquad \mathbf{E}\Big[\sup_{0\leqslant s\leqslant t} |X_s|^p\Big] \leqslant \Big(\frac{p}{p-1}\Big)^p \mathbf{E}\big[|X_t|^p\big].$$

Proof. (i) If D is a finite set in $[0,t]$ with $0,t \in D$, then the discrete-time inequality yields

$$\lambda\, \mathbf{P}\Big[\sup_{s\in D} |X_s| > \lambda\Big] \leqslant \mathbf{E}\big[|X_0|\big] + 2\,\mathbf{E}\big[|X_t|\big].$$

Now take D to be countable and dense in $[0,t]$ with $0,t \in D$ and write D as an increasing union of finite sets D_m with $0,t \in D_m$. Then take the limit in the above inequality for D_m as m goes to ∞ to get

$$\lambda\, \mathbf{P}\Big[\sup_{0\leqslant s\leqslant t} |X_s| > \lambda\Big] \leqslant \mathbf{E}\big[|X_0|\big] + 2\,\mathbf{E}\big[|X_t|\big].$$

Finally, use this inequality for a sequence λ_n increasing to λ.

(ii) The proof is similar; now we invoke the monotone convergence theorem. ◂

Remark. If we did not assume right-continuity, we would get the same results for $\sup_{s\in D} |X_s|$ for any countable $D \subseteq [0,t]$: we may add $\{0,t\}$ to D if necessary. In particular, by letting $\lambda \to \infty$, we get from (i) that $\sup_{s\in D} |X_s| < \infty$ almost surely.

We call a function ***càdlàg*** or ***RCLL*** if it is right-continuous with left-limits everywhere.

Exercise (due 10/12). Let $(X_t)_{t\geqslant 0}$ be a process with càdlàg sample paths. Let A_t be the event that $s \mapsto X_s$ is continuous for $s \in [0,t]$. Show that $\forall t \geqslant 0\;\; A_t \in \mathscr{F}_t^X$.

As for discrete times, we prove convergence using upcrossings, where for a function $f\colon I \to \mathbb{R}$ ($I \subseteq \mathbb{R}$) and $a < b$, the upcrossing number of f along $[a,b]$ is

$$M_{ab}^f(I) := \sup\left\{k \geqslant 0\,;\; \begin{array}{l} \exists s_i, t_i \in I \text{ with } s_1 < t_1 < \cdots < s_k < t_k \\ \text{and } f(s_i) \leqslant a,\ f(t_i) \geqslant b \text{ for } 1 \leqslant i \leqslant k \end{array}\right\}.$$

In Section 3.4, we use this, as in discrete time, to study convergence as $t \to \infty$. Here, we study right and left limits at finite times. We will use the following easy lemma:

Lemma 3.16. *Let $D \subseteq \mathbb{R}_+$ be dense and countable. Let $f\colon D \to \mathbb{R}$ satisfy*

$$\forall u \in D \qquad \sup_{t\in D\cap[0,u]} \big|f(t)\big| < \infty$$

and

$$\forall a, b \in \mathbb{Q} \text{ with } a < b \qquad M_{ab}^f\big(D \cap [0,u]\big) < \infty.$$

Then $\forall t \geqslant 0\;\; f(t^+) := \lim_{s\downarrow t\in D} f(t)$ exists and $f(t^-) := \lim_{s\uparrow t \in D} f(t)$ exists. In addition, $t \mapsto f(t^+)$ is càdlàg on $\mathbb{R}_+$. ◂

(Note that $t \mapsto f(t^+)$ has left limits by the upcrossing condition.)

In the proof of the next theorem, we will use the fact that a backward (sub)(super)martingale $\big((Y_n, \mathscr{F}_n)\big)_{n\leqslant 0}$ with $\sup_{n\leqslant 0} \mathbf{E}\big[|Y_n|\big] < \infty$ is uniformly integrable. (If you haven't seen this, here's a hint to prove it. Note that it does not matter whether the process is a submartingale or a supermartingale.

Assume the former. Show that there is a backward martingale $(Z_n)_{n\leqslant 0}$ and nonnegative random variables $W_n \in \mathscr{F}_{n-1}$ such that $Y_n = Z_n + W_n$ and $W_n \geqslant W_{n-1}$. Use that every backward martingale is uniformly integrable.)

Recall that if a sequence converges almost surely, then it converges in L^1 if and only if it is uniformly integrable. The same follows for continuous time.

Theorem 3.17. *Let $(X_t)_{t\geqslant 0}$ be a supermartingale and D be a countable dense subset of $\mathbb{R}_+$.*

(i) $\exists N \subseteq \Omega$ *such that* $\mathbf{P}(N) = 0$ *and* $\forall \omega \notin N$

$$\forall t \geqslant 0 \qquad X_{t^+}(\omega) := \lim_{D\ni s\downarrow t} X_s(\omega) \quad \textit{and} \quad X_{t^-}(\omega) := \lim_{D\ni s\uparrow t} X_s(\omega)$$

exist.

(ii) $\forall t \in \mathbb{R}_+ \;\; X_{t^+} \in L^1$ *and satisfies*

$$X_t \geqslant \mathbf{E}[X_{t^+} \mid \mathscr{F}_t]$$

with equality if $t \mapsto \mathbf{E}[X_t]$ is right-continuous (e.g., if X is a martingale). The process $(X_{t^+})_{t\geqslant 0}$ is indistinguishable from a process that is an $(\mathscr{F}_{t^+})$-supermartingale and, if X is a martingale, an $(\mathscr{F}_{t^+})$-martingale.

Proof. (i) Fix $u \in D$. We saw in the remark that

$$\sup_{s\in D\cap[0,u]} |X_s| < \infty \quad \text{a.s.}$$

For finite $D' \subseteq D \cap [0,u]$, Doob's upcrossing inequality yields

$$\forall a < b \qquad \mathbf{E}\big[M_{ab}^X(D')\big] \leqslant \frac{\mathbf{E}\big[(X_u - a)^-\big]}{b-a}.$$

Thus, using an increasing sequence of finite subsets of $D \cap [0,u]$ whose union is $D \cap [0,u]$ and using the monotone convergence theorem, we get

$$\mathbf{E}\Big[M_{ab}^X(D \cap [0,u])\Big] \leqslant \frac{\mathbf{E}\big[(X_u - a)^-\big]}{b-a} < \infty,$$

whence $M_{ab}^X\big(D \cap [0,u]\big) < \infty$ almost surely. Let

$$N := \Big[\exists u \in D, \sup_{s\in D\cap[0,u]} |X_s| = \infty \text{ or } \exists a,b \in \mathbb{Q} \text{ with } a < b \text{ and } M_{ab}^X\big(D\cap[0,u]\big) = \infty\Big].$$

We have seen that $\mathbf{P}(N) = 0$ as a countable union of sets of probability 0. For $\omega \notin N$, we may apply Lemma 3.16 to get (i).

(ii) Fix $t \in \mathbb{R}_+$. Choose $D \ni t_n \downarrow t$ monotonically, so $X_{t^+} := \lim_{n\to\infty} X_{t_n}$ almost surely. If we re-index time, $Y_n := X_{t_{-n}}$ $(n \leqslant 0)$, then $(Y_n)_{n\leqslant 0}$ is a backward supermartingale. By Proposition 3.13,

$$\sup_{n\leqslant 0} \mathbf{E}\big[|Y_n|\big] < \infty,$$

whence $X_{t_n} \to X_{t^+}$ in L^1 by uniform integrability. In particular, $X_{t^+} \in L^1$.

Since L^1 convergence implies L^1 convergence of conditional expectations (because of the inequality $\mathbf{E}\Big[\big|\mathbf{E}[Z_1 \mid \mathscr{G}] - \mathbf{E}[Z_2 \mid \mathscr{G}]\big|\Big] \leqslant \mathbf{E}\Big[\mathbf{E}\big[|Z_1 - Z_2| \bigm| \mathscr{G}\big]\Big] = \mathbf{E}\big[|Z_1 - Z_2|\big]$), we get

$$X_t \geqslant \mathbf{E}[X_{t_n} \mid \mathscr{F}_t] \xrightarrow{L^1} \mathbf{E}[X_{t^+} \mid \mathscr{F}_t].$$

If $s \mapsto \mathbf{E}[X_s]$ is right-continuous, then the expectation of the right-hand side equals $\mathbf{E}[X_t]$, which is the expectation of the left-hand side, whence the two sides agree almost surely.

Now redefine X_{t^+} to be $\lim_{D \ni s \downarrow t} X_s$ when the limit exists and 0 elsewhere. This changes X_{t^+} on a subset of N, whence it is indistinguishable from its definition in (i). Furthermore, $X_{t^+} \in \mathscr{F}_{t^+}$ now. Consider $s < t$ and choose $s_n \downarrow s$ and $t_n \downarrow t$ with $s_n, t_n \in D$ and $s_n \leqslant t_n$. To show that $X_{s^+} \geqslant \mathbf{E}[X_{t^+} \mid \mathscr{F}_{s^+}]$, it suffices to show that

$$\forall A \in \mathscr{F}_{s^+} \qquad \mathbf{E}[X_{s^+}\mathbf{1}_A] \geqslant \mathbf{E}\big[\mathbf{E}[X_{t^+} \mid \mathscr{F}_{s^+}]\mathbf{1}_A\big] = \mathbf{E}[X_{t^+}\mathbf{1}_A]$$

(by considering $A := \big[X_{s^+} < \mathbf{E}[X_{t^+} \mid \mathscr{F}_{s^+}]\big]$). Indeed, L^1 convergence yields

$$\mathbf{E}[X_{s^+}\mathbf{1}_A] = \lim_{n\to\infty} \mathbf{E}[X_{s_n}\mathbf{1}_A] \geqslant \lim_{n\to\infty} \mathbf{E}[X_{t_n}\mathbf{1}_A] = \mathbf{E}[X_{t^+}\mathbf{1}_A].$$

$$\hookrightarrow [X_{s_n} \geqslant \mathbf{E}[X_{t_n} \mid \mathscr{F}_{s_n}] \text{ and } A \in \mathscr{F}_{s^+} \subseteq \mathscr{F}_{s_n}]$$

Thus, (X_{t^+}) is an $(\mathscr{F}_{t^+})$-supermartingale. If (X_t) had been a submartingale, then we would have concluded (X_{t^+}) is an $(\mathscr{F}_{t^+})$-submartingale, whence (X_t) is a martingale implies (X_{t^+}) is an $(\mathscr{F}_{t^+})$-martingale. ◂

Exercise (due 10/19). Let $(X_t)_{t\geqslant 0}$ be a supermartingale with càdlàg sample paths. Show that $t \mapsto \mathbf{E}[X_t]$ is càdlàg.

There is a kind of converse to this exercise:

Theorem 3.18. *Let $(\mathscr{F}_t)$ be right-continuous and complete (often called "the usual conditions"). If (X_t) is a supermartingale such that $t \mapsto \mathbf{E}[X_t]$ is right-continuous, then (X_t) has a modification that is a supermartingale with càdlàg sample paths.*

Proof. Consider the modification of (X_{t^+}) that we used in the proof of Theorem 3.17(ii). We saw there that $X_{t^+} \in \mathscr{F}_{t^+}$, which now equals $\mathscr{F}_t$. The theorem showed now that $X_t = \mathbf{E}[X_{t^+} \mid \mathscr{F}_t] = X_{t^+}$ almost surely. Thus, use (X_{t^+}) as the modification of (X_t). Lemma 3.16 shows that X_{t^+} is càdlàg. ◂

3.4. Optional Stopping Theorems

Our first two results really belong in Section 3.3. They are about convergence as time $t \to \infty$.

Theorem 3.19. *Let X be a right-continuous submartingale or supermartingale bounded in L^1. Then there exists $X_\infty \in L^1$ such that $\lim_{t\to\infty} X_t = X_\infty$ almost surely.*

Proof. We may assume X is a supermartingale. Let D be a countable dense subset of $\mathbb{R}_+$. In the proof of Theorem 3.17, we saw that

$$\forall s \in \mathbb{R}_+ \;\; \forall a < b \qquad \mathbf{E}\Big[M^X_{ab}\big(D \cap [0,s]\big)\Big] \leqslant \frac{\mathbf{E}\big[(X_s - a)^-\big]}{b-a}.$$

Taking the supremum over s and using the monotone convergence theorem yields

$$\mathbf{E}\big[M_{ab}^X(D)\big] \leqslant \frac{1}{b-a} \sup_{s \geqslant 0} \mathbf{E}\big[(X_s - a)^-\big] \underset{\text{[by hypothesis]}}{<} \infty.$$

Apply this to $a, b \in \mathbb{Q}$ to get $M_{ab}^X(D) < \infty$ a.s. simultaneously in $a, b \in \mathbb{Q}$, whence $X_\infty := \lim_{D \ni t \to \infty} X_t$ exists in $[-\infty, \infty]$ almost surely. By Fatou's lemma, we have

$$\mathbf{E}\big[|X_\infty|\big] \leqslant \varliminf_{D \ni t \to \infty} \mathbf{E}\big[|X_t|\big] < \infty,$$

whence $X_\infty \in L^1$. Finally, right-continuity shows the conclusion. ◀

Whether convergence holds in L^1 is just as in the case of deterministic time:

Definition 3.20. *A martingale X is **closed** if there exists $Z \in L^1$ such that*

$$\forall t \geqslant 0 \qquad X_t = \mathbf{E}[Z \mid \mathscr{F}_t] \quad \textit{a.s.}$$

Theorem 3.21. *Let X be a right-continuous martingale. The following are equivalent:*

(i) *X is closed;*

(ii) *X is uniformly integrable;*

(iii) *X_t converges almost surely and in L^1 as $t \to \infty$.*

In this case

$$\forall t \geqslant 0 \qquad X_t = \mathbf{E}[X_\infty \mid \mathscr{F}_t] \quad \textit{a.s.},$$

where $X_\infty := \lim_{t\to\infty} X_t$. ◀

Since we are now interested in X_∞, we will define X_T even where $T = \infty$: If $\lim_{t\to\infty} X_t = X_\infty$ almost surely and T is a $[0, \infty]$-valued random variable, then we write

$$X_T(\omega) := X_{T(\omega)}(\omega),$$

defined almost surely. We saw in Theorem 3.7 that if X is progressive and T is a stopping time, then X_T is $\mathscr{F}_T$ measurable on $[T < \infty]$. If X is adapted, then $X_\infty \in \mathscr{F}_\infty$, whence X_T is $\mathscr{F}_T$-measurable on $[T = \infty]$. Therefore, if X is a right-continuous submartingale or supermartingale and T is a stopping time, then $X_T \in \mathscr{F}_T$ (by Proposition 3.4).

One of the main reasons martingales are useful is:

Theorem 3.22 (Optional stopping theorem for martingales)**.** *Let X be a uniformly integrable, right-continuous martingale. Let $S \leqslant T$ be stopping times. Then $X_S, X_T \in L^1$ and*

$$\begin{aligned} X_S &= \mathbf{E}[X_T \mid \mathscr{F}_S], \qquad (*) \\ X_T &= \mathbf{E}[X_\infty \mid \mathscr{F}_T], \\ \mathbf{E}[X_0] &= \mathbf{E}[X_T] = \mathbf{E}[X_\infty]. \end{aligned}$$

Remark. This extends to uniformly integrable, right-continuous supermartingales with "$\geqslant$" in the conclusions; see *Stochastic Calculus and Applications*, second edition, by Samuel N. Cohen and Robert J. Elliott, Theorem 5.3.1.

Note: In the case that S and T are constants, the first equation is the definition of martingale and the rest are from Theorem 3.21.

Proof. We use the approximations of S and T from Proposition 3.8, now defined where $[S = \infty]$ or $[T = \infty]$:

$$S_n := \lceil 2^n S \rceil / 2^n \text{ and } T_n := \lceil 2^n T \rceil / 2^n.$$

These are stopping times that decrease to S and T and also satisfy $S_n \leqslant T_n$. Thus, we may apply the discrete-time version of this theorem to get

$$\forall n \qquad X_{S_n} = \mathbf{E}[X_{T_n} \mid \mathscr{F}_{S_n}].$$

We want to let $n \to \infty$ to get Eq. ($*$), i.e., that

$$\forall A \in \mathscr{F}_S \qquad \mathbf{E}[\mathbf{1}_A X_S] = \mathbf{E}[\mathbf{1}_A X_T]$$

(because indeed $X_S \in \mathscr{F}_S$). Now, right-continuity yields $X_{S_n} \to X_S$ and $X_{T_n} \to X_T$ almost surely and in L^1, the latter since

$$X_{S_n} = \mathbf{E}[X_\infty \mid \mathscr{F}_{S_n}] \text{ and } X_{T_n} = \mathbf{E}[X_\infty \mid \mathscr{F}_{T_n}]$$

by the discrete-time theorem. In particular, $X_S, X_T \in L^1$. Thus, for each $A \in \mathscr{F}_S \subseteq \mathscr{F}_{S_n}$, we have

$$\begin{array}{ccc} \mathbf{E}[\mathbf{1}_A X_{S_n}] & = & \mathbf{E}[\mathbf{1}_A X_{T_n}] \\ \downarrow & & \downarrow \\ \mathbf{E}[\mathbf{1}_A X_S] & & \mathbf{E}[\mathbf{1}_A X_T]. \end{array}$$

This shows Eq. ($*$), and the rest is immediate from the fact that $T \leqslant \infty$ is a stopping time. ◀

Uniform integrability is a key assumption (e.g., the double-or-nothing martingale). To make it easier to use, we have the following two corollaries.

Corollary 3.23. *Let X be a right-continuous martingale and $S \leqslant T$ be* bounded *stopping times. Then $X_S, X_T \in L^1$ and $X_S = \mathbf{E}[X_T \mid \mathscr{F}_S]$.*

Proof. Suppose $T \leqslant r$ almost surely. Note that $(X_{t\wedge r})_{t\geqslant 0}$ is a martingale, closed by X_r. Thus, it is uniformly integrable. Since $S \wedge r = S$ and $T \wedge r = T$, the result follows from applying Theorem 3.22 to $(X_{t\wedge r})_{t\geqslant 0}$. ◀

A trivial fact is that if $Z \in L^1$, then

$$\forall s, t \geqslant 0 \qquad Z \in \mathscr{F}_s \quad \Longrightarrow \quad \mathbf{E}[Z \mid \mathscr{F}_t] = \mathbf{E}[Z \mid \mathscr{F}_{t\wedge s}].$$

We now replace s by a stopping time, T:

Proposition. *If T is a stopping time, then $\forall Z \in L^1$*

$$\forall t \geqslant 0 \qquad Z \in \mathscr{F}_T \quad \Longrightarrow \quad \mathbf{E}[Z \mid \mathscr{F}_t] = \mathbf{E}[Z \mid \mathscr{F}_{t\wedge T}].$$

Proof. Let $Y := \mathbf{E}[Z \mid \mathscr{F}_{t\wedge T}]$. Since $Y \in \mathscr{F}_{t\wedge T} \subseteq \mathscr{F}_t$, it suffices to show that

$$\forall A \in \mathscr{F}_t \qquad \mathbf{E}[\mathbf{1}_A Z] = \mathbf{E}[\mathbf{1}_A Y].$$

Consider $A = \big(A \cap [T \leqslant t]\big) \cup \big(A \cap [T > t]\big)$. Now, $Z\mathbf{1}_{[T\leqslant t]} \in \mathscr{F}_t$ (property (j) of stopping times) and since $Z \in \mathscr{F}_T$ and $T \in \mathscr{F}_T$, also $Z\mathbf{1}_{[T\leqslant t]} \in \mathscr{F}_T$, whence $Z\mathbf{1}_{[T\leqslant t]} \in \mathscr{F}_t \cap \mathscr{F}_T = \mathscr{F}_{t\wedge T}$ (property (f)). Therefore,

$$Z\mathbf{1}_{[T\leqslant t]} = \mathbf{E}[Z\mathbf{1}_{[T\leqslant t]} \mid \mathscr{F}_{t\wedge T}] = \mathbf{E}[Z \mid \mathscr{F}_{t\wedge T}]\underbrace{\mathbf{1}_{[T\leqslant t]}}_{\in \mathscr{F}_{t\wedge T}} = Y\mathbf{1}_{[T\leqslant t]},$$

so

$$\mathbf{E}[Z\mathbf{1}_A\mathbf{1}_{[T\leqslant t]}] = \mathbf{E}[Y\mathbf{1}_A\mathbf{1}_{[T\leqslant t]}].$$

Also, $A \cap [T > t] \in \mathscr{F}_t$ and since we have that $\forall s \geqslant 0\ \ A \cap [T > t] \cap [T \leqslant s] \in \mathscr{F}_s$, also $A \cap [T > t] \in \mathscr{F}_T$, so again $A \cap [T > t] \in \mathscr{F}_{t\wedge T}$, whence by definition of Y,

$$\mathbf{E}\big[Z\mathbf{1}_{A\cap[T>t]}\big] = \mathbf{E}\big[Y\mathbf{1}_{A\cap[T>t]}\big].$$

Adding these last two displays gives the result. ◀

We apply this to stopping a process, i.e., if X is a process, the stopped process $(X_{t\wedge T})_{t\geqslant 0}$.

Corollary 3.24. *Let X be a right-continuous martingale and T be a stopping time.*

(i) *The process $(X_{t\wedge T})_{t\geqslant 0}$ is a martingale.*
(ii) *If X is uniformly integrable, then so is $(X_{t\wedge T})_{t\geqslant 0}$, which is closed by X_T:*

$$X_{t\wedge T} = \mathbf{E}[X_T \mid \mathscr{F}_t] \quad a.s. \tag{$*$}$$

Proof. (ii) By Theorem 3.22, $X_T \in L^1$, so we may apply the proposition to $Z := X_T$ to obtain

$$\mathbf{E}[X_T \mid \mathscr{F}_t] = \mathbf{E}[X_T \mid \mathscr{F}_{t\wedge T}].$$

Also, $t \wedge T$ is a stopping time (property (f)), so Theorem 3.22 gives $\mathbf{E}[X_T \mid \mathscr{F}_{t\wedge T}] = X_{t\wedge T}$. This gives Eq. ($*$)—which, by the way, also implies uniform integrability.

(i) Recall that $\forall s \geqslant 0\ \ (X_{t\wedge s})_{t\geqslant 0}$ is a uniformly integrable martingale. Applying Eq. ($*$) to this process, we get $\forall t \leqslant s$

$$\underbrace{X_{(t\wedge s)\wedge T}}_{=X_{t\wedge T}} = \mathbf{E}[X_{T\wedge s} \mid \mathscr{F}_t]. \qquad ◀$$

Exercise (due 10/26). Deduce the proposition from Corollary 3.24 for $(\mathscr{F}_t)$ that is right-continuous and complete.

Exercise. There is a kind of converse to Theorem 3.22. Suppose that X_t is defined for all $t \in [0, \infty]$, including $t = \infty$. Show that if X is progressive and for every finite or infinite stopping time, T, X_T is integrable with mean 0, then X is a uniformly integrable martingale. *Hint:* consider $A \in \mathscr{F}_t$ and define $T := t\mathbf{1}_A + \infty\mathbf{1}_{A^c}$ to deduce that $X_t = \mathbf{E}[X_\infty \mid \mathscr{F}_t]$.

Applications

Let B be a real Brownian motion. For $a \in \mathbb{R}$, let $T_a := \inf\{t \geqslant 0\,;\ B_t = a\}$.

(a) For $a, b > 0$, consider $T := T_{-a} \wedge T_b$ and the stopped martingale $M_t := B_{t \wedge T}$. Because $|M_t| \leqslant a \vee b$, M is uniformly integrable, whence

$$0 = \mathbf{E}[M_0] = \mathbf{E}[M_T] = \mathbf{E}[B_T] = b\,\mathbf{P}[T_b < T_{-a}] - a\,\mathbf{P}[T_{-a} < T_b].$$

Since the two probabilities add to 1, we can solve to find

$$\mathbf{P}[T_b < T_{-a}] = \frac{a}{a+b}.$$

We needed only that Brownian motion is a continuous martingale from 0 that leaves $(-a, b)$.

(b) For $a, b > 0$ and $T := T_{-a} \wedge T_b$, consider the martingale $M_t := B_t^2 - t$. Again, $M_{t \wedge T}$ is a martingale, though no longer bounded. Still, we have from the martingale property that

$$\forall t \geqslant 0 \qquad 0 = \mathbf{E}[M_0] = \mathbf{E}[M_{t \wedge T}],$$

i.e.,

$$\mathbf{E}\big[B_{t \wedge T}^2\big] = \mathbf{E}[t \wedge T].$$

We may let $t \to \infty$ and use the bounded convergence theorem on the left-hand side and monotone convergence theorem on the right-hand side to obtain

$$\mathbf{E}[B_T^2] = \mathbf{E}[T].$$

Using (a), we find that $\mathbf{E}[T] = ab$.

(c) For $a > 0$, and $\theta > 0$, consider the martingale

$$N_t^\theta := \mathrm{e}^{\theta B_t - \theta^2 t/2}.$$

The stopped process, $N_{t \wedge T_a}^\theta$, takes values in $(0, \mathrm{e}^{\theta a})$, so is uniformly integrable, whence

$$1 = \mathbf{E}[N_0^\theta] = \mathbf{E}[N_{T_a}^\theta] = \mathrm{e}^{\theta a}\,\mathbf{E}[\mathrm{e}^{-\theta^2 T_a/2}]. \tag{$*$}$$

Taking $\theta := \sqrt{2\lambda}$ gives the Laplace transform of T_a:

$$\mathbf{E}[\mathrm{e}^{-\lambda T_a}] = \mathrm{e}^{-a\sqrt{2\lambda}} \quad (\lambda > 0). \tag{3.7}$$

Note that if we used $\theta = -\sqrt{2\lambda}$ in Eq. ($*$), we would get a different result. The reason is that when $\theta < 0$, N^θ is not uniformly integrable.

Exercise (due 10/26). Exercise 3.26.

Exercise (due 10/26). For $x \in \mathbb{R}^d$ and $R > |x|$, let $T_{d,R} := \inf\{t \geqslant 0\,;\ |B_t| = R\}$, where $(B_t)_{t \geqslant 0}$ is a d-dimensional Brownian motion started from x. Show that

$$\mathbf{E}[T_{d,R}] = \frac{R^2 - |x|^2}{d}.$$

Extra credit: compute the Laplace transform of $T_{d,R}$.

Exercise. Let B be 1-dimensional Brownian motion and $T := T_{-a} \wedge T_b$ for $a, b > 0$. Use an earlier martingale to compute $\mathbf{E}[TB_T]$.

We'll later need the continuous-time analogue of the discrete-time optional stopping theorem for nonnegative supermartingales.

Theorem 3.25. *Let X be a nonnegative right-continuous supermartingale and $S \leqslant T$ be stopping times. Then $X_S, X_T \in L^1$ and*

$$X_S \geqslant \mathbf{E}[X_T \mid \mathscr{F}_S].$$

Also,

$$\mathbf{E}[X_S] \geqslant \mathbf{E}[X_T]$$

and

$$\mathbf{E}\big[X_S \mathbf{1}_{[S<\infty]}\big] \geqslant \mathbf{E}\big[X_T \mathbf{1}_{[S<\infty]}\big] \geqslant \mathbf{E}\big[X_T \mathbf{1}_{[T<\infty]}\big].$$

Proof. The strategy of proof for the martingale case (Theorem 3.22) mostly works, but now we need some extra arguments. We are not assuming uniform integrability, and even for nonnegative martingales, equality need not hold in the conclusion.

We first claim that if T is bounded, then $\mathbf{E}[X_S] \geqslant \mathbf{E}[X_T]$. Let $S_n := \lceil 2^n S \rceil / 2^n$ and $T_n := \lceil 2^n T \rceil / 2^n$. Right-continuity ensures that $X_{S_n} \to X_S$ and $X_{T_n} \to X_T$ as $n \to \infty$. The optional stopping theorem in discrete time for bounded stopping times gives

$$\forall n \geqslant 0 \qquad X_{S_{n+1}} \geqslant \mathbf{E}[X_{S_n} \mid \mathscr{F}_{S_{n+1}}]$$

(note $S_{n+1} \leqslant S_n$). This means that $\big((X_{S_{-n}}, \mathscr{F}_{S_{-n}})\big)_{n \leqslant 0}$ is a backward supermartingale. The optional stopping theorem also yields $\mathbf{E}[X_{S_n}] \leqslant \mathbf{E}[X_0]$, so this backward supermartingale is L^1-bounded, whence converges in L^1 to X_S. Likewise, $X_{T_n} \to X_T$ in L^1.

Since $S_n \leqslant T_n$, the optional stopping theorem also implies that

$$\mathbf{E}[X_{S_n}] \geqslant \mathbf{E}[X_{T_n}].$$

Taking $n \to \infty$ gives the claim.

Now, we prove the theorem. For any $m > 0$, we may apply the first part to the bounded stopping times $0 \leqslant S \wedge m$ to get $\mathbf{E}[X_{S \wedge m}] \leqslant \mathbf{E}[X_0]$. Fatou's lemma then yields $\mathbf{E}[X_S] \leqslant \mathbf{E}[X_0] < \infty$ and similarly $X_T \in L^1$.

Property (d) of stopping times says that for $A \in \mathscr{F}_S$,

$$S^A(\omega) := \begin{cases} S(\omega) & \text{for } \omega \in A, \\ \infty & \text{for } \omega \notin A \end{cases}$$

is a stopping time. Likewise, T^A is a stopping time because $\mathscr{F}_S \subseteq \mathscr{F}_T$. Applying the first part of the proof to the bounded stopping times $S^A \wedge m \leqslant T^A \wedge m$ gives

$$\forall m > 0 \qquad \mathbf{E}[X_{S^A \wedge m}] \geqslant \mathbf{E}[X_{T^A \wedge m}].$$

Now if $S > m$, then $T > m$, so

$$X_{S^A \wedge m} \mathbf{1}_{[S>m]} = X_{T^A \wedge m} \mathbf{1}_{[S>m]},$$

whence

$$\mathbf{E}\big[X_S \mathbf{1}_{A\cap[S\leqslant m]}\big] \geqslant \mathbf{E}\big[X_{T\wedge m}\mathbf{1}_{A\cap[S\leqslant m]}\big].$$

Apply the monotone convergence theorem to the left-hand side and Fatou's lemma to the right-hand side to obtain

$$\mathbf{E}\big[X_S \mathbf{1}_{A\cap[S<\infty]}\big] \geqslant \mathbf{E}\big[X_T \mathbf{1}_{A\cap[S<\infty]}\big].$$

Since $X_S \mathbf{1}_{A\cap[S=\infty]} = X_T \mathbf{1}_{A\cap[S=\infty]}$, we get

$$\mathbf{E}[X_S \mathbf{1}_A] \geqslant \mathbf{E}[X_T \mathbf{1}_A] = \mathbf{E}\big[\mathbf{E}[X_T \mid \mathscr{F}_S]\mathbf{1}_A\big].$$

Therefore, $X_S \geqslant \mathbf{E}[X_T \mid \mathscr{F}_S]$. ◀

Exercise. Prove that if X is a nonnegative right-continuous supermartingale and $\lambda > 0$, then

$$\lambda\, \mathbf{P}\Big[\sup_{t\geqslant 0} X_t \geqslant \lambda\Big] \leqslant \mathbf{E}[X_0].$$

Chapter 4

Continuous Semimartingales

A semimartingale is by definition the sum of a local martingale and a finite-variation process, both of which we define and study here. The next chapter studies stochastic integration with respect to continuous semi-martingales. A key process studied here will be the quadratic variation process of a continuous local martingale.

4.1. Finite-Variation Processes

4.1.1. Functions with Finite Variation

This is a review from real analysis. Let $I \subseteq \mathbb{R}$ be an interval. We say $a\colon I \to \mathbb{R}$ has ***finite*** (or ***bounded***) ***variation*** if

$$\sup\left\{\sum_{i=1}^{p} \bigl|a(t_i) - a(t_{i-1})\bigr|\,;\ t_0 < t_1 < \cdots < t_p \in I\right\} < \infty. \tag{$*$}$$

This is equivalent to the property that there exists a signed Borel measure μ on I such that

$$\forall s < t \in I \quad a(t) - a(s) = \mu\bigl((s, t]\bigr).$$

Such a μ is uniquely determined by a. A signed measure μ has a Hahn–Jordan decomposition $\mu = \mu^+ - \mu^-$, where $\mu^+, \mu^- \geqslant 0$ and $\mu^+ \perp \mu^-$. We write $|\mu| := \mu^+ + \mu^-$. A function of finite variation is thus the difference of two bounded increasing functions and conversely.

For $f \in L^1(\mu)$, we write

$$\int_I f\,\mathrm{d}a := \int_I f\,\mathrm{d}\mu \quad \text{and} \quad \int_I f\,|\mathrm{d}a| := \int_I f\,|\mathrm{d}\mu|.$$

We have

$$\left|\int_I f\,\mathrm{d}a\right| \leqslant \int_I |f|\,|\mathrm{d}a|.$$

Furthermore, the function $t \mapsto \int_{I\cap(-\infty,t]} f\,|\mathrm{d}a|$ on I has finite variation, represented by the measure $f \cdot \mu$.

Proposition 4.2. *If a has finite variation on $[s,t]$, then*

$$\int_s^t |\mathrm{d}a| = \text{supremum in Eq. } (*)$$

$$= \lim_{\varepsilon\downarrow 0}\left\{\sum_{i=1}^{p}\left|a(t_i)-a(t_{i-1})\right|;\ s=t_0<t_1<\cdots<t_p=t,\ \max_i|t_i-t_{i-1}|<\varepsilon\right\}. \qquad \blacktriangleleft$$

We also have

Lemma 4.3. *If a has finite variation on $[s,t]$ and $f\colon [s,t]\to\mathbb{R}$ is continuous, then*

$$\int_s^t f\,\mathrm{d}a = \lim_{\varepsilon\downarrow 0}\left\{\sum_{i=1}^{p} f(t_{i-1})\left[a(t_i)-a(t_{i-1})\right];\quad \begin{array}{c} s=t_0<t_1<\cdots<t_p=t, \\ \max_i|t_{i+1}-t_1|<\varepsilon \end{array}\right\}. \qquad \blacktriangleleft$$

Corollary. *If a has finite variation on $[s,t]$ and $f\colon [s,t]\to\mathbb{R}$ is continuous, then*

$$\int_s^t f\,|\mathrm{d}a| = \lim_{\varepsilon\downarrow 0}\left\{\sum_{i=1}^{p} f(t_{i-1})\left|a(t_i)-a(t_{i-1})\right|;\quad \begin{array}{c} s=t_0<t_1<\cdots<t_p=t, \\ \max_i|t_{i+1}-t_1|<\varepsilon \end{array}\right\}.$$

We will not use this, so we skip the proof.

We call a function $a\colon\mathbb{R}_+\to\mathbb{R}$ a ***finite-variation function*** if $a\restriction I$ has finite variation for all bounded $I\subseteq\mathbb{R}_+$. In this case, there is a σ-finite positive measure μ such that

$$\forall s<t\in\mathbb{R}_+ \qquad \int_{(s,t]}|\mathrm{d}a| = \mu\big((s,t]\big).$$

For $f\in L^1(\mu)$, we write

$$\int_0^\infty f\,\mathrm{d}a := \lim_{t\to\infty}\int_0^t f\,\mathrm{d}a.$$

4.1.2. Finite-Variation Processes

Fix a filtered probability space.

Definition 4.4. *A process A is a **finite-variation process** if A is adapted, all its sample paths are finite-variation functions on $\mathbb{R}_+$, and, for us, $A_0=0$ and all sample paths are continuous. If also the sample paths are increasing, then A is an **increasing process**.*

If A is a finite-variation process, then the process

$$V_t := \int_0^t |\mathrm{d}A_s|$$

is an increasing process: adaptedness follows from Proposition 4.2. As for the case of functions, A is the difference of two increasing processes:

$$A_t = \frac{V_t+A_t}{2}-\frac{V_t-A_t}{2}.$$

Integration with respect to a finite-variation process can be done pointwise, but we need something to guarantee the result will be adapted:

Proposition 4.5. *Let A be a finite-variation process and H be a progressive process satisfying*

$$\forall t \geqslant 0 \ \ \forall \omega \in \Omega \qquad \underbrace{\int_0^t \left|H_s(\omega)\right| \left|\mathrm{d}A_s(\omega)\right|}_{\text{integration with respect to } s, \text{ not } \omega} < \infty.$$

Then the process $H \cdot A$ defined by

$$(H \cdot A)_t \coloneqq \int_0^t H_s \,\mathrm{d}A_s$$

is a finite-variation process.

Proof. We already saw that $H \cdot A$ has sample paths that are finite-variation functions, so it remains to check that $H \cdot A$ is adapted. Recall that for all t, $H\colon \Omega \times [0,t] \to \mathbb{R}$ is measurable with respect to $\mathscr{F} \otimes \mathscr{B}\big([0,t]\big)$. Thus, it suffices to show that if

$$h\colon \Big(\Omega \times [0,t], \mathscr{F}_t \otimes \mathscr{B}\big([0,t]\big)\Big) \to \big(\mathbb{R}, \mathscr{B}(\mathbb{R})\big)$$

and

$$\forall \omega \in \Omega \qquad \int_0^t \left|h(\omega,s)\right| \left|\mathrm{d}A_s(\omega)\right| < \infty,$$

then

$$\Big(\omega \mapsto \int_0^t h(\omega,s)\,\mathrm{d}A_s(\omega)\Big) \in \mathscr{F}_t.$$

This is like Fubini's theorem. We start with h of the form $h(\omega,s) = \mathbf{1}_\Gamma(\omega)\mathbf{1}_{(u,v]}(s)$ for $\Gamma \in \mathscr{F}_t$ and $0 \leqslant u < v \leqslant t$. In this case,

$$\int_0^t h(\omega,s)\,\mathrm{d}A_s(\omega) = \underbrace{\mathbf{1}_\Gamma(\omega)}_{\in \mathscr{F}_t}\big[\underbrace{A_v(\omega)}_{\in \mathscr{F}_v \subseteq \mathscr{F}_t} - \underbrace{A_u(\omega)}_{\in \mathscr{F}_u \subseteq \mathscr{F}_t}\big] \in \mathscr{F}_t.$$

The class of such $\Gamma \times (u,v]$ is closed under finite intersections, so forms a π-system. Also the class of $G \in \mathscr{F}_t \otimes \mathscr{B}\big([0,t]\big)$ such that $h = \mathbf{1}_G$ satisfies the conclusion is closed under complements and countable disjoint unions, so forms a λ-system. Therefore, the class is exactly $\mathscr{F}_t \otimes \mathscr{B}\big([0,t]\big)$. Taking a limit of simple functions dominated by $|h|$ gives the desired result. ◄

Suppose we have only

$$\exists \text{ negligible } N \subseteq \Omega \ \ \forall \omega \notin N \ \ \forall t \geqslant 0 \qquad \int_0^t \left|H_s(\omega)\right| \cdot \left|\mathrm{d}A_s(\omega)\right| < \infty.$$

If the filtration is complete, then we can define $H' \cdot A$, where

$$H'_t(\omega) = \begin{cases} H_t(\omega) & \text{if } \omega \notin N, \\ 0 & \text{if } \omega \in N. \end{cases}$$

Because of completeness, H' is progressive. We then define $H \cdot A := H' \cdot A$.

If H and K are progressive, then so is HK, where $(HK)_t := H_t K_t$. If H and HK satisfy the integrability condition of Proposition 4.5, then

$$K \cdot (H \cdot A) = (KH) \cdot A$$

because for functions and measures,

$$k(h\mu) = (kh)\mu.$$

In the simple case $A_t \equiv t$ and H is progressive with

$$\forall \omega \in \Omega \ \ \forall t \geqslant 0 \qquad \int_0^t |H_s(\omega)| \, \mathrm{d}s < \infty,$$

we obtain a finite-variation process $\int_0^t H_s \, \mathrm{d}s$.

4.2. Continuous Local Martingales

For a process $X = (X_t)_{t \geqslant 0}$ and a stopping time T, we write $X^T := (X_{t \wedge T})_{t \geqslant 0}$ for ***the process*** X ***stopped at*** T. Note that if S is also a stopping time, then

$$(X^T)^S = X^{T \wedge S} = (X^S)^T.$$

Recall from Corollary 3.24 that if X is a martingale and T is a bounded stopping time, then X^T is a uniformly integrable martingale. But there are other processes that have this property besides martingales.

Like local integrability on $\mathbb{R}_+$, but instead of $[0, t_n]$, we use $[0, T_n]$ in the following definition.

Definition 4.6. *An adapted process M with continuous sample paths and $M_0 = 0$ a.s. is called a* ***continuous local martingale*** *if there exist stopping times $T_1 \leqslant T_2 \leqslant \cdots \to \infty$ such that for all n, M^{T_n} is a uniformly integrable martingale. If we do not assume $M_0 = 0$ a.s. but $(M_t - M_0)_{t \geqslant 0}$ satisfies the preceding condition, then we still call M a* ***continuous local martingale****. Stopping times T_n that witness the definition are said to* ***reduce*** *M.*

One need not assume sample paths are continuous in order to define local martingales, but we will.

Note that it is *not* assumed that $M_t \in L^1$. In particular, M_0 need only be $\mathscr{F}_0$-measurable.

To distinguish martingales from local martingales, we may speak of ***true martingales***. Some examples of the difference:

Example. Let B be an $(\mathscr{F}_t)$-Brownian motion from 0 and $Z \in \mathscr{F}_0$, $Z \notin L^1$. Then $M_t := ZB_t$ is a continuous local martingale but not a true martingale by Exercise 4.22.

Exercise (due 11/2). Exercise 4.22.

Example. Let B be a real Brownian motion and $T := \inf\{t \geqslant 0\,;\ B_t = -1\}$. Define

$$X_t := \begin{cases} B_{\frac{t}{1-t} \wedge T} & \text{if } t < 1, \\ -1 & \text{if } t \geqslant 1. \end{cases}$$

By Corollary 3.24(i), B^T is a martingale. We claim that X is a continuous local martingale with respect to $(\mathscr{F}_t^X)_{t \geqslant 0}$ but not a martingale. To see this, let

$$T_n := \inf\{t \geqslant 0\,;\ X_t = n\}.$$

These are stopping times that increase to infinity. We want to show that $0 \leqslant s < t$ implies

$$X_s^{T_n} = \mathbf{E}\big[X_t^{T_n} \bigm| \mathscr{F}_s^X\big].$$

Write

$$\varphi(s) := \begin{cases} \frac{s}{1-s} & \text{if } s < 1, \\ \infty & \text{if } s \geqslant 1. \end{cases}$$

Since $X^t = B^T_{\varphi(t)}$, we have

$$X_s^{T_n} = B_{\varphi(s)}^{\varphi(T_n) \wedge T} \quad \text{and} \quad \mathscr{F}_s^X = \mathscr{F}^B_{\varphi(s) \wedge T},$$

so the desired equation is

$$B_{\varphi(s)}^{\varphi(T_n) \wedge T} = \mathbf{E}\big[B_{\varphi(t)}^{\varphi(T_n) \wedge T} \bigm| \mathscr{F}^B_{\varphi(s) \wedge T}\big].$$

Since $\varphi(T_n)$ is an $(\mathscr{F}_\bullet^B)$-stopping time and $B^{\varphi(T_n) \wedge T}$ is a bounded martingale, the equation follows from the optional stopping theorem. To see that X is not a true martingale, we note that $X_0 = 0 \neq -1 = \mathbf{E}[X_1]$. In effect, B^T is not closed, but is still a martingale.

Some properties of continuous local martingales:

(b) If M is a continuous adapted process with $M_0 = 0$ and T_n are increasing stopping times going to infinity with M^{T_n} a martingale, then M is a continuous local martingale: we may replace T_n by $T_n \wedge n$ to get stopping times that reduce M because $(M^{T_n})^n$ is a uniformly integrable martingale by Theorem 3.21.

(c) If M is a continuous local martingale and T is any stopping time, then M^T is a continuous local martingale, because if T_n reduce M, then $(M^T)^{T_n} = (M^{T_n})^T$ and Corollary 3.24(ii) shows that T_n reduce M^T.

(d) Similarly, if T_n reduce M and $S_n \to \infty$ are stopping times, then $T_n \wedge S_n$ reduce M.

(e) If T_n reduce M and T_n' reduce M', then $T_n \wedge T_n'$ reduce both M and M' by (d), whence also reduce $M + M'$ (the sum of two uniformly integrable classes is uniformly integrable). Thus, the space of continuous local martingales is a vector space.

Exercise (due 11/9). Show that if X is a discrete-time adapted process in L^1 and T_n are stopping times going to infinity such that X^{T_n} is a martingale for each n, then X is a martingale.

If M is a continuous local martingale reduced by (T_n) and $M_0 \in L^1$ then M^{T_n} is a uniformly integrable martingale, since adding an L^1 function to a uniformly integrable class results in a uniformly integrable class.

Proposition 4.7. *Let M be a continuous local martingale with $M_0 \in L^1$.*

(i) *If $M \geqslant 0$, then M is a supermartingale.*

(ii) *If M is dominated (i.e., $\exists Z \in L^1$ with $|M_t| \leqslant Z$ for all $t \geqslant 0$), then M is a uniformly integrable martingale.*

(iii) *M is reduced by*

$$T_n := \inf\{t \geqslant 0\,;\ |M_t| \geqslant n + |M_0|\}.$$

Proof. (i) Let T_n reduce M. Then $\mathbf{E}[M_0] = \mathbf{E}[M_{t\wedge T_n}]$ for $n \geqslant 0$. By Fatou's lemma,

$$\mathbf{E}[M_t] \leqslant \mathbf{E}[M_0] < \infty.$$

Furthermore,

$$s \leqslant t \quad \Longrightarrow \quad \forall n \quad M_{s\wedge T_n} = \mathbf{E}[M_{t\wedge T_n} \mid \mathscr{F}_s]. \tag{$*$}$$

By Fatou's lemma for conditional expectation, we get

$$M_s \geqslant \mathbf{E}[M_t \mid \mathscr{F}_s].$$

(ii) Combined with Eq. ($*$), the Lebesgue dominated convergence theorem implies that $M_s = \mathbf{E}[M_t \mid \mathscr{F}_s]$.

(iii) Proposition 3.9 shows that T_n are stopping times. By property (c) of local martingales, M^{T_n} is a continuous local martingale. Since it is dominated by $n + |M_0|$, part (ii) shows that it is a uniformly integrable martingale, as required. ◀

It is *not* true that a uniformly integrable continuous local martingale is necessarily a martingale, even if it is bounded in L^2. A natural example appears in Exercise 5.33 (which was historically the first example, due to Johnson and Helms in 1963; two years later, Itô and Watanabe introduced local martingales).

Recall Corollary 2.17 that Brownian motion has infinite variation on every non-trivial interval. This came from the fact that the quadratic variation was positive on every interval (indeed, it equals the length of interval), which was a simple consequence of Brownian motion coming from Gaussian white noise. Every continuous local martingale M also has infinite variation on every interval where it is "changing". To prove this, we could try to use Proposition 3.14: if M is a square-integrable martingale, then for $0 = t_0 < t_1 < \cdots < t_p$, we have

$$\mathbf{E}[M_{t_p}^2 - M_{t_0}^2] = \mathbf{E}\Big[\sum_{i=1}^{p} (M_{t_i} - M_{t_{i-1}})^2\Big].$$

However, M need not even be a martingale, nor square-integrable. In addition, we don't have almost sure convergence of the quadratic variation to a something greater than 0 (though we will prove this convergence in probability in the next section). Instead, localization will allow us to get a proof, using a proper stopping time derived from the assumption to the contrary that the variation is finite.

Theorem 4.8. *Let M be a continuous local martingale that is also a finite-variation process. Then* $\mathbf{P}[\forall t \geqslant 0 \;\; M_t = 0] = 1$.

Proof. Since $t \mapsto \int_0^t |\mathrm{d}M_s|$ is an increasing process, for $n \in \mathbb{N}$

$$T_n := \inf\Big\{t \geqslant 0\,;\; \int_0^t |\mathrm{d}M_s| \geqslant n\Big\}$$

is a stopping time by Proposition 3.9. It is enough to show that

$$\forall n \qquad M^{T_n} = 0 \text{ a.s.},$$

since $T_n \to \infty$.

Fix n and write $N := M^{T_n}$. Then

$$\forall t \geqslant 0 \qquad |N_t| = |M_{t\wedge T_n}| = \left|\int_0^{t\wedge T_n} \mathrm{d}M_s\right| \leqslant \int_0^{t\wedge T_n} |\mathrm{d}M_s| \leqslant n.$$

By property (c) and Proposition 4.7(ii), N is a bounded martingale. For $t > 0$, consider $0 = t_0 < t_1 < \cdots < t_p = t$. By Proposition 3.14,

$$\begin{aligned}\mathbf{E}[N_t^2] &= \mathbf{E}\Big[\sum_{i=1}^{p} (N_{t_i} - N_{t_{i-1}})^2\Big] \\ &\leqslant \mathbf{E}\Big[\underbrace{\sup_i |N_{t_i} - N_{t_{i-1}}|}_{\leqslant 2n,\text{ small by continuity}} \cdot \overbrace{\sum_{i=1}^{p} |N_{t_i} - N_{t_{i-1}}|}^{\leqslant n \text{ by Proposition 4.2}}\Big].\end{aligned}$$

Thus, the bounded convergence theorem yields $\mathbf{E}[N_t^2] = 0$, whence $N_t = 0$ a.s. Because N is continuous, it follows that $N = 0$ a.s., as desired. ◀

Exercise (due 11/9). Let $p > 1$ and X be a right-continuous martingale satisfying $\sup_t \mathbf{E}\big[|X_t|^p\big] < \infty$. Show that for all measurable $T \colon \Omega \to [0, \infty]$, $X_T \in L^p$. Show that this is not always true for $p = 1$.

4.3. The Quadratic Variation of a Continuous Local Martingale

For the rest of the chapter, we assume $(\mathscr{F}_t)$ is complete.

Again like Brownian motion, continuous local martingales have finite quadratic variation on every bounded interval. This will be a crucial result and is the main result of Chapter 4.

Theorem 4.9. *Let M be a continuous local martingale. There is an increasing process $\langle M, M\rangle = \big(\langle M, M\rangle_t\big)_{t\geqslant 0}$ such that $\big(M_t^2 - \langle M, M\rangle_t\big)_{t\geqslant 0}$ is a continuous local martingale. Such a process $\langle M, M\rangle$ is unique up to indistinguishability and has the following form: if $t > 0$ and $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$ is an increasing sequence of subdivisions of $[0, t]$ with mesh going to zero, then*

$$\langle M, M\rangle_t = \lim_{n\to\infty} \sum_{i=1}^{p_n} (M_{t_i^n} - M_{t_{i-1}^n})^2 \quad \textit{in probability}. \tag{4.3}$$

Remark. The sum in Eq. (4.3) is not monotone in n, unlike for total variation.

We call $\langle M, M\rangle$ the ***quadratic variation*** of M. For example, if B is a Brownian motion, then $\langle B, B\rangle_t = t$.

From Eq. (4.3), we see that $\langle M, M\rangle$ does not depend on M_0, nor on $(\mathscr{F}_t)_{t\geqslant 0}$. Also, Eq. (4.3) holds even if the subdivisions are not increasing, but we will prove that only in Chapter 5.

The proof of Theorem 4.9 relies on some calculations.

Lemma A. *Let $(X_k)_{k\in\mathbb{N}}$ be a martingale, $(k_i)_{i\in\mathbb{N}}$ be an increasing sequence with $k_0 = 0$, and $i(k) \coloneqq \min\{i\,;\ k \leqslant k_i\}$. Define*

$$Y_m \coloneqq \sum_{k=1}^{m} X_{k-1}(X_k - X_{k-1})$$

and

$$Z_\ell \coloneqq \sum_{i=1}^{\ell} X_{k_{i-1}}(X_{k_i} - X_{k_{i-1}}).$$

Then

$$\mathbf{E}\big[(Y_{k_\ell} - Z_\ell)^2\big] = \mathbf{E}\Bigg[\sum_{k=1}^{k_\ell} (X_{k_{i(k)-1}} - X_{k-1})^2 (X_k - X_{k-1})^2\Bigg].$$

Proof. We have

$$\mathbf{E}[Y_{k_\ell} Z_\ell] = \sum_{i=1}^{\ell}\sum_{k=1}^{k_\ell} \mathbf{E}\big[X_{k_{i-1}}(X_{k_i} - X_{k_{i-1}}) X_{k-1}(X_k - X_{k-1})\big].$$

If $k_i \leqslant k-1$, then by conditioning on $\mathscr{F}_{k-1}$, we see that the (i,k) summand is 0. Similarly, if $k \leqslant k_{i-1}$, then by conditioning on $\mathscr{F}_{k_{i-1}}$, we get that the (i,k) summand is 0. Therefore,

$$\mathbf{E}[Y_{k_\ell} Z_\ell] = \sum_{k=1}^{k_\ell} \mathbf{E}\big[X_{k_{i(k)-1}}(X_{k_{i(k)}} - X_{k_{i(k)-1}}) X_{k-1}(X_k - X_{k-1})\big].$$

Writing

$$X_{k_{i(k)}} - X_{k_{i(k)-1}} = \sum_{j=k_{i(k)-1}+1}^{k_{i(k)}} (X_j - X_{j-1}),$$

we get that the (k,j) summand is 0 unless $j = k$: if $j < k$, condition on $\mathscr{F}_{k-1}$, whereas if $j > k$, condition on $\mathscr{F}_{j-1}$. Thus, we have

$$\mathbf{E}[Y_{k_\ell} Z_\ell] = \sum_{k=1}^{k_\ell} \mathbf{E}\big[X_{k_{i(k)-1}} X_{k-1}(X_k - X_{k-1})^2\big]. \tag{1}$$

By choosing $k_i \equiv i$, we obtain

$$\mathbf{E}[Y_{k_\ell}^2] = \sum_{k=1}^{k_\ell} \mathbf{E}\big[X_{k-1}^2 (X_k - X_{k-1})^2\big]. \tag{2}$$

If we apply Eq. (2) to the martingale $(X_{k_i})_{i\in\mathbb{N}}$, we obtain

$$\mathbf{E}[Z_\ell^2] = \sum_{i=1}^{\ell} \mathbf{E}\big[X_{k_{i-1}}^2 (X_{k_i} - X_{k_{i-1}})^2\big].$$

Now condition on $\mathscr{F}_{k_{i-1}}$ and use Proposition 3.14 to write

$$\mathbf{E}\big[(X_{k_i} - X_{k_{i-1}})^2 \mid \mathscr{F}_{k_{i-1}}\big] = \sum_{k=k_{i-1}+1}^{k_i} \mathbf{E}\big[(X_k - X_{k-1})^2 \mid \mathscr{F}_{k_{i-1}}\big],$$

whence

$$\begin{aligned}\mathbf{E}[Z_\ell^2] &= \sum_{i=1}^{\ell} \mathbf{E}\Big[X_{k_{i-1}}^2 \sum_{k=k_{i-1}+1}^{k_i} \mathbf{E}\big[(X_k - X_{k-1})^2 \mid \mathscr{F}_{k_{i-1}}\big]\Big] \\ &= \sum_{k=1}^{k_\ell} \mathbf{E}\big[X_{k_{i(k)-1}}^2 (X_k - X_{k-1})^2\big].\end{aligned} \tag{3}$$

Using Eqs. (1) to (3), we get the desired result. ◂

Lemma B. *If $(X_k)_{k\in\mathbb{N}}$ is a martingale, then*

$$\forall m \in \mathbb{N} \qquad \mathbf{E}\Bigg[\bigg(\sum_{k=1}^{m} (X_k - X_{k-1})^2\bigg)^2\Bigg] \leqslant 6 \cdot \max_{0\leqslant k\leqslant m} \|X_k\|_\infty^4.$$

Proof. Write $A := \max_{0\leqslant k\leqslant m} \|X_k\|_\infty$. Note that

$$\begin{aligned}&\sum_{1\leqslant k<j\leqslant m} \mathbf{E}\big[(X_k - X_{k-1})^2 (X_j - X_{j-1})^2\big] \\ &\qquad = \sum_{k=1}^{m-1} \mathbf{E}\Big[(X_k - X_{k-1})^2\, \mathbf{E}\Big[\sum_{j=k+1}^{m} (X_j - X_{j-1})^2 \,\Big|\, \mathscr{F}_k\Big]\Big] \\ &\qquad = \sum_{k=1}^{m-1} \mathbf{E}\Big[(X_k - X_{k-1})^2\, \mathbf{E}\big[X_m^2 - X_k^2 \mid \mathscr{F}_k\big]\Big] \quad \text{[by Proposition 3.14]} \\ &\qquad \leqslant A^2 \cdot \sum_{k=1}^{m-1} \mathbf{E}\big[(X_k - X_{k-1})^2\big].\end{aligned}$$

In addition,

$$\mathbf{E}\big[(X_k - X_{k-1})^4\big] \leqslant 4A^2\, \mathbf{E}\big[(X_k - X_{k-1})^2\big].$$

Therefore,

$$\begin{aligned}\mathbf{E}\Big[\Big(\sum_{k=1}^{m} (X_k - X_{k-1})^2\Big)^2\Big] &\leqslant 6A^2 \sum_{k=1}^{m} \mathbf{E}\big[(X_k - X_{k-1})^2\big] \\ &= 6A^2\, \mathbf{E}\big[X_m^2 - X_0^2\big] \quad \text{[by Proposition 3.14]} \\ &\leqslant 6A^4.\end{aligned}$$

◂

Lemma C. *If* $\forall n \in \mathbb{N}\ \ X^n = (X_t^n)_{t\in I}$ *is a process with continuous sample paths and*

$$\lim_{n,m\to\infty} \mathbf{E}\big[\sup_{t\in I}(X_t^n - X_t^m)^2\big] = 0,$$

then there exists $n_k \to \infty$ *and* $Y = (Y_t)_{t\in I}$ *with continuous sample paths such that almost surely*

$$\forall t \in I \qquad \lim_{k\to\infty} X_t^{n_k} = Y_t.$$

Proof. Choose $n_k \to \infty$ such that

$$\sum_{k=1}^{\infty} \mathbf{E}\big[\sup_{t\in I}(X_t^{n_k} - X_t^{n_{k+1}})^2\big]^{1/2} < \infty.$$

Then $\mathbf{E}\big[\sum_{k=1}^{\infty}\sup_{t\in I}|X_t^{n_k} - X_t^{n_{k+1}}|\big] = \sum_{k=1}^{\infty}\mathbf{E}\big[\sup_{t\in I}|X_t^{n_k} - X_t^{n_{k+1}}|\big] < \infty$, so

$$\sum_{k=1}^{\infty}\sup_{t\in I}|X_t^{n_k} - X_t^{n_{k+1}}| < \infty \quad \text{almost surely.}$$

Off a negligible set N, we have uniform convergence of $X_t^{n_k}$, so for $\omega \notin N$ one may define $Y(\omega) := \lim_{k\to\infty} X_t^{n_k}(\omega)$, whereas for $\omega \in N$, define $Y(\omega) := 0$. (Note that N depends on $X_t^{(n_k)}$ for all $t \in I$—or at least a dense subset of such t—and all k, so that we cannot conclude that $Y_t \in \sigma(X_t^{(n_k)}, k \geqslant 1)$.) ◂

Proof of Theorem 4.9. We first show uniqueness. Suppose that A and A' are increasing processes such that $(M_t^2 - A_t)$ and $(M_t^2 - A_t')$ are both continuous local martingales. Then their difference, $A_t' - A_t$, is a continuous local martingale and a finite-variation process, whence is 0 by Theorem 4.8 (up to indistinguishability).

To prove existence, first assume $M_0 = 0$ and M is bounded. By Proposition 4.7(ii), M is a true martingale. Fix $K > 0$ and an increasing sequence of subdivisions of $[0, K]$ with mesh going to 0, $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = K$.

It is easy to see that if $0 \leqslant r < s$ and $Z \in L^\infty(\mathscr{F}_r)$, then $t \mapsto Z(M_{s\wedge t} - M_{r\wedge t})$ is a martingale. Therefore, for all n, the process

$$X_t^n := \sum_{i=1}^{p_n} M_{t_{i-1}^n}\big(M_{t_i^n\wedge t} - M_{t_{i-1}^n\wedge t}\big)$$

is a bounded martingale. Now for $0 \leqslant j \leqslant p_n$,

$$
\begin{aligned}
(M_{t_j^n})^2 - 2X_{t_j^n}^n &= M_{t_j^n}^2 - 2\sum_{i=1}^{j} M_{t_{i-1}^n}(M_{t_i^n} - M_{t_{i-1}^n}) \\
&= M_{t_j^n}^2 - 2M_{t_j^n}M_{t_{j-1}^n} + M_{t_{j-1}^n}^2 \\
&\quad + M_{t_{j-1}^n}^2 - 2M_{t_{j-1}^n}M_{t_{j-2}^n} + M_{t_{j-2}^n}^2 \\
&\quad + \cdots \\
&\quad + M_{t_1^n}^2 - 2M_{t_1^n}M_{t_0^n} + M_{t_0^n}^2 \\
&\quad + M_{t_0^n}^2 \\
&= \sum_{i=1}^{j}(M_{t_i}^n - M_{t_{i-1}^n})^2
\end{aligned}
\tag{4.4}
$$

since $M_{t_0^n} = M_0 = 0$. (In Chapter 5, we will see this implies $M_t^2 - \langle M, M\rangle_t = 2\int_0^t M_s \,\mathrm{d}M_s$. Note that if M is a finite-variation process, then $\langle M, M\rangle = 0$ and this is ordinary calculus.)

By Lemma A, if $n \leqslant m$,

$$\mathbf{E}\big[(X_K^n - X_K^m)^2\big] = \mathbf{E}\Bigg[\sum_{j=1}^{p_m}(M_{t_{i_n(j)-1}^n} - M_{t_{j-1}^m})^2 \cdot (M_{t_j^m} - M_{t_{j-1}^m})^2\Bigg],$$

where $i_n(j) \coloneqq \min\{i\,;\ t_j^m \leqslant t_i^n\}$. The right-hand side is

$$
\begin{aligned}
&\leqslant \mathbf{E}\Bigg[\max_{1\leqslant j\leqslant p_m}\Big|M_{t_{i_n(j)-1}^n} - M_{t_{j-1}^m}\Big|^2 \cdot \sum_{i=1}^{p_m}(M_{t_j^m} - M_{t_{j-1}^m})^2\Bigg] \\
&\leqslant \mathbf{E}\Big[\max_j \Big|M_{t_{i_n(j)-1}^n} - M_{t_{j-1}^m}\Big|^4\Big]^{1/2} \cdot \mathbf{E}\Big[\Big(\sum_j (M_{t_j^n} - M_{t_{j-1}^n})^2\Big)^2\Big]^{1/2}.
\end{aligned}
$$

[Cauchy–Schwarz inequality]

The first term converges to 0 as $m, n \to \infty$ by continuity of sample paths and boundedness of M. The second term is less than or equal to $\sqrt{6}\cdot\sup_{0\leqslant t\leqslant K}\|M_t\|_\infty^2$ by Lemma B. Therefore,

$$\lim_{n,m\to\infty}\mathbf{E}\big[(X_K^n - X_K^m)^2\big] = 0.$$

By Doob's L^2-inequality (Proposition 3.15(ii)), we get

$$\lim_{m,n\to\infty}\mathbf{E}\big[\sup_{t\leqslant K}(X_t^n - X_t^m)^2\big] = 0.$$

By Lemma C, there exists $Y = (Y_t)_{0\leqslant t\leqslant K}$ with continuous sample paths and $n_k \to \infty$ such that almost surely,

$$\forall t \in [0, K] \qquad \lim_{k\to\infty} X_t^{n_k} = Y_t.$$

Also,

$$\forall t \in [0, K] \qquad \lim_{n\to\infty} X_t^{n_k} = Y_t \quad \text{in } L^2.$$

Because the filtration is complete, $Y_t \in \mathscr{F}_t$. Since $s \leqslant t$ implies that $\mathbf{E}[X_t^n \mid \mathscr{F}_s] = X_s^n$, we obtain that $\mathbf{E}[Y_t \mid \mathscr{F}_s] = Y_s$ for $0 \leqslant s \leqslant t \leqslant K$, i.e., $(Y_{t\wedge K})_{t\geqslant 0}$ is a continuous martingale.

By Eq. (4.4), the sample paths of $M_t^2 - 2X_t^n$ are increasing along the sequence $t_0^n < t_1^n < \cdots < t_{p_n}^n$. Therefore, the sample paths of $M_t^2 - 2Y_t$ are increasing on $[0, K]$ off a negligible set N. Thus, define the increasing process $A^{(K)}$ on $[0, K]$ by

$$A_t^{(K)} := \begin{cases} M_t^2 - 2\,Y_t & \text{on } \Omega \setminus N, \\ 0 & \text{on } N. \end{cases}$$

Then $A^{(K)}$ is an increasing process and $\big(M_{t\wedge K}^2 - A_{t\wedge K}^{(K)}\big)_{t\geqslant 0}$ is a continuous martingale.

In this manner, for all $\ell \in \mathbb{N}$, we obtain a process $A^{(\ell)}$ on $[0, \ell]$. By the uniqueness argument at the beginning of this proof, $\big(A_{t\wedge\ell}^{(\ell+1)}\big)_{t\geqslant 0}$ and $\big(A_{t\wedge\ell}^{(\ell)}\big)_{t\geqslant 0}$ are indistinguishable. This allows us to define an increasing process $\langle M, M\rangle$ such that $\big(\langle M, M\rangle_{t\wedge\ell}\big)_{t\geqslant 0}$ is indistinguishable from $\big(A_{t\wedge\ell}^{(\ell)}\big)_{t\geqslant 0}$ for each $\ell \in \mathbb{N}$. It satisfies that $\big(M_t^2 - \langle M, M\rangle_t\big)_{t\geqslant 0}$ is a martingale.

This is not quite Eq. (4.3) because there t was arbitrary and the subdivisions were of $[0, t]$. However, call "t" there now by "K". As before, $\big(A_{t\wedge K}^{(K)}\big)_{t\geqslant 0}$ is indistinguishable from $\big(\langle M, M\rangle_{t\wedge K}\big)_{t\geqslant 0}$. In particular, $\langle M, M\rangle_K = A_{(K)}^K$ almost surely. As we saw, this gives L^2-convergence in Eq. (4.3), which is stronger than convergence in probability. This completes the proof when $M_0 = 0$ and M is bounded.

For the general case, write $M_t = M_0 + N_t$. Then $M_t^2 = M_0^2 + 2M_0N_t + N_t^2$. By Exercise 4.22, $(M_0N_t)_t$ is a continuous local martingale, so by uniqueness, $\langle M, M\rangle = \langle N, N\rangle$. Thus, we may take $M_0 = 0$ without loss of generality.

Now use the stopping times $T_n := \inf\big\{t \geqslant 0\,;\ |M_t| \geqslant n\big\}$. The case we proved applies to M^{T_n}: Write $A^{[n]} := \langle M^{T_n}, M^{T_n}\rangle$. By uniqueness, for all n, $\big(A_{t\wedge T_n}^{[n+1]}\big)_t$ and $\big(A_t^{[n]}\big)_t$ are indistinguishable, so there exists an increasing process A such that for all $n \in \mathbb{N}$, A^{T_n} and $A^{[n]}$ are indistinguishable. By construction, for all n, $(M_{t\wedge T_n}^2 - A_{t\wedge T_n})_t$ is a martingale, whence $(M_t^2 - A_t)_t$ is a continuous local martingale. Thus, we may define $\langle M, M\rangle := A$. Now, Eq. (4.3) in the bounded case says that

$$\forall n \ \ \forall t \qquad \lim_{m\to\infty} Z_{t\wedge T_n}^{(m)} = \langle M, M\rangle_{t\wedge T_n} \quad \text{in probability,}$$

where

$$Z_t^{(m)} := \sum_{i=1}^{p_m} (M_{t_i^m\wedge t} - M_{t_{i-1}^m\wedge t})^2.$$

That is,

$$\forall n \ \ \forall t \ \ \forall \varepsilon > 0 \ \ \exists m_0 \ \ \forall m \geqslant m_0 \qquad \mathbf{P}\Big[\big|Z_{t\wedge T_n}^{(m)} - \langle M, M\rangle_{t\wedge T_n}\big| > \varepsilon\Big] < \varepsilon.$$

In addition, there exists n_0 such that $\mathbf{P}[T_{n_0} < t] < \varepsilon$. Therefore,

$$\forall m \geqslant m_0 \qquad \mathbf{P}\Big[\big|Z_t^{(m)} - \langle M, M\rangle_t\big| > \varepsilon\Big] < 2\varepsilon,$$

so $Z_t^{(m)} \xrightarrow{\mathbf{P}} \langle M, M\rangle_t$, as desired. ◂

Exercise (due 11/30). Exercise 4.23.

Proposition 4.11. *If M is a continuous local martingale and T is a stopping time, then almost surely,*

$$\forall t \geqslant 0 \qquad \langle M^T, M^T \rangle_t = \langle M, M \rangle_t^T .$$

Proof. By property (c) of continuous local martingales, $\left(M_{t \wedge T}^2 - \langle M, M \rangle_{t \wedge T}\right)_t$ is a continuous local martingale. By uniqueness, $\left(\langle M, M \rangle_{t \wedge T}\right)_t$ is indistinguishable from $\langle M^T, M^T \rangle_t$. ◀

Exercise (due 11/30). Show that if M is a continuous local martingale and T is a stopping time, then almost surely,

$$\forall t \geqslant 0 \quad \langle M - M^T, M - M^T \rangle_t = \langle M, M \rangle_t - \langle M, M \rangle_t^T .$$

Exercise (due 11/30). Let B be an $(\mathscr{F}_t)$-Brownian motion and S, T be stopping times. Calculate

$$\langle B^T - B^S, B^T - B^S \rangle.$$

Hint: Do first the case $S \leqslant T$.

We next show how various properties of M are reflected in $\langle M, M \rangle$. Our first result is that M changes only where $\langle M, M \rangle$ changes.

Proposition 4.12. *Let M be a continuous local martingale and $0 \leqslant t_1 < t_2 \leqslant \infty$. Then a.s. $\forall t \in [t_1, t_2]$ $M_t = M_{t_1}$ if and only if a.s. $\forall t \in [t_1, t_2]$ $\langle M, M \rangle_t = \langle M, M \rangle_{t_1}$.*

Proof. $\Rightarrow$: By Eq. (4.3), $\langle M, M \rangle_{t_2} = \langle M, M \rangle_{t_1}$. Since $\langle M, M \rangle$ is increasing, we get the result.

$\Leftarrow$: M^{t_i} is a continuous local martingale by property (c), whence so is $M^{t_2} - M^{t_1}$ by property (e). By Eq. (4.3), $\langle M^{t_2} - M^{t_1}, M^{t_2} - M^{t_1} \rangle = 0$. Therefore, $(M^{t_2} - M^{t_1})^2$ is a continuous local martingale. By Proposition 4.7(i), it is a supermartingale. Thus, $\mathbf{E}\left[(M_t - M_{t_1})^2\right] \leqslant \mathbf{E}\left[(M_{t_1} - M_{t_1})^2\right] = 0$ for $t \in [t_1, t_2]$. ◀

For an increasing process A, we define $A_\infty := \lim_{t \to \infty} A_t \in [0, \infty]$.

Theorem 4.13. *Let M be a continuous local martingale with $M_0 \in L^2$.*

(i) *The following are equivalent:*

(a) *M is a true martingale bounded in L^2.*

(b) $\mathbf{E}\left[\langle M, M \rangle_\infty\right] < \infty$.

If these hold, then $\left(M_t^2 - \langle M, M \rangle_t\right)_{t \geqslant 0}$ is a uniformly integrable martingale and so

$$\mathbf{E}[M_\infty^2] = \mathbf{E}[M_0^2] + \mathbf{E}\left[\langle M, M \rangle_\infty\right].$$

(ii) *The following are equivalent:*

(a) *M is a true martingale and $\forall t \geqslant 0$ $M_t \in L^2$.*

(b) $\forall t \in [0, \infty)$ $\mathbf{E}\left[\langle M, M \rangle_t\right] < \infty$.

If these hold, then $\left(M_t^2 - \langle M, M \rangle_t\right)_{t \geqslant 0}$ is a true martingale.

Proof. (i) Without loss of generality, $M_0 = 0$.

(a) $\Rightarrow$ (b): By Doob's L^2-inequality,

$$\forall s \geqslant 0 \quad \Big\| \sup_{0 \leqslant t \leqslant s} |M_t| \Big\|_2 \leqslant 2\|M_s\|_2,$$

whence $\sup_{t \geqslant 0} |M_t| \in L^2$ if (a) holds. Let $S_n := \inf\{t \geqslant 0\,;\ \langle M, M\rangle_t \geqslant n\}$. By property (c) of continuous local martingales, $\big(M^2_{t \wedge S_n} - \langle M, M\rangle_{t \wedge S_n}\big)_{t \geqslant 0}$ is a continuous local martingale. Also, it is dominated by $(\sup_{t \geqslant 0} M_t^2 + n) \in L^1$, so by Proposition 4.7(ii), it is a uniformly integrable martingale. Therefore,

$$\forall t \geqslant 0 \quad \mathbf{E}\big[\langle M, M\rangle_{t \wedge S_n}\big] = \mathbf{E}\big[M^2_{t \wedge S_n}\big] \leqslant \mathbf{E}\big[\sup_{s \geqslant 0} M_s^2\big] < \infty.$$

Take $n \to \infty$ and $t \to \infty$ to get (b).

(b) $\Rightarrow$ (a): If (b) holds, then set

$$T_n := \inf\{t \geqslant 0\,;\ |M_t| \geqslant n\}.$$

Now

$$\big|M^2_{t \wedge T_n} - \langle M, M\rangle_{t \wedge T_n}\big| \leqslant n^2 + \langle M, M\rangle_\infty \in L^1,$$

so again $\big(M^2_{t \wedge T_n} - \langle M, M\rangle_{t \wedge T_n}\big)_{t \geqslant 0}$ is a uniformly integrable martingale and

$$\forall t \geqslant 0 \quad \mathbf{E}[M^2_{t \wedge T_n}] = \mathbf{E}\big[\langle M, M\rangle_{t \wedge T_n}\big] \leqslant \mathbf{E}\big[\langle M, M\rangle_\infty\big] < \infty. \tag{$*$}$$

Take $n \to \infty$ to get $(M_t)_{t \geqslant 0}$ is bounded in L^2 by Fatou's Lemma.

To see that M is a martingale, note that Eq. ($*$) implies $(M_{t \wedge T_n})_{n \geqslant 1}$ is uniformly integrable, so converges in L^1 to M_t as $n \to \infty$. By Proposition 4.7(iii), M^{T_n} is a martingale, whence so is its L^1-limit, M.

Lastly, if (a) and (b) hold, then

$$\big|M_t^2 - \langle M, M\rangle_t\big| \leqslant \sup_{s \geqslant 0} M_s^2 + \langle M, M\rangle_\infty \in L^1,$$

so by Proposition 4.7(ii), $M^2 - \langle M, M\rangle$ is a uniformly integrable martingale.

(ii) Apply (i) to M^a for each $a \geqslant 0$. ◀

Exercise (due 11/30). Exercise 4.24.

4.4. The Bracket of Two Continuous Local Martingales

The reason for our notation $\langle M, M\rangle$ is that it leads to:

Definition 4.14. *If M and N are continuous local martingales, the **bracket** (or **covariation**) $\langle M, N\rangle$ is the finite-variation process*

$$\langle M, N\rangle_t := \frac{1}{2}\Big(\langle M+N, M+N\rangle_t - \langle M, M\rangle_t - \langle N, N\rangle_t\Big).$$

Proposition 4.15. *Let M and N be continuous local martingales.*

(i) *$\langle M, N\rangle$ is the unique (up to indistinguishability) finite-variation process such that $\big(M_t N_t - \langle M, N\rangle_t\big)_{t\geqslant 0}$ is a continuous local martingale.*

(ii) *The map $(M, N) \mapsto \langle M, N\rangle$ is bilinear and symmetric.*

(iii) *If $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$ is an increasing sequence of subdivisions of $[0, t]$ with mesh going to 0, then*

$$\lim_{n\to\infty} \sum_{i=1}^{p_n} (M_{t_i^n} - M_{t_{i-1}^n})(N_{t_i^n} - N_{t_{i-1}^n}) = \langle M, N\rangle_t \quad \textit{in probability}.$$

(iv) *If T is a stopping time, then*

$$\langle M^T, N^T\rangle = \langle M, N\rangle^T = \langle M^T, N\rangle.$$

(v) *If M and N are both true martingales bounded in L^2, then $MN - \langle M, N\rangle$ is a uniformly integrable martingale, whence $\langle M, N\rangle_\infty$ exists as the almost sure and L^1 limit of $\langle M, N\rangle_t$ as $t \to \infty$ and satisfies*

$$\mathbf{E}[M_\infty N_\infty] = \mathbf{E}[M_0 N_0] + \mathbf{E}\big[\langle M, N\rangle_\infty\big].$$

Proof. (i) This follows from Theorem 4.9, with uniqueness from Theorem 4.8.

(ii) This follows from uniqueness in (i).

(iii) This follows from Eq. (4.3) applied to M, N and $M + N$.

(iv) The first equality follows from (i) as in the proof of Proposition 4.11. By (iii), given $0 \leqslant s \leqslant t$, we may take the subdivisions of $[0, t]$ to include s in order to deduce that

$$\langle M^T, N\rangle_t = \langle M, N\rangle_t \quad \text{a.s. on } [T \geqslant t]$$

and

$$\langle M^T, N\rangle_t = \langle M^T, N\rangle_s \quad \text{a.s. on } [T \leqslant s],$$

whence $\langle M^T, N\rangle_t = \langle M, N\rangle_t^T$ almost surely (consider $s \in \mathbb{Q}_+$).

(v) Apply Theorem 4.13(i) to M, N and $M + N$ to get three uniformly integrable martingales. Combining them gives the result. ◂

Exercise (due 11/30). Give another proof that

$$\langle M - M^T, M - M^T\rangle = \langle M, M\rangle - \langle M, M\rangle^T$$

by using Proposition 4.15.

Proposition 4.16. *If B and B' independent $(\mathscr{F}_t)$-Brownian motions, then $\langle B, B'\rangle = 0$.*

We will skip the proof in favor of the following exercise:

Exercise (due 11/30). Let M and N be independent continuous local martingales (so $\sigma(M) \perp\!\!\!\perp \sigma(N)$). Give two proofs as follows that $\langle M, N\rangle = 0$:

(1) Assume first that M and N are bounded. For $0 = t_0 < t_1 < \cdots < t_n = t$, show that

$$\mathbf{E}\Big[\Big(\sum_{i=1}^{n}(M_{t_i} - M_{t_{i-1}})(N_{t_i} - N_{t_{i-1}})\Big)^2\Big] \leqslant \max_{1\leqslant i\leqslant n}\big(\mathbf{E}[M_{t_i}^2] - \mathbf{E}[M_{t_{i-1}}^2]\big)\cdot\big(\mathbf{E}[N_t^2] - \mathbf{E}[N_0^2]\big).$$

Deduce that $\langle M, N\rangle = 0$.

In the general case, localize M and N and use Proposition 4.15.

(2) Assume first that M and N are martingales. Show that for $0 \leqslant s \leqslant t$, $A \in \mathscr{F}_s^M$, $B \in \mathscr{F}_s^N$, we have

$$\mathbf{E}[M_t N_t \mathbf{1}_{A\cap B}] = \mathbf{E}[M_s N_s \mathbf{1}_{A\cap B}].$$

Deduce that MN is an $\big(\mathscr{F}_t^M \vee \mathscr{F}_t^N\big)_{t\geqslant 0}$-martingale and that $\langle M, N\rangle = 0$.

In the general case, localize M and N.

Definition 4.17. *We say that two continuous local martingales are* ***orthogonal*** *if their bracket is* 0*; this is equivalent to their product being a continuous local martingale.*

Thus, $M \perp\!\!\!\perp N$ implies M and N are orthogonal. The converse is false:

Exercise (due 11/30). Show that if B is an $(\mathscr{F}_t)$-Brownian motion and T is a stopping time, then $\langle B^T, B - B^T\rangle = 0$. Give an example where B^T and $B - B^T$ are *not* independent.

If $\langle M, N\rangle = 0$ and M, N are true martingales bounded in L^2, then

$$\mathbf{E}[M_T N_T] = \mathbf{E}[M_0 N_0]$$

for all stopping times T by Proposition 4.15(v) and Theorem 3.22 (the optional stopping theorem).

Exercise. Prove, conversely, that if M and N are true martingales bounded in L^2 and $\mathbf{E}[M_T N_T] = \mathbf{E}[M_0 N_0]$ for all finite or infinite stopping times T, then $\langle M, N\rangle = 0$.

We are next going to prove a Cauchy–Schwarz type inequality that involves integrating with respect to the bracket of two continuous local martingales. The proof involves the usual Cauchy–Schwarz inequality a couple of times, including via the following:

Lemma. *Let $a\colon \mathbb{R}_+ \to \mathbb{R}$ be a finite-variation function and $v_1, v_2\colon \mathbb{R}_+ \to \mathbb{R}$ be increasing functions. If*

$$\forall 0 \leqslant s < t < \infty \qquad \big|a(t) - a(s)\big| \leqslant \sqrt{v_1(t) - v_1(s)}\cdot\sqrt{v_2(t) - v_2(s)},$$

then for any Borel functions $h, k\colon \mathbb{R}_+ \to \mathbb{R}_+$,

$$\int_0^\infty h\cdot k\,|\mathrm{d}a| \leqslant \Big(\int_0^\infty h^2\,\mathrm{d}v_1\Big)^{1/2}\Big(\int_0^\infty k^2\,\mathrm{d}v_2\Big)^{1/2}. \tag{$*$}$$

Proof. Suppose that $(*)$ holds for functions h_i and k_i that are both are 0 outside a Borel set A_i, with A_i disjoint for different i. Write $h := \sum h_i$ and $k := \sum k_i$. Then

$$\int_0^\infty h\cdot k\,|\mathrm{d}a| = \int_0^\infty \sum h_i k_i\,|\mathrm{d}a| \leqslant \sum\Big(\int_0^\infty h_i^2\,\mathrm{d}v_1\Big)^{1/2}\Big(\int_0^\infty k_i^2\,\mathrm{d}v_2\Big)^{1/2}$$

$$\leqslant \Big(\int_0^\infty \sum h_i^2\,\mathrm{d}v_1\Big)^{1/2}\Big(\int_0^\infty \sum k_i^2\,\mathrm{d}v_2\Big)^{1/2} = \Big(\int_0^\infty h^2\,\mathrm{d}v_1\Big)^{1/2}\Big(\int_0^\infty k^2\,\mathrm{d}v_2\Big)^{1/2}$$

by the Cauchy–Schwarz inequality. That is, ($*$) then holds also for the pair h, k. A similar computation using the hypothesis shows that if $s = t_0 < t_1 < \cdots < t_p = t$, then

$$\sum_{i=1}^{p} \left|a(t_i) - a(t_{i-1})\right| \leqslant \sqrt{v_1(t) - v_1(s)} \cdot \sqrt{v_2(t) - v_2(s)}.$$

(Actually, this is a special case of the initial computation with $h_i = \mathbf{1}_{(t_{i-1},t_i]} \frac{\mathrm{d}a}{|\mathrm{d}a|}$ and $k_i = \mathbf{1}_{(t_{i-1},t_i]} \operatorname{sgn}\big(a(t_i) - a(t_{i-1})\big)$.) Taking a limit of such subdivisions and using Proposition 4.2, we obtain

$$\int_{(s,t]} |\mathrm{d}a| \leqslant \Big(\int_{(s,t]} \mathrm{d}v_1\Big)^{1/2}\Big(\int_{(s,t]} \mathrm{d}v_2\Big)^{1/2},$$

in other words, ($*$) holds for functions of the form $\mathbf{1}_{(s,t]}$. By our first computation, it follows that if B is a finite disjoint union of intervals $(s_i, t_i]$, then

$$\int_B |\mathrm{d}a| \leqslant \Big(\int_B \mathrm{d}v_1\Big)^{1/2}\Big(\int_B \mathrm{d}v_2\Big)^{1/2}. \tag{$**$}$$

The class of B such that ($**$) holds is closed under countable increasing unions and decreasing intersections. Furthermore, the class of finite disjoint unions of intervals $(s, t]$ is an algebra. By Halmos' monotone class lemma, it follows that ($**$) holds for all $B \in \mathscr{B}(\mathbb{R}_+)$. Therefore, ($*$) holds when h and k are multiples of the same indicator, and thus when they are simple functions. We may take monotone increasing limits of simple functions to get the full result. ◀

Proposition 4.18 (Kunita–Watanabe)**.** *If M and N are continuous local martingales and H and K are measurable processes, then almost surely,*

$$\int_0^\infty |H_s| \cdot |K_s| \left|\mathrm{d}\langle M, N\rangle_s\right| \leqslant \Big(\int_0^\infty H_s^2 \,\mathrm{d}\langle M, M\rangle_s\Big)^{1/2}\Big(\int_0^\infty K_s^2 \,\mathrm{d}\langle N, N\rangle_s\Big)^{1/2}.$$

Proof. For $s = t_0 < t_1 \cdots < t_p = t$, we have

$$\Big|\sum_{i=1}^{p} (M_{t_i} - M_{t_{i-1}})(N_{t_i} - N_{t_{i-1}})\Big| \leqslant \Big(\sum_{i=1}^{p} (M_{t_i} - M_{t_{i-1}})^2\Big)^{1/2}\Big(\sum_{i=1}^{p} (N_{t_i} - N_{t_{i-1}})^2\Big)^{1/2}.$$

Taking a limit and using Theorem 4.9 and Proposition 4.15, we get almost surely

$$\left|\langle M, N\rangle_t - \langle M, N\rangle_s\right| \leqslant \big(\langle M, M\rangle_t - \langle M, M\rangle_s\big)^{1/2}\big(\langle N, N\rangle_t - \langle N, N\rangle_s\big)^{1/2}.$$

By taking $s, t \in \mathbb{Q}_+$ and using continuity, we obtain that this holds almost surely simultaneously in $0 \leqslant s < t < \infty$. The result now follows from the lemma. ◀

4.5. Continuous Semimartingales

Definition 4.19. *A process X is a* ***continuous semimartingale*** *if there is a continuous local martingale M and a finite-variation process A such that*

$$\forall t \geqslant 0 \quad X_t = M_t + A_t.$$

By Theorem 4.8, the decomposition $X = M + A$ is unique up to indistinguishability; it is called the ***canonical decomposition*** *of X.*

Definition 4.20. *Let $X = M + A$ and $X' = M' + A'$ be the canonical decompositions of two continuous semimartingales, X and X'. The* ***bracket*** *of X and X' is*

$$\langle X, X' \rangle := \langle M, M' \rangle.$$

Proposition 4.21. *Let X and X' be continuous semimartingales. Given an increasing sequence of subdivisions $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$ of $[0, t]$ whose mesh tends to 0, we have*

$$\lim_{n\to\infty} \sum_{i=1}^{p_n} (X_{t_i^n} - X_{t_{i-1}^n})(X'_{t_i^n} - X'_{t_{i-1}^n}) = \langle X, X' \rangle_t \quad \textit{in probability}.$$

Proof. We have

$$(X_{t_i^n} - X_{t_{i-1}^n})(X'_{t_i^n} - X'_{t_{i-1}^n}) = (M_{t_i^n} - M_{t_{i-1}^n})(M'_{t_i^n} - M'_{t_{i-1}^n}) + \text{terms involving } A \text{ or } A'.$$

The sums of the first terms converge in probability to $\langle M, M' \rangle = \langle X, X' \rangle$ by Proposition 4.15(iii). The other terms have sums going to 0 almost surely by continuity; e.g.,

$$\Big|\sum_{i=1}^{p_n} (M_{t_i^n} - M_{t_{i-1}^n})(A'_{t_i^n} - A'_{t_{i-1}^n})\Big| \leqslant \underbrace{\max_{1\leqslant i\leqslant p_n} |M_{t_i^n} - M_{t_{i-1}^n}|}_{\to 0} \cdot \int_0^t |\mathrm{d}A'_s|.$$ ◂

Chapter 5

Stochastic Integration

This chapter is the heart of the course.

5.1. The Construction of Stochastic Integrals

We fix a complete filtered probability space.

We construct stochastic integrals in stages. In Section 5.1.1, we begin with the analogue of step functions and proceed to integrate with respect to L^2-bounded martingales. We'll find that all the hard work was done earlier, especially Theorem 4.9. In Section 5.1.2, we extend to integrate with respect to continuous local martingales and in Section 5.1.3, to continuous semimartingales; these extensions will be easy. In Section 5.1.4, we prove some limit theorems about stochastic integrals.

5.1.1. Stochastic Integrals for Martingales Bounded in L^2

Every L^2-bounded martingale is closed, so one can think of the space of L^2-bounded martingales as a subspace of $L^2(\Omega)$. However, they are certainly not necessarily all continuous. Thus, we define $\mathbb{H}^2$ to be the space of continuous L^2-bounded martingales M with $M_0 = 0$ and identify it with a subspace of $L^2(\Omega)$, so for $M, N \in \mathbb{H}^2$, we have

$$(M, N)_{\mathbb{H}^2} := (M_\infty, N_\infty)_{L^2(\Omega)} = \mathbf{E}[M_\infty N_\infty] = \mathbf{E}\big[\langle M, N\rangle_\infty\big]$$

by Proposition 4.15(v) and the fact that $M_0 = N_0 = 0$.

This subspace $\mathbb{H}^2$ is closed:

Proposition 5.1. *The space $\mathbb{H}^2$ is a Hilbert space.*

Proof. Suppose that $(M^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{H}^2$, i.e., $(M^n_\infty)_{n\in\mathbb{N}}$ is Cauchy in $L^2(\Omega)$. By Doob's L^2-inequality, $\mathbf{E}\big[\sup_{t\geqslant 0}(M^n_t - M^m_t)^2\big] \leqslant 4\,\mathbf{E}\big[(M^n_\infty - M^m_\infty)^2\big]$, so by Lemma C in Chapter 4 for the proof of Theorem 4.9, there exists a subsequence $(n_k)_{k\in\mathbb{N}}$ and a Y with continuous sample paths such that almost surely,

$$\forall t \geqslant 0 \quad M^{n_k}_t \to Y_t.$$

Also, there exists $Z \in L^2(\Omega)$ such that $M^n_\infty \to Z$ in L^2. This implies

$$M^n_t = \mathbf{E}[M^n_\infty \mid \mathscr{F}_t] \to \mathbf{E}[Z \mid \mathscr{F}_t],$$

whence
$$\forall t \qquad Y_t = \mathbf{E}[Z \mid \mathscr{F}_t].$$
Therefore, Y is an L^2-bounded continuous martingale. Since $M^n_\infty \to Z$ in L^2, we get that Y is the limit of M^n in $\mathbb{H}^2$. ◂

Recall that $\mathscr{P}$ denotes the progressive σ-field. For $M \in \mathbb{H}^2$, write $\langle M, M\rangle\, \mathbf{P}$ for the measure on $\mathscr{P}$ given by
$$A \mapsto \mathbf{E}\Big[\int_0^\infty \mathbf{1}_A(\omega, s)\, \mathrm{d}\langle M, M\rangle_s\Big];$$
the total mass of $\langle M, M\rangle\, \mathbf{P}$ is $\mathbf{E}\big[\langle M, M\rangle_\infty\big] = \|M\|^2_{\mathbb{H}^2}$. Then
$$L^2(M) := L^2\big(\Omega \times \mathbb{R}_+, \mathscr{P}, \langle M, M\rangle\, \mathbf{P}\big) = \Big\{H \in \mathscr{P}\,;\; \mathbf{E}\Big[\int_0^\infty H_s^2\, \mathrm{d}\langle M, M\rangle_s\Big] < \infty\Big\}.$$
This has the usual inner product
$$(H, K)_{L^2(M)} = \mathbf{E}\Big[\int_0^\infty H_s K_s\, \mathrm{d}\langle M, M\rangle_s\Big].$$
Note that
$$\int_0^\infty H_s K_s\, \mathrm{d}\langle M, M\rangle_s \in L^1(\Omega, \mathbf{P})$$
for $H, K \in L^2(M)$.

The analogue of step function is:

Definition 5.2. *An **elementary process** is a process H of the form*
$$H_s(\omega) = \sum_{i=0}^{p-1} H_{(i)}(\omega)\mathbf{1}_{(t_i, t_{i+1}]}(s)$$
for $0 = t_0 < t_1 < \cdots < t_p$ and $H_{(i)} \in L^\infty(\mathscr{F}_{t_i}, \mathbf{P})$. We denote this class by $\mathscr{E}$.

It is straightforward to check that $\mathscr{E} \subseteq \mathscr{P}$; for this, we could even use $\mathbf{1}_{[t_i, t_{i+1})}$. The stricter measurability requirement from using $(t_i, t_{i+1}]$ makes $\mathscr{E}$ a smaller class. We have $\mathscr{E} \subseteq L^2(M)$ for $M \in \mathbb{H}^2$. In fact, $\mathscr{E}$ is dense in $L^2(M)$:

Proposition 5.3. $\forall M \in \mathbb{H}^2$ *$\mathscr{E}$ is dense in $L^2(M)$.*

Proof. This is equivalent to showing that $\mathscr{E}^\perp = \{0\}$. Let $K \perp \mathscr{E}$. Then for $0 \leqslant s < t$ and $F \in L^\infty(\mathscr{F}_s)$, we have
$$0 = \big(K, F \otimes \mathbf{1}_{(s,t]}\big)_{L^2(M)} = \mathbf{E}\Big[F \int_s^t K_u\, \mathrm{d}\langle M, M\rangle_u\Big] = \mathbf{E}\big[F(X_t - X_s)\big],$$
where $X_t := \int_0^t K_u\, \mathrm{d}\langle M, M\rangle_u \in L^1(\Omega, \mathbf{P})$. That is, by Proposition 4.5, $X = K \cdot \langle M, M\rangle$ is a finite-variation process that is also a martingale, whence by Theorem 4.8, $X = 0$. This means almost surely, $K = 0 \;\; \mathrm{d}\langle M, M\rangle$-a.e., i.e., $K = 0$ in $L^2(M)$. ◂

Exercise (due 12/7). Prove that if M is a bounded continuous martingale and A is a bounded increasing process, then

$$\mathbf{E}[M_\infty A_\infty] = \mathbf{E}\Big[\int_0^\infty M_t \,\mathrm{d}A_t\Big].$$

If $M \in \mathbb{H}^2$ and T is a stopping time, then

$$\langle M^T, M^T\rangle_\infty = \langle M, M\rangle_\infty^T = \langle M, M\rangle_T \leqslant \langle M, M\rangle_\infty,$$

so $M^T \in \mathbb{H}^2$.

Exercise (due 12/7). Derive $M^T \in \mathbb{H}^2$ from the optional stopping theorem (Theorem 3.22) instead.

Let $\mathbf{1}_{[0,T]}$ denote the process $(\omega, t) \mapsto \mathbf{1}_{[0,T(\omega)]}(t)$. If T is a stopping time, then $\mathbf{1}_{[0,T]}$ is adapted and left-continuous, so progressive by Proposition 3.4. Therefore, if $H \in L^2(M)$, also $\mathbf{1}_{[0,T]}H \in L^2(M)$.

Here is our first definition of stochastic integral.

Theorem 5.4. *Let $M \in \mathbb{H}^2$. Given an $H \in \mathscr{E}$ as in Definition 5.2, the formula*

$$(H \cdot M)_t := \sum_{i=0}^{p-1} H_{(i)}\big(M_{t_{i+1}\wedge t} - M_{t_i \wedge t}\big)$$

defines a process $H \cdot M \in \mathbb{H}^2$. The map $H \mapsto H \cdot M$ from $\mathscr{E} \to \mathbb{H}^2$ extends uniquely to a linear isometry $L^2(M) \to \mathbb{H}^2$, also denoted $H \mapsto H \cdot M$. For all $H \in L^2(M)$, $H \cdot M$ is the unique element of $\mathbb{H}^2$ such that

$$\forall N \in \mathbb{H}^2 \qquad \langle H \cdot M, N\rangle = H \cdot \langle M, N\rangle. \tag{5.2}$$

(Recall $\big(H \cdot \langle M, N\rangle\big)_t = \int_0^t H_s \,\mathrm{d}\langle M, N\rangle_s$.) If T is a stopping time, then

$$\forall H \in L^2(M) \qquad (\mathbf{1}_{[0,T]}H) \cdot M = (H \cdot M)^T = H \cdot M^T. \tag{5.3}$$

We call $H \cdot M$ the ***stochastic integral*** *of H with respect to M and write*

$$(H \cdot M)_t =: \int_0^t H_s \,\mathrm{d}M_s.$$

Note that the two uses of $\cdot$ in Eq. (5.2) are unambiguous because every finite-variation martingale is 0 by Theorem 4.8.

More abstractly, one could alternatively use Eq. (5.2) to define $H \cdot M$ as follows: Given $M \in \mathbb{H}^2$ and $H \in L^2(M)$, the map

$$\mathbb{H}^2 \ni N \mapsto \mathbf{E}\big[\big(H \cdot \langle M, N\rangle\big)_\infty\big]$$

satisfies

$$\begin{aligned}
\big|\mathbf{E}\big[\big(H\cdot\langle M,N\rangle\big)_\infty\big]\big| &\leqslant \mathbf{E}\Big|\int_0^\infty H_s\,\mathrm{d}\langle M,N\rangle_s\Big| \leqslant \mathbf{E}\Big[\int_0^\infty |H_s|\,|\mathrm{d}\langle M,N\rangle_s|\Big]\\
&\leqslant \mathbf{E}\Big[\Big(\int_0^\infty H_s^2\,\mathrm{d}\langle M,M\rangle_s\Big)^{1/2}\cdot\Big(\int_0^\infty \mathrm{d}\langle N,N\rangle_s\Big)^{1/2}\Big]\\
&\quad \text{[Kunita–Watanabe]}\\
&\leqslant \mathbf{E}\Big[\int_0^\infty H_s^2\,\mathrm{d}\langle M,M\rangle_s\Big]^{1/2}\mathbf{E}\Big[\int_0^\infty \mathrm{d}\langle N,N\rangle_s\Big]^{1/2}\\
&\quad \text{[Cauchy–Schwarz inequality]}\\
&= \|H\|_{L^2(M)}\cdot\|N\|_{\mathbb{H}^2},
\end{aligned} \tag{$*$}$$

and thus is a continuous linear functional on $\mathbb{H}^2$. Hence, there is a unique $H\cdot M\in\mathbb{H}^2$ such that

$$\mathbf{E}\big[\big(H\cdot\langle M,N\rangle\big)_\infty\big] = (H\cdot M,N)_{\mathbb{H}^2} = \mathbf{E}\big[\langle H\cdot M,N\rangle_\infty\big].$$

One can then deduce Eq. (5.2) and everything else.

In fact, we should have verified $H\in L^1\big(\big|\mathrm{d}\langle M,N\rangle\big|\big)$ almost surely and

$$\big(H\cdot\langle M,N\rangle\big)_\infty\in L^1(\mathbf{P}),$$

but this follows as in ($*$), starting instead with

$$\mathbf{E}\Big[\int_0^\infty |H_s|\cdot|\mathrm{d}\langle M,N\rangle_s|\Big].$$

Note that a special case of ($*$), with $H=\mathbf{1}$, is

$$\mathbf{E}\big[|\langle M,N\rangle_\infty|\big]\leqslant\|M\|_{\mathbb{H}^2}\cdot\|N\|_{\mathbb{H}^2}. \tag{$**$}$$

Proof of Theorem 5.4. It is easy to see that the definition of $H\cdot M$ for $H\in\mathscr{E}$ does not depend on its representation as in Definition 5.2. It follows that $H\mapsto H\cdot M$ is linear on $\mathscr{E}$. To see that the map is an isometry into $\mathbb{H}^2$, write

$$M^{(i)} := H_{(i)}(M^{t_{i+1}}-M^{t_i}),$$

so that $H\cdot M=\sum_{i=0}^{p-1}M^{(i)}$. We saw in the proof of Theorem 4.9 that $M^{(i)}$ is a continuous martingale, so $H\cdot M\in\mathbb{H}^2$. By Proposition 4.15(iv), we have

$$\forall s,t\geqslant 0\qquad \langle M^s,M^t\rangle=\langle M,M\rangle^{s\wedge t}.$$

Thus, $\langle M^{(i)},M^{(j)}\rangle=0$ for $i\neq j$ and

$$\langle M^{(i)},M^{(i)}\rangle=H_{(i)}^2\big(\langle M,M\rangle^{t_{i+1}}-\langle M,M\rangle^{t_i}\big).$$

Therefore,

$$\langle H\cdot M,H\cdot M\rangle=\sum_{i=0}^{p-1}H_{(i)}^2\big(\langle M,M\rangle^{t_{i+1}}-\langle M,M\rangle^{t_i}\big)=\int_0^\bullet H_s^2\,\mathrm{d}\langle M,M\rangle_s$$

(i.e., $\langle H \cdot M, H \cdot M \rangle_t = \int_0^t H_s^2 \,\mathrm{d}\langle M, M \rangle_s$). In particular,

$$\|H \cdot M\|_{\mathbb{H}^2}^2 = \mathbf{E}\big[\langle H \cdot M, H \cdot M \rangle_\infty\big] = \mathbf{E}\Big[\int_0^\infty H_s^2 \,\mathrm{d}\langle M, M \rangle_s\Big] = \|H\|_{L^2(M)}^2,$$

as desired.

By Proposition 5.3, $\mathscr{E}$ is dense in $L^2(M)$, so $H \mapsto H \cdot M$ has a unique continuous extension to a map from $L^2(M)$ to $\mathbb{H}^2$.

If $H \in \mathscr{E}$, we have, similarly to above,

$$\begin{aligned} \langle H \cdot M, N \rangle &= \sum_{i=1}^{p-1} \langle M^{(i)}, N \rangle = \sum_{i=0}^{p-1} H_{(i)} \big(\langle M, N \rangle^{t_{i+1}} - \langle M, N \rangle^{t_i} \big) \\ &= \int_0^\bullet H_s \,\mathrm{d}\langle M, N \rangle_s = H \cdot \langle M, N \rangle, \end{aligned}$$

i.e., Eq. (5.2) holds for $H \in \mathscr{E}$. Now, Eq. ($*$) shows that $H \mapsto \big(H \cdot \langle M, N \rangle\big)_\infty$ is continuous as a map from $\mathbb{H}^2$ to $L^1(\mathbf{P})$ and Eq. ($**$) that $X \mapsto \langle X, N \rangle_\infty$ is continuous from $\mathbb{H}^2$ to $L^1(\mathbf{P})$. Since $H \mapsto H \cdot M$ is an isometry from $L^2(M) \to \mathbb{H}^2$, it follows that $H \mapsto \langle H \cdot M, N \rangle_\infty$ is continuous from $L^2(M) \to L^1(\mathbf{P})$. Therefore,

$$\langle H \cdot M, N \rangle_\infty = \big(H \cdot \langle M, N \rangle\big)_\infty$$

for all $H \in L^2(M)$, $N \in \mathbb{H}^2$. Replace N by N^t to obtain Eq. (5.2) in general.

We've already seen that something weaker than Eq. (5.2) characterizes $H \cdot M$ among elements of $\mathbb{H}^2$.

To see Eq. (5.3), let $N \in \mathbb{H}^2$ and note that

$$\big\langle (H \cdot M)^T, N \big\rangle \underset{\text{[Proposition 4.15(iv)]}}{=} \langle H \cdot M, N \rangle^T \overset{\text{[Eq. (5.2)]}}{=} \big(H \cdot \langle M, N \rangle\big)^T \underset{\text{[deterministic]}}{=} \mathbf{1}_{[0,T]} H \cdot \langle M, N \rangle \overset{\text{[Eq. (5.2)]}}{=} \big\langle (\mathbf{1}_{[0,T]} H) \cdot M, N \big\rangle,$$

whence by the uniqueness of Eq. (5.2), $\big(\mathbf{1}_{[0,T]} H\big) \cdot M = (H \cdot M)^T$. Similarly,

$$\langle H \cdot M^T, N \rangle \underset{\text{[Eq. (5.2)]}}{=} H \cdot \langle M^T, N \rangle = H \cdot \langle M, N \rangle^T = \mathbf{1}_{[0,T]} H \cdot \langle M, N \rangle,$$

so $(H \cdot M)^T = H \cdot M^T$. ◂

We could rewrite Eq. (5.2) as

$$\Big\langle \int_0^\bullet H_s \,\mathrm{d}M_s, N \Big\rangle_t = \int_0^t H_s \,\mathrm{d}\langle M, N \rangle_s.$$

If $M \in \mathbb{H}^2$ and $H \in L^2(M)$, then by Eq. (5.2),

$$\langle H \cdot M, H \cdot M \rangle = H \cdot \langle M, H \cdot M \rangle = H \cdot \langle H \cdot M, M \rangle = H^2 \cdot \langle M, M \rangle. \tag{5.4}$$

If also $N \in \mathbb{H}^2$ and $K \in L^2(N)$, then similarly we obtain

$$\langle H \cdot M, K \cdot N \rangle = HK \cdot \langle M, N \rangle.$$

Proposition 5.5. *Let $M \in \mathbb{H}^2$, $H \in L^2(M)$, and K be progressive. Then*
$$KH \in L^2(M) \quad \Longleftrightarrow \quad K \in L^2(H \cdot M),$$
in which case
$$(KH) \cdot M = K \cdot (H \cdot M).$$

Proof. By Eq. (5.4), we have
$$\mathbf{E}\big[\big(K^2H^2 \cdot \langle M, M\rangle\big)_\infty\big] = \mathbf{E}\big[\big(K^2 \cdot \langle H \cdot M, H \cdot M\rangle\big)_\infty\big],$$
which gives $\|KH\|_{L^2(M)} = \|K\|_{L^2(H \cdot M)}$. If this is finite, then for $N \in \mathbb{H}^2$, we have
$$\begin{aligned}\big\langle (KH) \cdot M, N\big\rangle &\underset{\text{[Eq. (5.2)]}}{=} KH \cdot \langle M, N\rangle \underset{\text{[see below]}}{=} K \cdot \big(H \cdot \langle M, N\rangle\big) \\ &\underset{\text{[Eq. (5.2)]}}{=} K \cdot \langle H \cdot M, N\rangle \underset{\text{[Eq. (5.2)]}}{=} \big\langle K \cdot (H \cdot M), N\big\rangle,\end{aligned}$$
where the second equality is justified as follows: by the Kunita–Watanabe inequality,
$$\int_0^\infty H^2 \big|\mathrm{d}\langle M, M\rangle\big| < \infty \quad \text{and} \quad \int_0^\infty K^2H^2 \big|\mathrm{d}\langle M, M\rangle\big| < \infty$$
implies
$$\forall t \qquad \int_0^t |H_s| \big|\mathrm{d}\langle M, N\rangle_s\big| < \infty \quad \text{and} \quad \int_0^t |K_sH_s| \big|\mathrm{d}\langle M, N\rangle_s\big| < \infty.$$
By the uniqueness part of Eq. (5.2), we conclude that $(KH) \cdot M = K \cdot (H \cdot M)$. ◀

Recall that for $M, N \in \mathbb{H}^2$, $(M, N)_{\mathbb{H}^2} = \mathbf{E}[M_\infty N_\infty] = \mathbf{E}[\langle M, N\rangle_\infty]$. By considering M^t and N^t, this implies that $\mathbf{E}[M_tN_t] = \mathbf{E}[\langle M, N\rangle_t]$ for $t \in [0, \infty]$.

Suppose that $M, N \in \mathbb{H}^2$, $H \in L^2(M)$, and $K \in L^2(N)$. Since $H \cdot M$, $K \cdot N \in \mathbb{H}^2$, we get
$$\forall t \in [0, \infty] \quad \mathbf{E}\Big[\int_0^t H_s \, \mathrm{d}M_s\Big] = \mathbf{E}\big[(H \cdot M)_t\big] \underset{\text{[martingale]}}{=} \mathbf{E}\big[(H \cdot M)_0\big] = 0 \tag{5.6}$$
and
$$\begin{aligned}\mathbf{E}\Big[\int_0^t H_s \, \mathrm{d}M_s \cdot \int_0^t K_s \, \mathrm{d}N_s\Big] &= \mathbf{E}\big[(H \cdot M)_t (K \cdot N)_t\big] \\ &= \mathbf{E}\big[\langle H \cdot M, K \cdot N\rangle_t\big] \\ &= \mathbf{E}\big[\big(HK \cdot \langle M, N\rangle\big)_t\big] \\ &= \mathbf{E}\Big[\int_0^t H_sK_s \, \mathrm{d}\langle M, N\rangle_s\Big].\end{aligned}$$
In particular,
$$\mathbf{E}\Big[\Big(\int_0^t H_s \, \mathrm{d}M_s\Big)^2\Big] = \mathbf{E}\Big[\int_0^t H_s^2 \, \mathrm{d}\langle M, M\rangle_s\Big]; \tag{5.8}$$
this equality of norms is referred to as the ***Itô isometry***.

Note that we have defined $\int_0^t H_s \, \mathrm{d}B_s$ for progressive H with $\int_0^\infty \mathbf{E}[H_s^2] \, \mathrm{d}s < \infty$ by stopping B at t. If H is deterministic, this agrees with the Wiener integral almost surely: check first for step functions. Thus, when H is deterministic, $H \cdot B$ is a Wiener-integral process.

Exercise (due 1/18). Exercise 5.25; assume $\sup_{t,\omega}\bigl|H_t(\omega)\bigr| < \infty$.

We may rewrite the martingale condition for $H \cdot M$ as follows:

$$0 \leqslant s < t \leqslant \infty \quad \Longrightarrow \quad \mathbf{E}\Bigl[\int_0^t H_r \,\mathrm{d}M_r \Bigm| \mathscr{F}_s\Bigr] = \int_0^s H_r \,\mathrm{d}M_r, \tag{5.9}$$

or, with $\int_s^t H_r \,\mathrm{d}M_r := \int_0^t H_r \,\mathrm{d}M_r - \int_0^s H_r \,\mathrm{d}M_r$,

$$\mathbf{E}\Bigl[\int_s^t H_r \,\mathrm{d}M_r \Bigm| \mathscr{F}_s\Bigr] = 0.$$

5.1.2. Stochastic Integrals for Local Martingales

The stopping time identities Eq. (5.3) will allow us to extend stochastic integrals to continuous local martingales. If M is a continuous local martingale, we again write $L^2(M)$ for the set of progressive processes H in $L^2\bigl(\langle M, M\rangle\,\mathbf{P}\bigr)$. We write $L^2_{\mathrm{loc}}(M)$ for the set of the progressive H such that

$$\text{a.s.} \quad \forall t \geqslant 0 \qquad \int_0^t H_s^2 \,\mathrm{d}\langle M, M\rangle_s < \infty.$$

Theorem 5.6. *Let M be a continuous local martingale. If $H \in L^2_{\mathrm{loc}}(M)$, then there exists a unique continuous local martingale with initial value 0, denoted $H \cdot M$, such that for all continuous local martingales N,*

$$\langle H \cdot M, N\rangle = H \cdot \langle M, N\rangle. \tag{5.10}$$

If T is a stopping time, then for all $H \in L^2_{\mathrm{loc}}(M)$,

$$\bigl(\mathbf{1}_{[0,T]} H \cdot M\bigr) = (H \cdot M)^T = H \cdot M^T. \tag{5.11}$$

If $H \in L^2_{\mathrm{loc}}(M)$ and K is progressive, then $K \in L^2_{\mathrm{loc}}(H \cdot M)$ if and only if $HK \in L^2_{\mathrm{loc}}(M)$, in which case

$$K \cdot (H \cdot M) = (KH) \cdot M. \tag{5.12}$$

If $M \in \mathbb{H}^2$ and $H \in L^2(M)$, then this definition of $H \cdot M$ agrees with that of Theorem 5.4.

Proof. Since $\langle M - M_0, N\rangle = \langle M, N\rangle$ for every continuous local martingale N, we may set $H \cdot M := H \cdot (M - M_0)$ (to be defined) and assume that $M_0 = 0$. Also, we may take H to be 0 on the negligible set where for some $t \geqslant 0$, $\int_0^t H_s^2 \,\mathrm{d}\langle M, M\rangle_s = \infty$.

The idea is to localize and put together the resulting definitions.

For $n \geqslant 1$, let

$$T_n := \inf\Bigl\{t \geqslant 0\,;\ \int_0^t (1 + H_s^2)\,\mathrm{d}\langle M, M\rangle_s \geqslant n\Bigr\}.$$

This gives a sequence of stopping times that increase to infinity. Since

$$\forall t \geqslant 0 \quad \langle M^{T_n}, M^{T_n}\rangle_t = \underbrace{\langle M, M\rangle_{t\wedge T_n}}_{\text{[Proposition 4.11]}} \leqslant n,$$

Theorem 4.13 tells us that $M^{T_n} \in \mathbb{H}^2$. Also,

$$\int_0^\infty H_s^2 \,\mathrm{d}\langle M^{T_n}, M^{T_n}\rangle_s = \int_0^{T_n} H_s^2 \,\mathrm{d}\langle M, M\rangle_s \leqslant n,$$

so $H \in L^2(M^{T_n})$. Therefore, Theorem 5.4 defines $H \cdot M^{T_n}$. These are consistent: if $m > n$, then

$$(H \cdot M^{T_m})^{T_n} = H \cdot (M^{T_m})^{T_n} = H \cdot M^{T_n}.$$

[Eq. (5.3)]

Thus, there exists a unique process, $H \cdot M$, such that $\forall n \;\; (H \cdot M)^{T_n} = H \cdot M^{T_n}$. Since $H \cdot M^{T_n}$ has continuous sample paths, so does $H \cdot M$. Since $(H \cdot M)_t = \lim_{n\to\infty} (H \cdot M^{T_n})_t$, we get that $H \cdot M$ is adapted. Since $(H \cdot M)^{T_n}$ is a martingale (in $\mathbb{H}^2$, even), we get that $H \cdot M$ is a continuous local martingale.

Now we verify the properties (5.10)–(5.12).

To prove Eq. (5.10), we may assume that $N_0 = 0$. For $n \geqslant 1$, write

$$T_n' = \inf\big\{t \geqslant 0\,;\; |N_t| \geqslant n\big\}, \quad S_n := T_n \wedge T_n'.$$

As before, $N^{T_n'} \in \mathbb{H}^2$, so

$$\langle H \cdot M, N\rangle^{S_n} = \big(\langle H \cdot M, N\rangle^{T_n}\big)^{T_n'} = \big\langle (H \cdot M)^{T_n}, N\big\rangle^{T_n'} = \big\langle (H \cdot M)^{T_n}, N^{T_n'}\big\rangle$$

[Proposition 4.15(iv)]

$$= \langle H \cdot M^{T_n}, N^{T_n'}\rangle = H \cdot \langle M^{T_n}, N^{T_n'}\rangle = H \cdot \langle M, N\rangle^{S_n}$$

[definition] [Eq. (5.2)] [Proposition 4.15(iv)]

$$= \big(H \cdot \langle M, N\rangle\big)^{S_n}.$$

[deterministic]

Since $S_n \to \infty$, this gives $\langle H \cdot M, N\rangle = H \cdot \langle M, N\rangle$, as desired. If X is also a continuous local martingale with $X_0 = 0$ and $\langle X, N\rangle = H \cdot \langle M, N\rangle$ for all continuous local martingales N, then $\langle H \cdot M - X, N\rangle = 0$, so choosing $N := H \cdot M - X$, we get $X = H \cdot M$ from Proposition 4.12.

The proof of Eq. (5.11) is like that of Eq. (5.3), and proof of Eq. (5.12) is like that of Proposition 5.5.

If $M \in \mathbb{H}^2$ and $H \in L^2(M)$, then $\langle H \cdot M, H \cdot M\rangle = H \cdot \langle M, H \cdot M\rangle = H^2 \cdot \langle M, M\rangle$ by two uses of Eq. (5.10). This shows that $H \cdot M \in \mathbb{H}^2$, so the characteristic property Eq. (5.12) (which holds by Eq. (5.10)) shows that the definitions agree. ◀

We again write

$$\int_0^t H_s \,\mathrm{d}M_s := (H \cdot M)_t.$$

We can then rewrite Eq. (5.10) as

$$\Big\langle \int_0^\bullet H_s \,\mathrm{d}M_s, N_\bullet \Big\rangle_t = \int_0^t H_s \,\mathrm{d}\langle M, N\rangle_s.$$

If $H \in L^2_{\mathrm{loc}}(M),\ 0 \leqslant t \leqslant \infty$, and $\mathbf{E}\big[\int_0^t H_s^2 \,\mathrm{d}\langle M, M\rangle_s\big] < \infty$, then $(H \cdot M)^t \in \mathbb{H}^2$ by Theorem 4.13, so we have the analogues of Eqs. (5.6), (5.8) and (5.9):

$$\mathbf{E}\Big[\int_0^t H_s \,\mathrm{d}M_s\Big] = 0,$$
$$\mathbf{E}\Big[\Big(\int_0^t H_s \,\mathrm{d}M_s\Big)^2\Big] = \mathbf{E}\Big[\int_0^t H_s^2 \,\mathrm{d}\langle M, M\rangle_s\Big].$$

In particular, if $H \in L^2(M)$ (the case $t = \infty$), then $H \cdot M \in \mathbb{H}^2$ (even though M need not be in $\mathbb{H}^2$).

Exercise (due 1/18). Give an example of a continuous local martingale M and a process $H \in L^2_{\mathrm{loc}}(M)$ such that

$$\mathbf{E}\Big[\int_0^1 H_s \,\mathrm{d}M_s\Big] \neq 0 \quad \text{and} \quad \mathbf{E}\Big[\Big(\int_0^1 H_s \,\mathrm{d}M_s\Big)^2\Big] \neq \mathbf{E}\Big[\int_0^1 H_s^2 \,\mathrm{d}\langle M, M\rangle_s\Big].$$

Hint: use M that is not a true martingale.

Exercise (due 1/18). Exercise 5.25 (in general).

5.1.3. Stochastic Integrals for Semimartingales

We call a progressive process H ***locally bounded*** if

$$\text{a.s.} \quad \forall t \geqslant 0 \qquad \sup_{0 \leqslant s \leqslant t} |H_s| < \infty.$$

This is equivalent to the existence of stopping times $T_n \uparrow \infty$ such that $\mathbf{1}_{[0,T_n]} H$ is bounded, i.e., that there exists a negligible set $\mathscr{N}$ such that

$$\sup_{\omega \notin \mathscr{N},\, t \geqslant 0} \big|\mathbf{1}_{[0,T_n(\omega)]}(t) H_t(\omega)\big| < \infty.$$

Note that if H is adapted and continuous, then H is locally bounded.

The assumption that H is locally bounded is convenient, because then for each finite-variation process V,

$$\text{a.s.} \quad t \geqslant 0 \qquad \int_0^t |H_s|\, |\mathrm{d}V_s| < \infty,$$

i.e., $H \in L^1_{\mathrm{loc}}\big(|\mathrm{d}V|\big)$, and for each continuous local martingale, M, we have $H \in L^2_{\mathrm{loc}}(M)$.

Definition 5.7. *Let $X = M + V$ be the canonical decomposition of a continuous semimartingale, X, and H be locally bounded. We define the* ***stochastic integral*** *$H \cdot X$ to be the continuous semimartingale with canonical decomposition*

$$H \cdot X := H \cdot M + H \cdot V.$$

We also write $\int_0^t H_s \,\mathrm{d}X_s := (H \cdot X)_t$.

Remark. We could have done the same as long as $H \in L^2_{\mathrm{loc}}(M) \cap L^1_{\mathrm{loc}}\big(|\mathrm{d}V|\big)$.

The following properties are evident:

(i) $(H, X) \mapsto H \cdot X$ is bilinear.

(ii) If H and K are locally bounded, then $H \cdot (K \cdot X) = (HK) \cdot X$. Rewritten: if $Y_t = \int_0^t K_s \, \mathrm{d}X_s$, then $\int_0^t H_s \, \mathrm{d}Y_s = \int_0^t H_s K_s \, \mathrm{d}X_s$.

(iii) For all stopping times T, $(H \cdot X)^T = (\mathbf{1}_{[0,T]} H) \cdot X = H \cdot X^T$.

(iv) If X is a continuous local martingale, then so is $H \cdot X$; if X is a finite-variation process, then so is $H \cdot X$.

The next property is less evident:

(v) If $H_s(\omega) = \sum_{i=0}^{p-1} H_{(i)}(\omega) \mathbf{1}_{(t_i, t_{i+1}]}(s)$, $0 = t_0 < t_1 < \cdots < t_p$, and $H_{(i)} \in \mathscr{F}_{t_i}$ is locally bounded, then

$$(H \cdot X)_t = \sum_{i=1}^{p-1} H_{(i)} (X_{t_{i+1} \wedge t} - X_{t_i \wedge t}).$$

This is clear if $M = 0$, so it suffices to prove it when $V = 0$. If H is bounded and $M \in \mathbb{H}^2$, then this is the definition of $H \cdot M$. In general, we may assume $M_0 = 0$; let

$$T_n := \inf\{t \geqslant 0\,;\ |H_t| \geqslant n\} = \min\{t_i\,;\ |H_{(i)}| \geqslant n\}$$

and

$$S_n := \inf\{t \geqslant 0\,;\ \langle M, M\rangle_s \geqslant n\}.$$

Note that

$$\mathbf{1}_{[0,T_n]}(s) H_s = \sum_{i=0}^{p-1} H_i^{(n)} \mathbf{1}_{(t_i, t_{i+1}]}(s),$$

where $H_{(i)}^{(n)} := \mathbf{1}_{[T_n > t_i]} H_{(i)} \in \mathscr{F}_{t_i}$. Therefore,

$$(H \cdot M)_{t \wedge T_n \wedge S_n} \underset{\text{[Eq. (5.11)]}}{=} (\mathbf{1}_{[0,T_n]} H \cdot M^{S_n})_t \underset{\text{[definition]}}{=} \sum_{i=0}^{p-1} H_{(i)}^{(n)} (M_{t_{i+1} \wedge t}^{S_n} - M_{t_i \wedge t}^{S_n}).$$

Now let $n \to \infty$.

Exercise. Show that if $Z \in \mathscr{F}_0$ and X is a continuous semimartingale, then $H \cdot X = ZX$, where $H_t := Z$ for all t.

5.1.4. Convergence of Stochastic Integrals

Proposition 5.8 (Dominated Convergence Theorem). *Let $X = M+V$ be the canonical decomposition of a continuous semimartingale and $t > 0$. Suppose that $H, H^{(1)}, H^{(2)}, \ldots$ are locally bounded, progressive processes and that K is a nonnegative, progressive process such that almost surely,*

(i) $\forall s \in [0,t]\ \ \lim_{n\to\infty} H_s^{(n)} = H_s$,

(ii) $\forall s \in [0,t]\ \ \forall n \geqslant 1\ \ |H_s^{(n)}| \leqslant K_s$, *and*

(iii) $\int_0^t (K_s)^2\,\mathrm{d}\langle M,M\rangle_s < \infty$ *and* $\int_0^t K_s\,|\mathrm{d}V_s| < \infty$.

Then

$$\lim_{n\to\infty} \int_0^t H_s^{(n)}\,\mathrm{d}X_s = \int_0^t H_s\,\mathrm{d}X_s \quad \textit{in probability}.$$

That H and $H^{(n)}$ be locally bounded can be weakened. In (i) and (ii), we can weaken "$\forall s \in [0,t]$" to "$\forall s \in [0,t]$ outside a set of $\mathrm{d}\langle M,M\rangle$-measure 0 and of $|\mathrm{d}V|$-measure 0"; this will be clear from the proof. Part (iii) is automatic if K is locally bounded.

Proof. The Lebesgue dominated convergence theorem gives $\int_0^t H_s^{(n)}\,\mathrm{d}V_s \to \int_0^t H_s\,\mathrm{d}V_s$ where (i)–(iii) hold, hence almost surely. It remains to show that

$$\int_0^t H_s^{(n)}\,\mathrm{d}M_s \xrightarrow{\mathbf{P}} \int_0^t H_s\,\mathrm{d}M_s.$$

For $p \geqslant 1$, let

$$T_p := \inf\Big\{r \in [0,t]\,;\ \int_0^r (K_s)^2\,\mathrm{d}\langle M,M\rangle_s \geqslant p\Big\} \wedge t.$$

Then almost surely, by (iii), for all large p, $T_p = t$. Now,

$$\mathbf{E}\bigg[\int_0^{T_p} (H_s^{(n)} - H_s)^2\,\mathrm{d}\langle M,M\rangle_s\bigg] \leqslant \mathbf{E}\bigg[\int_0^{T_p} (2K_s)^2\,\mathrm{d}\langle M,M\rangle_s\bigg] \leqslant 4p < \infty,$$

whence

$$\mathbf{E}\Big[\Big(((H^{(n)} - H)\cdot M)_{T_p}\Big)^2\Big] = \mathbf{E}\bigg[\int_0^{T_p} (H_s^{(n)} - H_s)^2\,\mathrm{d}\langle M,M\rangle_s\bigg] \to 0$$

as $n \to \infty$ by Lebesgue's dominated convergence theorem applied to $\langle M,M\rangle^{T_p}\,\mathbf{P}$. Because $\mathbf{P}[T_p = t] \to 1$ as $p \to \infty$, we get the result. ◀

We can deduce the following Riemann-integral type of result:

Proposition 5.9. *Let X be a continuous semimartingale and H be a continuous adapted process. If $t > 0$ and $0 = t_0^n < t_1^n < \cdots < t_{p_n}^n = t$ is* any *sequence of subdivisions with mesh going to* 0, *then*

$$\lim_{n\to\infty} \sum_{i=0}^{p_n-1} H_{t_i^n}(X_{t_{i+1}^n} - X_{t_i^n}) = \int_0^t H_s\,\mathrm{d}X_s \quad \textit{in probability}.$$

Proof. Note that the sum on the left-hand side equals $\int_0^t H_s^{(n)}\,\mathrm{d}X_s$ for the step function progressive process

$$H_s^{(n)} := \sum_{i=0}^{p_n-1} H_{t_i^n}\mathbf{1}_{(t_i^n,t_{i+1}^n]}(s) + H_0\mathbf{1}_{\{0\}}(s)$$

by property (v) in Section 5.1.3. Since $K_s := \max_{0\leqslant r\leqslant s}|H_r|$ is locally bounded, the result follows. ◂

It is crucial that the Riemann sum used the left-hand endpoints. For example, let $H = X$. Then

$$\sum_{i=0}^{p_n-1} X_{t_{i+1}^n}\big(X_{t_{i+1}^n} - X_{t_i^n}\big) = \underset{\substack{\Big\downarrow\ \text{Proposition 5.9} \\ \int_0^t X_s\,\mathrm{d}X_s}}{\sum_{i=0}^{p_n-1} X_{t_i^n}\big(X_{t_{i+1}^n} - X_{t_i^n}\big)} + \underset{\substack{\Big\downarrow\ \text{Proposition 4.21 if subdivisions are increasing} \\ \langle X,X\rangle_t}}{\sum_{i=0}^{p_n-1}\big(X_{t_{i+1}^n} - X_{t_i^n}\big)^2}.$$

Thus, we get a different limit when we use the right-hand endpoints unless the martingale part of X is constant on $[0,t]$. On the other hand, this calculation is useful: if we add to it that of Proposition 5.9, we get

$$(X_t)^2 - (X_0)^2 = 2\int_0^t X_s\,\mathrm{d}X_s + \langle X,X\rangle_t.$$

This can also be derived from Itô's formula (in the next section).

Exercise (due 1/25). Show that if X is a continuous semimartingale, $t > 0$, and $0 = t_0^n < \cdots < t_{p_n}^n = t$ is *any* sequence of subdivisions of $[0,t]$ with mesh going to 0, then

$$\lim_{n\to\infty}\sum_{i=1}^{p_n}\big(X_{t_i^n} - X_{t_{i-1}^n}\big)^2 = \langle X,X\rangle_t \quad \text{in probability.}$$

5.2. Itô's Formula

This is analogous to the fundamental theorem of calculus: in order to calculate an integral, it helps to know how to differentiate. However, there is no stochastic derivative. The formula also shows that the class of continuous semimartingales is closed under compositions with C^2 functions.

Theorem 5.10 (Itô's Formula)**.** *Let X be a continuous semimartingale and $F \in C^2(\mathbb{R})$ (i.e., twice continuously differentiable). Then*

$$\forall t \geqslant 0 \quad \underset{\substack{\uparrow \\ \textit{[antiderivative]}}}{F(X_t)} = \underset{\substack{\uparrow \\ \textit{[value at 0]}}}{F(X_0)} + \int_0^t \underset{\substack{\uparrow \\ \textit{[integrand]}}}{F'(X_s)}\,\mathrm{d}X_s + \frac{1}{2}\int_0^t \underset{\substack{\uparrow \\ \textit{[derivative]}}}{F''(X_s)}\,\mathrm{d}\langle X,X\rangle_s.$$

We may also write this as

$$F(X) = F(X_0) + \underset{\substack{\uparrow \\ \text{[CLM+FV]}}}{F'(X) \cdot X} + \underset{\substack{\uparrow \\ \text{[FV]}}}{\tfrac{1}{2}F''(X) \cdot \langle X, X\rangle},$$

showing that $F(X)$ is a continuous semimartingale and giving its canonical decomposition. More generally, if $X^1, \ldots, X^p$ are continuous semimartingales and $F \in C^2(\mathbb{R}^p)$, then

$$\forall t \geqslant 0 \quad F(X_t^1, \ldots, X_t^p) = F(X_0^1, \ldots, X_0^p) + \sum_{i=1}^{p} \int_0^t F_i(X_s^1, \ldots, X_s^p)\, \mathrm{d}X_s^i + \frac{1}{2}\sum_{i,j=1}^{p} \int_0^t F_{ij}(X_s^1, \ldots, X_s^p)\, \mathrm{d}\langle X^i, X^j\rangle_s,$$

which we may write for $X = (X^1, \ldots, X^p)$ as

$$F(X) = F(X_0) + \nabla F(X) \cdot X + \frac{1}{2}\Big\langle X, \big(\nabla^2 F(X)\big) \cdot X\Big\rangle,$$

where

$$H \cdot X := \left(\sum_{k=1}^{p} H_{jk} \cdot X^k\right)_{j=1}^{q} \quad \text{and} \quad \langle X, Y\rangle := \sum_{j,k=1}^{p} \langle X^j, Y^k\rangle$$

when H is a $(q \times p)$-matrix-valued process and Y is a p-dimensional process.

Lemma. *If $Y_n \xrightarrow{\mathbf{P}} 0$ and $Z_n \xrightarrow{\mathbf{P}} Z$, then $Y_n Z_n \xrightarrow{\mathbf{P}} 0$.*

This lemma is a special case of Slutsky's theorem, Exercise 25.7 in Billingsley's book, *Probability and Measure*.

Proof. Let $\varepsilon > 0$. Choose K such that

$$\mathbf{P}\big[|Z| > K\big] < \varepsilon.$$

Choose N such that

$$\mathbf{P}\big[|Z_n - Z| > 1\big] < \varepsilon \text{ and } \mathbf{P}\big[|Y_n| > \varepsilon/(K+1)\big] < \varepsilon$$

for $n \geqslant N$. Then for $n \geqslant N$,

$$\begin{aligned}\mathbf{P}\big[|Y_n Z_n| > \varepsilon\big] &\leqslant \mathbf{P}\big[|Y_n| > \tfrac{\varepsilon}{K+1}\big] + \mathbf{P}\big[|Z_n| > K+1\big] \\ &< \varepsilon + \mathbf{P}\big[|Z| > K\big] + \mathbf{P}\big[|Z_n - Z| > 1\big] < 3\varepsilon.\end{aligned}$$ ◀

Lemma. *If X is a continuous semimartingale and $0 = t_0^n < \cdots < t_{p_n}^n = t$ form an increasing sequence of subdivisions of $[0,t]$ whose mesh goes to zero, then there exists a subsequence $(n_k)_{k\geqslant 1}$ such that almost surely,*

$$\sum_{i=0}^{p_{n_k}-1} (X_{t_{i+1}^{n_k}} - X_{t_i^{n_k}})^2 \delta_{t_i^{n_k}} \Rightarrow \mathbf{1}_{[0,t]}\, \mathrm{d}\langle X, X\rangle$$

as $k \to \infty$.

Proof. Write $\mu_n := \sum_{i=0}^{p_n-1} (X_{t_{i+1}^n} - X_{t_i^n})^2 \delta_{t_i^n}$. Let $D := \{t_i^n\,;\ n \geqslant 1, 0 \leqslant i \leqslant p_n\}$. We have by Proposition 4.21 that

$$\mu_n\big([0,r]\big) \xrightarrow{\mathbf{P}} \langle X, X\rangle_r$$

as $n \to \infty$ for each $r \in D$. Choose (n_k) such that this converges almost surely for all $r \in D$. Then it also converges for all $r \in [0,t]$. ◀

Proof of Theorem 5.10. Suppose first $p = 1$. Let $(t_i^n)_{i=0}^{p_n}$ be an increasing sequence of subdivisions of $[0,t]$ with mesh going to 0. Then

$$\begin{aligned}
F(X_t) &= F(X_0) + \sum_{i=0}^{p_n-1} \big(F(X_{t_{i+1}^n}) - F(X_{t_i^n})\big) \\
&= F(X_0) + \sum_{i=0}^{p_n-1} \Big(F'(X_{t_i^n})(X_{t_{i+1}^n} - X_{t_i^n}) + \frac{1}{2}F''(\xi_{n,i})(X_{t_{i+1}^n} - X_{t_i^n})^2\Big)
\end{aligned}$$

for some $\xi_{n,i}$ between $X_{t_i^n}$ and $X_{t_{i+1}^n}$. By Proposition 5.9,

$$\sum_{i=0}^{p_n-1} F'(X_{t_i^n})(X_{t_{i+1}^n} - X_{t_i^n}) \xrightarrow{\mathbf{P}} \int_0^t F'(X_s)\,\mathrm{d}X_s.$$

Since $\max_i \big|F''(\xi_{n,i}) - F''(X_{t_i^n})\big| \to 0$ as $n \to \infty$ (because $F'' \circ X$ is uniformly continuous on $[0,t]$ and the mesh goes to 0), the first lemma in combination with Proposition 4.21 shows that it suffices to prove that

$$\sum_{i=0}^{p_n-1} F''(X_{t_i^n})(X_{t_{i+1}^n} - X_{t_i^n})^2 \xrightarrow{\mathbf{P}} \int_0^t F''(X_s)\,\mathrm{d}\langle X, X\rangle_s.$$

In fact, we prove this holds almost surely along a subsequence [which could be taken to be a subsequence of any given sequence, so the claim of the display does hold]. Note that the left-hand side equals

$$\int_0^t F''(X_s)\,\mathrm{d}\mu_n(s),$$

where μ_n is as in the proof of the second lemma. Since $F'' \circ X$ is continuous on $[0,t]$, the result follows from the lemma (with weak convergence applied to $F'' \circ X$).

For $p > 1$, we consider F on the broken line from X_0 to X_t that is linear between $X_{t_i^n}$ and $X_{t_{i+1}^n}$. We may again choose $\xi_{n,i}$ on that broken line between $X_{t_i^n}$ and $X_{t_{i+1}^n}$ to write

$$\begin{aligned}
F(X_t) = F(X_0) &+ \sum_{i=0}^{p_n-1} \nabla F(X_{t_i^n}) \overset{\text{[dot product]}}{\cdot} (X_{t_{i+1}^n} - X_{t_i^n}) \\
&+ \frac{1}{2}\sum_{i=0}^{p_n-1} (X_{t_{i+1}^n} - X_{t_i^n}) \underset{\text{[dot product]}}{\cdot} \nabla^2 F(\xi_{n,i})(X_{t_{i+1}^n} - X_{t_i^n}).
\end{aligned}$$

Proposition 5.9 shows that the first sum converges to $\int_0^t \nabla F(X_s) \cdot \mathrm{d}X_s$ in probability. We can apply the second lemma to X^j, X^ℓ and $X^j + X^\ell$ to get a subsequence (n_k) such that almost surely,

$$\sum_{i=0}^{p_{n_k}-1} (X^j_{t^{n_k}_{i+1}} - X^j_{t^{n_k}_i})(X^l_{t^{n_k}_{i+1}} - X^l_{t^{n_k}_i})\delta_{t^{n_k}_i} \Rightarrow \mathbf{1}_{[0,t]}\mathrm{d}\langle X^j, X^\ell\rangle,$$

and this completes the proof. ◀

Exercise (due 1/25). Exercise 5.26.

If we use $F(x,y) := xy$, then we get a formula for ***integration by parts***:

$$X_tY_t = X_0Y_0 + \int_0^t X_s\,\mathrm{d}Y_s + \int_0^t Y_s\,\mathrm{d}X_s + \langle X, Y\rangle_t.$$

If we use $Y = X$, then

$$X_t^2 = X_0^2 + 2\int_0^t X_s\,\mathrm{d}X_s + \langle X, X\rangle_t. \tag{$*$}$$

When X is a continuous local martingale, we get the formula promised during the proof of Theorem 4.9 and seen at the end of Section 5.1.3:

$$X^2 - \langle X, X\rangle = X_0^2 + 2\int_0^\bullet X_s\,\mathrm{d}X_s.$$

Also, Eq. ($*$) implies the integration by parts formula by applying Eq. ($*$) to X, Y, and $X+Y$. In fact, we can prove Itô's formula from integration by parts (and therefore from Eq. ($*$)):

Exercise (due 1/25).
(1) Use integration by parts to show that if Itô's formula holds for some $F \in C^2(\mathbb{R}^p)$, then it also holds for all G of the form

$$G(x_1, \dots, x_p) = x_i F(x_1, \dots, x_p).$$

(2) Deduce that Itô's formula holds for all polynomials, F.
(3) Show that if $K \subseteq \mathbb{R}^p$ is compact, then for each $F \in C^2(\mathbb{R}^p)$, there exist polynomials P_n such that

$$\lim_{n\to\infty}\left(\|F - P_n\|_{C(K)} + \sum_{i=1}^p \|F_i - (P_n)_i\|_{C(K)} + \sum_{i,j=1}^p \|F_{ij} - (P_n)_{i,j}\|_{C(K)}\right) = 0.$$

(4) Deduce that if X takes values only in K, then Itô's formula holds for all $F \in C^2(\mathbb{R}^p)$.
(5) By using stopping times, deduce the full Itô's formula.

Exercise. Show that if X and Y are continuous semimartingales, then $\langle XY, XY\rangle = X^2 \cdot \langle Y, Y\rangle + 2(XY)\cdot\langle Y, X\rangle + Y^2 \cdot \langle X, X\rangle$.

If B is a d-dimensional Brownian motion, then the components of $B - B_0$ are independent, whence $\langle B^i, B^j \rangle = 0$ for $i \neq j$, and Itô's formula becomes

$$F(B_t) = F(B_0) + \int_0^t \nabla F(B_s) \cdot \mathrm{d}B_s + \frac{1}{2} \int_0^t \Delta F(B_s)\,\mathrm{d}s.$$

If $F \in C^2(\mathbb{R}_+ \times \mathbb{R}^p)$ instead, then we get

$$F(t, B_t) = F(0, B_0) + \int_0^t \underset{\text{[gradient for space variables]}}{\nabla F(s, B_s)} \cdot \mathrm{d}B_s + \int_0^t \Big(\frac{\partial F}{\partial t} + \frac{1}{2}\,\underset{\text{[Laplacian for space variables]}}{\Delta F}\Big)(s, B_s)\,\mathrm{d}s.$$

We actually do not need F to have a second derivative in t. Indeed, in general, if $X^1, \ldots, X^p$ are continuous semimartingales and X^i ($i \in I$) are finite-variation, then we need only F_i continuous ($1 \leqslant i \leqslant p$) and $F_{i,j}$ continuous ($i, j \notin I$).

Suppose $U \subseteq \mathbb{R}^p$ is open and $F \in C^2(U)$. If $X_s \in U$ almost surely for $0 \leqslant s < T$ (T random) and $X_0 \in K$, where K is compact, then we may still apply Itô's formula to $F(X_T)$. To see this, let $K \subseteq V_1 \subseteq V_2 \subseteq \cdots$ be open with $\overline{V_n} \subseteq U$ and $\bigcup_n V_n = U$. Let

$$T_n := \inf\{t \geqslant 0\,;\ X_t \notin V_n\},$$

which is a stopping time by Proposition 3.9. By using a partition of unity, we may construct $G_n \in C^2(\mathbb{R}^p)$ such that $G_n \restriction \overline{V}_n = F \restriction \overline{V}_n$. Itô's formula applied to $G_n(X^{T_n})$ involves only F and its derivatives. Then we may let $n \to \infty$, noting that $T_n \wedge T \to T$.

Exercise (due 1/25). Exercise 5.28 ("to be determined" means you should give it).

A $\mathbb{C}$-valued random process whose real and imaginary parts are continuous local martingales is called a ***complex continuous local martingale***.

Proposition 5.11. *Let M be a continuous local martingale and $\lambda \in \mathbb{C}$. The **(stochastic) exponential process***

$$\mathscr{E}(\lambda M) := \exp\left\{\lambda M - \langle \lambda M, \lambda M \rangle / 2\right\}$$

is a complex continuous local martingale that satisfies

$$\mathscr{E}(\lambda M) = \mathrm{e}^{\lambda M_0} + \lambda \mathscr{E}(\lambda M) \cdot M,$$

where

$$\lambda \mathscr{E}(\lambda M) \cdot M := \Big(\mathrm{Re}\big(\lambda \mathscr{E}(\lambda M)\big)\Big) \cdot M + \mathrm{i}\Big(\mathrm{Im}\big(\lambda \mathscr{E}(\lambda M)\big)\Big) \cdot M.$$

We saw some examples in Section 3.3 that were true martingales.

Proof. The function

$$F(r, x) := \exp\{\lambda x - \lambda^2 r/2\}$$

satisfies the time-reversed heat equation, $F_1 + F_{22}/2 = 0$. Applying Itô's formula to $\mathrm{Re}\, F$ and $\mathrm{Im}\, F$ gives

$$\begin{aligned} F\big(\langle M, M \rangle, M\big) &= F(0, M_0) + F_2\big(\langle M, M \rangle, M\big) \cdot M + (F_1 + F_{22}/2)\big(\langle M, M \rangle, M\big) \cdot \langle M, M \rangle \\ &= F(0, M_0) + F_2\big(\langle M, M \rangle, M\big) \cdot M = \mathrm{e}^{\lambda M_0} + \lambda \mathscr{E}(\lambda M) \cdot M. \end{aligned}$$ ◀

Exercise (due 2/1). Let M be a continuous local martingale with $M_0 = 0$.

(1) Show that

$$\forall t, a, b > 0 \quad \mathbf{P}\big[M_t \geqslant a, \langle M, M\rangle_t \leqslant b\big] \leqslant \exp\Big\{-\frac{a^2}{2b}\Big\}.$$

(2) Show that

$$\forall a, b > 0 \quad \mathbf{P}\big[\exists t > 0 \;\; M_t \geqslant a, \langle M, M\rangle_t \leqslant b\big] \leqslant \exp\Big\{-\frac{a^2}{2b}\Big\}.$$

5.3. A Few Consequences of Itô's Formula

5.3.1. Lévy's Characterization of Brownian Motion

We know that if B is a real Brownian motion, then $\langle B, B\rangle_t = t$. In fact, B is the only continuous local martingale with this property:

Theorem 5.12 (Lévy)**.** *Let $X = (X^1, \ldots, X^d)$ be an adapted continuous process. The following are equivalent:*

(i) *X is a d-dimensional $(\mathscr{F}_t)$-Brownian motion.*

(ii) *Each X^i is a continuous local martingale and $\forall i, j, t \;\; \langle X^i, X^j\rangle_t = \delta_{ij} t$.*

Note this implies that if X is a Brownian motion and each coordinate is an $(\mathscr{F}_t)$-martingale, then X is an $(\mathscr{F}_t)$-Brownian motion. Also, if H is progressive and ± 1-valued and B is a 1-dimensional $(\mathscr{F}_t)$-Brownian motion, then $H \cdot B$ is also an $(\mathscr{F}_t)$-Brownian motion, an extension of the symmetry used in the reflection principle. More generally, if $B = (B^1, \ldots, B^d)$ is a d-dimensional $(\mathscr{F}_t)$-Brownian motion and H is a $(d \times d)$-matrix-valued process whose entries are in $L^2_{\mathrm{loc}}(B^1)$, then $H \cdot B$ is an $(\mathscr{F}_t)$-Brownian motion iff H is a.s. an orthogonal matrix.

Proof. We have seen (i) $\Rightarrow$ (ii) in Chapter 4. Assume (ii). Then for all $\xi \in \mathbb{R}^d$, the process

$$\xi \cdot X_t = \sum_{j=1}^{d} \xi_j X_t^j$$

is a continuous local martingale with

$$\langle \xi \cdot X, \xi \cdot X\rangle_t = \sum_{j,k} \xi_j \xi_k \langle X^j, X^k\rangle_t = |\xi|^2 t.$$

Use $\lambda := \mathrm{i}$ in Proposition 5.11 to conclude that $\big(\mathrm{e}^{\mathrm{i}\xi\cdot X_t + |\xi|^2 t/2}\big)_t$ is a complex continuous local martingale. Since it is bounded on every finite interval, it is a true complex martingale. That is, for $0 \leqslant s < t < \infty$,

$$\mathbf{E}\big[\mathrm{e}^{\mathrm{i}\xi\cdot X_t + |\xi|^2 t/2} \mid \mathscr{F}_s\big] = \mathrm{e}^{\mathrm{i}\xi\cdot X_s + |\xi|^2 s/2},$$

or

$$\mathbf{E}\big[\mathrm{e}^{\mathrm{i}\xi\cdot (X_t - X_s)} \mid \mathscr{F}_s\big] = \mathrm{e}^{-|\xi|^2 (t-s)/2}.$$

This means that for all $A \in \mathscr{F}_s$, the $P(\,\cdot \mid A)$-distribution of $X_t - X_s$ is $\mathscr{N}\big(\mathbf{0}, (t-s)I\big)$, and thus $X_t - X_s \perp\!\!\!\perp \mathscr{F}_s$. Hence, all X^j have independent increments with respect to $(\mathscr{F}_t)$. Furthermore,

$j \neq k$ implies that $X_t^j - X_s^j$ and $X_t^k - X_s^k$ are independent given $\mathscr{F}_s$. It follows that $X - X_0$ is a d-dimensional $(\mathscr{F}_t)$-Brownian motion started from $\mathbf{0}$; since $X - X_0 \perp\!\!\!\perp \mathscr{F}_0$ and $X_0 \in \mathscr{F}_0$, also $X - X_0 \perp\!\!\!\perp X_0$, whence X is a d-dimensional $(\mathscr{F}_t)$-Brownian motion. ◀

Exercise. Let X^i be continuous square-integrable martingales for $1 \leqslant i \leqslant d$. Show that $X := (X^1, \ldots, X^d)$ is a d-dimensional $(\mathscr{F}_t)$-Brownian motion if and only if for all i, j, and $s < t$, $\mathbf{E}\big[(X_t^i - X_s^i)(X_t^j - X_s^j) \bigm| \mathscr{F}_s\big] = \delta_{ij}(t - s)$.

Exercise (due 2/1). Show that a continuous local martingale M is an $(\mathscr{F}_t)$-Brownian motion if and only if for all $f \in C^2(\mathbb{R})$,

$$\Big(f(M_t) - \frac{1}{2}\int_0^t f''(M_s)\,\mathrm{d}s\Big)_{t \geqslant 0}$$

is a continuous local martingale.

Exercise. Let B be a Brownian motion. Suppose that $H \in L^2_{\mathrm{loc}}(B)$ is such that $H \cdot B$ is a Gaussian process. Show that $H \cdot B$ is indistinguishable from a Wiener-integral process.

Exercise. Let B be a Brownian motion. Suppose that H is a measurable process (not necessarily adapted) such that $\mathbf{E}\big[|H|\big] \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$. Show that the Brownian motion with random drift defined by $X_t := B_t + \int_0^t H_s\,\mathrm{d}s$ also satisfies $X_t = \beta_t + \int_0^t \mathbf{E}[H_s \mid \mathscr{F}_s^X]\,\mathrm{d}s$ for some $(\mathscr{F}_t^X)$-Brownian motion, β.

5.3.2. Continuous Martingales as Time-Changed Brownian Motions

We have seen several quantitative similarities between continuous local martingales and Brownian motion. This is not a coincidence. In fact, a continuous local martingale M is a Brownian motion β with time process $\langle M, M\rangle$:

$$M_t = \beta_{\langle M,M\rangle_t}.$$

This is similar in spirit to other ways of representing random walks or random variables via Brownian motion. For example, for simple random walk on $\mathbb{Z}$, we could let B be a Brownian motion from 0, $\tau_1 := \inf\{t\,;\ |B_t| = 1\}$, $\tau_2 := \inf\{t > \tau_1\,;\ |B_t - B_{\tau_1}| = 1\}$, etc. Then $\big(B_{\tau_n}\big)_{n \geqslant 0}$ has the law of simple random walk with $\tau_0 := 0$, and also has the nice property that $\mathbf{E}[\tau_n - \tau_{n-1}] = 1$. If the steps have mean 0 and finite variance more generally, this is a bit harder to achieve:

Exercise (due 2/8). (Skorokhod) Let B be a Brownian motion and Z be a random variable with $\mathbf{E}[Z] = 0$ and $\mathbf{E}\big[Z^2\big] < \infty$. Let $p := \mathbf{E}\big[Z\mathbf{1}_{[Z>0]}\big]$.

(1) Show that

$$\frac{x - y}{p}\mathbf{1}_{(0,\infty)}(x)\mathbf{1}_{(-\infty,0]}(y)\,\mathrm{d}F_Z(x)\,\mathrm{d}F_Z(y)$$

is a probability measure on $\mathbb{R}^2$, where F_Z is the c.d.f. of Z.

(2) Let (X, Y) have the law of (1), independent of B. Write $T_a := \inf\{t\,;\ B_t = a\}$. Show that $B_{T_X \wedge T_Y} \sim F_Z$ and

$$\mathbf{E}[T_X \wedge T_Y] = \mathbf{E}\big[Z^2\big].$$

(3) Show there exists a continuous closed martingale M on some filtered probability space such that $M_0 = 0$ and $M_\infty \sim F_Z$.

Skorokhod's embedding can be done even with an $\big(\mathscr{F}_t^B\big)$-stopping time, but that is much harder; see, e.g., Billingsley.

We need a strengthening of Proposition 4.12. For a function $f\colon \mathbb{R}_+ \to \mathbb{R}$, write

$$C_f := \bigcup\Big\{[s,t]\,;\; s<t,\; f\restriction[s,t] \text{ is constant}\Big\}$$

for its intervals of constancy.

Lemma 5.14. *Let M be a continuous local martingale. Then $C_M = C_{\langle M,M\rangle}$ almost surely.*

Proof. By continuity of M and $\langle M,M\rangle$, this will follow from the statement that for $0 \leqslant a < b$,

$$\mathbf{P}\Big[\big([a,b] \subseteq C_M\big) \,\triangle\, \big([a,b] \subseteq C_{\langle M,M\rangle}\big)\Big] = 0.$$

(This will also show that $\big[C_M = C_{\langle M,M\rangle}\big]$ is measurable.)

Fix $a<b$. By Eq. (4.3) of Theorem 4.9, it follows directly that

$$\mathbf{P}\Big[\big([a,b] \subseteq C_M\big) \setminus \big([a,b] \subseteq C_{\langle M,M\rangle}\big)\Big] = 0.$$

For the other direction, let $N := M - M^a$. By exercise, we have

$$\langle N,N\rangle = \langle M,M\rangle - \langle M,M\rangle^a.$$

Define $T_0 := \inf\big\{t \geqslant 0\,;\; \langle N,N\rangle_t > 0\big\}$. Now, this may not be a stopping time. However, let us change to the filtration $(\mathscr{F}_{t^+})_t$, with respect to which T_0 *is* a stopping time by Proposition 3.9(i). In addition, N is still a continuous local martingale by Theorem 3.17. Since $\langle N,N\rangle^{T_0} = 0$, it follows from Proposition 4.12 that $N^{T_0} = 0$ a.s. If $[a,b] \subseteq C_{\langle M,M\rangle}(\omega)$, then $T_0(\omega) \geqslant b$, whence $\forall t \leqslant b\;\; N_t(\omega) = 0$ for a.e. such ω. This proves the other direction. ◀

Exercise (due 2/1). Show that if $X = M + V$ is the canonical decomposition of a continuous semimartingale, then $C_X = C_M \cap C_V$ almost surely.

Exercise. Let X be a continuous semimartingale and H be a locally bounded, progressive process. Show that almost surely,

$$C_H^0 \cup C_X \subseteq C_{H\cdot X},$$

where $C_H^0 := C_H \cap H^{-1}\big[\{0\}\big]$.

Theorem 5.13 (Dambis–Dubins–Schwarz)**.** *If M is a continuous local martingale with $\langle M,M\rangle_\infty = \infty$ almost surely, then there exists a Brownian motion β such that*

$$\textit{a.s.}\quad \forall t \geqslant 0 \qquad M_t = \beta_{\langle M,M\rangle_t}.$$

Remarks. (1) If $\langle M,M\rangle_\infty < \infty$ with positive probability, one can do the same, but one may need a larger probability space to define β after time $\langle M,M\rangle_\infty$. It follows that for every $t>0$, up to a set of probability 0, we have $\sup_{s<t} M(s) > 0$ iff $\inf_{s<t} M(s) < 0$ iff $\sup_{s<t}|M(s)| > 0$ iff $\langle M,M\rangle_t > 0$ by Theorem 2.13.

(2) β is not adapted to $(\mathscr{F}_t)$, but to a "time-changed" filtration.

(3) $\beta \perp\!\!\!\perp \langle M, M\rangle$ if and only if M is an ***Ocone*** continuous local martingale if and only if the conditional law of $(M_t - M_s)_{t\geqslant s}$ given $(M_t)_{t\leqslant s}$ is symmetric for all $s \geqslant 0$.

Proof. We will define β by the conclusion, "inverting" $\langle M, M\rangle$. Assume first $M_0 = 0$ almost surely. For $r \geqslant 0$, set

$$\tau_r := \inf\bigl\{t \geqslant 0\,;\ \langle M, M\rangle_t \geqslant r\bigr\}.$$

By Proposition 3.9, τ_r is a stopping time. Except on the event $\mathscr{N} := \bigl[\langle M, M\rangle_\infty < \infty\bigr]$, we have $\tau_r < \infty$ for all r. Since $\mathbf{P}[\mathscr{N}] = 0$, we may redefine τ_r to be 0 on $\mathscr{N}$. Recall that $(\mathscr{F}_t)$ is complete by assumption, so still τ_r is a stopping time.

Note that $\langle M, M\rangle$ can be constant on intervals (where a.s., M is constant by Lemma 5.14). Still, $r \mapsto \tau_r$ is increasing and left-continuous, so has right limits, namely,

$$\lim_{s\downarrow r} \tau_s = \tau_{r^+} = \inf\{t \geqslant 0\,;\ \langle M, M\rangle_t > r\},$$

except on $\mathscr{N}$, where $\tau_{r^+} = 0$.

Define $\beta_r := M_{\tau_r}$ for $r \geqslant 0$. By Theorem 3.7, $\beta_r \in \mathscr{F}_{\tau_r}$, i.e., β is adapted to $(\mathscr{G}_r)$, where $\mathscr{G}_r := \mathscr{F}_{\tau_r}$ and $\mathscr{G}_\infty := \mathscr{F}_\infty$. Because $(\mathscr{F}_t)$ is complete, so is $(\mathscr{G}_r)$.

Let $\mathscr{N}'$ be the set of probability 0 where M is non-constant on some interval where $\langle M, M\rangle$ is constant. Then off $\mathscr{N}'$, we have $M_{\tau_r} = M_{\tau_{r^+}}$, whence β is continuous. Redefine $\beta := 0$ on $\mathscr{N}'$. We have off $\mathscr{N} \cup \mathscr{N}'$,

$$\beta_{\langle M,M\rangle_t} = M_{\tau_{\langle M,M\rangle_t}}$$

and

$$\tau_{\langle M,M\rangle_t} \leqslant t \leqslant \tau_{\langle M,M\rangle_t^+}.$$

Because M is constant on that interval, we get $\beta_{\langle M,M\rangle_t} = M_t$.

It remains to show that β is a Brownian motion. We use Lévy's characterization, i.e., we prove that β and $(\beta_s^2 - s)_{s\geqslant 0}$ are continuous $(\mathscr{G}_r)$-martingales. Consider $n \in \mathbb{N}$. Since $\langle M, M\rangle_\infty^{\tau_n} = \langle M, M\rangle_{\tau_n} = n$ almost surely, Theorem 4.13(i) yields that M^{τ_n} and $(M^{\tau_n})^2 - \langle M, M\rangle^{\tau_n}$ are uniformly integrable martingales. The optional stopping theorem thus gives

$$0 \leqslant r \leqslant s \leqslant n \quad\Longrightarrow\quad \mathbf{E}\bigl[\beta_s \bigm| \mathscr{G}_r\bigr] = \mathbf{E}\bigl[M_{\tau_s}^{\tau_n} \bigm| \mathscr{F}_{\tau_r}\bigr] = M_{\tau_r}^{\tau_n} = \beta_r.$$

Similarly,

$$\mathbf{E}\bigl[\beta_s^2 - s \bigm| \mathscr{G}_r\bigr] = \mathbf{E}\bigl[(M_{\tau_s}^{\tau_n})^2 - \langle M, M\rangle_{\tau_s}^{\tau_n} \mid \mathscr{F}_{\tau_r}\bigr] = (M_{\tau_r}^{\tau_n})^2 - \langle M, M\rangle_{\tau_r}^{\tau_n} = \beta_r^2 - r.$$

This finishes the proof when $M_0 = 0$.

If $M_0 \neq 0$, write $M_t = M_0 + M_t'$. The previous argument gives a Brownian motion β' such that

$$\text{a.s.}\quad \forall t \geqslant 0 \quad M_t' = \beta'_{\langle M',M'\rangle_t}.$$

We actually showed that β' is a $(\mathscr{G}_r)$-Brownian motion, so $\beta' \perp\!\!\!\perp \mathscr{G}_0 = \mathscr{F}_0 \ni M_0$. Therefore, $\beta_s := M_0 + \beta_s'$ is a Brownian motion. ◂

Exercise (due 2/8). Exercise 5.27.

The following additional result will be useful in Chapter 7 when we show conformal invariance of complex Brownian motion.

Proposition 5.15. *Let M, N be continuous local martingales such that $M_0 = N_0 = 0$, $\langle M, M\rangle = \langle N, N\rangle$, $\langle M, N\rangle = 0$, and $\langle M, M\rangle_\infty = \langle N, N\rangle_\infty = \infty$. Let β, γ be the real Brownian motions such that $M = \beta_{\langle M,M\rangle}$ and $N = \gamma_{\langle N,N\rangle}$. Then $\beta \perp\!\!\!\perp \gamma$. Thus, (M, N) is a time change of a 2-dimensional Brownian motion.*

(Note: if $\langle M, M\rangle$ is not deterministic, then M is not independent of N.)

Proof. Again, let $\tau_r := \inf\{t \geqslant 0\,;\ \langle M, M\rangle \geqslant r\}$, so $\beta_r = M_{\tau_r}$, $\gamma_r = N_{\tau_r}$ and β, γ are $(\mathscr{G}_r)$-Brownian motions, where $\mathscr{G}_r := \mathscr{F}_{\tau_r}$. Since $\langle M, N\rangle = 0$, we have MN is a continuous local martingale. As before, we get $(MN)^{\tau_n}$ is a uniformly integrable martingale for $n \geqslant 1$ (now using Proposition 4.15(v)), whence

$$0 \leqslant r \leqslant s \leqslant n \quad\Longrightarrow\quad \mathbf{E}\big[\beta_s\gamma_s \bigm| \mathscr{G}_r\big] = \mathbf{E}\big[M^{\tau_n}_{\tau_s} N^{\tau_n}_{\tau_s} \bigm| \mathscr{F}_{\tau_r}\big] = M^{\tau_n}_{\tau_r} N^{\tau_n}_{\tau_r} = \beta_r\gamma_r.$$

Thus, $\beta\gamma$ is a $(\mathscr{G}_r)$-martingale and so $\langle \beta, \gamma\rangle = 0$. By Theorem 5.12, (β, γ) is a 2-dimensional Brownian motion. Since $\beta_0 = \gamma_0 = 0$, it follows that $\beta \perp\!\!\!\perp \gamma$. ◀

The proposition also holds without the assumption that $\langle M, M\rangle_\infty = \langle N, N\rangle_\infty = \infty$; see the first remark after Theorem 5.13. In addition, there is an extension due to Knight when $\langle M, M\rangle \neq \langle N, N\rangle$, but one loses the filtration $(\mathscr{G}_r)$; it is more difficult.

Exercise (due 2/8). (1) Let M be a continuous local martingale such that $M_0 = 0$ and $\langle M, M\rangle$ is deterministic with $\langle M, M\rangle_\infty = \infty$. Show that M is a Gaussian process.

(2) Show that if M and N are continuous local martingales such that $M_0 = N_0 = 0$, $\langle M, M\rangle$ and $\langle N, N\rangle$ are deterministic, $\langle M, M\rangle_\infty = \langle N, N\rangle_\infty = \infty$ and $\langle M, N\rangle = 0$, then $M \perp\!\!\!\perp N$. Do not use Knight's theorem. *Hint:* modify the proof of Lévy's theorem.

Exercise (due 2/15). Exercise 5.33.

Exercise. Let M be a continuous martingale with $M_t \in L^2(\mathbf{P})$ for all $t \in \mathbb{R}_+$ and $\langle M, M\rangle_\infty = \infty$ almost surely. Write $d(s, t) := \|M_s - M_t\|_2$ for $s, t \in \mathbb{R}_+$. Show that the stochastic process M is almost surely locally Hölder continuous of order α for all $\alpha < 1$ with respect to the metric d on $\mathbb{R}_+$.

5.3.3. The Burkholder–Davis–Gundy Inequalities

Here we give yet another relation between a continuous local martingale and its quadratic variation. This will be useful in Chapter 8. For a process X, write

$$X^*_t := \sup_{s \leqslant t} |X_s|.$$

Theorem 5.16 (Burkholder–Davis–Gundy)**.** *There exist $c, C\colon (0, \infty) \to (0, \infty)$ such that for all continuous local martingales M with $M_0 = 0$, for all stopping times T,*

$$\forall p \in \mathbb{R}_+ \qquad c(p)\,\mathbf{E}\Big[\sqrt{\langle M, M\rangle_T}^{\,p}\Big] \leqslant \mathbf{E}\big[(M^*_T)^p\big] \leqslant C(p)\,\mathbf{E}\Big[\sqrt{\langle M, M\rangle_T}^{\,p}\Big].$$

Remark. Note that the case $p = 2$ is immediate from Doob's L^2-inequality and Theorem 4.13: if $M^*_T \in L^2$, then $M^T \in \mathbb{H}^2$ and

$$\mathbf{E}\big[(M^*_T)^2\big] \leqslant 4\,\mathbf{E}\big[(M_T)^2\big] = 4\,\mathbf{E}\big[\langle M, M\rangle_T\big] \leqslant 4\,\mathbf{E}\big[(M^*_T)^2\big],$$

whereas if $\langle M, M\rangle_T \in L^1$, then $M^*_T \in L^2$.

We need some preliminary results.

Proposition. *Let M be a continuous local martingale with $M_0 = 0$. Write*

$$T_x := \inf\{t \geqslant 0\,;\ M_t = x\}.$$

We have

$$\forall a, b > 0 \qquad \mathbf{P}\big[T_a < T_{-b}\big] \leqslant \frac{b}{a+b}\,\mathbf{P}[M_\infty^* > 0].$$

Proof. Because $M^{T_a \wedge T_{-b}}$ is a bounded martingale, we have

$$\begin{aligned} 0 = \mathbf{E}[M_{T_a \wedge T_{-b}}] &= a\,\mathbf{P}[T_a < T_{-b}] - b\,\mathbf{P}[T_{-b} < T_a] + \mathbf{E}\big[M_\infty \mathbf{1}_{[T_a = T_{-b} = \infty, M_\infty^* > 0]}\big] \\ &\geqslant a\,\mathbf{P}[T_a < T_{-b}] - b\,\mathbf{P}[T_{-b} \leqslant T_a,\ M_\infty^* > 0] \\ &= -b\,\mathbf{P}[M_\infty^* > 0] + (a+b)\,\mathbf{P}[T_a < T_{-b}]. \end{aligned}$$ ◄

Corollary. *Let $X, Y \geqslant 0$, $X_0 = Y_0 = 0$, and $X - Y$ be a continuous local martingale. Then*

$$\forall 0 < b < a \qquad \mathbf{P}\big[X_\infty^* \geqslant a,\ Y_\infty^* < b\big] \leqslant \frac{b}{a}\,\mathbf{P}\big[(X-Y)_\infty^* > 0\big].$$

Proof. Since

$$\big[X_\infty^* \geqslant a,\ Y_\infty^* < b\big] \subseteq \big[\sup(X-Y) > a-b\big] \cap \big[\inf(X-Y) > -b\big],$$

it follows that on this event, $X - Y$ hits $a - b$ before $-b$, so the proposition applies. ◄

Corollary. *Let M be a continuous local martingale with $M_0 = 0$ and $r > 0$. Then*

$$\forall b \in (0,1) \quad \mathbf{P}\big[(M_\infty^*)^2 \geqslant 4r,\ \langle M, M\rangle_\infty < br\big] \leqslant b\,\mathbf{P}\big[(M_\infty^*)^2 > r\big]$$

and

$$\forall b \in (0, \tfrac{1}{4}) \quad \mathbf{P}\big[\langle M, M\rangle_\infty \geqslant 2r,\ (M_\infty^*)^2 < br\big] \leqslant 4b\,\mathbf{P}\big[\langle M, M\rangle_\infty > r\big].$$

Proof. Since $M^2 - \langle M, M\rangle$ is a continuous local martingale, the previous corollary gives

$$\forall b \in (0,1) \quad \mathbf{P}\big[(M_\infty^*)^2 \geqslant r,\ \langle M, M\rangle_\infty < br\big] \leqslant b\,\mathbf{P}\big[M_\infty^* > 0\big].$$

Here, we have used the fact that $M_\infty^* > 0$ iff $\langle M, M\rangle_\infty > 0$ by Proposition 4.12. Apply this to $N := M - M^T$, where $T := \inf\big\{t \geqslant 0\,;\ M_t^* \geqslant \sqrt{r}\big\}$. Then

$$\begin{gathered} [N_\infty^* > 0] = [(M_\infty^*)^2 > r], \\ \langle N, N\rangle = \langle M, M\rangle - \langle M, M\rangle^T \leqslant \langle M, M\rangle, \end{gathered}$$

and

$$[M_\infty^* \geqslant 2\sqrt{r}\,] \subseteq [N_\infty^* \geqslant \sqrt{r}\,]$$

since $N_\infty^* \geqslant M_\infty^* - \sqrt{r}$ in that case. This gives the first inequality.

Likewise, $\langle M, M\rangle - M^2$ is a continuous local martingale, so

$$\forall b \in (0, \tfrac{1}{4}) \quad \mathbf{P}\big[\langle M, M\rangle_\infty \geqslant r,\ (M_\infty^*)^2 < 4br\big] \leqslant 4b\,\mathbf{P}\big[\langle M, M\rangle_\infty > 0\big].$$

Apply this to $N := M - M^T$, where $T := \inf\{t \geqslant 0\,;\ \langle M, M\rangle_t \geqslant \sqrt{r}\}$. Then

$$\big[\langle N, N\rangle_\infty > 0\big] = \big[\langle M, M\rangle_\infty > r\big],$$
$$\big[\langle M, M\rangle_\infty \geqslant 2r\big] \subseteq \big[\langle N, N\rangle_\infty \geqslant r\big],$$

and

$$\big[(M^*_\infty)^2 < br\big] \subseteq \big[(N^*_\infty)^2 < 4br\big]$$

since $\forall t \geqslant 0\ \ M_t \in (-\sqrt{br}, \sqrt{br}\,)$ in that case. This gives the second inequality. ◀

Proof of Theorem 5.16. Recall that for $X \geqslant 0$ and $p > 0$,

$$\mathbf{E}\big[X^p\big] = \int_0^\infty p\,\mathbf{P}[X \geqslant r]r^{p-1}\,\mathrm{d}r = b^p \int_0^\infty p\,\mathbf{P}[X \geqslant br]r^{p-1}\,\mathrm{d}r$$

for $b > 0$. By the corollary, for $b \in (0, 1)$,

$$\begin{aligned}\mathbf{P}\big[(M^*_\infty)^2 \geqslant 4r\big] &\leqslant \mathbf{P}\big[\langle M, M\rangle_\infty \geqslant br\big] + \mathbf{P}\big[(M^*_\infty)^2 \geqslant 4r,\ \langle M, M\rangle_\infty < br\big]\\ &\leqslant \mathbf{P}\big[\langle M, M\rangle_\infty \geqslant br\big] + b\,\mathbf{P}\big[(M^*_\infty)^2 \geqslant r\big].\end{aligned}$$

Multiply by $\frac{p}{2}r^{\frac{p}{2}-1}$ and integrate from $r = 0$ to ∞:

$$2^{-p}\,\mathbf{E}\big[(M^*_\infty)^p\big] \leqslant b^{-p/2}\,\mathbf{E}\big[\langle M, M\rangle_\infty^{p/2}\big] + b\,\mathbf{E}\big[(M^*_\infty)^p\big].$$

Choose $b \in (0, 2^{-p})$ to obtain $C(p)$ for $T = \infty$.

Similarly, for $b \in (0, \frac{1}{4})$, we have

$$\mathbf{P}\big[\langle M, M\rangle_\infty \geqslant 2r\big] \leqslant \mathbf{P}\big[(M^*_\infty)^2 \geqslant br\big] + 4b\,\mathbf{P}\big[\langle M, M\rangle_\infty \geqslant r\big],$$

so

$$2^{-p/2}\,\mathbf{E}\big[\langle M, M\rangle_\infty^{p/2}\big] \leqslant b^{-p/2}\,\mathbf{E}\big[(M^*_\infty)^p\big] + 4b\,\mathbf{E}\big[\langle M, M\rangle_\infty^{p/2}\big].$$

Choose $b \in (0, 2^{-p/2}/4)$ to obtain $c(p)$ for $T = \infty$.

Finally, as usual, apply these inequalities to M^T to obtain them for any T, not just $T = \infty$. ◀

Exercise. Let M be a continuous local martingale with $M_0 = 0$ and $a, b > 0$. Applying the previous exercise on page 73 to both M and $-M$, we see that

$$\mathbf{P}\big[(M^*_\infty)^2 \geqslant a,\ \langle M, M\rangle_\infty \leqslant b\big] \leqslant 2\mathrm{e}^{-\frac{a}{2b}}.$$

Show that

$$\mathbf{P}\big[\langle M, M\rangle_\infty \geqslant a,\ (M^*_\infty)^2 \leqslant b\big] \leqslant \mathbf{P}\big[(B^*_a)^2 \leqslant b\big],$$

where B is a real Brownian motion starting at 0. You may use the extension of Theorem 5.13 to the case $\langle M, M\rangle_\infty < \infty$ with positive probability.

An exponential bound on $\mathbf{P}\big[(B^*_a)^2 \leqslant b\big]$ is shown in Exercise 6.29(6); it can also be bounded by using an alternating infinite series expression for its exact value (e.g., page 342 of Feller, volume 2, or (7.15) of Mörters and Peres). Taking its first term gives the bound $\frac{4}{\pi}\exp\{-\frac{\pi^2 a}{8b}\}$. As a third alternative, one can get an exponential bound by a direct iterated martingale argument using a sequence of stopping times and conditioning on the associated σ-fields to bound the conditional probabilities.

Exercise. Let H be a bounded, continuous, adapted process with $H_0 \equiv 0$ and B be a Brownian motion. Show that $\|(H \cdot B)_t^*/B_t\|_p \to 0$ as $t \downarrow 0$ for all $p \in (0,1)$. Find an H so that $\|(H \cdot B)_t/B_t\|_1 \not\to 0$ as $t \downarrow 0$.

Corollary 5.17. *If M be a continuous local martingale with $M_0 = 0$ and*

$$\mathbf{E}\Big[\sqrt{\langle M, M\rangle_\infty}\,\Big] < \infty,$$

then M is a uniformly integrable martingale.

Proof. By Theorem 5.16 with $p = 1$, we have $|M| \leqslant M^*_\infty \in L^1$, so Proposition 4.7(ii) applies. ◂

Note that this condition is weaker than $\mathbf{E}\big[\langle M, M\rangle_\infty\big] < \infty$, which is the condition for $M \in \mathbb{H}^2$.

Exercise (due 2/15). Let B be a Brownian motion with $B_0 = 0$ and T be a stopping time with $\mathbf{E}\big[\sqrt{T}\,\big] < \infty$. Show that $\mathbf{E}[B_T] = 0$ and $\mathbf{E}\big[B_T^2\big] = \mathbf{E}[T]$ (this is trickier than it looks). Give a stopping time T where both equalities fail yet $\mathbf{E}[T^c] < \infty$ for all $c < 1/2$.

Exercise. Let B be a Brownian motion with $B_0 = 0$ and T be a finite stopping time. Show that for all $\mu \neq 0$, we have $\mathbf{E}[B_T + \mu T] = \mu\,\mathbf{E}[T]$.

Appendix: The Cameron–Martin and Girsanov Theorems

While the proof of the Cameron–Martin theorem given in the appendix to Chapter 2 is short and elementary, it is instructive to see how stochastic calculus can also be used for a short proof. This will lead us to extensions of the theorem. It is convenient here to change the sign of the drift. Let B be an $(\mathscr{F}_t)_t$-Brownian motion. We begin with linear drift up to some finite time, r, after which we add no further drift: $B_t - \theta(t \wedge r)$. Note that the quadratic variation does not change with a deterministic drift. Let $M := \mathscr{E}(\theta B)$ be the exponential martingale corresponding to B. Recall from the exercise on page 28 that we may differentiate M with respect to θ to get another martingale, namely, $(B_t - \theta t)M_t$. We claim this means that with respect to the probability measure $\mathbf{Q} := M_r\,\mathbf{P}$, the process $\big(B_t - \theta(t \wedge r)\big)_t$ is an $(\mathscr{F}_t)_t$-martingale. First note the following general principle: if X is an adapted process and N is a uniformly integrable nonnegative martingale such that $(X_t N_t)_t$ is a martingale, then $(X_t N_\infty)_t$ is a martingale. Indeed, $X_t N_\infty$ is integrable because $\mathbf{E}[|X_t| N_\infty] = \mathbf{E}\big[\mathbf{E}[|X_t| N_\infty \mid \mathscr{F}_t]\big] = \mathbf{E}[|X_t| N_t]$. Now we may calculate for $0 \leqslant s \leqslant t$ and $A \in \mathscr{F}_s$ that

$$\mathbf{E}[X_t N_\infty \mathbf{1}_A] = \mathbf{E}\big[\mathbf{E}[X_t N_\infty \mathbf{1}_A \mid \mathscr{F}_t]\big] = \mathbf{E}[X_t N_t \mathbf{1}_A] = \mathbf{E}[X_s N_s \mathbf{1}_A] = \mathbf{E}[X_s N_\infty \mathbf{1}_A].$$

Using this principle, we see that $(B_t - \theta t)_{0 \leqslant t \leqslant r}$ is a martingale with respect to $\mathbf{Q}$, and it is obvious that $(B_t - \theta r)_{r \leqslant t < \infty}$ is a martingale with respect to $\mathbf{Q}$. This proves our claim. Now the quadratic variation of $\big(B_t - \theta(t \wedge r)\big)_t$ is the same with respect to $\mathbf{Q}$ as with respect to $\mathbf{P}$ because $\mathbf{Q} \ll \mathbf{P}$. Hence, it follows from Lévy's theorem that $\big(B_t - \theta(t \wedge r)\big)_t$ is an $(\mathscr{F}_t)_t$-Brownian motion with respect to $\mathbf{Q}$. Writing Q for the $\mathbf{P}$-law of $(B_t - \theta t)_t$ and W for Wiener measure, we may conclude that $Q\!\restriction\!\mathscr{G}_r \ll W\!\restriction\!\mathscr{G}_r$ for all $r \in \mathbb{R}_+$, where $\mathscr{G}_\bullet$ is the natural filtration on $C(\mathbb{R}_+, \mathbb{R})$; we say that Q is ***locally absolutely continuous*** with respect to W, written $Q \overset{\text{loc}}{\ll} W$. Of course, if $\theta \neq 0$, then $Q \perp W$.

Exercise. Let B be a d-dimensional $(\mathscr{F}_t)_t$-Brownian motion and $\mu \in \mathbb{R}^d$. Define $X_t := B_t + \mu t$ and $T := \inf\{t \geqslant 0\,;\ |X_t| = 1\}$. Write $M_t := \exp\{-\mu \cdot B_t - |\mu|^2 t/2\}$ and $\mathbf{Q} := M_T\,\mathbf{P}$.

(1) Verify that M^T is a uniformly integrable $\mathbf{P}$-martingale.

(2) Show that the $\mathbf{Q}$-law of X^T is Brownian motion up to time T.

(3) Let P be the $\mathbf{Q}$-law of (X^T, T) on $C(\mathbb{R}_+, \mathbb{R}^d) \times \mathbb{R}_+$. Show that the $\mathbf{P}$-law of (X^T, T) is mP, where $m(w,t) = \mathrm{e}^{\mu \cdot w_t - |\mu|^2 t/2}$. Deduce that X_T and T are independent with respect to $\mathbf{P}$; that for some constant c_μ, the $\mathbf{P}$-law of X_T has density $x \mapsto c_\mu \mathrm{e}^{\mu \cdot x}$ with respect to hypersurface measure on the sphere of radius 1; and that the $\mathbf{P}$-law of T has density $t \mapsto c_\mu^{-1} \mathrm{e}^{-|\mu|^2 t/2}$ with respect to the $\mathbf{Q}$-law of T, the hitting time for ordinary Brownian motion.

For more general drift functions, suppose that $f \in L^2(\mathbb{R}_+)$ and $F_t := \int_0^t f(s)\,\mathrm{d}s$. We will consider $X := B - F$. To analyze this, let $L_t := \int_0^t f(s)\,\mathrm{d}B_s$, which is just a Wiener integral. By Proposition 5.11, we have $\mathrm{d}\mathscr{E}(L) = f\mathscr{E}(L)\,\mathrm{d}B$. Thus, integration by parts gives us

$$\mathrm{d}\big(X\mathscr{E}(L)\big) = Xf\mathscr{E}(L)\,\mathrm{d}B + \mathscr{E}(L)\big(\mathrm{d}B - f\,\mathrm{d}t\big) + f\mathscr{E}(L)\,\mathrm{d}t = Xf\mathscr{E}(L)\,\mathrm{d}B + \mathscr{E}(L)\,\mathrm{d}B,$$

whence $X\mathscr{E}(L)$ is a continuous local martingale. For $t \in [0,\infty]$, we have $L_t \sim \mathscr{N}\big(0, \|f\mathbf{1}_{[0,t]}\|^2\big)$, whence $\mathbf{E}\big[\mathrm{e}^{L_t}\big] = \mathrm{e}^{\|f\mathbf{1}_{[0,t]}\|^2/2} = \mathrm{e}^{\langle L,L\rangle_t}$, so $\mathbf{E}\big[\mathscr{E}(L)_t\big] = 1$. This implies that $\mathscr{E}(L)$ is a uniformly integrable martingale: by Proposition 4.7, it is a supermartingale, and so

$$1 = \mathbf{E}\big[\mathscr{E}(L)_t\big] \geqslant \mathbf{E}\big[\mathbf{E}[\mathscr{E}(L)_\infty \mid \mathscr{F}_t]\big] = \mathbf{E}\big[\mathscr{E}(L)_\infty\big] = 1,$$

which implies $\mathscr{E}(L)_t = \mathbf{E}\big[\mathscr{E}(L)_\infty \;\big|\; \mathscr{F}_t\big]$ a.s., as desired. Furthermore, $\langle X\mathscr{E}(L), X\mathscr{E}(L)\rangle_t = \int_0^t \big(X(s)f(s)+1\big)^2 \mathscr{E}(L)_s^2\,\mathrm{d}s$ has finite expectation, whence $X\mathscr{E}(L)$ is a true martingale by Theorem 4.13(ii). As above, it follows that X is a martingale with respect to $\mathbf{Q} := \mathscr{E}(L)_\infty\,\mathbf{P}$. Again, the quadratic variation of X is the same as that of B, whence X is an $(\mathscr{F}_t)_t$-Brownian motion with respect to $\mathbf{Q}$. The explicit form of $\mathscr{E}(L)_\infty$ is $\exp\big\{\int_0^\infty f(s)\,\mathrm{d}B_s - \int_0^\infty f(s)^2\,\mathrm{d}s/2\big\}$.

In fact, we may add random drifts as well: Suppose that L is a continuous local martingale such that $\mathscr{E}(L)$ is a uniformly integrable martingale with mean 1. Then $B - \langle B, L\rangle$ is an $(\mathscr{F}_t)_t$-Brownian motion with respect to $\mathscr{E}(L)_\infty\,\mathbf{P}$, whence the $\mathbf{P}$-law of $B - \langle B, L\rangle$ is mutually absolutely continuous with Wiener measure. This follows just as above, with the following extension of the "general principle" we used.

Proposition. *If N is a nonnegative uniformly integrable martingale, X is adapted, and XN is a local martingale, then $(X_t N_\infty)_t$ is a local martingale.*

Proof. We claim that a sequence of stopping times that reduces XN also reduces XN_∞. Indeed, let T be a stopping time such that $X^T N^T$ is a martingale; it suffices to show that $X^T N_\infty$ is a martingale. Integrability follows as before: $\mathbf{E}[|X_t^T|N_\infty] = \mathbf{E}\big[\mathbf{E}[|X_t^T|N_\infty \mid \mathscr{F}_{T\wedge t}]\big] = \mathbf{E}[|X_t^T|N_t^T]$. Let $0 \leqslant s \leqslant t$ and $A \in \mathscr{F}_s$. Similar to the calculation near the end of the proof of the proposition on page 35, we have

$$\begin{aligned}\mathbf{E}[X_t^T \mathbf{1}_{A\cap[T>s]} N_\infty] &= \mathbf{E}\big[\mathbf{E}[X_t^T \mathbf{1}_{A\cap[T>s]} N_\infty \mid \mathscr{F}_{T\wedge t}]\big] \\ &= \mathbf{E}\big[X_t^T \mathbf{1}_{A\cap[T>s]}\,\mathbf{E}[N_\infty \mid \mathscr{F}_{T\wedge t}]\big] = \mathbf{E}[X_t^T \mathbf{1}_{A\cap[T>s]} N_t^T].\end{aligned}$$

Applying this to $t = s$, we obtain $\mathbf{E}[X_s^T \mathbf{1}_{A\cap[T>s]} N_\infty] = \mathbf{E}[X_s^T \mathbf{1}_{A\cap[T>s]} N_s^T]$. Because $X^T N^T$ is a martingale, we conclude that $\mathbf{E}[X_t^T \mathbf{1}_{A\cap[T>s]} N_\infty] = \mathbf{E}[X_s^T \mathbf{1}_{A\cap[T>s]} N_\infty]$. On the other hand, $X_t^T = X_s^T$ on the event $[T \leqslant s]$, whence $\mathbf{E}[X_t^T \mathbf{1}_{A\cap[T\leqslant s]} N_\infty] = \mathbf{E}[X_s^T \mathbf{1}_{A\cap[T\leqslant s]} N_\infty]$. Adding these equations gives $\mathbf{E}[X_t^T \mathbf{1}_A N_\infty] = \mathbf{E}[X_s^T \mathbf{1}_A N_\infty]$, as desired. ◀

This part of the proof has nothing to do with Brownian motion, so we may deduce this theorem of Girsanov:

Theorem 5.22 (Girsanov)**.** *Let M and L be continuous local martingales such that $\mathscr{E}(L)$ is a uniformly integrable martingale with mean 1. Then $M - \langle M, L\rangle$ is a continuous local martingale with respect to $\mathscr{E}(L)_\infty\,\mathbf{P}$.*

Proof. Let $X := M - \langle M, L\rangle$. By Proposition 5.11, we have $\mathrm{d}\mathscr{E}(L) = \mathscr{E}(L)\,\mathrm{d}L$, so that from integration by parts,

$$\mathrm{d}\big(X\mathscr{E}(L)\big) = X\mathscr{E}(L)\,\mathrm{d}L + \mathscr{E}(L)\big(\mathrm{d}M - \mathrm{d}\langle M, L\rangle\big) + \mathscr{E}(L)\,\mathrm{d}\langle M, L\rangle = X\mathscr{E}(L)\,\mathrm{d}L + \mathscr{E}(L)\,\mathrm{d}M,$$

whence $X\mathscr{E}(L)$ is a continuous local martingale. ◀

Exercise. Find M and L as in Girsanov's theorem such that the $\mathbf{P}$-law of M is not equal to the $\mathscr{E}(L)_\infty\,\mathbf{P}$-law of $M - \langle M, L\rangle$.

Exercise. Show that if L is a continuous local martingale such that $\langle L, L\rangle_\infty = \infty$ a.s., then $\mathscr{E}(L)$ is not uniformly integrable. Show that for each $\varepsilon > 0$, there exists a continuous martingale L such that $\mathbf{P}[\langle L, L\rangle_\infty < \infty] < \varepsilon$ and $\mathscr{E}(L) + \mathscr{E}(-L)$ is uniformly integrable.

Exercise. Show that if L is a continuous local martingale such that $\langle L, L\rangle_\infty \leqslant \alpha < \infty$ a.s. for some constant, α, then $\mathbf{E}\big[\mathrm{e}^{c(L_\infty^*)^2}\big] < \infty$ for all constants $c < 1/(2\alpha)$.

Returning to Brownian motion with random drift, suppose that L is not only a continuous local martingale such that $\mathscr{E}(L)$ is a uniformly integrable martingale with mean 1, but also is adapted to the completed canonical filtration $\overline{\mathscr{F}_\bullet^B}$. In other words, there are Borel functions $f_t \colon C\big([0, t], \mathbb{R}\big) \to \mathbb{R}$ such that $L_t = f_t\big((B_s)_{0\leqslant s\leqslant t}\big)$ a.s. Write this relation as $L = f(B)$. Since $\beta := B - \langle B, L\rangle$ is a Brownian motion with respect to $\mathbf{Q} := \mathscr{E}(L)_\infty\,\mathbf{P}$, we have that the process $X = B$ satisfies the equation $X = \beta + \langle X, f(X)\rangle$, i.e., X is a Brownian motion with drift that depends on X.

To give a concrete example satisfying all these assumptions, suppose that $b \colon \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ is Borel with $g := \sup_x |b(\cdot, x)| \in L^2_{\mathrm{loc}}(\mathbb{R}_+)$. Then $(\omega, s) \mapsto b\big(s, B_s(\omega)\big) \in L^2_{\mathrm{loc}}(B)$, so we may define $L_t := \int_0^t b(s, B_s)\,\mathrm{d}B_s$. Because $\langle L, L\rangle_t \leqslant \int_0^t g(s)^2\,\mathrm{d}s < \infty$, the preceding exercise implies that $\mathscr{E}(L)$ is a martingale. We conclude that for each $t_0 < \infty$, the pair (X, β) on $\big(\Omega, (\mathscr{F}_t)_{t\leqslant t_0}, \mathscr{E}(L)_{t_0}\,\mathbf{P}\big)$ solves the stochastic differential equation $\mathrm{d}X_t = \mathrm{d}\beta_t + b(t, X_t)\,\mathrm{d}t$ for $0 \leqslant t \leqslant t_0$, and β is a Brownian motion up to time t_0. Because $\mathscr{E}(L)$ is a martingale, the laws for pairs (X, β) corresponding to two endings times t_0 and t_0' are consistent. Therefore, Kolmogorov's consistency theorem yields a global solution for all $t \geqslant 0$. Write Q for the resulting law of X on $C(\mathbb{R}_+, \mathbb{R})$. Because the law of X^{t_0} is the pushforward of $\mathscr{E}(L)_{t_0}\,\mathbf{P}$ by $X^{t_0} = B^{t_0}$, we have that $Q \overset{\mathrm{loc}}{\ll} W$ (and $W \overset{\mathrm{loc}}{\ll} Q$). We have that $Q \ll W$ iff $\mathscr{E}(L)$ is uniformly integrable. In Chapter 8, we will discuss solutions to SDEs, but only with restrictive regularity assumptions on the function b. We will also allow a function σ in front of $\mathrm{d}\beta_t$.

Chapter 6

General Theory of Markov Processes

The Markov property allows one to make many more calculations than one can for a general stochastic process. Also, it is desirable for modeling, analogous to not having a time-lag in a differential equation.

6.1. General Definitions and the Problem of Existence

Let $(E, \mathscr{E})$ be a measurable space. A ***(Markovian) transition kernel*** from E to E is a map $Q \colon E \times \mathscr{E} \to [0, 1]$ such that

(i) $\forall x \in E \;\; A \mapsto Q(x, A)$ is a probability measure on $(E, \mathscr{E})$, and

(ii) $\forall A \in \mathscr{E} \;\; x \mapsto Q(x, A)$ is $\mathscr{E}$-measurable.

This looks like a regular conditional probability, and indeed will be one. When E is countable, Q is determined by all $Q(x, \{y\})$, the transition matrix.

Let $B(E)$ be the space of bounded (real) $\mathscr{E}$-measurable functions on E with the supremum norm. For $f \in B(E)$, we write Qf for the function whose value at $x \in E$ is the integral of f with respect to $Q(x, \cdot)$; we write

$$(Qf)(x) = \int Q(x, \mathrm{d}y)\, f(y).$$

Obviously, $f \geqslant 0$ implies $Qf \geqslant 0$ and $\forall f \in B(E) \;\; Qf \in B(E)$ (note this is $\mathscr{E}$-measurable by approximating by simple functions) with $\|Qf\| \leqslant \|f\|$ (Q is a contraction on $B(E)$). Thus, Q defines a bounded, positive, linear operator on $B(E)$.

Definition 6.1. *A collection $(Q_t)_{t \geqslant 0}$ of transition kernels on E is called a* ***transition semigroup*** *if*

(i) $Q_0 = \mathrm{Id}$, *i.e.,* $\forall x \in E \;\; Q_0(x, \cdot) = \delta_x$,

(ii) $\forall s, t \geqslant 0 \;\; Q_t Q_s = Q_{t+s}$, *i.e.,* $\forall x \in E \;\; \forall A \in \mathscr{E}$

$$\int_E Q_t(x, \mathrm{d}y)\, Q_s(y, A) = Q_{t+s}(x, A),$$

called the ***Chapman–Kolmogorov identity****, and*

(iii) $\forall A \in \mathscr{E} \;\; (t, x) \mapsto Q_t(x, A)$ *is* $\mathscr{B}(\mathbb{R}_+) \otimes \mathscr{E}$*-measurable.*

Definition 6.2. *Given a filtered probability space* $\big(\Omega, \mathscr{F}, (\mathscr{F}_t)_{0\leqslant t\leqslant\infty}, \mathbf{P}\big)$ *and a transition semigroup* $(Q_t)_{t\geqslant 0}$ *on* E*, an* $(\mathscr{F}_t)$*-adapted* E*-valued process* $(X_t)_{t\geqslant 0}$ *is called a* ***Markov process*** *(with respect to* $(\mathscr{F}_t)$*)* ***with transition semigroup*** $(Q_t)_{t\geqslant 0}$ *if*

$$\forall s,t \geqslant 0 \;\; \forall A \in \mathscr{E} \qquad \mathbf{P}[X_{s+t} \in A \mid \mathscr{F}_s] = Q_t(X_s, A). \tag{$*$}$$

If we do not specify $(\mathscr{F}_t)$*, then we mean the canonical filtration* $(\mathscr{F}_t^X)$*.*

Thus, Q_t gives many regular conditional probabilities. Inherent in ($*$) is the assumption of time-homogeneity. Note that ($*$) gives

$$\forall s,t \geqslant 0 \;\; \forall A \in \mathscr{E} \qquad \mathbf{P}\big[X_{s+t} \in A \;\big|\; (X_r)_{0\leqslant r\leqslant s}\big] = Q_t(X_s, A).$$

This can also be stated as saying that X is Markov with respect to its canonical filtration $(\mathscr{F}_t^X)_t$.

Note also that ($*$) gives

$$\forall s,t \geqslant 0 \;\; \forall f \in B(E) \qquad \mathbf{E}\big[f(X_{s+t}) \mid \mathscr{F}_s\big] = (Q_t f)(X_s) : \tag{$**$}$$

the definition gives this when f is an indicator, from which it follows when f is simple and then, by taking a limit, for general f.

One can calculate as follows. Let $X_0 \sim \gamma$. We claim that if $0 < t_1 < t_2 < \cdots < t_p$ and $A_0, A_1, \ldots, A_p \in \mathscr{E}$, then

$$\begin{aligned} &\mathbf{P}\big[X_0 \in A_0, X_{t_1} \in A_1, \ldots, X_{t_p} \in A_p\big] \\ &\quad = \int_{A_0} \gamma(\mathrm{d}x_0) \int_{A_1} Q_{t_1}(x_0, \mathrm{d}x_1) \cdots \int_{A_p} Q_{t_p - t_{p-1}}(x_{p-1}, \mathrm{d}x_p). \end{aligned} \tag{$***$}$$

To show this, we show the more general formula,

$$\begin{aligned} &\forall f_0, f_1, \ldots, f_p \in B(E) \\ &\qquad \mathbf{E}\big[f_0(X_0) f_1(X_{t_1}) \cdots f_p(X_{t_p})\big] \\ &\qquad = \int \gamma(\mathrm{d}x_0)\, f_0(x_0) \int Q_{t_1}(x_0, \mathrm{d}x_1)\, f_1(x_1) \int Q_{t_2 - t_1}(x_1, \mathrm{d}x_2)\, f_2(x_2) \cdots \\ &\qquad\qquad \int Q_{t_p - t_{p-1}}(x_{p-1}, \mathrm{d}x_p)\, f_p(x_p). \end{aligned}$$

For $p = 0$, this is the definition of γ. Suppose $p \geqslant 1$ and the formula holds for $p - 1$. Then

$$\begin{aligned} &\mathbf{E}\big[f_0(X_0) f_1(X_{t_1}) \cdots f_p(X_{t_p})\big] \\ &\quad = \mathbf{E}\Big[f_0(X_0) f_1(X_{t_1}) \cdots f_{p-1}(X_{t_{p-1}})\, \mathbf{E}\big[f_p(X_{t_p}) \;\big|\; \mathscr{F}_{t_{p-1}}\big]\Big] \\ &\quad = \mathbf{E}\big[f_0(X_0) f_1(X_{t_1}) \cdots f_{p-1}(X_{t_{p-1}}) (Q_{t_p - t_{p-1}} f_p)(X_{t_{p-1}})\big] \end{aligned}$$

by Eq. ($**$), so we may apply the induction hypothesis with functions $f_0, \ldots, f_{p-2}, f_{p-1} \cdot (Q_{t_p - t_{p-1}} f_p)$ to get

$$\int \gamma(\mathrm{d}x_0)\, f_0(x_0) \cdots \int Q_{t_{p-1} - t_{p-2}}(x_{p-2}, \mathrm{d}x_{p-1})\, f_{p-1}(x_{p-1})\, \underbrace{(Q_{t_p - t_{p-1}} f_p)(x_{p-1})}_{= \int Q_{t_p - t_{p-1}}(x_{p-1}, \mathrm{d}x_p)\, f_p(x_p)},$$

with slightly different notation in case $p = 1$.

Conversely, if (∗∗∗) holds and (Q_t) is a transition semigroup, then $(X_t)_{t\geqslant 0}$ is a Markov process with semigroup $(Q_t)_{t\geqslant 0}$ with respect to $(\mathscr{F}_t^X)_{t\geqslant 0}$: use a π-λ argument as on page 262 of the book, number 3.

Note that (∗∗∗) shows that γ and $(Q_t)_{t\geqslant 0}$ determine the finite-dimensional distributions of X.

If Definition 6.1(i) and (iii) hold and $\forall x \in E\ \ \exists(\mathscr{F}_t)$-adapted X such that Eq. (∗) holds and $Q_t f(x) = \mathbf{E}\big[f(X_t)\big]$ for $f \in B(E)$, then $(Q_t)_t$ is a transition semigroup:

$$\begin{aligned} Q_{t+s} f(x) &= \mathbf{E}\big[f(X_{t+s})\big] = \mathbf{E}\Big[\mathbf{E}\big[f(X_{t+s}) \bigm| \mathscr{F}_s\big]\Big] \\ &= \mathbf{E}\big[Q_t f(X_s)\big] = Q_s(Q_t f)(x). \end{aligned}$$

Thus, if Eq. (∗∗∗) holds for all γ (or all δ_x), the Chapman–Kolmogorov identity holds.

Example. Let X be d-dimensional Brownian motion. Then X_t has density

$$p_t(x) := \frac{1}{(2\pi t)^{d/2}} \mathrm{e}^{-|x|^2/(2t)} \qquad (t > 0,\ x \in \mathbb{R}^d).$$

Let $Q_t(x, \cdot)$ have density $y \mapsto p_t(y - x)$ for $t > 0$. The Markov property of Brownian motion shows that X is a Markov process with transition semigroup $(Q_t)_{t\geqslant 0}$—in particular, (Q_t) is a semigroup.

Exercise (due 2/22). Exercise 6.24.

Given a transition semigroup, is there a Markov process with that semigroup? We show the answer is yes under a topological condition on E.

First, we recall a version of Kolmogorov's extension theorem. Let $\Omega^* := E^{\mathbb{R}_+}$ with the σ-field $\mathscr{F}^*$ generated by the coordinate maps $\omega \mapsto \omega(t)$ $(t \in \mathbb{R}_+)$. For $U \subseteq \mathbb{R}_+$, write $\pi_U : \Omega^* \to E^U$ for the map $\omega \mapsto \omega\restriction U$. For $U \subseteq V \subseteq \mathbb{R}_+$, write $\pi_U^V : E^V \to E^U$ for the map $\omega \mapsto \omega\restriction U$. Let $F(\mathbb{R}_+)$ be the collection of finite sets in $\mathbb{R}_+$.

A topological space is ***Polish*** if it is separable (there exists a countable dense subset) and its topology is generated by a complete metric.

Theorem 6.3 (Kolmogorov)**.** *Let $(E, \mathscr{E})$ be a Polish space with its Borel σ-field. Suppose that $\forall U \in F(\mathbb{R}_+)$ $\ \mu_U$ is a probability measure on E^U. If $(\mu_U)_{U\in F(\mathbb{R}_+)}$ is consistent in the sense that $\forall U \subseteq V \in F(\mathbb{R}_+)$ $\ (\pi_U^V)_*\mu_V = \mu_U$, then there exists a unique probability measure μ on $(\Omega^*, \mathscr{F}^*)$ such that $\forall U \in F(\mathbb{R}_+)$ $\ (\pi_U)_*\mu = \mu_U$.* ◀

In words: consistent finite-dimensional distributions determine a probability measure.

Corollary 6.4. *Let $(E, \mathscr{E})$ be a Polish space with its Borel σ-field. If $(Q_t)_{t\geqslant 0}$ is a transition semigroup on E and γ is a probability measure on E, then there exists a unique probability measure P on Ω^* such that the canonical process*

$$X_t(\omega) := \omega(t) \qquad (t \in \mathbb{R}_+,\ \omega \in \Omega^*)$$

is a Markov process with transition semigroup $(Q_t)_{t\geqslant 0}$ and $X_0 \sim \gamma$.

Proof. We make sure (∗∗∗) holds. Given $0 \leqslant t_1 < t_2 < \cdots < t_p$, we define $\mu_{\{t_1,\ldots,t_p\}}$ on $E^{\{t_1,\ldots,t_p\}}$ by

$$\mu_{\{t_1,\ldots,t_p\}}(A_1 \times \cdots \times A_p) := \int_E \gamma(\mathrm{d}x_0) \int_{A_1} Q_{t_1}(x_0, \mathrm{d}x_1) \cdots \int_{A_p} Q_{t_p - t_{p-1}}(x_{p-1}, \mathrm{d}x_p)$$

for $A_i \in \mathscr{E}$. The consistency condition amounts to putting some $A_i = E$ and verifying that those coordinates can be eliminated via Chapman–Kolmogorov (details are left to the reader). ◂

In particular, for $x \in E$ and $\gamma = \delta_x$, we write $\mathbf{P}_x$ for the associated measure. Note that $x \mapsto \mathbf{P}_x$ is measurable in the sense that for all $A \in \mathscr{F}^*$, $x \mapsto \mathbf{P}_x(A)$ is measurable: when A depends on only finitely many coordinates, this follows from the measurability assumption in Definition 6.1, and then the π-λ theorem gives it for all A. We may express the general measure $\mathbf{P}_{(\gamma)}$ associated to any γ by

$$\mathbf{P}_{(\gamma)}(A) = \int_E \gamma(\mathrm{d}x)\ \mathbf{P}_x(A) :$$

the integral makes sense by measurability of $x \mapsto \mathbf{P}_x$ and the integral is a probability measure by the monotone convergence theorem. By uniqueness, this is the measure from Corollary 6.4. Under additional assumptions (of Section 6.2), we prove there is a càdlàg modification of X in Section 6.3. There is a lot of operator theory one can develop related to semigroups, but we will avoid most of it. However, to motivate the next definition, suppose that $Q_t = \mathrm{e}^{tL}$ (reasonable from the Chapman–Kolmogorov identity). The resolvent of L is the operator-valued function $\lambda \mapsto (\lambda - L)^{-1}$ for $\lambda \notin \sigma(L)$. Formally, for $\lambda > 0$ and thinking of $L \leqslant 0$, we have

$$\int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{e}^{tL}\,\mathrm{d}t = (\lambda - L)^{-1}.$$

Definition 6.5. *For $\lambda > 0$, the **λ-resolvent** of the semigroup $(Q_t)_{t \geqslant 0}$ is the linear operator $R_\lambda \colon B(E) \to B(E)$ defined by*

$$(R_\lambda f)(x) := \int_0^\infty \mathrm{e}^{-\lambda t}(Q_t f)(x)\,\mathrm{d}t \quad (f \in B(E),\ x \in E).$$

Note that Definition 6.1(iii) shows that $t \mapsto (Q_t f)(x)$ is measurable and $R_\lambda f \in \mathscr{E}$ (if $g \in \mathscr{B}(\mathbb{R}_+) \otimes \mathscr{E}$ is bounded, then $\big(x \mapsto \int_{\mathbb{R}_+} \mathrm{e}^{-\lambda t} g(t,x)\,\mathrm{d}t\big) \in \mathscr{E}$ by the usual progression starting from g being an indicator).

Clearly, $f \geqslant 0$ implies $R_\lambda f \geqslant 0$ and $\forall f \in B(E)$ $\|R_\lambda f\| \leqslant \|f\|/\lambda$. We also have the ***resolvent equation***

$$\lambda \neq \mu \quad \Longrightarrow \quad R_\lambda R_\mu = -\frac{R_\lambda - R_\mu}{\lambda - \mu}.$$

To see this, we first show

Lemma. $\forall \mu > 0\ \ \forall t \geqslant 0\ \ Q_t R_\mu = R_\mu Q_t$.

Proof. We show this lemma by direct calculation:

$$\begin{aligned}
(Q_t R_\mu f)(x) &= \int_E Q_t(x, \mathrm{d}y)\, R_\mu f(y) \\
&= \int_E Q_t(x, \mathrm{d}y) \int_0^\infty \mathrm{e}^{-\mu s} Q_s f(y)\, \mathrm{d}s \\
&= \int_0^\infty \mathrm{e}^{-\mu s} \int_E Q_t(x, \mathrm{d}y)\, Q_s f(y)\, \mathrm{d}s \\
&= \int_0^\infty \mathrm{e}^{-\mu s} Q_t(Q_s f)(x)\, \mathrm{d}s \\
&= \int_0^\infty \mathrm{e}^{-\mu s} Q_s(Q_t f)(x)\, \mathrm{d}s = (R_\mu Q_t f)(x)
\end{aligned}$$

◂

Now, using the above lemma, we can verify the resolvent equation:

$$\begin{aligned}
(R_\lambda R_\mu f)(x) &= \int_0^\infty \mathrm{e}^{-\lambda t} (Q_t R_\mu f)(x)\, \mathrm{d}t \\
&= \int_0^\infty \mathrm{e}^{-\lambda t} (R_\mu Q_t f)(x)\, \mathrm{d}t \\
&= \int_0^\infty \mathrm{e}^{-\lambda t} \int_0^\infty \mathrm{e}^{-\mu s} Q_s(Q_t f)(x)\, \mathrm{d}s\, \mathrm{d}t \\
&= \int_0^\infty \mathrm{e}^{-\lambda t} \int_0^\infty \mathrm{e}^{-\mu s} Q_{t+s} f(x)\, \mathrm{d}s\, \mathrm{d}t \\
&= \int_0^\infty \mathrm{e}^{-\lambda t} \mathrm{e}^{\mu t} \int_t^\infty \mathrm{e}^{-\mu r} Q_r f(x)\, \mathrm{d}r\, \mathrm{d}t \\
&= \int_0^\infty \mathrm{e}^{-\mu r} Q_r f(x) \int_0^r \mathrm{e}^{-(\lambda-\mu)t}\, \mathrm{d}t\, \mathrm{d}r \\
&\quad \hookrightarrow \text{[Fubini's theorem]} \\
&= \int_0^\infty \mathrm{e}^{-\mu r} Q_r f(x) \Big(\frac{1 - \mathrm{e}^{-(\lambda-\mu)r}}{\lambda - \mu} \Big)\, \mathrm{d}r \\
&= \int_0^\infty Q_r f(x) \Big(\frac{\mathrm{e}^{-\mu r} - \mathrm{e}^{-\lambda r}}{\lambda - \mu} \Big)\, \mathrm{d}r \\
&= \frac{R_\mu f(x) - R_\lambda f(x)}{\lambda - \mu}.
\end{aligned}$$

Example. For real Brownian motion, we have

$$R_\lambda f(x) = \int_{\mathbb{R}} r_\lambda(y - x) f(y)\, \mathrm{d}y,$$

where

$$r_\lambda(z) := \frac{1}{\sqrt{2\lambda}} \exp\{-|z|\sqrt{2\lambda}\}.$$

See page 157 of the book for a proof.

The resolvent provides useful supermartingales:

Lemma 6.6. *Let X be a Markov process with semigroup $(Q_t)_{t\geqslant 0}$ and filtration $(\mathscr{F}_t)$, $0 \leqslant h \in B(E)$, and $\lambda > 0$. Then*

$$t \mapsto \mathrm{e}^{-\lambda t} R_\lambda h(X_t)$$

is an $(\mathscr{F}_t)$-supermartingale.

Proof. Since $R_\lambda h \in B(E)$, integrability of the random variables $\mathrm{e}^{-\lambda t} R_\lambda h(X_t)$ is ensured. We want to bound, for $s, t \geqslant 0$,

$$\mathbf{E}\Big[\mathrm{e}^{-\lambda(t+s)} R_\lambda h(X_{t+s}) \;\Big|\; \mathscr{F}_t\Big] \underset{\text{[Definition 6.2]}}{=} \mathrm{e}^{-\lambda(t+s)} Q_s R_\lambda h(X_t).$$

It suffices to show that $\mathrm{e}^{-\lambda s} Q_s R_\lambda h \leqslant R_\lambda h$. Indeed, we have

$$\mathrm{e}^{-\lambda s} Q_s R_\lambda h \underset{\text{[by the Lemma for resolvent equation]}}{=} \int_s^\infty \mathrm{e}^{-\lambda t} Q_t h \,\mathrm{d}t \overset{[h \geqslant 0]}{\leqslant} R_\lambda h. \qquad ◀$$

6.2. Feller Semigroups

A topological space is ***locally compact*** if every point has a neighborhood with compact closure. A locally compact Polish space has the property that there exist compact K_n such that for all n, $K_n \subseteq K_{n+1}$ and every compact set is contained in some K_n. See the appendix to this chapter.

In the rest of this section, let E be locally compact and Polish. We say $f\colon E \to \mathbb{R}$ ***vanishes at infinity*** if for all $\varepsilon > 0$, there exists a compact set K such that $|f(x)| < \varepsilon$ for $x \notin K$. Write $C_0(E)$ for the continuous functions that vanish at infinity, and give it the supremum norm.

Definition 6.7. *A transition semigroup $(Q_t)_{t\geqslant 0}$ on E is called **Feller** if*

(i) $\forall t \geqslant 0 \;\; \forall f \in C_0(E) \;\; Q_t f \in C_0(E)$, *and*

(ii) $\forall f \in C_0(E) \;\; \lim_{t\to 0} \|Q_t f - f\| = 0$.

*A Markov process with a Feller semigroup is called **Feller**.*

Part (i) says that $Q_t(x, \cdot)$ depends continuously on x and for all compact K, $Q_t(\cdot, K)$ vanishes at infinity. Part(ii) says that $Q_t(x, \cdot)$ has most of its mass near x for small t. Be aware that different authors use different definitions for "Feller".

We also see that a Feller semigroup has the property that $\forall f \in C_0(E) \;\; t \mapsto Q_t f$ is uniformly continuous:

$$\forall s, t \geqslant 0 \quad \|Q_s f - Q_t f\| = \big\|Q_{s\wedge t}(Q_{|s-t|} f - f)\big\| \leqslant \|Q_{|s-t|} f - f\|.$$

Lebesgue's dominated convergence theorem gives that $\forall f \in C_0 E \;\; \forall \lambda > 0 \;\; R_\lambda f \in C_0(E)$.

For the rest of the section, let $(Q_t)_{t\geqslant 0}$ be a Feller semigroup on E.

We are now going to show how R_λ is an inverse. However, the operator of which it is the inverse is not defined on all of $C_0(E)$, but only a dense subspace, which we can get as the range of R_λ and which does not depend on λ.

Proposition 6.8. *For $\lambda > 0$, let $\mathcal{R} := \{R_\lambda f\,;\ f \in C_0(E)\}$. Then $\mathcal{R}$ does not depend on λ and is dense in $C_0(E)$.*

Proof. We can write the resolvent equation this way:

$$R_\lambda f = R_\mu\big(f + (\mu - \lambda)R_\lambda f\big).$$

Therefore, every $R_\lambda f$ has the form $R_\mu g$ for some $g \in C_0(E)$, as desired.

To show $\mathcal{R}$ is dense, we write

$$\lambda R_\lambda f = \lambda \int_0^\infty \mathrm{e}^{-\lambda t} Q_t f \,\mathrm{d}t = \int_0^\infty \mathrm{e}^{-t} Q_{t/\lambda} f \,\mathrm{d}t,$$

so

$$\begin{aligned} \|\lambda R_\lambda f - f\| &= \left\| \int_0^\infty \mathrm{e}^{-t}(Q_{t/\lambda} f - f)\,\mathrm{d}t \right\| \\ &\leqslant \int_0^\infty \mathrm{e}^{-t}\|Q_{t/\lambda} f - f\|\,\mathrm{d}t \to 0 \end{aligned}$$

by Lebesgue's dominated convergence theorem. ◂

If $Q_t = \mathrm{e}^{tL}$, then $L = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} Q_t$. This motivates

Definition 6.9. *Write*

$$D(L) := \left\{ f \in C_0(E)\,;\ \frac{Q_t f - f}{t} \text{ converges in } C_0(E) \text{ as } t \downarrow 0 \right\}$$

and, for $f \in D(L)$,

$$Lf := \lim_{t \downarrow 0} \frac{Q_t f - f}{t}.$$

The set $D(L)$ is the ***domain*** *of the* ***(infinitesimal) generator*** *L of $(Q_t)_{t \geqslant 0}$.*

Of course, $D(L)$ is a linear subspace and L is a linear map from $D(L)$ to $C_0(E)$.

We can express differential and integral equations for $(Q_t)_{t\geqslant 0}$ at times other than $t = 0$:

Proposition 6.10. *If $f \in D(L)$ and $s > 0$, then $Q_s f \in D(L)$ with*

$$L(Q_s f) = Q_s(Lf).$$

Proof. We have, for $t > 0$,

$$\frac{Q_t(Q_s f) - Q_s f}{t} = Q_s\Big(\frac{Q_t f - f}{t}\Big). \tag{$*$}$$

Since every Markovian kernel is a contraction, the right-hand side converges to $Q_s(Lf)$ in $C_0(E)$. ◂

Exercise (due 3/8). Prove that

$$\forall s > 0 \ \ \forall f \in D(L) \qquad \lim_{t \uparrow 0} \frac{Q_{t+s} f - Q_s f}{t} = LQ_s f \quad \text{in } C_0(E).$$

Proposition 6.11. *For $f \in D(L)$ and $t \geqslant 0$, we have*

$$Q_t f = f + \int_0^t Q_s(Lf)\,\mathrm{d}s = f + \int_0^t L(Q_s f)\,\mathrm{d}s.$$

Proof. Another way of writing ($*$), Proposition 6.10, and the exercise is that for all $x \in E$, $t \mapsto Q_t f(x)$ has a derivative $Q_t L f(x)$, which is a continuous function of t. [Indeed, repeating shows that $\big(t \mapsto Q_t f(x)\big) \in C^\infty(\mathbb{R}_+)$.] Thus, the result follows from the fundamental theorem of calculus. ◀

We are ready to justify the name "resolvent".

Proposition 6.12. *Let $\lambda > 0$. Then $D(L) = \mathcal{R}$ and*

$$R_\lambda \colon C_0(E) \to D(L), \quad \lambda - L \colon D(L) \to C_0(E)$$

are inverses. That is,

(i) $\forall g \in C_0(E)\ \ R_\lambda g \in D(L)$ *and* $(\lambda - L)R_\lambda g = g$, *and*

(ii) $\forall f \in D(L)\ \ R_\lambda(\lambda - L)f = f$.

Proof. (i) We want to show that $R_\lambda g \in D(L)$ with $LR_\lambda g = \lambda R_\lambda g - g$. We calculate for all $\varepsilon > 0$,

[1st term by Lemma for resolvent equation;
decompose the 2nd term to get the next step]

$$\begin{aligned} \varepsilon^{-1}(Q_\varepsilon R_\lambda g - R_\lambda g) &= \varepsilon^{-1}\Big(\int_0^\infty \mathrm{e}^{-\lambda t} Q_{\varepsilon+t} g\,\mathrm{d}t - \int_0^\infty \mathrm{e}^{-\lambda t} Q_t g\,\mathrm{d}t\Big) \\ &= \underset{\substack{\downarrow \\ \lambda}}{\frac{1 - \mathrm{e}^{-\lambda\varepsilon}}{\varepsilon}} \ \underset{\substack{\downarrow \\ R_\lambda g}}{\int_0^\infty \mathrm{e}^{-\lambda t} Q_{\varepsilon+t} g\,\mathrm{d}t} \ - \ \underset{\substack{\downarrow \\ g}}{\frac{1}{\varepsilon}\int_0^\varepsilon \mathrm{e}^{-\lambda t} Q_t g\,\mathrm{d}t}\ , \end{aligned}$$

the last two convergences as $\varepsilon \downarrow 0$ being in $C_0(E)$.

This also shows that $\mathcal{R} \subseteq D(L)$.

(ii) Now we want $\lambda R_\lambda f = f + R_\lambda L f$. Using Proposition 6.11, we get

$$\begin{aligned} \lambda R_\lambda f &= \lambda \int_0^\infty \mathrm{e}^{-\lambda t} Q_t f\,\mathrm{d}t \\ &= \lambda \int_0^\infty \mathrm{e}^{-\lambda t}\Big(f + \int_0^t Q_s L f\,\mathrm{d}s\Big)\,\mathrm{d}t \\ &= f + \int_0^\infty \mathrm{e}^{-\lambda s} Q_s L f\,\mathrm{d}s = f + R_\lambda L f. \end{aligned}$$

This also shows that $D(L) \subseteq \mathcal{R}$. ◀

Exercise (due 3/8). Show that if $f_n, f, g \in C_0(E)$ and $f_n \to f$ and $Lf_n \to g$ in $C_0(E)$, then $g = Lf$.

Corollary 6.13. *The map $L \colon D(L) \to C_0(E)$ determines $(Q_t)_{t \geqslant 0}$.*

This justifies the name "generator".

Proof. Given L, we know R_λ for each $\lambda > 0$, whence we know the Laplace transform of $t \mapsto Q_t f(x)$ for each $f \in C_0(E)$ and $x \in E$. The uniqueness of the Laplace transform shows that we then know $Q_t f(x)$. Since $Q_t \restriction C_0(E)$ determines Q_t (Riesz representation theorem—regularity gives uniqueness), this completes the proof. ◂

Exercise (due 3/8). Fix $a \in \mathbb{R} \setminus \{0\}$. Let $Q_t(x, \cdot) := \delta_{x+at}$.

(1) Show that $(Q_t)_{t\geqslant 0}$ is a Feller semigroup on $\mathbb{R}$.
(2) Given a probability measure γ on $\mathbb{R}$, find a Markov process with semigroup $(Q_t)_{t\geqslant 0}$ and initial distribution γ.
(3) Find the generator of $(Q_t)_{t\geqslant 0}$ and its domain.

It is easy to see that the semigroup $(Q_t)_t$ of Brownian motion is Feller. It is also intuitive that its generator is $Lf = \frac{1}{2} f''$ in some sense:

$$\begin{aligned} Q_t f(x) - f(x) &= \mathbf{E}_x f(X_t) - f(x) \\ &= \mathbf{E}_x\Big[f'(x)(X_t - x) + \frac{1}{2} f''(\xi_t)(X_t - x)^2\Big] \\ &\approx \frac{t}{2} f''(x) \end{aligned}$$

for some ξ_t between x and X_t, except that we would need ξ_t to be measurable for this argument to work.

Instead, we use the resolvent. We saw that

$$\forall \lambda > 0 \;\; \forall f \in C_0(\mathbb{R}) \qquad R_\lambda f(x) = \int \frac{1}{\sqrt{2\lambda}} \exp\big\{-\sqrt{2\lambda}|y - x|\big\} f(y)\, \mathrm{d}y.$$

Take $\lambda := \frac{1}{2}$. If $h \in D(L)$, then $\exists f \in C_0(\mathbb{R})$ such that $h = R_{\frac{1}{2}} f$ and $f = (\frac{1}{2} - L)h$. If we differentiate

$$h(x) = \int \mathrm{e}^{-|y-x|} f(y)\, \mathrm{d}y$$

twice (see page 161 of the book for details), we get

$$\begin{aligned} h''(x) &= \int (-2\delta_x + \mathrm{e}^{-|y-x|}) f(y)\, \mathrm{d}y \text{ (informally)} \\ &= -2f(x) + h(x) \quad (\in C_0(\mathbb{R})) \\ &= 2Lh \text{ (by above).} \end{aligned}$$

In particular, $D(L) \subseteq \big\{h \in C^2(\mathbb{R})\,;\; h, h'' \in C_0(\mathbb{R})\big\}$.

In fact, this equals $D(L)$. If $g \in C^2(\mathbb{R})$ with $g, g'' \in C_0(\mathbb{R})$, then set

$$f := (g - g'')/2 \in C_0(\mathbb{R})$$

and $h := R_{\frac{1}{2}} f \in D(L)$.We saw that $h \in C^2(\mathbb{R})$ and $h'' = -2f + h$, i.e., $(h - h'')/2 = f$. Therefore, $(h - g)'' = h - g$. Since $h - g \in C_0(\mathbb{R})$, it follows that $h - g = 0$. Thus, $g = h \in D(L)$, as desired.

Exercise (due 3/8). Exercise 6.23.

Exercise. We call a probability measure γ ***stationary*** if for all $f \in C_0(E)$ and all $t \geqslant 0$, we have $\int Q_t f \, \mathrm{d}\gamma = \int f \, \mathrm{d}\gamma$. Show that γ is stationary iff for all $f \in D(L)$, we have $\int Lf \, \mathrm{d}\gamma = 0$.

It is usually very difficult to determine $D(L)$ exactly, but we can find a subset of $D(L)$ by using the following martingales.

Theorem 6.14. *Suppose that for all $x \in E$, there is a càdlàg process X that is Markov with semigroup (Q_t) for the probability measure $\mathbf{P}_x$. Let $h, g \in C_0(E)$. The following are equivalent:*

(i) $h \in D(L)$ *and* $Lh = g$.

(ii) $\forall x \in E$

$$t \mapsto h(X_t) - \int_0^t g(X_s)\,\mathrm{d}s$$

is a $\mathbf{P}_x$-martingale.

(iii) $\forall x \in E \ \ \forall t \geqslant 0 \ \ \mathbf{E}_x\big[h(X_t) - \int_0^t g(X_s)\,\mathrm{d}s\big] = h(x)$.

Proof. Assume (i). We have

$$\begin{aligned}
&\mathbf{E}_x\Big[h(X_{t+s}) - \int_0^{t+s} g(X_r)\,\mathrm{d}r \;\Big|\; \mathscr{F}_t\Big] \\
&\quad = \mathbf{E}_x[h(X_{t+s}) \mid \mathscr{F}_t] - \int_0^t g(X_r)\,\mathrm{d}r - \mathbf{E}_x\Big[\int_t^{t+s} g(X_r)\,\mathrm{d}r \;\Big|\; \mathscr{F}_t\Big] \\
&\quad = Q_s h(X_t) - \int_0^t g(X_r)\,\mathrm{d}r - \int_t^{t+s} \mathbf{E}_x[g(X_r) \mid \mathscr{F}_t]\,\mathrm{d}r \\
&\qquad \hookrightarrow \textbf{[}\text{for the third term, take } \mathbf{E}[\mathbf{1}_A \cdots] \text{ for } A \in \mathscr{F}_t\textbf{]} \\
&\quad = Q_s h(X_t) - \int_0^t g(X_r)\,\mathrm{d}r - \int_t^{t+s} Q_{r-t} g(X_t)\,\mathrm{d}r \\
&\quad = Q_s h(X_t) - \int_0^t g(X_r)\,\mathrm{d}r - \int_0^s Q_r g(X_t)\,\mathrm{d}r \\
&\quad = h(X_t) - \int_0^t g(X_r)\,\mathrm{d}r
\end{aligned}$$

because $Q_s h = h + \int_0^s Q_r g \,\mathrm{d}r$ by Proposition 6.11. This gives (ii).

Obviously, (ii) implies (iii).

Assume (iii). We also have

$$\mathbf{E}_x\Big[h(X_t) - \int_0^t g(X_s)\,\mathrm{d}s\Big] = Q_t h(x) - \int_0^t Q_s g(x)\,\mathrm{d}s.$$

Therefore,

$$\frac{Q_t h - h}{t} = \frac{1}{t}\int_0^t Q_r g\,\mathrm{d}r,$$

which converges to g in $C_0(E)$. ◂

For example, consider d-dimensional Brownian motion. By Itô's formula,

$$\forall h \in C^2(\mathbb{R}^d) \quad t \mapsto h(X_t) - \frac{1}{2}\int_0^t \Delta h(X_s)\,\mathrm{d}s$$

is a continuous local martingale. If h and Δh are bounded, then this is a true martingale. In particular, this holds if $h, \Delta h \in C_0(\mathbb{R}^d)$. Therefore, Theorem 6.14 tells us that

$$D(L) \supseteq \left\{h \in C^2(\mathbb{R}^d)\,;\; h, \Delta h \in C_0(\mathbb{R}^d)\right\}$$

and $Lh = \frac{1}{2}\Delta h$ for such h. Equality does *not* hold for $d \geqslant 2$, where, in fact,

$$D(L) = \left\{h \in C_0(\mathbb{R}^d)\,;\; \Delta h \in C_0(\mathbb{R}^d) \text{ in the sense of distributions}\right\}$$

(see page 288 of the book by Revuz and Yor).

Example. If $d = 2$, let $h(x) := \frac{x_1^2 - x_2^2}{|x| \log |x|}$. For $d > 2$, multiply this by a smooth function. See B. Epstein, *Partial Differential Equations*, pp. 162–163.

Exercise (due 3/8). Exercise 6.27 (note the hypotheses on page 180 of the book).

6.3. The Regularity of Sample Paths

Let E be a locally compact Polish space and $(Q_t)_{t\geqslant 0}$ be a Feller semigroup on E.

Theorem 6.15. *Suppose (X_t) is a process and $(\mathbf{P}_x)_{x\in E}$ are probability measures such that $\forall x \in E$ $(X_t)_{t\geqslant 0}$ is a $\mathbf{P}_x$-Markov process with semigroup $(Q_t)_{t\geqslant 0}$ with respect to $(\mathscr{F}_t)_t$ and $\mathbf{P}_x[X_0 = x] = 1$. Set $\widetilde{\mathscr{F}}_\infty := \mathscr{F}_\infty$ and $\forall t \geqslant 0$ $\widetilde{\mathscr{F}}_t := \mathscr{F}_{t^+} \vee \sigma(\mathscr{N})$, where*

$$\mathscr{N} := \left\{A \in \mathscr{F}_\infty\,;\; \forall x \in E \;\; \mathbf{P}_x(A) = 0\right\}.$$

Then there exists a process $(\widetilde{X}_t)$ that is càdlàg, adapted to $(\widetilde{\mathscr{F}}_t)_t$, and for all probability measures γ on E, $(\widetilde{X}_t)_{t\geqslant 0}$ is a $\mathbf{P}_{(\gamma)}$-modification of $(X_t)_{t\geqslant 0}$, $\mathbf{P}_{(\gamma)}$-Markov with semigroup $(Q_t)_{t\geqslant 0}$ with respect to $(\widetilde{\mathscr{F}}_t)_t$, and $\forall A \in \mathscr{E}$ $\mathbf{P}_{(\gamma)}[\widetilde{X}_0 \in A] = \gamma(A)$.

Remark. The hypothesis implies that X itself is $\mathbf{P}_{(\gamma)}$-Markov, etc., for all γ on E.

Exercise (due 3/22). Show that $(\widetilde{\mathscr{F}}_t)_t$ is right-continuous. *Hint:* check that for all $B \in \widetilde{\mathscr{F}}_t$, there exists $C \in \mathscr{F}_{t^+}$ and $N \in \mathscr{N}$ such that $B = C \triangle N$.

Lemma. *Let Y be a right-continuous, nonnegative supermartingale with respect to a right-continuous filtration. Define*

$$T := \inf\{t \geqslant 0\,;\; Y_t = 0\} \wedge \inf\{t > 0\,;\; Y_{t^-} = 0\}.$$

Then

$$\mathbf{P}\Big[\forall t \in [T, \infty) \;\; Y_t = 0\Big] = 1.$$

Remark. The right-continuity of the filtration is not necessary.

Proof. By Proposition 3.9(i),

$$T_n := \inf\{t \geqslant 0\,;\ Y_t < 1/n\}$$

is a stopping time. By property (g) of stopping times in Chapter 3, $T := \lim_{n\to\infty} T_n$ is also a stopping time. We are concerned only with what happens on $[T < \infty]$. For $0 < q \in \mathbb{Q}$, apply Theorem 3.25 to the stopping times $T_n < T + q$:

$$\mathbf{E}\big[Y_{T+q}\mathbf{1}_{[T<\infty]}\big] \leqslant \mathbf{E}\big[Y_{T_n}\mathbf{1}_{[T_n<\infty]}\big] \leqslant \tfrac{1}{n}.$$

This gives $\mathbf{E}\big[Y_{T+q}\mathbf{1}_{[T<\infty]}\big] = 0$, so $Y_{T+q} = 0$ almost surely on $[T < \infty]$. By right-continuity, we get the result. ◂

Proof of Theorem 6.15. If E is not compact, then let

$$E_\Delta := E \cup \{\Delta\}$$

be its one-point compactification; otherwise, let $E_\Delta := E$. Every function in $C_0(E)$ extends to a function in $C(E_\Delta)$ by defining it to be 0 at Δ.

Step 1: (define $\widetilde{X}$ on E_Δ)

Let $(f_n)_{n\geqslant 0}$ be a sequence of nonnegative functions in $C_0(E)$ that separates points of E_Δ, i.e., $\forall x \neq y \in E_\Delta\ \exists n\ \ f_n(x) \neq f_n(y)$. Let

$$\mathcal{H} := \big\{R_p f_n\,;\ p \in \mathbb{N}^+,\, n \in \mathbb{N}\big\}.$$

Then $\mathcal{H}$ also separates points of E_Δ because $\lim_{p\to\infty}\|pR_p f_n - f_n\| = 0$, as we saw in the proof of Proposition 6.8.

For $h \in \mathcal{H}$ with $h = R_p f_n$, the process $\big(\mathrm{e}^{-pt}h(X_t)\big)_{t\geqslant 0}$ is, for all x, a $\mathbf{P}_x$-supermartingale by Lemma 6.6. Let N_h be the event that for some $k \in \mathbb{N}$ and some $a, b \in \mathbb{Q}$ with $a < b$, $\big(\mathrm{e}^{-ps}h(X_s)\big)_{s\in\mathbb{Q}_+\cap[0,k]}$ makes an infinite number of upcrossings of $[a, b]$. In the proof of Theorem 3.17, we saw that $\mathbf{P}_x(N_h) = 0$. Put $N := \bigcup_{h\in\mathcal{H}} N_h \in \mathscr{N}$. Then for all γ, $\mathbf{P}_{(\gamma)}(N) = 0$, and

$$\forall \omega \notin N\ \ \forall h \in \mathcal{H}\ \ \forall t \geqslant 0 \qquad \lim_{\mathbb{Q}_+\ni s\downarrow t} h\big(X_s(\omega)\big) \text{ exists}$$

and

$$\forall \omega \notin N\ \ \forall h \in \mathcal{H}\ \ \forall t > 0 \qquad \lim_{\mathbb{Q}_+\ni s\uparrow t} h\big(X_s(\omega)\big) \text{ exists.}$$

Because $\mathcal{H}$ separates points, it follows that $\forall \omega \notin N\ \ \big(X_s(\omega)\big)_{s\in\mathbb{Q}_+}$ has right and left limits in E_Δ (not necessarily in E).

Thus, we may define

$$\forall \omega \notin N\ \ \forall t \geqslant 0 \qquad \widetilde{X}_t(\omega) := \lim_{\mathbb{Q}_+\ni s\downarrow t} X_s(\omega).$$

If $\omega \in N$, put $\widetilde{X}_t(\omega) := x_0$ for some fixed $x_0 \in E$ and all $t \geqslant 0$. Then $\widetilde{X}$ is E_Δ-valued and $(\mathscr{F}_t)_t$-adapted. Lemma 3.16 shows that $\big(h(\widetilde{X}_t)\big)_t$ is càdlàg for all $h \in \mathcal{H}$, whence so is $\widetilde{X}$.

Step 2: (show $\forall t \geqslant 0\ \forall\gamma\ \ \mathbf{P}_{(\gamma)}\big[X_t = \widetilde{X}_t\big] = 1$)
Let $t \geqslant 0$. For all $f, g \in C_0(E)$, we have

$$\begin{aligned}
\mathbf{E}_{(\gamma)}\big[f(X_t)g(\widetilde{X}_t)\big] &= \lim_{\mathbb{Q}\ni s\downarrow t} \mathbf{E}_{(\gamma)}\big[f(X_t)g(X_s)\big] \\
&\qquad \text{[bounded convergence theorem]} \\
&= \lim \mathbf{E}_{(\gamma)}\big[f(X_t)\,\mathbf{E}_{(\gamma)}[g(X_s) \mid \mathscr{F}_t]\big] \\
&= \lim \mathbf{E}_{(\gamma)}\big[f(X_t)Q_{s-t}g(X_t)\big] \\
&= \mathbf{E}_{(\gamma)}\big[f(X_t)g(X_t)\big]. \\
&\qquad \text{[Feller property and} \\
&\qquad \text{bounded convergence theorem]}
\end{aligned}$$

As in Exercise 6.27, this means that $(X_t, \widetilde{X}_t) \overset{\mathscr{D}}{=} (X_t, X_t)$ under $\mathbf{P}_{(\gamma)}$, whence we have $\mathbf{P}_{(\gamma)}\big[X_t = \widetilde{X}_t\big] = 1$.

Step 3: (show that $\forall\gamma\ \ \widetilde{X}$ is $\mathbf{P}_{(\gamma)}$-Markov with semigroup $(Q_t)_{t\geqslant 0}$ with respect to $(\widetilde{\mathscr{F}}_t)$)
We want to verify that

$$\forall s \geqslant 0\ \ \forall t > 0\ \ \forall f \in B(E) \qquad \mathbf{E}_{(\gamma)}\big[f(\widetilde{X}_{s+t}) \,\big|\, \widetilde{\mathscr{F}}_s\big] = Q_t f(\widetilde{X}_s),$$

i.e.,

$$\forall A \in \widetilde{\mathscr{F}}_s \qquad \mathbf{E}_{(\gamma)}\big[\mathbf{1}_A f(\widetilde{X}_{s+t})\big] = \mathbf{E}_{(\gamma)}\big[\mathbf{1}_A Q_t f(\widetilde{X}_s)\big].$$

By regarding each side as a linear functional on $B(E)$, we see that it suffices to establish the equality for $f \in C_0(E)$. In addition, since s and t are fixed, we may replace $\widetilde{X}_{s+t}$ by X_{s+t}. Furthermore, we may assume $A \in \mathscr{F}_{s^+}$. Then for $r \in \mathbb{Q} \cap (s, s+t)$,

$$\begin{aligned}
\mathbf{E}_{(\gamma)}\big[\mathbf{1}_A f(X_{s+t})\big] &= \mathbf{E}_{(\gamma)}\Big[\mathbf{1}_A\,\mathbf{E}_{(\gamma)}\big[f(X_{s+t}) \,\big|\, \mathscr{F}_r\big]\Big] \\
&= \mathbf{E}_{(\gamma)}\big[\mathbf{1}_A Q_{s+t-r} f(X_r)\big].
\end{aligned}$$

Since $\|Q_{s+t-r}f - Q_t f\| \to 0$ as $r \downarrow s$ and $X_r \to \widetilde{X}_s\ \ \mathbf{P}_{(\gamma)}$ -a.s. as $r \downarrow s$, we get the result.

Step 4: ($\widetilde{X}$ is càdlàg as an E-valued process off another set in $\mathscr{N}$; this step is not needed if E is compact)

Note that $\widetilde{X}$ being càdlàg in E_Δ and Step 2 do not ensure this (even if $\widetilde{X} = X$).
Choose $0 < g \in C_0(E)$, put $C_0(E) \ni h := R_1 g > 0$, and

$$Y_t := \mathrm{e}^{-t} h(\widetilde{X}_t).$$

Then Y is a nonnegative $(\widetilde{\mathscr{F}}_t)$-supermartingale by Lemma 6.6. Also, Y is càdlàg (recall that $h(\Delta) := 0$). Define T as in the lemma. Let

$$N_1 := \big[\exists t \in [T, \infty)\ \ Y_t \neq 0\big].$$

By the lemma, $N_1 \in \mathscr{N}$. Let

$$N_2 := \big[\exists k \in \mathbb{N}\ \ X_k \neq \widetilde{X}_k\big];$$

by Step 2, $N_2 \in \mathscr{N}$. Off $N_1 \cup N_2$, we have $T = \infty$ because $Y_t = 0$ iff $\widetilde{X}_t = \Delta$, and $X_k \neq \Delta$. That is, off $N_1 \cup N_2$, $\widetilde{X}$ is E-valued. ◀

6.4. The Strong Markov Property

Theorem 6.16 (Simple Markov Property)**.** *Let E be a measurable space, $(X_t)_{t\geqslant 0}$ be an E-valued process, and $(\mathbf{P}_x)_{x\in E}$ be probability measures such that $\forall x \in E\ \ (X_t)_{t\geqslant 0}$ is a $\mathbf{P}_x$-Markov process with semigroup $(Q_t)_{t\geqslant 0}$ with respect to $(\mathscr{F}_t)_{t\geqslant 0}$ and $\mathbf{P}_x[X_0 = x] = 1$. Let γ be a probability measure on E. Let*

$$\Phi\colon E^{\mathbb{R}_+} \longrightarrow \mathbb{R}_+$$

be measurable. Then

$$\forall s \geqslant 0 \qquad \mathbf{E}_{(\gamma)}\big[\Phi\big((X_{s+t})_{t\geqslant 0}\big) \bigm| \mathscr{F}_s\big] = \mathbf{E}_{X_s}[\Phi],$$

where $\mathbf{E}_{X_s}[\Phi]$ denotes the composition of $\omega \mapsto X_s(\omega)$ and $x \mapsto \mathbf{E}_x\big[\Phi(X)\big]$.

Proof. We saw in Section 6.1 that $x \mapsto \mathbf{P}_x$ is measurable, whence so is $x \mapsto \mathbf{E}_x\big[\Phi(X)\big]$.

To prove the theorem, it suffices to prove the case when Φ is an indicator of an elementary cylinder set, or, more generally,

$$\Phi(X) = \prod_{i=1}^{p} \varphi_i(X_{s+t_i}), \quad 0 \leqslant t_1 < t_2 < \cdots < t_p, \quad \varphi_i \in B(E).$$

The proof is just like that of (∗∗∗), but with an extra conditioning: We want

the left-hand side of the conclusion of the theorem

$$= \int Q_{t_1}(X_s, \mathrm{d}x_1)\varphi_1(x_1) \int Q_{t_2-t_1}(x_1, \mathrm{d}x_2)\varphi_2(x_2) \cdots \int Q_{t_p-t_{p-1}}(x_{p-1}, \mathrm{d}x_p)\varphi_p(x_p).$$

For $p = 1$, this is Definition 6.2 (of a Markov process with semigroup). The induction step is: the left-hand side of the conclusion of the theorem equals

$$\begin{aligned}\mathbf{E}_{(\gamma)}\Big[\prod_{i=1}^{p-1} \varphi_i(X_{s+t_i}) \cdot \mathbf{E}_{(\gamma)}\big[\varphi_p(X_{s+t_p}) \bigm| \mathscr{F}_{s+t_{p-1}}\big] \Bigm| \mathscr{F}_s\Big]& \\ = \mathbf{E}_{(\gamma)}\Big[\prod_{i=1}^{p-1} \varphi_i(X_{s+t_i}) \cdot Q_{t_p-t_{p-1}}\varphi_p(X_{s+t_{p-1}}) \Bigm| \mathscr{F}_s\Big].&\end{aligned}$$

◀

Exercise (due 3/22)*.* Exercise 6.26 (note that the derivative in part 3 is in the norm sense).

If E is a topological space,we write

$$\mathbb{D}(E) := \big\{f \in E^{\mathbb{R}_+};\ f \text{ is càdlàg}\big\},$$

and give $\mathbb{D}(E)$ the σ-field $\mathscr{D}$ induced from $E^{\mathbb{R}_+}$, even though $\mathbb{D}(E)$ need not be measurable. We call $\mathbb{D}(E)$ the ***Skorokhod space***.

Theorem 6.17 (Strong Markov Property)**.** *Let $(X_t)_{t\geqslant 0}$ be an E-valued càdlàg process, and $(\mathbf{P}_x)_{x\in E}$ be probability measures such that $\forall x \in E$ $(X_t)_{t\geqslant 0}$ is a $\mathbf{P}_x$-Markov process with semigroup $(Q_t)_{t\geqslant 0}$ with respect to $(\mathscr{F}_t)_{t\geqslant 0}$ and $\mathbf{P}_x[X_0 = x] = 1$. Assume that E is locally compact Hausdorff, Q_t maps $C_0(E)$ to $C(E)$ [e.g., E is also Polish and $(Q_t)_{t\geqslant 0}$ is Feller],*

$$\Phi\colon (\mathbb{D}(E), \mathscr{D}) \to \mathbb{R}_+$$

is measurable, and T is an $(\mathscr{F}_{t^+})_t$-stopping time [e.g., T is a stopping time]. Then, for all probability measures γ on E,

$$\mathbf{E}_{(\gamma)}\big[\mathbf{1}_{[T<\infty]}\Phi\big((X_{T+t})_{t\geqslant 0}\big) \bigm| \mathscr{F}_T\big] = \mathbf{1}_{[T<\infty]}\,\mathbf{E}_{X_T}[\Phi].$$

Proof. Theorem 3.7 guarantees that X_T is measurable on $[T < \infty]$, so the right-hand side is $\mathscr{F}_T$-measurable. Thus, it suffices to show that

$$\forall A \in \mathscr{F}_T \qquad \mathbf{E}_{(\gamma)}\big[\mathbf{1}_{A\cap[T<\infty]}\Phi\big((X_{T+t})_{t\geqslant 0}\big)\big] = \mathbf{E}_{(\gamma)}\big[\mathbf{1}_{A\cap[T<\infty]}\,\mathbf{E}_{X_T}[\Phi]\big].$$

As in the proof of Theorem 6.16, it suffices to do this for Φ of the form

$$\Phi(f) = \prod_{i=1}^{p} \varphi_i\big(f(t_i)\big), \quad 0 \leqslant t_1 < t_2 < \cdots < t_p, \quad \varphi_i \in B(E).$$

We again use induction, but this time the case $p = 1$ requires work; the induction step is like before:

$$\begin{aligned}
&\mathbf{E}_{(\gamma)}\Big[\mathbf{1}_{A\cap[T<\infty]}\prod_{i=1}^{p}\varphi_i(X_{T+t_i})\Big]\\
&\quad= \mathbf{E}_{(\gamma)}\Big[\mathbf{1}_{A\cap[T<\infty]}\prod_{i=1}^{p-1}\varphi_i(X_{T+t_i})\cdot \mathbf{E}_{(\gamma)}\big[\varphi_p(X_{T+t_p}) \bigm| \mathscr{F}_{T+t_{p-1}}\big]\Big]\\
&\quad= \mathbf{E}_{(\gamma)}\Big[\mathbf{1}_{A\cap[T<\infty]}\prod_{i=1}^{p-1}\varphi_i(X_{T+t_i})\cdot Q_{t_p-t_{p-1}}\varphi_p(X_{T+t_{p-1}})\Big].
\end{aligned}$$

So it remains to prove that

$$\begin{aligned}
&\forall t \geqslant 0 \;\; \forall \varphi \in B(E) \;\; \forall A \in \mathscr{F}_T\\
&\qquad \mathbf{E}_{(\gamma)}\big[\mathbf{1}_{A\cap[T<\infty]}\varphi(X_{T+t})\big] = \mathbf{E}_{(\gamma)}\big[\mathbf{1}_{A\cap[T<\infty]}Q_t\varphi(X_T)\big].
\end{aligned}$$

Because E is locally compact Hausdorff, it suffices to prove this for all $\varphi \in C_0(E)$ [regularity ensures uniqueness again]. Write

$$T_n := \lfloor nT + 1\rfloor / n.$$

Then

$$
\begin{aligned}
\mathbf{E}_{(\gamma)}\big[\mathbf{1}_{A\cap[T<\infty]}\varphi(X_{T+t})\big] &= \lim_{n\to\infty}\mathbf{E}_{(\gamma)}\big[\mathbf{1}_{A\cap[T<\infty]}\varphi(X_{T_n+t})\big] \\
&\qquad \text{[right-continuity]} \\
&= \lim_{n\to\infty}\sum_{i=1}^{\infty}\mathbf{E}_{(\gamma)}\big[\mathbf{1}_{A\cap[(i-1)/n\leqslant T<i/n]}\varphi(X_{\frac{i}{n}+t})\big] \\
&= \lim_{n\to\infty}\sum_{i=1}^{\infty}\mathbf{E}_{(\gamma)}\big[\mathbf{1}_{A\cap[(i-1)/n\leqslant T<i/n]}Q_t\varphi(X_{i/n})\big] \\
&\qquad \text{[conditioning on } \mathscr{F}_{i/n}\text{]} \\
&= \lim_{n\to\infty}\mathbf{E}_{(\gamma)}\big[\mathbf{1}_{A\cap[T<\infty]}Q_t\varphi(X_{T_n})\big] \\
&= \mathbf{E}_{(\gamma)}\big[\mathbf{1}_{A\cap[T<\infty]}Q_t\varphi(X_T)\big],
\end{aligned}
$$

because X is right-continuous and $Q_t\varphi$ is continuous. ◂

The formulation of the strong Markov property for Brownian motion, Theorem 2.20, was different, though equivalent, because it used that Brownian motion has independent, stationary increments.

Exercise (due 3/29). Exercise 6.25, Exercise 6.29.

Exercise. Derive Eq. (3.7), the Laplace transform of the hitting time for Brownian motion, from Dynkin's formula, Exercise 6.29(3).

Appendix: Locally Compact Polish Spaces are σ-compact

We prove the standard result that if E is a locally compact Polish space, then there is an increasing sequence $(K_n)_{n\geqslant 1}$ of compact subsets of E such that every compact subset of E is contained in some K_n.

First note that every separable metric space is second countable, i.e., has a countable basis for its topology. Indeed, if D is a countable dense set, then the balls centered at points of D with rational radii form a countable basis.

Second, we claim that every second countable, locally compact space has a countable basis each of whose members has compact closure. To see this, let $\mathcal{U}$ be a countable basis. Write $\mathcal{U}'$ for the collection of members of $\mathcal{U}$ whose closure is compact. Let O be open. For each $x \in O$, there is a neighborhood $V \subseteq O$ of x with compact closure. Every element of $\mathcal{U}$ contained in V lies in $\mathcal{U}'$, whence V is a union of sets in $\mathcal{U}'$. Since this holds for all $x \in O$, also O is a union of sets in $\mathcal{U}'$. Therefore $\mathcal{U}'$ satisfies the requirements.

Putting together these two facts, we see that E has a countable basis each of whose members has compact closure. Order such a basis as $(V_n)_{n\geqslant 1}$. Define $K_n := \bigcup_{k\leqslant n}\overline{V_k}$. Then K_n is compact and $K_n \subseteq K_{n+1}$. If K is any compact subset of E, then since $E = \bigcup_n V_n$, the definition of compactness provides a finite subcover of K, whence $K \subseteq K_n$ for some n.

Although we did not use completeness of E, every locally compact, separable metric space is Polish. For a proof, see Theorem 5.3 of *Classical Descriptive Set Theory* by Alexander S. Kechris.

Chapter 8

Stochastic Differential Equations

We treat mainly the case of Lipschitz coefficients, where we prove existence, uniqueness, and that the solution is a Feller Markov process whose generator is a second-order differential operator.

8.1. Motivation and General Definitions

A real-valued ordinary differential equation has the form

$$y'(t) = b\big(t, y(t)\big),$$

also written as

$$\mathrm{d}y_t = b(t, y_t)\,\mathrm{d}t.$$

Here, we are writing the subscript t not to indicate derivative, but the time variable. We may wish to model noise by adding a term on the right, $\sigma\,\mathrm{d}B_t$ or $\sigma(t, y_t)\,\mathrm{d}B_t$, where B is Brownian motion. The equation

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

means, by definition, that

$$y_t = y_0 + \int_0^t b(s, y_s)\,\mathrm{d}s + \int_0^t \sigma(s, y_s)\,\mathrm{d}B_s.$$

We have a similar notion for vector-valued processes:

Definition 8.1. *Let $d, m \in \mathbb{N}^+$. Denote the set of $d \times m$ real matrices by $M_{d\times m}(\mathbb{R})$ and give it the product topology. Let $\sigma\colon \mathbb{R}_+ \times \mathbb{R}^d \to M_{d\times m}(\mathbb{R})$ and $b\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ be locally bounded, measurable functions. Write their coordinates as $\sigma = (\sigma_{ij})_{1\leqslant i\leqslant d,\, 1\leqslant j\leqslant m}$ and $b = (b_i)_{1\leqslant i\leqslant d}$. By a* ***solution of the stochastic differential equation***

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

we mean

- *a filtered probability space $\big(\Omega, \mathscr{F}, (\mathscr{F}_t)_{0\leqslant t\leqslant\infty}, \mathbf{P}\big)$ with a complete filtration,*
- *an m-dimensional $(\mathscr{F}_t)$-Brownian motion $B = (B^1, \ldots, B^m)$ started from 0, and*

- *an* $(\mathscr{F}_t)$*-adapted continuous* $\mathbb{R}^d$*-valued process* $X = (X^1, \ldots, X^d)$ *such that*

$$\forall t \geqslant 0 \quad X_t = X_0 + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s + \int_0^t b(s, X_s)\,\mathrm{d}s,$$

which means

$$\forall i \in [1, d] \quad X_t^i = X_0^i + \sum_{j=1}^m \int_0^t \sigma_{ij}(s, X_s)\,\mathrm{d}B_s^j + \int_0^t b_i(s, X_s)\,\mathrm{d}s.$$

If also $\mathbf{P}[X_0 = x] = 1$*, then we say* X *is a solution of* $E_x(\sigma, b)$*.*

Note that $X_0 \perp\!\!\!\perp B$ because $X_0 \in \mathscr{F}_0 \perp\!\!\!\perp B$.

Existence and uniqueness can be defined in various ways probabilistically:

Definition 8.2. *We say* $E(\sigma, b)$ *has* ***weak existence*** *if for all* $x \in \mathbb{R}^d$*, there exists a solution to* $E_x(\sigma, b)$*. We say* $E(\sigma, b)$ *has* ***weak uniqueness*** *if for each* $x \in \mathbb{R}^d$*, over all solutions to* $E_x(\sigma, b)$ *(including varying the filtered probability space), the law of* X *is the same. We say* $E(\sigma, b)$ *has* ***pathwise uniqueness*** *if for each filtered probability space* $\big(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbf{P}\big)$ *and for each* $(\mathscr{F}_t)$*-Brownian motion* B*, any pair,* X *and* X'*, of solutions to* $E(\sigma, b)$ *such that* $\mathbf{P}[X_0 = X_0'] = 1$ *are indistinguishable. We say a solution of* $E_x(\sigma, b)$ *is a* ***strong solution*** *if it is adapted to the completed canonical filtration of* B*.*

Exercise (due 3/29). Exercise 8.9.

Example (Section 8.4.1)**.** To model the motion of a physical Brownian particle, we should not consider the forces as changing the position of the particle, but its momentum or, equivalently, velocity. Furthermore, there is a frictional drag (viscosity). This leads to the stochastic differential equation

$$\mathrm{d}X_t = \mathrm{d}B_t - \lambda X_t\,\mathrm{d}t \quad \text{(1-dimension)} \tag{$*$}$$

for the velocity X, where $\lambda > 0$. This is the Langevin equation, up to constants, historically the first stochastic differential equation. It is also exponential decay with noise. We can solve this by applying integration by parts to $\mathrm{e}^{\lambda t} X_t$:

$$\mathrm{d}(\mathrm{e}^{\lambda t} X_t) = \lambda \mathrm{e}^{\lambda t} X_t\,\mathrm{d}t + \mathrm{e}^{\lambda t}\,\mathrm{d}X_t \overset{\text{want}}{=} \mathrm{e}^{\lambda t}\,\mathrm{d}B_t. \tag{$**$}$$

Before continuing, note that the equality labeled "want" can be written $\mathrm{e}^{\lambda t}(\lambda X_t\,\mathrm{d}t + \mathrm{d}X_t) = \mathrm{e}^{\lambda t}\,\mathrm{d}B_t$, which has the form $H \cdot Y = H \cdot Z$, where Y and Z are continuous semimartingales. Provided H and H^{-1} are locally bounded, progressive processes, Section 5.1.3 shows that this equation is equivalent to $Y - Y(0) = Z - Z(0)$. Therefore, ($**$) is equivalent to ($*$). Hence

$$X_t := X_0 \mathrm{e}^{-\lambda t} + \int_0^t \mathrm{e}^{-\lambda(t-s)}\,\mathrm{d}B_s$$

solves the Langevin equation. This solution is called the ***Ornstein–Uhlenbeck process***. We have proved weak existence, weak uniqueness, and pathwise uniqueness. It is also a strong solution: the integral is a Wiener integral, whence it belongs to the Gaussian space generated by B. We consider two special cases.

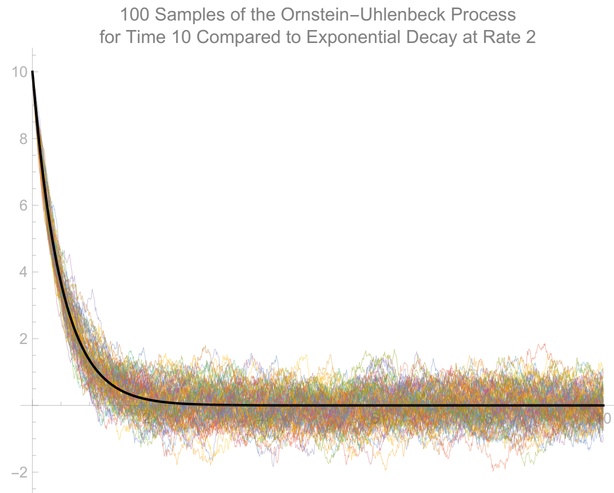


Figure 8.1: Simulation of the Ornstein–Uhlenbeck process

(1) Suppose $\mathbf{P}[X_0 = x] = 1$. Then X is a non-centered Gaussian process with mean function

$$m(t) := \mathbf{E}[X_t] = x\mathrm{e}^{-\lambda t}.$$

Thus, X is gotten by adding $m(t)$ to a centered Gaussian process with covariance function, for $0 \leqslant s \leqslant t$,

$$\begin{aligned} K(s,t) := \operatorname{Cov}(X_s, X_t) &= \mathbf{E}\Big[\int_0^t \mathrm{e}^{-\lambda(t-u)}\,\mathrm{d}B_u \cdot \int_0^s \mathrm{e}^{-\lambda(s-u)}\,\mathrm{d}B_u\Big] \\ &= \int_0^s \mathrm{e}^{-\lambda(t-u)}\mathrm{e}^{-\lambda(s-u)}\,\mathrm{d}u = \mathrm{e}^{-\lambda(t+s)} \underbrace{\int_0^s \mathrm{e}^{2\lambda u}\,\mathrm{d}u}_{\frac{\mathrm{e}^{2\lambda s}-1}{2\lambda}} \\ &\quad \text{[isometry]} \\ &= \frac{\mathrm{e}^{-\lambda|t-s|} - \mathrm{e}^{-\lambda(t+s)}}{2\lambda}. \end{aligned}$$

Thus, we see decay of the initial condition and convergence to a stationary process.

(2) Suppose $X_0 \sim \mathcal{N}(0, \frac{1}{2\lambda})$. Then X is a centered Gaussian process with covariance function

$$K(s,t) + \mathbf{E}\big[X_0\mathrm{e}^{-\lambda t} \cdot X_0\mathrm{e}^{-\lambda s}\big] = K(s,t) + \frac{\mathrm{e}^{-\lambda(t+s)}}{2\lambda} = \frac{\mathrm{e}^{-\lambda|t-s|}}{2\lambda}.$$

We see that the Ornstein–Uhlenbeck process in this case is stationary. Our later theory will show it is Markov, but this can be shown directly:

Exercise (due 4/5). Show that a centered stationary Gaussian process on $\mathbb{R}_+$ is Markov if and only if there exist $\lambda \in [0, \infty]$ and $a \geqslant 0$ such that the covariance function is

$$(s,t) \mapsto a\mathrm{e}^{-\lambda|s-t|}.$$

In this case, give the transition semigroup.

Exercise. Calculate the quadratic variation of an Ornstein–Uhlenbeck process in two ways, one from the defining SDE and the other from the solution of the SDE as a stochastic integral.

A simple transformation of Brownian motion gives another description of the stationary Ornstein–Uhlenbeck process: Suppose that $(\beta_t)_t$ is an $(\mathscr{F}_t)_t$-Brownian motion and $\lambda > 0$. Then $\big((2\lambda)^{-1/2}\mathrm{e}^{-\lambda t}\beta_{\mathrm{e}^{2\lambda t}}\big)_{t\geqslant 0}$ is a centered Gaussian process with the same initial distribution and covariance function as in (2) above. Since it is continuous, this process has the same law as a stationary Ornstein–Uhlenbeck process, but it is adapted to the filtration $(\mathscr{F}_{\mathrm{e}^{2\lambda t}})_t$. A slightly different description comes from applying Theorem 5.13 (the Dambis–Dubins–Schwarz theorem) to $\big(\mathrm{e}^{-\lambda t}(X_t - X_0)\big)_t$, yielding $\big(\mathrm{e}^{-\lambda t}(X_0 + \beta_{(\mathrm{e}^{2\lambda t}-1)/(2\lambda)})\big)_{t\geqslant 0}$, which works for any $X_0 \in \mathscr{F}_0$.

Exercise. Brownian bridges can be defined in many ways. One such is in Exercise 2.27. We will define a ***standard Brownian bridge*** as a continuous Gaussian process X on $[0, 1]$ with covariance function $(s, t) \mapsto (s \wedge t)(1 - s \vee t)$ such that $X_0 = X_1 = 0$. Let B be a standard Brownian motion.

(1) Show that $\mathrm{d}X_t = \mathrm{d}B_t - \frac{X_t}{1-t}\,\mathrm{d}t$ has a unique strong solution on $[0, 1)$ with $X_0 = 0$ given by

$$X_t = (1-t)\int_0^t \frac{\mathrm{d}B_s}{1-s} = B_t - (1-t)\int_0^t \frac{B_s\,\mathrm{d}s}{(1-s)^2}.$$

(2) Show that the solution in part (1) extends continuously to $X_1 = 0$ and then is a standard Brownian bridge.

(3) Show that every standard Brownian bridge X has the property that $\big(X_t/(1-t)\big)_{t\in[0,1)}$ has independent increments.

We now give an example (due to Tanaka) of a stochastic differential equation where weak existence and weak uniqueness hold, but pathwise uniqueness fails and there is no strong solution:

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

where

$$\sigma(x) := \begin{cases} 1 & \text{if } x > 0, \\ -1 & \text{if } x \leqslant 0. \end{cases}$$

Recall that Theorem 5.12 (of Lévy) implies that if B is an $(\mathscr{F}_t)$-Brownian motion, H is progressive, and $|H| = 1$, then $H \cdot B$ is an $(\mathscr{F}_t)$-Brownian motion. Therefore, weak uniqueness holds (since X is progressive, so is $\sigma(X)$). This also suggests how to get weak existence: Let X be a Brownian motion starting from $x \in \mathbb{R}$ and define

$$B_t := \int_0^t \sigma(X_s)\,\mathrm{d}X_s.$$

Then B is a Brownian motion, and $\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t$ because $\sigma^2 = 1$.

We claim that $\mathrm{d}(-X_t) = \sigma(-X_t)\,\mathrm{d}B_t$, which means that pathwise uniqueness fails. It suffices to see that

$$\int_0^t \mathbf{1}_{\{0\}}(X_s)\,\mathrm{d}B_s = 0,$$

which follows from the fact that its expected quadratic variation is

$$\mathbf{E}_x\Big[\int_0^t 1_{\{0\}}(X_s)\,\mathrm{d}s\Big] = \int_0^t \mathbf{P}_x[X_s = 0]\,\mathrm{d}s = 0.$$

One can show that $\overline{\mathscr{F}^B_\bullet} = \overline{\mathscr{F}^{|X|}_\bullet} \subsetneq \overline{\mathscr{F}^X_\bullet}$ (here, $\overline{}$ denotes completion), so X is not a strong solution. Similarly, one shows there does not exist any strong solution. This relies on Tanaka's formula for local time (Chapter 9).

Barlow (1982) gave, for each $\beta \in (0, 1/2)$, a function $\sigma\colon \mathbb{R} \to \mathbb{R}$ that is Hölder-continuous of order β and bounded above and below by positive constants such that

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

has a weak solution but no strong solution and no pathwise uniqueness.

Exercise (due 4/5). Let M be a continuous semimartingale with $M_0 = 0$. The proof of Proposition 5.11 shows that for all $\lambda \in \mathbb{C}$,

$$\mathscr{E}(\lambda M) := \exp\big\{\lambda M - \frac{1}{2}\langle \lambda M, \lambda M\rangle\big\}$$

satisfies

$$\mathrm{d}X_t = \lambda X_t\,\mathrm{d}M_t, \quad X_0 = 1.$$

Show that there is no other solution. *Hint:* compute $X\mathscr{E}(\lambda M)^{-1}$ using Itô's formula.

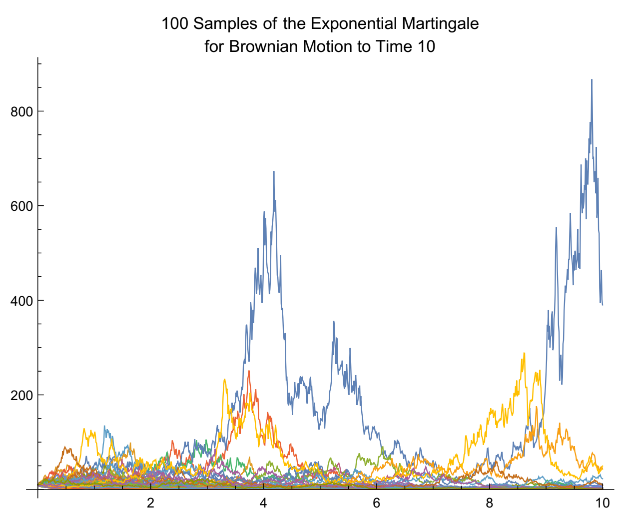


Figure 8.2: Simulation of the exponential martingale. Note: $\mathscr{E}(B)_t \to 0$ almost surely as $t \to \infty$.

Example (Section 8.4.2)**.** A combination of the SDEs of both the Ornstein–Uhlenbeck process and the preceding exercise is

$$\mathrm{d}X_t = \sigma X_t\,\mathrm{d}B_t + rX_t\,\mathrm{d}t$$

for constants $\sigma > 0$ and $r \in \mathbb{R}$. To solve this, calculate (if $X_0 > 0$)

$$\mathrm{d}\log X_t = X_t^{-1}\,\mathrm{d}X_t - \frac{1}{2}X_t^{-2}\,\mathrm{d}\langle X, X\rangle_t = \sigma\,\mathrm{d}B_t + r\,\mathrm{d}t - \frac{\sigma^2}{2}\,\mathrm{d}t,$$

whence

$$X_t = X_0\exp\bigl\{\sigma B_t + (r - \frac{\sigma^2}{2})t\bigr\} = X_0\mathscr{E}(\sigma B)_t\mathrm{e}^{rt};$$

one checks this is indeed a solution. One can also show uniqueness as in the exercise. This is known as ***geometric Brownian motion*** with parameters σ and r. It is fundamental in financial mathematics; r represents interest rate.

In fact, this example is itself an example as in the exercise: take $\lambda := 1$ and $M_t := \sigma B_t + rt$.

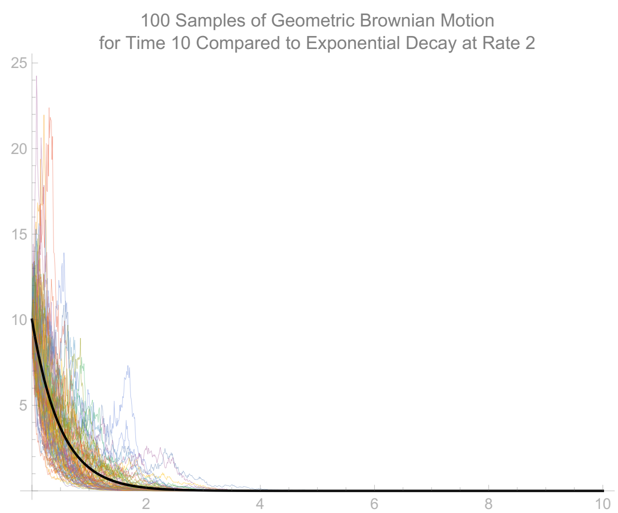


Figure 8.3: Simulation of geometric Brownian motion

There is a very general relation between existence and uniqueness:

Theorem (Yamada–Watanabe)**.** *If $E(\sigma, b)$ has pathwise uniqueness, then $E(\sigma, b)$ has weak uniqueness. If $E(\sigma, b)$ also has weak existence, then for every filtered probability space $\bigl(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbf{P}\bigr)$ and $(\mathscr{F}_t)$-Brownian motion, for all $x \in \mathbb{R}^d$, $E_x(\sigma, b)$ has a strong solution.*

Theorem (Gikhman–Skorokhod)**.** *If $E(\sigma, b)$ has weak uniqueness and a strong solution, then it has pathwise uniqueness.*

We will not prove these because we will establish these properties for the case of Lipschitz coefficients.

8.2. The Lipschitz Case

Here we show that when σ and b are Lipschitz in space uniformly in time, then all the existence and uniqueness properties hold. (The hypothesis is the same as in the Picard–Lindelöf theorem for ordinary differential equations, which gives existence and uniqueness there.)

Lemma 8.4 (Gronwall's Lemma)**.** *Let $T > 0$ be a constant and g be a measurable function on $[0, T]$. If there exist $a \in \mathbb{R}$ and a measurable, nonnegative function b such that $b \cdot g$ is Lebesgue-integrable on $[0, T]$ such that*

$$\forall t \in [0, T] \quad g(t) \leqslant a + \int_0^t b(s)g(s)\,\mathrm{d}s,$$

then

$$\forall t \in [0, T] \quad g(t) \leqslant a \cdot \mathrm{e}^{\int_0^t b(s)\,\mathrm{d}s}.$$

We will use only the case that b is constant, in which case the upper bound is $a \cdot \mathrm{e}^{bt}$.

Proof. Let $G(t)$ denote the right-hand side of the hypothesized inequality, $g \leqslant G$. Then G is an absolutely continuous function with $G' \overset{\text{a.e.}}{=} b \cdot g \leqslant b \cdot G$ on $[0, T]$. It follows that, with $h(t) := \int_0^t b(s)\,\mathrm{d}s$,

$$\big(\mathrm{e}^{-h}G\big)' \overset{\text{a.e.}}{=} (G' - bG)\mathrm{e}^{-h} \leqslant (bG - bG)\mathrm{e}^{-h} = 0.$$

Therefore, $\mathrm{e}^{-h}G \leqslant \mathrm{e}^{-h(0)}G(0) = a$ on $[0, T]$, whence $g \leqslant G \leqslant a\mathrm{e}^h$, as desired. ◂

In the rest of this section, we assume $\sigma\colon \mathbb{R}_+ \times \mathbb{R}^d \to M_{d\times m}(\mathbb{R})$ ***and*** $b\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ ***are continuous and***

$$\exists K \ \ \forall t \geqslant 0 \ \ \forall x, y \in \mathbb{R}^d \qquad \begin{cases} |\sigma(t, x) - \sigma(t, y)| \leqslant K|x - y|, \\ |b(t, x) - b(t, y)| \leqslant K|x - y|. \end{cases}$$

Theorem 8.3. *$E(\sigma, b)$ has pathwise uniqueness and for every filtered probability space and associated Brownian motion, for all $x \in \mathbb{R}^d$, $E_x(\sigma, b)$ has a (unique) strong solution.*

In particular, we have weak existence. Theorem 8.5 will imply weak uniqueness.

We will prove this using $d = m = 1$ to simplify the notation.

Lemma. *Suppose that for all $t \geqslant 0$,*

$$\widetilde{X}_t = \widetilde{X}_0 + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s + \int_0^t b(s, X_s)\,\mathrm{d}s$$

and

$$\widetilde{Y}_t = \widetilde{Y}_0 + \int_0^t \sigma(s, Y_s)\,\mathrm{d}B_s + \int_0^t b(s, Y_s)\,\mathrm{d}s.$$

Then for all stopping times τ, for all $t \geqslant 0$,

$$\mathbf{E}\Big[\sup_{0\leqslant s\leqslant t} (\widetilde{X}_{s\wedge\tau} - \widetilde{Y}_{s\wedge\tau})^2\Big] \leqslant 3\,\mathbf{E}\big[(\widetilde{X}_0 - \widetilde{Y}_0)^2\big] + 3(4 + t)K^2 \int_0^t \mathbf{E}\big[(X_{r\wedge\tau} - Y_{r\wedge\tau})^2\big]\,\mathrm{d}r.$$

Proof. By the arithmetic mean-quadratic mean inequality, we see that the left-hand side is

$$\leqslant 3\Bigg(\mathbf{E}\big[(\widetilde{X}_0 - \widetilde{Y}_0)^2\big] + \mathbf{E}\Big[\sup_{0\leqslant s\leqslant t}\Big|\int_0^s \big[\sigma(r, X_{r\wedge\tau}) - \sigma(r, Y_{r\wedge\tau})\big]\mathbf{1}_{[s\leqslant\tau]}\,\mathrm{d}B_r\Big|^2\Big]$$
$$+\,\mathbf{E}\Big[\sup_{0\leqslant s\leqslant t}\Big|\int_0^s \big[b(r, X_{r\wedge\tau}) - b(r, Y_{r\wedge\tau})\mathbf{1}_{[s\leqslant\tau]}\,\mathrm{d}r\big]\Big|^2\Big]\Bigg).$$

The second term can be bounded by the Doob's L^2-inequality: it is

$$\leqslant 4\,\mathbf{E}\Big[\int_0^t \big[\sigma(r, X_{r\wedge\tau}) - \sigma(r, Y_{r\wedge\tau})\big]^2\,\mathrm{d}r\Big] \leqslant 4K^2\int_0^t \mathbf{E}\big[(X_{r\wedge\tau} - Y_{r\wedge\tau})^2\big]\,\mathrm{d}r.$$

The third term can be bounded by the arithmetic mean-quadratic mean inequality: it is

$$\leqslant \mathbf{E}\Big[t\cdot\int_0^t \big[b(r, X_{r\wedge\tau}) - b(r, Y_{r\wedge\tau})\big]^2\,\mathrm{d}r\Big] \leqslant tK^2\int_0^t \mathbf{E}\big[(X_{r\wedge\tau} - Y_{r\wedge\tau})^2\big]\,\mathrm{d}r.$$

Adding these gives the result. ◀

Proof of Theorem 8.3. We first show uniqueness. Fix a filtered probability space and a Brownian motion. Suppose that X and X' are both solutions of $E(\sigma, b)$ with $X_0 = X'_0$. Fix $M > 0$ and define

$$\tau := \inf\big\{t \geqslant 0\,;\ |X_t| \geqslant M \text{ or } |X'_t| \geqslant M\big\}.$$

Then by the lemma, $\forall T > 0\ \ \forall t \in [0, T]$

$$\mathbf{E}\big[\sup_{0\leqslant s\leqslant t}(X_{s\wedge\tau} - X'_{s\wedge\tau})^2\big] \leqslant 3K^2(4+T)\int_0^t \mathbf{E}\big[(X_{r\wedge\tau} - X'_{r\wedge\tau})^2\big]\,\mathrm{d}r.$$

Thus, Gronwall's lemma applies to

$$g(t) := \mathbf{E}\big[\sup_{0\leqslant s\leqslant t}(X_{s\wedge\tau} - X'_{s\wedge\tau})^2\big],$$

yielding $g = 0$. Now let $M \to \infty$ to get $X = X'$ (i.e., indistinguishable).

To show existence, we use Picard's approximation method. Fix $x \in \mathbb{R}$ and define recursively

$$X_t^0 := x,$$
$$\forall n \geqslant 0 \qquad X_t^{n+1} := x + \int_0^t \sigma(s, X_s^n)\,\mathrm{d}B_s + \int_0^t b(s, X_s^n)\,\mathrm{d}s.$$

Note that by induction, X^n is adapted to $\overline{\mathscr{F}_\bullet^B}$ and continuous, so the stochastic integrals are defined.

Next, fix $T > 0$. For $n \geqslant 1$, set

$$g_n(t) := \mathbf{E}\big[\sup_{0\leqslant s\leqslant t}(X_s^n - X_s^{n-1})^2\big].$$

Because $\sigma(\cdot, x)$ is continuous, it is bounded on $[0, T]$, whence by Doob's L^2-inequality,

$$\mathbf{E}\Big[\sup_{0\leqslant s\leqslant t}\Big|\int_0^t \sigma(s, x)\,\mathrm{d}B_s\Big|^2\Big]$$

is bounded on $[0,T]$. Therefore,

$$\exists C'_T \ \forall t \leqslant T \qquad g_1(t) \leqslant C'_T.$$

The lemma shows that

$$\exists C_T \ \forall n \geqslant 1 \ \forall t \in [0,T] \qquad g_{n+1}(t) \leqslant C_T \int_0^t g_n(s)\,\mathrm{d}s.$$

Induction shows that

$$\forall n \geqslant 1 \qquad g_n(t) \leqslant C'_T (C_T)^{n-1} \frac{t^{n-1}}{(n-1)!}.$$

In particular, $\sum_{n=1}^{\infty} g_n(T)^{1/2} < \infty$, whence the arithmetic mean-quadratic mean inequality yields

$$\mathbf{E}\Bigg[\sum_{n=1}^{\infty} \sup_{0\leqslant t\leqslant T} |X_t^{n+1} - X_t^n|\Bigg] < \infty,$$

so the sum is finite almost surely. Therefore, $X^n \restriction [0,T]$ almost surely converges uniformly; let its limit be X on $[0,T]$, which necessarily has continuous sample paths. As the almost sure limit of X^n, X is also adapted to $\overline{\mathscr{F}_\bullet^B}$.

Since $X^n \to X$ in $L^2\big(\Omega \times [0,T]\big)$ and $b(\cdot)$ is Lipschitz, we have

$$\int_0^t b(x, X_s^n)\,\mathrm{d}s \xrightarrow{\mathbf{P}} \int_0^t b(x, X_s)\,\mathrm{d}s.$$

Similarly,

$$\mathbf{E}\Big[\Big(\int_0^t \sigma(s, X_s^n)\,\mathrm{d}B_s - \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s\Big)^2\Big] = \mathbf{E}\Big[\int_0^t \big(\sigma(s, X_s^n) - \sigma(s, X_s)\big)^2\,\mathrm{d}s\Big] \to 0,$$

so

$$\int_0^t \sigma(s, X_s^n)\,\mathrm{d}B_s \xrightarrow{\mathbf{P}} \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s.$$

Therefore, X satisfies $E_x(\sigma, b)$ on $[0,T]$. Because T is arbitrary, X^n has an almost sure limit, X, on $\mathbb{R}_+$ and X is a strong solution to $E_x(\sigma, b)$. ◂

Exercise. Let σ, σ', b, b' all satisfy the Lipschitz conditions of this section. Suppose that on some open set $U \subseteq \mathbb{R}^d$, we have $(\sigma, b) = (\sigma', b')$ on $\mathbb{R}_+ \times U$. Fix $x \in U$ and let X and X' be the corresponding solutions to $E_x(\sigma, b)$ and $E_x(\sigma', b')$. Let $T := \inf\{t \geq 0\,;\ X_t \notin U\}$ and $T' := \inf\{t \geq 0\,;\ X'_t \notin U\}$. Show that $T = T'$ a.s. and X^T is indistinguishable from $(X')^{T'}$.

Exercise (due 4/5). Exercise 8.10, Exercise 8.12.

Exercise. Suppose that X is a solution of $E_0(\sigma, b)$, where $\sigma, b\colon \mathbb{R} \to \mathbb{R}$ are Borel and such that b/σ^2 is continuous and locally integrable and $\sigma(x) \neq 0$ for all $x \in \mathbb{R}$. Let $T_x := \inf\{t \geqslant 0\,;\ X_t = x\}$ and $c < 0 < d$.

(1) Show that $T_c \wedge T_d < \infty$ a.s.

(2) Calculate $\mathbf{P}[T_c < T_d]$.

(3) Show that the answers do not change if σ is replaced by $g \cdot \sigma$ and b is replaced by $g^2 \cdot b$, where g is a strictly positive, Borel function.

Exercise. Let $\sigma, b\colon \mathbb{R} \to \mathbb{R}$ be bounded, σ be continuous, b be Borel, and $\inf \sigma > 0$. Let X solve $E(\sigma, 0)$. Define $L_t := \int_0^t b(X_s)\sigma(X_s)^{-1}\,\mathrm{d}B_s$, $\widetilde{X} := X - \langle X, L\rangle$, and $\beta_t := \int_0^t \sigma(X_s)^{-1}\,\mathrm{d}\widetilde{X}_s$.

(1) Show that $\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}\beta_t + b(X_t)\,\mathrm{d}t$.

(2) Use Girsanov's theorem to show that $E(\sigma, b)$ has a weak solution whose law is mutually locally absolutely continuous with respect to the $\mathbf{P}$-law of X.

We now show continuity of the solution to $E_x(\sigma, b)$ as a function of x. The space $C(\mathbb{R}_+, \mathbb{R}^m)$ of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^m$ has the topology of uniform convergence on compact subsets of $\mathbb{R}_+$, whose Borel σ-field $\mathscr{C}_m$ is the one generated by coordinate maps. Note $C(\mathbb{R}_+, \mathbb{R}^m)$ is complete. The law of m-dimensional Brownian motion started at 0 is Wiener measure W on $C(\mathbb{R}_+, \mathbb{R}^m)$. The idea is to look at the solution of $E_x(\sigma, b)$ as a function F_x of the path w of Brownian motion. Write $\overline{\mathscr{C}_m}$ for the W-completion of $\mathscr{C}_m$.

Theorem 8.5. *There exists a map $\mathbb{R}^d \ni x \mapsto F_x\colon \big(C(\mathbb{R}_+, \mathbb{R}^m), \overline{\mathscr{C}_m}\big) \to \big(C(\mathbb{R}_+, \mathbb{R}^d), \mathscr{C}_d\big)$ such that*

(i) *for all $t \geqslant 0$ and $x \in \mathbb{R}^d$, there exists $\varphi_t^x\colon \big(C([0,t], \mathbb{R}^m), \mathscr{C}_m\big) \to (\mathbb{R}^d, \mathscr{R}^d)$ such that*

$$F_x(\omega)_t = \varphi_t^x\big(\omega\restriction[0,t]\big) \quad W\text{-a.s.};$$

(ii) *for all $\omega \in C(\mathbb{R}_+, \mathbb{R}^m)$, $x \mapsto F_x(\omega)$ is continuous from $\mathbb{R}^d$ to $C(\mathbb{R}_+, \mathbb{R}^d)$; and*

(iii) *for every complete filtered $\big(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbf{P}\big)$ and m-dimensional $(\mathscr{F}_t)$-Brownian motion B with $B_0 = 0$, for all $\mathbb{R}^d$-valued $U \in \mathscr{F}_0$, the process*

$$t \mapsto F_U(B)_t$$

is the pathwise unique solution to $E(\sigma, b)$ with initial value U.

Remark. Note that (iii) implies weak uniqueness: each solution of $E_x(\sigma, b)$ has the form $F_x(B)$ for some Brownian motion B, whence its law is $(F_x)_*(W)$, the pushforward of W under F_x.

Proof. Again for simplicity of notation, we prove only the case $m = d = 1$. Let $\mathscr{G}_t$ be the W-completion of the σ-field on $C(\mathbb{R}_+, \mathbb{R})$ generated by the coordinate maps $s \mapsto w(s)$ for $s \in [0, t]$, and $\mathscr{G}_\infty := \bigvee_{t \geqslant 0} \mathscr{G}_t$ (which we denoted by $\overline{\mathscr{C}_1}$ above).

The topology on $C(\mathbb{R}_+, \mathbb{R})$ can be defined by a metric of the form

$$\rho(w, w') := \sum_{k=1}^{\infty} \alpha_k \Big(\sup_{s \in [0,k]} |w(s) - w'(s)| \wedge 1 \Big)$$

for any sequence $\alpha_k > 0$ with $\sum_{k=1}^\infty \alpha_k < \infty$.

Let X^x be the solution to $E_x(\sigma, b)$ corresponding to $\big(C(\mathbb{R}_+, \mathbb{R}), \mathscr{G}_\infty, (\mathscr{G}_t)_t, W\big)$ and the Brownian motion $t \mapsto w(t)$: such a solution exists and is unique (up to indistinguishability) by Theorem 8.3.

Fix $x, y \in \mathbb{R}$. Let

$$T_n := \inf\big\{t \geqslant 0\,;\ |X_t^x| \geqslant n \text{ or } |X_t^y| \geqslant n\big\}.$$

By the lemma,

$$\forall t \geqslant 0 \qquad \mathbf{E}\Big[\sup_{s\leqslant t}(X^x_{s\wedge T_n} - X^y_{s\wedge T_n})^2\Big] \leqslant 3(x-y)^2 + 3K^2(4+t)\int_0^t \mathbf{E}\big[(X^x_{s\wedge T_n} - X^y_{s\wedge T_n})^2\big]\,\mathrm{d}s.$$

Note that $(X^x_{s\wedge T_n} - X^y_{s\wedge T_n})^2 \leqslant 4n^2$, so Gronwall's lemma implies that

$$\forall T > 0 \qquad \mathbf{E}\Big[\sup_{s\leqslant T}(X^x_{s\wedge T_n} - X^y_{s\wedge T_n})^2\Big] \leqslant 3(x-y)^2 \mathrm{e}^{3K^2(4+T)T}.$$

Choose $\alpha_k > 0$ such that

$$\sum_{k=1}^{\infty} \alpha_k \mathrm{e}^{3K^2(4+k)k} =: C < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} \alpha_k = 1.$$

Then by the arithmetic mean-quadratic mean inequality,

$$\mathbf{E}\big[\rho(X^x, X^y)^2\big] \leqslant \mathbf{E}\Big[\sum_{k=1}^{\infty} \alpha_k \sup_{s\in[0,k]} (X^x_s - X^y_s)^2\Big] \leqslant 3C(x-y)^2.$$

By Kolmogorov's lemma (Theorem 2.9) applied to the process $x \mapsto X^x$ with values in $\big(C(\mathbb{R}_+, \mathbb{R}), \rho\big)$, we get a modification $\widetilde{X}$ of X with continuous (in $x \in \mathbb{R}$) sample paths. (Note: $\forall x \in \mathbb{R}\quad X^x$ and $\widetilde{X}^x$ are indistinguishable.) Define

$$F_x(w) := \widetilde{X}^x(w) = \big(\widetilde{X}^x_t(w)\big)_{t\geqslant 0}.$$

This makes (ii) satisfied.

Since $X^x_t \in \mathscr{G}_t$, there exists

$$\varphi^x_t\colon \Big(C\big([0,t], \mathbb{R}\big), \sigma\big(w(s), s \in [0,t]\big)\Big) \to (\mathbb{R}, \mathscr{R})$$

such that $X^x_t(w) = \varphi^x_t\big(w\restriction[0,t]\big)$ W-a.s. Because

$$F_x(w)_t = \widetilde{X}^x_t(w) \overset{W\text{-a.s.}}{=} X^x_t(w),$$

we obtain (i).

Because X^x solves $E_x(\sigma, b)$, we have

$$\forall t \geqslant 0 \qquad X^x_t(w) = x + \int_0^t \sigma\big(s, X^x_s(w)\big)\,\mathrm{d}w(s) + \int_0^t b\big(s, X^x_s(w)\big)\,\mathrm{d}s$$

for W-a.e. w. By definition of F_x, we obtain

$$F_x(w)_t = x + \int_0^t \sigma\big(s, F_x(w)_s\big)\,\mathrm{d}w(s) + \int_0^t b\big(s, F_x(w)_s\big)\,\mathrm{d}s$$

for W-a.e. w. We want to substitute U for x and B for w. The stochastic integral is not defined pointwise, so we must be careful.

First, consider the map $(x, \omega) \mapsto F_x\big(B(\omega)\big)_t$. For fixed ω, this is continuous in x. Now

$$F_x\big(B(\omega)\big)_t = \varphi_t^x\big(B(\omega)\restriction[0,t]\big) \qquad \mathbf{P}\text{-a.s.}$$

by (i) and the fact that $W = B_* \mathbf{P}$. The right-hand side belongs to $\mathscr{F}_t$, whence so does the left-hand side by completeness. In other words, for fixed x, it is $\mathscr{F}_t$-measurable in ω. Therefore, the map is the limit as $n \to \infty$ of the functions

$$(x, \omega) \mapsto \sum_{k \in \mathbb{Z}} \mathbf{1}_{[\frac{k}{n}, \frac{k+1}{n})}(x) F_{\frac{k}{n}}\big(B(\omega)\big)_t,$$

which shows that the map is measurable with respect to $\mathscr{R} \otimes \mathscr{F}_t$. Because $\omega \mapsto \big(U(\omega), \omega\big)$ is measurable from $\mathscr{F}_t$ to $\mathscr{R} \otimes \mathscr{F}_t$, the composition $\omega \mapsto F_{U(\omega)}\big(B(\omega)\big)_t$ is $\mathscr{F}_t$-measurable. Thus, the process $F_U(B)$ is adapted.

Write

$$H(x, w) := \int_0^t \sigma\big(s, F_x(w)_s\big)\,\mathrm{d}w(s)$$

and

$$H_n(x, w) := \sum_{i=0}^{n-1} \sigma\big(\tfrac{it}{n}, F_x(w)_{it/n}\big)\big[w\big(\tfrac{(i+1)t}{n}\big) - w\big(\tfrac{it}{n}\big)\big].$$

By Proposition 5.9, $\forall x \;\; H_n(x, w) \xrightarrow{W} H(x, w)$ (i.e., in probability). Therefore,

$$\forall x \qquad H_n(x, B) \xrightarrow{\mathbf{P}} H(x, B).$$

Because $U \perp\!\!\!\perp B$, it follows (by conditioning on U) that

$$H_n(U, B) \xrightarrow{\mathbf{P}} H(U, B).$$

By Proposition 5.9 again, $H_n(U, B) \xrightarrow{\mathbf{P}} \int_0^t \sigma\big(s, F_U(B)_s\big)\,\mathrm{d}B_s$, whence $H(U, B)$ is that stochastic integral. Because

$$H(x, w) = F_x(w)_t - x - \int_0^t b\big(s, F_x(w)_s\big)\,\mathrm{d}s,$$

it follows that $F_U(B)$ solves $E(\sigma, b)$. ◀

In the lemma, one may use powers $p > 1$, not only $p = 2$, and this allows us to show (from our version of Kolmolgorov's lemma) that

$$\forall A > 0 \;\; \forall T > 0 \;\; \forall \varepsilon \in (0,1) \;\; \forall p > 0 \qquad \mathbf{E}\Bigg[\sup_{\substack{t \in [0,T], x \neq y \\ |x|, |y| \leqslant A}} \bigg(\frac{|X_t^x - X_t^y|}{|x - y|^{1-\varepsilon}}\bigg)^p\Bigg] < \infty.$$

Exercise (due 4/12). Exercise 8.14 (the inequality $0 < |Z_s|$ in (1) should be $0 \leqslant |Z_s|$; the conclusion in (4) is that X and X' are indistinguishable).

8.3. Solutions of Stochastic Differential Equations as Markov Processes

We now suppose that σ and b do not depend on time, but still are Lipschitz. Let F_x be as in Theorem 8.5.

Theorem 8.6. *Let X be a solution of $E(\sigma, b)$ on a complete filtered probability space $\big(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbf{P}\big)$ with $(\mathscr{F}_t)$-Brownian motion B. Then X is a Markov process with respect to $(\mathscr{F}_t)$ with semigroup*

$$Q_t f(x) := \int f\big(F_x(w)_t\big)\, W(\mathrm{d}w) \qquad \big(f \in B(\mathbb{R}^d),\ t \geqslant 0,\ x \in \mathbb{R}^d\big).$$

Proof. We first show that

$$\forall f \in B(\mathbb{R}^d)\ \ \forall s, t \geqslant 0 \qquad \mathbf{E}\big[f(X_{s+t}) \,\big|\, \mathscr{F}_s\big] = Q_t f(X_s).$$

Write

$$X_{s+t} = X_s + \int_s^{s+t} \sigma(X_r)\, \mathrm{d}B_r + \int_s^{s+t} b(X_r)\, \mathrm{d}r.$$

Recall that the stochastic integral is defined as $(\sigma(X)\cdot B)_{s+t} - (\sigma(X)\cdot B)_s$. However, by Proposition 5.9, we may also write it as

$$\int_0^t \sigma(X'_u)\, \mathrm{d}B'_u,$$

where $X'_u := X_{s+u}$, $B'_u := B_{s+u} - B_s$, and we use the complete filtration $\mathscr{F}'_u := \mathscr{F}_{s+u}$: X' is adapted to $\mathscr{F}'_\bullet$ and B' is an m-dimensional $\mathscr{F}'_\bullet$-Brownian motion. Thus,

$$X'_t = X_s + \int_0^t \sigma(X'_u)\, \mathrm{d}B'_u + \int_0^t b(X'_u)\, \mathrm{d}u,$$

i.e., X' solves $E(\sigma, b)$ on $(\Omega, \mathscr{F}, \mathscr{F}'_\bullet, \mathbf{P})$ with Brownian motion B' and initial value

$$X'_0 = X_s \in \mathscr{F}'_0.$$

By Theorem 8.5(iii), we have $X' = F_{X_s}(B')$. It follows that

$$\begin{aligned}\mathbf{E}\big[f(X_{s+t}) \,\big|\, \mathscr{F}_s\big] &= \mathbf{E}\big[f(X'_t) \,\big|\, \mathscr{F}_s\big] = \mathbf{E}\big[f\big(F_{X_s}(B')\big) \,\big|\, \mathscr{F}_s\big] \\ &= \int f\big(F_{X_s}(w)_t\big)\, W(\mathrm{d}w) = Q_t f(X_s),\end{aligned}$$

[$B' \sim W$ and $B' \perp\!\!\!\perp \mathscr{F}_s \ni X_s$; this is *not* a stochastic integral, so we may substitute X_s for x]

as desired.

It remains to show that (Q_t) is a transition semigroup, so we check Definition 6.1 step by step:

(i) $(Q_0 = \mathrm{Id})$ $F_x(w)_0 = x$

(ii) $(Q_t Q_s = Q_{s+t})$ Let X^x solve $E_x(\sigma, b)$. Then

[by the above]

$$Q_{s+t} f(x) = \mathbf{E}\big[f(X^x_{s+t})\big] = \mathbf{E}\big[Q_t f(X^x_s)\big] = Q_s(Q_t f)(x).$$

[$X^x_{s+t} \overset{\mathscr{D}}{=} F_x(w)_{s+t}$] [the first equality with f replaced by $Q_t f$ and $s+t$ by s]

(iii) ($(t, x) \mapsto Q_t(x, A)$ is measurable) By the topology of uniform convergence on compact sets for $C(\mathbb{R}_+, \mathbb{R}^d)$, Theorem 8.5 gives that for all $f \in C_{\mathrm{b}}(\mathbb{R}^d)$,

$$(t, x) \mapsto \int f\big(F_x(w)_t\big)\, W(\mathrm{d}w)$$

is continuous (since $(t, x) \mapsto F_x(\cdot)_t$ is continuous). Thus, for all $f \in C_{\mathrm{b}}(\mathbb{R}^d)$,

$$(t, x) \mapsto Q_t f(x)$$

is continuous. By regularity, it follows that if A is closed, $(t, x) \mapsto Q_t(x, A)$ is measurable. By the π-λ theorem, the same holds for all $A \in \mathscr{B}^d$. ◀

Exercise (due 4/19). Let E be locally compact Polish and $(Q_t)_{t \geqslant 0}$ be a transition semigroup on E that satisfies

(i) $\forall t \geqslant 0 \;\; \forall f \in C_0(E) \;\; Q_t f \in C_0(E)$, and

(ii) $\forall f \in C_0(E) \;\; \forall x \in E \;\; \lim_{t \downarrow 0} Q_t f(x) = f(x)$.

This exercise will show that (Q_t) is Feller.

(1) The proof of Proposition 6.8 shows that the range $\mathcal{R}$ of R_λ on $C_0(E)$ is the same for all $\lambda > 0$ and is contained in $C_0(E)$; also,

$$\forall x \in E \;\; \forall f \in C_0(E) \qquad \lim_{\lambda \to \infty} \lambda R_\lambda f(x) = f(x).$$

Use the Hahn–Banach theorem to deduce that $\mathcal{R}$ is dense in $C_0(E)$.

(2) The proof of Lemma 6.6 showed (used) that

$$\forall s > 0 \;\; \forall h \in B(E) \qquad \mathrm{e}^{-s} Q_s R_1 h = \int_s^\infty \mathrm{e}^{-t} Q_t h \, \mathrm{d}t.$$

Deduce that

$$\forall h \in B(E) \qquad \lim_{s \downarrow 0} \|Q_s R_1 h - R_1 h\| = 0.$$

Infer that

$$\forall f \in C_0(E) \qquad \lim_{s \downarrow 0} \|Q_s f - f\| = 0,$$

whence $(Q_t)_t$ is Feller.

Let $C_{\mathrm{c}}^2(\mathbb{R}^d)$ denote the space of $f \in C^2(\mathbb{R}^d)$ with compact support.

Theorem 8.7. *The semigroup $(Q_t)_t$ of Theorem 8.6 is Feller. Its generator L satisfies*

(i) *$C_{\mathrm{c}}^2(\mathbb{R}^d) \subseteq D(L)$, and*

(ii) *$\forall f \in C_{\mathrm{c}}^2(\mathbb{R}^d) \;\; \forall x \in \mathbb{R}^d$*

$$\begin{aligned} Lf(x) &= \frac{1}{2} \sum_{i,j=1}^d (\sigma\sigma^*)_{ij}(x) f_{ij}(x) + \sum_{i=1}^d b_i(x) f_i(x) \\ &= \underbrace{\tfrac{1}{2}\big(\sigma\sigma^*(x), \nabla^2 f(x)\big)}_{\textit{natural inner product}} + b(x) \cdot \nabla f(x), \end{aligned}$$

where σ^* is the transpose of $\sigma \in M_{d \times m}(\mathbb{R}^d)$.

Proof. Let $f \in C_0(\mathbb{R}^d)$. Theorem 8.5 (continuity of $x \mapsto F_x(w)$), the definition of Q_t, and the bounded convergence theorem show that $Q_t f \in C(\mathbb{R}^d)$. Similarly, since $t \mapsto F_x(w)_t$ is continuous, we obtain

$$\forall x \qquad \lim_{t \downarrow 0} Q_t f(x) = f(x).$$

Thus, by the exercise, it suffices to show that

$$Q_t f \in C_0(\mathbb{R}^d)$$

in order to conclude that $(Q_t)_t$ is Feller. We assume σ and b are bounded and leave the general case to another exercise.

Suppose

$$\forall i, j \qquad |\sigma_{i,j}| \leqslant C \quad \text{and} \quad |b_i| \leqslant C.$$

For a solution X^x of $E_x(\sigma, b)$, we obtain, for all $t \geqslant 0$,

$$\begin{aligned}
\mathbf{E}\big[|X_t^x - x|^2\big] &\leqslant \sum_{i=1}^{d} (m+1)\Bigg\{ \sum_{j=1}^{m} \mathbf{E}\bigg[\Big(\int_0^t \sigma_{ij}(X_s^x)\, \mathrm{d}B_s^j\Big)^2\bigg] + \mathbf{E}\bigg[\Big(\int_0^t b_i(X_s^x)\, \mathrm{d}s\Big)^2\bigg]\Bigg\} \\
&\leqslant d(m+1)C^2(t + t^2).
\end{aligned}$$

Therefore, for all $A > 0$, Chebyshev's inequality gives

$$\begin{aligned}
|Q_t f(x)| &= \big|\mathbf{E}\big[f(X_t^x)\big]\big| \\
&\leqslant \big|\mathbf{E}\big[f(X_t^x)\mathbf{1}_{[|X_t^x - x| \leqslant A]}\big]\big| + \|f\| \frac{d(m+1)C^2(t+t^2)}{A^2}.
\end{aligned}$$

Now let $x \to \infty$ and then $A \to \infty$ to get $Q_t f \in C_0(\mathbb{R}^d)$.

Finally, let $f \in C_c^2(\mathbb{R}^d)$ and set g to be the desired value of Lf. By Theorem 6.14, if we show that

$$M := f(X^x) - \int_0^{\bullet} g(X_s^x)\, \mathrm{d}s$$

is a martingale, then it will follow that $f \in D(L)$ and $g = Lf$. Apply Itô's formula to $f(X^x)$:

$$\begin{aligned}
f(X_t^x) &= f(x) + \text{stochastic integral} \\
&\quad + \sum_{i=1}^{d} \int_0^t b_i(X_s^x) f_i(X_s^x)\, \mathrm{d}s + \frac{1}{2} \sum_{i,j=1}^{d} \int_0^t f_{ij}(X_s^x)\, \mathrm{d}\langle X^{x,i}, X^{x,j}\rangle_s \\
&= \underbrace{f(x) + \text{stochastic integral}}_{\text{continuous local martingale}} + \int_0^t g(X_s^x)\, \mathrm{d}s.
\end{aligned}$$

Therefore, M is a continuous local martingale. Because f and g are bounded, M is a martingale by Proposition 4.7(ii). ◀

Exercise (due 4/19). Complete the proof of Theorem 8.7 in the general Lipschitz case as follows. By what we have shown at the start of the proof of Theorem 8.7 and our version of Theorem 6.17, X^x has the strong Markov property.

(1) Show that there exists $C < \infty$ such that

$$\forall |x| \geqslant 1 \qquad |\sigma(x)| \leqslant C|x| \quad \text{and} \quad |b(x)| \leqslant C|x|.$$

For $A > 0$, let

$$T_A := \inf\{t \geqslant 0\,;\ |X_t| = A\}.$$

For $k \in \mathbb{N}^+$, let $S_k := T_{A\cdot 2^{k-1}} \wedge T_{A\cdot 2^{k+1}}$. Show that there exists $C' < \infty$ such that

$$\forall t \geqslant 0\ \ \forall A \geqslant 1\ \ \forall |x| = A\cdot 2^k$$
$$\mathbf{E}_x\big[|X_{t\wedge S_k} - x|^2\big] \leqslant C'\cdot A^2\cdot 2^{2k}(t+t^2) \quad \text{and} \quad \mathbf{P}_x[S_k \leqslant t] \leqslant 4C'(t+t^2).$$

(2) Show that there exists $t_0 > 0$ such that

$$\forall k \in \mathbb{N}^+\ \ \forall A \geqslant 1\ \ \forall |x| \geqslant A\cdot 2^k \qquad \mathbf{P}_x[T_{A\cdot 2^{k-1}} > t_0] \geqslant \frac{1}{2}.$$

Use the strong Markov property to deduce that for all $A \geqslant 1$,

$$\lim_{k\to\infty}\ \sup_{|x|\geqslant A\cdot 2^k} \mathbf{P}_x\Big[\big|\{j \in [1,k]\,;\ T_{A\cdot 2^j} = \infty \text{ or } T_{A\cdot 2^{j-1}} - T_{A\cdot 2^j} > t_0\}\big| < \frac{k}{4}\Big] = 0.$$

(3) Show that

$$\forall A \geqslant 1 \qquad \lim_{k\to\infty}\ \sup_{|x|\geqslant A\cdot 2^k} \mathbf{P}_x\Big[T_A < \frac{kt_0}{4}\Big] = 0.$$

Deduce that $\forall f \in C_0(\mathbb{R}^d)\ \ \forall t \geqslant 0\ \ Q_t f \in C_0(\mathbb{R}^d)$.

The words "diffusion process" are assigned different meanings by different authors, but one common meaning is a process with continuous sample paths that is Markov and solves a stochastic differential equation. The study of L as a second-order differential operator can be done via probability, as we will see (for Brownian motion) in Chapter 7.

Exercise (due 4/19). (1) What is the generator on $C_c^2(\mathbb{R})$ of the Ornstein–Uhlenbeck process? of geometric Brownian motion?

(2) Find Lipschitz σ and b such that the solution of $E(\sigma, b)$ has the generator

$$\begin{aligned} Lf(x_1,x_2) &= 2x_2 f_1(x_1,x_2) + \ln(1+x_1^2+x_2^2) f_2(x_1,x_2) \\ &\quad + \tfrac{1}{2}(1+x_1^2) f_{11}(x_1,x_2) + x_1 f_{12}(x_1,x_2) + \tfrac{1}{2} f_{22}(x_1,x_2) \end{aligned}$$

on $C_c^2(\mathbb{R}^2)$.

Exercise (due 4/19). Exercise 8.11: for (1), we consider only $\lambda > 0$; for (2), this means to show the Laplace transform of $Q_t(x,\cdot)$ is as on the bottom of page 178 of the book.

Exercise (due 4/19). Let B be 1-dimensional Brownian motion. Define

$$X_t^x := (x^{1/3} + \tfrac{1}{3}B_t)^3$$

for $t \geqslant 0$. Let $\sigma(x) := x^{2/3}$ and $b(x) = x^{1/3}/3$.

(1) Show that X^x solves $E_x(\sigma, b)$.
(2) Let $T := \inf\{t \geqslant 0\,;\ X_t = 0\}$. Show that $(X^x)^T$ is also a strong solution to $E_x(\sigma, b)$.

Exercise (due 4/26). Let σ and b be Lipschitz. Let

$$H := \{y \in \mathbb{R}^d\,;\ \sigma(y) = 0 \text{ and } b(y) = 0\}.$$

Show that if X^x solves $E_x(\sigma, b)$, then H is absorbing for X^x, i.e., if

$$T := \inf\{t \geqslant 0\,;\ X_t^x \in H\},$$

then $X^x = (X^x)^T$.

Extra credit: Show that if $x \notin H$, then $T = \infty$ almost surely.

Exercise. Suppose that $\mathrm{d}X_t = \mathrm{d}B_t + b(X_t)\,\mathrm{d}t$ and $X_0 = 0$, where $b(x) := c'(x)/\big(2c(x)\big)$ for some strictly positive function $c \in C^2(\mathbb{R})$. Define $\tau_0 := 0$ and recursively set $\tau_{k+1} := \inf\{t > \tau_k\,;\ |X_{\tau_{k+1}} - X_{\tau_k}| = 1\}$ for $k \geqslant 0$. Show that the discrete-time process $(X_{\tau_k}\,;\ k \geqslant 0)$ is a nearest-neighbor random walk on $\mathbb{Z}$ with transition probabilities $p_{n,n+1} = r_n/\big(r_n + r_{n+1}\big)$, where $r_n := \int_{n-1}^n \mathrm{d}x/c(x)$. (If we interpret $c(x)$ as the conductivity at x, then r_n is the resistance of the edge between $n-1$ and n in an electrical network on $\mathbb{Z}$.)

Chapter 7

Brownian Motion and Partial Differential Equations

We discuss the heat equation and especially Laplace's equation and how they can be solved using Brownian motion. This is a model for other partial differential equations and diffusions. We then discuss some path properties of Brownian motion.

7.1. Brownian Motion and the Heat Equation

For the whole chapter, B denotes d-dimensional Brownian motion with $\mathbf{P}_x$ the measure such that $\mathbf{P}_x[B_0 = x] = 1$, $Q_\bullet$ is its transition semigroup, and L is its generator. Recall that

$$D(L) \supseteq \left\{\psi \in C^2(\mathbb{R}^d)\,;\ \psi, \Delta\psi \in C_0(\mathbb{R}^d)\right\}$$

and on that set,

$$L\psi = \tfrac{1}{2}\Delta\psi.$$

For all $\varphi \in B(\mathbb{R}^d)$

$$\forall t > 0 \qquad Q_t\varphi = p_t * \varphi,$$

where p_t is the density of $\mathcal{N}(0, t\,I)$, a C^∞ function. Thus, $Q_t\varphi \in C^\infty$. If $\varphi \in C_0(\mathbb{R}^d)$, then all derivatives of $Q_t\varphi$ lie in $C_0(\mathbb{R}^d)$, as we may see by differentiating under the integral:

$$\partial_i Q_t\varphi = (\partial_i p_t) * \varphi.$$

Thus, for $\varphi \in C_0(\mathbb{R}^d)$, we have $Q_t\varphi \in D(L)$ and $L(Q_t\varphi) = \frac{1}{2}\Delta(Q_t\varphi)$ for $t > 0$.

Exercise. Exercise 7.28; add part (0): Let $s \mapsto \varphi_s$ be continuous from $\mathbb{R}_+$ to $C_0(\mathbb{R}^d)$. Show that $\int_0^t Q_s\varphi_s \,\mathrm{d}s \in C_0(\mathbb{R}^d) \cap C^1(\mathbb{R}^d)$ with gradient in $C_0(\mathbb{R}^d)$ and continuous in $t \in \mathbb{R}_+$, where $(Q_t)_t$ is the transition semigroup of Brownian motion. *Hint for (1):* write $\int_0^t = \int_0^{t/2} + \int_{t/2}^t$ for the second derivatives.

The following shows how Brownian motion solves the heat equation with initial value $\varphi \in C_0(\mathbb{R}^d)$. (Direct calculation shows more, but our proof extends to other equations involving the generator of a Feller process.)

Theorem 7.1. *Let $\varphi \in C_0(\mathbb{R}^d)$. For $t > 0$ and $x \in \mathbb{R}^d$, set*

$$u_t(x) := Q_t\varphi(x) = \mathbf{E}_x\big[\varphi(B_t)\big].$$

Then $(t, x) \mapsto u_t(x)$ on $(0, \infty) \times \mathbb{R}^d$ satisfies

$$\frac{\partial u_t}{\partial t} = \frac{1}{2}\Delta u_t \quad \text{and} \quad \lim_{\substack{t\downarrow 0\\ y\to x}} u_t(y) = \varphi(x).$$

Proof. Recall Proposition 6.11:

$$\forall f \in D(L)\ \ \forall t \geqslant 0 \qquad Q_t f = f + \int_0^t L(Q_s f)\,\mathrm{d}s.$$

We do not necessarily have $\varphi \in D(L)$, but we do have $Q_\varepsilon\varphi \in D(L)$ for $\varepsilon > 0$. Thus,

$$\forall t \geqslant \varepsilon > 0 \qquad u_t = u_\varepsilon + \int_0^{t-\varepsilon} L(Q_s u_\varepsilon)\,\mathrm{d}s = u_\varepsilon + \int_\varepsilon^t Lu_s\,\mathrm{d}s.$$

Now, we can also write

$$Lu_s = Q_{s-\varepsilon}(Lu_\varepsilon)$$

by Proposition 6.10 to see that $s \mapsto Lu_s$ is continuous on $[\varepsilon, \infty)$. Thus,

$$\forall t \geqslant \varepsilon > 0 \qquad \frac{\partial u_t}{\partial t} = Lu_t = \frac{1}{2}\Delta u_t.$$

The initial condition follows from the Feller property $Q_t\varphi \to \varphi$ as $t \downarrow 0$. ◀

Exercise. Let X be a Markov process on a locally compact Polish space E with Feller semigroup $(Q_t)_t$ and generator L. Let $\varphi \in C_0(E)$. Show that if $u_t(x) := Q_t\varphi(x)$, then

$$\frac{\partial u_t}{\partial t} = Lu_t \quad \text{and} \quad \lim_{\substack{t\downarrow 0\\ y\to x}} u_t(y) = \varphi(x)$$

for $t > 0$ and $x \in E$.

7.2. Brownian Motion and Harmonic Functions

Definition 7.2. *A **domain** of $\mathbb{R}^d$ is a non-empty, open, connected set. For a domain $D \subseteq \mathbb{R}^d$, a function $u\colon D \to \mathbb{R}$ is **harmonic** on D if $u \in C^2(D)$ and $\Delta u = 0$ on D.*

Suppose D' is a subdomain of D with $\overline{D'} \subseteq D$. Let

$$T := \inf\{t \geqslant 0\,;\ B_t \notin D'\}.$$

By Itô's formula, if u is harmonic on D and $B_0 \in D'$, then $u(B)^T$ is a continuous local martingale with

$$u(B_{t\wedge T}) = u(B_0) + \int_0^{t\wedge T} \nabla u(B_s)\cdot \mathrm{d}B_s,$$

like a line integral. In fact, if D' is bounded, then u is bounded on D', so $u(B)^T$ is a true martingale. Conversely, if $u \in C^2(D)$ and for all subdomains $D' \subseteq D$ with $\overline{D'} \subset D$, $u(B)^T$ is a continuous local martingale, then by Itô's formula, u is harmonic in D. We will weaken the hypothesis "$u \in C^2(D)$" with Lemma 7.5.

Proposition 7.3. *Let u be harmonic on a domain D. For a bounded subdomain D' of D with $\overline{D'} \subset D$, let*

$$T := \inf\{t \geqslant 0\,;\; B_t \notin D'\}.$$

Then

$$\forall x \in D' \qquad u(x) = \mathbf{E}_x\big[u(B_T)\big].$$

Proof. Because D' is bounded, $u(B)^T$ is a bounded $\mathbf{P}_x$-continuous local martingale, so it is a true martingale. Thus,

$$\forall t \geqslant 0 \qquad u(x) = \mathbf{E}_x\big[u(B_{t\wedge T})\big].$$

Since $T < \infty$ $\mathbf{P}_x$-a.s., we may let $t \to \infty$ and use the bounded convergence theorem to obtain the desired formula. ◂

Rotational symmetry of Brownian motion shows that if D' is a ball centered at x, then B_T has the uniform distribution on $\partial D'$. Let $\sigma_{x,r}$ denote the uniform measure on the sphere of radius r centered at x.

Proposition 7.4 (Mean-Value Property)**.** *If u is harmonic on a neighborhood of the closed ball of radius r centered at x, then*

$$u(x) = \int u(y)\,\mathrm{d}\sigma_{x,r}(y).$$ ◂

Exercise (due 4/26). Let u be harmonic on $\mathbb{R}^d$ and B be Brownian motion.

(1) Let $t > 0$. Show that

$$\forall y \in \mathbb{R}^d \qquad u(B)^t \text{ is a true } \mathbf{P}_y\text{-martingale}$$

if and only if

$$\forall y \in \mathbb{R}^d \qquad \int_{\mathbb{R}^d} \big|u(x)\big| \mathrm{e}^{-|x-y|^2/2t}\,\mathrm{d}x < \infty.$$

(2) Find u on $\mathbb{R}^2$ such that for all $y \in \mathbb{R}^2$, $\big(u(B_s)\big)_{0\leqslant s<1}$ is a true $\mathbf{P}_y$-martingale, but $\big(u(B_s)\big)_{0\leqslant s\leqslant 1}$ is not a true $\mathbf{P}_y$-martingale.

Exercise (due 4/26). (1) Show that every bounded harmonic function on $\mathbb{R}^2$ is constant by using Exercise 5.33(5).

(2) Show the same on $\mathbb{R}^d$ for $d > 2$. *Hint:* Let $x \neq y$ and let H be the hyperplane

$$\big\{z \in \mathbb{R}^d\,;\; |z-x| = |z-y|\big\}.$$

Let $T := \inf\{t \geqslant 0\,;\; B_t \in H\}$. Show that

$$\mathbf{E}_x\big[u(B_T)\big] = \mathbf{E}_y\big[u(B_T)\big].$$

(3) Let u be a nonconstant harmonic function on $\mathbb{R}^d$ ($d \geqslant 2$). Show that

$$\forall p > 1 \qquad \sup_{t>0} \int_{\mathbb{R}^d} \big|u(tx)\big|^p \mathrm{e}^{-|x|^2}\,\mathrm{d}x = \infty.$$

(4) (Extra credit) Does (3) hold for $p = 1$?

(5) (Extra credit) Show that every positive harmonic function on $\mathbb{R}^d$ ($d \geqslant 1$) is constant.

We say that a locally bounded, measurable function u on D satisfies the ***mean-value property*** if the equation of Proposition 7.4 holds for all closed balls in D.

Lemma 7.5. *If u satisfies the mean-value property on a domain D, then u is harmonic on D.*

Thus, if u is locally bounded and measurable, and $u(B)^T$ is a continuous local martingale for every exit time T of a closed ball centered at the starting point and contained in D, then u is harmonic.

Proof. It suffices to show that for all $r_0 > 0$, if $D' := \{x \in D\,;\ |x - D^{\mathrm{c}}| > r_0\}$, then u is harmonic in D'.

We first show that $u \in C^\infty(D')$. Choose any $h\colon \mathbb{R}_+ \to \mathbb{R}_+$ that is C^∞, has support in $(0, r_0)$, and is not identically zero. For $0 < r < r_0$, multiply both sides of

$$u(x) = \int u(y)\,\mathrm{d}\sigma_{x,r}(y) \quad (x \in D')$$

by $r^{d-1}h(r)$ and integrate from $r = 0$ to r_0. We get, for some constant $c > 0$, that

$$\forall x \in D' \qquad cu(x) = \int_{|y|<r_0} u(x+y)h\big(|y|\big)\,\mathrm{d}y = \int_{\mathbb{R}^d} u(x+y)h\big(|y|\big)\,\mathrm{d}y$$

if we set u to be 0 outside D. We can rewrite this as a convolution:

$$cu(x) = \int_{\mathbb{R}^d} u(z)h\big(|z-x|\big)\,\mathrm{d}z.$$

Since $z \mapsto h\big(|z|\big) \in C^\infty(\mathbb{R}^d)$, we get $u \in C^\infty(D')$.

To show $\Delta u = 0$ in D', we may now apply Itô's formula to $u(B)$: Let

$$T_{x,r} := \inf\{t \geqslant 0\,;\ |x - B_t| = r\}.$$

Then

$$\forall x \in D' \ \ \forall r \in (0, r_0) \qquad \mathbf{E}_x\big[u(B_{t\wedge T_{x,r}})\big] = u(x) + \tfrac{1}{2}\,\mathbf{E}_x\Big[\int_0^{t\wedge T_{x,r}} \Delta u(B_s)\,\mathrm{d}s\Big].$$

Recall that $\mathbf{E}_x[T_{x,r}] < \infty$ (in fact, $\mathbf{E}_x[T_{x,r}] = r^2/d$). Thus, we may let $t \to \infty$ and apply Lebesgue's dominated convergence theorem to get

$$\mathbf{E}_x\big[u(B_{T_{x,r}})\big] = u(x) + \tfrac{1}{2}\,\mathbf{E}_x\Big[\int_0^{T_{x,r}} \Delta u(B_s)\,\mathrm{d}s\Big].$$

The left-hand side equals $u(x)$, therefore the second term on the right-hand side equals 0. Now, let $r \downarrow 0$ to get $\Delta u(x) = 0$. ◀

Definition 7.6. *Let D be a domain and $g \in C(\partial D)$. We say that $u\colon D \to \mathbb{R}$* ***solves the Dirichlet problem in*** D ***with boundary condition*** g *if u is harmonic in D and*

$$\forall y \in \partial D \qquad \lim_{D \ni x \to y} u(x) = g(y).$$

Thus, if

$$\widetilde{u}(x) := \begin{cases} u(x) & \text{if } x \in D, \\ g(x) & \text{if } x \in \partial D, \end{cases}$$

then $\widetilde{u} \in C(\overline{D})$. If D is bounded, then u is bounded.

Exercise (due 4/26). Let D be a bounded domain, $g \in B(\partial D)$, and

$$T := \inf\{t \geqslant 0\,;\ B_t \notin D\}.$$

Define

$$v(x) := \mathbf{E}_x\big[g(B_T)\big] \quad \text{for } x \in D.$$

Show that $v \in C(D)$.

Proposition 7.7. *Keep the notation of the preceding exercise.*

(i) *If $g \in C(\partial D)$ and u solves the Dirichlet problem in D with boundary condition g, then $u = v$.*

(ii) *The function v is harmonic in D and*

$$\forall x \in D \qquad \lim_{t \uparrow T} v(B_t) = g(B_T) \quad \mathbf{P}_x\text{-}a.s.$$

Proof. (i) In Proposition 7.3, we saw this formula for subdomains D' with $\overline{D'} \subseteq D$. Take an increasing sequence $D_n \subseteq D$ with $\overline{D_n} \subseteq D$ and $\bigcup_n D_n = D$. Apply continuity of sample paths and of $\widetilde{u}$.

(ii) The exercise shows $v \in C(D)$, and obviously v is bounded. (Or: measurability follows from the start of Theorem 6.16 and measurability of g and B_T.) The mean-value property is a consequence of the strong Markov property: If $|x - D^{\mathrm{c}}| > r$, then

$$v(x) = \mathbf{E}_x\Big[\mathbf{E}_x\big[g(B_T) \mid \mathscr{F}_{T_{x,r}}\big]\Big] = \mathbf{E}_x\big[v(B_{T_{x,r}})\big].$$

Here, we use the strong Markov property in the forms of both Theorem 6.17 and Theorem 2.20. Thus, v is harmonic in D. The proof of the rest of (ii), which we won't use, is in an appendix. ◂

These results do not say when the Dirichlet problem has a solution. In fact, it need not:

Exercise (due 4/26). Exercise 7.24, Exercise 7.25.

However, convex domains have a solution for all g. In fact, it suffices that every point of ∂D satisfy the ***exterior cone condition***, where $y \in \partial D$ satisfies this if there exists a non-empty open cone C with apex y and there exists $r > 0$ such that $C \cap \big\{z\,;\ |z - y| < r\big\} \subseteq D^{\mathrm{c}}$. The idea is that if $x \in D$ is close to ∂D, then it is $\mathbf{P}_x$-likely that B_t leaves D close to x.

Lemma 7.9. *Let D be a domain that satisfies the exterior cone condition at some $y \in \partial D$. Let $T := \inf\{t \geqslant 0\,;\ B_t \notin D\}$. Then the $\mathbf{P}_x$-law of T tends weakly to δ_0 as $D \ni x \to y$.*

Proof. Let $\mathscr{B}_r := \{z \in \mathbb{R}^d\,;\ |z| < r\}$. Let C be an open circular cone whose apex is 0 and whose intersection with the unit sphere has normalized measure $\alpha > 0$ such that $y + (C \cap \mathscr{B}_r) \subseteq D^{\mathrm{c}}$ for some $r > 0$. Then $\lim_{t\to 0} \mathbf{P}_0[B_t \in C \cap \mathcal{B}_r] = \alpha$. Blumenthal's 0-1 law (Theorem 2.13) extends to higher-dimensional Brownian motion with the same proof, whence $\mathbf{P}_0[T_{C\cap\mathcal{B}_r} = 0] = 1$, where $T_F := \inf\{t \geqslant 0\,;\ B_t \in F\}$.

Let $C' \subseteq C$ be an open circular cone with apex 0 and opening $\alpha/2$ and the same axis of symmetry as C. Then $\mathbf{P}_0[T_{C'\cap\mathcal{B}_r} = 0] = 1$ as well. Given $\eta > 0$, there exists $a > 0$ such that $\mathbf{P}_0[T_{C'_a\cap\mathcal{B}_{r/2}} \leqslant \eta] \geqslant 1-\eta$, where $C'_a := \{z \in C'\,;\ |z| > a\}$ (because $C'_a \uparrow C'$ as $a \downarrow 0$). Choose $\varepsilon > 0$ such that

$$|z| < \varepsilon \quad \Longrightarrow \quad C'_a \cap \mathcal{B}_{r/2} \subseteq z + C \cap \mathcal{B}_r.$$

Then for $|y - x| < \varepsilon$,

$$\begin{aligned}\mathbf{P}_x[T \leqslant \eta] \geqslant \mathbf{P}_x[T_{y+C\cap\mathcal{B}_r} \leqslant \eta] &= \mathbf{P}_0[T_{y-x+C\cap\mathcal{B}_r} \leqslant \eta]\\ &\geqslant \mathbf{P}_0[T_{C'_a\cap\mathcal{B}_{r/2}} \leqslant \eta] \geqslant 1-\eta.\end{aligned}$$ ◀

Theorem 7.8. *Let D be a bounded domain in $\mathbb{R}^d$ that satisfies the exterior cone condition at every point of ∂D. Then for all $g \in C(\partial D)$, the Dirichlet problem in D with boundary condition g has a solution.*

Proof. Let v be as in the exercise. By Proposition 7.7(ii), we need only show that

$$\forall y \in \partial D \qquad \lim_{D\ni x\to y} v(x) = g(y).$$

In fact, we show this holds for each y where the exterior cone condition holds and where g is continuous, regardless of other points on ∂D. The idea is that for x close to y, it is $\mathbf{P}_x$-likely that T is small (from Lemma 7.9) and thus that B_T is close to y, whence that $g(B_T)$ is close to $g(y)$.

Let $\varepsilon > 0$. Choose $\delta > 0$ such that $|g(z) - g(y)| < \varepsilon/3$ for $|z - y| < \delta$, $z \in \partial D$. Choose $\eta > 0$ such that

$$2\|g\|\,\mathbf{P}_0\Big[\sup_{t\leqslant\eta} |B_t| > \tfrac{\delta}{2}\Big] < \tfrac{\varepsilon}{3}.$$

By Lemma 7.9, we may choose $\alpha \in (0, \frac{\delta}{2})$ such that

$$2\|g\|\,\mathbf{P}_x[T > \eta] < \tfrac{\varepsilon}{3} \quad \text{for } |x-y| < \alpha,\ x \in D.$$

We obtain that for $|x - y| < \alpha$, $x \in D$,

$$\begin{aligned}|v(x) - g(y)| &\leqslant \mathbf{E}_x\big[|g(B_T) - g(y)|\mathbf{1}_{[T\leqslant\eta]}\big] + \mathbf{E}_x\big[|g(B_T) - g(y)|\mathbf{1}_{[T>\eta]}\big]\\ &\leqslant \mathbf{E}_x\big[|g(B_T) - g(y)|\mathbf{1}_{[T\leqslant\eta]\cap[\sup_{t\leqslant\eta}|B_t - x|\leqslant\delta/2]}\big]\\ &\quad + 2\|g\|\,\mathbf{P}_x\Big[\sup_{t\leqslant\eta}|B_t - x| > \delta/2\Big] + 2\|g\|\,\mathbf{P}_x[T > \eta]\\ &< \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} + \tfrac{\varepsilon}{3} = \varepsilon.\end{aligned}$$ ◀

For $d = 2$, another sufficient condition for Lemma 7.9 is that y belongs to a nonconstant curve contained in ∂D. To see this, make Brownian motion behave like in the following figure.

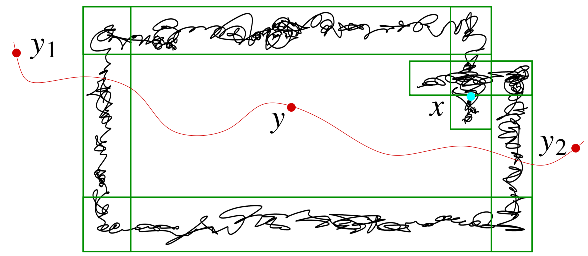


Figure: Choose two points, y_1 and y_2, on the curve near y that are separated by y. A Brownian motion started at x has a probability that is bounded below over all x near y that it will create a curve (up to some time) surrounding y and hitting ∂D only near y. For example, consider the event that it stays within the union of the green rectangles, moving successively from one of the 6 rectangles to the next until it is guaranteed to cross itself.

7.3. Harmonic Functions in a Ball and the Poisson Kernel

The $\mathbf{P}_x$-law of $B_T \in \partial D$ is called the ***harmonic measure of*** D ***relative to*** x. When D is a ball, this has a very simple expression: It has a density $K(x, \cdot)$ with respect to normalized surface measure, σ. Note that the strong Markov property shows that $K(x, y)$ is harmonic in x for each $y \in \partial D$. Also, for $z \in \partial D$,

$$\lim_{D \ni x \to z} K(x, \cdot)\sigma = \delta_z \quad \text{weakly.}$$

This suggests some properties to look for when finding K.

For the rest of this section, let $D = \mathcal{B}_1$, the open unit ball in $\mathbb{R}^d$, $d \geqslant 2$.

Definition 7.10. *The* ***Poisson kernel*** *is the function* $K\colon \mathcal{B}_1 \times \partial\mathcal{B}_1 \to \mathbb{R}_+$ *defined by*

$$K(x, y) := \frac{1 - |x|^2}{|y - x|^d}.$$

Lemma 7.11. *For all* $y \in \partial\mathcal{B}_1$, $K(\cdot, y)$ *is harmonic on* $\mathcal{B}_1$.

Proof. Clearly, $K(\cdot, y) \in C^\infty(\mathcal{B}_1)$. A direct calculation shows that

$$\Delta K(\cdot, y) = 0$$

off $\partial\mathcal{B}_1$ (see an appendix to this chapter). ◀

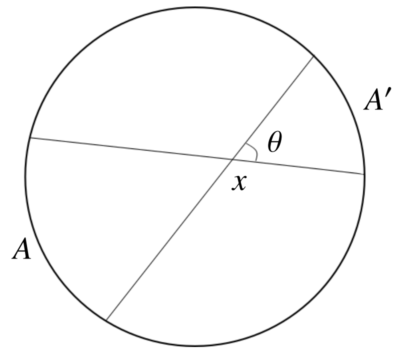
There is a beautiful geometric representation of $K(x, \cdot)\sigma_1$ due to Malmheden in 1934, where $\sigma_1 := \sigma_{0,1}$. Namely, given $g \in C(\partial\mathcal{B}_1)$, for each line L through $x \in \mathcal{B}_1$, let $f(L)$ denote the value at $x \in L$ of the linear function on L whose values at $L \cap \partial\mathcal{B}_1$ are g. Then the harmonic extension u of g satisfies

$$u(x) = \int f(L)\, \mathrm{d}L,$$

where $\mathrm{d}L$ denotes a uniform direction for lines L that pass through x.

Equivalently, if $A \in B(\partial\mathcal{B}_1)$ and A' is its image on $\partial\mathcal{B}_1$ reflected in x, then the harmonic measure of A equals $\sigma_1(A')$.

In $\mathbb{R}^2$, this is due to Schwarz. It is easy to prove in another form. Recall Euclid's theorem that if A is an arc, then $\sigma_1(A) + \sigma_1(A') = 2\theta$, where θ is the angle at x of the chords giving A and A':



Therefore, $\sigma_1(A') = 2\theta - \sigma_1(A)$. It is not hard to check that this is indeed the harmonic measure of A from x by checking boundary values and by representing θ using the imaginary part of a holomorphic function.

Lemma 7.13. *We have*

$$\forall x \in \mathcal{B}_1 \qquad \int_{\partial\mathcal{B}_1} K(x, y)\,\mathrm{d}\sigma_1(y) = 1,$$

where $\sigma_1 \coloneqq \sigma_{0,1}$.

Proof. For $x \in \mathcal{B}_1$, write $F(x)$ for the integral in the lemma. We claim that F satisfies the mean-value property because $K(\cdot, y)$ does. Clearly, F is locally bounded and measurable. Now if $0 < r < 1 - |x|$, then

$$\begin{aligned}\int F(z)\,\mathrm{d}\sigma_{x,r}(z) &= \iint K(z, y)\,\mathrm{d}\sigma_1(y)\,\mathrm{d}\sigma_{x,r}(z)\\ &= \iint K(z, y)\,\mathrm{d}\sigma_{x,r}(z)\,\mathrm{d}\sigma_1(y) = \int K(x, y)\,\mathrm{d}\sigma_1(y)\\ &= F(x),\end{aligned}$$

as desired. Also, F is rotationally symmetric because K is diagonally invariant under rotations and σ_1 is invariant. Therefore,

$$1 = F(0) = \int F(x)\,\mathrm{d}\sigma_{0,r}(x) \quad \text{for } 0 < r < 1$$

implies

$$F(x) = F(0) \quad \text{for } |x| = r,$$

i.e., $F \equiv 1$. ◂

Theorem 7.14. *If $g \in C(\partial\mathcal{B}_1)$, then the solution to the Dirichlet problem in $\mathcal{B}_1$ with boundary condition g is*

$$u(x) \coloneqq \int_{\partial\mathcal{B}_1} K(x, y) g(y)\,\mathrm{d}\sigma_1(y) \qquad (x \in \mathcal{B}_1).$$

Proof. A Fubini argument as in the proof of Lemma 7.13 shows that u satisfies the mean-value property, so is harmonic. It is clear that the probability measures $K(x,\cdot)\sigma_1 \Rightarrow \delta_z$ as $\mathcal{B}_1 \ni x \to z \in \partial\mathcal{B}_1$. Therefore, $u(x) \to g(z)$ as $\mathcal{B}_1 \ni x \to z \in \partial\mathcal{B}_1$. ◀

Corollary 7.15. *The harmonic measure of $\mathcal{B}_1$ relative to $x \in \mathcal{B}_1$ is $K(x,\cdot)\sigma_1$.* ◀

Exercise. Exercise 7.26.

7.4. Transience and Recurrence of Brownian Motion

Theorem 7.17. *The following hold.*

(i) *For $d = 2$, Brownian motion is **(neighborhood) recurrent**, meaning that almost surely, for all open $U \subseteq \mathbb{R}^2$, $\{t \geqslant 0\,;\ B_t \in U\}$ is unbounded.*

(ii) *For $d \geqslant 3$, Brownian motion is **transient**, meaning that almost surely, $|B_t| \to \infty$.*

Proof. We saw (ii) in Exercise 5.33(7). We also saw in Exercise 5.33(5) that for $x \neq 0$, $\mathbf{P}_x[\forall t \geqslant 0\ \ B_t \neq 0] = 1$, while from the same formula as there, for $d = 2$,

$$\forall \varepsilon > 0 \qquad \mathbf{P}_x\big[\exists t\ \ |B_t| < \varepsilon\big] = 1.$$

Combining these two facts, we get that for every neighborhood U of 0, $\{t \geqslant 0\,;\ B_t \in U\}$ is unbounded $\mathbf{P}_x$-a.s. The same holds for every ball with rational center and rational radius simultaneously almost surely by a similar argument, whence (i) holds. ◀

7.5. Planar Brownian Motion and Holomorphic Functions

Let $d = 2$, identify $\mathbb{R}^2$ with $\mathbb{C}$, and write $B_t = X_t + \mathrm{i}Y_t$, where X and Y are independent real Brownian motions. We call B a ***complex Brownian motion***. Let $D \subseteq \mathbb{C}$ be a domain and $\Phi\colon D \to \mathbb{C}$ be analytic. Since $\operatorname{Re}\Phi$ and $\operatorname{Im}\Phi$ are harmonic, $\Phi(B)^T$ is a continuous local martingale, where T is the exit time of D. Much more is true:

Theorem 7.18 (Lévy)**.** *Suppose $\mathbb{C} \setminus D$ is polar. Write*

$$C_t := \int_0^t \left|\Phi'(B_s)\right|^2 \mathrm{d}s \quad (t \geqslant 0).$$

Then for each $z \in D$, there exists a complex Brownian motion Γ started from $\Phi(z)$ such that $\mathbf{P}_z$-a.s.,

$$\forall t \geqslant 0 \qquad \Phi(B_t) = \Gamma_{C_t}.$$

That is, $\Phi(B)$ is a time-changed complex Brownian motion with clock C; this is called the ***conformal invariance property***. The case $\Phi(z) = az$ is the usual Brownian scaling for $a \in \mathbb{R}$ and is rotation invariance for $|a| = 1$. This shows why Theorem 7.18 is true on an infinitesimal level. If $\mathbb{C} \setminus D$ is not polar, then a similar conclusion holds for the process $\Phi(B)^T$.

Proof. Let $\Phi = g + \mathrm{i}h$, where g and h are real harmonic. Write $M := g(B)$ and $N := h(B)$. By Itô's formula,

$$\mathrm{d}M = g_x(B)\,\mathrm{d}X + g_y(B)\,\mathrm{d}Y, \quad \mathrm{d}N = h_x(B)\,\mathrm{d}X + h_y(B)\,\mathrm{d}Y,$$

so M and N are continuous local martingales. The Cauchy–Riemann equations,

$$g_x = h_y, \quad g_y = -h_x,$$

imply that $\langle M, N\rangle = 0$ and $\langle M, M\rangle = \langle N, N\rangle = C$.

Recall Exercise 4.24, which shows that

$$[C_\infty < \infty] \overset{\text{a.s.}}{=} [\lim_{t\to\infty} M_t \text{ exists in } \mathbb{R}] \overset{\text{a.s.}}{=} [\lim_{t\to\infty} N_t \text{ exists in } \mathbb{R}].$$

By neighborhood recurrence of B, if Φ is not constant, then $\Phi(B_t)$ does not have a finite limit as $t \to \infty$, whence $C_\infty = \infty$ almost surely. Of course, if Φ is constant, then $C \equiv 0$ and nothing more is needed. Thus, in the nonconstant case, we may apply Proposition 5.15 to $M - g(z)$ and $N - h(z)$ under $\mathbf{P}_z$ to obtain independent real Brownian motions β and γ started from 0 such that

$$M_t - g(z) = \beta_{C_t} \quad \text{and} \quad N_t - h(z) = \gamma_{C_t}.$$

Thus, the result holds with $\Gamma := \Phi(z) + \beta + \mathrm{i}\gamma$. ◀

In this proof, one could also write in complex notation $\mathrm{d}\Phi(B) = \Phi'(B)\,\mathrm{d}B$. More generally, if $X^1, \ldots, X^p$ are continuous semimartingales taking values in $\mathbb{C}$ and $F\colon \mathbb{C}^p \to \mathbb{C}$ is analytic in each variable, then Itô's formula takes exactly the same form as in Theorem 5.10, where the bracket is now complex valued and still bilinear, not sesquilinear. (Hartog's theorem guarantees that such an F has a multivariable power series expansion in a neighborhood of each point.) In applications of this, note that for complex Brownian motion, we have $\langle B, B\rangle = 0$. Bilinearity also guarantees that all parts of Proposition 4.15 hold for complex-valued, continuous local martingales. Complex-valued, continuous local martingales Z that satisfy $\langle Z, Z\rangle = 0$ are called ***conformal***; they are time changes of complex Brownian motion, as we can see by the second half of the proof of Lévy's theorem.

Exercise. Determine all Φ such that in the preceding proof, M and N are independent.

Exercise. Exercise 7.27.

Exercise. Use the result of Exercise 7.27 to show that every nonconstant complex polynomial has a root. *Hint:* note that $\{z\,;\ |P(z)| \leqslant \varepsilon\}$ is compact if P is a polynomial.

Exercise. Suppose that H is complex-valued and progressive and that B is complex Brownian motion. Show that if $|H|^2$ is locally integrable with respect to Lebesgue measure and $Z = H \cdot B$, then there exists a complex Brownian motion Γ such that $Z_t = \Gamma_{C_t}$ for $t \geqslant 0$, where $C_t := \int_0^t |H_s|^2\,\mathrm{d}s$ for $t \geqslant 0$. In particular, if $|H| = 1$, then Z is a complex Brownian motion.

It is also interesting to look at Brownian motion in polar coordinates. This yields the ***skew-product representation*** (or ***decomposition***):

Theorem 7.19. *Fix* $z \in \mathbb{C} \setminus \{0\}$ *and choose any* $w \in \mathbb{C}$ *with* $z = \mathrm{e}^w$. *There exists a complex Brownian motion* β *starting at* w *such that* $\mathbf{P}_z$*-a.s.,*

$$\forall t \geqslant 0 \qquad B_t = \exp\{\beta_{H_t}\},$$

where

$$H_t := \int_0^t \frac{\mathrm{d}s}{|B_s|^2}.$$

The point is that $\operatorname{Re}\beta$ describes the radial motion of B while $\operatorname{Im}\beta$ describes the angular motion of B.

Proof. This is intuitive from Theorem 7.18 by using

$$\Phi(\zeta) := \log \zeta.$$

Note $\Phi'(\zeta) = 1/\zeta$. However, this Φ is multiple valued and would require an extension to Riemann surfaces.

Instead, let us start with a complex Brownian motion Γ that starts from w and use $\Phi(\zeta) := \mathrm{e}^{\zeta}$. Then by Theorem 7.18, we have

$$\mathrm{e}^{\Gamma_t} = Z_{C_t},$$

where Z is a complex Brownian motion from z and

$$C_t = \int_0^t |\mathrm{e}^{\Gamma_s}|^2 \,\mathrm{d}s = \int_0^t \mathrm{e}^{2\operatorname{Re}\Gamma_s}\,\mathrm{d}s.$$

Let $K_\bullet$ be the inverse function of $C_\bullet$: by calculus,

$$K_s = \int_0^s \exp\{-2\operatorname{Re}\Gamma_{K_u}\}\,\mathrm{d}u = \int_0^s \frac{\mathrm{d}u}{|Z_u|^2}.$$

Then $Z_s = \mathrm{e}^{\Gamma_{K_s}}$. This is what we want, except it is for Z rather than for B. But the formula $B_t = \exp\{\beta_{H_t}\}$ together with the formula for H gives β as a deterministic function of B. When applied to Z, it gives Γ. Since $B \overset{\mathscr{D}}{=} Z$, it follows that $\beta \overset{\mathscr{D}}{=} \Gamma$, as desired. ◀

Exercise. For which $t > 0$ is $\mathbf{E}[H_t^{1/2}] < \infty$? *Hint:* is $\log|B|$ a true martingale?

Exercise. Exercise 7.29.

Exercise. Let $B := (B_t)_{t\geqslant 0}$ be Brownian motion in the complex plane. Suppose that $B_0 = 1$.

(1) Let T_1 be the first time that B hits the imaginary axis, T_2 be the first time after T_1 that B hits the real axis, T_3 be the first time after T_2 that B hits the imaginary axis, etc. Prove that for each $n \geqslant 1$, the probability that $|B_{T_n}| \leqslant 1$ is 1/2.

(2) More generally, let ℓ_n be lines through 0 for $n \geqslant 1$ such that $1 \notin \ell_1$. Let $T_1 := \inf\{t \geqslant 0\,;\ B_t \in \ell_1\}$, and recursively define $T_{n+1} := \inf\{t > T_n\,;\ B_t \in \ell_{n+1}\}$ for $n \geqslant 1$. Prove that for each $n \geqslant 1$, the probability that $|B_{T_n}| \leqslant 1$ is 1/2.

(3) (Extra credit) In the context of part (2), let α_n be the smaller of the two angles between ℓ_n and ℓ_{n+1}. Show that $\sum_{n=1}^{\infty} \alpha_n = \infty$ iff for all $\varepsilon > 0$, the probability that $\varepsilon \leqslant |B_{T_n}| \leqslant 1/\varepsilon$ tends to 0 as $n \to \infty$.

(4) (Extra credit) In the context of part (1), show that

$$\lim_{n\to\infty} \mathbf{P}\Big[\exp\big(-\delta_n\sqrt{n}\,\big) \leqslant |B_{T_n}| \leqslant \exp\big(\delta_n\sqrt{n}\,\big)\Big] = \int_{-2\delta/\pi}^{2\delta/\pi} \frac{\mathrm{e}^{-u^2/2}}{\sqrt{2\pi}}\,\mathrm{d}u$$

if $\delta_n \geqslant 0$ tend to $\delta \in [0, \infty]$.

Exercise. Let B be a complex Brownian motion not starting from 0. Let A_t be an argument of B_t, so that $B_t = |B_t|\,\mathrm{e}^{\mathrm{i}A_t}$. Assume that A has been chosen to be continuous. Show how to reconstruct B from A; more formally, show that B is adapted to $\overline{\mathscr{F}_\bullet^A}$.

7.6. Asymptotic Laws of Planar Brownian Motion

Let B be complex Brownian motion. Our first result is not asymptotic. If $B_0 = a \cdot \mathrm{i}$, $a > 0$, and $T := \inf\{t \geqslant 0\,;\ \operatorname{Im} B_t = 0\}$, what is the distribution of B_T? By scaling, it is a times the distribution when $a = 1$: the $\mathbf{P}_{a\mathrm{i}}$-law of B_T equals the $\mathbf{P}_{\mathrm{i}}$-law of $a \cdot B_T$.

Proposition. *If $B_0 = \mathrm{i}$ and $T := \inf\{t \geqslant 0\,;\ \operatorname{Im} B_t = 0\}$, then B_T has the standard symmetric Cauchy distribution,*

$$\mathbf{P}_{\mathrm{i}}[B_T \leqslant x] = \int_{-\infty}^{x} \frac{\mathrm{d}y}{\pi(1+y^2)} = \frac{1}{2} + \frac{1}{\pi}\arctan(x).$$

Proof. The function

$$\varphi(z) := \mathrm{i} \cdot \frac{1-z}{1+z}$$

maps the unit disk to the upper half plane with $\varphi(0) = \mathrm{i}$ and $\varphi(\mathrm{e}^{2\pi\mathrm{i}\alpha}) = \tan(\pi\alpha)$. Let

$$S := \inf\{t \geqslant 0\,;\ |B_t| = 1\}.$$

The $\mathbf{P}_{\mathrm{i}}$-law of B_T equals the $\mathbf{P}_0$-law of B_S pushed forward by φ, in view of Theorem 7.18. Because the $\mathbf{P}_0$-law of B_S is the uniform measure, it follows that

$$\mathbf{P}_{\mathrm{i}}[B_T \leqslant x] = \mathbf{P}_0\big[\varphi(B_S) \leqslant x\big] = \mathbf{P}_0\big[\arg B_S \in (-\pi, 2\pi\alpha]\big],$$

where $\tan(\pi\alpha) = x$. This gives the result. ◀

Exercise. (1) Let $C_1 \perp\!\!\!\perp C_2$ be standard Cauchy and $a_1, a_2 > 0$. Show that $a_1C_1 + a_2C_2$ has the law of $a_1 + a_2$ times standard Cauchy.

(2) Let B be complex Brownian motion starting at 0. For $s \geq 0$, write $T_s := \inf\{t \geqslant 0\,;\ \operatorname{Im} B_t = s\}$ and $C_s := \operatorname{Re} B_{T_s}$. Show that the process C has independent, stationary increments and that C_s has the law of s times standard Cauchy.

Exercise. Let $(B^1, B^2, \ldots, B^{d+1})$ be Brownian motion in $\mathbb{R}^{d+1}$ starting at $(0, 0, \ldots, 0, 1)$. Let $T := \inf\{t \geqslant 0\,;\ B_t^{d+1} = 0\}$. Use Corollary 2.22 to show that $(B_T^1, \ldots, B_T^d)$ has density

$$x \mapsto \frac{\Gamma\left(\frac{d+1}{2}\right)}{\left(\pi(1+|x|^2)\right)^{(d+1)/2}}$$

on $\mathbb{R}^d$, where $\Gamma(a) := \int_0^\infty s^{a-1} \mathrm{e}^{-s}\, \mathrm{d}s$ is the usual Gamma function. This is called the ***standard d-dimensional (multivariate) Cauchy distribution***. Because of its connection to Brownian motion, it follows that when $d' < d$, every d'-dimensional marginal of a standard d-dimensional Cauchy distribution is a standard d'-dimensional Cauchy distribution. Use the fact that the characteristic function of the standard 1-dimensional Cauchy distribution is $\xi \mapsto \mathrm{e}^{-|\xi|}$ ($\xi \in \mathbb{R}$) to deduce that the characteristic function of the standard d-dimensional Cauchy distribution is $\xi \mapsto \mathrm{e}^{-|\xi|}$ ($\xi \in \mathbb{R}^d$). *Hint:* For $\xi_1, \ldots, \xi_d \in \mathbb{R}$, what is the law of $\sum_{j=1}^d \xi_j B_T^j$?

Now we look at the winding of Brownian motion about 0 and its distance from 0 separately. Let θ_t be a continuous process such that

$$\frac{B_t}{|B_t|} = \mathrm{e}^{\mathrm{i}\theta_t} \qquad (B_t \neq 0,\ \theta_0 \in (-\pi, \pi]).$$

If $B_t = \mathrm{e}^{\beta_{H_t}}$ as in Theorem 7.19, then $\theta_t = \operatorname{Im} \beta_{H_t}$. Because $H_t \to \infty$ almost surely and $\operatorname{Im} \beta$ is recurrent, it follows that θ is recurrent as well. Winding happens both when $|B|$ is small (when it is fast) and when $|B|$ is large (infinitely often—consider $1/B$). How large is θ_t typically?

Theorem 7.20 (Spitzer)**.** *For all $z \neq 0$*

$$\text{the } \mathbf{P}_z\text{-law of } \frac{\theta_t}{\log \sqrt{t}} \Rightarrow \text{standard symmetric Cauchy distribution}$$

as $t \to \infty$.

Proof. (David Williams) Write $B_t = z + \beta_t$, so that when $B_0 = z$, we have $\beta_0 = 0$. In particular, by Brownian scaling, the $\mathbf{P}_z$-law of B_{1/δ^2} equals the law of $z + \delta^{-1}\beta_1$ when $\beta_0 = 0$. Because the angle does not change when we multiply the location by δ, it follows that the conclusion is equivalent to:

$$\lim_{\delta \downarrow 0} \mathbf{P}_z\text{-law of } \frac{\theta_{1/\delta^2}}{\log \frac{1}{\delta}} = \lim_{\delta \downarrow 0} \mathbf{P}_{\delta z}\text{-law of } \frac{\theta_1}{\log \frac{1}{\delta}}$$

is standard Cauchy.

Fix (small) $a > 0$ such that $\mathbf{P}_0\big[|B_1| > a\big]$ is close to 1. Let $T := \inf\big\{t \geqslant 0\,;\ |B_t| \geqslant a\big\}$. Let α satisfy the property that for all z with $|z| = a$, $\mathbf{P}_z\big[|\theta_1| > \alpha\big]$ is small; then also for all δ with $\delta|z| < a$, $\mathbf{P}_{\delta z}\big[|\theta_1 - \theta_T| > \alpha\big]$ is small, so we need concern ourselves only with the winding between times 0 and T, rather than between 0 and 1, i.e., with θ_T.

Consider $B_t = \mathrm{e}^{\beta_{H_t}}$; the law of θ_T is the law of $\operatorname{Im} \beta_{H_T}$, where H_T is the time that $\operatorname{Re} \beta_{H_T}$ goes from $\log(\delta|z|)$ to $\log a$. Thus, the law of θ_T is $\big|\log \delta|z| - \log a\big|$ times standard Cauchy. Since we are dividing θ_1 by $\log \frac{1}{\delta}$, this gives the result. ◀

For the radial part, we know $\min\big\{|B_s|\,;\ 0 \leqslant s \leqslant t\big\} \to 0$ as $t \to \infty$; how fast?

Proposition 7.22. *For every $z \neq 0$, we have*

$$\forall a > 0 \qquad \lim_{t\to\infty} \mathbf{P}_z\Big[\min_{0\leqslant s\leqslant t} |B_s| \leqslant t^{-a/2}\Big] = \frac{1}{1+a}.$$

Proof. Choose $b > 0$ so large that $\mathbf{P}_0[\frac{1}{b} < \max_{0\leqslant s\leqslant 1}|B_s| < b]$ is close to 1. Then for fixed z,

$$\mathbf{P}_z\Big[\frac{\sqrt{t}}{b} < \max_{0\leqslant s\leqslant t} |B_s| < b\sqrt{t}\,\Big] = \mathbf{P}_z\big[T_{\sqrt{t}/b} < t < T_{b\sqrt{t}}\big]$$

is close to 1 uniformly in t for t sufficiently large. (Note that s is the time variable, not t.) Now, $\min_{s\leqslant t} |B_s| \leqslant t^{-a/2}$ if and only if $T_{t^{-a/2}} \leqslant t$. For $c > 0$, we have (if $t^{-a/2} \leqslant |z| \leqslant c\sqrt{t}$)

$$\mathbf{P}_z\big[T_{t^{-a/2}} < T_{c\sqrt{t}}\big] = \frac{\log(c\sqrt{t}) - \log|z|}{\log(c\sqrt{t}) - \log t^{-a/2}}$$

by Exercise 5.33(5) [or use optional stopping, Section 3.4, on $\log|B|$]. As $t \to \infty$, this goes to $\frac{1}{1+a}$. Use $c = \frac{1}{b}$ and $c = b$ to get the result:

$$\mathbf{P}_z\big[T_{t^{-a/2}} \leqslant t\big] \geqslant \mathbf{P}_z\big[T_{t^{-a/2}} < T_{\sqrt{t}/b}\big] - \mathbf{P}_z\big[t \leqslant T_{\sqrt{t}/b}\big]$$

and

$$\mathbf{P}_z\big[T_{t^{-a/2}} \leqslant t\big] \leqslant \mathbf{P}_z\big[T_{t^{-a/2}} < T_{b\sqrt{t}}\big] + \mathbf{P}_z\big[t \geqslant T_{b\sqrt{t}}\big].$$ ◂

It is also interesting to know how quickly H_t grows.

Lemma 7.21. *For all z, the $\mathbf{P}_z$-law of $\frac{H_t}{(\log\sqrt{t})^2}$ converges to that of $\frac{1}{Z^2}$, where Z is standard normal.*

Le Gall uses this to prove the two preceding results; he also formulates it differently—in particular, see Corollary 2.22.

Proof idea. We have

$$H_t := \inf\Big\{s\,;\ \int_0^s e^{2\,\mathrm{Re}\,\beta_u}\,\mathrm{d}u > t\Big\};$$

this is $K_\bullet = C_\bullet^{-1}$ in the proof of Theorem 7.19. For large s,

$$\log \int_0^s e^{2\,\mathrm{Re}\,\beta_u}\,\mathrm{d}u \approx \max_{0\leqslant u\leqslant s} 2\,\mathrm{Re}\,\beta_u,$$

so

$$H_t \approx \inf\big\{s\,;\ \mathrm{Re}\,\beta_s \geqslant \log\sqrt{t}\big\} \overset{\mathscr{D}}{=} \frac{(\log\sqrt{t})^2}{Z^2}$$

by Corollary 2.22. ◂

Appendix: The Poisson Kernel is Harmonic

We give the calculations for Lemma 7.11, which states that the Poisson kernel in $\mathbb{R}^d$,

$$K(x,y) := \frac{1-|x|^2}{|y-x|^d},$$

is harmonic in x for $x \neq y$ and $|y| = 1$. I took this from https://math.stackexchange.com/q/569481.

Recall the following from calculus, where $u\colon \mathbb{R}^d \to \mathbb{R}$, $\varphi\colon \mathbb{R} \to \mathbb{R}$, and $\mathbf{F}\colon \mathbb{R}^d \to \mathbb{R}^d$:

$$\nabla\big(\varphi(u)\big) = \varphi'(u)\nabla u, \tag{A1}$$

$$\Delta(u) = \operatorname{div} \nabla u, \tag{A2}$$

$$\operatorname{div}(u\,\mathbf{F}) = \nabla u \cdot \mathbf{F} + u \operatorname{div} \mathbf{F}, \tag{A3}$$

$$\Delta(uv) = u\,\Delta v + v\,\Delta u + 2\,\nabla u \cdot \nabla v. \tag{A4}$$

For fixed y, we have $K(x,y) = u(x)v(x)$, where $u(x) := 1 - |x|^2$ and $v(x) := |x-y|^{-d}$. We calculate that

$$\nabla u = -2x, \text{ and so } \Delta u = -2d.$$

Using (A1), we get

$$\begin{aligned}\nabla v &= -d\,|x-y|^{-d-1}\nabla|x-y| \\ &= -d\,|x-y|^{-d-1}\frac{x-y}{|x-y|} \\ &= -d\,|x-y|^{-d-2}(x-y).\end{aligned}$$

Using (A2) and then (A3), we get

$$\begin{aligned}\Delta v &= -d\operatorname{div}\big(|x-y|^{-d-2}(x-y)\big) \\ &= -d(-d-2)|x-y|^{-d-3}\frac{x-y}{|x-y|}\cdot(x-y) - d\,|x-y|^{-d-2}d \\ &= 2d\,|x-y|^{-d-2}.\end{aligned}$$

Finally, combine the results using (A4) and the fact that $|y| = 1$:

$$\begin{aligned}|x-y|^{d+2}\Delta(uv) &= (1-|x|^2)2d - 2d\,|x-y|^2 + 4dx\cdot(x-y) \\ &= 0.\end{aligned}$$

Appendix: Convergence of Harmonic Functions to Boundary Values

The following standard result (see, e.g., Doob's 1984 book, *Classical Potential Theory and Its Probabilistic Counterpart*, Theorem 2.IX.13, p. 651) is not as obvious as it looks:

Theorem A.1. *Let D be a bounded domain in $\mathbb{R}^d$ and g be a bounded Borel function on ∂D. Let B be Brownian motion in $\mathbb{R}^d$ and $T := \inf\{t \geqslant 0 ;\ B_t \notin D\}$. For $x \in \overline{D}$, define*

$$u(x) := \mathbf{E}_x\big[g(B_T)\big].$$

Then for all $x \in D$,

$$\forall x \in D \qquad \lim_{t \uparrow T} v(B_t) = g(B_T) \quad \mathbf{P}_x\text{-a.s.}$$

This is part of Proposition 7.7.

We give two proofs, the first being essentially the standard one; thanks to Michael Damron for conversations on this. The result can easily be extended to unbounded domains by defining $g(B_\infty)$ to be a constant.

Our first proof uses the following standard result (a version of the "predictable stopping theorem"):

Theorem A.2. *Let $(\mathscr{F}_t)_t$ be a filtration. Let T be a* ***predictable stopping time****, i.e., there are stopping times $T_n < T$ that increase to T (such T_n* ***announce*** *T). Suppose that $\mathscr{F}_T = \bigvee_n \mathscr{F}_{T_n}$. Let M be a uniformly integrable right-continuous martingale. Then M is left-continuous at time T.*

An example where the conclusion fails for a nonpredictable stopping time is given by continuous-time simple random walk on $\{0, -1, 1\}$ started at 0 and stopped at the time T of its first jump.

Proof. Because M is right-continuous, the optional-stopping theorem gives $M_{T_n} = \mathbf{E}[M_T \mid \mathscr{F}_{T_n}]$. By the convergence of closed martingales in discrete time, we may deduce that $\lim_{n\to\infty} M_{T_n} = M_T$. For $\epsilon > 0$, let A_ϵ be the event that $\overline{\lim}_{t\uparrow T} |M_t - M_T| > \epsilon$. For $n \geqslant 1$, define the stopping times

$$S_n := T_{n+1} \wedge \inf\big\{t \geqslant T_n ;\ |M_{T_n} - M_t| > \epsilon\big\}.$$

Since $T_n \leqslant S_n < T$, we have $\mathscr{F}_T = \bigvee_n \mathscr{F}_{S_n}$. Since $(S_n)_n$ announce T, it follows that $M_{S_n} \to M_T$ a.s. Thus, $M_{T_n} - M_{S_n} \to 0$ a.s., whence $\mathbf{P}(A_\epsilon) = 0$. Because this holds for every $\epsilon > 0$, we obtain the desired result. ◀

Proof of Theorem A.1. Let $(\mathscr{F}_t)$ be the completed canonical filtration of B. Let $M_t := u(B_{t\wedge T})$. Clearly M is bounded and right-continuous. By the strong Markov property, M is a martingale. The stopping times

$$T_n := \inf\big\{t \geqslant 0 ;\ |B_t - D^c| \leqslant 1/n\big\}$$

announce T. Since $\mathscr{F}_T = \bigvee_n \mathscr{F}_{T_n}$ by continuity of B, we may apply Theorem A.2. ◀

Our second proof uses some auxiliary results.

Theorem A.3. *Let $d \in \mathbb{N}_+$. Let B be d-dimensional Brownian motion. Let $\sigma \in C^2(\mathbb{R}^d)$ with bounded first and second derivatives. Suppose that X solves $E_x(\sigma, 0)$, i.e., $x \in \mathbb{R}^d$ and X is an adapted process with values in $\mathbb{R}^d$ such that*

$$X_t = x + \int_0^t \sigma(X_s)\ \mathrm{d}B_s \quad (t \geqslant 0).$$

If $\sigma(x) \neq 0$, then

$$\mathbf{P}\big[\forall t \geqslant 0\ \ \sigma(X_t) \neq 0\big] = 1.$$

Proof. By Itô's formula,

$$\begin{aligned} \mathrm{d}\sigma^2(X_t) &= 2\sigma(X_t)\sigma'(X_t)\ \mathrm{d}X_t + \big(\sigma'(X_t)^2 + \sigma(X_t)\sigma''(X_t)\big)\ \mathrm{d}\langle X_t, X_t\rangle \\ &= 2\sigma^2(X_t)\sigma'(X_t)\ \mathrm{d}B_t + \big(\sigma'(X_t)^2 + \sigma(X_t)\sigma''(X_t)\big)\sigma^2(X_t) d\ \mathrm{d}t \\ &= \sigma^2(X_t)\ \mathrm{d}M_t \end{aligned}$$

for a continuous semimartingale M with $M_0 = 0$. It follows (say, by the exercise on page 104 near the end of Section 8.1) that

$$\sigma^2(X) = \sigma^2(x)\mathscr{E}(M) = \sigma^2(x)\ \exp\big\{M - \langle M, M\rangle/2\big\}$$

is never 0 if $\sigma(x) \neq 0$. ◂

Theorem A.4. *Let D be a bounded domain in $\mathbb{R}^d$. Let B be Brownian motion in $\mathbb{R}^d$ and $T := \inf\{t \geqslant 0\,;\ B_t \notin D\}$. Let $\sigma \in C^2(\mathbb{R}^d)$ be such that $\sigma(x) > 0$ if $x \in D$ and $\sigma(x) = 0$ if $x \in \partial D$. If $x \in D$ and X solves $E_x(\sigma, 0)$, then $\mathbf{P}[\forall t \geqslant 0\ \ X_t \in D] = 1$ and X is a time change of $(B_t)_{0 \leqslant t < T}$ (in law).*

Proof. The first statement is immediate from Theorem A.3, while the second is proved just as the conformal invariance of Brownian motion (Theorem 7.18) is proved, but simpler (use Proposition 5.15). Note that X is a continuous bounded martingale, so $\int_0^\infty \sigma^2(X_t)\ \mathrm{d}t = \langle X, X\rangle_\infty < \infty$ a.s., whence $X_\infty \in \partial D$ a.s. ◂

Remark. A special case of Theorem A.4 is the following: Let D be the unit disk when $d = 2$ and $\sigma(x) := \big(1 - |x|^2\big)/2$ for $|x| \leqslant 1$. Then X is Brownian motion in the Poincaré model of the hyperbolic plane. The law of X is the same as the law of $\varphi(X)$ for every Möbius transformation φ of the unit disk to itself. For Brownian motion X in the Poincaré model of d-dimensional hyperbolic space, there is a drift term: X solves $E_x(\sigma, b)$ with $\sigma(x) := \big(1 - |x|^2\big)/2$ and $b(x) := (d/2 - 1)\sigma(x)x$ for $|x| \leqslant 1$.

We are now ready to give a second proof of Theorem A.1.

Second proof of Theorem A.1. Let $(\mathscr{F}_t)$ be the completed canonical filtration of B. Construct σ as in the statement of Theorem A.4 by, say, summing a countable collection of small bump functions. Fix $x \in D$. Let X solve $E_x(\sigma, 0)$. Then the path $\big(u(B_t)\big)_{0 \leqslant t < T}$ is the same in law as the path $\big(u(X_t)\big)_{0 \leqslant t < \infty}$ but with a different parametrization. Also, X is a continuous Markov process. Let $X_\infty := \lim_{t\to\infty} X_t$. By the strong Markov property, $u(X_t) = \mathbf{E}\big[g(X_\infty)\ \big|\ \mathscr{F}_t\big]$, whence the result follows from the convergence of closed martingales (in continuous time). ◂

A generalization of Theorem A.3 is in the extra credit exercise on page 116 at the end of Chapter 8. Our proof of Theorem A.3 is modelled on one we heard from Étienne Pardoux for $d = 1$.