Chapter 2
Introduction to stochastic calculus

In this chapter, we focus on the essential aspects of *stochastic calculus*, a theory of integration with respect to Brownian-motion-type processes and their transformation properties. We select a "standard approach" to such questions and don't make shortcuts even though they might be possible under some additional assumptions. We present here the proofs that are necessary for understanding the theory and leave the less important ones to the appendices.

2.1 Brownian motion

Let's use the following notation: $\mathbb{Z}_{\geq 0} = \{k \in \mathbb{Z} : k \geq 0\}$ and $\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\}$.

Suppose throughout this text that we are given a probability space $(\Omega, \mathscr{F}, \mathsf{P})$ where Ω is a measurable space, $\mathscr{F}$ a σ-algebra on Ω and P a probability measure on $\mathscr{F}$. For more details, consult any book on probability theory, for instance, Durrett's book [2].

Definition 2.1. A *stochastic process* is a collection of random variables[1] X_t indexed by a variable t which we call time and which belongs to an ordered set I. A notation $(X_t)_{t \in I}$ is used for a stochastic process.

Almost always $I = \mathbb{R}_{\geq 0}$ or $I = \mathbb{Z}_{\geq 0}$. Since t is regarded as time, we call the process in those cases *continuous time stochastic process* and *discrete time stochastic process*, respectively. In this text usually $I = \mathbb{R}_{\geq 0}$.

The mapping $t \mapsto X_t(\omega)$ is called the *path* of $(X_t)_{t \in I}$. For continuous time processes the path regularity properties are usually essential already when defining the process (as in the definition of Brownian motion below).

Remember that X is a normally distributed with mean μ and variance σ^2 if and only if

[1] We use the standard notation $X(\omega)$, where $\omega \in \Omega$ and (usually) $X(\omega) \in \mathbb{R}$, for a random variable on $(\Omega, \mathscr{F}, \mathsf{P})$.

$$\mathsf{P}[X \in A] = \int_A \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathrm{d}x$$

for any Borel subset A of $\mathbb{R}$.

Definition 2.2. A stochastic process $(B_t)_{t\geq 0}$ is called a *(standard one-dimensional) Brownian motion* if $B_0 = 0$ and

1. $B_{t_1} - B_{s_1}, B_{t_2} - B_{s_2}, \ldots, B_{t_n} - B_{s_n}$ are independent for any $n \in \mathbb{N}$ and for any $0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \ldots \leq s_n < t_n$.
2. For any $s,t \geq 0$, $B_{s+t} - B_s$ is normally distributed with mean 0 and variance t.
3. With probability one, $t \mapsto B_t$ is continuous.

Remark 2.1. We say that the process has *independent* and *stationary* increments, if the properties 1. and 2. hold, respectively.

Remark 2.2. Brownian motion is a Gaussian process meaning that all finite dimensional distributions are multivariate Gaussians. By the above definition, for any $0 \leq s_1 < t_1 \leq s_2 < t_2 \leq \ldots \leq s_n < t_n$ and any Borel sets $A_k \subset \mathbb{R}$, $k = 1, 2, \ldots, n$, it holds that

$$\mathsf{P}[B_{t_k} - B_{s_k} \in A_k, \forall k = 1, 2, \ldots, n] = \prod_{k=1}^{n} \int_{A_k} \frac{1}{\sqrt{2\pi(t_k - s_k)}} \exp\left(-\frac{x^2}{2(t_k - s_k)}\right) \mathrm{d}x.$$

The "canonical" probability space for Brownian motion is the space of continuous functions $C(\mathbb{R}_{\geq 0})$ with a certain Borel probability measure P and where the Brownian motion is the coordinate map $B_t(\omega) = \omega_t$. If a Brownian motion exists in some probability space, its distribution in $C(\mathbb{R}_{\geq 0})$ defines the 'canonical" Brownian motion. The typical rough appearance of a Brownian motion path is illustrated in Figure 1.4.

There are many ways to construct a Brownian motion. One of them is given to the reader as an exercise.[2]

Theorem 2.1. *A probability space with a Brownian motion exists.*

A related result is the following on its regularity.

Theorem 2.2. *For each $\gamma \in (0,1)$ and $T > 0$, there exists a random variable $K > 0$ such that almost surely*

$$|B_t - B_s| \leq K\,|t-s|^\gamma$$

for all $s,t \in [0,T]$.

A *standard d-dimensional Brownian motion* is a $\mathbb{R}^d$-valued stochastic process $(B_t^{(1)}, \ldots, B_t^{(d)})_{t\in\mathbb{R}_{\geq 0}}$ where $B_t^{(1)}, \ldots, B_t^{(d)}$ are independent standard one-dimensional Brownian motions.

The following theorem shows that the assumption that the increments are normal is partly redundant in the definition of Brownian motion.

---

[2] The exercises have been written in a separate document(s).

Theorem 2.3. *If $(X_t)_{t\in\mathbb{R}_{\geq 0}}$ is a continuous stochastic process which has independent and stationary increments, then there exists a standard one-dimensional Brownian motion $(B_t)_{t\in\mathbb{R}_{\geq 0}}$ and real numbers $\alpha \geq 0$ and β such that $X_t = \alpha B_t + \beta t$.*

Remark 2.3. The process $X_t = \alpha B_t + \beta t$ is called a *Brownian motion with a linear drift*.

Definition 2.3. A *filtration* on $(\Omega, \mathscr{F})$ is a collection $(\mathscr{F}_t)_{t\in\mathbb{R}_{\geq 0}}$ of sub-σ-algebras $\mathscr{F}_t \subset \mathscr{F}$ such that for each $0 \leq s < t$, $\mathscr{F}_s \subset \mathscr{F}_t$.

A filtration can be though as refining information on the probability space and $\mathscr{F}_t$ as the information available at time t. For example, the σ-algebras generated by a Brownian motion $(B_t)_{t\in\mathbb{R}_+}$, i.e. $\mathscr{F}_t^B = \sigma(B_s, s \in [0,t])$, form a filtration $(\mathscr{F}_t^B)_{t\in\mathbb{R}_+}$.[3]

Definition 2.4. A stochastic process $(X_t)_{t\in\mathbb{R}_+}$ on $(\Omega, \mathscr{F})$ is *adapted* to the filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ if for each $t \geq 0$, X_t is $\mathscr{F}_t$-measurable.

We will make the following more restrictive definition of Brownian motion.

Definition 2.5. A process $(B_t)_{t\geq 0}$ is called a *(standard one-dimensional) Brownian motion with respect to the filtration* $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ if it is adapted to $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$, $B_0 = 0$ and

1. $B_t - B_s$ are independent from $\mathscr{F}_s$ for any $0 \leq s < t$,
2. $B_t - B_s$, $0 \leq s < t$, is normally distributed with mean 0 and variance $t - s$
3. With probability one, $t \mapsto B_t$ is continuous.

Remark 2.4. In the definition $(\mathscr{F}_t)_{t\in\mathbb{R}_{\geq 0}}$ can be weakened to $(\mathscr{F}_t^B)_{t\in\mathbb{R}_+}$ and therefore Definition 2.5 implies Definition 2.2. This definition is useful for instance when two Brownian motions $B^{(1)}$ and $B^{(2)}$ are considered on the same probability space.

2.1.1 Quadratic variation of Brownian motion

Definition 2.6. Let $p \geq 1$. Define the *p'th variation* of a process $(X_t)_{t\in\mathbb{R}_+}$ as the process

$$V_X^{(p)}(t) = \lim_{\mathrm{mesh}(\pi)\to 0} \sum_{k=0}^{m(\pi)-1} |X_{t_{k+1}} - X_{t_k}|^p$$

where π is a partitions of $[0,t]$ of the form $\pi = \{0 = t_0 < t_1 < \ldots < t_{m(\pi)} = t\}$ and the limit is in terms of *convergence in probability* as $\mathrm{mesh}(\pi) = \max_k (t_{k+1} - t_k) \to 0$ in the sense that that for each $\varepsilon > 0$ there exist $\delta > 0$ such that

$$\mathsf{P}\left[\left|\sum_{k=0}^{m(\pi)-1} |X_{t_{k+1}} - X_{t_k}|^p - V_X^{(p)}(t)\right| \geq \varepsilon\right] < \varepsilon$$

[3] We use the standard notations $\sigma(A, B, \ldots)$ and $\sigma(A_i, i \in I)$ for the σ-algebra generated by the random variables $A, B, \ldots$ and A_i, $i \in I$, respectively.

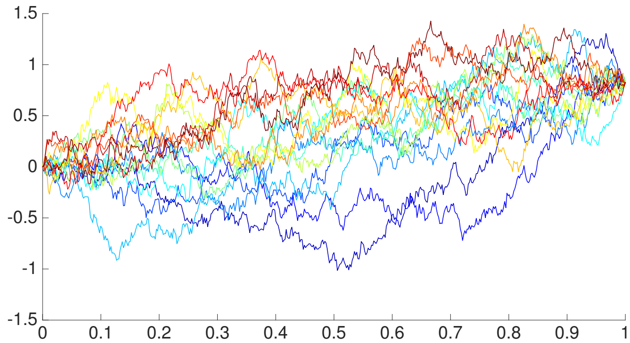


Fig. 2.1 If Brownian motion is conditioned for a particular end value at T, the resulting process is called Brownian bridge. In the figure, 16 instances of Brownian bridge paths are plotted.

when $\text{mesh}(\pi) < \delta$. We call the first variation ($p = 1$) as *total variation* and the second variation ($p = 2$) as *quadratic variation*.

Proposition 2.1. *The quadratic variation of a Brownian motion exist and* $V_B^{(2)}(t) = t$.

Proof. Let $\varepsilon > 0$ and π be a partitioning with $\text{mesh}(\pi) < (2t)^{-1}\varepsilon^3$. Let $\Delta_k = (B_{t_{k+1}} - B_{t_k})^2 - (t_{t+1} - t_k)$

$$\mathsf{E}\left[\left(\sum_{k=0}^{m(\pi)-1}(B_{t_{k+1}} - B_{t_k})^2 - t\right)^2\right] = \sum_k \mathsf{E}\left[\Delta_k^2\right] + 2\sum_{j<k} \underbrace{\mathsf{E}\left[\Delta_j \Delta_k\right]}_{=0,\text{ by independence}}$$
$$= \mathsf{E}[(N^2-1)^2]\sum_k (t_{k+1} - t_k)^2 \leq 2\text{mesh}(\pi)t.$$

Here N is normally distributed with mean zero and variance one and we used the scaling property of Brownian motion. Hence by Chebyshev's inequality [2]

$$\mathsf{P}\left[\left|\sum_{k=0}^{m(\pi)-1}(B_{t_{k+1}} - B_{t_k})^2 - t\right| \geq \varepsilon\right] \leq \frac{2\text{mesh}(\pi)t}{\varepsilon^2} < \varepsilon \tag{2.1}$$

and the convergence in probability follows. □

The above proof and the Borel–Cantelli lemma, see e.g. [2], gives that the total variation of a Brownian motion is almost surely infinite in the sense that if take the limit along the sequence of dyadic partitions $\pi_n = \{t\,k\,2^{-n} : k = 0,1,2,\ldots,2^n\} = \{t_0 < t_1 < \ldots < t_{2^n}\}$ of $[0,t]$, then

$$\lim_{n\to\infty} \sum_{t_k \in \pi_n, k \leq 2^n - 1} |B_{t_{k+1}} - B_{t_k}| = \infty \tag{2.2}$$

almost surely. Namely, if we denote $\mathsf{P}[E(\pi)]$ the left-hand side of (2.1), then $\sum_n \mathsf{P}[E(\pi_n)] < \infty$ and hence

$$\sum_{t_k \in \pi_n, k \leq 2^n - 1} (B_{t_{k+1}} - B_{t_k})^2 \to t$$

almost surely by the Borel–Cantelli lemma. Take any ω for which the convergence occurs. Then (2.2) is implied by the fact that as $n \to \infty$

$$\underbrace{\sum_{t_k \in \pi_n, k \leq 2^n - 1} (B_{t_{k+1}}(\omega) - B_{t_k}(\omega))^2}_{\to t} \leq \underbrace{\mathrm{mesh}(\pi_n)}_{\to 0} \sum_{t_k \in \pi_n, k \leq 2^n - 1} |B_{t_{k+1}}(\omega) - B_{t_k}(\omega)|.$$

2.2 Stochastic integration

The goal of this section is to define a process X_t which can be interpreted as the integral

$$X_t(\omega) = \int_0^t f(t, \omega) \, \mathrm{d}B_t(\omega).$$

It is important because of the following reasons:

- It is tool for generating new stochastic processes out of Brownian motion.
- Coordinate changes such as $f(B_t)$ turn out to have extremely useful representation using the above integral.
- It appears in many applications, since $\mathrm{d}B_t$ models independent and stationary noise.

The integral doesn't exist as a pathwise Riemann–Stieltjes (or similar) integral even for a continuous f, because the total variation of the Brownian motion is infinite. For instance, for the definition that we use, it holds that $\int_0^t B_s \mathrm{d}B_s \neq (1/2)B_t^2$ and therefore the usual integration by parts formula can't hold.

2.2.1 Stochastic integral as L^2-extension

In this section $(\mathscr{F}_t)_{t \in \mathbb{R}_{\geq 0}}$ is a filtration and $(B_t)_{t \in \mathbb{R}_{\geq 0}}$ is a standard one-dimensional Brownian motion with respect to $(\mathscr{F}_t)_{t \in \mathbb{R}_{\geq 0}}$.

We need to define the correct set of integrands f for the stochastic integral. In this subsection, they'll be the measurable, adapted and square-integrable processes.

Definition 2.7. A stochastic process $(X_t)_{t \in \mathbb{R}_{\geq 0}}$ is *measurable* if the mapping $(t, \omega) \mapsto X_t(\omega)$ is $\mathscr{B}_{\mathbb{R}} \times \mathscr{F}$-measurable.

Definition 2.8. Let $T > 0$. We define $\mathscr{L}^2$ to be the set of measurable, adapted processes f that satisfy

$$\mathsf{E}\left[\int_0^T f(t,\cdot)^2 \mathrm{d}t\right] < \infty \tag{2.3}$$

and we call $f \in \mathscr{L}^2$ *simple* if f can be written in the form

$$f(t,\omega) = \sum_{k=0}^{n-1} X_k(\omega) \mathbb{1}_{[t_k,t_{k+1})}(t) \tag{2.4}$$

where $0 \le t_0 < t_1 < t_2 \ldots < t_n \le T$ and X_k is a $\mathscr{F}_{t_k}$-measurable, square integrable random variable.

Remark 2.5. The above class could be called as $\mathscr{L}^2(T)$ and then we could set $f \in \mathscr{L}^2$ if and only if $f \in \mathscr{L}^2(T)$ for any $T > 0$. However, we don't make a big difference between $\mathscr{L}^2(T)$ and $\mathscr{L}^2$, and consequently, we use the notation $\mathscr{L}^2$ for both classes.

Remark 2.6. Notice that $\mathscr{L}^2$ is a closed subspace of $L^2(\mathrm{d}t \times \mathrm{dP})$.

We would like to define a mapping $f \mapsto I[f]$ which we later denote by

$$I[f](\omega) =: \int_0^T f(t,\omega) \mathrm{d}B_t(\omega).$$

If that notation makes any sense, we need to define

$$I[\mathbb{1}_{[s,t)}] = B_t - B_s$$

for any $0 \le s < t \le T$. Therefore for any f which is of the form (2.4) it is natural to define by linearity that

$$I[f] = \sum_{k=0}^{n-1} X_k (B_{t_{k+1}} - B_{t_k}).$$

It turns out that this definition that works for any simple $f \in \mathscr{L}^2$ has a unique L^2-continuous extension to the whole $\mathscr{L}^2$. Namely, we first observe that the following isometry holds.

Proposition 2.2 (Itô isometry for simple processes). *For any bounded, simple $f \in \mathscr{L}^2$*

$$\mathsf{E}\left[I[f]^2\right] = \mathsf{E}\left[\int_0^T f(t,\cdot)^2 \mathrm{d}t\right].$$

Proof. Let's calculate both sides explicitly for a bounded, simple $f \in \mathscr{L}^2$ of the form (2.4). Notice that $f^2 = \sum_{k=0}^{n-1} X_k^2 \mathbb{1}_{[t_k,t_{k+1})}$ and hence

$$\mathsf{E}\left[\int_0^T f(t,\cdot)^2 \mathrm{d}t\right] = \sum_{k=0}^{n-1} \mathsf{E}[X_k^2](t_{k+1} - t_k).$$

On the other hand

$$\mathsf{E}[I[f]^2] = \sum_k \mathsf{E}\left[X_k^2 (B_{t_{k+1}} - B_{t_k})^2\right] + 2\sum_{k<l} \mathsf{E}[X_k X_l (B_{t_{k+1}} - B_{t_k})(B_{t_{l+1}} - B_{t_l})].$$

The facts that f is adapted, and thus X_k is $\mathscr{F}_{t_k}$-measurable, and that $B_{t_{k+1}} - B_{t_k}$ is independent from $\mathscr{F}_{t_k}$ imply that

$$\mathsf{E}\left[X_k^2 (B_{t_{k+1}} - B_{t_k})^2\right] = \mathsf{E}[X_k^2]\,\mathsf{E}\left[(B_{t_{k+1}} - B_{t_k})^2\right] = \mathsf{E}[X_k^2]\,(t_{k+1} - t_k)$$
$$\mathsf{E}[X_k X_l (B_{t_{k+1}} - B_{t_k})(B_{t_{l+1}} - B_{t_l})] = \mathsf{E}[X_k X_l (B_{t_{k+1}} - B_{t_k})]\mathsf{E}[B_{t_{l+1}} - B_{t_l}] = 0$$

for $k < l$. The claim follows. □

The simple processes are dense in $\mathscr{L}^2$ by the next result. A sketch of its proof is given in Appendix B.

Proposition 2.3. *For each $f \in \mathscr{L}^2$, there exist a sequence of bounded, simple $f_n \in \mathscr{L}^2$ such that*

$$\mathsf{E}\left[\int_0^T (f(t,\cdot) - f_n(t,\cdot))^2 \mathrm{d}t\right] \to 0,$$

i.e. f_n converges to f in $L^2(\mathrm{d}t \times \mathrm{dP})$.

If $f_n \in \mathscr{L}^2$ is a sequence of simple, bounded processes converging to f, then f_n is a Cauchy sequence in $L^2(\mathrm{d}t \times \mathrm{dP})$ and hence by the isometry property $I[f_n]$ is a Cauchy sequence if $L^2(\mathrm{dP})$ and hence it converges. Therefore we can define $I[f] = \lim_n I[f_n]$. Notice that this limit doesn't depend on the choice of f_n: if f_n and f_n' are two such sequences, then $f_n - f_n'$ goes to zero in $L^2(\mathrm{d}t \times \mathrm{dP})$ and hence by isometry, $\lim_n I(f_n) = \lim_n I(f_n')$ almost surely. This is summarized in the following definition.

Definition 2.9. For any $f \in \mathscr{L}^2$, the *stochastic integral* (or *Itô integral*) is defined to be

$$\int_0^T f\,\mathrm{d}B_t(\omega) := I[f](\omega) := (\lim_n I[f_n)])(\omega) \tag{2.5}$$

where the limit is in $L^2(\mathsf{P})$ and $f_n \in \mathscr{L}^2$ is any sequence of bounded, simple processes converging to f in $L^2(\mathrm{d}t \times \mathrm{dP})$. The integral is defined almost surely.

Corollary 2.1 (Itô isometry for $\mathscr{L}^2$). *For any $f \in \mathscr{L}^2$*

$$\mathsf{E}\left[\left(\int_0^T f\,\mathrm{d}B_t\right)^2\right] = \mathsf{E}\left[\int_0^T f^2 \mathrm{d}t\right].$$

Corollary 2.2. *If $f_n \in \mathscr{L}^2$, $f \in \mathscr{L}^2$ and $f_n \to f$ in $L^2(\mathrm{d}t \times \mathrm{dP})$ then $\int_0^T f_n \mathrm{d}B_t \to \int_0^T f \mathrm{d}B_t$ in $L^2(\mathsf{P})$.*

Example 2.1. We'll show that

$$\int_0^t B_s \mathrm{d}B_s = \frac{1}{2}B_t - \frac{1}{2}t$$

Let π_n be a sequence of partitions of $[0,t]$ such that $\mathrm{mesh}(\pi_n) \to 0$. By the above, the sequence of processes $f_n(s,\omega) = \sum_{t_j \in \pi_n} B_{t_j}(\omega) \mathbb{1}_{[t_j,t_{j+1})}(s)$ is a reasonable choice for a discretization of the integrand. Since

$$\mathsf{E}\left[\int_0^t (B_s - f_n(s,\cdot))^2 \mathrm{d}s\right] = \mathsf{E}\left[\sum_j \int_{t_j}^{t_{j+1}} (B_s - B_{t_j})^2 \mathrm{d}s\right] = \sum_j \frac{1}{2}(t_{j+1}-t_j)^2 \to 0$$

as $n \to \infty$, then by Corollary 2.2, $\int_0^t B_s \mathrm{d}B_s = \lim \int_0^t f_n \mathrm{d}B_s = \lim \sum_j B_{t_j}(B_{t_{j+1}} - B_{t_j})$. Now notice that

$$B_{t_{j+1}}^2 - B_{t_j}^2 = (B_{t_{j+1}} - B_{t_j})^2 + 2B_{t_j}(B_{t_{j+1}} - B_{t_j})$$

and thus

$$\sum_j B_j(B_{t_{j+1}} - B_{t_j}) = \frac{1}{2}B_t^2 - \frac{1}{2}\sum_j (B_{t_{j+1}} - B_{t_j})^2$$

and the second term on the right converges in L^2 to the quadratic variation of Brownian motion which we already showed to be t.

The following proposition states some properties of the stochastic integral. Those properties hold for the simple processes and hence hold also for any limit of a sequence of simple processes.

Proposition 2.4. *Let $f,g \in \mathscr{L}^2$, $a,b \in \mathbb{R}$ and let $0 \le S < U < T$. Then*

1. $\int_S^T f \mathrm{d}B_t = \int_S^U f \mathrm{d}B_t + \int_U^T f \mathrm{d}B_t$
2. $\int_S^T (af + bg) \mathrm{d}B_t = a\int_S^T f \mathrm{d}B_t + b\int_S^T g \mathrm{d}B_t$
3. $\mathsf{E}[\int_S^T f \mathrm{d}B_t] = 0$
4. $\int_S^T f \mathrm{d}B_t$ *is $\mathscr{F}_T$-measurable*

2.2.2 Stochastic integral as a process

2.2.2.1 Stochastic integral as a continuous martingale

Based on the results of Section 2.2.1, we try to define a process X_t such that $X_t = \int_0^t f \,\mathrm{d}B_s$ for every t. The problem in defining $X_t = I[f \mathbb{1}_{[0,t]}]$ is that the for each fixed t, X_t is defined in a set of probability one, say, in Ω_t, but it is possible that the probability of the uncountable intersection $\bigcap_t \Omega_t$ is strictly less than 1 or even that $\bigcap_t \Omega_t$ is not an event (a measurable set). Therefore we define X_t in this way in a countable set of t and then extend by continuity of $t \mapsto X_t$ to other values of t as in the following theorem. For the definition of a martingale consult Appendix A.

Theorem 2.4. *For each $f \in \mathscr{L}^2$ there exists a continuous square integrable martingale $(X_t)_{t\in\mathbb{R}_+}$ such that for each t, $X_t = \int_0^t f(s,\cdot)\mathrm{d}B_s$ almost surely.*

Remark 2.7. The process $(X_t)_{t\in\mathbb{R}_{\geq 0}}$ is unique in the sense that if there is another process $(X_t')_{t\in\mathbb{R}_{\geq 0}}$ with the same properties, then almost surely $X_t = X_t'$ for all t.

Proof. Fix some $T > 0$. Take a sequence of simple (and bounded) $f_n \in \mathscr{L}^2$ such that $f_n \to f$ in $L^2(\mathrm{d}t \times \mathsf{dP}, [0,T]\times\Omega)$ and define $X_t^{(n)} = I[f_n \mathbb{1}_{[0,t]}]$ which is well-defined in whole Ω. If $f_n = \sum a_k \mathbb{1}_{[t_k,t_{k+1})}$, then for $t_l \leq t < t_{l+1}$ we have an explicit formula

$$X_t^{(n)} = a_l \cdot (B_t - B_{t_l}) + \sum_{k=0}^{l-1} a_k \cdot (B_{t_{k+1}} - B_{t_k}). \tag{2.6}$$

Clearly $t \mapsto X_t$ is continuous. To show that it is a martingale, notice first that it is adapted because all the random variables on the right of (2.6) are $\mathscr{F}_t$-measurable. Next notice that $\mathsf{E}[|X_t^{(n)}|] < \infty$, because it is a finite sum of integrable random variables. Finally, for $0 \leq s < t \leq T$ we can assume that $s = t_l$ and $t = t_m$ for some l and m (redefine the "partitioning" of f_n again if necessary) and then

$$\mathsf{E}[X_t^{(n)}|\mathscr{F}_s] = \mathsf{E}[X_s^{(n)}|\mathscr{F}_s] + \mathsf{E}[\sum_{k=l}^{m-1} a_k \cdot (B_{t_{k+1}} - B_{t_k})|\mathscr{F}_s] = X_s^{(n)}$$

because $X_s^{(n)}$ is $\mathscr{F}_s$-measurable and by the properties of conditional expectation (see Appendix A)

$$\begin{aligned}\mathsf{E}[a_k \cdot (B_{t_{k+1}} - B_{t_k})|\mathscr{F}_s] &= \mathsf{E}[\mathsf{E}[a_k \cdot (B_{t_{k+1}} - B_{t_k})|\mathscr{F}_{t_k}]|\mathscr{F}_s] \\ &= \mathsf{E}[a_k \cdot \mathsf{E}((B_{t_{k+1}} - B_{t_k})|\mathscr{F}_{t_k}]|\mathscr{F}_s] = 0.\end{aligned} \tag{2.7}$$

Since $X_t^{(n)} - X_t^{(m)}$ is a martingale, by Doob's maximal inequality

$$\begin{aligned}\mathsf{P}\left[\sup_{t\in[0,T]}\left|X_t^{(n)} - X_t^{(m)}\right| \geq \varepsilon\right] &\leq \frac{1}{\varepsilon^2}\mathsf{E}[|X_T^{(n)} - X_T^{(m)}|^2] \\ &= \frac{1}{\varepsilon^2}\|f_n - f_m\|^2_{L^2(\mathrm{d}t\times\mathsf{dP})}\end{aligned}$$

for any $\varepsilon > 0$. Choose a subsequence n_k such that $\|f_{n_{k+1}} - f_{n_k}\|^2_{L^2(\mathrm{d}t\times\mathsf{dP})} \leq 2^{-3k}$ and use the previous estimate for $\varepsilon = 2^{-k}$ to get

$$\mathsf{P}\left[\sup_{t\in[0,T]}\left|X_t^{(n_{k+1})} - X_t^{(n_k)}\right| \geq 2^{-k}\right] \leq 2^{-k}$$

By the Borel–Cantelli lemma, there exist random variable N which is almost surely finite and for $k \geq N(\omega)$

$$\sup_{t\in[0,T]}\left|X_t^{(n_{k+1})} - X_t^{(n_k)}\right| < 2^{-k}.$$

Hence the sequence of the continuous processes $(X_t^{(n_k)})$ converges almost surely uniformly to a continuous process (X_t). Since for any fixed t, $\lim X_t^{(n_k)}$ in $L^2(\mathsf{P})$ is $\int_0^t f \mathrm{d}B_s$ then

$$X_t = \int_0^t f \,\mathrm{d}B_s$$

almost surely. This also shows that (X_t) is adapted and square integrable.

Finally the martingale property of $(X_t^{(n)})$, for any $0 \leq s < t \leq T$

$$X_s^{(n)} = \mathsf{E}[X_t^{(n)} | \mathscr{F}_s].$$

Since the random variables $X_s^{(n)}$ and $X_s^{(n)}$ converge in $L^2(\mathsf{P})$ to X_s and X_t, respectively, then by the properties of conditional expectation (see Appendix A)

$$X_s = \mathsf{E}[X_t | \mathscr{F}_s].$$

for any $0 \leq s < t \leq T$. For the whole $\mathbb{R}_+$, the claim follows from the above by taking a countable sequence $T \nearrow \infty$ and using the uniqueness. □

Remark 2.8. The property that we used in (2.7) could be reformulated in the following way: if $(M_t)_{t \in \mathbb{R}_+}$ is a martingale and if $0 \leq s \leq t \leq u$ and Y is a $\mathscr{F}_t$-measurable bounded random variable, then

$$\mathsf{E}[Y(M_u - M_t) | \mathscr{F}_s] = 0.$$

We say that martingale increments $M_u - M_t$ are orthogonal to $\mathscr{F}_t$.

Definition 2.10. For any $f \in \mathscr{L}^2$, the *stochastic integral* (or *Itô integral*) is redefined to be a continuous version of $\int_0^t f \mathrm{d}B_s$, which exists by the previous Theorem.

Remark 2.9. The processes $(X_t)_{t \in \mathbb{R}_+}$ and $(Y_t)_{t \in \mathbb{R}_+}$ are versions of each other if $\mathsf{P}[X_t = Y_t] = 1$ for each t.

Definition 2.11. For any process $X_t = \int_0^t f \mathrm{d}B_s$, define the *quadratic variation process* as

$$\langle X \rangle_t(\omega) = \int_0^t f(s,\omega)^2 \mathrm{d}t.$$

The process $\langle X \rangle$ is the quadratic variation in the sense of Definition 2.6. We postpone the statement of that result. The following result gives a second interpretation of the quadratic variation process.

Theorem 2.5. *Let $f \in \mathscr{L}^2$, $X_t = \int_0^t f \mathrm{d}B_s$ and $\langle X \rangle_t$ as above. Then $X_t^2 - \langle X \rangle_t$ is a martingale.*

Proof. We leave as an exercise to check this for bounded, simple $f \in \mathscr{L}^2$. In the general case take a sequence of bounded, simple $f_n \in \mathscr{L}^2$ and define $X_t^{(n)} = \int_0^t f_n \mathrm{d}B_s$. The claim follows easily from the $L^1(\mathsf{P})$ convergence of $(X_t^{(n)})^2 - \langle X^{(n)} \rangle_t$ which implies the $L^1(\mathsf{P})$ convergence of $\mathsf{E}[(X_t^{(n)})^2 - \langle X^{(n)} \rangle_t \,|\, \mathscr{F}_s]$ by the properties of conditional expectation (see Appendix A). □

Next we define a stopping time which can be taught as the time when "some event occurs" so that for each time instant, the question whether this event has already occurred or not before or at that time is a "measurable question".

Definition 2.12. A random variable $\tau : \Omega \to [0,\infty]$ is called a *stopping time* with respect to the filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ if for all $t \geq 0, \{\omega : \tau(\omega) \leq t\} \in \mathscr{F}_t$.

An example of a stopping time is $\tau_A = \inf\{t \in \mathbb{R}_{\geq 0} : B_t \in A\}$ where A is a closed or open set in $\mathbb{R}$.

One way to describe the following result is that by that proposition, the pathwise interpretation of the Itô integral makes sense: if two integrands have the same paths up to a stopping time, then the integrals also agree up to that stopping time.

Proposition 2.5. *If τ is a stopping time and $f \in \mathscr{L}^2$ and $g \in \mathscr{L}^2$ processes such that $f(t,\omega) = g(t,\omega)$ for any (t,ω) such that $t \leq \tau(\omega)$, then for almost all ω*

$$\int_0^t f\,\mathrm{d}B_s(\omega) = \int_0^t g\,\mathrm{d}B_s(\omega)$$

for all $t \leq \tau(\omega)$.

For the proof see Appendix B.

2.2.2.2 Localization and a general class of integrands

At this point, we have the Itô integral defined for any measurable, adapted process f such that

$$\mathsf{E}\left[\int_0^T f^2 \mathrm{d}t\right] < \infty$$

for any $T \in (0,\infty)$. However, we would like to have a larger class of processes that includes at least all the continuous processes, such as $f(t,\omega) = \exp(B_t(\omega)^3)$ which is an example of a process that doesn't belong to $\mathscr{L}^2$.

Definition 2.13. $\mathscr{L}^2_{\mathrm{loc}}$ is definied to be the set of measurable, adapted process f such that

$$\int_0^T f(t,\cdot)^2\,\mathrm{d}t < \infty$$

almost surely for any $T \in (0,\infty)$.

Fix some $f \in \mathscr{L}^2_{\mathrm{loc}}$. Define a stopping time

$$\tau_n(\omega) = \inf\left\{t \in \mathbb{R}_+ : \int_0^t f(s,\omega)^2 \mathrm{d}s \geq n\right\}.$$

It follows from $f \in \mathscr{L}^2_{\mathrm{loc}}$, that $\tau_n \nearrow \infty$ almost surely as $n \to \infty$.

Let $f_n(t,\omega) = f(t,\omega)\mathbb{1}_{t\le\tau_n(\omega)}$. Then $f_n \in \mathscr{L}^2$ and we can define the Itô integral $X_t^{(n)} = \int_0^t f_n \mathrm{d}B_s$. Since $f_n(t,\omega) = f_m(t,\omega)$ for all (t,ω) such that $t \le (\tau_n \wedge \tau_m)(\omega)$ and since $\tau_n \wedge \tau_m$ is a stopping time[4], by Proposition 2.5 for almost all ω,

$$X_t^{(n)}(\omega) = X_t^{(m)}(\omega)$$

for $t \le (\tau_n \wedge \tau_m)(\omega)$.

For each fixed ω, this is a strong mode of convergence: there is a finite $n_0(\omega)$ such that $X_t^{(n)}(\omega) = X_t^{(m)}(\omega)$ for any $n,m \ge n_0(\omega)$. Define now a process $(X_t)_{t\in\mathbb{R}_+}$ on the event $\{\tau_n \nearrow \infty\}$

$$X_t(\omega) = X_t^{(n)}(\omega)$$

where $n \in \mathbb{N}$ is any number satisfying $\tau_n(\omega) \ge t$. The complement of the event $\{\tau_n \nearrow \infty\}$ has zero probability and there we can define $X_t = 0$ identically, say.

Definition 2.14. The Itô integral of $f \in \mathscr{L}^2_{\text{loc}}$ is defined as

$$\int_0^t f \mathrm{d}B_s(\omega) = X_t(\omega) = X_t^n(\omega)$$

where $n \in \mathbb{N}$ is any number satisfying $\tau_n(\omega) \ge t$ and $X_t^n(\omega)$ is as above.

For any continuous process $(X_t)_{t\in\mathbb{R}_{\ge 0}}$ and for any stopping time τ, define a stopped process $(X_t^\tau)_{t\in\mathbb{R}_{\ge 0}}$ by $X_t^\tau = X_{t\wedge\tau}$. The continuity of $(X_t)_{t\in\mathbb{R}_{\ge 0}}$ guarantees that X_t^τ is measurable.

Definition 2.15. A continuous process $(M_t)_{t\in\mathbb{R}_{\ge 0}}$ adapted to $(\mathscr{F}_t)_{t\in\mathbb{R}_{\ge 0}}$ is called *local martingale* if there exist a sequence of stopping times $0 \le \tau_1 \le \tau_2 \le \dots$ such that $\mathsf{P}(\tau_k \nearrow \infty) = 1$ and for each k, M^{τ_k} is a martingale. It is a *local square integrable martingale*, if each $(M_t^{\tau_k})_{t\in\mathbb{R}_{\ge 0}}$ is a square integrable martingale.

Remark 2.10. The use of stopping times of similar to τ_n, in Definitions 2.14 and 2.15, is called *localization* of the processes.

The next theorem lists the properties of Itô integral in its general form.

Theorem 2.6. *For any $f \in \mathscr{L}^2_{\text{loc}}$, the processes $X_t = \int_0^t f \,\mathrm{d}B_s$ and $X_t^2 - \langle X\rangle_t$ are continuous local martingale. Furthermore, $(X_t)_{t\in\mathbb{R}_{\ge 0}}$ has finite quadratic variation and almost surely for any t*

$$V_X^{(2)}(t) = \langle X\rangle_t.$$

The theorem follows from the properties of the Itô integral for $\mathscr{L}^2$-integrands and the construction of the integral using localization. For the proof last statement see Appendix B.

[4] The minimum of two stopping times is a stopping time.

2.3 Itô's formula

2.3.1 Itô's formula for a Brownian motion

Itô's formula is a result of central importance in stochastic calculus. We present first the version of it for a Brownian motion. Itô's formula shows that functions of Brownian motion can be written as sum of a stochastic integral and a integral with respect to dt.

Theorem 2.7 (Itô's formula for a Brownian motion). *Let $F : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ be a continuous function such that $\dot{F}, F', F''$ exist and are continuous, where*

$$\dot{F}(t,x) = \frac{\partial F}{\partial t}(t,x), \quad F'(t,x) = \frac{\partial F}{\partial x}(t,x) \quad \text{and } F''(t,x) = \frac{\partial^2 F}{\partial x^2}(t,x).$$

Then almost surely

$$F(t,B_t) = F(0,B_0) + \int_0^t \dot{F}(s,B_s)\mathrm{d}s + \int_0^t F'(s,B_s)\mathrm{d}B_s + \frac{1}{2}\int_0^t F''(s,B_s)\mathrm{d}s \tag{2.8}$$

for any $t \in \mathbb{R}_+$. For the previous equation we will use the shorthand notation

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

Proof (A sketch). The proof is based on the Taylor expansion of $F(t,x)$ in both of its variables.

Take a partition π of $[0,t]$ and write a telescoping sum

$$F(t,B_t) - F(0,B_0) = \sum_{k=0}^{m(\pi)-1} (F(t_{k+1}, B_{t_{k+1}}) - F(t_k, B_{t_k})).$$

By the mean value theorem

$$F(t_{k+1}, B_{t_{k+1}}) - F(t_k, B_{t_k}) = [F(t_{k+1}, B_{t_{k+1}}) - F(t_k, B_{t_{k+1}})] + [F(t_k, B_{t_{k+1}}) - F(t_k, B_{t_k})]$$
$$= \underbrace{[F(t_{k+1}, B_{t_{k+1}}) - F(t_k, B_{t_{k+1}})]}_{=a_k} + \underbrace{F'(t_k, B_{t_k})(B_{t_{k+1}} - B_{t_k})}_{=b_k} + \frac{1}{2}\underbrace{F''(t_k, \eta_k)(B_{t_{k+1}} - B_{t_k})^2}_{=c_k}$$

where η_k is a $\mathscr{F}_{t_{k+1}}$-measurable random variable that lies on the interval between B_{t_k} and $B_{t_{k+1}}$. Take a sequence of partitions π_n such that $\mathrm{mesh}(\pi_n) \to 0$ as $n \to \infty$. The claim is that the sums $\sum a_k$, $\sum b_k$ and $\sum c_k$ will converge to each of the three integrals in (2.8), respectively. The convergence will be almost sure along suitable subsequences of π_n. For the rest of the proof see Appendix B. □

2.3.1.1 An example

Example 2.2. Let $F(x) = x^2/2$ and let $(B_t)_{t\in\mathbb{R}_+}$ be a one-dimensional Brownian motion with $B_0 = 0$, then by Theorem 2.7

$$\frac{1}{2}B_t^2 = \int_0^t B_s \mathrm{d}B_s + \frac{1}{2}\int_0^t \mathrm{d}s$$

and hence after rearranging the terms

$$\int_0^t B_s \mathrm{d}B_s = \frac{1}{2}B_t^2 - \frac{1}{2}t$$

which is in agreement with the result we obtained by directly applying the definition of Itô integral.

2.3.2 *Itô's formula for semimartingales*

Henceforth, we'll write the time parameter of the integrands explicitly. Let's first study two stochastic integrals with respect to a common Brownian motion

$$X_t = \int_0^t f_s \mathrm{d}B_s, \quad Y_t = \int_0^t g_s \mathrm{d}B_s$$

where $f, g \in \mathscr{L}^2_{\mathrm{loc}}$. Their *(quadratic) covariation process* is defined as

$$\langle X, Y\rangle_t = \int_0^t f_s g_s \mathrm{d}s.$$

Then we notice that it satisfies the relation

$$4\langle X, Y\rangle_t = \langle X+Y\rangle_t - \langle X-Y\rangle_t.$$

A similar relation can be written for the product $X_t Y_t$ and a sum of the form $\sum_k (X_{t_{k+1}} - X_{t_k})(Y_{t_{k+1}} - Y_{t_k})$. Consequently, $X_t Y_t - \langle X, Y\rangle_t$ is a local martingale and along partitions of $[0,t]$

$$\lim_{\mathrm{mesh}(\pi)\to 0} \sum_{k=0}^{m(\pi)-1} (X_{t_{k+1}} - X_{t_k})(Y_{t_{k+1}} - Y_{t_k}) = \langle X, Y\rangle_t \tag{2.9}$$

in probability.

Let's then consider the case of two stochastic integrals with respect to independent Brownian motion. If $(B^{(1)}, B^{(2)})$ is a standard two-dimensional Brownian motion and

$$X_t = \int_0^t f_s \,\mathrm{d}B_s^{(1)}, \quad Y_t = \int_0^t g_s \,\mathrm{d}B_s^{(2)}$$

where $f,g \in \mathscr{L}^2_{\text{loc}}$, then X_tY_t is a local martingale. The covariation process is

$$\langle X,Y\rangle_t = 0$$

and it satisfies (2.9) together with $(X_t)_{t\in\mathbb{R}_{\geq 0}}$ and $(Y_t)_{t\in\mathbb{R}_{\geq 0}}$. These statements can be verified in the same manner as Theorem 2.6.

In the most general case, let $(B_t^{(1)}, B_t^{(2)}, \dots, B_t^{(m)})$ be a standard m-dimensional Brownian motion. Let

$$X_t = X_0 + \int_0^t f_s \mathrm{d}s + \sum_{k=1}^m \int_0^t g_s^{(k)}\,\mathrm{d}B_s^{(k)} \tag{2.10}$$
$$Y_t = Y_0 + \int_0^t \hat{f}_s \mathrm{d}s + \sum_{k=1}^m \int_0^t \hat{g}_s^{(k)}\,\mathrm{d}B_s^{(k)}$$

where X_0 and Y_0 are $\mathscr{F}_0$-measurable random variables, $g^{(k)}, \hat{g}^{(k)} \in \mathscr{L}^2_{\text{loc}}$ and $f, \hat{f}$ are measurable, adapted to $(F_t)_{t\in\mathbb{R}_+}$ and satisfy

$$\mathsf{P}\left[\int_0^t |f_s|\mathrm{d}s < \infty \text{ for all } t \in \mathbb{R}_+\right] = 1.$$

Then since integrals $\int_0^t f_s \mathrm{d}s$ have (locally) finite total variation, by the above it is natural to define

$$\langle X\rangle_t = \sum_{k=1}^m \int_0^t \left(g_s^{(k)}\right)^2 \mathrm{d}s, \quad \langle Y\rangle_t = \sum_{k=1}^m \int_0^t \left(\hat{g}_s^{(k)}\right)^2 \mathrm{d}s, \quad \langle X,Y\rangle_t = \sum_{k=1}^m \int_0^t g_s^{(k)}\hat{g}_s^{(k)}\,\mathrm{d}s$$

which are the quadratic variation and covariation processes also in the sense of Definition 2.6 and (2.9).

Definition 2.16. We call a process of the form (2.10) a *semimartingale* and use a shorthand notation

$$\mathrm{d}X_t = f_t \mathrm{d}t + \sum_{k=1}^m g_t^{(k)}\,\mathrm{d}B_t^{(k)}.$$

Remark 2.11. This is a slight abuse of standard terminology. More generally, semimartingale is any process that is sum of an adapted finite variation process and a local martingale.

Next we present a version of Itô's formula for semimartingales. An interesting viewpoint to this result is that the class of semimartingales is closed under forming new processes of the form $F(X_t^{(1)}, \dots, X_t^{(n)})$ from semimartingales $(X_t^{(k)})_{t\in\mathbb{R}_{\geq 0}}$, $k = 1, 2, \dots, n$.

Theorem 2.8 (Itô's formula for semimartingales). *Let $1 \leq l \leq n$. Let $X_t^{(j)}$ be semimartingales*

$$\mathrm{d}X_t^{(j)} = f_t^{(j)}\,\mathrm{d}t + \sum_{k=1}^m g_t^{(j,k)}\,\mathrm{d}B_t^{(k)}$$

for $1 \le j \le n$ where $f^{(j)}$ and $g^{(j,k)}$ are as above. Assume that $g^{(j,k)} = 0$ identically for $j > l$. Let $F : \mathbb{R}^n \to \mathbb{R}$ be continuous function such that $\partial_{x_i} F$ exists and is continuous for all $1 \le i \le n$ and that $\partial_{x_i x_j} F$ exists and is continuous for all $1 \le i, j \le l$.

Then $Y_t = F(X_t^{(1)}, \ldots, X_t^{(n)})$ is a semimartingale and almost surely

$$\mathrm{d}Y_t = \sum_{j=1}^{n} \left\{ \partial_{x_j} F(X_t^{(1)}, \ldots, X_t^{(n)}) f_t^{(j)}\, \mathrm{d}t + \sum_{k=1}^{m} \partial_{x_j} F(X_t^{(1)}, \ldots, X_t^{(n)}) g_t^{(j,k)}\, \mathrm{d}B_t^{(k)} \right\}$$
$$+ \frac{1}{2} \sum_{i,j=1}^{l} \sum_{k=1}^{m} \partial_{x_i, x_j} F(X_t^{(1)}, \ldots, X_t^{(n)})\, g_t^{(i,k)} g_t^{(j,k)}\, \mathrm{d}t$$

for all $t \in \mathbb{R}_+$. This is written shortly as

$$\mathrm{d}Y_t = \sum_{j=1}^{n} (\partial_j F)\, \mathrm{d}X_t^{(j)} + \frac{1}{2} \sum_{i,j=1}^{l} (\partial_{ij} F)\ \mathrm{d} \left\langle X^{(i)}, X^{(j)} \right\rangle_t .$$

Remark 2.12. Note that the theorem includes the case when F depends explicitly on time: let $l < n$ and take $X_t^{(n)} = t$. Theorem 2.7 is a special case of Theorem 2.8.

Remark 2.13 (Rules for stochastic calculus). Let $Y_t = F(X_t^{(1)}, \ldots, X_t^{(n)})$. Then the reader can memorize Itô's formula for Y_t by writing formally $Z_{t+\mathrm{d}t} = Z_t + \mathrm{d}Z_t$ for any semimartingale Z_t and then take the Taylor expansion of F at $(X_t^{(1)}, \ldots, X_t^{(n)})$ and then use the rules

$$\mathrm{d}t^2 = 0, \qquad \mathrm{d}t\, \mathrm{d}B_t^{(i)} = 0, \qquad \mathrm{d}B_t^{(i)}\, \mathrm{d}B_t^{(j)} = \delta_{ij} \mathrm{d}t.$$

2.3.2.1 An example

Example 2.3. Let $(B_t^{(1)}, \ldots, B_t^{(m)})$ be m-dimensional standard Brownian motion, $m \ge 2$, started from $(B_0^{(1)}, \ldots, B_0^{(m)}) \ne 0$ and let $F(x_1, \ldots, x_m) = \left(\sum_{k=1}^m x_k^2\right)^{1/2}$. Then by Itô's formula $Y_t = F(B_t^{(1)}, \ldots, B_t^{(n)})$ satisfies

$$\mathrm{d}Y_t = \sum_k \frac{B_t^{(k)} \mathrm{d}B_t^{(k)}}{Y_t} + \frac{m-1}{2Y_t} \mathrm{d}t.$$

2.4 Further topics in stochastic calculus

2.4.1 When is a semimartingale a local martingale?

The following result is based on the observation that $\int_0^t f_s \mathrm{d}s$ has finite total variation, where as for every continuous martingale it is infinite. The proof is given in

Appendix B. The result turns out to be extremely useful in conjunction with Itô's formula.

Lemma 2.1. *Let* $\mathrm{d}X_t = \sum g_t^{(k)} \mathrm{d}B_t^{(k)} + f_t \mathrm{d}t$ *be a semimartingale. Then it is a local martingale if and only if almost surely* $f_t = 0$ *for almost all* t.

Example 2.4. Let $F : \mathbb{R} \to \mathbb{R}$ be smooth, $\lambda \in \mathbb{R} \setminus \{0\}$ and suppose that $F(B_t)\, e^{\lambda t}$ is a martingale. By Itô's formula

$$\mathrm{d}\left(F(B_t)\, e^{\lambda t}\right) = F'(B_t)\, e^{\lambda t} \mathrm{d}B_t + \left(\lambda F(B_t) + \frac{1}{2} F''(B_t)\right) e^{\lambda t} \mathrm{d}t.$$

By Lemma 2.1, it holds that $\lambda F(B_t) + \frac{1}{2} F''(B_t) = 0$ for all t. This is possible only if F satisfies $\lambda F(x) + \frac{1}{2} F''(x) = 0$ for all x. Thus

$$\begin{aligned} F(x) &= C_1 \exp\left(\sqrt{-2\lambda} t\right) + C_2 \exp\left(-\sqrt{-2\lambda} t\right), && \text{when } \lambda < 0, \text{ and} \\ F(x) &= C_1 \sin(\sqrt{2\lambda} t) + C_2 \cos(\sqrt{2\lambda} t), && \text{when } \lambda > 0. \end{aligned}$$

Here $C_1, C_2 \in \mathbb{R}$ are constants.

2.4.2 *Time-changes*

2.4.2.1 Time-change of local martingales to Brownian motion

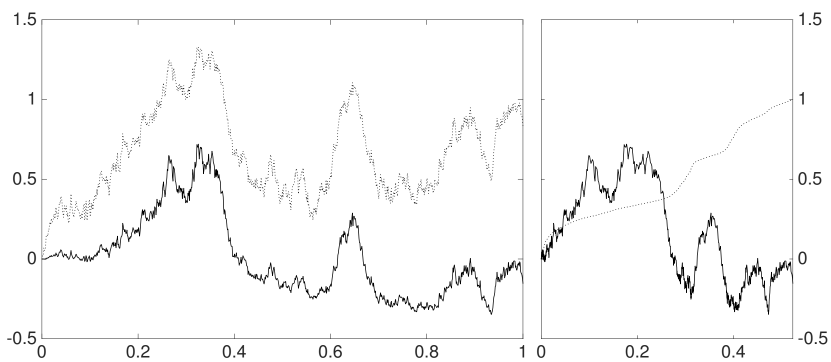


Fig. 2.2 On the left, an instance of Brownian motion $(B_t)_{t \in \mathbb{R}_{\geq 0}}$ is plotted with dots and the corresponding instance of $(X_t)_{t \in \mathbb{R}_{\geq 0}} = ((1/2)\,(B_t^2 - t))_{t \in \mathbb{R}_{\geq 0}}$ with a solid line. On the right, the time change of $(X_t)_{t \in \mathbb{R}_{\geq 0}}$ plotted with a solid line and the change of time, which is the inverse of the map $t \mapsto \langle X \rangle_t$, is plotted with dots.

As usual, let $(\Omega, \mathscr{F}, \mathsf{P})$ be a probability space with a filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_{\geq 0}}$ and let $(B_t)_{t\in\mathbb{R}_{\geq 0}}$ be a standard one-dimensional Brownian motion with respect to $(\mathscr{F}_t)_{t\in\mathbb{R}_{\geq 0}}$. Let's start this section by making the following definition.

Definition 2.17. If τ is a stopping time with respect to $(\mathscr{F}_t)_{t\in\mathbb{R}_{\geq 0}}$, define the *stopping time σ-algebra* as

$$\mathscr{F}_\tau = \{A \in \mathscr{F} \,:\, A \cap \{\tau \leq t\} \in \mathscr{F}_t \text{ for all } t \in \mathbb{R}_{\geq 0}\}$$

Remark 2.14. If $s \in \mathbb{R}_{\geq 0}$ is a constant and $\tau = s$ almost surely, then it's easy to check that $\mathscr{F}_\tau = \mathscr{F}_s$. So the notation $\mathscr{F}_\tau$ and the concept of stopping time σ-algebra is consistent with the earlier definitions.

In the same way, as $\mathscr{F}_t$ can be thought as the information available at time t, a stopping time σ-algebra $\mathscr{F}_\tau$ can be thought as the information available at a random time τ. The main reason to introduce the stopping time σ-algebra is time changes and analysis of martingales under time changes. See Appendix A for the optional stopping theorem which a result that that extends the martingale property from non-random time instances to stopping times.

The following theorem is an application of Itô's formula. It is a special case of more general result that *any continuous local martingale is a time-change of a Brownian motion*. The proof of the general result would follow the same lines if we had established the theory of the stochastic integral with respect to local martingales and we had corresponding Itô's formula available.

Theorem 2.9. *Let $(X_t)_{t\in\mathbb{R}_{\geq 0}}$ be a local martingale defined by*

$$X_t = \sum_{k=1}^{m} \int_0^t g_s^{(k)}\, \mathrm{d}B_s^{(k)}$$

where $g_t^{(k)} \in \mathscr{L}^2_{\mathrm{loc}}$. Let $(\sigma_r)_{r\in\mathbb{R}_{\geq 0}}$ be the set of stopping times

$$\sigma_r = \inf\{t \geq 0 \,:\, \langle X \rangle_t \geq r\}$$

where

$$\langle X \rangle_t = \sum_{k=1}^{m} \int_0^t \left(g_s^{(k)}\right)^2 \mathrm{d}s$$

is the quadratic variation process as before. Assume that almost surely $\langle X \rangle_t \to \infty$ as $t \to \infty$. Then the process $(Y_t)_{t\in\mathbb{R}_{\geq 0}}$ defined by

$$Y_t = X_{\sigma_t}$$

is a standard one-dimensional Brownian motion with respect to the filtration $(\mathscr{F}_{\sigma_t})_{t\in\mathbb{R}_{\geq 0}}$.

Proof. Since $\langle X \rangle_t \to \infty$ as $t \to \infty$, each σ_r is almost surely finite. By the continuity of the mapping $t \mapsto \langle X \rangle_t$, we have that $\langle X \rangle_{\sigma_r} = r$.

Let

$$M_t = \exp\left(i\theta X_t + \frac{\theta^2}{2}\langle X\rangle_t\right).$$

By Itô's formula $(M_t)_{t\in\mathbb{R}_{\geq 0}}$ is a continuous local martingale, see also Example 2.4. Note that $(M_t)_{t\in\mathbb{R}_{\geq 0}}$ is a complex valued process, but this causes no problems: we can apply Itô's formula separately for its real and imaginary parts. The statement that it is a local martingale means that both its real and imaginary parts are local martingales. Since $M_t^{\sigma_r} = M_{t\wedge\sigma_r}$ is bounded, $(M_t^{\sigma_r})_{t\in\mathbb{R}_{\geq 0}}$ is a martingale. Namely, if τ_n is the localizing sequence of $(M_t)_{t\in\mathbb{R}_{\geq 0}}$, then $(M_t^{\sigma_r\wedge\tau_n})_{t\in\mathbb{R}_{\geq 0}}$ is a martingale. Hence by boundedness of $(M_t^{\sigma_r})_{t\in\mathbb{R}_{\geq 0}}$ and by the fact that $\tau_n \nearrow \infty$ almost surely as $n\to\infty$,

$$\mathsf{E}[\underbrace{M_t^{\sigma_r\wedge\tau_n}}_{\to M_t^{\sigma_r} \text{ in } L^1} |\mathscr{F}_s] = \underbrace{M_s^{\sigma_r\wedge\tau_n}}_{\to M_s^{\sigma_r} \text{ in } L^1\text{, as } n\to\infty}$$

and therefore by properties of conditional expected value, see Appendix A,

$$\mathsf{E}[M_t^{\sigma_r}|\mathscr{F}_s] = M_s^{\sigma_r}.$$

Thus $(M_t^{\sigma_r})_{t\in\mathbb{R}_{\geq 0}}$ is a continuous bounded martingale.

Next we apply the optional stopping theorem for stopping times $\sigma_s \leq \sigma_r$, where $0\leq s\leq r$, to show that

$$\mathsf{E}[M_{\sigma_r}|\mathscr{F}_{\sigma_s}] = M_{\sigma_s}.$$

This implies that for any $0\leq s\leq r$ and for any $\theta\in\mathbb{R}$,

$$\mathsf{E}\left[\exp\left(i\theta(X_{\sigma_r}-X_{\sigma_s})\right)|\mathscr{F}_{\sigma_s}\right] = \exp\left(-\frac{\theta^2}{2}(r-s)\right).$$

The right-hand side of this equation is the characteristic function of a normal random variable with mean 0 and variance $r-s$. The left-hand side is a conditional version of characteristic function of $X_{\sigma_r}-X_{\sigma_s}$. That characteristic function is now constant as a $\mathscr{F}_{\sigma_s}$-measurable random variable. Therefore the fact that the characteristic function determines the distribution uniquely shows that $X_{\sigma_r}-X_{\sigma_s}$ is independent from $\mathscr{F}_{\sigma_s}$ and that $X_{\sigma_r}-X_{\sigma_s}$ is normally distributed with mean 0 and variance $r-s$.[5] □

Example 2.5. Let us continue the setup of Example 2.3. Let $(W_t)_{t\in\mathbb{R}_{\geq 0}}$ be a process defined by W_0 and $\mathrm{d}W_t = \sum_{k=1}^n (B_t^{(k)}/Y_t)\mathrm{d}B_t^{(k)}$, where $Y_t = F(B_t^{(1)},\ldots,B_t^{(n)})$. Then

$$\langle W\rangle_t = \sum_{k=1}^n \int_0^t \frac{(B_s^{(k)})^2}{Y_s^2}\mathrm{d}s = t.$$

[5] If $\mathsf{E}[\exp(\mathrm{i}\theta_1 X)|\mathscr{G}] = \psi(\theta_1)$ is a constant as a function of $\omega\in\Omega$, then for any $\mathscr{G}$-measurable random variable Z, it holds that $\phi_{(X,Z)}(\theta_1,\theta_2) = \mathsf{E}[\exp(\mathrm{i}\theta_1 X + \mathrm{i}\theta_2 Z)] = \mathsf{E}[\exp(\mathrm{i}\theta_2 Z)\,\mathsf{E}[\exp(\mathrm{i}\theta_1 X)|\mathscr{G}]] = \psi(\theta_1)\phi_Z(\theta_2)$ by properties of conditional expected value. Here $\phi_{\underline{\theta}} = \mathsf{E}[\exp(\mathrm{i}\underline{\theta}\cdot\underline{Y})]$, $\underline{\theta}\in\mathbb{R}^n$ is the characteristic function of a $\mathbb{R}^n$-valued random variable $\underline{Y}$. Thus by the uniqueness theorem of the characteristic function, it follows that X and Z are independent and that the law of X is the unique probability measure on $\mathbb{R}$ such that $\phi_X = \psi$. Since this holds in particular for any random variable $Z = \mathbb{1}_E$, $E\in\mathscr{G}$, it follows that X is independent from $\mathscr{G}$.

By Theorem 2.9, $(W_t)_{t\in\mathbb{R}_{\geq 0}}$ is a $(\mathscr{F}_t)_{t\in\mathbb{R}_{\geq 0}}$-Brownian motion.

2.4.2.2 Time-change of semimartingales

The next result gives a general form of a time-change for semimartingales. The proof is left as an exercise.

Proposition 2.6. *Let $a_t(\omega)$ be a continuous, positive, adapted process. Define a random time change by setting:*

$$S(t,\omega) = \int_0^t a_r(\omega)^2 \,\mathrm{d}r, \qquad \sigma(s,\omega) = \inf\{t \in \mathbb{R}_{\geq 0} : S(t,\omega) \geq s\}$$

Let $(\tilde{B}_s)_{s\in\mathbb{R}_{\geq 0}}$ be the process defined by

$$\tilde{B}_s(\omega) = \int_0^{\sigma(s)} a_r \,\mathrm{d}B_r(\omega).$$

Then $(\tilde{B}_s)_{s\in\mathbb{R}_{\geq 0}}$ is a standard one-dimensional Brownian motion with respect to $(\mathscr{F}_{\sigma(s)})_{s\in\mathbb{R}_{\geq 0}}$, and for any continuous, adapted process $v_t(\omega)$ the following time-change formula holds

$$\int_0^s v_{\sigma(q)} \,\mathrm{d}\tilde{B}_q = \int_0^{\sigma(s)} v_r a_r \,\mathrm{d}B_r.$$

Moreover if X_t is a semimartingale $\mathrm{d}X_t = u_t \mathrm{d}t + v_t \mathrm{d}B_t$ then the process $(\tilde{X}_s)_{s\in\mathbb{R}_{\geq 0}}$ defined by $\tilde{X}_s = X_{\sigma(s)}$ is a semimartingale with respect to $(\mathscr{F}_{\sigma(s)})_{s\in\mathbb{R}_{\geq 0}}$ and $(\tilde{B}_s)_{s\in\mathbb{R}_{\geq 0}}$ and satisfies

$$\mathrm{d}\tilde{X}_s = \frac{u_{\sigma(s)}}{a^2_{\sigma(s)}}\mathrm{d}s + \frac{v_{\sigma(s)}}{a_{\sigma(s)}}\mathrm{d}\tilde{B}_s.$$

2.4.3 Stochastic differential equations

We present here the rudiments of stochastic differential equations for single variable. The multidimensional version, which is a straightforward extension, is presented in Appendix B. For one-dimensional stochastic differential equations, one can obtain results under weaker conditions, see, for instance, [6] Section IX.3.

Let $(X_t)_{t\in[0,T]}$ be an $\mathbb{R}$-valued continuous stochastic process and let $(B_t)_{t\in\mathbb{R}_{\geq 0}}$ be a standard one-dimensional Brownian motion. We say that X_t satisfies the *stochastic differential equation* (SDE)

$$\mathrm{d}X_t = F(t,X_t)\mathrm{d}t + G(t,X_t)\mathrm{d}B_t \tag{2.11}$$

with initial value $X_0 = Z$ if for each $t \in [0,T]$

$$X_t = Z + \int_0^t F(s, X_s)\,\mathrm{d}s + \int_0^t G(s, X_s)\,\mathrm{d}B_s.$$

If the process can be constructed in a given probability space with a given filtration and Brownian motion, then $(X_t)_{t\in[0,T]}$ is called *strong solution* of the SDE.

Theorem 2.10. *Let $(B_t)_{t\in\mathbb{R}_{\geq 0}}$ be one-dimensional Brownian motion and let*

$$F : [0,T] \times \mathbb{R} \to \mathbb{R}$$
$$G : [0,T] \times \mathbb{R} \to \mathbb{R}$$

be measurable maps. Let Z be $\mathbb{R}$-valued square integrable random variable which is independent from $\sigma(B_t, t \in \mathbb{R}_{\geq 0})$. Suppose that

$$|F(t,x)| + |G(t,x)| \leq C(1+|x|)$$
$$|F(t,x) - F(t,y)| + |G(t,x) - G(t,y)| \leq D\,|x-y|.$$

Then there exist a unique continuous solution $(X_t)_{t\in[0,T]}$ to the stochastic differential equation (2.11) *with initial value $X_0 = Z$ with the property that X_t is adapted to the filtration $\mathscr{F}_t^{(B,Z)}$ generated by Z and B_s, $s \in [0,t]$. Furthermore*

$$\mathsf{E}\left[\int_0^T |X_t|^2 \mathrm{d}t\right] < \infty.$$

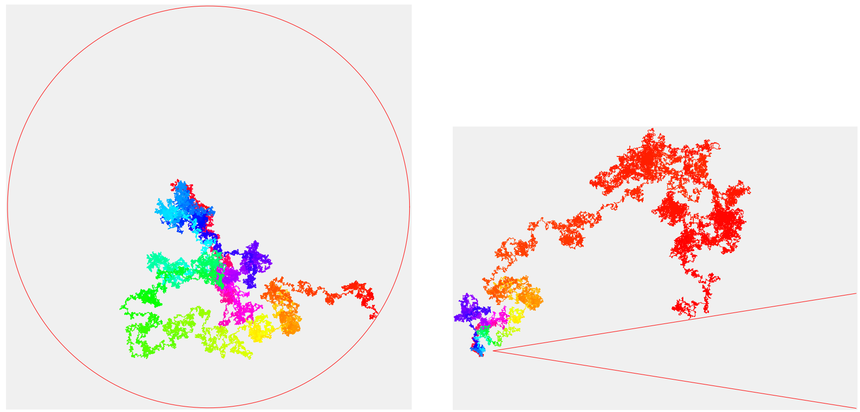
The proof of the theorem is very similar to the proofs of existence and uniqueness of solutions of ordinary differential equations and is based on Picard–Lindelöf iteration. We leave the details to the reader.

Example 2.6. Let us continue Examples 2.3 and 2.5. The solution of the SDE

$$\mathrm{d}X_t = \mathrm{d}B_t + \frac{\delta - 1}{2X_t}\mathrm{d}t,$$

$X_0 = x$ is called a *Bessel process* of dimension $\delta \in \mathbb{R}$ sent from x. In Examples 2.3 and 2.5 we saw that the norm of n-dimensional Brownian motion is a n-dimensional Bessel process. We can use Theorem 2.10 with Proposition 2.5 to show that the solution exists and is unique for all $\delta \in \mathbb{R}$ up to the time $\tau = \inf\{t \in \mathbb{R}_{\geq 0} : \inf_{s\in[0,t)} X_s = 0\}$ which is the hitting time of 0. Using other methods, we could define it beyond the hitting of 0 for the parameter values $\delta > 0$.

Remark 2.15. In the time-homogeneous case, $F(t,x) = F(x)$ and $G(t,x) = G(x)$, these solutions X_t are called *diffusions*. Another viewpoint to diffusions is that it is a family of processes with one element for each starting point $x \in \mathbb{R}$.



(a) A sample of complex Brownian motion stopped at its exit time from the disc.

(b) Conformal image of the Brownian sample under a map that takes the disc onto a sector.

Fig. 2.3 Illustration of conformal invariance of Brownian motion. The colors indicate time. We notice that the appearances of the paths are similar in both pictures except that the time is changed by a local factor when we move from the first picture to the second one.

2.5 Conformal invariance of two-dimensional Brownian motion

As usual complex number z is represented in terms of its real and imaginary parts as $z = x + \mathrm{i}\, y$, similarly complex valued function of a complex variable is divided into its real and imaginary parts as $f(z) = u(z) + \mathrm{i}\, v(z)$. Define as usual the following partial differential operators

$$\partial = \frac{1}{2}(\partial_x - \mathrm{i}\partial_y), \qquad \overline{\partial} = \frac{1}{2}(\partial_x + \mathrm{i}\partial_y).$$

Let U be an open set in the complex plane $\mathbb{C}$ and let $z_0 \in U$. The starting point of complex analysis is that the following statements about a function $f : U \to \mathbb{C}$ are equivalent:

- The function f is *holomorphic* near z_0: the complex derivative

  $$f'(z) = \lim_{h \to 0} \frac{f(z+h) - f(z)}{h}$$

  exists and is continuous in a neighborhood of z_0. This is equivalent to that the statement that f has continuous partial derivatives $\partial_x f$, $\partial_y f$ and satisfies $\overline{\partial} f(z) = 0$ in a neighborhood of z_0. The complex derivative f' satisfies $f'(z) = \partial f(z) = \partial_x f(z) = -\mathrm{i}\partial_y f(z)$.
- The real and imaginary parts of f satisfies *Cauchy–Riemann equations* near z_0:

  $$\partial_x u = \partial_y v, \qquad \partial_x v = -\partial_y u$$

- The function f is *(complex) analytic* at z_0: $f(z) = \sum_{n=0}^{\infty} c_n(z-z_0)^n$ which converges absolutely when $|z-z_0| \leq r$ for some $r > 0$.

Remember that u and v are *harmonic*: $\Delta u = 0 = \Delta v$.[6]

We conclude this chapter by showing that the complex Brownian motion is conformally invariant (up to a time-change). This justifies more or less all the time that we invested on the technical steps in this chapter and also works as motivation for the treatise of conformally invariant scaling limit in later chapters. Define a complex Brownian motion send from $z_0 \in \mathbb{C}$ as

$$B_t = B_t^{\mathbb{C}} = z_0 + B_t^{(1)} + \mathrm{i}\, B_t^{(2)}$$

Theorem 2.11. *Let $U \subset \mathbb{C}$ be a domain (non-empty connected open set) and let $f : U \to \mathbb{C}$ be analytic. Let $z_0 \in U$ and let B_t be a complex Brownian motion send from $z_0 \in \mathbb{C}$. Let $\tau = \inf\{t \geq 0 : B_t \notin U\}$. Let $Z_t = f(B_{\sigma(t)})$ for $0 \leq t < S(\tau)$ where $\sigma(t) = S^{-1}(t)$ and*

$$S(t) = \int_0^t |f'(B_s)|^2 \mathrm{d}s$$

for $0 \leq 0 < \tau$. Then Z_t is a complex Brownian motion send from $f(z_0)$ and stopped at $S(\tau)$.

Proof. As above write $f = u + iv$. Define

$$X_t = u(B_t), \qquad Y_t = v(B_t)$$

Since u and v are harmonic and satisfy the Cauchy–Riemann equations,

$$\mathrm{d}X_t = u_1(B_t)\mathrm{d}B_t^{(1)} + u_2(B_t)\mathrm{d}B_t^{(2)}$$
$$\mathrm{d}Y_t = -u_2(B_t)\mathrm{d}B_t^{(1)} + u_1(B_t)\mathrm{d}B_t^{(2)}$$

by Itô's formula, where $u_1 = \partial_x u$ and $u_2 = \partial_y u$ are the partial derivatives of u. The dt-terms vanished by $\Delta u = 0 = \Delta v$. Therefore $(X_t)_{t\in\mathbb{R}_+}$ and $(Y_t)_{t\in\mathbb{R}_+}$ are local martingales.

Now $\langle X, Y\rangle_t = 0$ and

$$\langle X\rangle_t = \langle Y\rangle_t = \int_0^t u_1(B_s)^2 + u_2(B_s)^2 \,\mathrm{d}s = \int_0^t |f'(B_s)|^2 \mathrm{d}s.$$

Here we used the fact that $f'(z) = u_1(z) - i u_2(z)$. A slight modification of the proof of Theorem 2.9 shows that for any $\theta_1, \theta_2 \in \mathbb{R}$

$$\exp\left(i\theta_1 X_t + \frac{\theta_1^2}{2}\langle X\rangle_t\right) \exp\left(i\theta_2 Y_t + \frac{\theta_2^2}{2}\langle Y\rangle_t\right)$$

[6] By the Cauchy–Riemann equations, $u_{xx} + u_{yy} = v_{xy} - v_{xy} = 0$ and similarly for v.

is a local martingale, and consequently, $(X_{\sigma_t})_{t\in\mathbb{R}_+}$ and $(Y_{\sigma_t})_{t\in\mathbb{R}_+}$ are independent Brownian motions. □

Remark 2.16. In the previous proof, it was crucial that $(X_t)_{t\in\mathbb{R}_+}$ and $(Y_t)_{t\in\mathbb{R}_+}$ had the same quadratic variation. There is no general time-change result for multidimensional continuous local martingales of the form of Theorem 2.9.

References

1. Durrett, R.: Stochastic calculus. a practical introduction. CRC (1996)
2. Durrett, R.: Probability: theory and examples, fourth edn. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge (2010)
3. Kallenberg, O.: Foundations of modern probability. Springer Verlag (2002)
4. Karatzas, I., Shreve, S.E.: Brownian motion and stochastic calculus, *Graduate Texts in Mathematics*, vol. 113, second edn. Springer-Verlag, New York, New York, NY (1991)
5. Øksendal, B.: Stochastic differential equations, sixth edn. Universitext. Springer-Verlag, Berlin, Berlin, Heidelberg (2003)
6. Revuz, D., Yor, M.: Continuous martingales and Brownian motion, *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 293, third edn. Springer-Verlag, Berlin (1999)
7. Steele, J.M.: Stochastic calculus and financial applications, *Applications of Mathematics (New York)*, vol. 45. Springer-Verlag, New York, New York, NY (2001)