Stochastic Calculus
Math 7880-1; Spring 2008

Davar Khoshnevisan

155 South 1400 East JWB 233, Department of Mathematics, University of Utah, Salt Lake City UT 84112–0090

E-mail address: davar@math.utah.edu

URL: http://www.math.utah.edu/~davar

Contents

Chapter 1

Background Material

1. Basic Brownian Motion

Definition 1.1. A *one-dimensional Brownian motion* $B := \{B_t\}_{t\geq 0}$ started at zero is a stochastic process with the following properties:

- $B_0 = 0$;
- $t \mapsto B_t$ is continuous with probability one;
- $B_{t+s} - B_t = N(0\,, s)$ for all $s > 0$ and $t \geq 0$;
- $B_{t+s} - B_t$ is independent of $\{B_r\}_{r\in[0,t]}$ for all $t \geq 0$.

A *one-dimensional Brownian motion* B starting at $x \in \mathbf{R}$ is defined by $B_t := W_t + x$, where W is a one-dimensional Brownian motion started at zero. Note that x could in principle be a random variable independent of $\{B_t\}_{t>0}$.

We recall some basic facts about one-dimensional Brownian motion B:

(1) B exists (Wiener);

(2) $t \mapsto B_t$ is nowhere differentiable a.s.;

(3) $t \mapsto B_t$ is a Gaussian process with mean function $\mathrm{E}B_t := \mathrm{E}B_0$ and $\mathrm{E}(B_s B_t) = \min(s\,, t)$;

(4) $t \mapsto B_t$ is a martingale. That is,

$$\mathrm{E}(B_{t+s}\,|\,\mathscr{F}^B_s) = B_s \quad \text{a.s.},$$

where $\mathscr{F}^B := \{\mathscr{F}^B_s\}_{s\geq 0}$ is the *Brownian filtration*. That is:
  (a) Set $\mathscr{F}^0_t$ to be the σ-algebra generated by B_s for all $s \in [0\,, t]$;
  (b) define $\bar{\mathscr{F}}_t$ to be the P-completion of $\mathscr{F}^0_t$;
  (c) set $\mathscr{F}^B_t := \cap_{s>t}\bar{\mathscr{F}}_s$.

(5) $t \mapsto B_t$ is a strong Markov process. That is, whenever T is a stopping time with respect to $\mathscr{F}^B$, the "post-T process" $\{B_{t+T} - B_t\}_{t\geq 0}$ is a Brownian motion independent of the σ-algebra $\mathscr{F}^B_T$, where

$$\mathscr{F}^B_T := \left\{A \in \mathscr{F}^B_\infty : \ A \cap \{T \leq t\} \in \mathscr{F}^B_t \text{ for all } t \geq 0\right\},$$

and $\mathscr{F}^B_\infty := \vee_{t\geq 0}\mathscr{F}^B_t$ is the underlying σ-algebra in our probability space.

Exercise 1.2 (Kolmogorov zero-one for Brownian motion)**.** For each $t \geq 0$ define $\mathscr{T}_t$ to be the σ-algebra generated by $\{B_u\}_{u\geq t}$. Let $\mathscr{T} := \cap_t \mathscr{T}_t$; this is the socalled *tail σ-algebra* for Brownian motion. Prove that $\mathscr{T}$ is *trivial*. That is, $\mathrm{P}(A) = 0$ or 1 for all $A \in \mathscr{T}$.

Exercise 1.3 (Blumenthal's zero-one for Brownian motion)**.** Prove that $\mathscr{F}_0 := \cap_{t>0}\mathscr{F}_t$ is trivial. That is, $\mathrm{P}(A) = 0$ or 1 for all $A \in \mathscr{F}_0$. (Hint: You can prove this directly. Alternatively, prove first that $t \mapsto tB(1/t)$ defines a Brownian motion also, and then appeal to the Kolmogorov zero-one law.)

2. General Brownian Motion

There are occasions where we need a slightly more general definition of Brownian motion. Let $(\Omega\,, \mathscr{F}_\infty, \mathrm{P})$ denote a complete probability space. Suppose $\mathscr{F} := \{\mathscr{F}_t\}_{t\geq 0}$ is a filtration that satisfies the "usual hypothesis." That is,

(1) Each $\mathscr{F}_t$ is a sub-σ-algebra of $\mathscr{F}_\infty$;

(2) $\mathscr{F}_s \subseteq \mathscr{F}_t$ if $s \leq t$;

(3) $\mathscr{F}_t$ is P-complete for all $t \geq 0$; and

(4) $\mathscr{F}$ is "right continuous." That is, $\mathscr{F}_t = \cap_{s>t}\mathscr{F}_s$ for all $t \geq 0$.

Recall that if $(S\,, \mathscr{S})$ is a measure space, then an S-valued random variable $X : \Omega \to S$ is just a measurable function. And an S-valued *stochastic process* [or sometimes just a *process*] is a family $X := \{X_t\}_{t\geq 0}$ of S-valued random variables, one for every $t \geq 0$. And an S-valued process X is *adapted* [to $\mathscr{F}$] if X_t is $\mathscr{F}_t$-measurable for every $t \geq 0$.

Definition 2.1. Let $B := \{B_t\}_{t\geq 0}$ denote an real-valued stochastic process. Then we say that B is a *Brownian motion*, started at $x \in \mathbf{R}$, with respect to the filtration $\mathscr{F}$ if:

(1) $B_0 = x$;

(2) $B_{t+s} - B_s = N(0\,, t)$ for all $t > 0$ and $s \geq 0$;

(3) B is adapted to $\mathscr{F}$; and

(4) $B_{t+s} - B_s$ is independent of $\mathscr{F}_s$.

Exercise 2.2. Prove that if B is a Brownian motion with respect to some filtration $\mathscr{F}$, then B is also a Brownian motion with respect to its own filtration $\mathscr{F}^B$.

Exercise 2.3. Let B denote a Brownian motion. Suppose $\{\mathscr{G}_t\}_{t\geq 0}$ is a filtration that is independent of $\mathscr{F}^B$. [That is, whenever $A_1 \in \mathscr{G}_t$ and $A_2 \in \mathscr{F}^B_s$ for some $s, t \geq 0$, then $\mathrm{P}(A_1 \cap A_2) = \mathrm{P}(A_1)\mathrm{P}(A_2)$.] Let $\bar{\mathscr{F}}_t$ denote the P-completion of $\mathscr{F}^B_t \vee \mathscr{G}_t$, and define $\mathscr{F}_t := \cap_{s>t} \bar{\mathscr{F}}_s$. Then prove that B is a Brownian motion with respect to $\mathscr{F} := \{\mathscr{F}_t\}_{t\geq 0}$.

Definition 2.4. Let $\boldsymbol{B} := \{\boldsymbol{B}_t\}_{t\geq 0}$ denote an $\mathbf{R}^d$-valued stochastic process. We say that $\boldsymbol{B}$ is a d-dimensional *Brownian motion* with respect to a given filtration $\mathscr{F} := \{\mathscr{F}_t\}_{t\geq 0}$ when the coordinate processes $B^1, \ldots, B^d$ are i.i.d. one-dimensional Brownian motions with respect to $\mathscr{F}$.

Let $\mathscr{F}^{\boldsymbol{B}} := \{\mathscr{F}^{\boldsymbol{B}}_t\}_{t\geq 0}$ denote the natural filtration of d-dimensional Brownian motion $\boldsymbol{B}$. Then the preceding discussion shows that B^1 [say] is a one-dimensional Brownian motion with respect to $\mathscr{F}^{\boldsymbol{B}}$. And more generally, $(B^1, \ldots, B^n)$ is n-dimensional Brownian motion with respect to $\mathscr{F}^{\boldsymbol{B}}$ for every fixed integer $n \in \{1, \ldots, d\}$.

3. More Abstract State Spaces

3.1. Brownian Motion as a Random Variable. Recall that a set S is a *Banach space* if it is a complete normed linear space in norm $\|\cdot\|_S$. Consider d-dimensional Brownian motion $B := \{B_t\}_{t\in[0,1]}$ run for one unit of time, and started at $x \in \mathbf{R}^d$. One can think of B as a random element of S for various choices of S. Here are some natural choices of S. We will need the following technical result.

Lemma 3.1. *If X is an S-valued random variable, then $\|X\|_S$ is an $\mathbf{R}_+$-valued random variable.*

Proof. By a monotone class argument it is enough to prove that $A := \{\|X\| < \lambda\}$ is measurable for all $\lambda > 0$. But A is the inverse image of the open ball in S, about the origin of S, of radius λ. The lemma follows. □

Now let us consider some examples.

Example 3.2. Let $S := C([0\,,1]\,,\mathbf{R}^d)$ denote the collection of all continuous functions mapping $[0\,,1]$ into $\mathbf{R}^d$. Then S is a Banach space with norm $\|f\|_{C([0,1],\mathbf{R}^d)} := \sup_{x\in S} |f(x)|$. Thus, we can think of B as a random variable with values in the Banach space $C([0\,,1]\,,\mathbf{R}^d)$. □

Example 3.3. Let $S := L^p([0\,,1]\,,\mathbf{R}^d)$ denote the collection of all Borel-measurable functions that map $[0\,,1]$ to $\mathbf{R}^d$ with $\|f\|_{L^p([0,1],\mathbf{R}^d)} < \infty$, where $p \in [1\,,\infty)$ and $\|f\|_{L^p([0,1],\mathbf{R}^d)} := (\int_0^1 |f(t)|^p\,\mathrm{d}t)^{1/p}$, as usual. If we identify functions that are equal almost everywhere, then the resulting quotient space $\mathcal{L}^p([0\,,1]\,,\mathbf{R}^d)$ is a Banach space. Because $C([0\,,1]\,,\mathbf{R}^d) \subset \mathcal{L}^p([0\,,1]\,,\mathbf{R}^d)$, it follows then that [the equivalence class of] B is in $\mathcal{L}^p([0\,,1]\,,\mathbf{R}^d)$. Here is a more useful and slightly stronger result: Because $t \mapsto B_t(\omega)$ is continuous for all ω, it follows that $(\omega\,,t) \mapsto B_t(\omega)$ is measurable. Thanks to Fubini's theorem, $\|B\|_{L^p([0\,,1]\,,\mathbf{R}^d)}$ is a real-valued random variable and

$$\mathrm{E}\left(\|B\|^p_{L^p([0,1],\mathbf{R}^d)}\right) = \int_0^1 \mathrm{E}\left(|B_t|^p\right)\,\mathrm{d}t = c_p \int_0^1 t^{p/2}\,\mathrm{d}t < \infty.$$

This proves that $B \in L^p([0\,,1]\,,\mathbf{R}^d)$ a.s., which readily implies that the equivalence class of B is in $\mathcal{L}^p([0\,,1]\,,\mathbf{R}^d)$ a.s. □

Example 3.4. We can also consider d-dimensional Brownian motion $B := \{B_t\}_{t\geq 0}$ indexed by all nonnegative times. Define $C(\mathbf{R}_+\,,\mathbf{R}^d)$ to be the collection of all continuous functions from $\mathbf{R}_+$ to $\mathbf{R}^d$. This is a Banach space with norm,

$$\|f\|_{C(\mathbf{R}_+,\mathbf{R}^d)} := \sum_{k=1}^\infty 2^{-k}\,\frac{\|f\|_{C([0,k],\mathbf{R}^d)}}{1 + \|f\|_{C([0,k],\mathbf{R}^d)}}.$$

Other Banach spaces can be considered as well. □

3.2. Regularity of Processes. Let S be a Banach space, and suppose $X := \{X_t\}_{t\in I}$ is an S-valued process, where I is now some compact upright cube $\prod_{j=1}^n [a_j\,,b_j]$ in $\mathbf{R}^n$. When is X in $C(I\,,S)$? [The Banach space of space of all continuous functions mapping I into S, endowed with the norm $\|f\|_{C(I,S)} := \sup_{t\in I}\|f(t)\|_S$.]

Ordinarily we understand X via its *finite-dimensional distributions.* These are the collection of all probabilities,

$$\mathrm{P}\{X_{t_1} \in E_1\,,\ldots,X_{t_k} \in E_k\},$$

as k varies over positive integers, $t_1,\ldots,t_k$ vary over I, and $E_1,\ldots,E_k$ vary over Borel subsets of S. Suppose X is *a priori* in $C(I\,,S)$. Define U to be an independent random variable, chosen uniformly at random from I, and set $Y_t := X_t\mathbf{1}_{\{U\neq t\}} + (1 + X_t)\mathbf{1}_{\{U=t\}}$. Because X is continuous, Y is not. But Y has the same finite-dimensional distributions as X. This suggests that we cannot hope, in general, to prove that X is in $C(I\,,S)$ using only the finite-dimensional distributions of X. Our strategy will be to prove that in many cases X has a continuous "modification."

Definition 3.5. An S-valued process $\{Y_t\}_{t\in I}$ is a *modification* of X if $\mathrm{P}\{X_t = Y_t\} = 1$ for all $t \in I$.

Note that whenever Y is a modification of X, it has the same finite-dimensional distributions as X.

Theorem 3.6 (Kolmogorov Continuity Theorem). *Suppose $X := \{X_t\}_{t\in I}$ is an S-valued process, as above, where I is a fixed compact upright cube in $\mathbf{R}^n$. Suppose there exist constants $c, \epsilon > 0$ and $p \geq 1$ such that*

$$\mathrm{E}\left(\|X_t - X_s\|_S^p\right) \leq c|t-s|^{n+\epsilon} \quad \textit{for all } s,t \in I. \tag{3.1}$$

Then X has a modification Y that is continuous. In fact,

$$\mathrm{E}\left(\sup_{\substack{s,t\in I:\\ s\neq t}} \frac{\|Y_t - Y_s\|_S}{|t-s|^q}\right) < \infty \quad \textit{provided that } 0 \leq q < \frac{\epsilon}{p}.$$

Proof. Because the form of (3.1) does not change after we apply an affine transformation to I, we can assume without loss of generality that $I = [0\,,1]^n$.

Also, because all Euclidean norms are equivalent, we may assume without loss of generality that $|t| := \sum_{j=1}^n |t^j|$ denotes the ℓ^1-norm on $\mathbf{R}^n$.

For every integer $m \geq 1$, let $\mathcal{D}_m$ denote the collection all points $t \in I$ whose coordinates each have the form $j2^{-m}$ for some integer $j \in \{0\,,\ldots,2^m\}$. Clearly, $|\mathcal{D}_m| = (2^m+1)^n$, where $|\cdots|$ denotes cardinality. The collection $\mathcal{D} := \cup_{m=1}^\infty \mathcal{D}_m$ is the set of all "dyadic rationals" in I, and is manifestly countable and also dense [in I].

If $s,t \in \mathcal{D}$ and $|s-t| \leq 2^{-m}$, then we can find sequences $t_m, t_{m+1},\ldots$ and $s_m, s_{m+1},\ldots$ such that: (a) $t_j, s_j \in \mathcal{D}_j$ for all $j \geq m$; (b) $|t_m - s_m| = 2^{-m}$; and (c) $|t_{j+1} - t_j| = |s_{j+1} - s_j| = 2^{-j}$ for all $j \geq m$; and $\lim_{n\to\infty} t_n = t$ as well as $\lim_{n\to\infty} s_n = s$. By the triangle inequality,

$$\|X_t - X_s\|_S \leq 2\sum_{j=m}^\infty \max_{\substack{u,v\in\mathcal{D}_j:\\ |u-v|=2^{-j}}} \|X_u - X_v\|_S\,.$$

Therefore,

$$\mathrm{E}\left(\sup_{\substack{s,t\in\mathcal{D}:\\ |s-t|\leq 2^{-\ell}}} \|X_t - X_s\|_S\right) \leq 2\sum_{j=m}^\infty \mathrm{E}\left(\max_{\substack{u,v\in\mathcal{D}_j:\\ |u-v|=2^{-j}}} \|X_u - X_v\|_S\right).$$

There are $(2^j+1)^n$-many points in $\mathcal{D}_j$, and for each such point u, there are at most $2n$ points $v \in \mathcal{D}_j$ such that $|u-v| = 2^{-j}$. Therefore,

$$\begin{aligned}
\mathrm{E}\left(\max_{\substack{u,v\in\mathcal{D}_j:\\|u-v|=2^{-j}}}\|X_u - X_v\|_S^p\right) &\le \mathop{\sum\sum}_{\substack{u,v\in\mathcal{D}_j:\\|u-v|=2^{-j}}} \mathrm{E}\left(\|X_u-X_v\|_S^p\right)\\
&\le (2n)(2^j+1)^n c\left(2^{-j}\right)^{n+\epsilon}\\
&\le K_{n,c}2^{-j\epsilon}.
\end{aligned}$$

This and Jensen's inequality together implies that

$$\begin{aligned}
\mathrm{E}\left(\sup_{\substack{s,t\in\mathcal{D}_m:\\|s-t|\le 2^{-\ell}}}\|X_t - X_s\|_S\right) &\le K_{c,n}^{1/p}\sum_{j=m}^\infty 2^{-j\epsilon/p} \le K_{c,n,p,\epsilon}2^{-m\epsilon/p}\\
&\le K'_{c,n,p,\epsilon}2^{-\ell\epsilon/p}.
\end{aligned}$$

By the monotone convergence theorem,

$$\mathrm{E}\left(\sup_{\substack{s,t\in\mathcal{D}:\\|s-t|\le 2^{-\ell}}}\|X_t - X_s\|_S\right) \le K'_{c,n,p,\epsilon}2^{-\ell\epsilon/p}.$$

To finish we note that

$$\begin{aligned}
\mathrm{E}\left(\sup_{\substack{s,t\in\mathcal{D}:\\s\neq t}}\frac{\|X_t-X_s\|_S}{|t-s|^q}\right) &\le \sum_{\ell=0}^\infty 2^{\ell q}\mathrm{E}\left(\sup_{\substack{s,t\in\mathcal{D}:\\2^{-\ell}\le|s-t|\le 2^{-\ell-1}}}\|X_t-X_s\|_S\right)\\
&\le K'_{c,n,p,\epsilon}\sum_{\ell=0}^\infty 2^{\ell q-(\ell+1)\epsilon/p},
\end{aligned}$$

and this is finite if $q\in[0\,,\epsilon/p)$. To finish, we set $Y_t := \lim_{s\to t:\,s\in\mathcal{D}} X_s$. □

Example 3.7. Let B denote d-dimensional Brownian motion. We know that $\mathrm{E}(\|B_t - B_s\|^2) = d|t-s|$. General facts about Gaussian processes tell us that for all $p\ge 2$ there exists a constant c_p such that $\mathrm{E}(\|B_t-B_s\|^p) = c_p|t-s|^{p/2}$. Therefore, we can apply the Kolmogorov continuity theorem with $n=1$, $I:=[a\,,b]$, and $p\ge 2$ arbitrarily large, and deduce that

$$\mathrm{E}\left(\sup_{\substack{s,t\in[a,b]:\\s\neq t}}\frac{\|B_t-B_s\|}{|t-s|^q}\right) < \infty \quad \text{for all } q\in[0\,,1/2).$$

That is, there exists a random variable $V_{a,b} \in L^1(\mathrm{P})$ [in particular, $V_{a,b} < \infty$ a.s.] such that if $q \in [0\,,1/2)$ is held fixed, then outside a single null set,

$$\|B_t - B_s\| \le V_{a,b}|s-t|^q \quad \text{for all } s,t \in [a\,,b].$$

In the language of classical analysis, Brownian motion is a.s. Hölder continuous of any order $q < 1/2$. The following exercise shows that B is not Hölder continuous of order $q = 1/2$. □

Exercise 3.8. Show that for all $\lambda > 0$,

$$\mathrm{P}\left\{\max_{0\le j\le 2^n-1}\left\|B_{(j+1)/2^n} - B_{j/2^n}\right\| < \lambda 2^{-n/2}\right\} = \left(1 - \mathrm{P}\{\|Z\| > \lambda\}\right)^{2^n-1},$$

where Z is a d-vector of i.i.d. standard normal random variables. Conclude from this that with probability one, $\sup\|B_t - B_s\|/|t-s|^{1/2} = \infty$, where the supremum is taken over all unequal $s,t \in [a\,,b]$.

Example 3.9. Let B denote d-dimensional Brownian motion, and H_q the collection of all functions $f : \mathbf{R}_+ \to \mathbf{R}^d$ with $\|f\|_{H_q(\mathbf{R}_+,\mathbf{R}^d)} < \infty$, where

$$\|f\|_{H_q(\mathbf{R}_+,\mathbf{R}^d)} := \sum_{k=1}^\infty 2^{-k}\left(\frac{\|f\|_{H_q([0,k],\mathbf{R}^d)}}{1 + \|f\|_{H_q([0,k],\mathbf{R}^d)}}\right).$$

Here $q > 0$, and

$$\|f\|_{H_q([0,k],\mathbf{R}^d)} := \sup_{\substack{s,t\in[0,k]:\\ s\neq t}} \frac{\|f(t) - f(s)\|}{|t-s|^q}.$$

It is easy to see that $H_q(\mathbf{R}_+\,,\mathbf{R}^d)$ is a Banach space, and the preceding shows that $B \in H_q(\mathbf{R}_+\,,\mathbf{R}^d)$ a.s., as long as $q \in (0\,,1/2)$. The preceeding shows that with probability one, B is not an element of $H_{1/2}(\mathbf{R}_+\,,\mathbf{R}^d)$. □

4. Continuous Martingales

Suppose $\mathscr{F}$ is a filtration that satisfies the usual hypotheses. Recall that a real-valued process $M := \{M_t\}_{t\ge0}$ is a *martingale* if:

(1) M is adapted [to $\mathscr{F}$];
(2) $M_t \in L^1(\mathrm{P})$ for all $t \ge 0$; and
(3) $\mathrm{P}\{\mathrm{E}(M_{t+s}\,|\mathscr{F}_t) = M_s\} = 1$ for all $s,t \ge 0$.

Recall that M is [left- or right-] *continuous* if $t \mapsto M_t(\omega)$ is [left- or right-] continuous for every (sometimes almost every) $\omega \in \Omega$. Furthermore, M is *p-times integrable* [*square integrable* if $p = 2$] if $\mathrm{E}(|M_t|^p) < \infty$ for all $t \ge 0$.

Recall that an $[0\,,\infty]$-valued random variable T is a *stopping time* if

$$\{T \le t\} \in \mathscr{F}_t \quad \text{for all } t \ge 0.$$

[In the general theory of processes, stopping times are sometimes also called "optional times."]

If T is a stopping time, then we define

$$\mathscr{F}_T := \{A \in \mathscr{F}_\infty : \ A \cap \{T \le t\} \in \mathscr{F}_t \text{ for every } t \ge 0\}.$$

Lemma 4.1. *T is $\mathscr{F}_T$-measurable. Moreover, if X is a left- or right-continuous adapted process, then X_T is $\mathscr{F}_T$-measurable.*

Finally, let us recall three fundamental theorems of martingale theory.

Theorem 4.2 (Doob's optional stopping theorem)**.** *If M is a [left- or right-] continuous martingale, then for all finite stopping times T, $M^T := \{M_{t\wedge T}\}_{t\ge 0}$ is a [left- or right-] continuous martingale. In particular,*

$$\mathrm{E}(M_{t\wedge T}) = \mathrm{E}(M_0) \qquad \textit{for all } t \ge 0.$$

Theorem 4.3 (Doob's weak maximal inequality)**.** *Suppose M is a [left- or right-] continuous martingale. Consider the event A that $\sup_{s\in[0,t]} |M_s| \ge \lambda$, where $t \ge 0$ and $\lambda > 0$ are fixed and nonrandom. Then $\lambda\mathrm{P}(A) \le \mathrm{E}(|M_t|; A)$.*

Theorem 4.4 (Doob's strong maximal inequality)**.** *Suppose M is a [left- or right-] continuous martingale. Then for all $t \ge 0$ and $p \in [1\,,\infty)$,*

$$\mathrm{E}\left(\sup_{s\in[0,t]} |M_s|^p\right) \le q^p \mathrm{E}\left(|M_t|^p\right),$$

where $p^{-1} + q^{-1} = 1$.

All three are proved by first using the fact that $\{M_t\}_{t\in F}$ is a discrete-parameter martingale whenever F is a finite subset of $\mathbf{R}_+$. And then using continuity and a sequence of F's that exhaust $\mathbf{R}_+$. The following is another result from martingale theory. I will prove it, since its proof is slightly different from the argument described just now.

Theorem 4.5 (Doob's martingale convergence theorem)**.** *If M is a [left- or right-] continuous martingale that is bounded in $L^1(\mathrm{P})$, then $M_\infty := \lim_{t\to\infty} M_t$ exists almost surely, and $\mathrm{E}(|M_\infty|) < \infty$.*

Proof. Thanks to Fatou's lemma, it suffices to prove that M_∞ exists a.s. Because $\{M_t\}_{t\in\mathbf{Z}_+}$ is a discrete-parameter martingale that is bounded in $L^1(\mathrm{P})$, the discrete-parameter martingale convergence theorem implies that $M^n_\infty := \lim_{t\to\infty:\, t\in\mathbf{Z}_+} M_t$ exists a.s. For every fixed $s \ge 0$, the process

$t \mapsto M_{t+s} - M_s$ is a continuous martingale that is bounded in $L^1(\mathrm{P})$. [The filtration is the one generated by the process, for instance.] Therefore, Doob's maximal inequality (Theorem 4.3) implies that for all $\lambda, T, s > 0$,

$$\mathrm{P}\left\{\sup_{t\in[0,T]} |M_{t+s} - M_s| > \frac{\lambda}{2}\right\} \le \frac{2}{\lambda}\mathrm{E}\left(|M_{T+s} - M_s|\right) \le \frac{4C}{\lambda},$$

where $C := \sup_u \mathrm{E}(|M_u|)$. We can pick $s \in \mathbf{Z}_+$ so large that $\mathrm{P}\{|M_s - M_\infty| \ge \lambda/2\} \le \lambda^{-1}$. This and the monotone convergence theorem together imply that

$$\mathrm{P}\left\{\sup_{v\ge s} |M_v - M_\infty| \ge \lambda\right\} \le \frac{4C+1}{\lambda}. \tag{4.1}$$

Let $s \to \infty$ in $\mathbf{Z}_+$, and then $\lambda \to \infty$—in this order—to conclude that $\lim_{v\to\infty} M_v = M_\infty$ almost surely. $\square$

Exercise 4.6. Suppose M is a [left- or right-] continuous martingale that is bounded in $L^p(\mathrm{P})$ for some $p \in (1\,,\infty)$. Prove that $M_\infty := \lim_{t\to\infty} M_t$ exists a.s. and in $L^p(\mathrm{P})$. Prove that in addition, $\mathrm{P}\{M_t = \mathrm{E}(M_\infty \,|\, \mathscr{F}_t)\} = 1$ for all $t \ge 0$. (Hint: For the last part use the conditional Jensen's inequality: That is the fact that conditional expectations are contractions on $L^p(\mathrm{P})$. That is, $\|\mathrm{E}(Z \,|\, \mathscr{G})\|_{L^p(\mathrm{P})} \le \|Z\|_{L^p(\mathrm{P})}$.)

5. Itô Integrals

Now we wish to construct integrals of the form $\int G\,\mathrm{d}B$, where $\{G_t\}_{t\ge0}$ is a "nice" stochastic process, and B is a one-dimensional Brownian motion with respect to a filtration $\mathscr{F}$ that satisfies the usual hypotheses.

Definition 5.1. A process $G := \{G_t\}_{t\ge0}$ is said to be *progressively measurable*, if the map $\Omega \times [0\,,t] \ni (\omega\,,s) \mapsto G_s(\omega)$ is $\mathscr{F}_t \times \mathscr{B}([0\,,t])$-measurable for every $t \ge 0$. Here and throughout, "$\mathscr{B}(A)$" denotes the Borel subsets of A.

For the most part, we consider the case that G is real-valued. But we note in passing that the G_t's could be S-valued for any measure space $(S\,,\mathscr{S})$.

Progressively measurable processes are always adapted, but the converse is false. [Such strange cases usually appear as counterexamples in a standard discussion of the Fubini–Tonelli theorem.]

The following three exercises motivate why we define processes in this way. We will use them without further mention. So it is important to learn their content well.

Exercise 5.2. Suppose G is progressively measurable and T is a finite stopping time. Prove that $\{G_{t\wedge T}\}_{t\ge 0}$ is then progressively measurable. In particular, show that G_T is an $\mathscr{F}_T$-measurable random variable.

Exercise 5.3. Suppose $\phi : \mathbf{R}_+ \times \mathbf{R} \to \mathbf{R}_+$ is Borel measurable. Then prove that whenever G is real-valued and progressively meaurable, then $\int_0^\infty \phi(s\,, G_s)\,\mathrm{d}s$ is a random variable with values in $[0\,,\infty]$.

Exercise 5.4. Prove that if G is adapted and left-continuous [or right-continuous], then G is progressively measurable. We are tacitly assuming that G is S-valued for a topological space S.

Our immediate goal is to construct—in four steps—a useful theory of stochastic integration that allows us to define the stochastic integral-process $t \mapsto \int_0^t G\,\mathrm{d}B$ for all progressively measurable real-valued processes G that satisfy the following condition:

$$\mathrm{E}\left(\int_0^t |G_s|^2\,\mathrm{d}s\right) < \infty \quad \text{for all } t \ge 0. \tag{5.1}$$

Step 1. We say that G is *simple* if $G(\omega\,, s) = X(\omega)\mathbf{1}_{[a,b)}(s)$ for $0 \le a < b < \infty$ and $s \ge 0$, and a random variable $X \in L^2(\mathrm{P})$ that is $\mathscr{F}_a$-measurable. Clearly, such G's are progressively measurable and satisfy (5.1). We define

$$\int_0^t G\,\mathrm{d}B := X \cdot (B_{b\wedge t} - B_{a\wedge t}) \qquad \text{for all } t \ge 0.$$

You should check directly that $t \mapsto \int_0^t G\,\mathrm{d}B$ is a continuous square-integrable martingale that has mean zero and variance $(b\wedge t - a\wedge t)\mathrm{E}(X^2)$. We summarize this as follows: For all $t \ge 0$,

$$\begin{aligned} &\mathrm{E}\left(\int_0^t G\,\mathrm{d}B\right) = 0 \quad \text{and} \\ &\mathrm{E}\left(\left|\int_0^t G\,\mathrm{d}B\right|^2\right) = \mathrm{E}\left(\int_0^t |G_s|^2\,\mathrm{d}s\right). \end{aligned} \tag{5.2}$$

Finite [where the number is nonrandom] linear combinations of simple processes are called *elementary*, and the collection of all such processes is denoted by $\mathscr{E}$. When $G \in \mathscr{E}$ we can write $G = G^1 + \cdots + G^k$ when the G^j's are simple and have disjoint support. Evidently, elements of $\mathscr{E}$ are progressively measurable. Moreover, $G \in \mathscr{E}$ satisfies (5.1) if and only if each G^i does. Thus, in the case that G satisfies (5.1) we can define $\int_0^t G\,\mathrm{d}B := \sum_{j=1}^k \int_0^t G^j\,\mathrm{d}B$.

It can be shown that this is a well-defined continuous square-integrable martingale; for well-definedness one proceeds as one does for its counterpart for Lebesgue integrals. Also, $\mathrm{E}\int_0^t G\,\mathrm{d}B = 0$, and since the G^j's have disjoint supports, we can conclude that the variance formula (5.2) holds also when G is elementary. Let us emphasize also that the process $t \mapsto \int_0^t G\,\mathrm{d}B$ is progressively measurable (why?).

Step 2. Now we construct the stochastic integral in the case that the integrand G is bounded, continuous, and adapted [in particular, progressively measurable as well].

Define

$$G_t^n := \sum_{0\le j\le t2^n-1} G_{j/2^n}\mathbf{1}_{[j2^{-n},(j+1)2^{-n})}(t) \qquad \text{for all } t\ge 0.$$

Then, $\sup_{s\in[0,t]}|G_s^n - G_t| \to 0$ a.s. as $n\to\infty$. By the dominated convergence theorem,

$$\lim_{n\to\infty}\mathrm{E}\left(\int_0^t |G_s^n - G_s|^2\,\mathrm{d}s\right) = 0. \tag{5.3}$$

Because $G^n \in \mathscr{E}$, each $\int_0^t G^n\,\mathrm{d}B$ defines a continuous martingale that satisfies the "Itô isometry" (5.2). That isometry and (5.3) together show that

$$\lim_{m,n\to\infty}\mathrm{E}\left(\left|\int_0^t G^m\,\mathrm{d}B - \int_0^t G^n\,\mathrm{d}B\right|^2\right) = 0.$$

Because $\int_0^t (G^m - G^n)\,\mathrm{d}B$ is a continuous square-integrable martingale, Doob's maximal inequality (p. 8) asserts that

$$\mathrm{E}\left(\sup_{0\le u\le t}\left|\int_0^u G^m\,\mathrm{d}B - \int_0^u G^n\,\mathrm{d}B\right|^2\right) \le 4\mathrm{E}\left(\left|\int_0^t G^m\,\mathrm{d}B - \int_0^t G^n\,\mathrm{d}B\right|^2\right),$$

and this goes to zero as $m,n\to\infty$. We can use a subsequence to conclude that $I_t := \lim_{n\to\infty}\int_0^t G^n\,\mathrm{d}B$ exists a.s., and is a continuous adapted process. Because conditional expectations are contractions on $L^2(\mathrm{P})$, $\{I_t\}_{t\ge 0}$ is also a mean-zero martingale, and $\mathrm{E}(I_t^2) = \mathrm{E}(\int_0^t |G_s|^2\,\mathrm{d}s)$. Now define $\int_0^t G\,\mathrm{d}B := I_t$ to construct $\int_0^t G\,\mathrm{d}B$ as a continuous adapted martingale that has mean zero and satisfies the Itô isometry (5.2).

Step 3. Next suppose G is a bounded and progressively-measurable process.

For every integer $n\ge 1$, $G_t^n := n\int_{t-(1/n)}^t G_s\,\mathrm{d}s$ defines an adapted, bounded, and continuous process. By the Lebesgue differentiation theorem, $\lim_{n\to\infty} G_t^n = G_t$ for all but a Lebesgue-null set of t's. Boundedness implies that these G^n's satisfy (5.3) [dominated convergence theorem.] Another application of the Itô isometry [via Cauchy sequences] constructs $\int_0^t G\,\mathrm{d}B$ as a continuous martingale with mean zero that satisfies the Itô isometry.

Step 4. If G is a nonnegative progressively-measurable process that satisfies (5.1), then $G_s^n := G_s \wedge n$ defines a bounded and progressively-measurable process. Condition (5.1) and the dominated convergence theorem together imply that G^n satisfies (5.3). And hence we obtain the stochastic integral $\int_0^t G\,\mathrm{d}B$ as a continuous martingale with mean zero that satisfies the Itô isometry. Finally, if G is progressively measurable and satisfies (5.1), then $G_s^+ := G_s \vee 0$ and $G_s^- := -(0 \vee -G_s)$ are nonnegative progressively measurable process, each of which satisfies (5.1). Define $\int_0^t G\,\mathrm{d}B := \int_0^t G^+\,\mathrm{d}B - \int_0^t G^-\,\mathrm{d}B$; this defines a continuous martingale with mean zero that satisfies the Itô isometry. □

Let us end this discussion with a simple though important property of stochastic integrals. Suppose T is a stopping time. Then, whenever G is progressively measurable and satisfies (5.1),

$$\int_0^{T\wedge t} G\,\mathrm{d}B = \int_0^t G\mathbf{1}_{[0,T)}\,\mathrm{d}B \qquad \text{for all } t \geq 0, \tag{5.4}$$

almost surely. Indeed, we may note first that $\mathbf{1}_{[0,T)}$ is progressively measurable and satisfies (5.1). Therefore, all items in the previous formula are well defined.

In order to prove (5.4), it suffices to consider only the case that G is a simple process. That is, $G_t(\omega) = X(\omega)\mathbf{1}_{[a,b)}(t)$ where $X \in L^2(\mathrm{P})$ is $\mathscr{F}_a$-measurable. But in that case (5.4) holds vacuously.

6. Localization

It turns out that (5.1) is not the correct integrability condition that guarantess that a progressively-measurable process is be a good integrator for B. This idea leads us to "localization."

Definition 6.1. An adapted process $M := \{M_t\}_{t\geq 0}$ is a *local martingale* if we can find a sequence of stopping times $T_1 \leq T_2 \leq \ldots$ with following two properties:

(1) $\lim_{k\to\infty} T_k = \infty$ a.s.; and
(2) $t \mapsto M_{t\wedge T_k}$ defines a martingale for all $k \geq 1$.

Any such sequence $\{T_k\}_{k\geq 1}$ is called a *localizing sequence.*

Every [right-] continuous local martingale is manifestly a [right-] continuous martingale. In fact, if M is a [right-] continuous martingale then so is $\{M_{t\wedge T}\}_{t\geq 0}$ for *every* finite stopping time [the optional stopping theorem]. The converse need not be true because M_t need not be integrable for all

$t \geq 0$. Later on we will see natural examples of local martingales that are not martingales; see, for example, Exercise 2.9 on page 23.

Exercise 6.2. Prove that: (a) On one hand there is no unique localizing sequence; and yet (b) on the other hand, the following defines a canonical choice: $T_k := \inf\{s > 0 : \ |M_s| > k\}$, where $\inf \varnothing := \infty$, as usual.

Exercise 6.3. Prove that if M is a local martingale and $\sup_t \mathrm{E}(|M_t|^p) < \infty$ for some $p \in (1\,,\infty)$, then M is a martingale. In particular, bounded local martingales are martingales.

Now suppose that G is progressively measurable and satisfies the following [much] weaker version of (5.1):

$$\mathrm{P}\left\{\int_0^t |G_s|^2\,\mathrm{d}s < \infty\right\} = 1 \quad \text{for all } t \geq 0. \tag{6.1}$$

Define

$$T_k := \inf\left\{s > 0 : \ \left|\int_0^s G_r\,\mathrm{d}r\right| > k\right\} \qquad \text{for all integers } k \geq 1,$$

where $\inf \varnothing := \infty$. Evidently, $T_1 \leq T_2 \leq \cdots$ and $\lim_{k\to\infty} T_k = \infty$. If $t \mapsto G_t$ is [a.s.] right-continuous and adapted, then the T_k's are easily seen to be stopping times. In fact, these T_k's are *always* stopping times [assuming only progressive measurability]. But this is a *deep* theorem of G. Hunt that is too hard to prove here, and I will not reproduce the proof here. [For our purposes, you can think of G as being right-continuous adapted, in which case you can prove directly that T_k is a stopping time for each k.]

Because $t \mapsto G^k_t := G_{t\wedge T_k}$ is progressively measurable, the Itô integral process $t \mapsto \int_0^t G^k\,\mathrm{d}B$ is well defined as a continuous martingale. I claim that the following holds: With probability one,

$$\int_0^t G^k\,\mathrm{d}B = \int_0^t G^{k+1}\,\mathrm{d}B \qquad \text{whenever } 0 \leq t \leq T^k. \tag{6.2}$$

But this is not hard because: (a) $t \mapsto \mathbf{1}_{[0,T^k)}(t)$ is right-continuous and adapted (and hence progressively measurable); and (b) with probability one,

$$\int_0^{t\wedge T_k} (G^k - G^{k+1})\,\mathrm{d}B = \int_0^t (G^k - G^{k+1})\mathbf{1}_{[0,T^k)}\,\mathrm{d}B,$$

which is zero almost surely. [This last identity is a consequence of (5.4).] Equation (6.2) allows us to define, unambiguously, $\int_0^t G\,\mathrm{d}B$ as $\int_0^t G^k\,\mathrm{d}B$ on $\{t < T^k\}$. Because $T_k \to \infty$ a.s., this defines the entire process $t \mapsto \int_0^t G\,\mathrm{d}B$ as a continuous adapted process. Because $\int_0^{t\wedge T_k} G\,\mathrm{d}B = \int_0^t G^k\,\mathrm{d}B$, $t \mapsto \int_0^t G\,\mathrm{d}B$ is a local martingale (why?).

We summarize our findings.

Theorem 6.4. *Suppose G is progressively measurable. If G satisfies* (5.1), *then $t \mapsto \int_0^t G\,\mathrm{d}B$ is a continuous square-integrable martingale with mean zero, and satisfies the Itô isometry* (5.2). *If G satisfies* (6.1), *then $t \mapsto \int_0^t G\,\mathrm{d}B$ is a continuous local martingale such that whenever $\{\tau_k\}_{k\geq 1}$ is a localizing sequence for it,*

$$\mathrm{E}\left(\int_0^{t\wedge\tau_k} G\,\mathrm{d}B\right) = 0 \qquad \textit{for all } t \geq 0, \textit{ and}$$

$$\mathrm{E}\left(\left|\int_0^{t\wedge\tau_k} G\,\mathrm{d}B\right|^2\right) = \mathrm{E}\left(\int_0^{t\wedge\tau_k} |G(s)|^2\,\mathrm{d}s\right) \qquad \textit{for all } t \geq 0.$$

If F and G are progressively measurable and satisfy (6.1), *then for all $t \geq 0$ and $\alpha, \beta \in \mathbf{R}$,*

$$\int_0^t (\alpha F + \beta G)\,\mathrm{d}B = \alpha \int_0^t F\,\mathrm{d}B + \beta \int_0^t G\,\mathrm{d}B \qquad \textit{a.s.}$$

Convention: It is frequently more convenient to write $\int_0^t G_s\,\mathrm{d}B_s$ in place of $\int_0^t G\,\mathrm{d}B$, and $\int_r^t G_s\,\mathrm{d}B_s$ in place of $\int_0^t G_s \mathbf{1}_{[r,t)}(s)\,\mathrm{d}B_s$. The meanings are synonymous. In particular, $\int_r^t G\,\mathrm{d}B = \int_0^t G\,\mathrm{d}B - \int_0^r G\,\mathrm{d}B$. □

7. The Case of Several Dimensions

Now let us consider the more general case that $B := \{B_t\}_{t\geq 0}$ denotes a d-dimensional Brownian motion with respect to a filtration $\mathscr{F}$ that satisfies the usual hypotheses.

Let $G := \{G_t\}_{t\geq 0}$ be a matrix-valued progressively-measurable process with respect to $\mathscr{F}$. To be concrete, we suppose that $G_t(\omega) \in (\mathbf{R}^n)^d$; i.e., G takes values in the space of $(n \times d)$-dimensional matrices with real entries. Then we define $\int_0^t G\,\mathrm{d}B$—ω by ω—as follows:

$$\left(\int_0^t G\,\mathrm{d}B\right)_i (\omega) := \left(\sum_{k=1}^d \int_0^t G^{i,k}\mathrm{d}B^k\right)(\omega) \quad \text{for all } i = 1, \ldots, n.$$

Itô's Formula

1. The Basic Formula

Let B denote d-dimensional Brownian motion with respect to a filtration $\mathscr{F}$ that satisfies the usual hypotheses.

Theorem 1.1 (Itô)**.** *Choose and fix a bounded domain $D \subset \mathbf{R}^d$. Suppose B_0 is in the interior of D, $x \mapsto f(t\,,x) \in C^2(D)$, and $t \mapsto f(t\,,x)$ is C^1. Then outside a single null set the following holds:*

$$\begin{aligned} &f(t\,,B_t) - f(0\,,B_0) \\ &= \int_0^t (\nabla f)(s\,,B_s)\,\mathrm{d}B_s + \int_0^t (Lf)(s\,,B_s)\,\mathrm{d}s \quad (0 < t < T_D), \end{aligned} \tag{1.1}$$

where $\nabla := (\partial f/\partial x_1\,,\ldots,\partial f/\partial x_d)$ denotes the usual gradient operator acting on the spatial variable x, $(Lf)(s\,,x) := (\partial f/\partial t)(s\,,x) + \frac{1}{2}(\Delta f)(s\,,x)$, where $(\Delta f)(s\,,x) := \sum_{j=1}^d (\partial^2 f/\partial x_j^2)(s\,,x)$ denotes the [spatial] Laplacian, and $T_D := \inf\{s > 0 :\, B_s \notin D\}$.

Remark 1.2. Note that

$$\int_0^t (\nabla f)(s\,,B_s)\,\mathrm{d}B_s = \sum_{j=1}^d \int_0^t \frac{\partial f}{\partial x_j}(s\,,B_s)\,\mathrm{d}B_s^j.$$

Proof. I will prove the result in the case that $d = 1$ and f does not depend on time. The general case is derived by similar means, but uses a multidimensional Taylor expansion in place of the univariate one. Our goal is to prove that for all $t < T_D$,

$$(1.2) \qquad f(B_t) - f(B_0) = \int_0^t f'(B_s)\,\mathrm{d}B_s + \frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s.$$

We will tacitly use the fact that $T_D < \infty$ a.s. (why?).

First we consider the case that f is bounded and has at least 3 bounded derivatives on all of $\mathbf{R}$.

Choose and fix an integer $n \geq 1$ and a number $t \geq 0$, and define $s_j^n := tj/2^n$. We write $f(t\,,B_t)$ as a telescoping sum and rearrange as follows:

$$\begin{aligned}
&f(B_t) - f(B_0)\\
&= \sum_{0\leq j\leq 2^n-1}\left[f\left(B_{s_{j+1}^n}\right) - f\left(B_{s_j^n}\right)\right]\\
&= \sum_{0\leq j\leq 2^n-1}(\Delta B)_{j,n} f'(B_{s_j^n}) + \sum_{0\leq j\leq 2^n-1}\frac{(\Delta B)_{j,n}^2}{2} f''(B_{s_j^n}) + \sum_{0\leq j\leq 2^n-1} R_{j,n},
\end{aligned}$$

where

$$(\Delta B)_{j,n} := B_{s_{j+1}^n} - B_{s_j^n},$$

and $|R_{j,n}| \leq \frac{1}{6}\|f'''\|_\infty \cdot |(\Delta B)_{j,n}|^3$. Because $(\Delta B)_{j,n} = N(0, 2^{-n})$,

$$(1.3) \qquad \mathrm{E}\left(\sum_{0\leq j\leq 2^n-1}|R_{j,n}|\right) \leq \mathrm{const}\cdot 2^{-n/2}.$$

Thus, the third term in the expansion of $f(B_t) - f(B_0)$ goes to zero in $L^1(\mathrm{P})$—and hence in probability—as $n \to \infty$. The first term converges to $\int_0^t f'(B_s)\,\mathrm{d}B_s$ a.s.— hence in probability—as $n \to \infty$ [the construction of the Itô integral]. We claim that

$$(1.4) \qquad \sum_{0\leq j\leq 2^n-1}(\Delta B)_{j,n}^2\, f''(B_{s_j^n}) \to \int_0^t f''(B_s)\,\mathrm{d}s \quad \text{in probability.}$$

This proves that in the case that f and its first three derivatives are bounded and continuous, (1.1) holds a.s. for each $t \geq 0$. Because all quantities in (1.1) are continuous functions of t, it follows that (1.1) holds for all $t \geq 0$ outside a single null set (why?).

To prove (1.4) let Q_n denote the quantity on the left-hand side,and Q the one on the right-hand side. We can write

$$\begin{aligned} Q_n - Q &= T_1 + T_2, \quad \text{where} \\ T_1 &:= \sum_{0\le j\le 2^n-1} f''(B_{s_j^n}) \left[(\Delta B)_{j,n}^2 - 2^{-n}\right], \quad \text{and} \\ T_2 &:= 2^{-n} \sum_{0\le j\le 2^n-1} f''(B_{s_j^n}) - \int_0^t f''(B_s)\,\mathrm{d}s. \end{aligned}$$

Now $T_2 \to 0$ a.s., and hence in probability, by the theory of Riemann sums. As for T_1 we may note that: (a) 2^{-n} is the expectation of $(\Delta B)_{j,n}^2$; and (b) the summands in T_1 are independent. Therefore,

$$\mathrm{E}(T_1^2) \le \text{const} \cdot \|f''\|_\infty^2 \cdot \sum_{0\le j\le 2^n-1} \mathrm{E}\left[(\Delta B)_{j,n}^4\right] = \text{const} \cdot 2^{-n}.$$

This proves that when $f \in C_b^3(\mathbf{R})$ (1.2) holds for all $t \ge 0$. By localization, the general case follows once we derive (1.2) for $f \in C_b^2(\mathbf{R})$ and all $t \ge 0$. [Indeed we could consider the first time $|f(B_s)|+|f'(B_s)|+|f''(B_s)|$ exceeds k.] In that case, define $f_\epsilon := f * \phi_\epsilon$ where ϕ_ϵ denotes the density of $N(0\,,\epsilon)$. Itô's formula (1.1) is valid if we replace f by f_ϵ because the latter is in $C_b^\infty(\mathbf{R})$. By Fejer's theorem, $f_\epsilon(B_t) - f_\epsilon(B_0) \to f(B_t) - f(B_0)$ a.s. as $\epsilon \downarrow 0$. Also, by the dominated convergence theorem, $\int_0^t f_\epsilon(B_s)\,\mathrm{d}s \to \int_0^t f(B_s)\,\mathrm{d}s$ a.s. as $\epsilon \downarrow 0$. Finally, note that by the Itô isometry (p. 10),

$$\mathrm{E}\left(\left|\int_0^t \left(f_\epsilon'(B_s) - f'(B_s)\right)\,\mathrm{d}B_s\right|^2\right) = \mathrm{E}\left(\int_0^t \left|f_\epsilon'(B_s) - f'(B_s)\right|^2\,\mathrm{d}s\right).$$

And this tends to zero, as $\epsilon \downarrow 0$, by the dominated convergence theorem. This completes the proof. □

Itô's formula has many nice consequences. Here are a few examples.

Example 1.3 (Wiener's formula; $d = 1$)**.** We can apply the preceding form of the Itô formula with $f(t\,,x) = xg(t)$—for $g \in C^1(\mathbf{R})$—and $V_t := t$. In this case we obtain

$$g(t)B_t - g(0)B_0 = \int_0^t g(s)\,\mathrm{d}B_s + \int_0^t g'(s)B_s\,\mathrm{d}s.$$

Thus, in this special case the chain rule of ordinary calculus applies. □

Example 1.4 ($d = 1$)**.** In most cases, the rules of ordinary calculus do not hold for Itô integrals. For example, apply Itô's formula to $f(x) = x^2$ to find that

$$B_t^2 = B_0^2 + 2\int_0^t B\,\mathrm{d}B + t.$$

If f had finite variation, then the theory of Stieldjes integrals implies that $f_t^2 = f_0^2 + 2\int_0^t f\,\mathrm{d}f$. The extra t factor in the preceding Itô formula comes from the fact that Brownian motion does not have finite variation, and the Itô integral is not a random Stieldjes integral [nor is it the Lebesgue integral, for that matter]. $\square$

Recall that when $d = 1$ and $B_0 = 0$, the process $t \mapsto B_t^2 - t$ is a mean-zero martingale. The preceding tells us what that martingale is [namely, twice the Itô integral $\int_0^t B\,\mathrm{d}B$.] You might know also that $\exp(\lambda B_t - \frac{1}{2}t\lambda^2)$ defines a mean-one martingale, as long as λ is fixed. In order to identify that martingale we need a better Itô formula, which I leave for you to explore.

Exercise 1.5 (Important)**.** Suppose V is a nondecreasing, progressively-measurable stochastic process with values in $\mathbf{R}$. Then for all $f(t\,,x)$—continuously differentiable in time and twice continuously differentiable in space—we have the a.s. identity,

$$\begin{aligned}&f(V_t\,,B_t) - f(V_0\,,B_0)\\&= \int_0^t (\nabla f)(V_s\,,B_s)\,\mathrm{d}B_s + \int_0^t (\partial_t f)(V_s\,,B_s)\,\mathrm{d}V_s + \frac{1}{2}\int_0^t (\Delta f)(V_s\,,B_s)\,\mathrm{d}s,\end{aligned}$$

where ∇ and Δ act on the x variable, and $\partial_t := \partial/\partial t$.

Example 1.6 (McKean's exponential martingale; $d = 1$)**.** We apply the preceding form of the Itô formula to $f(t\,,x) := \exp(\lambda x - \lambda^2 t/2)$. Note that $\partial_x f = \lambda f$, $\partial^2_{xx} f = \lambda^2 f$, and $\partial_t f = -\frac{1}{2}\lambda^2 f$. Because $\frac{1}{2}\partial^2_{xx} f + \partial_t f = 0$, we have

$$e^{\lambda B_t - (\lambda^2 t/2)} - e^{\lambda B_0} = \lambda \int_0^t e^{\lambda B_s - (\lambda^2 s/2)}\,\mathrm{d}B_s.$$

In particular, $\exp(\lambda B_t - (\lambda^2 t/2))$ defines a martingale with mean $e^{-\lambda B_0}$. $\square$

Example 1.7 (Bessel processes)**.** Let B denote a d-dimensional Brownian motion started at $x \neq 0$, and define R to be its radial part; i.e.,

$$R_t := \|B_t\| \quad \text{for all } t \geq 0.$$

The process $R := \{R_t\}_{t\geq 0}$ is clearly progressively measurable, and in fact, continuous and adapted. Let $f(z) := \|z\| = (z_1^2 + \cdots + z_d^2)^{1/2}$ and note that $(\nabla f)(z) = z/\|z\|$ and $(\Delta f)(z) = (d-1)/\|z\|$, as long as $z \neq 0$. Itô's formula then implies that

$$R_t = \|x\| + \int_0^t \frac{B_s}{R_s}\,\mathrm{d}B_s + \frac{d-1}{2}\int_0^t \frac{\mathrm{d}s}{R_s} \quad \text{as long as } 0 < t < T_0,$$

where $T_0 := \inf\{s > 0 : B_s = 0\}$. In the next section we will see that $T_0 = \infty$ a.s. when $d \geq 2$. The process $R := \{R_t\}_{t\geq 0}$ is real-valued, and called the *Bessel process with dimension d*. The terminology will be justified next. $\square$

Example 1.8. Example 1.7 can be generalized in several ways. Here is an one that turns out to be interesting. Suppose B is d-dimensional Brownian motion started at $x \neq 0$, $R_t := \|B_t\|$, and h is C^2 on $(0\,,\infty)$. Define $f(z) := h(\|z\|)$ for all $z \in \mathbf{R}^d$. Then,

$$(\nabla f)(z) = \frac{z h'(\|z\|)}{\|z\|} \quad \text{and} \quad (\Delta f)(z) = (Lh)(\|z\|),$$

provided that $z \neq 0$. Here, L is the "d-dimensional Bessel operator,"

$$(Lh)(r) := \frac{d-1}{r}\, h'(r) + h''(r) \quad \text{for } r > 0.$$

Our discussion is one way to make precise the assertion that "L is the radial part of Δ." Itô's formula tells us that

$$h(R_t) = h(x) + \int_0^t \frac{h'(R_s)}{R_s} B_s\,\mathrm{d}B_s + \frac{1}{2}\int_0^t (Lh)(R_s)\,\mathrm{d}s, \tag{1.5}$$

for $t \in [0\,, T_0]$. We will return to this example soon. □

Exercise 1.9 (Bessel-squared processes)**.** Given a d-dimensional Brownian motion B, define $S_t := \|B_t\|^2$, and prove that for all $h \in C^2(\mathbf{R}_+)$ the following holds a.s.: For all $t \geq 0$,

$$h(S_t) = h(S_0) + \int_0^t (Gh)(S_s)\,\mathrm{d}s + \text{local martingale},$$

where $(Gh)(r) := 2rh''(r) + dh'(r)$. Identify the local martingale explicitly as a stochastic integral. The process $S := \{S_t\}_{t\geq 0}$ is called a *Bessel-squared process* of dimension d.

Exercise 1.10 (The Stratanovich integral)**.** Suppose $f : \mathbf{R} \to \mathbf{R}$ is continuously differentiable.

(1) Verify that for all $t \geq 0$ fixed, the following occurs uniformly for all $r \in [0\,, t]$: As $n \to \infty$,

$$\sum_{0\leq j<nr} f\left(B_{\tau_{j,n}}\right)\left[B_{(j+1)/n} - B_{j/n}\right] \xrightarrow{\text{P}} \int_0^r f(B_s)\,\mathrm{d}B_s + \frac{1}{2}\int_0^r f'(B_s)\,\mathrm{d}s,$$

where $\tau_{j,n} := \frac{j}{n} + (2n)^{-1}$ is the midpoint of $[j/n\,, (j+1)/n]$. The resulting "midpoint-rule integral" is sometimes written as $\int_0^t f(B)\,\delta B$, and referred to as the Stratanovich integral of $f \circ B$.

(2) Verify that $f(B_t) = f(B_0) + \int_0^t f'(B_s)\,\delta B_s$, provided that $f \in C^2(\mathbf{R})$; that is the fundamental theorem of calculus holds for the Stratanovich integral.

(3) More generally, suppose $\theta_{j,n} := \lambda(j+1)/n + (1-\lambda)j/n$ where $\lambda \in [0\,,1]$ is fixed. Then prove that if f is C^2 then as $n \to \infty$, $t \mapsto \sum_{j<nt} f(B_{\theta_{j,n}})\{B_{(j+1)/n} - B_{j/n}\}$ converges in $L^2(\mathrm{P})$ to a continuous adapted process $t \mapsto \int_0^t f(B_s)\,\delta_\lambda B_s$. Describe this integral.

2. Applications of Optional Stopping

We can couple the Itô calculus with the optional stopping theorem (p. 8), and obtain striking results. Below are some examples. But first, let me recall the principle of optional stopping in the least possible technical setting.

Example 2.1 (Gambler's ruin; $d = 1$)**.** Suppose $B_0 = 0$, and define T to be the first time Brownian motion exists $[-a\,,b]$, where $a, b > 0$ are nonrandom quantities. I will assume that you know the fact that $T < \infty$ a.s.

Because $T \wedge n$ is a bounded stopping time for every $n \geq 0$, and since B is a martingale with mean zero, $\mathrm{E}B_{T\wedge n} = \mathrm{E}B_0 = 0$. Because $\sup_n |B_{T\wedge n}| \leq \max(a\,,b)$, the dominated convergence theorem and the continuity of B together imply that $\mathrm{E}B_T = 0$. But B_T is a simple random variable [in the sense of integration theory]: It takes the value $-a$ with probability $\mathrm{P}\{B_T = -a\}$; and it takes on the value b with probability $1 - \mathrm{P}\{T = -a\}$. We can solve to see that $0 = -a\mathrm{P}\{B_T = -a\} + b(1 - \mathrm{P}\{B_T = -a\})$. It follows from this that

$$\mathrm{P}\left\{B \text{ hits } -a \text{ before } b\right\} = \frac{b}{a+b}.$$

We can also find $\mathrm{E}T$ easily: Since $B_t^2 - t$ is mean-zero martingale, $\mathrm{E}B_{T\wedge n}^2 = \mathrm{E}(T \wedge n)$. As $n \uparrow \infty$, $\mathrm{E}(T\wedge n) \uparrow \mathrm{E}(T)$ by the monotone convergence theorem. Since $\sup_n |B_{T\wedge n}|^2 < \infty$, $\mathrm{E}(B_{T\wedge n}^2)$ converges to $\mathrm{E}(B_T^2)$. Because B_T is a simple random variable,

$$\mathrm{E}(B_T^2) = a^2 \cdot \frac{b}{a+b} + b^2 \cdot \frac{a}{a+b} = ab.$$

Consequently, $\mathrm{E}(T) = ab$. $\square$

Example 2.2 (A formula of Lévy; $d = 1$)**.** Suppose $B_0 := y$, and let σ denote the first exit time of the interval $[x-a\,,x+a]$, where $a > 0$ and $x - a < y < x + a$. What is the distribution of σ?

In order to answer this question consider $f(t\,,z) := e^{-\lambda t} g(z)$ where $\lambda > 0$ and $g \in C^2(\mathbf{R})$. Itô's formula implies that

$$e^{-\lambda t} g(B_t) - g(B_0) = \int_0^t e^{-\lambda s} g'(B_s)\,\mathrm{d}B_s + \int_0^t e^{-\lambda s}\left(\frac{1}{2}g''(B_s) - \lambda g(B_s)\right)\mathrm{d}s.$$

If g solves the *eigenvalue problem* $g'' = 2\lambda g$ on $[x-a\,, x+a]$, then

$$e^{-\lambda(t\wedge\sigma)}g(B_{t\wedge\sigma}) - g(B_0) = \int_0^{t\wedge\sigma} e^{-\lambda s}g'(B_s)\,\mathrm{d}B_s.$$

By the optional stopping theorem, the right-hand side is a continuous mean-zero martingale. Hence, as long as $B_0 = y$, we have

$$\mathrm{E}\left[e^{-\lambda(t\wedge\sigma)}g(B_{t\wedge\sigma})\right] = g(y).$$

The aforementioned eigenvalue problem does not have a unique solution. But suppose g were a solution that is bounded on $[x-a\,, x+a]$. In that case we can let $t \to \infty$ and preserve the previous formula in the limit; i.e.,

$$\mathrm{E}\left[e^{-\lambda\sigma}g(B_\sigma)\right] = g(y).$$

The function $g(z) = \cosh((z-x)\sqrt{2\lambda})/\cosh(a\sqrt{2\lambda})$ is a bounded solution to the eigenvalue problem with $g(x\pm a) = g(B_\sigma) = 1$. Therefore, we obtain the elegant formula, due to Paul Lévy:

$$\mathrm{E}\left(e^{-\lambda\sigma}\;\middle|\; B_0 = y\right) = \frac{\cosh((y-x)\sqrt{2\lambda})}{\cosh(a\sqrt{2\lambda})}.$$

This describes the "Laplace transform" of T_a. □

Exercise 2.3. Prove that if $T_{-a,b} := \inf\{s > 0 :\ B_s = -a \text{ or } b\}$, then

$$\mathrm{E}\left(e^{-\lambda T_{-a,b}}\;\middle|\; B_0 = 0\right) = \frac{\cosh\left((b-a)\sqrt{\lambda/2}\right)}{\cosh\left((b+a)\sqrt{\lambda/2}\right)} \quad \text{for all } \lambda > 0.$$

In particular, conclude that $\mathrm{E}\exp(-\lambda T_z) = \exp(-|z|\sqrt{2\lambda})$ for all $z \in \mathbf{R}$, where $T_z := \inf\{s > 0 :\ B_s = z\}$.

Exercise 2.4 (Brownian motion with drift)**.** Define $Y_t := B_t + \mu t$, where B is linear Brownian motion started at $x \in \mathbf{R}$, and $\mu \in \mathbf{R}$ is fixed. The proces $Y := \{Y_t\}_{t\geq 0}$ is sometimes called *Brownian motion with drift* μ.

(1) Find an $f \in C^2(\mathbf{R})$ such that $f \circ Y$ defines a martingale.

(2) Use the preceding to compute the probability that Y hits $-a$ before b, where $0 < a, b$.

(3) Compute the probability that there exists no time $t \geq 0$ at which $B_t \geq b - \mu t$. (Warning: The case $\mu \geq 0$ is rather different from $\mu < 0$; why?)

Example 2.5. The preceding example can be generalized. Let B denote d-dimensional Brownian motion with $B_0 = 0$. Consider a bounded closed set $D \subset \mathbf{R}^d$, with C^∞ boundary, that contains 0 in its interior. Let $T_D := \inf\{s > 0 : \ B_s \notin D\}$. It is not hard to see that T is an a.s.-finite stopping time. Here is one way to prove this: Since D is bounded, there exists a [possibly large] constant $R > 0$ such that $D \subset B(0\,, R)$. Therefore, it is enough to show that $T_{B(0,R)} < \infty$ a.s. But $\mathrm{P}\{\|B_t\| > R\} = \mathrm{P}\{\|B_1\| > R/t^{1/2}\} \to 1$ as $t \to \infty$. Therefore, by the Kolmogorov zero-one law $T_{B(0,R)}$, and hence T_D, is finite a.s. (Why?) Consider the Dirichlet problem $\Delta g = \lambda g$ in the interior of D. Since D has a smooth boundary, standard elliptic theory tells us that this PDE has a unique smooth solution g that is bounded on D and one on ∂D. By the Itô formula (Exercise 1.5), $\exp(-\lambda T_D)f(B_{t\wedge T_D}) = f(0)+$ a mean-zero martingale. Since f is bounded in D, it follows readily from this that $\mathrm{E}\exp(-\lambda T_D) = f(0)$. This is the beginning of deep connection between PDEs and Brownian motion. □

Exercise 2.6 (The Dirichlet problem)**.** Consider a bounded closed domain $D \subset \mathbf{R}^d$, and suppose there exist $f \in C^2(D)$ and a bounded $F : \partial D \to \mathbf{R}$ that satisfy $(\Delta f)(x) = 0$ for $x \in \operatorname{int} D$ and $f(y) = F(y)$ for $y \in \partial D$. Define $T := \inf\{s > 0 : \ B_s \in \partial D\}$ $(\inf \varnothing := \infty)$, and prove that $f(x) = \mathrm{E}[F(B_T) \,|\, B_0 = x]$ for all $x \in D$, where B denotes d-dimensional Brownian motion.

Exercise 2.7 (An eigenvalue problem)**.** Consider a bounded domain $D \subset \mathbf{R}^d$, and suppose there exist $f \in C^2(\bar{D})$ and a bounded $F : \partial D \to \mathbf{R}$ that satisfy the following boundary value [eigenvalue] problem for some $\lambda \geq 0$: $(\Delta f)(x) = 2\lambda f(x)$ for all $x \in D$ and $f(x) = F(x)$ for all $x \in \partial D$. Let $T := \inf\{s > 0 : \ B_s \in \partial D\}$ $(\inf \varnothing := \infty)$ for a d-dimensional Brownian motion B, and prove that:

(1) $M_t := e^{-\lambda(t\wedge T)}f(B_{t\wedge T})$ defines a bounded martingale; and
(2) $f(x) = \mathrm{E}[e^{-\lambda T}F(B_T) \,|B_0 = x]$ for all $x \in D$. Thus, the boundary function F determines f everywhere.

Example 2.8 $(d \geq 3)$**.** Let B denote a d-dimensional Brownian motion started at $x \neq 0$, and $R_t := \|B_t\|$ is the correponding Bessel process (Example 1.7). According to Example 1.8, for all $f \in C^2(0\,,\infty)$,

$$f(R_t) - f(x) = \frac{1}{2}\int_0^t (Lf)(R_s)\,\mathrm{d}s + \text{local martingale} \quad \text{for } t < T_0,$$

where $(Lf)(r) = f''(r) + (d-1)f'(r)/r$. Note that if we set $f(r) := r^{2-d}$, then $Lf \equiv 0$, and therefore,

$$\|B_t\|^{2-d} = \|x\|^{2-d} + \text{local martingale} \quad \text{for } t < T_0.$$

Let $T_r := \inf\{s > 0 : R_s = r\}$ for $r > 0$, and note that as long as $\rho > \|x\| > r > 0$, $\|B_{t\wedge T_r\wedge T_\rho}\|^{2-d} = \|x\|^{2-d}+$ a mean-zero martingale and $\sup_t \|B_{t\wedge T_r\wedge T_\rho}\| < \infty$. The optional stopping theorem tells us that

$$\mathrm{E}\left(\|B_{T_r\wedge T_\rho}\|^{2-d}\right) = \|x\|^{2-d}.$$

But $\|B_{T_r\wedge T_\rho}\|^{2-d}$ is a simple random variable: It is equal to r^{2-d} with probability $\mathrm{P}\{T_r < T_\rho\}$, and ρ^{2-d} with probability $1-\mathrm{P}\{T_r < T_\rho\}$. Consequently,

$$\mathrm{P}\left\{T_r < T_\rho\right\} = \frac{\|x\|^{2-d} - \rho^{2-d}}{r^{2-d} - \rho^{2-d}}.$$

In particular, if we let $\rho \uparrow \infty$, we obtain the following: $\mathrm{P}\{T_r < \infty\} = (r/\|x\|)^{d-2}$ (why?). Let $r \downarrow 0$ to deduce that $T_0 = \infty$ a.s. (why?). □

Exercise 2.9 ($d \geq 3$)**.** Define $X_t := \|B_t\|^{2-d}$ where B is d-dimensional Brownian motion and $d \geq 3$. Prove that $X := \{X_t\}_{t\geq 0}$ is a local martingale, but not a martingale. (Hint: Try proving that the expectation of X_t is finite but vanishes as $t \to \infty$.)

Exercise 2.10 ($d = 2$)**.** Suppose B is planar Brownian motion started at $x \neq 0$. Prove that whenever $\rho > \|x\| > r > 0$,

$$\mathrm{P}\left\{T_r < T_\rho\right\} = \frac{\ln(1/\|x\|) - \ln(1/\rho)}{\ln(1/r) - \ln(1/\rho)}.$$

Conclude from this that $T_0 = \infty$ a.s.

Exercise 2.11 ($d \geq 2$)**.** Suppose B is Brownian motion started at an arbitrary point $x \in \mathbf{R}^d$. Prove that $\mathrm{P}\{B_t = y \text{ for some } t > 0\} = 0$, for all $y \in \mathbf{R}^d$.

Exercise 2.12. Prove that if B is d-dimensional Brownian motion, then $\lim_{t\to\infty} \|B_t\| = \infty$ a.s. That is, B is *transient* in dimensions greater than or equal to three. Also, prove that if $d = 2$, then $\{t > 0 : \|B_t\| < \epsilon\}$ is a.s. unbounded for all nonrandom $\epsilon > 0$. That is, planar Brownian motion is *neighborhood recurrent* (though not recurrent, since $T_0 := \infty$).

Exercise 2.13 (Hard)**.** Improve the last assertion to the following: With probability one, *planar Brownian motion defines a space-filling curve whose image has zero planar Lebesgue measure.* That is, with probability one, the closure of $B(\mathbf{R}_+)$ is all of $\mathbf{R}^2$, but the 2-dimensional Lebesgue measure of $B(\mathbf{R}_+)$ is zero. This is due to Paul Lévy.

3. A Glimpse at Harmonic Functions

The theory of harmonic functions has intimate connections with that of Brownian motion. We scratch the surface by mentioning three facts in this regard.

Definition 3.1. Let D be a domain in $\mathbf{R}^d$. We say that $f : D \to \mathbf{R}$ is *harmonic* in D if $\Delta f = 0$ on D, where $\Delta := \sum_{j=1}^d \partial^2_{jj}$ denotes the Laplacian.

Proposition 3.2 (Doob, Dynkin). *Suppose f is harmonic in $D \subset \mathbf{R}^d$. Let B be d-dimensional Brownian motion, and define $T_D := \inf\{s > 0 : B_s \notin D\}$. Then, $\{f(B_t)\}_{t\in[0,T_D)}$ is a local martingale.*

Proof. We apply Itô's formula to find that

$$f(B_t) = f(B_0) + \int_0^t (\nabla f)(B_s)\,\mathrm{d}B_s \quad \text{provided that } 0 \le t < T_D.$$

The proposition follows. □

Proposition 3.3 (Gauss). *If f is harmonic on a domain $D \subset \mathbf{R}^d$, then for all open balls $U(z)$ that are centered at z and lie in the interior $\mathrm{int}(D)$ of D,*

$$f(z) = \int_{\partial U(z)} f(y)\, m(\mathrm{d}y) \quad \textit{for all } z \in \mathrm{int}(D),$$

where m denotes the uniform distribution on $\partial U(z)$.

Proof. Let B denote d-dimensional Brownian motion started at z. Define $\tau := \inf\{s > 0 : B_s \notin U(z)\}$, and note that because B has continuous trajectories, $\tau < T_D$ a.s. Thanks to Proposition 3.2, and by the optional stopping theorem (p. 8), $\mathrm{E}f(B_{t\wedge\tau}) = f(z)$ for all $t \ge 0$. Because f is C^2 on $U(z)$, it is bounded there. Consequently, $\mathrm{E}f(B_\tau) = f(z)$. It remains to prove that the distribution of B_τ is m. We may observe the following two facts: (a) There is only one uniform probability measure on $\partial U(z)$; and (b) for all $(d \times d)$ rotation matrices Θ, $t \mapsto z + \Theta(B_t - z)$ is a a Brownian motion that starts at z. Consequently, the distribution of B_τ is the uniform measure on $\partial U(z)$. □

Proposition 3.4 (Liouville). *If f is a bounded harmonic function on $\mathbf{R}^d$, then f is a constant.*

Proof. We know from Proposition 3.2 that $t \mapsto f(B_t)$ is a local martingale. Because f is bounded, it follows that $f \circ B$ is in fact a bounded martingale; see, Exercise 6.3 on page 13, for instance. Thanks to the martingale convergence theorem (page 8), $Z := \lim_{t\to\infty} f(B_t)$ exists a.s. and in $L^p(\mathrm{P})$ for all $p \in [1\,,\infty)$. Moreover, $\mathrm{P}\{f(B_t) = \mathrm{E}(Z\,|\,\mathscr{F}_t)\} = 1$ for all $t \geq 0$. But according to the Kolmogorov zero-one, Z is a constant. This shows that $f(B_1)$ is a.s. constant. But given an arbitrary $x \in \mathbf{R}^d$ and $\epsilon > 0$, $\mathrm{P}\{\|B_1 - x\| < \epsilon\} > 0$. Therefore, $|f(x) - Z| < \epsilon$ with positive probability. Because Z is a constant and x and ϵ are arbitrary, this proves that $f(x) = Z$ for all $x \in \mathbf{R}^d$. □

Chapter 3

Martingale Calculus

1. Introduction

Let $M := \{M_t\}_{t\geq 0}$ denote a continuous martingale with respect to a filtration $\mathscr{F} := \{\mathscr{F}_t\}_{t\geq 0}$. Presently, we aim to mimic the work of Chapter 1, and construct stochastic integrals of the form $\int_0^t G\,\mathrm{d}M$ for a suitably-large family of processes G. Throughout we assume that $\mathscr{F}$ satisfies the usual hypotheses. In the beginning portions of the construction, we assume also that X is square integrable; that is, $\mathrm{E}(|M_t|^2) < \infty$ for all $t \geq 0$.

We say that G is *simple* if there exist $a < b$—both ≥ 0—such that

$$G_t(\omega) = X(\omega)\mathbf{1}_{[a,b)}(t), \tag{1.1}$$

for all $(\omega\,, t) \in \Omega \times \mathbf{R}_+$, and X is in $L^2(\mathrm{P})$ and $\mathscr{F}_a$-measurable. In this case we define $\int_0^t G\,\mathrm{d}M$, as in Chapter 1, viz.,

$$\int_0^t G\,\mathrm{d}M := X\left(M_{t\wedge b} - M_{t\wedge a}\right). \tag{1.2}$$

Clearly, $t \mapsto \int_0^t G\,\mathrm{d}M$ is a continuous martingale with

$$\mathrm{E}\left(\int_0^t G\,\mathrm{d}M\right) = 0 \quad \text{and} \quad \mathrm{E}\left(\left|\int_0^t G\,\mathrm{d}M\right|^2\right) = \mathrm{E}\left(X^2\,|M_{t\wedge b} - M_{t\wedge a}|^2\right),$$

though the second moment could, in principle, be infinity. In the case that $M := B$ is Brownian motion, the L^2-computation was simplified by the fact that $B_{t\wedge b} - B_{t\wedge a}$ is independent of X. And this was the starting point of the stochastic calculus of B. In the present, more general, case we need to inspect the second moment more carefully.

2. Stochastic Integrals against Continuous local Martingales

2.1. The Square-Integrable Case. Throughout this subsection, M is assumed to be a square integrable continuous martingale with respect to $\mathscr{F}$. In several spots we refer to the following, which is the result of a simple computation: Almost surely,

$$\mathrm{E}\left(|M_t - M_s|^2 \,\middle|\, \mathscr{F}_s\right) = \mathrm{E}\left(M_t^2 - M_s^2 \,\middle|\, \mathscr{F}_s\right) \quad \text{for all } t \geq s \geq 0. \tag{2.1}$$

Step 1. If $G_t(\omega) = X(\omega)\mathbf{1}_{[a,b)}(t)$ for a bounded $\mathscr{F}_a$-measurable random variable X, then we define

$$\int_0^t G\,\mathrm{d}M := X\left(M_{t\wedge b} - M_{t\wedge a}\right) \quad \text{for all } t \geq 0.$$

Clearly, $t \mapsto \int_0^t G\,\mathrm{d}M$ is a continuous square-integrable martingale with mean zero. Moreover. $\mathrm{E}(|\int_0^t G\,\mathrm{d}M|^2) = \mathrm{E}(X^2|M_{t\wedge b} - M_{t\wedge a}|^2)$.

Step 2. For all $\ell \geq 0$, let $\mathscr{E}_\ell$ denote the collection of all finite [the number being nonrandom] linear combination of bounded simple processes G of the form $X(\omega)\mathbf{1}_{[a,b)}(t)$ where X is $\mathscr{F}_a$-measurable and $|G_t| \leq \ell$ a.s.

Any $G \in \mathscr{E}_\ell$ can be written as $G = G^1 + \cdots + G^k$, where the G^k's are simple processes—uniformly bounded by ℓ in absolute value—and $G_s^i G_t^j = 0$ for all $s, t \geq 0$, provided that $i \neq j$. Given this representation, we can define $\int_0^t G\,\mathrm{d}M := \sum_{j=1}^k \int_0^t G^j\,\mathrm{d}M$. This is well defined, and in fact defines a continuous square-integrable martingale with mean zero. In addition,

$$\begin{aligned}\mathrm{E}\left(\left|\int_0^t G\,\mathrm{d}M\right|^2\right) &= \sum_{j=1}^k \mathrm{E}\left(\left|\int_0^t G^j\,\mathrm{d}M\right|^2\right) \\ &= \sum_{j=1}^k \mathrm{E}\left(X_j^2 \cdot \left|M_{t\wedge b_j} - M_{t\wedge a_j}\right|^2\right),\end{aligned}$$

where $G_t^j = X_j \mathbf{1}_{[a_j,b_j)}$, $\mathrm{P}\{|X_j| \leq \ell\} = 1$, and X_j is $\mathscr{F}_{a_j}$-measurable. Without loss of generality we may assume also that $a_j < b_j \leq a_{j+1}$, else we can reorder the G^j's suitably.

By (2.1),

$$\mathrm{E}\left(\left|\int_0^t G\,\mathrm{d}M\right|^2\right) \leq \left\|\max_{1\leq j\leq k} X_j\right\|_{L^\infty(\mathrm{P})}^2 \cdot \sum_{j=1}^k \mathrm{E}\left(M_{t\wedge b_j}^2 - M_{t\wedge a_j}^2\right),$$

where $\|\cdots\|_{L^\infty(\mathrm{P})}$ denotes the essential supremum with respect to P.

The conditional Jensen inequality implies that $\mathrm{E}(M_t^2 \,|\, \mathscr{F}_s) \geq M_s^2$ a.s. provided that $t \geq s \geq 0$. [I.e., M^2 is a submartingale in the same sense as the discrete-parameter theory.] Consequently, $t \mapsto \mathrm{E}(M_t^2)$ is nondecreasing, and hence

$$\mathrm{E}\left(\left|\int_0^t G\,\mathrm{d}M\right|^2\right) \le \left\|\sup_{s\in[0,t]}|G_s|\right\|_{L^\infty(\mathrm{P})}^2 \cdot \mathrm{E}\left(M_{t\wedge b_k}^2 - M_{t\wedge a_1}^2\right)$$
$$\le \left\|\sup_{s\in[0,t]}|G_s|\right\|_{L^\infty(\mathrm{P})}^2 \cdot \mathrm{E}(M_t^2).$$

Thanks to this, and Doob' inequality, for all $t \ge 0$,

$$\mathrm{E}\left(\sup_{r\in[0,t]}\left|\int_0^r G\,\mathrm{d}M\right|^2\right) \le 4\left\|\sup_{s\in[0,t]}|G_s|\right\|_{L^\infty(\mathrm{P})}^2 \cdot \mathrm{E}(M_t^2). \tag{2.2}$$

Step 2. If G is a continuous adapted process, then we can apply standard arguments from integration theory and find $G^1, G^2, \ldots \in \mathscr{E}_\ell$ such that

$$\lim_{n\to\infty}\left\|\sup_{s\in[0,\rho]}|G_s^n - G_s|\right\|_{L^\infty(\mathrm{P})} = 0 \quad \text{for all } \rho \in (0\,,\infty).$$

The preceding shows that $\int_0^t G\,\mathrm{d}M := \lim_{n\to\infty}\int_0^t G^n\,\mathrm{d}M$. Since $L^2(\mathrm{P})$-limits of martingales are themselves martingales, it follows that $t \mapsto \int_0^t G\,\mathrm{d}M$ is a mean-zero martingale. Thanks to (2.2), it is also continuous since it is the uniform-on-compact limit of continuous processes. Finally, $t \mapsto \int_0^t G\,\mathrm{d}M$ satisfies (2.2). We define $\int_r^t G\,\mathrm{d}M := \int_0^t G\,\mathrm{d}M - \int_0^r G\,\mathrm{d}M = \int_0^t \mathbf{1}_{[r,t)}G\,\mathrm{d}M$.

Step 3. Suppose M is a continuous martingale $[\mathscr{F}]$ that is a.s. bounded in absolute value by a nonrandom constant ℓ, and define

$$\langle M\rangle_t := M_t^2 - M_0^2 - 2\int_0^t M\,\mathrm{d}M \quad \text{for all } t \ge 0. \tag{2.3}$$

The process $\langle M\rangle$ is called the *quadratic variation* of M.

Note that for all $t \ge 0$,

$$\sum_{0\le j\le nt-1}\left(M_{(j+1)/n} - M_{j/n}\right)^2$$
$$= -2\sum_{0\le j\le nt-1} M_{j/n}\left(M_{(j+1)/n} - M_{j/n}\right) + M_{[nt]/n}^2 - M_0^2$$
$$= -2\int_0^t G^n\,\mathrm{d}M + M_t^2 - M_0^2 + \epsilon_n(t) \quad \text{a.s.},$$

where $|\epsilon_n(t)| \le \sup_{u,v\in[0,t]:\,|u-v|\le 1/n}|M_u - M_v| := \rho_n(t)$, and

$$G_s^n := \sum_{0\le j\le nt-1} M_{j/n}\mathbf{1}_{[j/n,(j+1)/n)}(s).$$

Note that each G^n is in $\mathscr{E}_\ell$, and $|G^n_s - M_s| \le \rho_n(t)$, uniformly for all $s \in [0\,, t]$. Because the G^n's, and M, are all uniformly bounded by ℓ is modulus, and since M is a.s. continuous, it follows that $\|G^n - M\|_{L^\infty(\mathrm{P})} \to 0$ as $n \to \infty$. This proves that for all $t \ge 0$,

$$\sup_{u\in[0,t]} \left| \langle M\rangle_u - \sum_{0\le j\le nu-1} \left(M_{(j+1)/n} - M_{j/n}\right)^2 \right| \to 0 \quad \text{as } n\to\infty, \tag{2.4}$$

where convergence takes place in $L^2(\mathrm{P})$. Because the sum is a nondecreasing [random] function on u, it follows that $u \mapsto \langle M\rangle_u$ is nondecreasing a.s. Moreover, (2.3) shows that $u \mapsto \langle M\rangle_u$ is continuous and adapted, as well.

If $t \ge s \ge 0$, then (2.3) shows that $\langle M\rangle_t - \langle M\rangle_s = M_t^2 - M_s^2 + 2\int_s^t M\,\mathrm{d}M$, and hence by the martingale property, and thanks to (2.1), whenever $t \ge s \ge 0$,

$$\mathrm{E}\left((M_t - M_s)^2 \,\middle|\, \mathscr{F}_s\right) = \mathrm{E}\left(\langle M\rangle_t - \langle M\rangle_s \mid \mathscr{F}_s\right) \quad \text{a.s.} \tag{2.5}$$

Finally, suppose $G \in \mathscr{E}_\ell$ as well, so that $G_t = \sum_{j=1}^k X_j \mathbf{1}_{[a_j,b_j)}(t)$, where $|X_j| \le \ell$, each X_j is $\mathscr{F}_{a_j}$-measurable, and the $[a_j\,, b_j)$'s are disjoint. We can retrace our steps in Step 2, and find that

$$\begin{aligned}
\mathrm{E}\left(\left|\int_0^t G\,\mathrm{d}M\right|^2\right) &= \sum_{j=1}^k \mathrm{E}\left(X_j^2 \cdot \left|M_{t\wedge b_j} - M_{t\wedge a_j}\right|^2\right)\\
&= \sum_{j=1}^k \mathrm{E}\left(X_j^2 \left\{\langle M\rangle_{t\wedge b_j} - \langle M\rangle_{t\wedge a_j}\right\}\right).
\end{aligned}$$

Equivalently,

$$\mathrm{E}\left(\left|\int_0^t G\,\mathrm{d}M\right|^2\right) = \mathrm{E}\left(\int_0^t |G|^2\,\mathrm{d}\langle M\rangle\right). \tag{2.6}$$

Step 4. If M is a continuous martingale that is square integrable, then consider the stopping times $T_k := \inf\{s > 0 : |M_s| > k\}$ and martingales $M_t^{T_k} := M_{t\wedge T_k}$. The quadratic variation $\langle M^{T_k}\rangle$ of M^{T_k} satisfies

$$\langle M^{T_k}\rangle_t = \int_0^t M^{T_k}\,\mathrm{d}M^{T_k} + M^2_{T_k\wedge t} - M_0^2.$$

The construction of the stochastic integral [Step 2] shows that $\int_0^t M^{T_k}\,\mathrm{d}M^{T_k} = \int_0^t M^{T_{k+1}}\,\mathrm{d}M^{T_{k+1}}$ for all $t \in [0\,, T_k]$. Thus, we can define $\int_0^t M\,\mathrm{d}M := \int_0^t M^{T_k}\,\mathrm{d}M^{T_k}$ for all $t \in [0\,, T_k]$ and $k \ge 1$. This defines $t \mapsto \int_0^t M\,\mathrm{d}M$ as a continuous local martingale. Furthermore, the preceding and (2.3) together show that $\langle M^{T_k}\rangle_t = \langle M^{T_{k+1}}\rangle_t$ for all $t \in [0\,, T_k]$. Thus, there exists a continuous nondecreasing adapted process $t \mapsto \langle M\rangle_t$—the *quadratic variation of*

M—that satisfies (2.3) and (2.5). Another approximation argument—very much like Steps 3–4 in Chapter 1—shows that we can define $t \mapsto \int_0^t G\,\mathrm{d}M$ as a continuous square-integrable martingale as long as G is a progressively measurable process that satisfies $\mathrm{E}(\int_0^t |G|^2\,\mathrm{d}\langle M\rangle) < \infty$. Moreover, the "Itô isometry" (2.6) continues to hold in that case.

We summarize our findings.

Theorem 2.1. *If M is a continuous square-integrable martingale with respect to $\mathscr{F}$, then there exists a continuous adapted nondecreasing process $\langle M\rangle$—the quadratic variation of M—that satisfies* (2.3) *and* (2.5). *Moreover, if G is progressively measurable and $\mathrm{E}(\int_0^t |G|^2\,\mathrm{d}\langle M\rangle) < \infty$ for all $t \geq 0$, then $t \mapsto \int_0^t G\,\mathrm{d}M$ can be defined as a continuous square-integrable martingale that satisfies the Itô isometry* (2.6). *Finally, if G and F are two such processes and $a, b \in \mathbf{R}$, then $\int_0^t (aF + bG)\,\mathrm{d}M = a\int_0^t F\,\mathrm{d}M + b\int_0^t G\,\mathrm{d}M$ for all $t \geq 0$, a.s.*

2.2. The General Case. If M is a continuous local martingale then it still has a quadratic variation $\langle M\rangle$ that is defined in Step 4 by localization.

If G is progressively measurable, then we can construct the stochastic integral $t \mapsto \int_0^t G\,\mathrm{d}M$ as a continuous local martingale, provided that

$$\mathrm{P}\left\{\int_0^t |G_s|^2\,\mathrm{d}\langle M\rangle_s < \infty\right\} = 1 \quad \text{for all } t \geq 0.$$

Moreover, (2.3) holds [also by localization]. We omit the details.

Exercise 2.2. Check that if B is a Brownian motion, then $\langle B\rangle_t = t$. Now (2.3) should seem familiar.

Exercise 2.3. Suppose M is a continuous local martingale [$\mathscr{F}$]. Prove that (2.4) is valid, with convergence taking place in probability. If $\mathrm{E}(|M_t|^2) < \infty$ for all $t \geq 0$, then prove that the said convergence holds also in $L^2(\mathrm{P})$.

Exercise 2.4. Suppose M is a continuous local martingale [$\mathscr{F}$]. Prove that if M has a.s.-finite variation, then $M_t \equiv M_0$ a.s. Use this to prove that if M is a continuous [local] martingale with $M_0 = 0$, then there always exist unique continuous adapted nondecreasing processes $A := \{A_t\}_{t\geq 0}$ and $\tilde{A} := \{\tilde{A}_t\}_{t\geq 0}$ such that $t \mapsto M_t^2 - A_t$ and $t \mapsto \exp(\lambda M_t - \lambda^2 \tilde{A}_t/2)$ are [local] martingales and $A_0 := 0$. Identify A_t and $\tilde{A}_t$ for all $t \geq 0$.

Exercise 2.5 (Important)**.** Suppose M is a continuous square-integrable local martingale [$\mathscr{F}$] and G is progressively measurable and $\int_0^t |G_s|^2\,\mathrm{d}\langle M\rangle_s <$

∞ a.s. for all $t \geq 0$. Prove that for all stopping times T,

$$\mathrm{P}\left\{\int_0^{T\wedge t} G\,\mathrm{d}M = \int_0^t G\mathbf{1}_{[0,T)}\,\mathrm{d}M \quad \text{for all } t \geq 0\right\} = 1.$$

3. Itô's Formula

Theorem 3.1. *Let M denote a continuous local martingale. Suppose V is an adapted nondecreasing process and $f = f(t\,,x)$ is C^1 in the t variable and C^2 in the x variable. Then a.s.,*

$$\begin{aligned}&f(V_t\,,M_t) - f(V_0\,,M_0)\\&= \int_0^t \frac{\partial f}{\partial t}(V_s\,,M_s)\,\mathrm{d}V_s + \int_0^t \frac{\partial f}{\partial x}(V_s\,,M_s)\,\mathrm{d}M_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(V_s\,,M_s)\,\mathrm{d}\langle M\rangle_s.\end{aligned}$$

The proof is modeled after that of the Itô formula for Brownian motion. So I will discuss only the novel differences.

Sketch of proof. We consider only that case that f depends on the x variable, for simplicity. Thus, Itô's formula becomes

$$f(M_t) - f(M_0) = \int_0^t f'(M_s)\,\mathrm{d}M_s + \frac{1}{2}\int_0^t f''(M_s)\,\mathrm{d}\langle M\rangle_s. \tag{3.1}$$

Let T_k denote the first time $t \geq 0$ when $|f(M_t)| + |f'(M_t)| + |f''(M_t)| + |M_t| + |\langle M\rangle_t|$ exceeds the integer $k \geq 0$. We need only prove (3.1) for all $t \in [0\,,T_k]$. Thanks to Step 4 (p. 30), $\int_0^{t\wedge T_k} G\,\mathrm{d}M = \int_0^t G\mathbf{1}_{[0,T_k]}\,\mathrm{d}M$. Consequently, it suffices to derive (3.1), for all $t \geq 0$, in the case that M, $\langle M\rangle$, f, f', and f'' are all bounded. In that case, the proof of (3.1) is very close to its Brownian-motion analogue. We need only add the following result: $\sum_{0\leq j\leq nt-1} f''(M_{j/n})\{\langle M\rangle_{(j+1)/n} - \langle M\rangle_{j/n}\} \to \int_0^t f''(M_s)\,\mathrm{d}\langle M\rangle_s$, uniformly on t-compacts, in probability. When $f'' \equiv$ constant, this is contained in (2.4). The general case follows by a similar argument used to derive the latter fact. □

4. Examples, Extensions, and Refinements

4.1. Brownian Motion. Let $M_t := B_t$ be a linear Brownian motion with respect to $\mathscr{F}$. Then, $B_t - B_s$ is independent of $\mathscr{F}_s$ and has variance $t - s$. That is, $\mathrm{E}((B_t - B_s)^2\,|\,\mathscr{F}_s) = t - s$, whence we can choose $\langle B\rangle_t := t$. Thus, the Itô isometry (2.6) is nothing but the Itô isometry (5.2) on page 10.

The following shows that the preceding constant-quadratic variation calculus is strictly a Brownian-motion story.

Theorem 4.1 (Lévy)**.** *If M is a continuous local martingale $[\mathscr{F}]$ with $\langle M\rangle_t := t$, then M is a Brownian motion.*

Proof. Without loss of generality we may assume that $M_0 = 0$, else we consider $M_t - M_0$ in place of M_t everywhere.

First let us prove that M is necessarily a square-integrable martingale. Since $\langle M\rangle_t = t$, it follows from our construction of quadratic variation that $\langle M^T\rangle_t = t\wedge T$ for all stopping times T, where $M_t^T := M_{t\wedge T}$. Thus, whenever $\{T_k\}_{k\geq 1}$ is a localizing sequence, $\mathrm{E}(|M_{t\wedge T_k}|^2) = \mathrm{E}(t\wedge T_k) \leq t$. Fatou's lemma then implies that $M_t \in L^2(\mathrm{P})$ for all $t \geq 0$, and $\lim_k M_{t\wedge T_k} = M_t$ in $L^2(\mathrm{P})$. Because L^2-limits of martingales are themselves martingales, this proves that M is a square-integrable martingale, as asserted.

An application of Itô's formula then shows that for all fixed constants $\lambda \in \mathbf{R}$,

$$\exp\left(\lambda M_t - \frac{\lambda^2}{2}t\right) = 1 + \lambda \int_0^t \exp\left(\lambda M_s - \frac{\lambda^2}{2}s\right) \mathrm{d}M_s.$$

In other words, if $\mathscr{E}_t$ defines the left-most quantity, then it satisfies $\mathscr{E}_t = 1 + \lambda \int_0^t \mathscr{E}_s \,\mathrm{d}M_s$, which in formal notation, reduces to the "stochastic differential equation," $\mathrm{d}\mathscr{E}_t = \lambda\mathscr{E}_t \,\mathrm{d}M_t$ subject to $\mathscr{E}_0 = 1$.

Let $\{T_k\}_{k=1}^\infty$ denote a localizing sequence for $\mathscr{E}$. We apply the optional stopping to the preceding and then appeal to Fatou's lemma to find that $\mathrm{E}\mathscr{E}_t \leq 1$ for all $t \geq 0$. This proves that $\exp(\lambda M_t)$ is integrable for all $t \geq 0$ and $\lambda \in \mathbf{R}$. Since $\lambda \in \mathbf{R}$ is arbitrary, $\exp(\lambda M_t)$ is square-integrable. Consequently, $\mathscr{E}$ is a mean-one continuous martingale for any and every $\lambda \in \mathbf{R}$. This implies then that $\mathrm{E}(\mathscr{E}_t \,|\mathscr{F}_s) = \mathscr{E}_s$ whenever $t \geq s \geq 0$. Equivalently,

$$\mathrm{E}\left(e^{\lambda(M_t - M_s)} \,\middle|\, \mathscr{F}_s\right) = e^{-\lambda^2(t-s)/2} \qquad \text{a.s. for } t \geq s \geq 0.$$

It follows that $M_t - M_s$ is independent of $\mathscr{F}_s$ and is $N(0\,, t-s)$. The theorem follows readily from this and the definition of Brownian motion. $\square$

Theorem 4.2 (Dambis, Dubins–Schwarz)**.** *Suppose M is a continuous local martingale with $M_0 := 0$ and $\langle M\rangle_\infty := \infty$. Define*

$$\tau(t) := \inf\{s > 0 : \langle M\rangle_s \geq t\} \quad \textit{for all } t \geq 0.$$

Then, $B_t := M_{\tau(t)}$ defines a Brownian motion in the filtration $\mathscr{F} \circ \tau := \{\mathscr{F}_{\tau(t)}\}_{t\geq 0}$.

Proof. By the optional stopping theorem, B defines a continuous local martingale in $\mathscr{F} \circ \tau$. Since its quadratic variation is t at time t, Lévy's theorem (Theorem 4.1) completes the proof. $\square$

4.2. Itô Integrals. Let $M_t := \int_0^t G\,\mathrm{d}N$, where N is a continuous square-integrable martingale $[\mathscr{F}]$, and G is progressively measurable and satisfies $\mathrm{E}(\int_0^t |G_s|^2\,\mathrm{d}\langle N\rangle_s) < \infty$ a.s. for all $t \geq 0$. We know that M is a continuous square-integrable martingale. Let us apply the Itô isometry (2.6) with G replaced by $G\mathbf{1}_{[0,T]}$—for a stopping time T—to conclude that

$$\mathrm{E}(|M_{T\wedge t}|^2) = \mathrm{E}\left(\int_0^{T\wedge t} |G_s|^2\,\mathrm{d}\langle N\rangle_s\right) \quad \text{for all stopping times } T.$$

Because $M_t^2 - \langle M\rangle_t$ is a continuous martingale, it follows that

$$\mathrm{E}\left(\langle M\rangle_{T\wedge t} - \int_0^{T\wedge t} |G_s|^2\,\mathrm{d}\langle N\rangle_s\right) = 0 \quad \text{for all stopping times } T.$$

We apply this with T being the first time $t \geq 0$ that $\langle M\rangle_t$ is at least $\int_0^t |G_s|^2\,\mathrm{d}\langle N\rangle_s + \epsilon$, where $\epsilon > 0$ is fixed. This proves that $T = \infty$ a.s., whence $\langle M\rangle_t \leq \int_0^t |G_s|^2\,\mathrm{d}\langle N\rangle_s$ for all $t \geq 0$, a.s. By symmetry,

$$\left\langle \int_0^\bullet G\,\mathrm{d}N \right\rangle_t = \int_0^t |G_s|^2\,\mathrm{d}\langle N\rangle_s \quad \text{for all } t \geq 0, \text{ a.s.}$$

By localization the preceding holds also when N is a continuous local martingale and $\int_0^t |G_s|^2\,\mathrm{d}\langle N\rangle_s < \infty$ a.s. for all $t \geq 0$ (why?). In that case, however, M is only a local martingale.

If H is progressively measurable, then $\int_0^t H_s^2\,\mathrm{d}\langle M\rangle_s = \int_0^t H_s^2 G_s^2\,\mathrm{d}\langle N\rangle_s$. If this is finite a.s. for all $t \geq 0$, then I claim that

$$\int_0^t H\,\mathrm{d}M = \int_0^t HG\,\mathrm{d}N.$$

It suffice to prove this when G and H are bounded simple processes. But in that case this identity is nearly a tautology.

4.3. Mutual Variation. If M and N are two continuous local martingales $[\mathscr{F}]$, then so are $M+N$ and $M-N$. By the definition of quadratic variation (or Itô's formula), $(M_t+N_t)^2 - (M_0+N_0)^2 = \langle M+N\rangle_t +$ local martingale and $(M_t-N_t)^2 - (M_0-N_0)^2 = \langle M-N\rangle_t +$ local martingale. We can take the difference to find that

$$M_tN_t - M_0N_0 - \frac{1}{4}\left(\langle M+N\rangle_t - \langle M-N\rangle_t\right) = \text{local martingale}. \tag{4.1}$$

Define the *mutual variation* between N and M to be the process $\langle N\,,M\rangle$, where

$$\langle N\,,M\rangle_t = \langle M\,,N\rangle_t := \frac{\langle M+N\rangle_t - \langle M-N\rangle_t}{4}.$$

Exercise 4.3. Check that:

(1) $\langle M\,,M\rangle_t = \langle M\rangle_t$; and

(2) $\langle M \pm N\rangle_t = \langle M\rangle_t + \langle N\rangle_t \pm 2\langle M\,,N\rangle_t$.

The following shows that (4.1) characterizes mutual variation.

Lemma 4.4. *If $M_tN_t - M_0N_0 = A_t +$ local martingale for a continuous adapted A of a.s.-finite variation such that $A_0 = 0$, then $A_t \equiv 0$ a.s. for all $t \geq 0$.*

Proof. Suppose $M_tN_t - M_0N_0 = A'_t +$ local martingale as well, where A' is a continuous adapted process of a.s.-finite variation, and $A'_0 = 0$ a.s. It then follows that $X_t := A_t - A'_t$ is a local martingale that starts at zero, and has finite variation a.s. A real-variable argument then shows that

$$\sum_{0\leq j\leq nt-1} \left(X_{(j+1)/n} - X_{j/n}\right)^2 \to 0 \quad \text{a.s. as } n\to\infty.$$

Consequently, $\langle X\rangle_t \equiv 0$ a.s., and hence X^2 is a local martingale with $X_0^2 = 0$. Thus, we can find a localizing sequence $T_1 \leq T_2 \leq \cdots$ of stopping times with $\lim_k T_k = \infty$ such that $\mathrm{E}(|X_{T_k\wedge t}|^2) = 0$. But then Doob's inequality implies that $X_t \equiv 0$ a.s. □

Here is an important consequence.

Lemma 4.5. *If $B := (B^1\,,\ldots,B^d)$ is a d-dimensional Brownian motion, then $\langle B^i\,,B^j\rangle_t \equiv 0$ for $i\neq j$.*

Proof. Let $\mathscr{F} := \mathscr{F}^B$ denote the Brownian filtration. Note that if $i\neq j$, then B^i and B^j are independent one-dimensional Brownian motions. By Lemma 4.4, it suffices to prove that $B^i_tB^j_t$ defines a martingale with respect to $\mathscr{F}^B$, as long as $i\neq j$.

Choose and fix $t\geq s\geq 0$. We can write $B^i_tB^j_t = (B^i_t - B^i_s)(B^j_t - B^j_s) + B^i_sB^j_t + B^j_sB^i_t - B^i_sB^j_s$ to deduce from the Markov property that $\mathrm{E}(B^i_tB^j_t\,|\,\mathscr{F}^B_s) = B^i_sB^j_s$. Therefore, $B^i_tB^j_t$ defines a martingale. □

Theorem 4.6 (Itô's Formula)**.** *Let $M^1,\ldots,M^d$ denote $d\geq 2$ continuous local martingales $[\mathscr{F}]$. Then, for all $f\in C^2(\mathbf{R}^d\,,\mathbf{R})$,*

$$\begin{aligned} &f\left(M^1_t\,,\ldots,M^d_t\right) - f\left(M^1_0\,,\ldots,M^d_0\right)\\ &\quad= \sum_{j=1}^d \int_0^t (\partial_j f)\left(M^1_s\,,\ldots,M^d_s\right)\,\mathrm{d}M^j_s\\ &\qquad\qquad + \frac12 \sum\sum_{1\leq i,j\leq d} (\partial^2_{ij}f)\left(M^1_s\,,\ldots,M^d_s\right)\,\mathrm{d}\left\langle M^i\,,M^j\right\rangle_s. \end{aligned}$$

Proof. The theorem holds by definition when $d = 2$ and $f(x\,,y) = xy$. By induction the result follows in the case that $d \geq 2$ and $f(x_1\,,\ldots,x_d) = x_i g(y)$ for $g \in C^2(\mathbf{R}^{d-1}\,,\mathbf{R})$, where y is the same as the vector x except it is missing the ith coordinate. In particular, the theorem holds for all polynomials $f : \mathbf{R}^d \to \mathbf{R}$. Let K denote a fixed compact subset of $\mathbf{R}^d$, and let T denote the first time that at least one of the M^j's exits K. The Stone–Weierstrass theorem shows that we can find polynomials that approximate f uniformly well on K. This proves that the theorem holds for all $f \in C^2(\mathbf{R}^d\,,\mathbf{R})$ for all $t \in [0\,,T]$. Since K is arbitrary, the result hold for all $t \geq 0$. □

Proposition 4.7. *Let M and N denote continuous local martingales $[\mathscr{F}]$, and assume that G and H are progressively measurable processes such that $\int_0^t G^2\,\mathrm{d}\langle M\rangle$ and $\int_0^t H^2\,\mathrm{d}\langle N\rangle$ are finite a.s. for all $t \geq 0$. Then, the mutual variation between $\int_0^t G\,\mathrm{d}M$ and $\int_0^t H\,\mathrm{d}N$ is $\int_0^t G_s H_s\,\mathrm{d}\langle M\,,N\rangle_s$.*

Proof. By symmetry, we need only consider the case that $H_s \equiv 1$. Also, localization reduces the problem to the case that M and N are bounded and G is simple and bounded.

If $c \geq b \geq a$, then

$$\mathrm{E}(N_c M_b \,|\, \mathscr{F}_a) = \mathrm{E}\left(\{N_c - N_b\} M_b \,|\, \mathscr{F}_a\right) + \mathrm{E}(N_b M_b \,|\, \mathscr{F}_a) = \mathrm{E}(N_b M_b \,|\, \mathscr{F}_a),$$

since $\mathrm{E}(N_c - N_b \,|\, \mathscr{F}_b) = 0$ by the martingale property. Consequently,

$$\mathrm{E}(N_c M_b \,|\, \mathscr{F}_a) = N_a M_a + \mathrm{E}\left(\langle M\,,N\rangle_b - \langle M\,,N\rangle_a \,|\, \mathscr{F}_a\right).$$

From this it follows that whenever $s \leq r < r + \epsilon < t$,

$$\mathrm{E}\left(N_t G_r M_{r+\epsilon} \,|\, \mathscr{F}_s\right) = \mathrm{E}\left(G_r \left\{N_r M_r + \langle N\,,M\rangle_{r+\epsilon} - \langle N\,,M\rangle_r\right\} \,|\, \mathscr{F}_s\right).$$

We also have $\mathrm{E}(N_t G_r M_r \,|\, \mathscr{F}_s) = \mathrm{E}(G_r N_r M_r \,|\, \mathscr{F}_s)$. Therefore,

$$\mathrm{E}\left(N_t G_r \left\{M_{r+\epsilon} - M_r\right\} \,|\, \mathscr{F}_s\right) = \mathrm{E}\left(G_r \left\{\langle N\,,M\rangle_{r+\epsilon} - \langle N\,,M\rangle_r\right\} \,|\, \mathscr{F}_s\right).$$

Because the process G of the proposition is bounded and simple, it has the form $G_r\{M_{r+\epsilon} - M_r\}$ for a bounded $\mathscr{F}_r$-measurable random variable G_r. The preceding proves

$$\mathrm{E}\left(N_t \cdot \int_0^t G\,\mathrm{d}M \,\middle|\, \mathscr{F}_s\right) = \mathrm{E}\left(\int_0^t G\,\mathrm{d}\langle M\,,N\rangle \,\middle|\, \mathscr{F}_s\right),$$

whence the result. □

Exercise 4.8 (Kunita–Watanabe)**.** Prove that under the conditions of Proposition 4.7,

$$\left|\int_0^t G_s H_s\,\mathrm{d}\langle M\,,N\rangle_s\right|^2 \leq \int_0^t |G_s|^2\,\mathrm{d}\langle M\rangle_s \cdot \int_0^t |H_s|^2\,\mathrm{d}\langle N\rangle_s.$$

(Hint: To begin, compute $\langle M+\lambda N\rangle_t - \langle M+\lambda N\rangle_s$ for $t \ge s \ge 0$ and $\lambda \in \mathbf{R}$, and optimize over λ.)

4.4. Semimartingales. A continuous [local] *semimartingale* [$\mathscr{F}$] X is a process that can be written as $X_t = M_t + A_t$, where M is a continuous [local] martingale and A is a continuous adapted process that has finite variation a.s., and $A_0 := 0$. It is an elementary real-variable consequence of finite variation (plus continuity) that for all $t > 0$,

$$\sup_{r\in[0,t]} \sum_{0\le j\le nr-1} \left(A_{(j+1)/n} - A_{j/n}\right)^2 \to 0 \quad \text{as } n\to\infty \text{, a.s.}$$

Equation (2.4) and localization together prove that for all $t \ge 0$,

$$\sup_{r\in[0,t]} \left| \sum_{0\le j\le nr-1} \left(X_{(j+1)/n} - X_{j/n}\right)^2 - \langle M\rangle_r \right| \to 0 \text{ in probability, as } n\to\infty.$$

Thus, we *define* the quadratic variation $\langle X\rangle$ of X as $\langle X\rangle_t := \langle M\rangle_t$. Also, $\int_0^t G\,\mathrm{d}X := \int_0^t G\,\mathrm{d}M + \int_0^t G\,\mathrm{d}A$, provided that both integrals make sense; i.e., that G is progressively measurable and $\int_0^t |G|^2\,\mathrm{d}\langle M\rangle < \infty$ a.s. for all $t \ge 0$. With these conventions in mind, one checks readily that Itô's formula (Theorem 3.1) extends to the following: for all continuous adapted processes V of bounded variation,

$$\begin{aligned}
&f(V_t\,, X_t) - f(V_0\,, X_0)\\
&= \int_0^t (\partial_x f)(V_s\,, X_s)\,\mathrm{d}X_s\\
&\qquad + \int_0^t (\partial_t f)(V_s\,, X_s)\,\mathrm{d}V_s + \frac12 \int_0^t (\partial_{xx} f)(V_s\,, X_s)\,\mathrm{d}\langle X\rangle_s\\
&= \int_0^t (\partial_x f)(V_s\,, X_s)\,\mathrm{d}M_s + \int_0^t (\partial_x f)(V_s\,, X_s)\,\mathrm{d}A_s\\
&\qquad + \int_0^t (\partial_t f)(V_s\,, X_s)\,\mathrm{d}V_s + \frac12 \int_0^t (\partial_{xx} f)(V_s\,, X_s)\,\mathrm{d}\langle M\rangle_s.
\end{aligned} \tag{4.2}$$

4.5. Bessel Processes. Let B denote a d-dimensional Brownian motion started at $x \ne 0$, and define $R_t := \|B_t\|$ to be the corresponding Bessel process in dimension d. In Example 1.7 (p. 18) we saw that

$$R_t = \|x\| + \sum_{j=1}^d \int_0^t \frac{B_s^j}{R_s}\,\mathrm{d}B_s^j + \frac{d-1}{2}\int_0^t \frac{\mathrm{d}s}{R_s}, \tag{4.3}$$

for all $t \in [0\,, T_0]$ where T_0 is the first hitting time of zero by R. In fact, $T_0 = \infty$ a.s.; see Example 2.8 and Exercise 2.10 on page 22. Thus, (4.3) holds for all $t \ge 0$, a.s., and R is a local semimartingale whose quadratic variation is the same as the quadratic variation of the stochastic integral. If $i \neq j$, then the mutual variation between $\int_0^t B_s^j R_s^{-1}\, \mathrm{d}B_s^j$ and $\int_0^t B_s^i R_s^{-1}\, \mathrm{d}B_s^i$ is zero (Lemma 4.5). Consequently,

$$\langle R\rangle_t = \sum_{j=1}^d \left\langle \int_0^\bullet \frac{B_s^j}{R_s}\, \mathrm{d}B_s^j \right\rangle_t = t.$$

Lévy's theorem proves then that

$$R_t = \|x\| + W_t + \frac{d-1}{2}\int_0^t \frac{\mathrm{d}s}{R_s},$$

where W is a linear Brownian motion with respect to $\mathscr{F}$ that starts at zero. Moreover, by Itô's formula,

$$\begin{aligned} h(R_t) &= h(x) + \int_0^t h'(R_s)\, \mathrm{d}R_s + \frac{1}{2}\int_0^t h''(R_s)\, \mathrm{d}\langle R\rangle_s \\ &= h(x) + \int_0^t h'(R_s)\, \mathrm{d}W_s + \frac{d-1}{2}\int_0^t \frac{h'(R_s)}{R_s}\, \mathrm{d}s + \frac{1}{2}\int_0^t h''(R_s)\, \mathrm{d}s. \end{aligned}$$

This identifies the stochastic integral of (1.5) (p. 19) as a stochastic integral against a one-dimensional Brownian motion.

5. Tanaka's Formula

Let $X = M + V$ be a continuous local semimartingale: M is a continuous local martingale $[\mathscr{F}]$; and V is a continuous adapted process of finite variation. By Itô's formula (4.2),

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\, \mathrm{d}M_s + \int_0^t f'(X_s)\, \mathrm{d}V_s + \frac{1}{2}\int_0^t f''(X_s)\, \mathrm{d}\langle M\rangle_s,$$

for all $t \geq 0$. We can try to extend this by asking what happens if $\mu = f''$ exists as a signed measure. This means that for all $\phi \in C_c^\infty(\mathbf{R})$,[1]

$$\int_{-\infty}^\infty \phi''(t) f(t)\, \mathrm{d}t = \int_{-\infty}^\infty \phi(t)\, \mu(\mathrm{d}t).$$

Integration by parts shows that if $f \in C^2(\mathbf{R})$, then the preceding holds with $\mu(\mathrm{d}t) = f''(t)\, \mathrm{d}t$. But the preceding can make sense when f'' does not exist as a function. A notable example is provided by the function $f(x) = |x|$. Clearly, $f'(x) = \mathrm{sign}(x)$ for all $x \neq 0$.

In order to find f'' we compute for $\phi \in C_c^\infty(\mathbf{R})$,

$$\int_{-\infty}^\infty \phi''(x)|x|\, \mathrm{d}x = \int_0^\infty x\left(\phi''(x) + \phi''(-x)\right)\, \mathrm{d}x.$$

[1] The subscript "c" refers to "compact support."

We integrate this by parts to find that

$$\int_{-\infty}^{\infty} \phi''(x)|x|\,\mathrm{d}x = -\int_0^{\infty} \big(\phi'(x) - \phi'(-x)\big)\,\mathrm{d}x = 2\phi(0).$$

That is, if $f(x) := |x|$, then $f'' = 2\delta_0$ in the sense of distributions.

Let us first consider the special case that M, $A_t := \int_0^t \mathrm{d}|V_s|$, and $\langle M\rangle$ are bounded processes, so that M is also a bounded continuous martingale. Let ϕ_ϵ denote the $N(0\,,\epsilon)$ density. Then $f_\epsilon := f * \phi_\epsilon \in C^\infty(\mathbf{R})$, $f'_\epsilon = f * \phi'_\epsilon$ and $f''_\epsilon = f * \phi''_\epsilon$. Therefore, by Itô's formula,

$$\begin{aligned} f_\epsilon(X_t - a) - f_\epsilon(X_0 - a) = \int_0^t f'_\epsilon(X_s - a)\,\mathrm{d}M_s + \int_0^t f'_\epsilon(X_s - a)\,\mathrm{d}V_s \\ + \frac{1}{2}\int_0^t f''_\epsilon(X_s - a)\,\mathrm{d}\langle M\rangle_s. \end{aligned} \tag{5.1}$$

Because f is continuous, Fejér's theorem tells us that $f_\epsilon(x) \to f(x) = |x|$ as $\epsilon \downarrow 0$, for all $x \in \mathbf{R}$. Thus, the left-hand side of the preceding display converges to $|X_t - a| - |X_0 - a|$ a.s., and the convergence holds uniformly for t-compacts.

Since $f'_\epsilon(x) = (f' * \phi_\epsilon)(x)$, a direct computation shows that $f'_\epsilon(x) = \operatorname{sign}(x)\mathrm{P}\{|N(0\,,\epsilon)| \le |x|\}$, and hence also that $f''_\epsilon(x) =$ twice density of $|N(0\,,\epsilon)|$ at x. If we define $\operatorname{sign}(0) := 0$, then it follows also that $f'_\epsilon \to \operatorname{sign}$ as $\epsilon \downarrow 0$, boundedly and pointwise.

Now we return to the stochastic integral and compute:

$$\begin{aligned} &\mathrm{E}\left(\left|\int_0^t f'_\epsilon(X_s - a)\,\mathrm{d}M_s - \int_0^t \operatorname{sign}(X_s - a)\,\mathrm{d}M_s\right|^2\right) \\ &\qquad = \mathrm{E}\left(\int_0^t \left|f'_\epsilon(X_s - a) - \operatorname{sign}(X_s - a)\right|^2\,\mathrm{d}\langle M\rangle_s\right) \\ &\qquad \to 0 \quad \text{as } \epsilon \downarrow 0, \end{aligned}$$

thanks to the dominated convergence theorem. Moreover, by Doob's inequality, this holds uniformly for t-compacts, as well. Similarly, the boundedness of the total variation A of V implies that

$$\mathrm{E}\left(\sup_{r\in[0,t]}\left|\int_0^r f'_\epsilon(X_s - a)\,\mathrm{d}V_s - \int_0^r \operatorname{sign}(X_s - a)\,\mathrm{d}V_s\right|^2\right) \to 0 \quad \text{as } \epsilon \downarrow 0.$$

Therefore, there exists a continuous adapted process $\{L^a_t(X)\}_{t\ge0}$—one for every $a \in \mathbf{R}$—such that

$$|X_t - a| = |X_0 - a| + \int_0^t \operatorname{sign}(X_s - a)\,\mathrm{d}X_s + L^a_t(X) \quad \text{for all } t \ge 0. \tag{5.2}$$

[Recall that $\int_0^t H\,\mathrm{d}X := \int_0^t H\,\mathrm{d}M + \int_0^t H\,\mathrm{d}V$.] Moreover, for all $t \geq 0$,

$$\lim_{\epsilon\downarrow 0}\sup_{r\in[0,t]}\left|\frac{1}{2}\int_0^r f_\epsilon''(X_s-a)\,\mathrm{d}\langle M\rangle_s - L_r^a(X)\right| = 0 \quad \text{in } L^2(\mathrm{P}). \tag{5.3}$$

Because $f_\epsilon''(x) \geq 0$ for all x, $t \mapsto L_t^a(X)$ is nondecreasing. Finally, note that for all $0 \leq u \leq v$ and $\delta > 0$,

$$\begin{aligned}\int_u^v f_\epsilon''(X_s-a)\mathbf{1}_{\{|X_s-a|>\delta\}}\,\mathrm{d}\langle M\rangle_s \\ &= \text{const}\cdot\int_u^v \frac{e^{-|X_s-a|^2/(2\epsilon)}}{\epsilon^{1/2}}\mathbf{1}_{\{|X_s-a|>\delta\}}\,\mathrm{d}\langle M\rangle_s \\ &\leq \text{const}\frac{e^{-\delta^2/(2\epsilon)}}{\epsilon^{1/2}}\left(\langle M\rangle_v - \langle M\rangle_u\right),\end{aligned}$$

and this converges to zero uniformly over all such choices of u and v. Consequently,

$$\lim_{\epsilon\downarrow 0}\sup_{0\leq u\leq v\leq t}\left|\frac{1}{2}\int_u^v f_\epsilon''(X_s-a)\mathbf{1}_{\{|X_s-a|\leq\delta\}}\,\mathrm{d}\langle M\rangle_s - \left(L_v^a(X)-L_u^a(X)\right)\right| = 0,$$

in $L^2(\mathrm{P})$. This completes our proof of the following in the case that X, M, A, and $\langle M\rangle$ are bounded; the general case follows from localization.

Theorem 5.1 (Tanaka)**.** *Choose and fix $a \in \mathbf{R}$, and let X be a continuous local semimartingale $[\mathscr{F}]$. There exists a continuous adapted process $t \mapsto L_t^a(X)$ that is nondecreasing, grows only on the level set $\{t \geq 0 : X_t = a\}$, and solves* (5.2)*.*

Remark 5.2. Equation (5.2) is called *Tanaka's formula.* □

The process $t \mapsto L_t^a(X)$ is called the *local time of X at a,* since it is a natural measure of how much time X spends at level a.

Exercise 5.3. Prove that if B is linear Brownian motion with $B_0 = 0$, then $|B_t| - L_t^0(B)$ is a Brownian motion.

The next exercise has deep consequences in Markov-process theory.

Exercise 5.4 (Wang)**.** Suppose f is C^1, and $f'' = \mu$, in the sense of distributions, where μ is a σ-finite Borel measure on $\mathbf{R}$. Prove then that there exists a nondecreasing continuous and adapted process $I^\mu := \{I_t^\mu\}_{t\geq 0}$ that grows only on the support of μ and solves

$$f(X_t) - f(X_0) = \int_0^t (f'\circ X)\,\mathrm{d}X + I_t^\mu.$$

Exercise 5.5. Prove that if f'' exists almost everywhere, then $f(B_t) = f(B_0) + \int_0^t (f' \circ B)\,\mathrm{d}B + \frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s$.

Next we choose a modification of local times that have improved measure-theoretic properties.

Theorem 5.6 (Occupation density formula)**.** *There exists a measurable map* $(t\,,a\,,\omega) \mapsto \hat{L}_t^a(X)(\omega)$ *such that*

$$\mathrm{P}\left\{\hat{L}_t^a(X) = L_t^a(X) \textit{ for all } t \geq 0\right\} = 1 \quad \textit{for all } a \in \mathbf{R}. \tag{5.4}$$

Moreover, for all nonnegative measurable functions $\phi : \mathbf{R} \to \mathbf{R}$*,*

$$\mathrm{P}\left\{\int_0^t \phi(X_s)\,\mathrm{d}\langle X\rangle_s = \int_{-\infty}^{\infty} \phi(a)\hat{L}_t^a(X)\,\mathrm{d}a \quad \textit{for all } t \geq 0\right\} = 1. \tag{5.5}$$

Remark 5.7. This is called the *occupation density formula* for historical reasons: Consider the case that $M = B$ is linear Brownian motion. In that case we obtain

$$\int_0^t \mathbf{1}_A(B_s)\,\mathrm{d}s = \int_A L_t^a\,\mathrm{d}a.$$

The left-hand side defines a measure called the *occupation density*, and we find that local time is the Radon–Nikodým density of the occupation density of Brownian motion. □

Proof. If T is a stopping time and $X_t^T := X_{T\wedge t}$, then

$$\mathrm{P}\left\{L_t^a(X^T) = L_{t\wedge T}^a(X) \quad \text{for all } t \geq 0\right\} = 1 \quad \text{for all } a \in \mathbf{R}.$$

Therefore, localization shows us that we can assume without loss of generality that M, $\langle M\rangle$, and $A_t := \int_0^t \mathrm{d}|V|_s$ are all bounded. We have derived (5.3) under the boundedness assumptions of the present proof. By the Borel–Cantelli lemma there exists a nonrandom sequence $\epsilon(1) > \epsilon(2) > \cdots$—decreasing to zero—such that

$$\mathrm{P}\left\{\lim_{k\to\infty} \frac{1}{2}\int_0^t f''_{\epsilon(k)}(X_s - a)\,\mathrm{d}\langle M\rangle_s = L_t^a(X) \text{ for all } t \geq 0\right\} = 1 \quad \text{for all } a \in \mathbf{R}.$$

Define

$$\hat{L}_t^a(X) := \limsup_{k\to\infty} \frac{1}{2}\int_0^t f''_{\epsilon(k)}(X_s - a)\,\mathrm{d}\langle M\rangle_s, \tag{5.6}$$

to readily obtain (5.4). Let Γ denote the [random] collection of all points $a \in \mathbf{R}$ for which the lim sup in (5.6) is in fact a limit. Because (5.6) asserts that $\mathrm{P}\{a \notin \Gamma\} = 0$ for all $a \in \mathbf{R}$, Tonelli's theorem shows that the expectation of the Lebesgue measure of Γ is zero. It follows immediately that, with probability one, Γ has zero Lebesgue measure. Therefore, for all nonrandom $\phi \in C_c^\infty(\mathbf{R})$, the following holds a.s.:

$$\lim_{k\to\infty} \frac{1}{2} \int_{-\infty}^{\infty} \int_0^t f''_{\epsilon(k)}(X_s - a)\phi(a)\,\mathrm{d}\langle M\rangle_s\,\mathrm{d}a = \int_{-\infty}^{\infty} \hat{L}_t^a(X)\phi(a)\,\mathrm{d}a.$$

The integrand on the left-hand side is $(f''_{\epsilon(k)} * \phi)(X_s)$. By Fejer's theorem, this converges to $2\phi(X_s)$ uniformly in s as $k \uparrow \infty$. This completes the proof. □

Our next goal is to derive good conditions that guarantee that $(t\,,a) \mapsto L_t^a(X)$ can be chosen to be a continuous process as well (after possibly making a modification). The following is a first step toward that goal.

Lemma 5.8. *Choose and fix $s \in \mathbf{R}$. Then,*

$$\mathrm{P}\left\{ \int_0^t \mathbf{1}_{\{X_s = a\}}\,\mathrm{d}\langle M\rangle_s = 0 \textit{ for all } t \geq 0 \right\} = 1.$$

Proof. Thanks to (5.2), $Y_t := |X_t - a|$ defines a semimartingale, and by Itô's formula, $Y_t^2 = Y_0^2 + 2\int_0^t Y\,\mathrm{d}Y + \langle Y\rangle_t$. This and (5.2) together yield

$$\begin{aligned}(X_t - a)^2 &- (X_0 - a)^2 \\ &= 2\int_0^t (X_s - a)\,\mathrm{d}X_s + 2\int_0^t |X_s - a|\,\mathrm{d}L_s^a(X) + \int_0^t \mathbf{1}_{\{X_s \neq a\}}\,\mathrm{d}\langle M\rangle_s.\end{aligned}$$

Because local time grows on $\{s:\ X_s = a\}$, we have $\int_0^t |X_s - a|\,\mathrm{d}L_s^a(X) = 0$ a.s., whence

$$(X_t - a)^2 - (X_0 - a)^2 = 2\int_0^t (X_s - a)\,\mathrm{d}X_s + \int_0^t \mathbf{1}_{\{X_s \neq a\}}\,\mathrm{d}\langle M\rangle_s.$$

By Itô's formula (4.2), the preceding is also equal to $2\int_0^t (X_s - a)\,\mathrm{d}X_s + \langle M\rangle_t$, since the quadratic variation of $X_t - a$ is $\langle M\rangle$. Comparing terms, we obtain

$$\mathrm{P}\left\{ \int_0^t \mathbf{1}_{\{X_s = a\}}\,\mathrm{d}\langle M\rangle_s = 0 \right\} = 1 \quad \text{for all } t \geq 0.$$

Consequently, $\int_0^t \mathbf{1}_{\{X_s = a\}}\,\mathrm{d}\langle M\rangle_s = 0$ for all rational t's a.s., and hence we obtain the result by continuity. □

Proposition 5.9. *The process I has a continuous modification J, where*

$$I_t^a := \int_0^t \operatorname{sign}(X_s - a)\,\mathrm{d}M_s.$$

If, in addition, we have

$$\mathrm{P}\left\{ \int_0^t \mathbf{1}_{\{X_s = a\}}\,\mathrm{d}V_s = 0 \textit{ for all } t \geq 0 \textit{ and } a \in \mathbf{R} \right\} = 1. \tag{5.7}$$

then $(t\,,a) \mapsto \int_0^t \operatorname{sign}(X_s)\,\mathrm{d}X_s$ *has a continuous modification.*

Proof. By localization we may, and will, assume that M, $\langle M\rangle$, and $A_t := \int_0^t \mathrm{d}|V_s|$ are bounded.

Given three real numbers $a < b$ and x,

$$|\operatorname{sign}(x-a) - \operatorname{sign}(x-b)|^2 = 4\mathbf{1}_{(a,b)}(x) + \mathbf{1}_{\{a,b\}}(x).$$

Consider the following square-integrable martingale:

$$D_t := I_t^a - I_t^b.$$

Lemma 5.8, and the occupation density formula (Theorem 5.6) together show that

$$\langle D\rangle_t = 4\int_0^t \mathbf{1}_{\{a<X_s<b\}}\,\mathrm{d}\langle M\rangle_s = 4\int_a^b L_t^z(X)\,\mathrm{d}z \quad \text{a.s.} \tag{5.8}$$

Itô's formula implies that $(D_t)^4 = 4\int_0^t (D_s)^3\,\mathrm{d}D_s + 6\int_0^t (D_s)^2\,\mathrm{d}\langle D\rangle_s$. Let $\{T_k\}_{k=1}^\infty$ localize D to find that

$$\begin{aligned}
\mathrm{E}\left(|D_{t\wedge T_k}|^4\right) &= 6\mathrm{E}\left(\int_0^{t\wedge T_k} D_s^2\,\mathrm{d}\langle D\rangle_s\right) \le 6\mathrm{E}\left(\sup_{s\le t\wedge T_k} D_s^2 \cdot \langle D\rangle_t\right)\\
&\le 6\sqrt{\mathrm{E}\left(\sup_{s\le t\wedge T_k} D_s^4\right)\cdot \mathrm{E}\left(\langle D\rangle_t^2\right)} \le \text{const}\cdot\sqrt{\mathrm{E}\left(|D_{t\wedge T_k}|^4\right)\cdot\mathrm{E}\left(\langle D\rangle_t^2\right)},
\end{aligned}$$

thanks to Doob's maximal inequality. Solve for the fourth moment, let $k\uparrow\infty$, and appeal to Fatou's lemma to find that

$$\mathrm{E}\left(\sup_{s\in[0,t]}|D_s|^4\right) \le \text{const}\cdot\mathrm{E}\left[\langle D\rangle_t^2\right] = \text{const}\cdot\int_a^b\int_a^b \mathrm{E}\left(L_t^z(X)L_t^w(X)\right)\,\mathrm{d}z\,\mathrm{d}w,$$

by (5.8). Consequently,

$$\mathrm{E}\left(\sup_{s\in[0,t]}|D_s|^4\right) \le \text{const}\cdot(b-a)^2\sup_z \mathrm{E}\left(|L_t^z(X)|^2\right).$$

The Kolmogorov continuity theorem (p. 5) proves that I has a continuous modification as soon as we prove that the second moment of $L_t^z(X)$ is bounded in z. But the Tanaka formula (5.2) implies that $L_t^z(X)$ is at most $|X_t - X_0| + |I_t^z|$, which is bounded by $\text{const} + |I_t^z|$. Hence, it suffices to prove that $\mathrm{E}(|I_t^z|^2)$ is bounded in z. We note further that $|I_t^z|$ is bounded above by

$$\left|\int_0^t \operatorname{sign}(X_s-z)\,\mathrm{d}M_s\right| + \left|\int_0^t \operatorname{sign}(X_s-z)\,\mathrm{d}V_s\right| \le \left|\int_0^t \operatorname{sign}(X_s-z)\,\mathrm{d}M_s\right| + A_t.$$

Because A is bounded, it suffices to prove that the second moment of the stochastic integral is bounded in z. But that follows readily because the quadratic variation of that integral is at most $\langle M\rangle_t$, which is bounded uniformly in z.

Now let us assume (5.7) holds, and define $J_t^a := \int_0^t \operatorname{sign}(X_s - a)\, \mathrm{d}V_s$. That is, for each $a \in \mathbf{R}$ fixed, the following holds a.s.:

$$J_t^a = \int_0^t \mathbf{1}_{\{X_s>a\}}\, \mathrm{d}V_s - \int_0^t \mathbf{1}_{\{X_s<a\}}\, \mathrm{d}V_s = 2\int_0^t \mathbf{1}_{\{X_s>a\}}\, \mathrm{d}V_s - V_t.$$

By localization we may assume that $A_t := \int_0^t \mathrm{d}|V_s|$ is bounded. But then it is clear from the dominated convergence theorem and (5.7) that the final integral is continuous in a, for all $t \geq 0$, outside a single null set. This proves the theorem since $\int_0^t \operatorname{sign}(X_s - a)\, \mathrm{d}X_s = I_t^a + J_t^a$ by definition. □

We finally arrive at the main result of this section.

Theorem 5.10. *If* (5.7) *holds, then there exist a continuous process* $(t\,,a) \mapsto \bar{L}_t^a(X)$ *such that*

$$\mathrm{P}\left\{\bar{L}_t^a(X) = \hat{L}_t^a(X) = L_t^a(X) \text{ for all } t \geq 0\right\} = 1 \quad \text{for all } a \in \mathbf{R}.$$

Proof. Let Q_t^a denote the continuous modification of the map $(t\,,a) \mapsto \int_0^t \operatorname{sign}(X_s - a)\, \mathrm{d}X_s$, and define

$$\bar{L}_t^a(X) := |X_t - a| - |X_0 - a| - Q_t^a.$$

The result follows from Tanaka's formula (5.2) and Proposition 5.9. □

Because $\bar{L}(X)$, $\hat{L}(X)$, and $L(X)$ are modifications of one another, we can justifiably discard the first two and use the better-behaved $\bar{L}(X)$ as our choice of local times. Thus, from now on we always mean $L(X)$ to be a continuous modification of local times. [If $L(X)$ is not continuous then it is replaced by $\bar{L}(X)$, and the "bar" is removed from the L by a slight abuse of notation.]

Exercise 5.11. Prove that L_t^a also satisfies

$$(M_t - a)^+ = (M_t - a)^+ + \int_0^t \mathbf{1}_{\{M_s>a\}}\, \mathrm{d}M_s + \frac{1}{2}L_t^a,$$

and

$$(M_t - a)^- = (M_t - a)^- - \int_0^t \mathbf{1}_{\{M_s\leq a\}}\, \mathrm{d}M_s + \frac{1}{2}L_t^a.$$

Exercise 5.12 (Burkholder)**.** Let M denote a continuous local martingale. Show that if $\langle M\rangle_t \in L^k(\mathrm{P})$ for an integer $k \geq 1$, then $M_t \in L^{2k}(\mathrm{P})$. Do this by proving that there exists a constant c_k such that for all $k \geq 1$ and $t \geq 0$,

$$\mathrm{E}\left[\sup_{s\in[0,t]} M_s^{2k}\right] \leq c_k \mathrm{E}\left[\langle M\rangle_t^k\right].$$

6. The Zero-Set of Brownian Motion

Now we apply the preceding local-time theory to describe the "fractal" structure of the level sets of a Brownian motion B. Without loss of too much generality we may assume that B_0, and because of the strong Markov property we can reduce attention to the zero set

$$B^{-1}\{0\} := \{s \geq 0 :\ B_s = 0\}.$$

Let us begin with a simple result on the geometry of $B^{-1}\{0\}$. More advanced results will follow after we develop some notions from geometric measure theory in the next subsection.

Proposition 6.1. *With probability one, $B^{-1}\{0\}$ is closed and Lebesgue-null, as well as nowhere dense in itself.*

Proof. The first two statements are easier to prove: $B^{-1}\{0\}$ is [a.s.] closed because it is the inverse image of a closed set under a map that is [a.s.] continuous. [To be perfectly honest, this statement requires that the assumed fact that the underlying probability measure is complete.] Also, by the Tonelli theorem,

$$\mathrm{E}\left(\int_0^\infty \mathbf{1}_{B^{-1}\{0\}}(s)\,\mathrm{d}s\right) = \mathrm{E}\left(\int_0^\infty \mathbf{1}_{\{0\}}(B_s)\,\mathrm{d}s\right) = \int_0^\infty \mathrm{P}\{B_s = 0\}\,\mathrm{d}s = 0.$$

It remains to prove that $B^{-1}\{0\}$ is nowhere dense. But this follows from the fact that the complement of $B^{-1}\{0\}$ in $\mathbf{R}_+$ is a countable union of intervals a.s. [This follows from the fact that B is never a constant a.s., which holds in turn because B is a.s. nowhere differentiable.] □

6.1. Hausdorff Measure and Dimension. Let G be a subset of $\mathbf{R}_+$, and consider a real number $s \geq 0$. For all $\epsilon > 0$ define $\mathcal{H}_s^\epsilon(G) := \inf \sum_{j=1}^\infty r_j^s$, where the infimum is taken over all left-closed, right-open intervals I_j of the form $[a\,, a + r_j)$, where: (i) $\cup_{j=1}^\infty I_j \supseteq G$; and (ii) $\sup_j r_j \leq \epsilon$. The $\{I_j\}_{j=1}^\infty$'s are various covers of G. And as we decrease ϵ, we are considering fewer and fewer such covers. Therefore, the following exists as an increasing limit:

$$\mathcal{H}_s(G) := \lim_{\epsilon\downarrow 0} \mathcal{H}_s^\epsilon(G).$$

This defines a set function $\mathcal{H}_s$ on all subsets of $\mathbf{R}_+$. The following is non-trivial, but we will not prove it here.

Theorem 6.2 (Hausdorff)**.** *The restriction of $\mathcal{H}_s$ to Borel subsets of $\mathbf{R}_+$ defines a Borel measure on $\mathbf{R}_+$.*

From now on $\mathcal{H}_s$ will denote the restriction of that quantity to Borel subsets of $\mathbf{R}_+$; then, $\mathcal{H}_s$ is called the *s-dimensional Hausdorff measure*. When $s = 1$, $\mathcal{H}_s$ and the Lebesgue measure on $\mathbf{R}_+$ agree. [It suffices to check that they agree on intervals, but that is easy.] We will use, without proof, the fact that $\mathcal{H}_s(G) = \sup \mathcal{H}_s(K)$, where the supremum is taken over all compact sets $K \subseteq G$ with $\mathcal{H}_s(K) < \infty$.

If $\mathcal{H}_s(G) < \infty$ for some Borel set $G \subset \mathbf{R}_+$, then $\mathcal{H}_s^\epsilon(G) \leq 2\mathcal{H}_s(G)$ for all small $\epsilon > 0$. Hence, for all left-closed right-open cover $I_1, I_2, \ldots$ of G such that $\sup_j |I_j| \leq \epsilon$, we have $\sum_{j=1}^\infty |I_j|^s \leq 2\mathcal{H}_s(G) < \infty$. If, in addition, $\sup_j |I_j| \leq \epsilon$, then $|I_j|^{s+\delta} \leq \epsilon^\delta |I_j|^s$ for $s, \delta \geq 0$. Consequently, we can optimize to deduce that $\mathcal{H}_{s+\delta}(G) \leq 2\epsilon^\delta \mathcal{H}_s(G)$ for all $\epsilon > 0$. That is,

$$\mathcal{H}_s(G) < \infty \quad \Longrightarrow \quad \mathcal{H}_{s+\delta}(G) = 0 \text{ for all } \delta > 0.$$

Similarly, one proves that

$$\mathcal{H}_s(G) > 0 \quad \Longrightarrow \quad \mathcal{H}_{s-\delta}(G) = \infty \text{ for all } \delta > 0.$$

Therefore, there is a unique critical number $s \geq 0$ such that $\mathcal{H}_t(G) = 0$ if $t > s$ and $\mathcal{H}_t(G) = \infty$ if $t < s$. This critical s is called the *Hausdorff dimension of* G. The preceding can be encapsulated, in symbols, as follows:

$$\begin{aligned} \dim_{\mathrm{H}} G &:= \inf \{ s \geq 0 : \mathcal{H}_s(G) = 0 \} \\ &= \sup \{ s \geq 0 : \mathcal{H}_s(G) = \infty \}, \end{aligned}$$

where $\sup \varnothing := 0$. Because $\mathcal{H}_1$ is the Lebesgue measure, we find that $0 \leq \dim_{\mathrm{H}} G \leq 1$; this is clear when G is bounded, and the general case follows from approximation of Hausdorff measure of G by $\mathcal{H}_s(K)$, where $K \subseteq G$ has finite measure. And one can construct examples of sets $G \subset \mathbf{R}_+$ whose Hausdorff dimension is any given number in $[0\,, 1]$. One can prove that

$$\dim_{\mathrm{H}} \bigcup_{n=1}^\infty G_n = \sup_{n \geq 1} \dim_{\mathrm{H}} G_n.$$

The utility of this is best seen as follows: If $\dim_{\mathrm{H}}(G \cap [-n\,, n]) = s$ for all $n \geq 1$, then $\dim_{\mathrm{H}} G = s$. Our goal is to prove

Theorem 6.3 (Lévy)**.** *With probability one,* $\dim_{\mathrm{H}} B^{-1}\{0\} = 1/2$.

6.2. Upper Bounds. Suppose that there exists a sequence of ϵ's—tending downward to zero—for which we could find a left-closed right-open cover $I_1, I_2, \ldots$ of G such that $|I_j| \le \epsilon$ for all j, and $\sum_{j=1}^\infty |I_j|^s \le C < \infty$. Then it follows easily from the definition of Hausdorff measure that $\mathcal{H}_s(G) \le C < \infty$. Thus, we find that $\dim_{\mathrm{H}} G \le s$.

Example 6.4 (The Cantor set)**.** The Cantor set on $[0\,,1]$ is obtained iteratively as follows: Define $C_0 := [0\,,1]$, $C_1 := [0\,,1/3] \cup [2/3\,,1]$, $\ldots$. Having defined $C_1 \supset C_2 \supset \cdots \supset C_{n-1}$, we find that C_{n-1} is comprised of 2^{n-1} intervals of length 3^{-n+1} each. Subdivide each interval into three equal parts; jettison the middle one; and combine the rest to form C_n. Since the C_n's are decreasing closed sets in $[0,,1]$, the Heine–Borel property of the real line tells us that $C := \cap_{n=1}^\infty C_n$ is a nonvoid compact subset of $[0\,,1]$. The set C is the socalled [ternary] Cantor set. Since $C \subset C_n$ for all $n \ge 1$, the Lebesgue measure of C is at most that of C_n; the latter is equal to $2^n \cdot 3^{-n}$. Therefore, C is Lebesgue null. Moreover, the intervals that comprise C_n form a closed cover of C with length $\le \epsilon := 3^{-n}$. We can extend them a little on the right-end to obtain left-closed, right-open intervals. Then, using more or less clear notation we have $\sum_j r_j^s = 2^n \cdot 3^{-ns}$. Therefore, we can choose s to be $\log_3(2)$ to find that $\mathcal{H}_{\log_3(2)}(C) \le 1$. Consequently, $\dim_{\mathrm{H}} C \le \log_3(2)$. The inequality is in fact a famous identity due to Hausdorff. $\square$

Proof of Theorem 6.3: Upper Bound. We will prove that the Hausdorff dimension of $B^{-1}\{0\} \cap [r\,,R)$ is at most 1/2 regardless of the values of $r < R$. Elementary considerations then show from this that $\dim_{\mathrm{H}} B^{-1}\{0\} \le 1/2$ a.s. To simplify the notation, I work with the case that $[r\,,R) = [1\,,2)$. It is not too hard to adapt the argument and derive the general case from this.

Consider $I_1, \ldots, I_m$, where $I_j := [(j-1)/m\,, j/m)$ for $j = m+1, \ldots, 2m$. If $B^{-1}\{0\} \cap I_j \ne \varnothing$, then we use I_j as a cover. Else, we do not. Consequently,

$$\mathcal{H}_s^{1/m}\left(B^{-1}\{0\} \cap [1\,,2)\right) \le \sum_{j=m+1}^{2m} m^{-s} \mathbf{1}_{\{I_j \cap B^{-1}\{0\} \ne \varnothing\}}.$$

Take expectations to find that

$$\mathrm{E}\left[\mathcal{H}_s^{1/m}\left(B^{-1}\{0\} \cap [1\,,2)\right)\right] \le \sum_{j=m+1}^{2m} m^{-s} \mathrm{P}\left\{I_j \cap B^{-1}\{0\} \ne \varnothing\right\}.$$

The probabilities are not hard to find. Indeed, by the Markov property,

$$\mathrm{P}\left\{I_j \cap B^{-1}\{0\} \ne \varnothing\right\} = \int_{-\infty}^{\infty} \mathrm{P}\left\{B_{(j-1)/m} \in \mathrm{d}x\right\} \cdot \mathrm{P}\left\{T_0 \le \frac{1}{m} \,\middle|\, B_0 = x\right\},$$

where T_0 denotes the first hitting time of zero. The probability density under the integral is easy to write down because $B_{(j-1)/m} = N(0\,,(j-1)/m)$. Thus, by symmetry,

$$\mathrm{P}\left\{I_j \cap B^{-1}\{0\} \neq \varnothing\right\} = 2\int_0^\infty f_m(x)\mathrm{P}\left\{T_0 \leq \frac{1}{m}\,\middle|\, B_0 = x\right\}\,\mathrm{d}x,$$

where

$$f_m(x) := \sqrt{\frac{m}{2\pi(j-1)}}\exp\left(-\frac{x^2 m}{2(j-1)}\right).$$

Over the range of j's, $m/(j-1) \leq 1$. Therefore, $f_m(x) \leq (2\pi)^{-1/2}$, and hence

$$\mathrm{P}\left\{I_j \cap B^{-1}\{0\} \neq \varnothing\right\} \leq \left(\frac{2}{\pi}\right)^{1/2}\int_0^\infty \mathrm{P}\left\{T_0 \leq \frac{1}{m}\,\middle|\, B_0 = x\right\}\,\mathrm{d}x,$$

The conditional probability that $T_0 \leq 1/m$ given $B_0 = x > 0$ is the same as the conditional probability that the probability that Brownian motion, started at zero, exceeds x by time $1/m$. By the reflection principle,

$$\mathrm{P}\left\{T_0 \leq \frac{1}{m}\,\middle|\, B_0 = x\right\} = \mathrm{P}\left\{\sup_{s\in[0,1/m]} B_s > x\right\} = \mathrm{P}\left\{|B_{1/m}| > x\right\}.$$

Therefore, $\mathrm{P}\{I_j \cap B^{-1}\{0\} \neq \varnothing\} \leq \text{const}\cdot \mathrm{E}(|B_{1/m}|) = \text{const}\cdot m^{-1/2}$, and hence

$$\mathrm{E}\left[\mathcal{H}_s^{1/m}\left(B^{-1}\{0\}\cap[1\,,2)\right)\right] \leq \text{const}\cdot\sum_{j=m+1}^{2m} m^{-s-(1/2)} \leq \text{const}\cdot m^{-s+(1/2)}.$$

In particular, we set $s := 1/2$ to find that

$$\sup_{m\geq 1}\mathrm{E}\left[\mathcal{H}_{1/2}^{1/m}\left(B^{-1}\{0\}\cap[1\,,2)\right)\right] < \infty.$$

By Fatou's lemma, the expectation of $\mathcal{H}_{1/2}(B^{-1}\{0\}\cap[1\,,2))$ is finite. This shows that $\mathcal{H}_{1/2}(B^{-1}\{0\}\cap[1\,,2))$ is finite a.s., and thus we obtain the desired bound, $\dim_{\mathrm{H}} B^{-1}\{0\}\cap[1\,,2) \leq 1/2$ a.s. □

6.3. Lower Bounds. The method of the previous section does not have a lower-bound counterpart because in the case of the lower bound we need to consider all possible coverings [and not just hand-picked ones]. Thus, we need a new idea for this second half. The surprising answer comes from another part of mathematics (potential theory) altogether.

Proposition 6.5 (Frostman)**.** *If G be a measurable set in $\mathbf{R}_+$, then $\dim_{\text{H}} G$ is at least $s > 0$, provided that there exists a nontrivial finite measure μ on G [i.e., $\mu(G^c) = 0$] such that*

$$c := \sup_{x \in \mathbf{R}} \sup_{\epsilon > 0} \frac{\mu([x\,, x+\epsilon])}{\epsilon^s} < \infty.$$

Proof. For all $\epsilon > 0$ and all left-closed, right-open covers $I_1, I_2, \ldots$ of G such that $\sup_j |I_j| \leq \epsilon$, and $\sum_{j=1}^\infty |I_j|^s \geq c^{-1} \sum_{j=1}^\infty \mu(I_j) = c^{-1}$. We can take the infimum over all such covers and let ϵ tend to zero to find that $\mathcal{H}_s(G) \geq c^{-1} > 0$. □

Proof of Theorem 6.3: Lower Bound. Define μ by

$$\mu([a\,, b)) := L^0_b(B) - L^0_a(B),$$

where $L^0_t(B)$ is Brownian local time at zero. Thanks to Tanaka's formula (Theorem 5.1),

$$\mu([a\,, b]) \leq |B_b - B_a| + |\beta_b - \beta_a|,$$

where $\beta_t := \int_0^t \operatorname{sign}(B_s)\,\mathrm{d}B_s$. The process β is a Brownian motion, as can be seen from its quadratic variation. Since B and β are Hölder continuous of any order $\alpha < 1/2$, it follows that

$$\mathrm{P}\left\{\sup_{x \in \mathbf{R}} \sup_{\epsilon > 0} \frac{\mu([x\,, x+\epsilon])}{\epsilon^\alpha} < \infty\right\} = 1 \quad \text{for all } \alpha < 1/2.$$

Therefore, Frostman's Proposition 6.5 would complete the proof provided that we could prove that $\mu([0\,, T])$ is positive and finite a.s. for some constant $T > 0$. It turns out that this statement is true for all $T > 0$.

By Tanaka's formula, $\mathrm{E}\,\mu([0\,, T]) = \mathrm{E}(|B_T|) < \infty$. Therefore, $\mu([0\,, T]) < \infty$ a.s. for all $T > 0$. For positivity we first recall that by the occupation density formula and the continuity of local times,

$$\mu([0\,, T]) = L^0_T(B) = \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} \int_0^T \mathbf{1}_{\{0 < B_s < \epsilon\}}\,\mathrm{d}s.$$

Therefore, scaling properties of B imply that $\mu([0\,, T])$ has the same distribution as $T^{1/2}\mu([0\,, 1])$. Thus,

$$\mathrm{P}\left\{\mu([0\,, T]) > 0\right\} = \mathrm{P}\left\{\mu([0\,, 1]) > 0\right\}.$$

Let $T \downarrow 0$ and appeal to the Blumenthal zero-one law to deduce that the left-hand side is zero or one. Since $\mathrm{E}\,\mu([0\,, 1]) = \mathrm{E}|B_1| > 0$, it follows that $\mathrm{P}\{\mu([0\,, T]) > 0\} = 1$ for all $T > 0$. This completes the proof. □

Exercise 6.6. Prove that $\dim_{\mathrm{H}} C \geq \log_3(2)$ for the Cantor set C by constructing an appropriate Frostman measure μ on C. (Hint: Stat by studying carefully the distribution function of the uniform distribution on C_n.)

Chapter 4

One-Dimensional Stochastic Differential Equations

1. Fundamental Notions

2. Equations with Lipschitz Coefficients

Suppose σ and b are real-valued Lipschitz continuous functions on $\mathbf{R}$. We wish to solve the following for a continuous and adapted process X, where x is nonrandom:

$$\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t \qquad \text{subject to } X_0 = x. \tag{2.1}$$

[The same procedure works if x is independent of B, but then we will require X to be continuous and adapted with respect to $\{\sigma(x) \vee \mathscr{F}_t\}$.]

Interpretation:

$$X_t = x + \int_0^t \sigma(X_s)\,\mathrm{d}B_s. \tag{2.2}$$

We proceed with a fixed-point argument.

Define $X_t^{(0)} := x$, and then iteratively define

$$X_t^{(n)} := x + \int_0^t \sigma(X_s^{(n-1)})\,\mathrm{d}B_s, \tag{2.3}$$

valid for all $n \geq 1$. The following will help us control the size of $X_t^{(n)}$.

Lemma 2.1 (Gronwall's lemma)**.** *Let $f_1, f_2, \ldots$ be a sequence of nondecreasing real-valued functions on $[0\,,T]$. Suppose there exist $\alpha, \beta > 0$ such that for all $t \in [0\,,T]$ and $n \geq 2$,*

$$f_n(t) \leq \alpha + \beta \int_0^t f_{n-1}(s)\,\mathrm{d}s. \tag{2.4}$$

Then,

$$f_n(t) \leq \alpha e^{\beta t} + \frac{(1+\beta)^n(1+t)^n}{n!} f_1(t) \qquad \text{for all } t \in [0\,,T]. \tag{2.5}$$

Proof. Clearly, (2.5) is valid for $n = 1$. The lemma is then proved by applying a simple induction argument. □

Theorem 2.2 (Itô)**.** *Suppose σ is Lipschitz continuous. Then,* (2.1) *has a unique adapted solution that is a.s. continuous.*

Proof. Because σ is Lipschitz, $|\sigma(z)| \leq |\sigma(0)| + \mathrm{Lip}_\sigma |z|$. Consequently,

$$\mathrm{E}\left(\left|X_t^{(1)} - x\right|^2\right) = |\sigma(x)|^2 t \leq 2t\left\{|\sigma(0)|^2 + x^2 \mathrm{Lip}_\sigma^2\right\}.$$

We have appealed to the inequality $(a+b)^2 \leq 2(a^2+b^2)$. This and Doob's inequality together prove that $\sup_{s \leq t} |X_s^{(1)}|$ is square integrable for all $t \geq 0$.

By Doob's maximal inequality,

$$\mathrm{E}\left(\sup_{0\leq s\leq t} |X_s^{(n)} - X_s^{(n-1)}|^2\right) \leq 4\mathrm{E}\left(\int_0^t \left|\sigma(X_s^{(n-1)}) - \sigma(X_s^{(n-2)})\right|^2 \mathrm{d}s\right).$$

Let $f_n(t)$ denote the left-hand side. Since σ is Lipschitz continuous, we have shown that $f_n(t) \leq 4\mathrm{Lip}_\sigma^2 \int_0^t f_{n-1}(s)\,\mathrm{d}s$. Gronwall's lemma then implies that

$$f_n(t) \leq \frac{(1+4\mathrm{Lip}_\sigma^2)^n(1+t)^n}{n!} f_1(t) \leq c f_1(t) 2^{-n}, \tag{2.6}$$

for some constant $c > 0$. [The last bound uses Stirling's formula, as well.] This makes sense provided that $f_1(t) < \infty$ for all t. But our choice of $X^{(0)}$ guarantess that $X_t^{(1)} = x + \sigma(x)B_t$, whence $f_1(t) < \infty$.

The preceding display tell us that $\sum_{n=1}^\infty |f_n(t)|^{1/2} < \infty$, and hence it follows that there exists a process $\{X_s\}_{s\geq 0}$ such that for all $t \geq 0$,

$$\sup_{0\leq s\leq t}\left|X_s^{(n)} - X_s\right| \xrightarrow{n\uparrow\infty} 0 \qquad \text{a.s. and in } L^2(\mathrm{P}). \tag{2.7}$$

Evidently, X is adapted, and

$$\begin{aligned}
&\mathrm{E}\left(\sup_{0\le s\le t}\left|\int_0^s \sigma(X_u^{(n)})\,\mathrm{d}B_u - \int_0^s \sigma(X_u)\,\mathrm{d}B_u\right|^2\right)\\
&\qquad\le 4\mathrm{E}\left(\int_0^t \left|\sigma(X_u^{(n)}) - \sigma(X_u)\right|^2\,\mathrm{d}u\right)\\
&\qquad\le 4\mathrm{Lip}_\sigma^2\; t\;\mathrm{E}\left(\sup_{0\le s\le t}\left|X_s^{(n)} - X_s\right|^2\right).
\end{aligned}\tag{2.8}$$

By (2.6) and Minkowski's inequality,

$$\begin{aligned}
\left\{\mathrm{E}\left(\sup_{0\le s\le t}\left|X_s^{(n)} - X_s\right|^2\right)\right\}^{1/2} &\le \sum_{j=n}^{\infty}|f_j(t)|^{1/2}\\
&\le \mathrm{const}\cdot|f_1(t)|^{1/2}2^{-n/2}.
\end{aligned}\tag{2.9}$$

This defines a summable series [in n]. Therefore,

$$\sup_{0\le s\le t}\left|\int_0^s \sigma(X_u^{(n)})\,\mathrm{d}B_u - \int_0^s \sigma(X_u)\,\mathrm{d}B_u\right| \stackrel{n\uparrow\infty}{\longrightarrow} 0 \quad \text{a.s. and in } L^2(\mathrm{P}).\tag{2.10}$$

Therefore, (2.3) implies that (2.1) has a continuous adapted solution X. To finish, suppose (2.1) has two continuous adapted solutions X and Y. Then,

$$\begin{aligned}
\mathrm{E}\left(\sup_{0\le s\le t}|X_s - Y_s|^2\right) &\le 4\mathrm{E}\left(\int_0^t |\sigma(X_u) - \sigma(Y_u)|^2\,\mathrm{d}u\right)\\
&\le 4\mathrm{Lip}_\sigma^2\mathrm{E}\left(\int_0^t |X_u - Y_u|^2\,\mathrm{d}u\right)
\end{aligned}\tag{2.11}$$

Call the left-hand side $f(t)$ to deduce that $f(t) \le \mathrm{const}\cdot\int_0^t f(s)\,\mathrm{d}s$. Thanks to Gronwall's lemma, $f(t) \le c_t^n/n!\cdot f(t)$ for a constant c_t and arbitrary integers $n \ge 1$. Let $n \to \infty$ to deduce that $f(t) = 0$ for all t. □

Exercise 2.3. Prove directly that $\sup_n \mathrm{E}(\sup_{s\le t}|X_s^{(n)}|^2) < \infty$ for all $t > 0$.

Exercise 2.4 (Important)**.** Let B denote d-dimensional Brownian motion. Suppose $b : \mathbf{R}^k \to \mathbf{R}^d$ and $\sigma : \mathbf{R}^k \to \mathbf{R}^{k\times d}$ are nonrandom Lipschitz functions. Prove that there exists a unique adapted and continuous solution to the SDE $\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t + b(X_t)\,\mathrm{d}t$, subject to $X_0 = x$ for a nonrandom $x \in \mathbf{R}^k$. This equation is interpreted as follows:

$$X_t = x + \int_0^t \sigma(X_s)\,\mathrm{d}B_s + \int_0^t b(X_s)\,\mathrm{d}s.\tag{2.12}$$

Exercise 2.5. Interpret $\mathrm{d}X_t = X_t^\alpha\,\mathrm{d}t$, subject to $X_0 = 0$, as a trivial one-dimensional SDE. For what values of α does this have a solution for all $t > 0$? When is the solution unique? This exercise helps you see why the assumption of Lipschitz continuity is important. [Hint: Try solutions that have are powers of $\nu(t-c)^+$ for various positive constants c and ν.]

3. Lévy's Identity in Law

Let B denote linear Brownian motion with $B_0 := 0$, and consider the non-decreasing continuous adapted process S defined by

$$S_t := \sup_{s\in[0,t]} B_s \quad \text{for all } t \geq 0.$$

Theorem 3.1 (Lévy). *As processes, $(S_t\,, S_t - B_t)$ and $(L_t^0(B)\,, |B_t|)$ have the same finite-dimensional distributions.*

Remark 3.2. Choose and fix a $t \geq 0$. The classical reflection principle tells us that the random variables S_t and $|B_t|$ have the same distibution. Because of this and Theorem 3.1, we can conclude that $L_t^0(B)$ and $|B_t|$ have the same distribution as well. That is,

$$\mathrm{P}\left\{L_t^0(B) \leq x\right\} = \left(\frac{2}{\pi t}\right)^{1/2} \int_0^x e^{-y^2/(2t)}\,\mathrm{d}y.$$

But of course Theorem 3.1 implies much more than this sort of statement about fixed times. □

Proof. Let $g : \mathbf{R}_+ \times \mathbf{R} \to \mathbf{R}$ be C^∞ in each variable (say), and consider the SDE

$$\begin{aligned} g(V_t\,, X_t) - g(0\,,0) &= -\int_0^t (\partial_x g)(V_s\,, X_s)\,\mathrm{d}\beta_s \\ &\quad + \int_0^t [(\partial_t g)(V_s\,, X_s) + (\partial_x g)(V_s\,, X_s)]\,\mathrm{d}V_s \\ &\quad + \frac{1}{2}\int_0^t (\partial_{xx} g)(V_s\,, X_s)\,\mathrm{d}s, \end{aligned} \tag{3.1}$$

where β is a Brownian motion, V is a nondecreasing adapted process, and X is an adapted and continuous semimartingale. Because g is smooth in all of its variables, this SDE can be shown to have a unique strong solution—which we write as $(V\,, X\,, \beta)$—and we can write $(V\,, X) = F(\beta)$ for a nonrandom function F that is obtained from the fixed-point proof of the basic theorem on SDEs.

Next we can apply Itô's formula to find that for all smooth functions $f : \mathbf{R}_+ \times \mathbf{R} \to \mathbf{R}$,

$$\begin{aligned} &f(S_t\,, B_t) - f(0\,,0) \\ &= \int_0^t (\partial_x f)(S_s\,, B_s)\,\mathrm{d}B_s + \int_0^t (\partial_t f)(S_s\,, B_s)\,\mathrm{d}S_s + \frac{1}{2}\int_0^t (\partial_{xx} f)(S_s\,, B_s)\,\mathrm{d}s. \end{aligned}$$

Define $f(t\,,x) := g(t\,,t-x)$ and $D_t := S_t - B_t$ to infer from this that $(S\,,D\,,B)$ solves (3.1), and consequently, $(S\,,D) = F(B)$.

Let $\ell_t := L_t^0(B)$ and note from Tanaka's formula (5.2) on page 39 that

$$\begin{aligned} &g(\ell_t\,, |B_t|) - g(0\,,0) \\ &= \int_0^t (\partial_x g)(\ell_s\,, |B_s|)\mathrm{sign}(B_s)\,\mathrm{d}B_s + \int_0^t \left[(\partial_t g)(\ell_s\,, |B_s|) + (\partial_x g)(\ell_s\,, |B_s|)\right]\,\mathrm{d}\ell_s \\ &\qquad + \frac{1}{2}\int_0^t (\partial_{xx} g)(\ell_s\,, |B_s|)\,\mathrm{d}s. \end{aligned}$$

Now consider the local martingale $W_t := \int_0^t \mathrm{sign}(B_s)\,\mathrm{d}B_s$. Thanks to Lemma 5.8 (p. 42), $\langle W\rangle_t = t$ and therefore W is a Brownian motion by Lévy's theorem (p. 33). This proves that $(\ell\,, |B|\,, W)$ solves (3.1), and consequently, $(\ell\,, |B|) = F(W)$. The theorem follows because the laws of $F(W)$ and $F(B)$ are the same. □

4. Girsanov's Counterexample

Suppose $\alpha > 0$ is a constant, and consider the equation

$$\mathrm{d}X_t = |X_t|^\alpha\,\mathrm{d}B_t, \qquad \text{subject to } X_0 = 0. \tag{4.1}$$

Obviously $X_t := 0$ is always a strong solution. The following shows that if we choose α suitably, then there also exist nonzero solutions.

Theorem 4.1 (Girsanov)**.** *If $\alpha \in (0\,,\frac{1}{2})$, then* (4.1) *does not have a weakly-unique solution. But if $\alpha \geq 1/2$ then* (4.1) *has an a.s.-unique strong solution.*

Note that the function $\sigma(x) := |x|^\alpha$ is not Lipschitz continuous for any $\alpha \in (0\,,1)$.

Proof. Consider the process

$$\beta_t := \int_0^t \frac{\mathrm{d}B_s}{|B_s|^\alpha}.$$

This is a well-defined continuous local martingale, since $\int_0^t \mathrm{d}s/|B_s|^{2\alpha}$ is a.s. finite. Indeed,

$$\mathrm{E}\left(\int_0^t \frac{\mathrm{d}s}{|B_s|^{2\alpha}}\right) = \mathrm{E}\left(|N(0\,,1)|^{-2\alpha}\right)\cdot\int_0^t \frac{\mathrm{d}s}{s^\alpha} < \infty.$$

Moreover, the quadratic variation of β is

$$\langle\beta\rangle_t = \int_0^t \frac{\mathrm{d}s}{|B_s|^{2\alpha}}.$$

I will show soon that $\lim_{t\to\infty}\langle\beta\rangle_t = \infty$ a.s. Let us use this to first complete the proof of the theorem. Then we verify the latter claim to complete the proof of nonuniqueness.

The theorem of Dambis, Dubins, and Schwarz (p. 33) implies that we can find a Brownian motion W such that $\beta_{\rho(t)} = W_t$, where $\rho(t)$ denotes the first instance $r > 0$ when $\langle\beta\rangle_r = t$. Because $\mathrm{d}\beta_t = |B_t|^{-2\alpha}\,\mathrm{d}B_t$, it follows that $\mathrm{d}B_t = |B_t|^{2\alpha}\,\mathrm{d}\beta_t = |B_t|^{2\alpha}\,\mathrm{d}(W\circ\langle\beta\rangle)_t$, and therefore $Y_t := B_{\rho(t)}$ solves $\mathrm{d}Y_t = |Y_t|^{2\alpha}\,\mathrm{d}W_t$ subject to $Y_0 = 0$. Because $t\mapsto\langle\beta\rangle_t$ is strictly increasing, and since the zero-set of B has zero Lebesgue measure, it follows that $Y \not\equiv 0$. Therefore, there are non-zero weak solutions to (4.1).

Now we prove the claim that $\langle\beta\rangle_\infty = \infty$ a.s. Fix $t > 0$ and note that $\langle\beta\rangle_t = t\int_0^1 |B_{ut}|^{-2\alpha}\,\mathrm{d}u$, and by the Brownian scaling property this has the same distribution as $t^{1-\alpha}\langle\beta\rangle_1$. Consequently,

$$\mathrm{P}\left\{\langle\beta\rangle_t \geq \lambda\right\} = \mathrm{P}\left\{\langle\beta\rangle_1 \geq \frac{\lambda}{t^{1-\alpha}}\right\} \to 1 \quad \text{as } t\to\infty,$$

for all $\lambda > 0$. That is, $\langle\beta\rangle_t \to \infty$ in probability, whence it follows that $\limsup_{t\to\infty}\langle\beta\rangle_t = \infty$ a.s. Since $\langle\beta\rangle$ is nondecreasing, this proves that $\langle\beta\rangle_\infty = \infty$ almost surely. The nonuniqueness portion follows.

For the other half, $X_t := 0$ is always a strong soltuion. We wish to prove that it is the only one. Suppose there is another strong solutions Z. By localization, we can assume without loss of generality that $|Z_t|$ is uniformly bounded by a constant; else, we obtain uniqueness up to $\inf\{s > 0 : |Z_s| > k\}$ for all $k \geq 1$, and hence the result in the general case.

We know that Z_t is a continuous local martingale with $\langle Z\rangle_t = \int_0^t |Z_s|^{2\alpha}\,\mathrm{d}s$. Thus, $\int_0^t |Z_s|^{-2\alpha}\mathrm{d}\langle Z\rangle_s = t < \infty$. According to Tanaka's formula,

$$|Z_t| = \int_0^t \operatorname{sign}(Z_s)\,\mathrm{d}Z_s + L_t^0, \tag{4.2}$$

where $\{L_t^a\}_{t\geq 0}$ denotes the local time of Z at a. The occupation density formula (p. 41) tells us that $\int_0^t |Z_s|^{-2\alpha}\mathrm{d}\langle Z\rangle_s = \int_{-\infty}^{\infty} |a|^{-2\alpha} L_t^a\,\mathrm{d}a$, which we know is finite a.s. Since $a\mapsto |a|^{-2\alpha}$ is not integrable, the continuity of local times implies that $L_t^0 \equiv 0$. Consequently, $|Z_t| = \int_0^t \operatorname{sign}(Z_s)|Z_s|^\alpha\,\mathrm{d}B_s$ defines a mean-zero square-integrable martingale. Thus, $\mathrm{E}(|Z_t|) = 0$, and uniqueness follows. □

Exercise 4.2. Suppose b is a Lipschitz function and there exists $C > 0$ and $\alpha \geq 1/2$ such that

$$|\sigma(w) - \sigma(z)| \leq C|w - z|^{\alpha}.$$

Then prove that if $\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t + b(X_t)\,\mathrm{d}t$ ($X_0 = x$) has a weak solution, then there is pathwise uniqueness, whence there is a unique strong solution. [Corollary 5.4 below implies that there is often a weak solution. Other results of this type also exist.]

5. Weak Solutions and the Martingale Problem

Suppose X is a continuous local semimartingale that satisfies

$$X_t = x + \int_0^t \sigma(X_s)\,\mathrm{d}B_s + \int_0^t b(X_s)\,\mathrm{d}s, \tag{5.1}$$

where $b, \sigma : \mathbf{R} \to \mathbf{R}$ are measurable functions, and $x \in \mathbf{R}$ is nonrandom and fixed. Such processes are called *Itô processes*. [We are implicitly assuming that both integrals make sense, and B is of course linear Brownian motion.] Define for all $f \in C^2(\mathbf{R})$,

$$M_t^f := f(X_t) - f(X_0) - \int_0^t (Lf)(X_s)\ ds, \tag{5.2}$$

where

$$(Lf)(z) := \frac{1}{2}\sigma^2(z) f''(z) + b(z) f'(z). \tag{5.3}$$

Then according to Itô's formula, M^f is a continuous local martingale. In fact, $M_t^f = \int_0^t \sigma(X_s) f'(X_s)\,\mathrm{d}B_s$. There is a kind of converse to this, which we begin to prove.

Lemma 5.1. *Suppose X is a continuous adapted $[\mathscr{F}]$ process. Then, M^f is a continuous local martingale for all $f \in C^2(\mathbf{R})$ if and only if $Z_t := X_t - \int_0^t b(X_s)\,ds$ defines a continuous local martingale with $\langle Z \rangle_t = \int_0^t \sigma^2(X_s)\,\mathrm{d}s$.*

Proof. First we suppose M^f is a continuous local martingale for all $f \in C^2(\mathbf{R})$. Choose $f(z) := z$, so that $(Lf)(z) := 0$, to find that Z is a continuous local martingale, and hence X is a continuous local semimartingale.

We choose $f(z) := z^2$, so that $(Lf)(z) = \sigma^2(z) + 2zb(z)$ and

$$X_t^2 = \int_0^t \sigma^2(X_s)\,\mathrm{d}s + 2\int_0^t X\,\mathrm{d}C + \text{ a local martingale,}$$

where $C_t := \int_0^t b(X_s)\,\mathrm{d}s$. By Itô's formula for semimartingales,

$$2X_tC_t = 2\int_0^t X\,\mathrm{d}C + 2\int_0^t C\,\mathrm{d}X = 2\int_0^t X\,\mathrm{d}C + 2\int_0^t C\,\mathrm{d}C + \text{ a loc. mart.}$$

In fact, the local martingale is $2\int_0^t C\,\mathrm{d}Z$. We take the difference of the previous two displays to find

$$X_t^2 - 2X_tC_t = \int_0^t \sigma^2(X_s)\,\mathrm{d}s - 2\int_0^t C\,\mathrm{d}C + \text{ a local martingale.}$$

But classical integration by parts tells us that $2\int_0^t C\,\mathrm{d}C = C_t^2$. Consequently, $Z_t^2 = \int_0^t \sigma^2(X_s)\,\mathrm{d}s +$ a local martingale, and this identifies $\langle Z\rangle$.

Conversely, suppose $Z_t := X_t - C_t$ is a continuous local martingale with quadratic variation $\int_0^t \sigma^2(X_s)\,\mathrm{d}s$. According to Itô's formula, $\mathscr{E}_t^\lambda := \exp\{\lambda Z_t - \frac{1}{2}\lambda^2\langle Z\rangle_t\}$ defines a continuous local martingale for all fixed $\lambda \in \mathbf{R}$. This proves that M_t^f is a continuous local martingale for $f(z) = \exp(\lambda z)$. Because exponentials are dense in $C^2(\mathbf{R})$ the result follows easily from the exponential case. □

Lemma 5.2. *Suppose X is a continuous adapted $[\mathscr{F}]$ process. such that $Z_t := X_t - \int_0^t b(X_s)\,ds$ defines a continuous local martingale with $\langle Z\rangle_t = \int_0^t \sigma^2(X_s)\,\mathrm{d}s$. Then we can enlarge the probability space such that on the enlarged space there exists a Brownian motion B that satisfies* (5.1).

Proof. First consider the case that $\sigma(z)$ is never zero. In that case, we can define $B_t := \int_0^t \mathrm{d}Z_s/\sigma(X_s)$. This is a well-defined Itô integral and $\langle B\rangle_t = t$. Therefore, B is a Brownian motion by Lévy's theorem (p. 33). Clearly, $\mathrm{d}B_t = \mathrm{d}Z_t/\sigma(X_t)$—equivalently, $\mathrm{d}Z_t = \sigma(X_t)\,\mathrm{d}B_t$, and no enlargement of the space is necessary.

In the case that σ can take on the value of zero, we introduce an independent Brownian motion $\{\beta_t\}_{t\geq 0}$ by enlarging the underlying space—if needed—and define

$$B_t := \int_0^t \frac{\mathbf{1}_{\{\sigma(X_s)\neq 0\}}}{\sigma(X_s)}\,\mathrm{d}Z_s + \int_0^t \mathbf{1}_{\{\sigma(X_s)=0\}}\,\mathrm{d}\beta_s,$$

where $0/0 := 0$. Then, B is a continuous local martingale with respect to the filtration generated by $(Z\,,\beta)$, and $\langle B\rangle_t = t$, whence follows the Brownian property of B. The fact that $\sigma(X_t)\,\mathrm{d}B_t = \mathrm{d}Z_t$ is obvious from the definitions. □

Now we use our theory of Itô processes, and construct weak solution to the stochastic differential equation (5.1) under very mild regularity conditions on σ and b.

Theorem 5.3. *If $\sigma, b : \mathbf{R} \to \mathbf{R}$ are bounded and continuous, then* (5.1) *has a weak solution.*

Proof. We need to prove only that there exists a continuous process X [adapted to a suitable filtration] such that M^f is a continuous local martingale [in that filtration] for all $f \in C^2(\mathbf{R})$.

If σ and b are Lipschitz continuous, then we know that there is in fact a strong solution X for every Brownian motion B (Theorem 2.2 and Exercise 2.4, pp. 52–53). If σ and b are bounded and continuous, then we choose a sequence of Lipschitz continuous functions $\{\sigma_n\}$ and $\{b_n\}$ such that $\sigma_n \to \sigma$ and $b_n \to b$. Because there exists a constant C such that $|b(z)| + |\sigma(z)| \leq C$ for all $z \in \mathbf{R}$, we can even choose the σ_n's and b_n's so that $|\sigma_n(z)| + |b_n(z)| \leq C$ for all $n \geq 1$ and $z \in \mathbf{R}$.

Define operators $L_1, L_2, \ldots$ by

$$(L_n f)(z) := \frac{1}{2}\sigma_n^2(z) f''(z) + b_n(z) f'(z).$$

We know already that there exist processes $X^1, X^2, \ldots$ such that for all $n \geq 1$ and $f \in C^2(\mathbf{R})$,

$$M_t^{f,n} := f(X_t^n) - f(x) - \int_0^t (L_n f)(X_s^n)\,\mathrm{d}s \quad \text{is a continuous loc. mart.}$$

In fact, we can choose a single Brownian motion B and define X^n as the [strong] solution to

$$X_t^n = x + \int_0^t \sigma_n(X_s^n)\,\mathrm{d}B_s + \int_0^t b_n(X_s^n)\,\mathrm{d}s.$$

Because σ_n^2 is bounded uniformly by C^2, Exercise 5.12 on page 45 tells us that

$$\mathrm{E}\left(|Y_t^n - Y_s^n|^4\right) \leq \text{const} \cdot |t-s|^2,$$

where the constant does not depend on n, t, or s, and $Y_t^n := X_t^n - \int_0^t b_n(X_s^n)\,\mathrm{d}s$. Kolmogorov's continuity theorem (p. 5) ensures that for all $T > 0$ there exists a constant $C_T \in (0\,,\infty)$ such that

$$\mathrm{E}\left(\sup_{\substack{0 \leq s,t \leq T:\\ |s-t| \leq \epsilon}} \frac{|Y_s^n - Y_t^n|}{|t-s|^{1/5}}\right) \leq C_T.$$

Since b_n is bounded this implies that

$$\mathrm{E}\left(\sup_{\substack{0 \leq s,t \leq T:\\ |s-t| \leq \epsilon}} \frac{|X_s^n - X_t^n|}{|t-s|^{1/5}}\right) \leq C_T'.$$

for a possibly-larger constant $C_T' \in (0\,,\infty)$. Let $\mathcal{F}_T(m)$ denote the collection of all continuous functions $f : [0\,,T] \to \mathbf{R}$ such that $f(0) = x$ and $|f(t) -$

$f(s)| \leq m|t-s|^{1/5}$ for all $s, t \in [0\,, T]$. Thanks to Arzéla–Ascoli theorem $\mathcal{F}_T(m)$ is compact in $C([0\,,T])$. By Chebyshev's inequality,

$$\sup_{n\geq 1} \mathrm{P}\left\{ X^n\big|_{[0,T]} \notin \mathcal{F}_T(m) \right\} \leq \frac{C'_T}{m}.$$

Let μ_n denote the distribution of X^n, restricted to the time-set $[0\,,T]$. The preceding proves that for all $\epsilon > 0$ there exists a compact $K_\epsilon \subseteq C([0\,,T])$, such that $\mu_n(K_\epsilon) \geq 1-\epsilon$ for all $n \geq 1$. That is, $\{\mu_n\}_{n=1}^\infty$ is a "tight" family. If all μ_n's supported the same compact set K_0 (say) then the Banach–Alaoglu theorem implies that there exists a subsequence $\{n(k)\}_{k=1}^\infty$—going to infinity as $k \to \infty$—such that $\mu_{n(k)}$ converges weakly to some probability measure μ_∞ supported on K_0.[1] Tightness is only a small complication and the same fact remains true (Prohorov's theorem), as can be seen by considering $\mu_n(\,\cdot\,|K_\epsilon)$ instead of μ_n. Let X denote the process in $C([0\,,T])$ whose law is μ_∞.

For all $f \in C^2(\mathbf{R})$, $\lim_{n\to\infty} L_n f = Lf$ pointwise and locally boundedly. Consequently, $M^{f,n}$ converges weakly to the continuous local martingale M^f for all $f \in C^2(\mathbf{R})$. Lemma 5.2 concludes the proof, and shows that there exists a Brownian motion B—on a possibly enlarged probability space—such that $(X\,,B)$ is a weak solution to (5.1). □

Corollary 5.4. *Suppose b is Lipschitz continuous and σ is Hölder-continuous with exponent $\alpha \geq 1/2$. Then* (5.1) *has a unique strong solution.*

Proof. Let $b_n(z) := b(z) \vee (-n) \wedge n$ and $\sigma_n(z) := \sigma(z) \vee (-n) \wedge n$. Then b_n and σ_n are bounded and continuous. Therefore, Theorem 5.3 assures us of the existence of a weak solution (X^n, B^n) to

$$\text{(5.4)} \qquad \mathrm{d}X_t = \sigma_n(X_t)\,\mathrm{d}B_t + b_n(X_t)\,\mathrm{d}t \quad \text{subject to } X_0 = x.$$

Because b_n is Lipschitz and σ_n is Hölder-continuous with index $\geq 1/2$, Exercise 4.2 (p. 57) assures us that there is in fact a unique strong solution to (5.4). Let X denote the solution and B the Brownian motion. Define T_n to be the first instance $s \geq 0$ at which $|\sigma(X_s)| + |b(X_s)| > n$. Clearly, $\sigma_n(X_s) = \sigma(X_s)$ and $b_n(X_s) = b(X_s)$ for all $s \leq T_n$. Therefore, strong uniqueness completes the proof. □

[1] Recall that this means that $\lim_n \int g \,\mathrm{d}\mu_{n(k)} = \int g\,\mathrm{d}\mu_\infty$ for all continuous and bounded functions $g : C([0\,,T]) \to \mathbf{R}$.

6. Some Interesting Examples

6.1. The Exponential Martingale. Let X denote the unique strong solution to the stochastic differential equation

$$\mathrm{d}X_t = \lambda X_t \,\mathrm{d}B_t \qquad \text{subject to } X_0 = x,$$

where $x > 0$ and $\lambda \in \mathbf{R}$ are constants. Then,

$$X_t = x \exp\left(\lambda B_t - \frac{\lambda^2}{2} t\right).$$

See the discussion on page 33.

6.2. The Ornstein–Uhlenbeck Process. Let X denote the unique strong solution to the *Langevin equation*

$$\mathrm{d}X_t = \mathrm{d}B_t - \nu X_t \,\mathrm{d}t \qquad \text{subject to } X_0 = x, \tag{6.1}$$

where $x, \nu \in \mathbf{R}$ are constants and B denotes a linear Brownian motion with $B_0 := 0$. Let f denote a C^1 nonrandom function. Then, we multiply both sides of (6.1) by f and "integrate" to find that

$$\int_0^t f(s)\,\mathrm{d}X_t = \int_0^t f(s)\,\mathrm{d}B_s - \nu \int_0^t f(s) X_s \,\mathrm{d}s.$$

(Why is this correct?) Because f is nonrandom and C^1, $\int_0^t f\,\mathrm{d}X = f(t)X_t - f(0)x - \int_0^t f'(s)X_s\,\mathrm{d}s$. (E.g., see Wiener's formula in Example 1.3, p. 17.) Consequently,

$$f(t)X_t - \int_0^t \left[f'(s) - \nu f(s)\right] X_s \,\mathrm{d}s = f(0)x + \int_0^t f\,\mathrm{d}B.$$

We choose any nonrandom C^1 function f that satisfies $f' = \nu f$; e.g., $f(t) := \exp(\nu t)$. This choice ensures that the first integral vanishes, and hence

$$X_t = e^{-\nu t}x + e^{-\nu t}\int_0^t e^{\nu s}\,\mathrm{d}B_s \quad \text{is the solution to (6.1).} \tag{6.2}$$

Exercise 6.1. Verify that the process X is a Gaussian process. Compute $\mathrm{E}X_t$ and $\mathrm{Cov}(X_r\,, X_{r+t})$, and verify that if h is C^2, then

$$h(X_t) = h(x) + \int_0^t h'(X_s)\,\mathrm{d}B_s + \int_0^t (Lh)(X_s)\,\mathrm{d}s,$$

where $(Lh)(x) := \frac{1}{2}h''(x) - \nu x h'(x)$. If $\beta > 0$, then X is called the [non-stationary] *Ornstein–Uhlenbeck process* of parameter β. Finally, check that if $\nu \geq 0$, then $\beta_t := e^{\nu t}X_{\sigma(t)} - x$ defines a Brownian motion (in its own filtration), provided that $\sigma(t) := (2\nu)^{-1}\ln(1 + 2\nu t)$ for all $t \geq 0$ if $\nu > 0$, and $\sigma(t) := t$ if $\nu = 0$.

Classically, B_t describes the displacement of a (here, one-dimensional) particle that is subject only to molecular motion. Because $\|B_t - B_s\|_{L^2(\mathrm{P})} = |t-s|^{1/2}$, Brownian motion is not "differentiable in square mean." This observation makes B an undesirable mathematical description for molecular motion. Ornstein and Uhlenbeck suggested that a better description might be had if we use B to describe the acceleration of the Brownian particle [instead of the displacement]. Langevin's equation is the physical description of the resulting velocity X_t at time t.

6.3. The Geometric/Exponential Brownian Motion. Given two constants $\mu \in \mathbf{R}$ and $\sigma > 0$, we seek to find the solution to the following SDE:

$$\mathrm{d}X_t = \mu X_t \, \mathrm{d}t + \sigma X_t \, \mathrm{d}B_t, \tag{6.3}$$

where B denotes linear Brownian motion. The strong solution to this SDE exists and is unique by Exercise 2.4 (p. 53). It is called the *geometric* (sometimes *exponential*) *Brownian motion.* In mathematical finance, X_t denotes the value of a risky security (such as a stock) at time t, and (6.3) states that $\mathrm{d}X_t/X_t = \mu \, \mathrm{d}t + \sigma \, \mathrm{d}B_t$, or roughly $(X_{t+\epsilon} - X_t)/X_t \approx \mu\epsilon + \sigma(B_{t+\epsilon} - B_t)$ for small $\epsilon > 0$; i.e., relative changes in the stock price are governed by the changes in a *Brownian motion with drift* μ *and volatility* σ.

We would like to think of $\mathrm{d}X_t/X_t$ as $\mathrm{d}(\ln X_t)/\mathrm{d}t$. This is not quite correct, but motivates the choice of $f(x) := \ln x$, to which we apply Itô's formula (for $x > 0$) and find that for all $0 \le t < T_0 := \inf\{s \ge 0 : X_s = 0\}$,

$$\ln(X_t/X_0) = \int_0^t \frac{\mathrm{d}X_s}{X_s} - \frac{1}{2}\int_0^t \frac{\mathrm{d}\langle X\rangle_s}{X_s^2} = \mu t + \sigma B_t - \frac{\sigma^2 t}{2}.$$

Or equivalently, geometric Brownian motion with parameters μ and σ is the process

$$X_t := X_0 \exp\left(t\left[\mu - \frac{\sigma^2}{2}\right] + \sigma B_t \right), \tag{6.4}$$

for all $t \le T_0$. Note that unless $X_0 \equiv 0$, X_t—as defined by (6.4)—is never zero, and therefore $T_0 = \infty$ a.s. Thus, (6.4) holds for all $t \ge 0$.

Exercise 6.2. Define $\kappa := \sigma^2/(2\mu)$ if $\mu \neq 0$. Then, prove that $f(X_t)$ is a local martingale, where

$$f(z) = \begin{cases} \left|\kappa\left(1 - e^{-z/\kappa}\right)\right| & \text{if } \mu \neq 0, \\ z & \text{if } \mu = 0. \end{cases}$$

Define $T_z := \inf\{s > 0 : X_s = z\}$ for all $z \ge 0$, where $\inf \varnothing := \infty$. If $r < X_0 < R$, then compute $\mathrm{P}\{T_R < T_r\}$; this is the probability that the value of our stock reaches the height of R before it plummets down to r.

6.4. The Basic Wright–Fisher Model. Suppose that there are many genes in a given [large] population, and the genes are of two types: Types $\mathfrak{a}$ and $\mathfrak{b}$. We are interested in the evolution of the fraction X_t of genes of type $\mathfrak{a}$ at time t, as t changes. A simple model for this evolution follows: First, let us write $X_0 := x \in (0\,,1)$, where x is a nonrandom constant.

First let us suppose there are no mutations, selections, etc. Then, given that by time j/n the fraction of $\mathfrak{a}$'s is $X_{j/n}$, and if there are $n \approx \infty$ genes at any time, then we want the conditional distribution of $nX_{(j+1)/n}$, given $\{X_{i/n}\}_{i=0}^{j}$, to be $\text{Bin}(n\,, X_{j/n})$. That is, every gene will be type $\mathfrak{a}$ with probability $X_{j/n}$; else, it will be type $\mathfrak{b}$. If $n \approx \infty$, then by the classical central limit theorem, $[\text{Bin}(n\,, X_{j/n})/n - X_{j/n}]$ is approximately $N(0\,, X_{j/n}(1 - X_{j/n})/n) = \sqrt{X_{j/n}(1 - X_{j/n})}N(0\,, 1/n)$, and we have

$$X_{(j+1)/n} - X_{j/n} \approx \sqrt{X_{j/n}(1 - X_{j/n})} \cdot N_{j/n},$$

where $N_0, N_{1/n}, N_{2/n}, \ldots$ are i.i.d. $N(0\,, 1/n)$'s. If B denotes a linear Brownian motion, then we can realize $N_{j/n}$ as $B_{(j+1)/n} - B_{j/n}$, whence we are led to the following: If the population size n is very large, then a reasonable continuous-time model for the evolution of type-$\mathfrak{a}$ genes is

$$\text{(6.5)} \qquad \mathrm{d}X_t = \sqrt{X_t(1 - X_t)}\,\mathrm{d}B_t, \quad \text{subject to } X_0 = x.$$

[Can you deduce from the preceding story that X is a natural candidate for a "continuous-time, continuous-state branching process"?]

Exercise 6.3. Define $T_z := \inf\{s > 0 : X_s = z\}$ for all $z \in [0\,, 1]$, where $\inf \varnothing := \infty$. Compute $\mathrm{P}\{T_1 < T_0\}$; this is the probability that genes of type $\mathfrak{b}$ ultimately die out.

7. Stochastic Flows

Let B denote d-dimensional Brownian motion and consider the unique solution X^x to the SDE

$$\text{(7.1)} \qquad X_t^x = x + \int_0^t \sigma(X_s^x)\,\mathrm{d}B_s + \int_0^t b(X_s^x)\,ds \quad \text{for all } t \geq 0,$$

where $\sigma : \mathbf{R}^k \to \mathbf{R}^{kd}$ and $b : \mathbf{R}^k \to \mathbf{R}^k$ are Lipschitz continuous and $x \in \mathbf{R}^k$ is fixed. Note that B is d-dimensional and X^x is k-dimensional.

So far we have solved (7.1) for each x. Next we prove that the previous, viewed as a system of SDEs indexed by x, can be solved simultaneously for all x and t, and the result is a continuous random function of $(x\,, t)$. If we let Y_t^x denote the "simultaneous solution," then it is easy to see that $\mathrm{P}\{Y_t^x = X_t^x \text{ for all } t \geq 0\} = 1$ for all $x \in \mathbf{R}^k$; i.e., the Y is a t-uniform modification of X.

The process $x \mapsto Y_t^x$ is called the *stochastic flow* associated to the SDE $\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t + b(X_t)\,\mathrm{d}t$.

Theorem 7.1. *There exists a jointly continuous process* $(x\,,t) \mapsto Y_t^x$ *that solves* (7.1) *off a null set, simultaneously for all* t *and* x.

Proof. First let X_t^x denote the solution for each fixed x. For all $x, y \in \mathbf{R}^d$ define $M_t := \int_0^t \sigma(X_s^x)\,\mathrm{d}B_s - \int_0^t \sigma(X_s^y)\,\mathrm{d}B_s$ and $V_t := \int_0^t b(X_s^x)\,\mathrm{d}s - \int_0^t b(X_s^y)\,\mathrm{d}s$. Since σ is a matrix-valued Lipschitz function,

$$\langle M^j \rangle_t = \int_0^t \|\sigma(X_s^x) - \sigma(X_s^y)\|^2 \,\mathrm{d}s \le \mathrm{const} \cdot \int_0^t |X_s^x - X_s^y|^2 \,\mathrm{d}s,$$

where we are writing $M_t := (M_t^1\,, \ldots, M_t^k)$ coordinatewise. Therefore, by Exercise 5.12 on page 45,

$$\mathrm{E}\left[\sup_{s\in[0,t]} |M_s|^{2n}\right] \le \mathrm{const} \cdot \mathrm{E}\left(\left|\int_0^t |X_s^x - X_s^y|^2 \,\mathrm{d}s\right|^n\right).$$

Also,

$$\begin{aligned}\mathrm{E}\left(\sup_{s\in[0,t]} |V_s|^{2n}\right) &\le \mathrm{E}\left(\left|\int_0^t |X_s^x - X_s^y| \,\mathrm{d}s\right|^{2n}\right)\\ &\le \mathrm{const} \cdot \mathrm{E}\left(\left|\int_0^t |X_s^x - X_s^y|^2 \,\mathrm{d}s\right|^n\right).\end{aligned}$$

Therefore, the Gronwall lemma (Lemma 2.1) implies that

$$\begin{aligned}\mathrm{E}\left(\sup_{s\in[0,t]} |X_s^x - X_s^y|^{2n}\right) &\le \mathrm{const}\left\{|x-y|^{2n} + \mathrm{E}\left(\left|\int_0^t |X_s^x - X_s^y|^2 \,\mathrm{d}s\right|^n\right)\right\}\\ &\le \mathrm{const} \cdot |x-y|^{2n}.\end{aligned}$$

By the Kolmogorov continuity theorem (p. 5), there exists for all $t \ge 0$ fixed a continuous process $\mathbf{R}^k \ni y \mapsto Y_\bullet^y \in C(\mathbf{R}^k, C([0\,,t]\,,\mathbf{R}^k))$ such that $\mathrm{P}\{Y_s^y = X_s^y \text{ for all } s \in [0\,,t]\} = 1$ [you need to parse this!]. We "let $t \uparrow \infty$" by using the Kolmogorov extension theorem of basic probability. This yields a continuous process $(t\,,y) \mapsto Y_t^y$ such that $\mathrm{P}\{Y_t^y = X_t^y \text{ for all } t \ge 0\} = 1$. Y solves (7.1) for all $t \ge 0$, a.s. for each fixed y, simply because X does. By countable additivity, the same is true simultaneously for all $(t\,,y) \in \mathbf{R}_+ \times \mathbf{Q}^k$. But what we just showed actually proves that, after making further modifications of the terms in (7.1), we can ensure that both sides of (7.1) are continuous. If two continuous functions agree on rationals they agree. This proves the result. □

Exercise 7.2. The preceding proof shows a little more. Namely, that Y_t^x can be chosen so that it is Hölder continuous of any index < 1 [in the x variable]. Prove that if σ and b are C^∞, then $x \mapsto Y_t^x$ can be chosen to be C^∞. [For simplicity, you may assume that $k = d = 1$.] Here is a hint. Start by showing that the following SDE has a solution:

$$\mathrm{d}Z_t^x = \sigma'(X_t^x) Z_t^x \,\mathrm{d}B_t + b'(X_t^x) Z_t^x \,\mathrm{d}t.$$

Then argue that: (a) $(t\,, x) \mapsto Z_t^x$ has a continuous modification; and (b) Z is a modification of $\mathrm{d}X_t^x/\mathrm{d}x$.

8. The Strong Markov Property

Let $\sigma, b : \mathbf{R} \to \mathbf{R}$ be two Lipschitz continuous nonrandom functions. Thanks to Theorem 7.1 we can solve the following simultaneously in all $t \geq 0$ and $x \in \mathbf{R}$:

$$X_t^x = x + \int_0^t \sigma(X_s^x)\,\mathrm{d}B_s + \int_0^t b(X_s^x)\,\mathrm{d}s, \tag{8.1}$$

and moreover $(t\,, x) \mapsto X_t^x$ is the a.s.-unique continuous [strong] solution.

Next we note that

$$X_{t+h}^x = X_t^x + \int_t^{t+h} \sigma(X_s^x)\,\mathrm{d}B_s + \int_t^{t+h} b(X_s^x)\,\mathrm{d}s. \tag{8.2}$$

This is valid simultaneously for all $t, h \geq 0$. Therefore, we may replace t by an a.s.-finite stopping time τ, and define $W_s := B_{\tau+s} - B_\tau$. By the strong Markov property of Brownian motion, the process W is a Brownian motion independent of $\mathscr{F}_\tau$. Moreover, if $Y_h := X_{\tau+h}^x$, then it is not hard to see that

$$Y_h = X_\tau^x + \int_0^h \sigma(Y_s)\,\mathrm{d}W_s + \int_0^h b(Y_s)\,\mathrm{d}s \qquad \text{for all } h \geq 0. \tag{8.3}$$

By uniqueness, the conditional finite-dimensional distributions of the process $\{Y_h\}_{h\geq 0}$, given $\mathscr{F}_\tau$, are the same as those of $\{\bar{X}_h^{X_\tau x}\}_{h\geq 0}$, where $\bar{X}$ is an independent copy of X. More precisely, for all real numbers $h_1, \ldots, h_k \geq 0$ and Borel sets $A_1, \ldots, A_k \subset \mathbf{R}^k$,

$$\mathrm{P}\left(\bigcap_{j=1}^k \left\{ \bar{X}_{\tau+h_j}^{X_\tau^x} \in A_j \right\} \middle| \mathscr{F}_\tau \right) = \mathrm{P}\left(\bigcap_{j=1}^k \left\{ Y_{h_j} \in A_j \right\} \middle| \mathscr{F}_\tau \right) \tag{8.4}$$

is a Borel-measurable function of X_τ^x. We have proved the following.

Proposition 8.1. *If σ and b are Lipschitz continuous, then the solution to* (8.1) *is a strong Markov process for each x fixed.*

In fact, the infinite-dimensional process $t \mapsto X_t^\bullet$—with values in the Banach space $C(\mathbf{R}^k, \mathbf{R}^k)$—is also strong Markov. The proof is the same.

9. A Comparison Principle

Theorem 9.1 (Yamada, …). *Choose and fix three Lipschitz continuous functions σ, b, and $\bar{b}$ on $\mathbf{R}$. Suppose $x \le \bar{x}$, $b(z) \le \bar{b}(z)$ for all $z \in \mathbf{R}$, $\mathrm{d}X_t = \sigma(X_t)\,\mathrm{d}B_t + b(X_t)\,\mathrm{d}t$ subject to $X_0 = x$, and $\mathrm{d}\bar{X}_t = \sigma(\bar{X}_t)\,\mathrm{d}B_t + \bar{b}(\bar{X}_t)\,\mathrm{d}t$ subject to $\bar{X}_0 = \bar{x}$. Then*

$$\mathrm{P}\left\{X_t \le \bar{X}_t \textit{ for all } t \ge 0\right\} = 1.$$

Proof. By localization we may assume without loss of generality that σ is bounded. Else, we prove that $X_{t\wedge T_k} \le \bar{X}_{t\wedge T_k}$ for all t and k, where T_k denotes the first time s such that $|\sigma(X_s)| + |\sigma(\bar{X}_s)| > k$.

If we define $D_t := X_t - \bar{X}_t$, then by Tanaka's formula, D^+ is a semimartingale that satisfies

$$(D_t)^+ = \int_0^t \mathbf{1}_{\{D_s>0\}}\,\mathrm{d}D_s + \frac{1}{2}L_t^0,$$

where L_t^0 denotes the local time of D at zero, at time t. See Exercise 5.11 on page 44). Because $\mathrm{d}D_t = \{\sigma(X_t) - \sigma(\bar{X}_t)\}\mathrm{d}B_t + \{b(X_t) - \bar{b}(\bar{X}_t)\}\,\mathrm{d}t$,

$$\mathrm{E}\left[(D_t)^+\right] = \mathrm{E}\left(\int_0^t \mathbf{1}_{\{X_s>\bar{X}_s\}}\left[b(X_s) - \bar{b}(\bar{X}_s)\right]\,\mathrm{d}s\right) + \frac{1}{2}\mathrm{E}\left[L_t^0\right].$$

I claim that $L_t^0 \equiv 0$. In order to prove this let $Z_t := (D_t)^+$ and note that $\mathrm{d}\langle Z\rangle_t = \mathbf{1}_{\{D_t>0\}}|\sigma(X_t) - \sigma(\bar{X}_t)|^2\,\mathrm{d}t$, and hence

$$\int_{-\infty}^{\infty} L_t^a\,\frac{\mathrm{d}a}{a^2} = \int_0^t \frac{\mathrm{d}\langle Z\rangle_s}{|Z_s|^2} \le \mathrm{const}\cdot t,$$

thanks to the occupation density formula (p. 41), and the fact that σ is Lipschitz. Since $a \mapsto L_t^a$ is continuous and a^{-2} is not integrable near $a = 0$, this proves that $L_t^0 \equiv 0$. Consequently,

$$\begin{aligned}
\mathrm{E}\left[\left(X_t - \bar{X}_t\right)^+\right] &= \mathrm{E}\left(\int_0^t \mathbf{1}_{\{X_s>\bar{X}_s\}}\left[b(X_s) - \bar{b}(\bar{X}_s)\right]\,\mathrm{d}s\right)\\
&\le \mathrm{E}\left(\int_0^t \mathbf{1}_{\{X_s>\bar{X}_s\}}\left[\bar{b}(X_s) - \bar{b}(\bar{X}_s)\right]\,\mathrm{d}s\right)\\
&\le \mathrm{const}\cdot \mathrm{E}\left(\int_0^t \left(X_s - \bar{X}_s\right)^+\,\mathrm{d}s\right),
\end{aligned}$$

since $\bar{b}$ is Lipschitz continuous. Gronwall's lemma (p. 51) finishes the proof. □

Remark 9.2. Closer inspection of the proof reveals that we only require the following conditions: (i) X and $\bar{X}$ exist as unique strong solutions to their respective SDEs; (ii) σ is Hölder continuous of some order $\alpha \ge 1/2$; and (iii) at least one of b and $\bar{b}$ is Lipschitz continuous. [The preceding works if $\bar{b}$ is; if b is Lipschitz continuous, then consider $(D_t)^-$ in place of $(D_t)^+$.] □

10. Bessel Processes

10.1. Bessel-Squared Processes. Consider the SDE

$$X_t^{(\delta)} = x + 2\int_0^t |X_s^{(\delta)}|^{1/2}\,\mathrm{d}B_s + \delta t, \tag{10.1}$$

where $\delta \in \mathbf{R}$ is fixed and B is linear Brownian motion. According to Corollary 5.4 (p. 60) the preceding SDE has a unique strong solution $X^{(\delta)}$, which is sometimes called the *Bessel-squared process* [*or squared Bessel process*] *of dimension* δ.

We are primarily interested in the case that $\delta \geq 0$ and $x \geq 0$. First, let us note that the solution to (10.1) with $\delta = x = 0$ is identically zero. This and the comparison theorem (page 66) together prove that with probability one, $X_t^{(\delta)} \geq 0$ for all $t \geq 0$. Therefore we can remove the absolute value in (10.1).

We apply Itô's formula to find that for all C^2 functions f, and all $t \geq 0$,

$$f(X_t^{(\delta)}) = f(x) + 2\int_0^t f'(X_s^{(\delta)})|X_s^{(\delta)}|^{1/2}\,\mathrm{d}B_s + \int_0^t (Lf)(X_s^{(\delta)})\,\mathrm{d}s,$$

where

$$(Lf)(z) := \delta f'(z) + 2zf''(z).$$

Define

$$\nu := 1 - \frac{\delta}{2}.$$

Then it is easy to check directly that $Lf \equiv 0$, where $f(z) := z^\nu$ if $\delta \neq 2$ and $f(z) := \ln z$ if $\delta = 2$. Consequently, a gambler's ruin computation shows that whenever $0 < a < x < b$,

$$\mathrm{P}\left(X^{(\delta)} \text{ hits } a \text{ before } b \,\middle|\, X_0^{(\delta)} = x \right) = \begin{cases} \dfrac{b^\nu - x^\nu}{b^\nu - a^\nu} & \text{if } \delta \neq 2, \\[2ex] \dfrac{\ln(b/x)}{\ln(b/a)} & \text{if } \delta = 2. \end{cases}$$

When $\delta \in [0\,, 2)$ we have $\nu > 0$. Thus, we can let $a \downarrow 0$ and then $b \uparrow \infty$—in this order—to deduce that $X^{(\delta)}$ hits zero at some finite time a.s. On the other hand, a similar analysis shows that for values of $\delta \geq 2$, $X_t^{(\delta)} \neq 0$ for any $t > 0$.

Exercise 10.1. Let $T_0 := \inf\{s > 0 : X_s^{(0)} = 0\}$, and prove that a.s., $X_{t+T_0}^{(0)} = 0$ for all $t \geq 0$.

Exercise 10.2. Prove that if $\delta > \eta$, then $X_t^{(\delta)} \geq X_t^{(\eta)}$ for all $t > 0$, provided that the inequality holds for $t = 0$. For a more challenging problem try the following variant: Suppose $X_0^{(\delta)} > X_0^{(\eta)}$; then prove that $X_t^{(\delta)} > X_t^{(\eta)}$ for all $t > 0$.

Exercise 10.3. Define $Z_t := \|W_t\|^2$ where W is n-dimensional Brownian motion ($n \geq 1$). Recall (or check) that Z is a Bessel-squared process of dimension $\delta = n$. That is, Bessel-squared processes can be obtained as the modulus-squared of multidimensional Brownian motion when δ is a positive integer. But of course Bessel-squared processes are defined when δ is nonintegral as well.

Exercise 10.4. Prove that $\mathrm{P}\{\lim_{t\to\infty} X_t^{(\delta)} = \infty\} = 1$ when $\delta > 2$, whereas $\mathrm{P}\{\liminf_{t\to\infty} X_t^{(\delta)} = 0\} = \mathrm{P}\{\limsup_{t\to\infty} X_t^{(\delta)} = \infty\} = 1$ when $\delta = 2$.

10.2. Bessel Processes. Let $X^{(\delta)}$ denote the Bessel-squared process of dimension $\delta \geq 0$ starting from $x \geq 0$, and define

$$R_t^{(\delta)} := |X_t^{(\delta)}|^{1/2}.$$

The process $R^{(\delta)}$ is called the *Bessel process of dimension* δ. According to Exercise 10.3, $R^{(\delta)}$ can be realized as the radial part of a δ-dimensional Brownian motion when δ is a positive integer.

We apply the Itô formula to $R^{(\delta)} := f \circ X^{(\delta)}$ with $f(x) := x^{1/2}$ and deduce that

$$R_t^{(\delta)} = R_0^{(\delta)} + B_t + \frac{\delta - 1}{2} \int_0^t \frac{\mathrm{d}s}{R_s^{(\delta)}},$$

for all $t < T_0 := \inf\{s > 0 : R_s^{(\delta)} = 0\}$ ($\inf \varnothing := \infty$). According to the preceding subsection $T_0 = \infty$ when $\delta \geq 2$, but $T_0 < \infty$ when $\delta \in [0\,,2)$. In that case, the preceding SDE has an "explosion" at time T_0.