Chapter 5

Itô Calculus

Itô calculus has profound applications in mathematics and mathematical finance. In this chapter, we construct the Itô integral on Brownian paths. The mathematical significance of Itô's work is to have given a rigorous meaning to an integral of random functions (like Brownian motion) whose paths do not have bounded variation; see Corollary 3.17. The construction is done in Section 5.3 as a limit of *martingale transforms*, which is the equivalent of the Riemann sum for stochastic integrals as explained in Section 5.2. One of the upshots is that the Itô integral gives a systematic way to construct Brownian martingales. We also derive Itô's formula in Section 5.4. The formula relates the Itô integral to explicit functions of Brownian motion. As such, it can be considered as the fundamental theorem of Itô calculus. As a point of comparison, it is useful to briefly go back to the integral of standard calculus: the Riemann integral.

5.1. Preliminaries

The construction of the classical Riemann integral goes as follows. Consider, for example, a continuous function g on $[0, t]$. We take a partition of $[0, t]$ in n intervals $(t_j, t_{j+1}]$ with $t_n = t$; for example $t_j = \frac{j}{n}t$. The Riemann integral is understood as the limit of Riemann sums:

$$\int_0^t g(s)\,\mathrm{d}s = \lim_{n\to\infty} \sum_{j=0}^{n-1} g(t_j)(t_{j+1} - t_j).$$

Note that the integral is a number for fixed t. The integral represents the area under the curve given by g on the interval $[0, t]$. It can also be seen as a continuous function of t as t varies on an interval. In fact, as a function of t, the integral is differentiable and its derivative is g. This is the fundamental theorem of calculus.

It is possible to modify the above definition slightly for more general increments. The construction is called the *Riemann-Stieltjes integral*. Let F be a function on $[0, t]$ of bounded variation, as in Example 3.6. It can be shown that the integral as a limit of

Riemann sums with the increments of F exists:

$$\int_0^t g(s)\,\mathrm{d}F(s) = \lim_{n\to\infty} \sum_{j=0}^{n-1} g(t_j)(F(t_{j+1}) - F(t_j)). \tag{5.1}$$

If F is a CDF of a random variable X, then $\int_{-\infty}^{\infty} g(s)\,\mathrm{d}F(s)$ represents the expectation of the random variable $g(X)$; see Remark 1.25. Note that $F(t_{j+1}) - F(t_j)$ is the probability that X falls in the interval $(t_j, t_{j+1}]$.

The goal is to make sense of the above when F is replaced by a Brownian motion $(B_t, t \geq 0)$:

$$\int_0^t g(s)\,\mathrm{d}B_s = \lim_{n\to\infty} \sum_{j=0}^{n-1} g(t_j)(B_{t_{j+1}} - B_{t_j}).$$

The major hurdle here is not the fact that the Brownian paths are random, but instead that these paths have *unbounded variation*, as proved in Corollary 3.17. This means that the classical construction does not apply for a given path. Therefore, another a priori construction is needed. The Poisson process has paths of bounded variation, as they are increasing. There is no problem in using the classical construction of the integral for Poisson paths.

Note that the sum $\sum_{j=0}^{n-1} g(t_j)(B_{t_{j+1}} - B_{t_j})$ above is a random variable. If the endpoint $t_n = t$ is varied, it can be seen as a stochastic process. Moreover, since the Brownian paths are continuous, this new stochastic process also has continuous paths. As we shall see, this stochastic process is in fact a continuous martingale like Brownian motion. It turns out that these properties remain in the limit $n \to \infty$.

What is the interpretation of the stochastic integral? If we think of $(B_t, t \geq 0)$ as modelling the price of a stock, then $\sum_{j=0}^{n-1} g(t_j)(B_{t_{j+1}} - B_{t_j})$ gives the value of a portfolio at time t that implements the following strategy: At t_j we buy $g(t_j)$ shares of the stock that we sell at time t_{j+1}. We do this for every $j \leq n-1$. The net gain or loss of this strategy is the sum over j of $g(t_j)(B_{t_{j+1}} - B_{t_j})$. Of course, in this interpretation, the number of shares $g(t_j)$ put in play could be random and depend on the past information of the path up to time t_j.

In the next section, we take a first step towards the Itô integral by defining the *martingale transform*. The construction makes sense for any square-integrable martingale.

5.2. Martingale Transform

Let $(M_t, t \leq T)$ be a continuous square-integrable martingale on $[0,T]$ for the filtration $(\mathcal{F}_t, t \leq T)$, defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The idea of the martingale transform is to modify the amplitude of each increment in such a way as to produce a martingale when these new increments are summed up. The martingale transforms are to the Itô integral what Riemann sums are for the Riemann integral.

More precisely, let $(t_j, j \leq n)$ be a sequence of partitions of $[0,T]$ with $t_0 = 0$ and $t_n = T$. For example, we can take $t_j = \frac{j}{n}T$. Consider n fixed numbers $(Y_0, Y_1, \ldots, Y_{n-1})$. It is convenient to construct a function of time X_t from these:

$$X_t = Y_j \qquad \text{if } t \in (t_j, t_{j+1}].$$

This can be written also as a sum of indicator functions:

$$X_t = \sum_{j=0}^{n-1} Y_j \mathbf{1}_{(t_j, t_{j+1}]}(t), \quad t \leq T. \tag{5.2}$$

The integral of $(X_t, t \leq T)$ with respect to the martingale M on $[0, T]$, also called a *martingale transform*, is the sum of the increments of the martingale modulated by X; i.e.,

$$I_T = Y_0(M_{t_1} - M_0) + \cdots + Y_{n-1}(M_T - M_{t_{n-1}}) = \sum_{j=0}^{n-1} Y_j(M_{t_{j+1}} - M_{t_j}). \tag{5.3}$$

This is a random variable in $L^2(\Omega, \mathcal{F}, \mathbf{P})$, since it is a linear combination of random variables in L^2. Note that we recover M_T when X_{t_j} is 1 for all intervals. We may think of $(M_t, s \leq T)$ as the price of an asset, say a stock, on a time interval $[0, T]$. Then the term

$$Y_j(M_{j+1} - M_{t_j})$$

can be seen as the gain or loss in the time interval $(t_j, t_{j+1}]$ of buying Y_j units of the asset at time t_j at price M_{t_j} and selling these at time t_{j+1} at price $M_{t_{j+1}}$. Summing these terms over time gives the value I_t of implementing the *investing strategy* X on the interval $[0, T]$. It is not hard to modify the definition to obtain a stochastic process on the whole interval $[0, T]$. For $t \leq T$, we simply sum the increments up to t. This can be written down as

$$I_t = Y_0(M_{t_1} - M_0) + Y_1(M_{t_2} - M_{t_1}) + \cdots + Y_j(M_t - M_{t_j}), \quad \text{if } t \in (t_j, t_{j+1}]. \tag{5.4}$$

Example 5.1 (Integral of a simple process)**.** Consider a standard Brownian motion $(B_t, t \in [0, 1])$ on the time interval $[0, 1]$. We know very well by now that it is a martingale. We look at a simple integral constructed from it. We take the following integrand:

$$X_t = \begin{cases} 10 & \text{if } t \in [0, 1/3], \\ 5 & \text{if } t \in (1/3, 2/3], \\ 2 & \text{if } t \in (2/3, 1]. \end{cases}$$

Then the integrals I_t as in equation (5.4) form a process $(I_t, t \in [0, 1])$ of the form

$$I_t = \begin{cases} 10B_t & \text{if } t \in [0, 1/3], \\ 10B_{1/3} + 5(B_t - B_{1/3}) & \text{if } t \in (1/3, 2/3], \\ 10B_{1/3} + 5(B_{2/3} - B_{1/3}) + 2(B_t - B_{2/3}) & \text{if } t \in (2/3, 1]. \end{cases}$$

We make three important observations. First, the paths of the process $(I_t, t \in [0, 1])$ are continuous, because Brownian paths are. Second, the process is a square-integrable martingale. It is easy to see that it is adapted and square-integrable, because I_t is a sum of square-integrable random variables. The martingale property is also not hard to verify. For example, we have for $t \in (2/3, 1]$,

$$\mathbf{E}[I_t | \mathcal{F}_{2/3}] = 10B_{1/3} + 5(B_{2/3} - B_{1/3}) + 2\mathbf{E}[B_t - B_{2/3} | \mathcal{F}_{2/3}] = I_{2/3},$$

since $\mathbf{E}[B_t - B_{2/3} | \mathcal{F}_{2/3}] = 0$ by the martingale property of Brownian motion. See Exercise 5.2 for more on this example.

We can generalize the integrand or investing strategy X by considering values X_{t_j} that depend on the process, hence are random, but in a *predictable way*. Namely, we can take X to be a random vector such that X_{t_j} is $\mathcal{F}_{t_j}$-measurable. In other words, X_{t_j} may be random but must only depend on the information up to time t_j. Common sense dictates that the number of shares you buy today should not depend on information in the future. With this in mind, for a given filtration, we define the space of *simple* (that is, discrete) *adapted* processes on $[0, T]$:

(5.5)
$$\mathcal{S}(T) = \left\{ (X_t, t \leq T) : X_t = \sum_{j=0}^{n-1} Y_j \mathbf{1}_{(t_j, t_{j+1}]}(t), Y_j \text{ is } \mathcal{F}_{t_j}\text{-measurable}, \mathbf{E}[Y_j^2] < \infty \right\}.$$

In words, the processes in $\mathcal{S}(T)$ have paths that are piecewise constant on a finite number of intervals of $[0, T]$. The values $Y_j(\omega)$ on each time interval might vary depending on the paths ω. As random variables, the Y_j's depend only on the information up to time t_j and have finite second moment: $\mathbf{E}[Y_j^2] < \infty$. Note that $\mathcal{S}(T)$ is a linear space: If $X, X' \in \mathcal{S}(T)$, then $aX + bX' \in \mathcal{S}(T)$ for $a, b \in \mathbb{R}$. Indeed, if the paths of X, X' take a finite number of values, then so are the ones of $aX + bX'$.

Example 5.2 (An example of simple adapted process)**.** Let $(B_t, t \leq 1)$ be a standard Brownian motion. For the interval $[0, 1]$, consider the investing strategy X in $\mathcal{S}(1)$ given by the position of the Brownian path at times $0, 1/3, 2/3$:

$$X_s = \begin{cases} 0 & \text{if } s \in [0, 1/3], \\ B_{1/3} & \text{if } s \in (1/3, 2/3], \\ B_{2/3} & \text{if } s \in (2/3, 1]. \end{cases}$$

Clearly, X is simple and adapted to the Brownian filtration. For example, the value at $s = 3/4$ is $B_{2/3}$. In particular, it depends only on the information prior to time 3/4. See Figure 5.1.

For a simple adapted process X, the integral I_t of X with respect to the martingale $(M_t, t \leq T)$ is the same as in equation (5.4).

Definition 5.3. Let $(M_t, t \leq T)$ be a continuous square-integrable martingale for the filtration $(\mathcal{F}_t, t \leq T)$. Let $X \in \mathcal{S}(T)$ be a simple, adapted process $X = \sum_{j=0}^{n-1} Y_j \mathbf{1}_{(t_j, t_{j+1}]}$ on $[0, T]$. The *martingale transform* I_t is

$$I_t = Y_0(M_{t_1} - M_0) + Y_1(M_{t_2} - M_{t_1}) + \cdots + Y_j(M_t - M_{t_j}), \quad \text{if } t \in (t_j, t_{j+1}].$$

It defines a process $(I_t, t \leq T)$ on $[0, T]$.

Example 5.4 (Another integral of a simple process)**.** Consider the simple process X of Example 5.2 defined on a Brownian motion. The integral of X as a process on $[0, 1]$ is

$$I_s = \begin{cases} 0 & \text{if } s \in [0, 1/3], \\ B_{1/3}(B_s - B_{1/3}) & \text{if } s \in (1/3, 2/3], \\ B_{1/3}(B_{2/3} - B_{1/3}) + B_{2/3}(B_s - B_{2/3}) & \text{if } s \in (2/3, 1]. \end{cases}$$

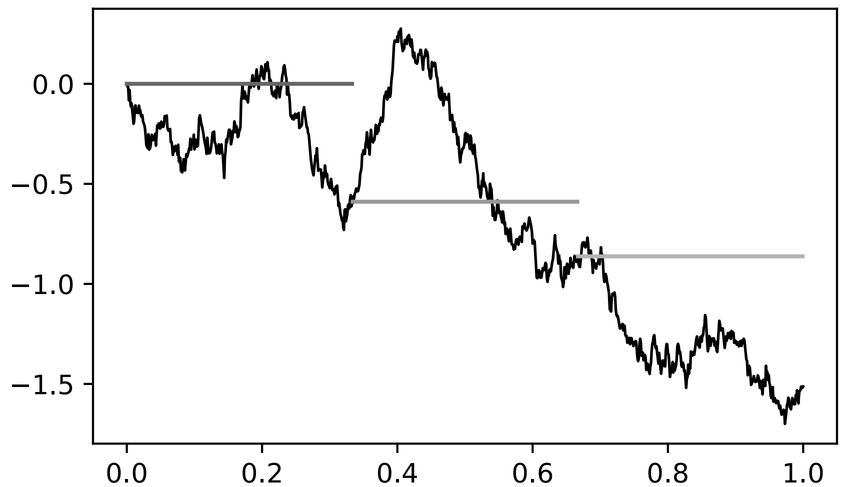

Figure 5.1. The simple process X constructed from a Brownian path in Example 5.2. Note that the path of X is piecewise constant. However, the value on each piece is random as it depends on the positions of the Brownian path at time 1/3 and 2/3.

As in Example 5.1, the paths of I_s are continuous for all $s \in [0,1]$, since the paths of B_s are continuous! This is also true at the integer times $s = 1/3, 2/3$, if we approached on the left or right. The process $(I_s, s \leq 1)$ is also a martingale for the Brownian filtration. The key here is that the value multiplying the increment on the interval $(t_j, t_{j+1}]$ is $\mathcal{F}_{t_j}$-measurable. For example, take $t > 2/3$ and $1/3 < s < 2/3$. The properties of conditional expectation in Proposition 4.19 and the fact that Brownian motion is a martingale give

$$\begin{aligned}
\mathbf{E}[I_t|\mathcal{F}_s] &= \mathbf{E}[B_{1/3}(B_{2/3} - B_{1/3}) + B_{2/3}(B_t - B_{2/3})|\mathcal{F}_s] \\
&= \mathbf{E}[B_{1/3}(B_{2/3} - B_{1/3})|\mathcal{F}_s] + \mathbf{E}[B_{2/3}(B_t - B_{2/3})|\mathcal{F}_s] \\
&= B_{1/3}(B_s - B_{1/3}) + \mathbf{E}[\ \mathbf{E}[B_{2/3}(B_t - B_{2/3})|\mathcal{F}_{2/3}]\ |\mathcal{F}_s] \\
&= B_{1/3}(B_s - B_{1/3}) + \mathbf{E}[B_{2/3}\ \mathbf{E}[(B_t - B_{2/3})|\mathcal{F}_{2/3}]\ |\mathcal{F}_s] \\
&= B_{1/3}(B_s - B_{1/3}) + 0 = I_s.
\end{aligned}$$

Note that it was crucial to use the tower property in the third equality and that we took out what is known at $t = 2/3$ in the fourth equality.

Martingale transforms are always themselves martingales. In particular, it is not possible in this setup to design an investing strategy whose value would be increasing on average.

Proposition 5.5 (Martingale transforms are martingales). *Let $(M_t, t \leq T)$ be a continuous square-integrable martingale for the filtration $(\mathcal{F}_t, t \leq T)$ and let $X \in \mathcal{S}(T)$ be a simple process as in equation* (5.5)*. Then the martingale transform $(I_t, t \leq T)$ is a continuous martingale on* $[0,T]$ *for the same filtration.*

Proof. The fact that I_t is $\mathcal{F}_t$-measurable for $t \leq T$ is clear from the construction in equation (5.4). Indeed, the increments $M_{t_{j+1}} - M_{t_j}$ are $\mathcal{F}_t$-measurable for $t_{j+1} \leq t$ since the martingale is adapted. The integrand X is also adapted. Moreover, I_t is integrable since

$$\mathbf{E}[|I_t|] \le \mathbf{E}[|I_T|] \le \sum_{j=0}^{n-1} \mathbf{E}[|Y_j||M_{t_{j+1}} - M_{t_j}|] \le \sum_{j=0}^{n-1} (\mathbf{E}[Y_j^2])^{1/2} (\mathbf{E}[(M_{t_{j+1}} - M_{t_j})^2])^{1/2},$$

by the Cauchy-Schwarz inequality. The last term is finite by assumption on X and M. As for continuity, since $(M_t, t \le T)$ is continuous, the only possible issue could be at the points t_j for some j. But in that case, we have for $t > t_j$ but close and any outcome ω,

$$I_t(\omega) = \sum_{i=0}^{j-1} Y_i(M_{t_{i+1}}(\omega) - M_{t_i}(\omega)) + Y_j(M_t(\omega) - M_{t_j}(\omega)),$$

which goes to I_{t_j} when $t \to t_j^+$ by continuity of $M_t(\omega)$. A similar argument holds for $t \to t_j^-$.

To prove the martingale property, consider $s < t$. We want to show that $\mathbf{E}[I_t|\mathcal{F}_s] = I_s$. Suppose $t \in (t_j, t_{j+1}]$ for some $t_j < T$. By linearity of conditional expectation in Proposition 4.19, we have

$$\mathbf{E}[I_t|F_s] = \sum_{i=0}^{j} \mathbf{E}[Y_i(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_s], \tag{5.6}$$

where it is understood that $t = t_{j+1}$ in the above to simplify notation. We can now handle each summand. There are three possibilities: $s \ge t_{i+1}$, $s \in (t_i, t_{i+1})$, and $s < t_i$. It all depends on Proposition 4.19. In the case $s \ge t_{i+1}$, we have

$$\mathbf{E}[Y_i(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_s] = Y_i(M_{t_{i+1}} - M_{t_i}),$$

since the whole summand is $\mathcal{F}_s$-measurable. In the case $s \in (t_i, t_{i+1})$, we have that Y_i is $\mathcal{F}_s$-measurable; therefore

$$\mathbf{E}[Y_i(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_s] = Y_i \mathbf{E}[M_{t_{i+1}} - M_{t_i}|\mathcal{F}_s] = Y_i(M_s - M_{t_i}),$$

by the martingale property. In the case $s < t_i$, we use the tower property to get

$$\begin{aligned}\mathbf{E}[Y_i(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_s] &= \mathbf{E}[\,\mathbf{E}[Y_i(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_{t_i}]\,|\mathcal{F}_s] \\ &= \mathbf{E}[Y_i\,\mathbf{E}[(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_{t_i}]\,|\mathcal{F}_s] = 0,\end{aligned}$$

since $\mathbf{E}[(M_{t_{i+1}} - M_{t_i})|\mathcal{F}_{t_i}] = 0$ by the martingale property. Putting all the cases together in (5.6) gives for $s \in (t_k, t_{k+1}]$, say,

$$\mathbf{E}[I_t|F_s] = Y_0(M_{t_1} - M_0) + Y_1(M_{t_2} - M_{t_1}) + \cdots + Y_k(M_s - M_{t_k}). \qquad \square$$

5.3. The Itô Integral

We now turn to martingale transforms where the underlying martingale is a standard Brownian motion $(B_t, t \ge 0)$. This gives our first definition of the Itô integral.

Definition 5.6 (Itô integral on $\mathcal{S}(T)$). Let $(B_t, t \le T)$ be a standard Brownian motion on $[0, T]$ and let $X \in \mathcal{S}(T)$ be a simple process $X = \sum_{j=0}^{n-1} Y_j \mathbf{1}_{(t_j, t_{j+1}]}$ on $[0, T]$ adapted to the Brownian filtration. The Itô integral of X with respect to the Brownian motion is defined as the martingale transform

$$\int_0^T X_s \, \mathrm{d}B_s = \sum_{j=0}^{n-1} Y_j (B_{t_{j+1}} - B_{t_j}),$$

and similarly for any $t \leq T$,

$$\int_0^t X_s \, \mathrm{d}B_s = Y_0(B_{t_1} - B_0) + Y_1(B_{t_2} - B_{t_1}) + \cdots + Y_j(B_t - B_{t_j}), \quad \text{if } t \in (t_j, t_{j+1}].$$

Note again the similarities with Riemann sums. The interpretation of the Itô integral is as follows:

> *the value of implementing the strategy X on the underlying asset with price given by the Brownian motion.*

The martingale transform with Brownian motion has more properties than with a generic martingale as given in Definition 5.3. This is because the Brownian increments are independent. We gather the properties of the Itô integral for $X \in \mathcal{S}(T)$ in an important proposition. The same exact result will hold for continuous strategies; see Theorem 5.12.

Proposition 5.7 (Properties of the Itô integral). *Let $(B_t, t \leq T)$ be a standard Brownian motion on $[0, T]$ defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. The Itô integral in Definition 5.6 has the following properties:*

- **Linearity:** *If $X, X' \in \mathcal{S}(T)$ and $a, b \in \mathbb{R}$, then for all $t \leq T$,*
$$\int_0^t (aX_s + bX'_s) \, \mathrm{d}B_s = a \int_0^t X_s \, \mathrm{d}B_s + b \int_0^t X'_s \, \mathrm{d}B_s.$$
- **Continuous martingale:** *The process $\left(\int_0^t X_s \, \mathrm{d}B_s, t \leq T\right)$ is a continuous martingale on $[0, T]$ for the Brownian filtration.*
- **Itô's isometry:** *The random variable $\int_0^t X_s \, \mathrm{d}B_s$ is in $L^2(\Omega, \mathcal{F}, \mathbf{P})$ with mean 0 and variance*
$$\mathbf{E}\left[\left(\int_0^t X_s \, \mathrm{d}B_s\right)^2\right] = \int_0^t \mathbf{E}[X_s^2] \, \mathrm{d}s = \mathbf{E}\left[\int_0^t X_s^2 \, \mathrm{d}s\right], \quad t \leq T.$$

It is very important for the understanding of the theory to keep in mind that $\int_0^t X_s \, \mathrm{d}B_s$ is a random variable. We should walk away from the temptation to use the reflexes of classical calculus to manipulate it as if it were a Riemann integral. The reason we use the integral sign to denote the random variable $\int_0^t X_s \, \mathrm{d}B_s$ is because it shares the linearity property with the Riemann integral.

It turns out that Itô's isometry not only yields the mean and the variance of the random variable $\int_0^t X_s \, \mathrm{d}B_s$, but also the covariances for these random variables at different times, and the covariance for two integrals built with two different strategies on the same Brownian motion; see Corollary 5.15. What about the distribution of $\int_0^t X_s \, \mathrm{d}B_s$? It turns out that the random variable $\int_0^t X_s \, \mathrm{d}B_s$ is not Gaussian in general. However, if the process X is not random, then it will be; see Corollary 5.18 below. For example, the process $(I_t, t \leq T)$ in Example 5.1 is Gaussian, but the one in Example 5.4 is not.

Proof of Proposition 5.7. The linearity is clear from the definition of the martingale transform. The continuity property and the martingale property follow generally from Proposition 5.5.

We now prove Itô's isometry. We will use the properties of conditional expectation in Proposition 4.19 many times, so the reader might quickly review it beforehand. To simplify notation, for fixed $t \in [0, T]$, we can suppose that the partition $(t_j, j \leq n)$ is a partition of $[0, t]$ with $t_n = t$. Since Y_j is $\mathcal{F}_{t_j}$-measurable, we have

$$\mathbf{E}[Y_j(B_{t_{j+1}} - B_{t_j})] = \mathbf{E}[\mathbf{E}[Y_j(B_{t_{j+1}} - B_{t_j})|\mathcal{F}_{t_j}]] = \mathbf{E}[Y_j\mathbf{E}[B_{t_{j+1}} - B_{t_j}|\mathcal{F}_{t_j}]] = 0,$$

since $\mathbf{E}[B_{t_{j+1}} - B_{t_j}|\mathcal{F}_{t_j}] = 0$, as Brownian motion is a martingale. Therefore, it follows that

$$\mathbf{E}\left[\int_0^t X_s \, \mathrm{d}B_s\right] = \sum_{j=0}^{n-1} \mathbf{E}[Y_j(B_{t_{j+1}} - B_{t_j})] = 0.$$

As for the variance, we have by conditioning on $\mathcal{F}_{t_j}$ that, for $t_i < t_j$,

$$\mathbf{E}[Y_jY_i(B_{t_{j+1}} - B_{t_j})(B_{t_{i+1}} - B_{t_i})] = \mathbf{E}[Y_jY_i(B_{t_{i+1}} - B_{t_i})\,\mathbf{E}[B_{t_{j+1}} - B_{t_j}|\mathcal{F}_{t_j}]] = 0,$$

since $\mathbf{E}[B_{t_{j+1}} - B_{t_j}|\mathcal{F}_{t_j}] = 0$ and since all factors but $B_{t_{j+1}} - B_{t_j}$ are $\mathcal{F}_{t_j}$-measurable. Thus, this yields

$$\begin{aligned}\mathbf{E}\left[\left(\int_0^t X_s \, \mathrm{d}B_s\right)^2\right] &= \sum_{i,j=0}^{n-1} \mathbf{E}[Y_jY_i(B_{t_{j+1}} - B_{t_j})(B_{t_{i+1}} - B_{t_i})] \\ &= \sum_{j=0}^{n-1} \mathbf{E}[Y_j^2\,\mathbf{E}[(B_{t_{j+1}} - B_{t_j})^2|\mathcal{F}_{t_j}]],\end{aligned}$$

by the previous equation and the fact that Y_j is $\mathcal{F}_{t_j}$-measurable. Since the increment $B_{t_{j+1}} - B_{t_j}$ is independent of $\mathcal{F}_{t_j}$, we have

$$\mathbf{E}[(B_{t_{j+1}} - B_{t_j})^2|\mathcal{F}_{t_j}] = \mathbf{E}[(B_{t_{j+1}} - B_{t_j})^2] = t_{j+1} - t_j.$$

Therefore, we conclude that

$$\mathbf{E}\left[\left(\int_0^t X_s \, \mathrm{d}B_s\right)^2\right] = \sum_{j=0}^{n-1} \mathbf{E}[Y_j^2](t_{j+1} - t_j).$$

From the definition of X as a simple process in equation (5.2), we have $\int_0^t \mathbf{E}[X_s^2]\,\mathrm{d}s = \sum_{j=0}^{n-1} \mathbf{E}[Y_j^2](t_{j+1} - t_j)$, since X equals Y_j on the whole interval $(t_j, t_{j+1}]$. □

Example 5.8. We go back to the Itô integral in Example 5.2. The mean of I_t is 0 by Proposition 5.7 or by direct computation. It is not hard to compute the variance. For example, at $t = 1$, it is

$$\mathbf{E}[I_1^2] = \int_0^1 \mathbf{E}[X_u^2]\,\mathrm{d}u = \mathbf{E}[B_0^2] \cdot \frac{1}{3} + \mathbf{E}[B_{1/3}^2] \cdot \frac{1}{3} + \mathbf{E}[B_{2/3}^2] \cdot \frac{1}{3} = 0 + \frac{1}{9} + \frac{2}{9} = \frac{1}{3}.$$

Consider now another process Y on $[0,1]$ defined on the same Brownian motion:

$$Y_t = B_0^2\mathbf{1}_{(0,1/3]}(t) + B_{1/3}^2\mathbf{1}_{(1/3,2/3]}(t) + B_{2/3}^2 1_{(2/3,1]}(t).$$

Again, the Itô integral $J_t = \int_0^t Y_s^2 \,\mathrm{d}B_s$ is well-defined as a process on $[0,1]$:

$$J_t = \begin{cases} 0 & \text{if } t \in [0,1/3], \\ B_{1/3}^2(B_t - B_{1/3}) & \text{if } t \in (1/3,2/3], \\ B_{1/3}^2(B_{2/3} - B_{1/3}) + B_{2/3}^2(B_t - B_{2/3}) & \text{if } t \in (2/3,1]. \end{cases}$$

The covariance between the random variables I_1 and J_1 can be computed easily by using the independence of the increments and suitable conditioning. Indeed, we have

$$\mathbf{E}\,[I_1 J_1] = \sum_{i,j=0}^{3} \mathbf{E}[B_{i/3}B_{j/3}^2(B_{(i+1)/3} - B_{i/3})(B_{(j+1)/3} - B_{j/3})].$$

If $j > i$, we can condition on $\mathcal{F}_{j/3}$ in the above summand to get

$$\begin{aligned} &\mathbf{E}[B_{i/3}B_{j/3}^2(B_{(i+1)/3} - B_{i/3})(B_{(j+1)/3} - B_{j/3})\,|\mathcal{F}_{j/3}] \\ &\quad = B_{i/3}B_{j/3}^2(B_{(i+1)/3} - B_{i/3})\mathbf{E}[B_{(j+1)/3} - B_{j/3}\,|\mathcal{F}_{j/3}] = 0. \end{aligned}$$

The same holds for $i > j$ by conditioning on $\mathcal{F}_{i/3}$. The only remaining terms are $i = j$:

$$E\,[I_1 J_1] = \sum_{i=0}^{3} \mathbf{E}[B_{i/3}^3(B_{(i+1)/3} - B_{i/3})^2] = \sum_{i=0}^{3} \mathbf{E}[B_{i/3}^3] \cdot \mathbf{E}[(B_{(i+1)/3} - B_{i/3})^2],$$

by independence of increments. The first factor of each term is zero (due to the nature of odd moments of a Gaussian centered at 0). Therefore, the variables I_1 and J_1 are uncorrelated. Corollary 5.15 gives a systematic way to compute covariances based on Itô's isometry.

Remark 5.9. An *isometry* is a mapping between metric spaces (i.e., with a distance) that actually preserves the distance between points. (It literally means *same measure* in Greek.) In the case of Itô's isometry, the mapping is the one that sends the integrand X to the square-integrable random variable given by the integral:

$$\begin{aligned} \mathcal{S}(T) &\to L^2(\Omega, \mathcal{F}, \mathbf{P}) \\ X &\mapsto \int_0^T X_s \,\mathrm{d}B_s. \end{aligned}$$

The L^2-norm of $\int_0^T X_s \,\mathrm{d}B_s$ is $(\mathbf{E}[(\int_0^T X_s \,\mathrm{d}B_s)^2])^{1/2}$. It turns out that the space $\mathcal{S}(T)$ is also a linear space with the norm $\|X\|_{\mathcal{S}} = (\int_0^T \mathbf{E}[X_s]^2 \,\mathrm{d}s)^{1/2}$. Itô's isometry says that these two norms (and hence the distance) are equal. In fact, this isometry extends in part to the L^2-space of functions on $\Omega \times [0,T]$, for which $\mathcal{S}(T)$ is a subspace. We will see that this isometry is central to the extension of the Itô integral in the limit $n \to \infty$.

The next goal is to extend the Itô integral to processes X other than simple processes. The integral will be defined as a limit of integrals of simple processes, very much like the Riemann integral is a limit of Riemann sums. But first, we need a good class of integrands.

Definition 5.10. For a given Brownian filtration $(\mathcal{F}_t, t \leq T)$, we consider the class of processes $\mathcal{L}^2_c(T)$ of processes $(X_t, t \leq T)$ such that the following hold:

(1) X is *adapted*; that is, X_t is $\mathcal{F}_t$-measurable for $t \leq T$.
(2) $\mathbf{E}[\int_0^T X_t^2 \,\mathrm{d}t] = \int_0^T \mathbf{E}[X_t^2] \,\mathrm{d}t < \infty$.
(3) X has continuous paths; that is, $t \mapsto X_t(\omega)$ is continuous on $[0, T]$ for a set of ω of probability one.

It is not hard to check that the processes $(B_t, t \leq T)$ and $(B_t^2, t \leq T)$ are in $\mathcal{L}^2_c(T)$. In fact, if f is a continuous function and $\int_0^T \mathbf{E}[f(B_t)^2] \,\mathrm{d}t < \infty$, then the process $(f(B_t), t \leq T)$ is in $\mathcal{L}^2_c(T)$. Indeed, $f(B_t)$ is $\mathcal{F}_t$-measurable, since it is an explicit function of B_t. Moreover, the second condition is by assumption. The third holds simply because the composition of two continuous functions is continuous. Example 5.17 describes a process that is in $\mathcal{L}^2_c(T)$ but is not an explicit function of Brownian motion. See Exercise 5.7 for an example of a process of the form $(f(B_t), t \leq T)$ that is not in $\mathcal{L}^2_c(T)$. The main advantage of processes in $\mathcal{L}^2_c(T)$ is that they are easily approximated by simple adapted processes.

Lemma 5.11. *Let $X \in \mathcal{L}^2_c(T)$. Then X can be approximated by simple adapted processes in $\mathcal{S}(T)$, in the sense that there exists a sequence $X^{(n)} \in \mathcal{S}(T)$ such that*

$$\lim_{n\to\infty} \int_0^T \mathbf{E}[(X_t^{(n)} - X_t)^2] \,\mathrm{d}t = 0.$$

Proof. For a given n, consider the partition $t_j = \frac{j}{n}T$ of $[0, T]$ and the simple adapted process given by

$$X_t^{(n)} = \sum_{j=0}^{n} X_{t_j} \mathbf{1}_{(t_j, t_{j+1}]}(t), \quad t \leq T.$$

In other words, we give the constant value X_{t_j} on the whole interval $(t_j, t_{j+1}]$. By continuity of the paths of X, it is clear that $X_t^{(n)}(\omega) \to X_t(\omega)$ at any $t \leq T$ and for any ω. Therefore, by Theorem 4.40, we have

$$\lim_{n\to\infty} \int_0^T \mathbf{E}[(X_t^{(n)} - X_t)^2] \,\mathrm{d}t = 0. \qquad \square$$

We are now ready to state the most important theorem of this section.

Theorem 5.12. *Let $(B_t, t \leq T)$ be a standard Brownian motion defined on $(\Omega, \mathcal{F}, \mathbf{P})$. Let $(X_t, t \leq T)$ be a process in $\mathcal{L}^2_c(T)$. There exist random variables $\int_0^t X_s \,\mathrm{d}B_s$, $t \leq T$, with the following properties:*

- **Linearity**: *If $X, Y \in \mathcal{L}^2_c(T)$ and $a, b \in \mathbb{R}$, then*

$$\int_0^t (aX_s + bY_s) \,\mathrm{d}B_s = a \int_0^t X_s \,\mathrm{d}B_s + b \int_0^t Y_s \,\mathrm{d}B_s, \quad t \leq T.$$

- **Continuous martingale:** *The process $\left(\int_0^t X_s \,\mathrm{d}B_s, t \leq T\right)$ is a continuous martingale for the Brownian filtration.*

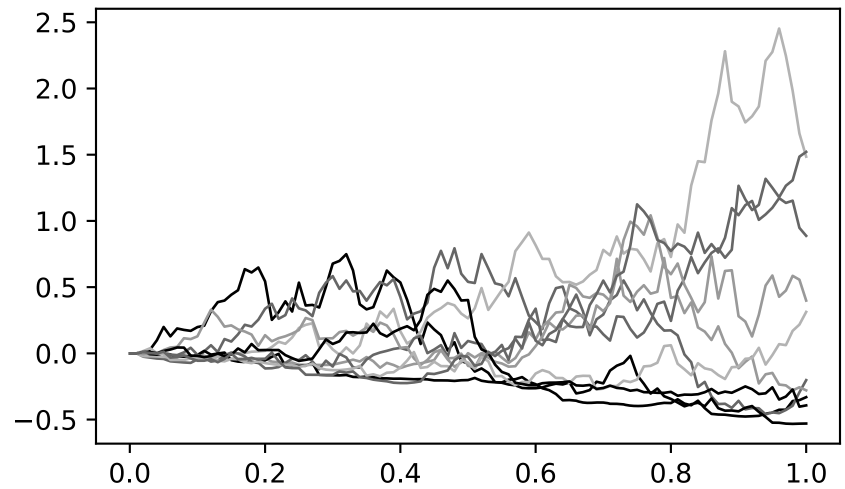

Figure 5.2. Simulation of 5 paths of the process $\int_0^t B_s \, \mathrm{d}B_s$, $t \leq 1$.

- **Itô's isometry:** *The random variable* $\int_0^t X_s \, \mathrm{d}B_s$ *is in* $L^2(\Omega, \mathcal{F}, \mathbf{P})$ *with mean* 0 *and variance*

$$\mathbf{E}\left[\left(\int_0^t X_s \, \mathrm{d}B_s\right)^2\right] = \int_0^t \mathbf{E}[X_s^2] \, \mathrm{d}s = \mathbf{E}\left[\int_0^t X_s^2 \, \mathrm{d}s\right], \quad t \leq T.$$

Example 5.13 (Sampling Itô integrals)**.** How can we sample paths of processes given by Itô integrals? A very simple method is to go back to the integral on simple processes. Consider the process $I_t = \int_0^t X_s \, \mathrm{d}B_s$, $t \leq T$, constructed from $X \in \mathcal{L}_c^2(T)$ and from a standard Brownian motion $(B_t, t \geq 0)$. To simulate the paths, we fix the endpoint, say T, and a step size, say $1/n$. Then we can generate the process at every $t_j = \frac{j}{n}T$ by taking

$$I_{t_j} = \sum_{i=0}^{j-1} X_{t_i}(B_{t_{i+1}} - B_{t_i}), \quad j \leq n.$$

Here are two observations that makes this expression more palatable. First, note that the increment $B_{t_{i+1}} - B_{t_i}$ is a Gaussian random variable of mean 0 and variance $\frac{1}{n}T$ for every i. Second, we have $I_{t_j} - I_{t_{j-1}} = X_{t_{j-1}}(B_{t_j} - B_{t_{j-1}})$, so the values I_{t_j} can be computed recursively. Numerical Project 5.1 is about implementing this method.

Remark 5.14 (L^2-spaces are complete)**.** The proof of the existence of the Itô integral is based on the completeness property of L^2-spaces: If $(X_n, n \geq 1)$ is a *Cauchy sequence* of elements in an L^2-space, then the sequence X_n converges to some element X in L^2. A sequence is Cauchy if for any choice of $\varepsilon > 0$, we can find n large enough so that

$$\|X_m - X_n\| < \varepsilon \text{ for any } m > n.$$

In other words, for an arbitrarily small distance ε, if we go further enough in the sequence the distances between increments are all smaller than ε. Another example of spaces that are complete is $\mathbb{R}$, endowed with the metric given by the absolute value.

However, the set of rational numbers $\mathbb{Q}$ as a subset of $\mathbb{R}$ is not complete, because there are sequences of rational numbers that converge to irrationals. A proof of the completeness of L^2 is outlined in Exercise 5.20.

Proof of Theorem 5.12. Consider a process $X = (X_t, t \leq T)$ in $\mathcal{L}_c^2(T)$. By Lemma 5.11, we can approximate it by a simple adapted processes $(X_t^{(n)}, t \leq T)$. In particular, this implies that the sequence is Cauchy for the metric

$$\|X^{(n)} - X^{(m)}\| = \left(\int_0^T \mathbf{E}[(X_t^{(n)} - X_t^{(m)})^2]\,\mathrm{d}t\right)^{1/2}. \tag{5.7}$$

The key step is the following. We know that the integral $I_t^{(n)} = \int_0^t X_s^{(n)}\,\mathrm{d}B_s$ is well-defined as a random variable in $L^2(\Omega, \mathcal{F}, \mathbf{P})$. Moreover, we know by the Itô isometry that the L^2-distance of the processes in equation (5.7) is the same as the L^2-distance of the $I^{(n)}$'s. This means that the sequence $(I^{(n)}, n \geq 1)$ must also be Cauchy in $L^2(\Omega, \mathcal{F}, \mathbf{P})$! We conclude by completeness of the space that $I^{(n)}$ converges in L^2 to a random variable that we denote by I_t or $\int_0^t X_s\,\mathrm{d}B_s$. Furthermore, the limit I_t does not depend on the approximating sequence $X^{(n)}$. We could have taken another sequence to approximate X and the isometry guarantees that the corresponding integrals will converge to the same random variable.

We now prove the properties:

- *Linearity*: It follows by using linearity in Proposition 5.7 for $X^{(n)}$ and $Y^{(n)}$, the two approximating processes for X and Y.
- *Isometry*: The variance follows from the following fact: If $I_t^{(n)} \to I_t$ in L^2, then $\mathbf{E}[(I_t^{(n)})^2] \to \mathbf{E}[I_t^2]$ and $\mathbf{E}[I_t^{(n)}] \to \mathbf{E}[I_t]$; see Exercise 5.3.
- *Continuous martingale*: Write $I_t = \int_0^t X_s\,\mathrm{d}B_s$. We must show that $\mathbf{E}[I_t|\mathcal{F}_s] = I_s$ for any $t > s$. To see this, we go back to Definition 4.14. The random variable I_t is $\mathcal{F}_t$-measurable by construction. Now for a bounded random variable W that is $\mathcal{F}_s$-measurable, we need to show

  $$\mathbf{E}[WI_t] = \mathbf{E}[WI_s].$$

  This is clear for $I_t^{(n)}$, the approximating integrals, because $(I_t^{(n)}, t \leq T)$ is a martingale. The above then follows from the fact that $I_s^{(n)}W$ converges in L^2 to I_sW (and thus the expectation converges) and the same way for t. The fact that the path $t \mapsto I_t(\omega)$ is continuous on $[0,t]$ with probability one is more involved. It uses Doob's maximal inequality; see Exercise 4.19.

□

Once the conclusions of Theorem 5.12 are accepted, we are free to explore the beauty and the power of Itô calculus. As a first step, we observe that with Itô's isometry, we can compute not only variances, but also covariances between integrals. This is because an isometry also preserves the inner product in L^2-spaces.

Corollary 5.15. *Let $(B_t, t \leq T)$ be a standard Brownian motion, and let $X \in \mathcal{L}_c^2(T)$. We have*

$$\mathbf{E}\left[\left(\int_0^t X_s \, \mathrm{d}B_s\right)\left(\int_0^{t'} X_s \, \mathrm{d}B_s\right)\right] = \int_0^{t \wedge t'} \mathbf{E}[X_s^2] \, \mathrm{d}s, \quad t, t' \leq T,$$

and for any $Y \in \mathcal{L}_c^2(T)$,

$$\mathbf{E}\left[\left(\int_0^t X_s \, \mathrm{d}B_s\right)\left(\int_0^t Y_s \, \mathrm{d}B_s\right)\right] = \int_0^t \mathbf{E}[X_s Y_s] \, \mathrm{d}s, \quad t \leq T.$$

Note that when X is just the constant 1, we recover from the first equation the covariance of Brownian motion.

Proof. The first assertion is Exercise 5.4. As for the second, we have on one hand by Itô's isometry

$$\begin{aligned}\mathbf{E}\left[\left(\int_0^t \{X_s + Y_s\} \, \mathrm{d}B_s\right)^2\right] &= \int_0^t \mathbf{E}[(X_s + Y_s)^2] \, \mathrm{d}s \\ &= \int_0^t \mathbf{E}[X_s^2] \, \mathrm{d}s + \int_0^t \mathbf{E}[Y_s^2] \, \mathrm{d}s + 2\int_0^t \mathbf{E}[X_s Y_s] \, \mathrm{d}s.\end{aligned}$$

On the other hand, by linearity of the Itô integral and of the expectation, we have

$$\begin{aligned}\mathbf{E}\left[\left(\int_0^t \{X_s + Y_s\} \, \mathrm{d}B_s\right)^2\right] &= \mathbf{E}\left[\left(\int_0^t X_s \, \mathrm{d}B_s + \int_0^t Y_s \, \mathrm{d}B_s\right)^2\right] \\ &= \mathbf{E}\left[\left(\int_0^t X_s \, \mathrm{d}B_s\right)^2\right] + \mathbf{E}\left[\left(\int_0^t Y_s \, \mathrm{d}B_s\right)^2\right] \\ &\quad + 2\mathbf{E}\left[\left(\int_0^t X_s \, \mathrm{d}B_s\right)\left(\int_0^t Y_s \, \mathrm{d}B_s\right)\right].\end{aligned}$$

By combining the two equations and by using Itô's isometry, we conclude that

$$\mathbf{E}\left[\left(\int_0^t X_s \, \mathrm{d}B_s\right)\left(\int_0^t Y_s \, \mathrm{d}B_s\right)\right] = \int_0^t \mathbf{E}[X_s Y_s] \, \mathrm{d}s. \tag{5.8}$$

□

Example 5.16. Consider the processes $(B_t, t \leq T)$ and $(B_t^2, t \leq T)$ for a given standard Brownian motion. Note that these two processes are in $\mathcal{L}_c^2(T)$ for any $T > 0$. By Theorem 5.12, the random variables

$$I_t = \int_0^t B_s \, \mathrm{d}B_s, \qquad J_t = \int_0^t B_s^2 \, \mathrm{d}B_s$$

exist and are in $L^2(\Omega, \mathcal{F}, \mathbf{P})$. Their mean is 0, and they have variances

$$\mathbf{E}[I_t^2] = \int_0^t \mathbf{E}[B_s^2] \, \mathrm{d}s = \int_0^t s \, \mathrm{d}s = \frac{t^2}{2}, \qquad \mathbf{E}[J_t^2] = \int_0^t \mathbf{E}[B_s^4] \, \mathrm{d}s = \int_0^t 3s^2 \, \mathrm{d}s = t^3.$$

(Recall the Gaussian moments in equation (1.8).) The covariance is by Corollary 5.15:

$$\mathbf{E}[I_t J_t] = \int_0^t \mathbf{E}[B_s B_s^2]\, \mathrm{d}s = 0.$$

The variables are uncorrelated.

Example 5.17 (A path-dependent integrand)**.** Consider the process $X_t = \int_0^t B_s \, \mathrm{d}B_s$ on $[0, T]$ as in Example 5.16. Note that the process $(X_t, t \leq T)$ is itself in $\mathcal{L}_c^2(T)$. In particular, the integral $\int_0^t X_s \, \mathrm{d}B_s$ is well-defined! (Note that the integrand X_t is $\mathcal{F}_t$-measurable but its value depends on the whole Brownian up to time t.) The mean of the integral is 0 and its variance is obtained by applying Itô's isometry twice:

$$\mathbf{E}\left[\left(\int_0^t X_s \, \mathrm{d}B_s\right)^2\right] = \int_0^t \mathbf{E}[X_s^2]\, \mathrm{d}s = \int_0^t \frac{s^2}{2}\, \mathrm{d}s = \frac{t^3}{6}.$$

See Numerical Project 5.4.

In general, the Itô integral is not Gaussian. However, if the integrand X is not random (as in Example 5.1), the process is actually Gaussian. In this particular case, the integral is sometimes called a *Wiener integral*.

Corollary 5.18 (Wiener integral)**.** *Let $(B_t, t \leq T)$ be a standard Brownian motion and let $f : [0, T] \to \mathbb{R}$ be a function such that $\int_0^T f^2(s)\, \mathrm{d}s < \infty$. Then the process $(\int_0^t f(s)\, \mathrm{d}B_s, t \leq T)$ is Gaussian with mean* 0 *and covariance*

$$\mathrm{Cov}\left(\int_0^t f(s)\, \mathrm{d}B_s, \int_0^{t'} f(s)\, \mathrm{d}B_s\right) = \int_0^{t \wedge t'} f(s)^2\, \mathrm{d}s.$$

Proof. We prove the case when f is continuous. In this case, we can use our proof of Lemma 5.11. Let $(t_j, j \leq n)$ be a partition of $[0, T]$ in n intervals. The lemma shows that the sequence of simple functions

$$f^{(n)}(t) = \sum_{j=0}^{n-1} f(t_j) \mathbf{1}_{(t_j, t_{j+1}]}(t), \quad t \leq T,$$

approximates f. The Itô integral of $f^{(n)}$ is

$$I_t^{(n)} = \sum_{i=0}^{j-1} f(t_j)(B_t - B_{t_j}), \quad t \in (t_j, t_{j+1}].$$

This is a Gaussian process for any n. This is because for any choice of times $s_1, \dots, s_m$, the vector $(I_{s_1}^{(n)}, \dots, I_{s_m}^{(n)})$ is Gaussian, since it reduces to linear combinations of Brownian motion at fixed times. Moreover, the random variable $\int_0^t f(s)\, \mathrm{d}B_s$ is the L^2-limit of $I_t^{(n)}$ by Theorem 5.12. It remains to show that an L^2-limit of a sequence of Gaussian vectors remains Gaussian. This is sketched in Exercise 5.19. The expression of the covariances is from Corollary 5.15. □

Example 5.19 (Ornstein-Uhlenbeck process as an Itô integral)**.** Consider the function $f(s) = e^s$. The Ornstein-Uhlenbeck process starting at X_0 defined in Example 2.29 can also be written as

$$Y_t = e^{-t} \int_0^t e^s \, \mathrm{d}B_s, \quad t \geq 0. \tag{5.9}$$

This is tested numerically in Numerical Project 5.2. To see this mathematically, note that $(Y_t, t \geq 0)$ is a Gaussian process by Corollary 5.18. The mean is 0 and the covariance is, by Corollary 5.15,

$$\mathbf{E}[Y_t Y_s] = e^{-t-s} \int_0^s e^{2u} \, \mathrm{d}u = \frac{1}{2}(e^{-(t-s)} - e^{-(t+s)}), \quad s \leq t.$$

We can also start the process at Y_0, a Gaussian random variable of mean 0 and variance $1/2$ independent of the Brownian motion $(B_t, t \geq 0)$. The process then takes the form

$$Y_t = Y_0 e^{-t} + e^{-t} \int_0^t e^s \, \mathrm{d}B_s.$$

Since Y_0 and the Itô integral are independent by assumption, the covariance is then

$$\mathbf{E}[Y_t Y_s] = \frac{1}{2} e^{-t-s} + \frac{1}{2}(e^{-(t-s)} - e^{-(t+s)}) = \frac{1}{2} e^{-(t-s)}, \quad s \leq t.$$

In this case, the process is stationary in the sense that $(Y_t, t \geq 0)$ has the same distribution as $(Y_{t+a}, t \geq 0)$ for any $a > 0$.

Example 5.20 (Brownian bridge as an Itô integral)**.** The Brownian bridge $(Z_t, t \in [0, 1])$ is the stochastic process with the distribution defined in Example 2.27. Another way to construct a Brownian bridge is as follows:

$$Z_t = (1 - t) \int_0^t \frac{1}{1 - s} \, \mathrm{d}B_s, \quad t < 1. \tag{5.10}$$

This is tested numerically in Numerical Project 5.2. It turns out that $Z_1 = 0$. This is done in Exercise 5.21. The process Z is a Gaussian process by Corollary 5.18. The mean is 0 and the covariance is, by Corollary 5.15,

$$\mathbf{E}[Z_t Z_s] = (1 - t)(1 - s)\mathbf{E}\left[\left(\int_0^s \frac{1}{1 - u} \, \mathrm{d}B_u\right)\left(\int_0^t \frac{1}{1 - u} \, \mathrm{d}B_u\right)\right] = s(1 - t), \quad s \leq t.$$

The above representations of the Orstein-Uhlenbeck and the Brownian bridge implies that they are not martingales; see Exercise 5.10.

Remark 5.21 (Fubini's theorem)**.** In Exercise 3.11 and Exercise 4.21, it was shown that we can interchange the expectation $\mathbf{E}$ and the sum $\sum$ if the random variables are positive or if $\sum_{n \geq 1} \mathbf{E}[|X_n|] < \infty$. This result holds in general when the integrands are positive or integrable. This is known as Fubini's theorem. This is applicable in particular when we calculate the variance using Itô's isometry. More precisely, we have

$$\int_0^t \mathbf{E}[X_s^2] \, \mathrm{d}s = \mathbf{E}\left[\int_0^t X_s^2 \, \mathrm{d}s\right].$$

Remark 5.22 (Extension to other processes)**.** Can we define the Itô integral for processes other than the ones in $\mathcal{L}_c^2(T)$? Of course, since simple adapted processes in $\mathcal{S}(T)$ given in equation (5.5) are not continuous. In fact, the Itô construction holds whenever X is a limit of simple adapted processes. Such processes will have the property that

$$\mathbf{E}\left[\int_0^T X_t^2 \, \mathrm{d}t\right] < \infty. \tag{5.11}$$

Theorem 5.12 is the same for these processes. In particular, they define continuous square-integrable martingales.

A further extension applies to processes such that

$$\int_0^T X_t^2(\omega) \, \mathrm{d}t < \infty \text{ for } \omega \text{ in a set of probability one.} \tag{5.12}$$

(Note that equation (5.11) implies the above by Exercise 1.15.) Equation (5.12) is a very weak condition, since any process $X_t = g(B_t)$ where g is continuous will satisfy it, because a continuous function is bounded on an interval. For example the process $X_t = e^{B_t^2}$ does not satisfy (5.11), but it satisfies (5.12); see Numerical Project 5.7 and Exercise 5.7. The construction of the Itô integral for such processes involves stopping times and will not be pursued here. The Itô integrals in this case are not martingales but are said to be *local martingales*; i.e., they are martingales when suitably stopped:

Definition 5.23. A process $(Y_t, t \geq 0)$ is said to be a *local martingale* for the filtration $(\mathcal{F}_t, t \geq 0)$ if there exists an increasing sequence of stopping times $(\tau_n, n \geq 1)$ for the same filtration such that $\tau_n \to +\infty$ as $n \to \infty$ almost surely, and the stopped processes $(M_{t\wedge\tau_n}, t \geq 0)$ are martingales for every $n \geq 1$.

5.4. Itô's Formula

The Itô integral was constructed in the last section in a rather abstract way. It is the limit of a sequence of random variables constructed from Brownian motion. It is good to remind ourselves that the classical Riemann integral is also very abstract! It is defined as the limit of the sequence of Riemann sums. It does not always have an explicit form. For example, the CDF of a Gaussian variable

$$\Phi(x) = \int_{-\infty}^x \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \, \mathrm{d}y$$

is a well-defined function of x, but the integral cannot be expressed in terms of the typical elementary functions of calculus. But in some cases, a Riemann integral can be written explicitly in terms of such functions. This is the content of the *fundamental theorem of calculus*. It is useful to recall the theorem, as Itô's formula is built upon it.

Let $f : [0, T] \to \mathbb{R}$ be a function for which the derivative f' exists and is a continuous function on $[0, T]$. We will say that such a function is in $\mathcal{C}^1([0, T])$. The fundamental theorem of calculus says that we can write

$$f(t) - f(0) = \int_0^t f'(s) \, \mathrm{d}s, \quad t \leq T. \tag{5.13}$$

Note that we often write this result in differential form:

$$\mathrm{d}f(t) = f'(t)\,\mathrm{d}t\,. \tag{5.14}$$

The differential form has no rigorous meaning in itself. It is simply a compact and convenient notation that encodes (5.13).

The stochastic equivalent of the fundamental theorem of calculus is Itô's formula provided below. It relates the Itô integral to an explicit function of Brownian motion. Note that the function f must be in $\mathcal{C}^2(\mathbb{R})$; i.e., f' and f'' exist and are continuous on the whole space $\mathbb{R}$.

Theorem 5.24 (Itô's formula)**.** *Let $(B_t, t \leq T)$ be a standard Brownian motion. Consider $f \in \mathcal{C}^2(\mathbb{R})$. Then, with probability one, we have*

$$f(B_t) - f(B_0) = \int_0^t f'(B_s)\,\mathrm{d}B_s + \frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s, \quad t \leq T. \tag{5.15}$$

We will see other variations in Proposition 5.28 and in Chapters 6 and 7. Before giving the idea of the proof, we make some important observations:

(i) Equation (5.15) is an *equality of processes*, which is much stronger than equality in distribution. In other words, if you take a path of the process on the left constructed on a given Brownian motion, then this path will be the same as the path of the process on the right constructed on the same Brownian motion. See Figure 5.3. The reader should verify this in Numerical Project 5.3. The equality holds in the limit where the mesh of the partition of the interval $[0, T]$ goes to 0. See Numerical Project 5.5.

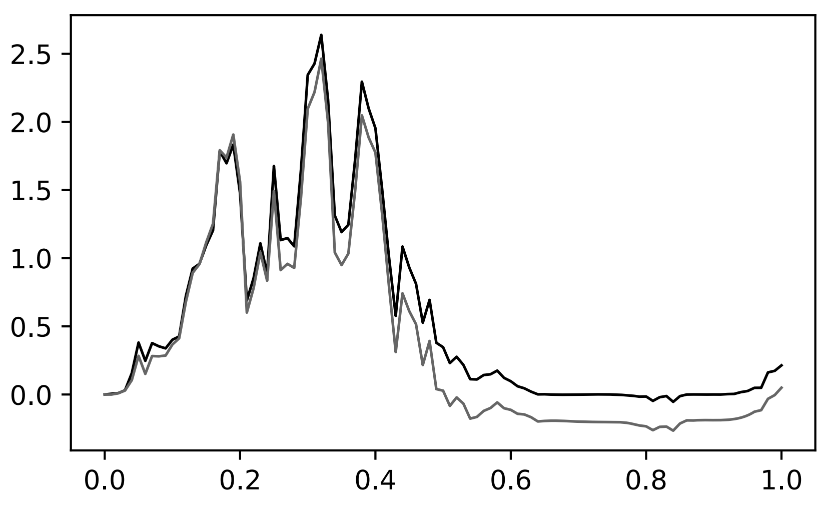

Figure 5.3. Simulation of a path of B_t^3 and of a path of $3\int_0^t B_s^2\,\mathrm{d}B_s + 3\int_0^t B_s\,\mathrm{d}s$ for a discretization of 0.01. See Numerical Project 5.5.

(ii) Note the similarity with the classical formulation in (5.13) if we replace the Riemann integral by Itô's integral. We do have the additional integral of $f''(B_s)$. As

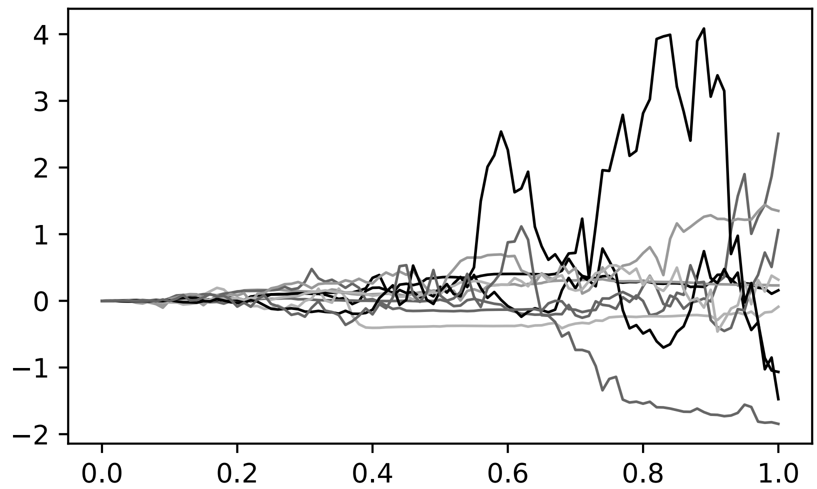

Figure 5.4. A sample of 10 paths of the martingale $B_t^3 - 3\int_0^t B_s\,\mathrm{d}s$.

we will see in the proof, this additional term comes from the quadratic term in the Taylor approximation and from the quadratic variation of Brownian motion seen in Theorem 3.8. As in the classical case (5.14), it is very convenient to summarize the conclusion of Itô's formula in differential form:

$$\boxed{\mathrm{d}f(B_t) = f'(B_t)\,\mathrm{d}B_t + \frac{1}{2}f''(B_t)\,\mathrm{d}t.} \tag{5.16}$$

We stress that the differential form has no meaning by itself. It is a compact way to express the two integrals in Itô's formula and a powerful device for computations.

(iii) An important consequence of Itô's formula is that it provides a systematic way to construct martingales as explicit functions of Brownian motion. To make sure that $\int_0^t f'(B_s)\,\mathrm{d}B_s$, $t \le T$, defines a continuous square-integrable martingale on $[0,T]$, we might need to check that $(f'(B_t), t \le T) \in \mathcal{L}_c^2(T)$. In general the Itô integral $\int_0^t f'(B_s)\,\mathrm{d}B_s$ makes sense as a local martingale; see Remark 5.22.

Corollary 5.25 (Brownian martingales)**.** *Let $(B_t, t \le T)$ be a standard Brownian motion. Consider $f \in \mathcal{C}^2(\mathbb{R})$ such that $\int_0^T \mathbf{E}[f'(B_s)^2]\,\mathrm{d}s < \infty$. Then the process*

$$\left(f(B_t) - \frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s,\ t \le T\right)$$

is a martingale for the Brownian filtration.

Proof. This is straightforward from Itô's formula

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

The first term is a constant and the second term is a continuous martingale by Proposition 5.7. □

The integral we subtract from $f(B_t)$ is called the *compensator*. A simple case is given by the function $f(x) = x^2$. For this function, the corollary gives that the process $B_t^2 - t$, $t \geq 0$, is a martingale, as we already observed in Example 4.28. The compensator was then simply t. In general, the compensator might be random.

(iv) The compensator is the Riemann integral $\int_0^t f''(B_s)\,\mathrm{d}s$. It might seem to be a strange object at first. The function $f''(B_s)$ is random (it depends on ω), so the integral is a random variable. There is no problem in integrating the random function $f''(B_s)$ since by assumption it is a continuous function of s, since f'' and $B_s(\omega)$ are continuous. In fact, the paths of $\int_0^t f''(B_s)\,\mathrm{d}s$ are much smoother than the ones of Brownian motion in general: The paths are differentiable everywhere (the derivative is $f''(B_t)$), and in particular, the paths have bounded variations (see Example 3.6). See Figure 5.5 for a sample of paths of the process $\int_0^t B_s\,\mathrm{d}s$.

To sum it up, Itô's formula says that $f(B_t)$ can be expressed as a sum of two processes: one with bounded variation (the Riemann integral) and a (local) martingale with finite quadratic variation (the Itô integral). In the next chapter, we will study Itô processes in more generality, which are processes that can be expressed as the sum of a Riemann integral and an Itô integral.

Example 5.26 ($f(x) = x^3$)**.**

In this case, Itô's formula yields

$$B_t^3 = \int_0^t 3B_s^2\,\mathrm{d}B_s + \frac{1}{2}\int_0^t 6B_s\,\mathrm{d}s = 3\int_0^t B_s^2\,\mathrm{d}B_s + 3\int_0^t B_s\,\mathrm{d}s. \tag{5.17}$$

Figure 5.3 shows a sample of a single path of each of these two processes constructed from the same Brownian path. Note that they are almost equal (the discrepancy is only due to the discretization in the numerics)! From the above equation, we conclude that the process $B_t^3 - 3\int_0^t B_s\,\mathrm{d}s$ is a martingale. See Figure 5.4 for a sample of its paths. The process $(\int_0^t B_s\,\mathrm{d}s, t \geq 0)$ is not complicated. It is a Gaussian process since the integral is the limit (almost sure and in L^2) of the Riemann sums

$$\sum_{j=0}^{n-1} B_{t_j}(t_{j+1} - t_j),$$

and each term of the sum is a Gaussian variable. (Why?) Clearly, the mean of $\int_0^t B_s\,\mathrm{d}s$ is 0. The covariance of the process can be calculated directly by interchanging the integrals and the expectation:

$$\mathbf{E}\left[\left(\int_0^t B_s\,\mathrm{d}s\right)\left(\int_0^{t'} B_s\,\mathrm{d}s\right)\right] = \int_0^t \int_0^{t'} \mathbf{E}[B_s B_{s'}]\,\mathrm{d}s\,\mathrm{d}s' = \int_0^t \int_0^{t'} (s \wedge s')\ \mathrm{d}s\,\mathrm{d}s'.$$

The integral equals $\frac{t't^2}{2} - \frac{t^3}{6}$, for $t \leq t'$. In particular, the variance at time t is $\frac{t^3}{3}$. The paths of this process are very smooth as can be observed in Figure 5.5. In fact, the paths are differentiable and the derivative at time t is B_t.

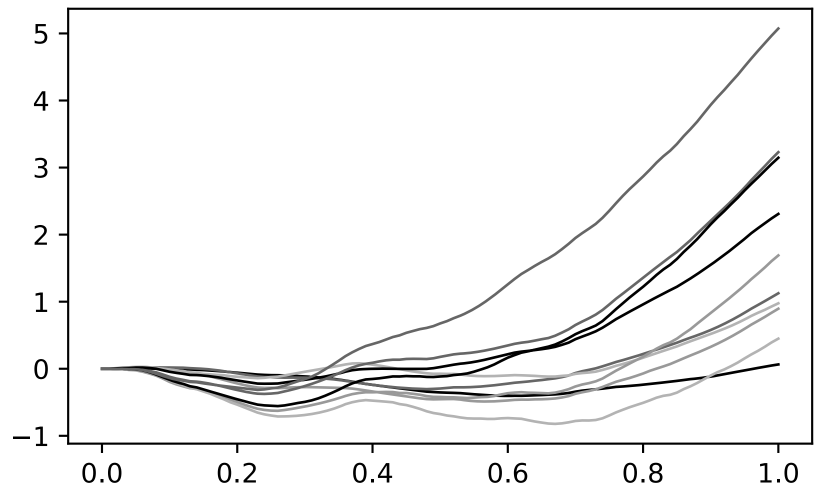

Figure 5.5. A sample of 10 paths of $3\int_0^t B_s\, \mathrm{d}s$.

Example 5.27 ($f(x) = \cos x$)**.**

In this case, Itô's formula gives

$$\cos B_t - \cos 0 = \int_0^t (-\sin B_s)\, \mathrm{d}B_s + \frac{1}{2}\int_0^t (-\cos B_s)\, \mathrm{d}s.$$

In particular, the process

$$M_t = \cos B_t + \frac{1}{2}\int_0^t \cos B_s\, \mathrm{d}s = 1 - \int_0^t \sin B_s\, \mathrm{d}B_s, \quad t \geq 0,$$

is a continuous martingale starting at $M_0 = 1$. It is easy to check that the process $(\sin B_t, t \leq T)$ is in $\mathcal{L}_c^2(T)$ for any T. A sample of the paths of $(M_t, t \leq 1)$ is depicted in Figure 5.6.

Where does Itô's formula come from? It is the same idea as for the proof of the fundamental theorem of calculus. Let's start with the latter. Suppose $f \in \mathcal{C}^1(\mathbb{R})$; that is, f is differentiable with a continuous derivative. Then f admits a Taylor approximation around s of the form

$$f(t) - f(s) = f'(s)(t-s) + \mathcal{E}(s,t). \tag{5.18}$$

(This is in the spirit of the *mean-value theorem*.) Here, $\mathcal{E}(s,t)$ is an error term that goes to 0 faster than $(t-s)$ as $s \to t$. Now, for a partition $(t_j, j \leq n)$ of $[0,t]$, say $t_j = \frac{j}{n}t$, we can trivially write for any n

$$f(t) - f(0) = \sum_{j=0}^{n} f(t_{j+1}) - f(t_j).$$

Now, we can use equation (5.18) at $s = t_j$:

$$f(t_{j+1}) - f(t_j) = f'(t_j)(t_{j+1} - t_j) + \mathcal{E}(t_j, t_{j+1}).$$

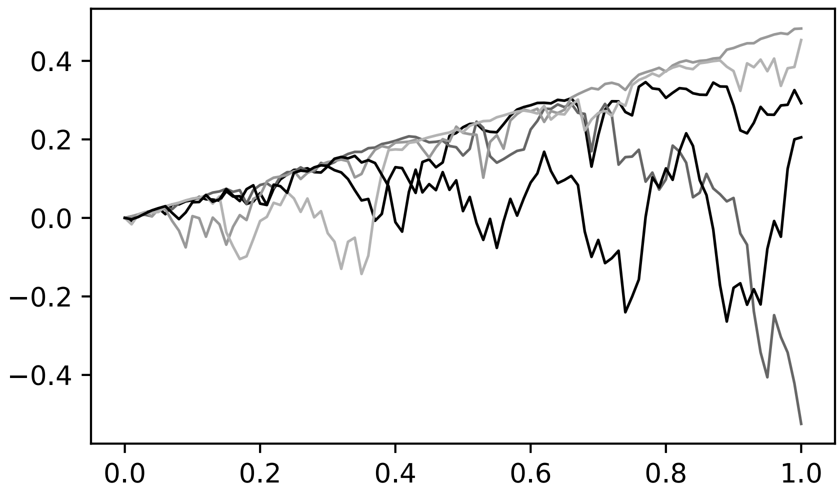

Figure 5.6. A sample of 5 paths of the martingale $\left(\cos B_t - 1 + \frac{1}{2}\int_0^t \cos B_s \, \mathrm{d}s, t \leq 1\right)$.

Therefore, we have by taking the limit of large n

$$f(t) - f(0) = \lim_{n\to\infty} \sum_{j=0}^{n} f'(t_j)(t_{j+1} - t_j) + \sum_{j=0}^{n} \mathcal{E}(t_j, t_{j+1}) = \int_0^t f'(s)\,\mathrm{d}s + 0.$$

The idea for Itô's formula is similar to the above with two big differences: First, we will consider a function f of *space* and not *time*. Second, we shall need a Taylor approximation to the second order around a point x: If $f \in \mathcal{C}^2(\mathbb{R})$, we have

$$f(y) - f(x) = f'(x)(y-x) + \frac{1}{2}f''(x)(x-y)^2 + \mathcal{E}(x,y), \tag{5.19}$$

where $\mathcal{E}(x,y)$ is an error term that now goes to 0 faster than $(x-y)^2$ as $y \to x$.

Proof of Theorem 5.24. Recall that by assumption $f \in \mathcal{C}^2(\mathbb{R})$. We will prove the particular case where f is 0 outside a bounded interval. This implies that both derivatives are bounded, since they are continuous functions on a bounded interval. We first prove the formula for a fixed t. Then we generalize to processes on $[0,T]$. Consider a partition $(t_j, j \leq n)$ of $[0,t]$. From equation (5.19), we get

$$\begin{aligned} &f(B_t) - f(B_0) \\ &= \sum_{j=0}^{n-1} f'(B_{t_j})(B_{t_{j+1}} - B_{t_j}) + \sum_{j=0}^{n-1} \frac{1}{2} f''(B_{t_j})(B_{t_{j+1}} - B_{t_j})^2 + \sum_{j=0}^{n} \mathcal{E}(B_{t_j}, B_{t_{j+1}}). \end{aligned} \tag{5.20}$$

As $n \to \infty$, the first term converges (as a random variable in L^2) to the Itô integral. This is how we proved Proposition 5.7 using simple processes. We claim the second term converges to the Riemann integral. To see this, consider the corresponding Riemann sum

$$\sum_{j=0}^{n-1} f''(B_{t_j})(t_{j+1} - t_j).$$

This term converges almost surely to the Riemann integral $\int_0^t f''(B_s)\,\mathrm{d}s$ since f'' is continuous. It also converges in L^2 by Theorem 4.40, since f'' is bounded by assumption. Therefore, to show the second term converges to the same limit, it suffices to show that the L^2-distance between the second term and the Riemann sum goes to 0; i.e.,

$$\lim_{n\to\infty} \mathbf{E}\left[\left(\sum_{j=0}^{n-1} f''(B_{t_j})\left\{(B_{t_{j+1}} - B_{t_j})^2 - (t_{j+1} - t_j)\right\}\right)^2\right] = 0. \tag{5.21}$$

This is in the same spirit as the proof of the quadratic variation of Brownian motion in Theorem 3.8. To lighten notation, define the variables $X_j = (B_{t_{j+1}} - B_{t_j})^2 - (t_{j+1} - t_j)$, $j \le n-1$. We expand the square in (5.21) to get

$$\sum_{j,k=0}^{n-1} \mathbf{E}\left[f''(B_{t_j})f''(B_{t_k})X_jX_k\right].$$

For $j < k$, we condition on $\mathcal{F}_{t_k}$ to get that the summand is 0 by Proposition 4.19 and since $\mathbf{E}[(B_{t_{k+1}} - B_{t_k})^2] = t_{k+1} - t_k$. For $j = k$, the sum is

$$\sum_{j=0}^{n-1} \mathbf{E}\left[(f''(B_{t_j}))^2\{(B_{t_{j+1}} - B_{t_j})^2 - (t_{j+1} - t_j)\}^2\right].$$

By expanding the square again and conditioning on $\mathcal{F}_{t_j}$, we have by independence of the increments

$$\begin{aligned}&\sum_j \mathbf{E}\left[(f''(B_{t_j}))^2\right]\{3(t_{j+1} - t_j)^2 + (t_{j+1} - t_j)^2 - 2(t_{j+1} - t_j)^2\}^2 \\ &\quad = 2\sum_j \mathbf{E}\left[(f''(B_{t_j}))^2\right](t_{j+1} - t_j)^2.\end{aligned}$$

Since f'' is bounded, this term goes to 0 exactly as in the proof of Theorem 3.8. It remains to handle the error term (5.20). This follows the same idea as for the second term and we omit it.

To extend the formula to the whole interval $[0, T]$, notice that the processes of both sides of equation (5.15) have continuous paths. Since they are equal (with probability one) at any fixed time by the above argument, they must be equal for any countable set of times with probability one; see Exercise 1.5. It suffices to consider the processes on the rational times in $[0, T]$, which are dense in $[0, T]$. Since the paths are continuous and they are equal on these times, they must be equal at all times on $[0, T]$. □

Recall from equation (5.16) that Itô's formula can be conveniently written in the *differential form*:

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

This notation has no meaning by itself. It is a compact way to write equation (5.15). This allows us to derive an easy and useful computational formula: If we blindly apply the classical differential to f to second order in the Taylor expansion, we formally obtain

$$\mathrm{d}f(B_t) = f'(B_t)\,\mathrm{d}B_t + \frac{1}{2}f''(B_t)(\mathrm{d}B_t)^2. \tag{5.22}$$

Therefore, Itô's formula is equivalent to applying the rule $\mathrm{d}t = \mathrm{d}B_t \cdot \mathrm{d}B_t$. In fact, it is counterproductive to learn Itô's formula by heart. It is much better to simply compute the differential up to the second order and apply the following simple *rules of Itô calculus*:

$$\begin{array}{c||c|c} \cdot & \mathrm{d}t & \mathrm{d}B_t \\ \hline\hline \mathrm{d}t & 0 & 0 \\ \hline \mathrm{d}B_t & 0 & \mathrm{d}t \end{array} \quad . \tag{5.23}$$

It is not hard to extend Itô's formula to a function $f(t,x)$ of both *time* and *space*:

$$\begin{aligned} f : [0,T]\times\mathbb{R} &\to \mathbb{R} \\ (t,x) &\mapsto f(t,x). \end{aligned} \tag{5.24}$$

Such functions have partial derivatives that are themselves functions of time and space. We will use the following notation for the partial derivatives:

$$\partial_0 f(t,x) = \frac{\partial f}{\partial t}(t,x), \qquad \partial_1 f(t,x) = \frac{\partial f}{\partial x}(t,x), \qquad \partial_1^2 f(t,x) = \frac{\partial^2 f}{\partial x^2}(t,x). \tag{5.25}$$

The reason for this notation is to avoid confusion between the variable that is *being differentiated* and the value of time and space at which the derivative is *being evaluated*. It might appear strange at first, but it will avoid confusion down the road (especially when dealing with several space variables in Chapter 6). To apply Itô's formula, we will need that the partial derivative with respect to time $\partial_0 f$ exists and is continuous as a function on $[0,T]\times\mathbb{R}$ and that the first and second partial derivatives in space $\partial_1 f$ and $\partial_1^2 f$ exist and are continuous. We say that such a function f is in $\mathcal{C}^{1,2}([0,T]\times\mathbb{R})$.

Proposition 5.28 (Itô's formula). *Let $(B_t, t \le T)$ be a standard Brownian motion on $[0,T]$. Consider a function f of time and space with $f \in \mathcal{C}^{1,2}([0,T]\times\mathbb{R})$. Then, with probability one, we have for every $t \in [0,T]$,*

$$f(t,B_t) - f(0,B_0) = \int_0^t \partial_1 f(s,B_s)\,\mathrm{d}B_s + \int_0^t \left\{\partial_0 f(s,B_s) + \frac{1}{2}\partial_1^2 f(s,B_s)\right\}\mathrm{d}s.$$

Or in differential form we have

$$\mathrm{d}f(t,B_t) = \partial_1 f(t,B_t)\,\mathrm{d}B_t + \left(\partial_0 f(t,B_t) + \frac{1}{2}\partial_1^2 f(t,B_t)\right)\mathrm{d}t.$$

Note that the notation $\partial_0 f(t,B_t)$ stands for the function $\partial_0 f$ evaluated at the point (t,B_t), and the notation $\partial_1^2 f(t,B_t)$ stands for the function $\partial_1^2 f$ evaluated at the point (t,B_t).

Proof. The idea of the proof is similar as for a function of space only, as it depends on a Taylor approximation and on the quadratic variation. Here, however, we need to apply Taylor approximation to second order in space and to first order in time. We then get something of the following form:

$$\begin{aligned} &f(t,B_t) - f(0,B_0) \\ &\quad= \sum_{j=0}^{n-1} \partial_1 f(t_j,B_{t_j})(B_{t_{j+1}} - B_{t_j}) + \partial_0 f(t_j,B_{t_j})(t_{j+1}-t_j) + \frac{1}{2}\partial_1^2 f(t_j,B_{t_j})(B_{t_{j+1}} - B_{t_j})^2 \\ &\quad\quad + \partial_1\partial_0 f(t_j,B_{t_j})(B_{t_{j+1}} - B_{t_j})(t_{j+1}-t_j) + \mathcal{E}. \end{aligned}$$

The first line becomes the integrals in Itô's formula. We see a new animal in the second line: the mixed derivative $\partial_0\partial_1 f$ in time and space. This term is related to the limit in the *cross variation* between B_t and t given by

$$\lim_{n\to\infty}\sum_{j=0}^{n-1}(B_{t_{j+1}}-B_{t_j})(t_{j+1}-t_j).$$

It can be shown that it goes to 0 in a suitable sense; see Exercise 5.15. This is the rigorous reason for the rule $\mathrm{d}t\cdot\mathrm{d}B_t=0$. Once this is known, the rest of the proof is done similarly to the one for a function of space only. We do notice though that the formula is easy to derive once we accept the rules of Itô calculus. By writing the differential to second order in space and to first order in time and applying the rules of Itô calculus, we get

$$\mathrm{d}f(t,B_t)=\partial_1 f(t,B_t)\,\mathrm{d}B_t+\left(\partial_0 f(t,B_t)+\frac{1}{2}\partial_1^2 f(t,B_t)\right)\mathrm{d}t. \qquad \square$$

As in the one variable case, we get a corollary to construct martingales:

Corollary 5.29 (Brownian martingales)**.** *Let $(B_t,t\le T)$ be a standard Brownian motion. Consider $f\in\mathcal{C}^{1,2}([0,T]\times\mathbb{R})$ such that the process $(\partial_1 f(t,B_t),t\le T)\in\mathcal{L}^2_c(T)$. Then the process*

$$\left(f(t,B_t)-\int_0^t\left\{\partial_0 f(s,B_s)+\frac{1}{2}\partial_1^2 f(s,B_s)\right\}\mathrm{d}s,t\le T\right)$$

is a martingale for the Brownian filtration. In particular, if $f(t,x)$ satisfies the partial differential equation $\partial_0 f=-\frac{1}{2}\partial_1^2 f$, then the process $(f(t,B_t),t\le T)$ is itself a martingale.

We now catch a glimpse of a powerful connection between two fields of mathematics: *The study of martingales is closely related to the study of differential equations.* We will see this connection in action in the gambler's ruin problem in Section 5.5. This is also explored further in Chapter 8.

Example 5.30. Consider the function $f(t,x)=tx$. In this case, we have $\partial_0 f=x$, $\partial_1 f=t$, and $\partial_1^2 f=0$. Itô's formula yields

$$\mathrm{d}(tB_t)=t\,\mathrm{d}B_t+B_t\,\mathrm{d}t.$$

Therefore, the process $M_t=tB_t-\int_0^t B_s\,\mathrm{d}s$ is a martingale for the Brownian filtration. It is also a Gaussian process by Corollary 5.18. The mean is 0 and the covariance is by Corollary 5.15

$$\mathbf{E}[M_tM_{t'}]=\int_0^{t\wedge t'}s^2\,\mathrm{d}s=\frac{(t\wedge t')^3}{3}.$$

Example 5.31 (Geometric Brownian motion revisited)**.** We know from Example 4.28 that geometric Brownian motion is a martingale for the choice $\mu=\frac{-1}{2}\sigma^2$. How does this translate in terms of Itô integrals? Note that $S_t=f(t,B_t)$ for the function of time and space $f(t,x)=e^{\sigma x+\mu t}$ with $S_0=1$. The relevant partial derivatives are $\partial_0 f=\mu f$,

$\partial_1 f = \sigma f$, and $\partial_1^2 f = \sigma^2 f$. Therefore, developing the function f to second order in space and first order in time and using the rules of Itô calculus yield

$$\begin{aligned} \mathrm{d}S_t = \ &\mathrm{d}f(t, B_t) = \partial_0 f(t, B_t)\,\mathrm{d}t + \partial_1 f(t, B_t)\,\mathrm{d}B_t + \frac{1}{2}\partial_1^2 f(t, B_t)(\mathrm{d}B_t)^2 \\ &= \sigma f(t, B_t)\,\mathrm{d}B_t + \left(\mu + \frac{1}{2}\sigma^2\right) f(t, B_t)\,\mathrm{d}t. \end{aligned}$$

In the integral notation, this is

$$S_t = 1 + \int_0^t \sigma f(s, B_s)\,\mathrm{d}B_s + \int_0^t \left(\mu + \frac{1}{2}\sigma^2\right) f(s, B_s)\,\mathrm{d}s.$$

We see that we have a martingale if $\mu = \frac{-1}{2}\sigma^2$ as expected. It is not hard to check that the integrand is in $\mathcal{L}_c^2(T)$ for any $T > 0$ (see Exercise 5.6).

5.5. Gambler's Ruin for Brownian Motion with Drift

We solved the gambler's ruin problem for standard Brownian motion in Example 4.41. We now deal with the case where a drift is present. Consider the Brownian motion with drift

$$X_t = \sigma B_t + \mu t,$$

where $(B_t, t \geq 0)$ is a standard Brownian motion. We assume that $\mu > 0$. Therefore, there is a bias upward. This is important!

We consider for $a, b > 0$ the first passage time of the level a or $-b$

$$\tau = \min\{t \geq 0 : X_\tau > a \text{ or } X_\tau < -b\}.$$

The problem consists of computing $\mathbf{P}(X_\tau = a)$. Recall that in the case of no drift, this probability was $b/(a+b)$. To solve the problem, we need to find a good martingale of X_t that gives us the desired probability using Doob's optional stopping theorem. It is not

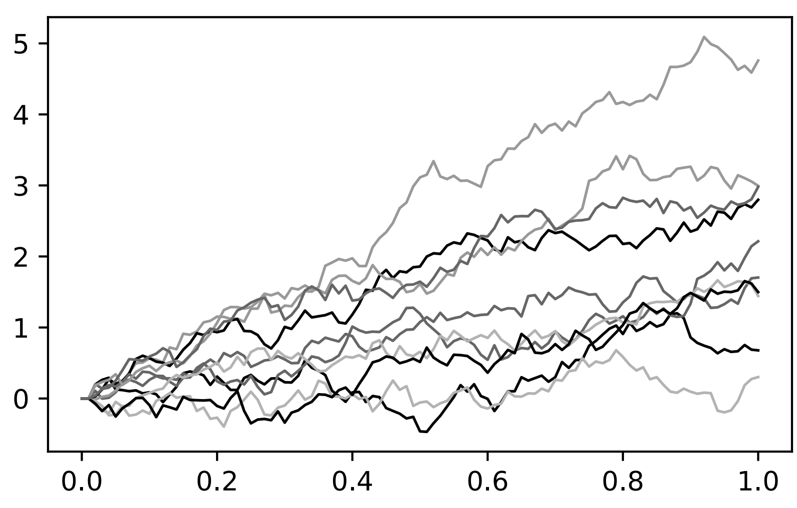

Figure 5.7. A sample of 10 paths of the process $X_t = B_t + 2t$.

hard to see here that $\tau < \infty$ (by the same argument as for standard Brownian motion). We assume the martingale is of the simplest form, that is, a function of X_t:

$$M_t = g(X_t),$$

for some function g to be found. The function needs to satisfy two properties:

- The process $(g(X_t), t \geq 0)$ is a martingale for the Brownian filtration.
- The values at $x = a$ or $x = -b$ are $g(a) = 1$ and $g(-b) = 0$.

The first condition implies by Corollary 4.38 that

$$\mathbf{E}[g(X_\tau)] = g(0).$$

The second condition is a convenient choice since we have

$$\mathbf{E}[g(X_\tau)] = g(a)\mathbf{P}(X_\tau = a) + g(-b)\mathbf{P}(X_\tau = -b) = \mathbf{P}(X_\tau = a).$$

Combining these two, we see that the ruin problem is reduced to finding $g(0)$ since

$$g(0) = \mathbf{E}[g(X_\tau)] = \mathbf{P}(X_\tau = a).$$

What are the conditions on g for $g(X_t)$ to be a martingale? Note that $g(X_t) = g(\sigma B_t + \mu t)$ is an explicit function of t and B_t: $f(t, x) = g(\sigma x + \mu t)$. By the chain rule, we have

$$\partial_0 f(t, x) = \mu g'(\sigma x + \mu t), \quad \partial_1 f(t, x) = \sigma g'(\sigma x + \mu t), \quad \partial_1^2 f(t, x) = \sigma^2 g''(\sigma x + \mu t).$$

By Corollary 5.29, for $g(X_t)$ to be martingale, we need g to satisfy the ordinary differential equation

$$\mu g' = \frac{-\sigma^2}{2} g''.$$

This is easy to solve just by integrating, and we get $g(y) = Ce^{-2\mu/\sigma^2} + C'$ for two constants C and C'. The boundary conditions $g(a) = 1$ and $g(-b) = 0$ determine those constants, and we finally have

$$g(y) = \frac{1 - e^{-2\mu(y+b)/\sigma^2}}{1 - e^{-2\mu(a+b)/\sigma^2}}. \tag{5.26}$$

(Notice that g is bounded, and so is the martingale $g(X_t)$. Hence, there is no problem in applying Corollary 4.38.) In particular, we get the answer to our initial question

$$\mathbf{P}(X_\tau = a) = g(0) = \frac{1 - e^{-2\mu b/\sigma^2}}{1 - e^{-2\mu(a+b)/\sigma^2}}. \tag{5.27}$$

This formula is tested numerically in Numerical Project 5.6. It is good to take a step back and look at what we have achieved:

- If we take the case $\mu = \sigma = 1$ and $a = b = 1$, then the probability is

  $$\mathbf{P}(X_\tau = a) = \frac{1 - e^{-2}}{1 - e^{-4}} = 0.881\ldots.$$

  Compare this to the case $\mu = 0$, where this probability is $1/2$!
- Notice that we reduced the problem of computing a probability to solving a *differential equation with boundary conditions*. This is amazing!
- Our answer is even more general. Had we started the process at $y \in [-b, a]$ instead of 0, then the probability would have been $g(y)$ given in equation (5.26).

- The identity

$$\mathbf{E}[g(X_\tau)] = \mathbf{P}(X_\tau = a)$$

is very intuitive. Since $g(a) = 1$ and $g(-b) = 0$, the paths that hit a (success) contribute to the expectation whereas the ones that hit $-b$ (failure) do not. Therefore, the *proportion of paths hitting* a, or in other words the probability, is given by averaging the Bernoulli variable $g(X_\tau)$ over all paths.

- Let's look at the limiting cases. If we take $b \to \infty$, then we get

$$\mathbf{P}(X_\tau = a) \to 1, \quad b \to \infty, \tag{5.28}$$

which makes sense since the drift is upward, and we already know that it is the case when $\mu = 0$. On the other hand, if $a \to \infty$, then we get

$$\mathbf{P}(X_\tau = -b) \to e^{\frac{-2\mu}{\sigma^2} b}. \tag{5.29}$$

It is not 1. The formula is telling us that even when $a \to \infty$, there are some paths that will never hit $-b$, because of the upward drift, no matter how small the drift is!

5.6. Tanaka's Formula

What happens to Itô's formula when f is not in $\mathcal{C}^2$? It turns out that in some cases we can still express $f(B_t)$ as a sum of a martingale and a process with bounded variation. The most famous example is when $f(B_t) = |B_t|$. (The absolute value is continuous, but the first and second derivative do not exist at 0.) Note that in this case, one can see the paths of the process $f(B_t)$ as the paths of a Brownian motion reflected on the x-axis. In this case, one recovers some, but not all, of Itô's formula as the following theorem shows.

Theorem 5.32 (Tanaka's formula)**.** *Let $(B_t, t \geq 0)$ be a standard Brownian motion. There exists an increasing adapted process $(L_t, t \geq 0)$, called the local time of the Brownian motion at* 0*, such that*

$$|B_t| = \int_0^t \operatorname{sgn}(B_s)\,\mathrm{d}B_s + L_t, \quad t \geq 0,$$

where $\operatorname{sgn}(x) = 1$ *if* $x \geq 0$ *and* $\operatorname{sgn}(x) = -1$ *if* $x < 0$.

As for the case of Itô's formula where $f \in \mathcal{C}^2(\mathbb{R})$, the function of Brownian motion is expressed as a sum of an Itô integral and of a process of bounded variation, since L_t is increasing in t. (The theorem is not surprising in view of the Doob-Meyer decomposition in Remark 4.31.) The theorem is illustrated in Figure 5.8. It turns out that the Itô integral has the distribution of Brownian motion; see Section 7.6. The integrand $\operatorname{sgn}(B_s)$ is not in $\mathcal{L}_c^2(T)$ but it can be shown that it falls in the first case in Remark 5.22, so that it is a martingale. It is the investing strategy that equals $+1$ when the Brownian motion is positive, and -1 when it is negative. The *local time at* 0, denoted by L_t, should be interpreted as the amount of time on $[0,t]$ that the Brownian motion has spent at 0. More precisely, it is equal to

$$L_t = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t \mathbf{1}_{\{s\in[0,t]\,:\,|B_s|\le\varepsilon\}}\, \mathrm{d}s \ \text{ in } L^2.$$

The existence of the process L_t is not obvious and is a consequence of the proof. The strategy of the proof is to use Itô's formula on an approximation of the absolute value that is in $\mathcal{C}^2$. The proof is technical and will be skipped. However, the result is not hard to simulate; see Numerical Project 5.8.

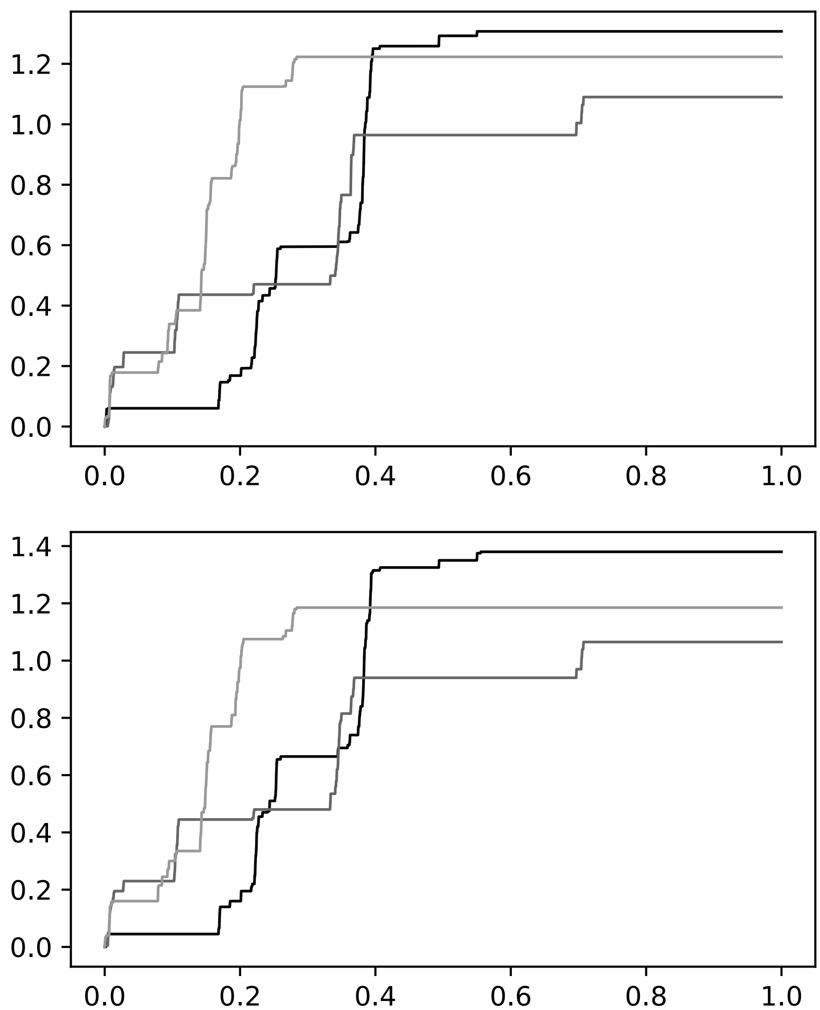

Figure 5.8. Sample of 3 paths of the processes $|B_t| - \int_0^t \operatorname{sgn}(B_s)\,\mathrm{d}B_s$ (top) and of L_t with the approximation $\varepsilon = 0.001$ (bottom) and step size 1/100,000.

5.7. Numerical Projects and Exercises

5.1. **Itô integrals.** Following the procedure outlined in Example 5.13, sample 10 paths of the following three processes on $[0,1]$ using a 0.001 discretization:
(a) $\int_0^t 4B_s^3 \, \mathrm{d}B_s$,
(b) $\int_0^t \cos B_s \, \mathrm{d}B_s$,
(c) $1 - \int_0^t e^{s/2} \sin B_s \, \mathrm{d}B_s$.

5.2. **Ornstein-Uhlenbeck process and Brownian bridge revisited.** We saw in Examples 5.19 and 5.20 that the Ornstein-Uhlenbeck process and the Brownian bridge can be expressed in terms of Wiener integrals; see equations (5.9) and (5.10). Use these representations to sample 100 paths of each process on $[0,1]$ with a discretization of 0.001. Compare this to the samples generated in Numerical Projects 2.3 and 2.5 using Cholesky decomposition.

5.3. **Itô's formula.** Sample a *single path* of the following three processes on $[0,1]$ using a 0.01 discretization and compare to the processes in Numerical Project 5.1 constructed *on the same Brownian path*:
(a) $B_t^4 - 6\int_0^t B_s^2 \, \mathrm{d}s$,
(b) $\sin B_t + \frac{1}{2}\int_0^t \sin B_s \, \mathrm{d}s$,
(c) $e^{t/2} \cos B_t$.
The command `random.seed` is useful to work on the same outcome.

5.4. **A path-dependent integrand.** Consider the process $(X_t, t \in [0,1])$ with $X_t = \int_0^t B_s \, \mathrm{d}B_s$. We construct the process $I_t = \int_0^t X_s \, \mathrm{d}B_s$ as in Example 5.17. Following the procedure outlined in Example 5.13, sample 10 paths of this process on $[0,1]$ using a 0.01 discretization.

5.5. **Convergence of Itô's formula.** Consider the two processes $(I_t, t \in [0,1])$ and $(J_t, t \in [0,1])$ in Example 5.26 defined by the two sides of equation (5.17) on the interval $[0,1]$. Sample 100 paths of these two processes for each of the discretization $0.1, 0.01, 0.001, 0.0001$. Estimate $\mathbf{E}[|I_1 - J_1|]$ for each of these time steps. What do you notice?

5.6. **Testing the solution to the gambler's ruin.** Let's test equation (5.29).
(a) Sample 10,000 paths of Brownian paths with drift $\mu = 1$ and volatility $\sigma = 1$ on $[0,5]$ for a step size of 0.01.
(b) Count the proportion of those paths that reach -1 on the time interval and compare with equation (5.29). Repeat the experiment for a step size of 0.001.
The experiment on $[0,5]$ *gives an approximation of the probability on* $[0,\infty)$. *It turns out that the probability on a finite interval can be computed exactly. See Exercise* 9.5.

5.7. **The integral of a process not in $\mathcal{L}_c^2(T)$.**
(a) Sample 100 paths of the process $Z_t = \exp B_t^2$, $t \in [0,10]$. This process in not in $\mathcal{L}_c^2(10)$ as shown in Exercise 5.7.
(b) Sample and plot 100 paths of the process $\int_0^t Z_s \, \mathrm{d}B_s$, $t \in [0,10]$. What do you notice?

5.8. **Tanaka's formula.** Generate 10 paths of Brownian motion on $[0,1]$ using a discretization of $1/1{,}000{,}000$.

(a) Plot the paths of the process $|B_t| - \int_0^t \operatorname{sgn}(B_s)\,\mathrm{d}B_s$ on $[0,1]$.

(b) Plot the paths of the process L_t^ε for $\varepsilon = 0.001$ where

$$L_t^\varepsilon = \frac{1}{2\varepsilon}|\{s \in [0,t] : |B_t| < \varepsilon\}|\,.$$

In other words, this is the amount of time before time t spent by Brownian motion in the interval $[-\varepsilon, \varepsilon]$ (rescaled by $1/2\varepsilon$).

Exercises

5.1. **Stopped martingales are martingales.** Let $(M_n, n = 0,1,2,\dots)$ be a martingale in discrete time for the filtration $(\mathcal{F}_n, n \geq 0)$. Let τ be a stopping time for the same filtration. Use the martingale transform with the process

$$X_n(\omega) = \begin{cases} +1 & \text{if } n < \tau(\omega), \\ 0 & \text{if } n \geq \tau(\omega) \end{cases}$$

to show that the stopped martingale $(M_{\tau \wedge n}, n \geq 0)$ is a martingale.

5.2. **Itô integral of a simple process**. Consider $(I_s, s \leq 1)$ the Itô integrals in Example 5.1.

(a) Argue that $(I_{1/3}, I_{2/3}, I_1)$ is a Gaussian vector.

(b) Compute the mean and the covariance matrix of $(I_{1/3}, I_{2/3}, I_1)$.

(c) Compute $\mathbf{E}[B_1 I_1]$. Are the random variables B_1 and I_1 independent? Briefly justify.

5.3. **Convergence in L^2 implies convergence of first and second moments.** Let $(X_n, n \geq 0)$ be a sequence of random variables that converge to X in $L^2(\Omega, \mathcal{F}, \mathbf{P})$.

(a) Show that $\mathbf{E}[X_n^2]$ converges to $\mathbf{E}[X^2]$.
Hint: Write $X = (X - X_n) + X_n$. The Cauchy-Schwarz inequality might be useful.

(b) Show that $\mathbf{E}[X_n]$ converges to $\mathbf{E}[X]$.
Hint: Write $|\mathbf{E}[X_n] - \mathbf{E}[X]|$ and use Jensen's inequality twice.

5.4. **Increments of martingales are uncorrelated.**

(a) Let $(M_t, t \geq 0)$ be a square-integrable martingale for the filtration $(\mathcal{F}_t, t \geq 0)$. Use the properties of conditional expectation to show that for $t_1 \leq t_2 \leq t_3 \leq t_4$, we have

$$\mathbf{E}[(M_{t_2} - M_{t_1})(M_{t_4} - M_{t_3})] = 0.$$

(b) Let $(B_t, t \geq 0)$ be a standard Brownian motion, and let $(X_t, t \leq T)$ be a process in $\mathcal{L}_c^2(T)$. Use part (a) to show that the covariance between integrals at different times $t < t'$ is

$$\mathbf{E}\left[\left(\int_0^t X_s\,\mathrm{d}B_s\right)\left(\int_0^{t'} X_s\,\mathrm{d}B_s\right)\right] = \int_0^t \mathbf{E}[X_s^2]\,\mathrm{d}s.$$

This motivates the natural notation

$$\int_t^{t'} X_s \, \mathrm{d}B_s = \int_0^{t'} X_s \, \mathrm{d}B_s - \int_0^t X_s \, \mathrm{d}B_s.$$

5.5. **Mean and variance of martingale transforms.** Let $(M_t, t \le T)$ be a square-integrable martingale for a filtration $(\mathcal{F}_t, t \le T)$, and let X be a simple process in $\mathcal{S}(T)$. Compute the mean and the variance of the martingale transform of X with respect to M on $[0, T]$.

5.6. **Geometric Brownian motion is in $\mathcal{L}_c^2$.** Let $M_t = \exp(\sigma B_t - \sigma^2 t/2)$ be a geometric Brownian motion. Verify that the process $(M_t, t \le T)$ is in $\mathcal{L}_c^2(T)$ for any $T > 0$.

5.7. **A process that is not in $\mathcal{L}_c^2(T)$.** Consider the process $(e^{B_t^2}, t \le T)$. Show that it is not in $\mathcal{L}_c^2(T)$ for $T > 1/4$.

5.8. **Practice on Itô integrals.** Consider the two processes

$$X_t = \int_0^t (1-s) \, \mathrm{d}B_s, \qquad Y_t = \int_0^t (1+s) \, \mathrm{d}B_s.$$

(a) Find the mean and the covariance of the process $(X_t, t \ge 0)$. What is its distribution?
(b) Find the mean and the covariance of the process $(Y_t, t \ge 0)$. What is its distribution?
(c) For which time t, if any, do we have that X_t and Y_t are uncorrelated? Are X_t and Y_t independent at these times?

5.9. **Practice on Itô integrals.** Consider the process $(X_t, t \ge 0)$ given by

$$X_t = \int_0^t \sin s \ \mathrm{d}B_s.$$

(a) Argue briefly that this process is Gaussian. Find the mean and the covariance matrix.
(b) Write the covariance matrix for $(X_{\pi/2}, X_\pi)$ (i.e., the process at time $t = \pi/2$ and $t = \pi$). Write down a double integral for the probability $\mathbf{P}(X_{\pi/2} > 1, X_\pi > 1)$.
(c) On the same Brownian motion, consider the process $Y_t = \int_0^t \cos s \ \mathrm{d}B_s$. Find for which time t the variables X_t and Y_t are independent.

5.10. **Not everything is a martingale.**
(a) Use the representation of the Ornstein-Uhlenbeck process in Example 5.19 to show that it is not a martingale for the Brownian filtration.
(b) Use the representation of the Brownian bridge in Example 5.20 to show that it is not a martingale for the Brownian filtration.
(c) Show that the process $(\int_0^t B_s \, \mathrm{d}s, t \ge 0)$ is not a martingale for the Brownian fltration.

5.11. **Practice on Itô integrals.** Let $(B_t, t \geq 0)$ be a Brownian motion defined on $(\Omega, \mathcal{F}, \mathbf{P})$. We define for $t \geq 0$ the process

$$X_t = \int_0^t \operatorname{sgn}(B_s)\, \mathrm{d}B_s,$$

where $\operatorname{sgn}(x) = -1$ if $x < 0$ and $\operatorname{sgn}(x) = +1$ if $x \geq 0$.
The integral is well-defined even though $s \mapsto \operatorname{sgn}(B_s)$ is not continuous.
(a) Compute the mean and the covariance of the process $(X_t, t \geq 0)$.
(b) Show that X_t and B_t are uncorrelated for all $t \geq 0$.
(c) Show that X_t and B_t are not independent. (Use $B_t^2 = 2\int_0^t B_s\, \mathrm{d}B_s + t$.)
It turns out that $(X_t, t \geq 0)$ is a standard Brownian motion. See Theorem 7.26.

5.12. **Integration by parts for some Itô integrals.** Let $g \in \mathcal{C}^2(\mathbb{R})$ and $(B_t, t \geq 0)$, a standard Brownian motion.
(a) Use Itô's formula to prove that for any $t \geq 0$

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

(b) Use the above to show that the process given by

$$X_t = t^2 B_t - 2\int_0^t s B_s\, \mathrm{d}s$$

is Gaussian. Find its mean and its covariance.

5.13. **Some practice with Itô's formula.** Let $(B_t, t \geq 0)$ be a standard Brownian motion. For each of the processes $(X_t, t \leq T)$ below:
- Determine if they are martingales for the Brownian filtration. If not, find a compensator for it.
- Find the mean, the variance, and the covariance.
- Is the process Gaussian? Argue briefly.

(a) $X_t = \int_0^t \cos s\ \mathrm{d}B_s$.
(b) $X_t = B_t^4$.
(c) $X_t = e^{t/2}\cos B_t$.
Hint: If Z is standard Gaussian, then $\mathbf{E}[\sin^2(\sigma Z)] = \frac{1-e^{-2\sigma^2}}{2}$.
(d) $Z_t = (B_t + t)\exp(-B_t - \frac{t}{2})$.

5.14. **Gaussian moments using Itô**. Let $(B_t, t \in [0,1])$ be a Brownian motion. Use Itô's formula to show that for $k \in \mathbb{N}$

$$\mathbf{E}[B_t^k] = \frac{1}{2}k(k-1)\int_0^t \mathbf{E}[B_s^{k-2}]\, \mathrm{d}s.$$

Conclude from this that $\mathbf{E}[B_t^4] = 3t^2$ and $\mathbf{E}[B_t^6] = 15t^3$.

5.15. **Cross-variation of t and B_t.** Let $(t_j, j \leq n)$ be a sequence of partitions of $[0,t]$ such that $\max_j |t_{j+1} - t_j| \to 0$ as $n \to \infty$. Prove that

$$\lim_{n\to\infty} \sum_{j=0}^{n-1} (t_{j+1} - t_j)(B_{t_{j+1}} - B_{t_j}) = 0 \quad \text{in } L^2.$$

This justifies the rule $\mathrm{d}t \cdot \mathrm{d}B_t = 0$. Can you also justify the rule $\mathrm{d}t \cdot \mathrm{d}t = 0$?

5.16. **Exercise on Itô's formula.** Consider for $t \geq 0$ the process

$$X_t = \exp(tB_t).$$

(a) Find the mean and the variance of this process.
(b) Use Itô's formula to write the process in terms of an Itô integral and a Riemann integral. Find a compensator C_t so that $X_t - C_t$ is a martingale.
(c) Argue that $(e^{tB_t}, t \leq T)$ is in $\mathcal{L}_c^2(T)$ for any $T > 0$, so that $\int_0^t e^{sB_s}\,\mathrm{d}B_s$ makes sense.
(d) Show that the covariance between B_t and $\int_0^t e^{sB_s}\,\mathrm{d}B_s$ is

$$\int_0^t e^{s^3/2}\,\mathrm{d}s.$$

5.17. **Itô's formula and optional stopping**. Let $(B_t, t \geq 0)$ be a standard Brownian motion. Consider for $a, b > 0$ the hitting time

$$\tau = \min_{t \geq 0}\{t : B_t \geq a \text{ or } B_t \leq -b\}.$$

The goal of this exercise is to compute $\mathbf{E}[\tau B_\tau]$.
(a) Let $f(t, x)$ be a function of the form

$$f(t, x) = tx + g(x).$$

Find an ODE for the function f for which $\partial_0 f = -\frac{1}{2}\partial_1^2 f$. Solve this ODE.
(b) Argue briefly that the process $(f(t, B_t), t \geq 0)$ is a continuous martingale.
(c) Use this to show that

$$\mathbf{E}[\tau B_\tau] = \frac{ab}{3}(a - b).$$

5.18. **A strange martingale.** Let $(B_t, t \geq 0)$ be a standard Brownian motion. Consider the process

$$M_t = \frac{1}{\sqrt{1-t}}\exp\left(\frac{-B_t^2}{2(1-t)}\right), \quad \text{for } 0 \leq t < 1.$$

(a) Show that M_t can be represented by

$$M_t = 1 + \int_0^t \frac{-B_s M_s}{1-s}\,\mathrm{d}B_s, \quad \text{for } 0 \leq t < 1.$$

(b) Deduce from the previous question that $(M_s, s \leq t)$ is a martingale for $t < 1$ and for the Brownian filtration.
(c) Show that $\mathbf{E}[M_t] = 1$ for all $t < 1$.
(d) Prove that $\lim_{t \to 1^-} M_t = 0$ almost surely.
(e) Argue (by contradiction) that $\mathbf{E}[\sup_{0 \leq t < 1} M_t] = +\infty$, where sup stands for the supremum.
Hint: Theorem 4.40 *is useful.*

5.19. ⋆ **L^2-limit of Gaussians is Gaussian.** Let $(X_n, n \geq 0)$ be a sequence of Gaussian random variables that converge to X in $L^2(\Omega, \mathcal{F}, \mathbf{P})$.

(a) Show that X is also Gaussian.
Hint: Use Exercise 1.14. *Use also the fact that there is a subsequence that converges almost surely; see Exercise* 3.14.

(b) Find its mean and variance in terms of X.

5.20. ⋆ **L^2 is complete.** We prove that the space $L^2(\Omega, \mathcal{F}, \mathbf{P})$ is complete; that is, if $(X_n, n \geq 1)$ is a Cauchy sequence in L^2 (see Remark 5.14), then there exists $X \in L^2(\Omega, \mathcal{F}, \mathbf{P})$ such that $X_n \to X$ in L^2.

(a) Argue from the definition of Cauchy sequence that we can find a subsequence $(X_{n_k}, k \geq 0)$ such that $\|X_m - X_{n_k}\| \leq 2^{-k}$ for all $m > n_k$, where $\|\cdot\|$ is the L^2-norm.

(b) Consider the candidate limit $\sum_{j=0}^{\infty}(X_{n_{j+1}} - X_{n_j})$ with $X_{n_0} = 0$. Show that this sum converges almost surely (so X is well-defined) by considering

$$\sum_{j=0}^{k} \mathbf{E}[|X_{n_{j+1}} - X_{n_j}|] .$$

(c) Show that $\|X - X_{n_k}\| \to 0$ as $k \to \infty$. Conclude that $\|X\| < \infty$. (This shows the convergence in L^2 along the subsequence!)

(d) Use again the Cauchy definition and the subsequence to show convergence of the whole sequence; i.e., $\|X_n - X\| \to 0$.

5.21. ⋆ **Another application of Doob's maximal inequality.** Let $(B_t, t \in [0,1])$ be a Brownian motion defined on $(\Omega, \mathcal{F}, \mathbf{P})$. Recall from Example 5.20 that the process

$$Z_t = (1-t)\int_0^t \frac{1}{1-s}\,\mathrm{d}B_s, \quad 0 \leq t < 1,$$

has the distribution of Brownian bridge on $[0,1)$. In this exercise we prove $\lim_{t\to 1} Z_t = 0$ almost surely as expected.

(a) Show that $\lim_{t\to 1} Z_t = 0$ in $L^2(\Omega, \mathcal{F}, \mathbf{P})$.

(b) Using Doob's maximal inequality of Exercise 4.19, show that

$$\mathbf{P}\left(\max_{t\in\left[1-\frac{1}{2^n}, 1-\frac{1}{2^{n+1}}\right]} |Z_t| > \delta\right) \leq \frac{1}{\delta^2}\frac{1}{2^{n-1}}.$$

(c) Deduce that $\lim_{t\to 1} Z_t = 0$ almost surely using the Borel-Cantelli lemma.

5.8. Historical and Bibliographical Notes

Stochastic integrals on Brownian motion were studied before Itô, notably by Wiener [**PW87**]. It was Kiyosi Itô who extended the definition to include integrands that were possibly dependent on the Brownian motion in a seminal paper during World War II [**Ito44**]. It is important to note that other definitions of stochastic integrals exist where the integrand is not necessarily adapted. The most famous one is arguably the

Stratonovich integral [**Str64**], for which the integrand depends symmetrically on the past and future. This definition has important applications in physics. The reader is referred to [**Øks03**] for an introduction to this integral and the comparison with the Itô integral. The proof of the continuity of the Itô integral in Theorem 5.12 is done in [**Ste01**]. Interestingly, Tanaka did not publish his formula. The first occurrence of the formula seems to have been in [**McK62**], giving credit to Tanaka.