6 Martingales in continuous time

Just as in discrete time, the notion of a martingale will play a key rôle in our continuous time models.

Recall that in discrete time, a sequence $X_0, X_1, \ldots, X_n$ for which $\mathbb{E}[|X_r|] < \infty$ for each r is a martingale if

$$\mathbb{E}[X_r | \mathcal{F}_{r-1}] = X_{r-1}.$$

The sequence $\{\mathcal{F}_r\}_{r=0}^n$ is called a *filtration*. In order to avoid abstract measure theory, we omit a detailed discussion of filtrations in continuous time. Instead we content ourselves with the following 'working definition'.

Definition 6.1 *The symbol $\mathcal{F}_t^X$ denotes the 'information generated by the stochastic process X on the interval $[0,t]$'. If, based upon observations of the trajectory $\{X(s), 0 \leq s \leq t\}$, it is possible to decide whether a given event A has occured or not, then we write this as*

$$A \in \mathcal{F}_t^X.$$

If the value of a stochastic variable can be completely determined given observations of the trajectory $\{X(s), 0 \leq s \leq t\}$ then we also write

$$Z \in \mathcal{F}_t^X.$$

If Y is a stochastic process such that we have $Y(t) \in \mathcal{F}_t^X$ for all $t \geq 0$, then we say that Y is adapted *to the filtration $\{\mathcal{F}_t^X\}_{t \geq 0}$.*

This definition is only intended to have intuitive content. Nevertheless, it is rather simple to use.

Example 6.2 *1. Define A by*

$$A = \{X(s) \leq 3.14, \forall s \leq 18\}.$$

Then $A \in \mathcal{F}_{18}^X$ but $A \notin \mathcal{F}_{17}^X$.

2. *For the event $A = \{X(10) > 8\}$, $A \in \mathcal{F}_s^X$ if and only if $s \geq 10$.*
3. *The stochastic variable*

$$Z(s) = \int_0^5 X(s) ds$$

is in $\mathcal{F}_s^X$ if and only if $s \geq 5$.

4. *If B_t is Brownian motion and $M_t = \max_{0 \leq s \leq t} B_s$, then M is adapted to the Brownian filtration.*
5. *If B_t is Brownian motion and $\tilde{M}_t = \max_{0 \leq s \leq t+1} B_s$, then $\tilde{M}$ is* not *adapted to the Brownian filtration.*

Definition 6.3 *Consider a probability space $(\Omega, \mathbb{P})$ and a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ on this space. An adapted family $\{M_t\}_{t \geq 0}$ of random variables on this space with $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 0$ is a* martingale *if, for any $s \leq t$,*

$$\mathbb{E}_{\mathbb{P}}[M_t | \mathcal{F}_s] = M_s.$$

Remarks:

1. Often we shall be sloppy about specifying the filtration. In all of our examples there will be a Brownian motion around and it will be implicit that the filtration is that generated by the Brownian motion.

2. Many stochastic calculus texts specify a probability space as $(\Omega, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, thereby specifying the filtration explicitly from the very beginning.

3. The *natural filtration* is the name given to the filtration for which $\mathcal{F}_t$ consists of those sets that can be decided by observing trajectories of the specified process up to time t. There are other possibilities, but we always use the natural filtration corresponding to Brownian motion in our examples.

Lemma 6.4 *If $\{B_t\}_{t\geq 0}$ is a standard Brownian motion generating the filtration $\{\mathcal{F}_t\}_{t\geq 0}$, then*

1. *B_t is an $\mathcal{F}_t$-martingale.*

2. *$B_t^2 - t$ is an $\mathcal{F}_t$-martingale.*

3. $$\exp\left(\sigma B_t - \frac{\sigma^2}{2}t\right)$$

   is an $\mathcal{F}_t$-martingale (called an exponential martingale*).*

Proof: The proofs are all rather similar. For example, consider $M_t = B_t^2 - t$. Evidently $\mathbb{E}[|M_t|] < \infty$. Now

$$\begin{aligned}
\mathbb{E}\left[\left. B_t^2 - B_s^2 \right| \mathcal{F}_s\right] &= \mathbb{E}\left[\left. (B_t - B_s)^2 + 2B_s(B_t - B_s) \right| \mathcal{F}_s\right] \\
&= \mathbb{E}\left[\left. (B_t - B_s)^2 \right| \mathcal{F}_s\right] + 2B_s \mathbb{E}\left[\left. (B_t - B_s) \right| \mathcal{F}_s\right] \\
&= t - s.
\end{aligned}$$

Thus

$$\begin{aligned}
\mathbb{E}\left[\left. B_t^2 - t \right| \mathcal{F}_s\right] = \mathbb{E}\left[\left. B_t^2 - B_s^2 + B_s^2 - (t-s) - s \right| \mathcal{F}_s\right] \\
= (t-s) + B_s^2 - (t-s) - s = B_s^2 - s.
\end{aligned}$$

□

Theorem 6.5 (Optional Sampling Theorem) *If $\{M_t\}_{t\geq 0}$ is a continuous martingale with respect to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ and if τ_1 and τ_2 are two stopping times such that $\tau_1 \leq \tau_2 \leq K$ where K is a finite real number, then M_{τ_2} is integrable (that is has finite expectation) and*

$$\mathbb{E}\left[\left. M_{\tau_2} \right| \mathcal{F}_{\tau_1}\right] = M_{\tau_1}, \qquad \mathbb{P} - a.s.$$

Remarks:

1. The term 'a.s.' (almost surely) means with ($\mathbb{P}$-) probability one.

2. Notice in particular that if τ is a bounded stopping time then $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$.

3. The name Optional Sampling Theorem is used because stopping times are sometimes also called *optional times*.

We illustrate the application of this by calculating the moment generating function for the hitting time T_a of level a by Brownian motion.

Proposition 6.6 *Let B_t be a Brownian motion and let $T_a = \inf\{s \geq 0 : B_s = a\}$ (or infinity if that set is empty). Then*

$$\mathbb{E}\left[e^{-\lambda T_a}\right] = e^{-\sqrt{2\lambda}|a|}.$$

Proof: We assume that $a \geq 0$. (The case $a < 0$ follows by symmetry.) We apply the Optional Sampling Theorem to the martingale

$$M_t = \exp\left(\sigma B_t - \frac{1}{2}\sigma^2 t\right).$$

We *cannot* apply it directly to T_a as it may not be bounded. Instead we take $\tau_1 = 0$ and $\tau_2 = T_a \wedge n$. This gives us that

$$\mathbb{E}[M_{T_a \wedge n}] = 1.$$

Now

$$0 \leq M_{T_a\wedge n} = \exp\left(\sigma B_{T_a\wedge n} - \frac{1}{2}\sigma^2 (T_a \wedge n)\right) \leq \exp(\sigma a).$$

On the other hand, if $T_a < \infty$, $\lim_{n\to\infty} M_{T_a\wedge n} = M_{T_a}$, and if $T_a = \infty$, $B_t \leq a$ for all t and so $\lim_{n\to\infty} M_{T_a\wedge n} = 0$. From the *Dominated Convergence Theorem*,

$$\mathbb{E}\left[\chi_{T_a<\infty} \exp\left(-\frac{1}{2}\sigma^2 T_a + \sigma a\right)\right] = \lim_{n\to\infty} \mathbb{E}[M_{T_a\wedge n}] = 1.$$

Taking $\sigma^2 = 2\lambda$ completes the proof. □

Warning: It was essential that there was a *dominating random variable* here. In this setting, the Dominated Convergence Theorem says that for a sequence of random variables Z_n, with $\lim_{n\to\infty} Z_n = Z$, if there is a random variable Y with $|Z_n| \leq Y$ for all n and $\mathbb{E}[Y] < \infty$, then we may deduce that $\mathbb{E}[Z] = \lim_{n\to\infty} \mathbb{E}[Z_n]$. In this example, the dominating random variable is just the *constant* $e^{\sigma a}$.

7 Stochastic integration and Itô's formula

We already saw (in Lemma 6.4) a number of random variables that are martingales with respect to the *same* probability measure and adapted to the *same* filtration. All those martingales were expressed in terms of the underlying Brownian motion. In this section, we shall see that this situation is generic: if there is a measure $\mathbb{Q}$ under which the process M_t is an $\mathcal{F}_t$-martingale, then any other $\mathbb{Q}$-martingale adapted to the same filtration $\mathcal{F}_t$ can be expressed in terms of M_t (modulo some technical assumptions). To understand this representation, we must first understand the notion of stochastic integral.

Processes that model stock prices are usually functions of one or more Brownian motions. Here, for simplicity, we restrict ourselves to functions of just one Brownian motion. The first thing that we should like to do is to write down a differential equation for the way in which the stock price evolves. The difficulty is that Brownian motion is 'too rough' for the familiar Newtonian calculus to be any help to us.

Suppose that the stock price is of the form $S_t = f(B_t)$. Formally, using Taylor's Theorem (and assuming that f at least is nice),

$$\begin{aligned} f(B_{t+\delta t}) - f(B_t) &= (B_{t+\delta t} - B_t) f'(B_t) \\ &+ \frac{1}{2!}(B_{t+\delta t} - B_t)^2 f''(B_t) + \cdots . \end{aligned} \tag{4}$$

Now in our usual derivation of the chain rule, when B_t is replaced by a Lipschitz function, the second term on the right hand side is order $O(\delta t^2)$. However, for Brownian motion, we know that $\mathbb{E}[(B_{t+\delta t} - B_t)^2]$ is δt. Consequently we cannot ignore the term involving the second derivative. Of course, now we have a problem, because we must interpret the term involving the *first* derivative. If $(B_{t+\delta t} - B_t)^2$ is $O(\delta t)$, then $(B_{t+\delta t} - B_t)$ should be $O(\sqrt{\delta t})$, which could lead to unbounded changes in y over a bounded time interval. However, things are not hopeless. The expected value of $B_{t+\delta t} - B_t$ is zero, and the fluctuations around zero are on the order of $\sqrt{\delta t}$. By comparison with the Central Limit Theorem, we see that it is possible that $S_t - S_0$ is a bounded random variable. Assuming that we can make this rigorous, the differential equation governing $S_t = f(B_t)$ will take the form

$$dS_t = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt.$$

It is convenient to write this in integrated form,

$$S_t = S_0 + \int_0^t f'(B_s)dB_s + \int_0^t \frac{1}{2}f''(B_s)ds. \tag{5}$$

We have to make rigorous mathematical sense of the *stochastic integral* (that is, the first integral) on the right hand side of this equation. The key is the following fact.

Brownian motion has finite quadratic variation.

Before proving this we define total variation and quadratic variation. For a function $f : [0,T] \to \mathbb{R}$, its variation is defined in terms of partitions.

Definition 7.1 *Let π be a partition of $[0,T]$, $N(\pi)$ the number of intervals that make up π and $\delta(\pi)$ be the* mesh *of π (that is the length of the largest interval in the partition). Write t_i, t_{i+1}, for the endpoints of a generic interval of the partition. Then the variation of f is*

$$\lim_{\delta\to 0}\left\{ \sup_{\pi:\delta(\pi)=\delta} \sum_{1}^{N(\pi)} |f(t_{j+1}) - f(t_j)| \right\}.$$

If a function is 'nice', for example differentiable, then it has bounded variation. Brownian motion has *unbounded* variation.

Definition 7.2 *The quadratic variation of a function f is defined as*

$$q.v.(f) = \lim_{\delta \to 0} \left\{ \sup_{\pi:\delta(\pi)=\delta} \sum_{1}^{N(\pi)} |f(t_{j+1}) - f(t_j))|^2 \right\}.$$

Notice that quadratic variation will be finite for functions that are much rougher than those for which the variation is bounded. Roughly speaking, finite quadratic variation will follow if the fluctuation of the function over an interval of length δ is order $\sqrt{\delta}$.

We can now be more precise about the quadratic variation of Brownian motion.

Theorem 7.3 *Let B_t denote Brownian motion and for a partition π of $[0, T]$ define*

$$S(\pi) = \sum_{j=1}^{N(\pi)} \left| B_{t_j} - B_{t_{j-1}} \right|^2.$$

Let π_n be a sequence of partitions with $\delta(\pi_n) \to 0$. Then

$$\mathbb{E}\left[|S(\pi_n) - T|^2 \right] \to 0 \qquad \text{as } n \to \infty.$$

Proof.

We expand the expression inside the expectation and make use of our knowledge of the normal distribution. So, first observe that

$$|S(\pi_n) - T|^2 = \left| \sum_{j=1}^{N(\pi_n)} \left\{ \left| B_{t_{n,j}} - B_{t_{n,j-1}} \right|^2 - (t_{n,j} - t_{n,j-1}) \right\} \right|^2.$$

Write $\delta_{n,j}$ for $\left| B_{t_{n,j}} - B_{t_{n,j-1}} \right|^2 - (t_{n,j} - t_{n,j-1})$. Then

$$|S(\pi_n) - T|^2 = \sum_{j=1}^{N(\pi_n)} \delta_{n,j}^2 + 2 \sum_{j<k} \delta_{n,j} \delta_{n,k}.$$

Note that since Brownian motion has independent increments,

$$\mathbb{E}[\delta_{n,j}\delta_{n,k}] = \mathbb{E}[\delta_{n,j}]\,\mathbb{E}[\delta_{n,k}] = 0 \qquad \text{if } j \neq k,$$

and

$$\mathbb{E}\left[\delta_{n,j}^2\right] = \mathbb{E}\left[\left| B_{t_{n,j}} - B_{t_{n,j-1}} \right|^4 - 2 \left| B_{t_{n,j}} - B_{t_{n,j-1}} \right|^2 (t_{n,j} - t_{n,j-1}) + (t_{n,j} - t_{n,j-1})^2 \right].$$

Now for a normally distributed random variable X with mean zero and variance λ, it is easy to check that $\mathbb{E}[|X|^4] = 3\lambda^2$, so that

$$\begin{aligned} \mathbb{E}\left[\delta_{n,j}^2\right] &= 3\left(t_{n,j} - t_{n,j-1}\right)^2 - 2\left(t_{n,j} - t_{n,j-1}\right)^2 + \left(t_{n,j} - t_{n,j-1}\right)^2 \\ &= 2\left(t_{n,j} - t_{n,j-1}\right)^2 \\ &\leq 2\delta(\pi_n)\left(t_{n,j} - t_{n,j-1}\right). \end{aligned}$$

Summing over j

$$\begin{aligned}\mathbb{E}\left[|S(\pi_n) - T|^2\right] &\leq 2 \sum_{j=1}^{N(\pi_n)} \delta(\pi_n)\left(t_{n,j} - t_{n,j-1}\right) \\ &= 2\delta(\pi_n)T \to 0 \qquad \text{as } n \to \infty.\end{aligned}$$

□

This result is not enough to define the integral $\int f(B_s)dB_s$ in the classical way, but it is enough to allow us to essentially mimic the construction of the (Lebesgue) integral, at least for functions for which $\mathbb{E}[f^2(B_\cdot)] \in L^1[0,T]$. However, although the construction of the integral may look familiar, its behaviour is far from familiar. We first illustrate this by defining $\int_0^T B_s dB_s$.

From classical integration theory we are used to the idea that

$$\int_0^T f(x_s)dx_s = \lim_{\delta(\pi)\to 0} \sum_{j=0}^{N(\pi)-1} f(x_{t_j})\left(x_{t_{j+1}} - x_{t_j}\right). \tag{6}$$

Let us define the stochastic integral in the same way, that is

$$\int_0^T B_s dB_s = \lim_{\delta(\pi)\to 0} \sum_{j=0}^{N(\pi)-1} B_{t_j}\left(B_{t_{j+1}} - B_{t_j}\right). \tag{7}$$

Consider again the quantity $S(\pi)$ of Theorem 7.3.

$$\begin{aligned}S(\pi) &= \sum_{j=1}^{N(\pi)} \left(B_{t_j} - B_{t_{j-1}}\right)^2 \\ &= \sum_{j=1}^{N(\pi)} \left\{\left(B_{t_j}^2 - B_{t_{j-1}}^2\right) - 2B_{t_{j-1}}\left(B_{t_j} - B_{t,j-1}\right)\right\} \\ &= B_T^2 - B_0^2 - 2\sum_{j=0}^{N(\pi)-1} B_{t_j}\left(B_{t_{j+1}} - B_{t_j}\right).\end{aligned}$$

The left hand side is T (by Theorem 7.3) and so letting $\delta(\pi) \to 0$ and rearranging we obtain

$$\int_0^T B_s dB_s = \frac{(B_T^2 - B_0^2 - T)}{2}$$

Remark. Notice that this is *not* what one would have predicted from classical integration theory. The extra term in the stochastic integral corresponds to $S(\pi)$.

In equation (6), we use $f(x_{t_j})$ to approximate the value of f on the interval (t_j, t_{j+1}), but in the classical theory we could equally have taken any other point in the interval in place of t_j and, in the limit, the result would have been the same. In the stochastic theory this is no longer the case. On the problem sheet you are asked to calculate two further limits

1. The limit as $\delta(\pi) \to 0$ of

$$\sum_{j=0}^{N(\pi)-1} B_{t_{j+1}} \left(B_{t_{j+1}} - B_{t_j} \right).$$

2.

$$\lim_{\delta(\pi)\to 0} \sum_{j=0}^{N(\pi)-1} \left(\frac{B_{t_j} + B_{t_{j+1}}}{2} \right) \left(B_{t_{j+1}} - B_{t_j} \right).$$

By choosing different points within each subinterval of the partition with which to approximate f over the subinterval we obtain *different* integrals. The *Itô integral* is defined as

$$\int_0^T f(B_s) dB_s = \lim_{\delta(\pi)\to 0} \sum_{j=1}^{N(\pi)} f(B_{t_j}) \left(B_{t_{j+1}} - B_{t_j} \right).$$

The *Stratonovich integral* is defined as

$$\int_0^T f(B_s) \circ dB_s = \lim_{\delta(\pi)\to 0} \sum_{j=1}^{N(\pi)} \left(\frac{f(B_{t_j}) + f(B_{t_{j+1}})}{2} \right) \left(B_{t_{j+1}} - B_{t_j} \right).$$

The Stratonovich integral has the advantage from the calculational point of view that the rules of Newtonian calculus hold good. From a modelling point of view, at least for our purposes, it is the wrong choice. To see why, think of what is happening over an infinitesimal time interval. We might be modelling, for example, the value of a portfolio. We readjust our portfolio at the *beginning* of the time interval and its change in value over the infinitesimal tick of the clock is beyond our control. A Stratonovich model would allow us to change our model *now* on the basis of the average of the value corresponding to current stock prices and the value corresponding to prices after the next tick. We don't have that information when we make our investment decisions.

Consider then the Itô integral. We have evaluated it in just one special case. We increase our repertoire in the same way as in the classical setting by first considering the value on simple functions.

Definition 7.4 *A* simple function *is one of the form*

$$f(B_s) = \sum_{i=1}^{n} a_i(B_s) \chi_{I_i}(s),$$

where $I_i = [s_i, s_{i+1})$, $\cup_{i=1}^n I_i = [0, T)$, $I_i \cap I_j = \{\emptyset\}$ *if* $i \neq j$ *and the functions* a_i *satisfy* $\mathbb{E}[a_i(B_s)^2] < \infty$.

By our definition,

$$\int_0^T f(B_s) dB_s = \sum_{i=1}^{n} a_i(B_s) \left(B_{s_{i+1}} - B_s \right).$$

Now, just as for regular integration, we approximate more general functions by simple functions and pass to a limit. We have to be sure, however, that the integrals converge when we pass to such a limit. This will not be true for all functions that can be approximated by simple functions. The next Lemma helps identify the space of functions for which we can reasonably expect a nice limit.

Lemma 7.5 *Suppose that f is a simple function, then*

1. $\int_0^t f_s(B_s)dB_s$ *is a continuous $\mathcal{F}_t$-martingale.*
2. $$\mathbb{E}\left[\left(\int_0^T f(B_s)dB_s\right)^2\right] = \int_0^T \mathbb{E}\left[f(B_s)^2\right] ds.$$
3. $$\mathbb{E}\left[\sup_{t\leq T}\left(\int_0^T f(B_s)dB_s\right)^2\right] \leq 4\int_0^T \mathbb{E}\left[f(B_s)^2\right] ds.$$

Remark: The second assertion is the famous Itô isometry. It suggests that we should be able to extend our definition of the integral to functions such that $\int_0^t \mathbb{E}[f_s(B_s)^2]ds < \infty$. Moreover, for such functions, all three assertions should remain true. In fact one can extend the definition a little further, but the integral may then fail to be a martingale and this property will be important to us.

Proof.

The third assertion follows from the second by an application of a famous result of Doob:

Theorem 7.6 (Doob's inequality) *If $\{M_t\}_{0\leq t\leq T}$ is a continuous martingale, then*

$$\mathbb{E}\left[\sup_{0\leq t\leq T} M_t^2\right] \leq 4\mathbb{E}\left[M_T^2\right].$$

We omit the proof of this remarkable Theorem.

We confine ourselves to proving the second assertion of Lemma 7.5. By our definition we have

$$\int_0^T f(B_s)dB_s = \sum_{i=1}^n a_i(B_s)\left(B_{s_{i+1}} - B_s\right),$$

and so

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^T f(B_s)dB_s\right)^2\right] &= \mathbb{E}\left[\left(\sum_{i=1}^n a_i(B_{s_i})\left(B_{s_{i+1}} - B_{s_i}\right)\right)^2\right] \\
&= \mathbb{E}\left[\sum_{i=1}^n a_i^2(B_{s_i})\left(B_{s_{i+1}} - B_{s_i}\right)^2\right] \\
&\quad +2\mathbb{E}\left[\sum_{i<j} a_i(B_{s_i})a_j(B_{s_j})\left(B_{s_{i+1}} - B_{s_i}\right)\left(B_{s_{j+1}} - B_{s_j}\right)\right].
\end{aligned}$$

Suppose that $j > i$, then

$$\begin{aligned}
&\mathbb{E}\left[a_i(B_{s_i})a_j(B_{s_j})\left(B_{s_{i+1}} - B_{s_i}\right)\left(B_{s_{j+1}} - B_{s_j}\right)\right] \\
&\qquad = \mathbb{E}\left[a_i(B_{s_i})a_j(B_{s_j})\left(B_{s_{i+1}} - B_{s_i}\right)\mathbb{E}\left[\left(B_{s_{j+1}} - B_{s_j}\right)\middle|\mathcal{F}_{s_j}\right]\right] = 0
\end{aligned}$$

Moreover,

$$\begin{aligned}
\mathbb{E}\left[a_i^2(B_{s_i})\left(B_{s_{i+1}} - B_{s_i}\right)^2\right] &= \mathbb{E}\left[a_i^2(B_{s_i})\mathbb{E}\left[\left(B_{s_{i+1}} - B_{s_i}\right)^2\middle|\mathcal{F}_{s_i}\right]\right] \\
&= \mathbb{E}\left[a_i^2(B_{s_i})\right](s_{i+1} - s_i).
\end{aligned}$$

Substituting we obtain

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^T f(B_s)dB_s\right)^2\right] &= \sum_{i=1}^n \mathbb{E}\left[a_i^2(B_{s_i})\right](s_{i+1} - s_i) \\
&= \int_0^T \mathbb{E}\left[f(B_s)^2\right]ds.
\end{aligned}$$

□

Notation: We write

$$\mathcal{H} = \{f : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R} : \int_0^T \mathbb{E}\left[f_s(B_s)^2\right]ds < \infty\}.$$

Theorem 7.7 *Let $\mathcal{F}_t$ denote the natural filtration generated by Brownian motion. There exists a unique linear mapping, J, from $\mathcal{H}$ to the space of continuous $\mathcal{F}_t$-martingales defined on $[0, T]$ such that*

1. *If f is simple,*

$$J(f)_t = \int_0^t f_s(B_s)dB_s,$$

2. *If $t \leq T$,*

$$\mathbb{E}\left[J(f)_t^2\right] = \int_0^t \mathbb{E}\left[f_s(B_s)^2\right]ds,$$

3.

$$\mathbb{E}\left[\sup_{0\leq t\leq T} J(f)_t^2\right] \leq 4\int_0^T \mathbb{E}\left[f_s(B_s)^2\right]ds.$$

Sketch of proof:

The last part follows from Doob's inequality once we know that $J(f)$ is a martingale.

The second assertion follows almost as our proof of existence of Brownian motion. We first take a sequence of simple functions such that

$$\mathbb{E}\left[\int_0^t |f_s - f_s^{(n)}|^2 ds\right] \to 0 \qquad \text{as } n \to \infty.$$

One then checks that (with probability one) the *uniform* limit of $J(f^{(n)})$ exists on $[0,T]$. (This is an easy consequence of Lemma 7.5.) $\square$

We write

$$J(f)_t = \int_0^t f_s(B_s)dB_s.$$

Having made some sense of the stochastic integral, we are now in a position to try to make sense of the chain rule for stochastic calculus.

Theorem 7.8 (Itô's formula) *For f such that $\frac{\partial f}{\partial x} \in \mathcal{H}$,*

$$f(t,B_t) - f(0,B_0) = \int_0^t \frac{\partial f}{\partial x}(s,B_s)dB_s + \int_0^t \frac{\partial f}{\partial s}(s,B_s)ds + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s,B_s)ds.$$

Notation: Often one writes this in differential notation as

$$df_t = f_t' dB_t + \dot{f}_t dt + \frac{1}{2} f_t'' dt.$$

Outline of proof:

To simplify notation, suppose that $\dot{f}_t \equiv 0$. The formula then becomes

$$f(B_t) - f(B_0) = \int_0^t \frac{\partial f}{\partial x}(B_s)dB_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(B_s)ds.$$

Let π be a partition of $[0,t]$ with mesh δ. Then

$$f(B_t) - f(B_0) = \sum_0^{N(\pi)-1} \left(f(B_{t_{j+1}}) - f(B_{t_j})\right).$$

We apply Taylor's Theorem on each interval of the partition.

$$\begin{aligned} f(B_t) - f(B_0) = \sum_0^{N(\pi)-1} f'(B_{t_j})\left(B_{t_{j+1}} - B_{t_j}\right) &+ \frac{1}{2}\sum_0^{N(\pi)-1} f''(B_{t_j})\left(B_{t_{j+1}} - B_{t_j}\right)^2 \\ &+ \frac{1}{3!}\sum_0^{N(\pi)-1} f'''(\xi_j)\left(B_{t_{j+1}} - B_{t_j}\right)^3 \end{aligned}$$

for some points $\xi_j \in [t_j, t_{j+1}]$. If f''' is uniformly bounded, then the third term tends to zero in probability (Exercise). The first converges to the Itô integral and the second, by Theorem 7.3, converges to $1/2\int_0^t f''(B_s)ds$. $\square$

Example 7.9 *Use Itô's formula to compute $\mathbb{E}[B_t^4]$.*

Let us define $Z_t = B_t^4$. Then by Itô's formula

$$dZ_t = 4B_t^3 dB_t + 6B_t^2 dt,$$

and, of course, $Z_0 = 0$. In integrated form,

$$Z_t - Z_0 = \int_0^t 4B_s^3 dB_s + \int_0^t 6B_s^2 ds.$$

Taking expectations, the expectation of the stochastic integral vanishes (by the martingale property) and so

$$\mathbb{E}[Z_t] = \int_0^t 6\mathbb{E}[B_s^2]ds = \int_0^t 6sds = 3t^2.$$

□

The most common model of stock price movements is given by *geometric Brownian motion*, defined by

$$S_t = \exp\left(\nu t + \sigma B_t\right).$$

Applying Itô's formula,

$$\begin{cases} dS_t = \sigma S_t dB_t + \left(\nu + \frac{1}{2}\sigma^2\right) S_t dt \\ S_0 = 1. \end{cases}$$

This expression is called the *stochastic differential equation* for S_t. It is common to write such symbolic equations even though it is the *integral* equation that makes sense.

Writing $\mu = \nu + \sigma^2/2$, geometric Brownian motion is a martingale if and only if $\mu = 0$ and $\mathbb{E}[S_t] = \exp(\mu t)$.

It is convenient to have a version of Itô's formula that allows us to work directly with S_t (that is to write down a stochastic differential equation for $f(S_t)$ for example). We now know how to make our original heuristic calculations rigorous, so with a clear conscience we proceed as follows:

$$\begin{aligned} f(S_{t+\delta t}) - f(S_t) &\approx f'(S_t)\left(S_{t+\delta t} - S_t\right) + \frac{1}{2}f''(S_t)\left(S_{t+\delta t} - S_t\right)^2 \\ &\approx f'(S_t)dS_t + \frac{1}{2}f''(S_t)\left\{\sigma^2 S_t^2 dB_t^2 + \mu^2 S_t^2 dt^2 + 2\sigma\mu S_t^2 dB_t dt\right\} \\ &= f'(S_t)dS_t + \frac{1}{2}f''(S_t)\sigma^2 S_t^2 dt. \end{aligned}$$

(We have used that $\delta t(B_{t+\delta t} - B_t) = o(\delta t)$ which is usually written symbolically as $dtdB_t = 0$.) As before, allowing f to also depend on t introduces an extra term $\dot{f}(S_t)dt$. Writing this version of Itô's formula in integrated form gives then:

$$\begin{aligned} f(t, S_t) - f(0, S_0) &= \int_0^t \frac{\partial f}{\partial x}(u, S_u)dS_u + \int_0^t \frac{\partial f}{\partial u}(u, S_u)du + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(u, S_u)\sigma^2 S_u^2 du \\ &= \int_0^t \frac{\partial f}{\partial x}(u, S_u)\sigma S_u dB_u + \int_0^t \frac{\partial f}{\partial x}(u, S_u)\mu S_u du \\ &\quad + \int_0^t \frac{\partial f}{\partial u}(u, S_u)du + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(u, S_u)\sigma^2 S_u^2 du \end{aligned}$$

Warning: Be aware that the stochastic integral with respect to S will not be a martingale with respect to the probability under which B_t is a martingale except in the special case when $\mu = 0$. To actually *calculate* it is often wise to separate the martingale part by expanding the 'stochastic' integral as in the last line.

It is left to the reader to justify the following more general version of Itô's formula.

Theorem 7.10 *If Y_t satisfies*

$$dY_t = a(Y_t)dB_t + b(Y_t)dt,$$

and

$$Z_t = f(t, Y_t),$$

then

$$dZ_t = f'(t, Y_t)dY_t + \dot{f}(t, Y_t)dt + \frac{1}{2}f''(t, Y_t)a(Y_t)^2 dt.$$

Remark: Notice that

$$M_t = Y_t - Y_0 - \int_0^t b(Y_s)ds$$

is a martingale with mean zero. From the Itô isometry, we know that the variance is

$$\mathbb{E}[M_t^2] = \mathbb{E}\left[\int_0^t a(Y_s)^2 ds\right].$$

The expression $\int_0^t a(Y_s)^2 ds$ is the *quadratic variation* of M_t, often denoted $\langle M\rangle_t$ or $[M]_t$.

Suppose now that we have two stochastic differential equations,

$$dY_t = a(Y_t)dB_t + b(Y_t)dt,$$

$$dZ_t = \tilde{a}(Z_t)dB_t + \tilde{b}(Z_t)dt.$$

Write

$$M_t^Y = \int_0^t a(Y_s)dB_s$$

and

$$M_t^Z = \int_0^t \tilde{a}(Z_s)dB_s.$$

Then the covariance is given by

$$\begin{aligned}\mathbb{E}\left[M_t^Y M_t^Z\right] &= \frac{1}{4}\mathbb{E}\left[\left(M_t^Y + M_t^Z\right)^2 - \left(M_t^Y - M_t^Z\right)^2\right] \\ &= \mathbb{E}\left[\int_0^t a(Y_s)\tilde{a}(Z_s)ds\right]\end{aligned}$$

The quantity $\int_0^t a(Y_s)\tilde{a}(Z_s)ds$, often denoted $\langle M^Y, M^Z\rangle_t$ or $[M^Y M^Z]_t$, is called the *covariation* of M^Y and M^Z.
Notice that

$$M_t^2 - \langle M\rangle_t$$

is an $\mathcal{F}_t$-martingale and so is

$$M_t^Y M_t^Z - \langle M^Y, M^Z\rangle_t.$$

In this notation we have

Theorem 7.11 (Integration by parts) *Let $X_t = Y_t Z_t$ with Y, Z as above, then*

$$dX_t = Y_t dZ_t + Z_t dY_t + d\langle M^Y, M^Z\rangle_t.$$

Proof: We apply the Itô formula to $(Y_t + Z_t)^2$ and Y_t^2 and Z_t^2, and subtract the second two from the first to obtain

$$Y_t Z_t - Y_0 Z_0 = \int_0^t Y_s dZ_s + \int_0^t Z_s dY_s + \int_0^t a(Y_s)\tilde{a}(Z_s)ds,$$

which is the integrated form of the result. □

We now have a very large number of continuous time martingales in our hands. For any reasonable function f,

$$M_t = \int f(s, B_s) dB_s$$

is a martingale with respect to the Brownian probability and adapted to the Brownian filtration. It is natural to ask if there are any others. The answer is provided by the martingale representation theorem which says, essentially, no.

Theorem 7.12 (Brownian martingale representation theorem) *Let $\{\mathcal{F}_t\}_{0\le t\le T}$ denote the natural filtration of Brownian motion. Let $\{M_t\}_{0\le t\le T}$ be a square-integrable $\mathcal{F}_t$-martingale. Then there exists an $\mathcal{F}_t$-adapted process ϕ_s such that with probability one,*

$$M_t = M_0 + \int_0^t \phi_s dB_s.$$

The process ϕ_s is essentially unique which leads, with a little work, to

Theorem 7.13 (Lévy's characterisation of Brownian motion) *If M_t is a continous (local) martingale with quadratic variation $\langle M\rangle_t = t$ (with probability one), then M_t is a standard Brownian motion.*

We'll use this to 'prove' one more important result. Recall that in the discrete setting we were able to reduce pricing options to calculating expectations once we had found a probability measure under which the discounted stock price was a martingale. The same will be true in the continuous world, but it will no longer be possible to find the martingale measure by linear algebra. The key now will be Girsanov's Theorem.

Theorem 7.14 (Girsanov's Theorem) *Suppose that B_t is a $\mathbb{P}$-Brownian motion with the natural filtration $\mathcal{F}_t$. Suppose that θ_t is an $\mathcal{F}_t$-adapted process such that*

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \theta_t^2 dt\right)\right] < \infty.$$

Define

$$L_t = \exp\left(-\int_0^t \theta_s dB_s - \frac{1}{2}\int_0^t \theta_s^2 ds\right)$$

and let $\mathbb{P}^{(L)}$ *be the probability measure defined by*

$$\mathbb{P}^{(L)}(A) = \int_A L_t(\omega)\mathbb{P}(d\omega).$$

Then under the probability measure $\mathbb{P}^{(L)}$*, the process* $\{W_t\}_{0\leq t\leq T}$*, defined by*

$$W_t = B_t + \int_0^t \theta_s ds,$$

is a standard Brownian motion.

Notation: We write

$$\left.\frac{d\mathbb{P}^{(L)}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = L_t.$$

(L_t is the *Radon-Nikodym derivative* of $\mathbb{P}^{(L)}$ with respect to $\mathbb{P}$.)

Remarks:

1. The condition

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \theta_t^2 dt\right)\right] < \infty$$

   is enough to guarantee that L_t is a martingale. It is clearly positive and has expectation one so that $\mathbb{P}^{(L)}$ really does define a probability measure.

2. Just as in the discrete world, two probability measures are equivalent if they have the same sets of probability zero. Evidently $\mathbb{P}$ and $\mathbb{P}^{(L)}$ are equivalent.

3. If we wish to calculate an expectation with respect to $\mathbb{P}^{(L)}$ we have

$$\mathbb{E}^{(L)}\left[\phi_t\right] = \mathbb{E}\left[\phi_t L_t\right].$$

   This will be fundamental in option pricing.

Outline of proof:

We have already said that L_t is a martingale. We don't prove this in full, but we find supporting evidence by finding the stochastic differential equation satisfied by L_t. We do this in two stages. First, define

$$Z_t = -\int_0^t \theta_s dB_s - \frac{1}{2}\int_0^t \theta_s^2 ds.$$

Then

$$dZ_t = -\theta_t dB_t - \frac{1}{2}\theta_t^2 dt.$$

Now we use Theorem 7.9 applied to $L_t = \exp(Z_t)$.

$$\begin{aligned} dL_t &= \exp(Z_t)dZ_t + \frac{1}{2}\exp(Z_t)\theta_t^2 dt \\ &= -\theta_t \exp(Z_t)dB_t = -\theta_t L_t dB_t. \end{aligned}$$

Now we integrate by parts (using Theorem 7.10) to find the stochastic differential equation for $W_t L_t$. Since

$$dW_t = dB_t + \theta_t dt,$$

$$\begin{aligned} d(W_t L_t) &= W_t dL_t + L_t dW_t + d\langle M^W, M^L\rangle_t \\ &= W_t dL_t + L_t dB_t + L_t \theta_t dt - \theta_t L_t dt \\ &= (L_t - \theta_t L_t W_t)\, dB_t. \end{aligned}$$

Granted enough boundedness (which is guaranteed by our assumptions), $W_t L_t$ is then a *martingale* and has expectation zero. Thus, under the measure $\mathbb{P}^{(L)}$, W_t is a martingale.

Now the quadratic variation of W_t is the same as that of B_t, and we proved in Theorem 7.3 that with $\mathbb{P}$-probability one, the quadratic variation of B_t is just t. Now $\mathbb{P}$ and $\mathbb{P}^{(L)}$ are equivalent and so have the same sets of probability one. Therefore W_t also has quadratic variation t with $\mathbb{P}^{(L)}$-probability one. Finally, by Lévy's characterisation of Brownian motion (Theorem 7.12) we have that W_t is a $\mathbb{P}^{(L)}$-Brownian motion as required. □

We now try this in practice

Example 7.15 *Let X_t be the drifting Brownian motion process*

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

where B_t is a $\mathbb{P}$-Brownian motion and σ and μ are constants. Then taking $\theta = \mu/\sigma$, under $\mathbb{P}^{(L)}$ of Theorem 7.13 we have that $W_t = B_t + \mu t/\sigma$ is a Brownian motion, and $X_t = \sigma W_t$ is then a scaled Brownian motion.

Notice that, for example,

$$\mathbb{E}_{\mathbb{P}}\left[X_t^2\right] = \mathbb{E}_{\mathbb{P}}\left[\sigma^2 B_t^2 + 2\sigma\mu t B_t + \mu^2 t^2\right] = \sigma^2 t + \mu^2 t^2,$$

whereas

$$\mathbb{E}_{\mathbb{P}^{(L)}}\left[X_t^2\right] = \mathbb{E}_{\mathbb{P}^{(L)}}\left[\sigma^2 W_t^2\right] = \sigma^2 t.$$

We are finally in a position to describe the Black-Scholes model for option pricing.