Chapter 7

Stochastic calculus

This chapter is split into two parts. First we will give meaning to the stochastic integral for a diffusion process

$$X_t = X_0 + \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s, \tag{7.1}$$

and also for the jump diffusion

$$X_t = X_0 + \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s + J_t.$$

The purpose of the other section is to provide analytic solution methods for some classes of stochastic differential equations (SDEs). Given an implicit definition for X_t, something like

$$dX_t = a(t, X_t) \, dt + b(t, X_t) \, dW_t,$$

our aim is to back out an expression for X_t of the form (7.1).

There are only a few SDEs which can be solved analytically. The others must be approximated numerically. A part of the course MN480 *Computational Techniques in Financial Mathematics* is devoted to the numerical solution to stochastic differential equations.

We recall that there are three primary components of Newtonian calculus:

- differentiation;
- integration; and
- the fundamental theorem.

Stochastic integration is defined as the limit of approximating sums, in the same way as classical integration. Unlike ordinary calculus, however, the derivative of a diffusion process cannot be computed as a limiting, instantaneous, rate of change. Indeed, the sample paths of Brownian motion are nowhere differentiable in this sense. In light of this fact, the role of the fundamental theorem changes. Instead of stating the relationship between integration and differentiation as it does in real analysis, the stochastic analysis version of the fundamental "theorem" actually defines what is meant by a stochastic differential; it might be better named the "fundamental stochastic definition".

Stochastic calculus has an additional component over ordinary calculus; Girsanov's Theorem. It allows the underlying probability measure of a Brownian driven stochastic process to be changed such that the process becomes a martingale. This result has some very useful applications, especially relating to finance and risk-neutral pricing.

Our approach follows that of Klebaner (1998). There are many other texts which also provide understandable introductions; search for any title which includes "stochastic calculus", "stochastic differential equations", or "stochastic integration". Reputable authors include Chung & Williams (1990), Durrett (1996), Karatzas & Shreve (1991), and Mikosch (1998). Technical treatments can be found in Jacod & Shiryaev (1987), Protter (1992), and Revuz & Yor (1990). And there must be something special in a book which is revised and republished five times by a top quality publisher; as has Øksendal (1998).

7.1 Stochastic integration

In Section 5.3 we talked of moulding Brownian motion to the shape of the process we were trying to model. Equivalent to the notion of moulding Brownian motion,

the modelling process may be regarded as adding random noise to a suitably chosen deterministic function. Here the deterministic function is chosen to reflect the underlying trends of the process of interest. Consider then the integral equation

$$X_t = x_0 + \int_0^t a(s, X_s)ds,$$

which defines X_t to be a process starting at x_0 and following a deterministic path. Its value at any time $t \geq 0$ is not random. Differentiating both sides of this equation yields the differential equation

$$\frac{dX_t}{dt} = a(t, X_t), \quad X_0 = x_0,$$

which, with a slight abuse of notation, can be more conveniently written

$$dX_t = a(t, X_t)dt, \quad X_0 = x_0. \tag{7.2}$$

Equation (7.2) specifies a first order ordinary differential equation with initial condition $X_0 = x_0$. The solution of which, if it exists, is a smooth function of t which exhibits no randomness at all.

Randomness is introduced via an additional random noise term:

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t. \tag{7.3}$$

Here, as usual, $W = (W_t, t \geq 0)$ denotes Brownian motion, and $a(t, x)$ and $b(t, x)$ are arbitrary functions. Provided it exists, the solution X of (7.3) is a stochastic process. The process starts at $X_0 = x_0$ and follows a path with trajectory based on $a(t, X_t)dt$ which has been perturbed by random noise. Equation (7.3) is a *stochastic differential equation.*

In addition to the diffusion process, we may also require a process to involve jumps. If, for example, we were modelling a process which has a tendency to change violently when unexpected news is broken. Sometimes such a process can be reproduced by incorporating a Poissonian jump term J_t into Equation (7.3),

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t + dJ_t.$$

The curious reader may ask whether (7.3) is the most general form for a stochastic process driven by Brownian motion to be written in. *Are there functions $X_t = g(W_t)$ that our modelling desires, but which don't fit the form of* (7.3)*?* This question is a current topic of debate. Alternative forms have been discussed including the McShane SDE (Blenman, Cantrell, Fennell, Parker, Wang & Womer 1995)

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t + \frac{1}{2}b_x(t, X_t)b(t, X_t)(dW_t)^2.$$

What follows is an argument suggesting the form (7.3) is sufficiently general. After which, we shall always be cautious that there may be room for further refinement.

We will be dealing extensively with discretisation of intervals in $\mathbb{R}$. Consider a *partition* of the interval $[0, T]$:

$$\tau_n : 0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T,$$

and define $\delta_i = \delta_i^{(n)} = t_i - t_{i-1}, \quad i = 1, \ldots, n$. The *mesh* of τ_n is defined as $\max_i \delta_i$. An *intermediate partition* γ_n of τ_n is given by any values y_i satisfying $t_{i-1} \leq y_i \leq t_i$ for $i = 1, \ldots, n$. We shall sometimes use ΔX_i to denote $(X_{t_i} - X_{t_{i-1}})$.

Suppose that the stock price is of the form $S_t = g(W_t)$. Formally, using Taylor's Theorem and assuming g is twice differentiable,

$$\begin{aligned} g(W_{t_{i+1}}) - g(W_{t_i}) &= (W_{t_{i+1}} - W_{t_i})g'(W_{t_i}) + \frac{1}{2!}(W_{t_{i+1}} - W_{t_i})^2 g''(W_{t_i}) \\ &\quad + \text{errors of order } (W_{t_{i+1}} - W_{t_i})^3,\ {\delta_{i+1}}^2, \text{ and } (W_{t_{i+1}} - W_{t_i})\delta_{i+1}. \end{aligned} \tag{7.4}$$

Now we know from the definition of Brownian motion that

$$\mathbb{E}\left[(W_{t+\delta t} - W_t)^2\right] = \delta t,$$

consequently we cannot ignore the second term in (7.4). Equation (7.4) can be written

$$g(W_{t_{i+1}}) - g(W_{t_i}) = (W_{t_{i+1}} - W_{t_i})g'(W_{t_i}) + \frac{1}{2!}(W_{t_{i+1}} - W_{t_i})^2 g''(W_{t_i}) + o(mesh(\tau_n)). \tag{7.5}$$

The next step is to replace $(W_{t_{i+1}} - W_{t_i})^2$ by dt in the above expression. This substitution is the source of much angst and is at the centre of the debate. The justification is encapsulated by the proof of the following result, which may be found in entirety in Klebaner (1998).

Lemma 8. *If g is a bounded continuous function, $\tau = (\tau_n; n \geq 1)$ is a sequence of partitions of $[0,t]$ with limiting mesh 0, and $\theta_i \in [W_{t_{i-1}}, W_{t_i}]$. Then*

$$\sum_{i=1}^{n} g(\theta_i)(W_{t_i} - W_{t_{i-1}})^2,$$

converges in probability to $\int_0^t g(W_s)ds$.

Proof. Any function of a continuous function is also continuous, therefore by the continuity of $g(W_t)$

$$\sum_{i=1}^{n} g(\theta_i)(t_i - t_{i-1}) \to \int_0^t g(W_s)ds.$$

Aiming to show

$$\sum_{i=1}^{n} g(\theta_i)(W_{t_i} - W_{t_{i-1}})^2 \to \int_0^t g(W_s)ds,$$

the crucial step in the proof is showing that

$$\sum_{i=1}^{n} g(W_{t_{i-1}})(W_{t_i} - W_{t_{i-1}})^2 - \sum_{i=1}^{n} g(W_{t_{i-1}})(t_i - t_{i-1}) \to 0,$$

in mean square. Put another way, we need to show

$$\mathbb{E}\left[\sum_{i=1}^{n} g(W_{t_{i-1}})\left((\Delta W_i)^2 - \delta_i\right)\right]^2 \to 0,$$

which is realised by noting that $\mathbb{E}\left((\Delta W_i)^2 - \delta_i\right)^2 \leq 2\delta_i^2$, and that $2\mathbb{E}\sum_{i=1}^{n} g^2(W_{t_{i-1}})\delta_i^2$ is $o(mesh(\tau_n))$.

Having shown that both sums $\sum_{i=1}^{n} g(\theta_i)(W_{t_i} - W_{t_{i-1}})^2$ and $\sum_{i=1}^{n} g(\theta_i)(t_i - t_{i-1})$ have the same limit in the case $\theta_i = W_{t_{i-1}}$, and that limit is $\int_0^t g(W_s)ds$. Using the continuity of g gives the generalisation from $\theta_i = W_{t_{i-1}}$ to an arbitrary θ chosen such that $\theta_i \in [W_{t_{i-1}}, W_{t_i}]$. $\square$

Taking the limit as $mesh(\tau_n) \to 0$ in (7.5) with dt substituted for $(W_{t+\delta t} - W_t)^2$ leads to

$$dS_t = g'(W_t)dW_t + \frac{1}{2}g''(W_t)dt.$$

The differential equation governing $S_t = g(W_t)$ will take the form

$$dS_t = a(t, S_t)dt + b(t, S_t)dW_t, \quad S_0 = s_0. \tag{7.6}$$

A naive interpretation of (7.6) tells us that the change $dS_t = S_{t+dt} - S_t$ is caused by a change dt of time, with factor $a(t, S_t)$, in combination with a change $dW_t = W_{t+dt} - W_t$ of Brownian motion, with factor $b(t, S_t)$. One of the strange properties of Brownian motion is that the sample paths are nowhere differentiable (w.p. 1). How then do we interpret dW_t? There is no easy (or even unique!) answer to this question. As a starting point we might consider the integral equation

$$S_t = s_0 + \int_0^t a(r, S_r)dr + \int_0^t b(r, S_r)dW_r, \tag{7.7}$$

which has equivalent meaning to Equation (7.6). The remainder of this section is largely devoted to finding a meaning for $\int_0^t b(r, S_r)dW_r$.

7.1.1 The Riemann integral

Everyone is familiar with the classical Riemann Integral defined as the limit of a Riemann sum. Suppose we have a real valued function defined on $[0, T]$. Given sequences of partitions τ and γ we can define the *Riemann sum*

$$S_n = S_n(\tau_n, \gamma_n) = \sum_{i=1}^{n} f(y_i)(t_i - t_{i-1}) = \sum_{i=1}^{n} f(y_i)\delta_i.$$

Essentially S_n is an approximation to the area between the curve f and the t axis. Choosing a sequence of partitions which have limiting mesh 0 we define

$$\int_0^T f(y)dy = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{i=1}^{n} f(y_i)\delta_i. \tag{7.8}$$

The integral in (7.8) is only defined if it is independent of the chosen sequences $\{\tau_n\}$ and $\{\gamma_n\}$. In this case we have the *Riemann integral* of f on $[0, T]$.

The Riemann integral will work fine for the deterministic part $\int_0^t a(r, S_r)dr$ of Equation (7.7), but we are still stuck for a definition for $\int_0^t b(r, S_r)dW_r$.

7.1.2 Lebesgue integration

The Riemann integral is actually a special case of the Lebesgue-Stieltjes integral,

$$\int_0^T f(y)dg(y).$$

This integral is constructed in the same fashion as the Riemann integral, only this time we are integrating f with respect to another function g. For this reason it promises to get us closer to the target of integrating with respect to a random process.

Let f and g be two real-valued functions on $[0, T]$ and define

$$\delta_i g = g(t_i) - g(t_{i-1}), \quad i = 1, \ldots, n.$$

The *Riemann-Stieltjes* sum is obtained by weighting the values $f(y_i)$ with the increments $\delta_i g$. If the limit $\lim_{n\to\infty} \sum_{i=1}^n f(y_i)\delta_i g$, exists as the mesh tends to zero, and is independent of the choice of the partitions $\{\tau_n\}$ and $\{\gamma_n\}$ then

$$\int_0^T f(y)dg(y) = \lim_{n\to\infty} \sum_{i=1}^n f(y_i)\delta_i g,$$

is called the *Lebesgue integral* of f with respect to g.

So, it appears that the integral $\int_0^t b(r, S_r)dW_r$ might fit this framework, provided b is nice enough that the integral exists. To quantify what we mean by "nice enough" we need the following definition.

Definition 1 (p-variation). *The function h is said to have* bounded p-variation *for some $p > 0$ if*

$$\sup_\tau \sum_{i=1}^n |h(t_i) - h(t_{i-1})|^p < \infty,$$

where the supremum is taken over all partitions τ of $[0,T]$. Special cases include $p=1$ corresponding to plain variation. Setting $p=2$ gives the definition of quadratic variation,

$$\langle h\rangle_T = \sup_\tau \sum_{i=1}^n |h(t_i)-h(t_{i-1})|^2,$$

Sufficient conditions for Lebesgue integrability of f with respect to g are

- the functions f and g do not have discontinuities at the same point $t\in[0,T]$;
- the function f has bounded p-variation and the function g has bounded q-variation for some $p>0$ and $q>0$ such that $p^{-1}+q^{-1}>1$.

Now Brownian motion has bounded p-variation on every fixed interval provided $p\geq 2$, and unbounded for $p<2$. Consider a deterministic function $f:[0,T]\mapsto\mathbb{R}$. According to the above theory, we can define the Lebesgue integral

$$\int_0^T f(t)dW_t,$$

provided f has bounded q-variation for some $q<2$. This is certainly satisfied if f has bounded variation; that is if $q=1$. For functions f which are differentiable with $f'(t)<\infty$, on $[0,T]$ it follows from the mean value theorem that

$$\exists K>0 \text{ such that } |f(t)-f(s)|\leq K(t-s), \quad \forall s<t,$$

then

$$\sup_\tau \sum_{i=1}^N |f(t_i)-f(t_{i-1})| \leq K\sum_{i=1}^N (t_i-t_{i-1}) = K\,T<\infty.$$

Hence f has bounded variation, and so $\int_0^T f(t)dW_t$ exists. In particular, within the Riemann-Stieltjes framework we can unambiguously define

$$\int_0^T e^t\,dW_t, \quad \int_0^T \sin(t)\,dW_t, \quad \int_0^T t^n\,dW_t, n\geq 0.$$

We need not look too hard to find an integrand which is not Lebesgue integrable though; consider the integral of Brownian motion with respect to itself

$$\int_0^T W_t \, dW_t.$$

Brownian motion has bounded p-variation for $p \geq 2$, not for $p < 2$, and so the sufficient condition $2p^{-1} > 1$ for existence of this integral is not satisfied. Moreover it can be shown that the value obtained by evaluating the limit of the Riemann-Stieltjes sum

$$\lim_{mesh(\tau_n) \to 0} \sum_{i=0}^{n-1} W_{y_i}(W_{t_{i+1}} - W_{t_i}),$$

is dependent on the particular sequence of intermediate partitions $(y_i : i = 0, \ldots, n-1)$ used. Therefore $\int_0^T W_t \, dW_t$ is not defined under this framework.
Incidentally,

Lemma 9. *the quadratic variation of Brownian motion over the interval* $[0, t]$ *is equal to* t.

Proof. Let $T_n = \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2$. For each n we have

$$\begin{aligned} \mathbb{E}(T_n) &= \sum_{i=0}^{n-1} \mathbb{E}(W_{t_{i+1}} - W_{t_i})^2, \\ &= \sum_{i=0}^{n-1} (t_{i+1} - t_i), \\ &= t, \end{aligned}$$

where the first equality holds because there are a finite number of terms in the sum, the second is because Brownian motion is a Gaussian process.

Using the fourth moment of a standard normal random variable, we obtain

$$\begin{aligned} Var(T_n) &= \sum_{i=0}^{n-1} Var(W_{t_{i+1}} - W_{t_i})^2, \\ &\leq 3t \, mesh(\tau_n). \end{aligned}$$

Now choose a sequence of partitions $\tau = (\tau_n; n \geq 1)$, such that $\sum_{n\geq 1} mesh(\tau_n) < \infty$. Such a sequence is obtained (for example) if $mesh(\tau_{n+1}) = \frac{1}{2} mesh(\tau_n)$ as it would if the successive refinements divided each interval in two. With this choice

$$\sum_{n\geq 1} Var(T_n) \leq 3t \sum_{n\geq 1} \frac{1}{2^n} < \infty.$$

Applying the Monotone Convergence Theorem to the sequence $S_k = \sum_{n=1}^{k} (T_n - \mathbb{E}(T_n))^2$, and using Fatou's Lemma

$$\mathbb{E}\left(\sum_{n\geq 1} (T_n - \mathbb{E}(T_n))^2\right) = \mathbb{E}(\liminf S_k) \leq \liminf \mathbb{E}(S_k) = \sum_{n\geq 1} Var(T_n) < \infty.$$

This implies $\sum_{n\geq 1} (T_n - \mathbb{E}(T_n))^2$ converges almost surely, which, in turn, leads to the conclusion $T_n \to \mathbb{E}(T_n)$ a.s., thus

$$\langle W \rangle_t = \lim_{mesh(\tau_n)\to 0} T_n \to t, \quad a.s.$$

□

This result is remarkable because, although the total of the sum

$$\sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2,$$

is a random variable for each n, the limiting value is non-random. The quadratic variation of Brownian motion is very important in the work to follow.

7.1.3 Itô calculus

We would like a framework in which all integrals, $\int_0^T f(t)dg(t)$, are unambiguously defined. The Lebesgue-Stieltjes integral does not exist for a large enough class of functions. The problem being that the Lebesgue integral only exists if the limit of the Riemann-Stieltjes sums are independent of the sequence of partitions chosen. We shall therefore aim to mimic the construction of the Lebesgue integral but drop the "independence of partitions" condition. So long as we consistently choose the same

partitions, we should be able to develop a calculus with rules analogous to those of standard integration; integration-by-parts, the chain rule, et cetera.

Consider a sequence of partitions $\tau = (\tau_n; n \geq 1)$ of $[0, T]$ with limiting mesh 0, and intermediate partitions $\gamma = (\gamma_n; n \geq 1)$ of τ. Define

$$\int_0^T f(y)dg(y) = \lim_{mesh(\tau_n)\to 0} \sum_{j=0}^{n-1} f(y_j)\left(g(t_{j+1}) - g(t_j)\right).$$

All that remains is to specify a choice of y_i, for $i = 0, \ldots, n-1$, and to be consistent with that choice throughout. The following two choices have turned out to be the most useful ones:

- $y_i = t_i$ (left end point) leading to the *Itô integral*, denoted by

$$\int_S^T f(t,\omega)\, dW_t(\omega),$$

  and

- $y_i = \frac{t_{i+1}+t_i}{2}$ (the midpoint) corresponding to the *Stratonovich integral*, denoted by

$$\int_S^T f(t,\omega) \circ dW_t(\omega).$$

Remark. *It follows from the discussion in the previous section that if f is deterministic and differentiable with $f'(t) < \infty$ on $[0, T]$ it doesn't matter whether we use Itô , or Stratonovich to integrate, as the result will agree with Lebesgue.*

Example 15 (The Itô integral of W_t w.r.t. itself). *We are aiming to evaluate*

$$\int_0^T W_s\, dW_s = \lim_{mesh(\tau_n)\to 0} \sum_{j=0}^{n-1} W_{t_j}\left(W_{t_{j+1}} - W_{t_j}\right).$$

Recall Lemma 9 where we proved that the sum $\sum_{j=0}^{n-1}(W_{t_{j+1}} - W_{t_j})^2$ *can be replaced by* T *in the limit. Now*

$$\begin{aligned}\sum_{j=0}^{n-1}(W_{t_{j+1}} - W_{t_j})^2 &= \sum_{j=0}^{n-1}\left[(W_{t_{j+1}}^2 - W_{t_j}^2) - 2W_{t_j}(W_{t_{j+1}} - W_{t_j})\right], \\ &= W_T^2 - W_0^2 - 2\sum_{j=0}^{n-1} W_{t_j}(W_{t_{j+1}} - W_{t_j}).\end{aligned}$$

Taking limits $mesh(\tau_n) \to 0$, *with* T *constant,*

$$T = W_T^2 - W_0^2 - 2\int_0^T W_s\, dW_s,$$

which when rearranged yields

$$\int_0^T W_s\, dW_s = \frac{W_T^2 - W_0^2 - T}{2},$$

with $W_0 = 0$ *this becomes*

$$\int_0^T W_s\, dW_s = \frac{W_T^2 - T}{2}.$$

The above example indicates that the classical rules of integration do not hold for Itô stochastic integration. If $w : [0, T] \mapsto \mathbb{R}$ is a deterministic differentiable function with $w(0) = 0$ then by standard calculus,

$$\int_0^T w(s)dw(s) = \frac{1}{2}w(T)^2.$$

This agrees with the result obtained using the Stratonovich integral,

$$\int_0^T W_s \circ dW_s = \frac{1}{2}W_T^2.$$

In fact, the Stratonovich stochastic integral obeys the standard calculus *as if* it were a regular integral. That is to say

$$\int_0^T g'(W_s) \circ dW_s = g(W_T) - g(W_0).$$

This statement does not mean that the Stratonovich stochastic integral is a classical (Riemann) integral. We only claim that similar rules hold.

Why bother considering any other integral type if all our established integration theory can be applied to the Stratonovich integral? From a modelling point of view it is often 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 each 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 based on the value midway along the next interval - sometime in the future. We don't have that information when we make our investment decisions.

The Itô integral, on the other hand, doesn't look into the future. It's value is determined using the values at the left-hand end points of each infinitesimal time interval. From the modelling perspective it is a more natural choice. To use Itô calculus we will be required to learn some new theory; Itô's formula, the analogue to the chain rule, for instance. Stratonovich integrals will act as important tools for evaluating Itô stochastic differential equations. The two integral types are related via the transformation formula. For now, we state some general properties of the Itô stochastic integral.

General properties of the Itô stochastic integral

Proposition 10 (Conditions on integrands). *Let C be adapted to the Brownian filtration on $[0,T]$; C_t may depend on W_s, for $s \leq t$ but not for $s > t$. Also suppose that the technical condition that $\int_0^T \mathbb{E}(C_s^2)ds < \infty$ is satisfied. These are the* conditions on integrands *for the Itô stochastic integral of C with respect to W to be well defined.*

When the conditions on integrands are satisfied the Itô stochastic integral has the properties:

1. *For any two processes C and D satisfying the conditions on integrands and any*

two constants c and d

$$\int_0^T [cC_s + dD_s]\, dW_s = c\int_0^T C_s dW_s + d\int_0^T D_s dW_s.$$

2. *For $0 \le t \le T$,*

$$\int_0^T C_s dW_s = \int_0^t C_s dW_s + \int_t^T C_s dW_s$$

3. *The Itô stochastic integral process $I_t(C) = \int_0^t C_s dW_s$, $t \in [0, T]$, has an expected value of zero. Furthermore $I_t(C)$ is a martingale with respect to the Brownian filtration $\mathcal{F}_t$.*

4. *The Itô stochastic integral satisfies the* isometry property*:*

$$\mathbb{E}\left[\left(\int_0^T C_s dW_s\right)^2\right] = \int_0^T \mathbb{E}(C_s^2)dt.$$

The first two are properties in common with the Riemann integral. Property 3 reminds us of the previous material we have covered regarding martingales and suggests that that theory will again be important. The process $I_t(C)$ will be of interest in its own right, as such we may wish to calculate its variance. Given $\mathbb{E}(I_t(C)) = 0$, the variance is $\mathbb{E}(I_t(C)^2)$. Property 4 is an identity which provides an alternate way of calculating the variance of $I_t(C)$.

Example 16. *Recall Lemma 6 where we showed $W_t^2 - t$ is an martingale with respect to the Brownian filtration. Notice that this is consistent with our statement that $I_t(C)$ is a martingale since*

$$I_t(W) = \int_0^t W_s\, dW_s = \frac{W_t^2 - t}{2}.$$

Itô's formula

Now that we've established conventions for stochastic integration in terms of Itô and Stratonovich, we need some higher level tools with which to work with. The first of

these is Itô's formula; sometimes presented as Itô's lemma. It is a rule corresponding to integration by substitution, roughly speaking:

$$\int f(g(s))ds = \int f(u)\left(\frac{dg}{ds}\right)^{-1} du.$$

Given the stochastic differential equation of some process $X = (X_t; t \geq 0)$ adapted to the Brownian filtration, and a smooth (C^2) function g, Itô's formula provides a method of extracting the stochastic differential and integral of the process Y defined as $Y_t = g(X_t)$.

For the moment, we shall limit our discussion to Itô processes. This class does not include all processes for which Itô's formula holds, but it is sufficiently broad to handle the majority of applications. An *Itô process* is any process which has the (Itô) stochastic integral equation form

$$X_t = X_0 + \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s, \tag{7.9}$$

with

$$\int_0^T |\mu_t| \, dt < \infty, \quad \int_0^T \sigma_t^2 \, dW_t < \infty.$$

The integral equation (7.9) is commonly written shorthand as

$$dX_t = \mu_t \, dt + \sigma_t \, dW_t. \tag{7.10}$$

Here, as usual, we are denoting by W Brownian motion. On its own the process $\int_0^t \sigma_s \, dW_s$ is called an *Itô integral process.*

Our presentation of Itô's formula will start by stating the result for any Itô process. This is a very useful formula. Itô's formula for functions of Brownian motion, for functions adapted to the Brownian filtration, and functions of two or more adapted processes can all be obtained from this first formula (7.11). We will then work through the above mentioned cases as examples.

We shall hold off proving Itô's formula until Section 7.1.4 where we shall provide a proof which holds for C^2 functions of any semimartingale. This general case encompasses not only Itô processes but also processes with possibly discontinuous sample paths.

Theorem 11 (Itô's formula for any Itô process). *Let X_t be an Itô process. If $g(x)$ is a twice continuously differentiable function, then the stochastic differential of $Y_t = g(X_t)$ is given by*

$$dY_t = g'(X_t)dX_t + \frac{1}{2}g''(X_t)d\langle X\rangle_t. \tag{7.11}$$

Recall that $\langle X\rangle_t$ denotes the quadratic variation of the process X up until time t. For X defined by (7.10) the quadratic variation is given by

$$\langle X\rangle_t = \int_0^t \sigma_s^2 ds.$$

By expanding the quadratic variation term we can evaluate Equation (7.11) two steps further:

$$\begin{aligned} dY_t &= g'(X_t)dX_t + \frac{1}{2}g''(X_t)\sigma_t^2 dt, \\ &= \sigma_t g'(X_t)dW_t + \left(\mu_t g'(X_t) + \frac{1}{2}\sigma_t^2 g''(X_t)\right)dt. \end{aligned}$$

The penultimate step involves replacing dX_t by $\mu_t\,dt + \sigma_t\,dW_t$ in the previous line.

Example 17 (Geometric Brownian motion). *Suppose $Y_t = \exp(\sigma W_t + \mu t)$. What stochastic differential does Y_t follow?*
Let $X_t = \sigma W_t + \mu t$ then $dX_t = \sigma\,dW_t + \mu\,dt$ and $Y_t = g(X_t)$ where $g(x) = e^x$. Applying Itô's formula

$$dY_t = \sigma g'(X_t)dW_t + \left(\mu g'(X_t) + \frac{1}{2}\sigma^2 g''(X_t)\right)dt,$$

and noticing that $g(x) = g'(x) = g''(x)$ and so $Y_t = g(X_t) = g'(X_t) = g''(X_t)$, yields

$$\begin{aligned} dY_t &= \sigma Y_t dW_t + \left(\mu Y_t + \frac{1}{2}\sigma^2 Y_t\right)dt, \\ &= Y_t\left(\sigma dW_t + (\mu + \frac{1}{2}\sigma^2)dt\right), \end{aligned}$$

Conversely, suppose we were given the stochastic differential

$$dY_t = Y_t(\sigma\,dW_t + \mu\,dt),$$

and asked to find the process which follows this equation. We know that the solution takes the form $\exp(\sigma W_t + \nu t)$ *and matching the parameters reveals* $\nu = \mu - \frac{1}{2}\sigma^2$. *Thus*

$$Y_t = \exp\left(\sigma W_t + (\mu - \frac{1}{2}\sigma^2)t\right),$$

which should be checked using Itô's formula.

There are some functions for which Itô's formula takes a particularly simple form. We shall now look at some of these cases and see how Equation (7.11) handles them. These results may also be established by inspecting the partial sums

$$g(X_t) = g(0) + \sum_{i=0}^{n-1} \left(g(X_{t_{i+1}}) - g(X_{t_i})\right),$$

and taking the limit as the mesh tends to zero.

Itô's formula for functions of Brownian motion

Corollary 12 (Itô's formula for functions of W_t). *If $g(x)$ is twice continuously differentiable, then for any t*

$$g(W_t) = g(0) + \int_0^t g'(W_s)dW_s + \frac{1}{2}\int_0^t g''(W_s)ds. \tag{7.12}$$

Proof. Let $Y_t = g(W_t)$, from (7.11)

$$dY_t = g'(W_t)dW_t + \frac{1}{2}g''(W_t)d\langle W\rangle_t.$$

Since $\langle W\rangle_t = \int_0^t ds = t$

$$dY_t = g'(W_t)dW_t + \frac{1}{2}g''(W_t)dt,$$

which in integrated form is (7.12). □

Example 18. *What is the stochastic differential for $Y_t = W_t^4$, hence compute $\mathbb{E}(W_t^4)$?*
By Itô's formula (7.12)

$$Y_t = Y_0 + \int_0^t 4W_s^3 dW_s + \int_0^t 6W_s^2 ds.$$

Taking expectations, the expectation of the stochastic integral vanishes due to the martingale property, and so

$$\mathbb{E}(Y_t) = \int_0^t 6\mathbb{E}(W_s^2)ds = \int_0^t 6s ds = 3t^2.$$

Itô's formula for vector processes

Like any great formula, Equation (7.11) works for vector processes as well. Let X_t be an n-dimensional Itô process,

$$\begin{cases} dX_t^{(1)} &= \mu_1\, dt + \sigma_{11} dW_t^{(1)} + \cdots + \sigma_{1m} dW_t^{(m)} \\ \vdots & \\ dX_t^{(n)} &= \mu_n\, dt + \sigma_{n1} dW_t^{(1)} + \cdots + \sigma_{nm} dW_t^{(m)} \end{cases}$$

Or, in matrix notation

$$dX_t = \mu\, dt + \sigma\, dW_t,$$

where

$$X_t = \begin{pmatrix} X_t^{(1)} \\ \vdots \\ X_t^{(n)} \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \quad \sigma = \begin{pmatrix} \sigma_{11}, \ldots, \sigma_{1m} \\ \vdots \\ \sigma_{n1}, \ldots, \sigma_{nm} \end{pmatrix}, \quad W_t = \begin{pmatrix} W_t^{(1)} \\ \vdots \\ W_t^{(n)} \end{pmatrix}$$

If $g = (g_1, \ldots, g_n)$ maps $\mathbb{R}^n$ into $\mathbb{R}^p$, and each component of g is twice continuously differentiable then Y, defined as $Y_t = g(X_t)$, is also an Itô process. The SDE that Y satisfies is

$$dY_t = Dg.dX_t + \frac{1}{2} D^2 g.d\langle X \rangle_t,$$

where the kth component is given by

$$dY_t^{(k)} = \sum_i \frac{\partial g_k}{\partial x_i} dX_t^{(i)} + \frac{1}{2} \sum_{i,j} \frac{\partial^2 g_k}{\partial x_i \partial x_j} d\left\langle X^{(i)}, X^{(j)} \right\rangle_t .$$

The last term involves the covariation between $X^{(i)}$ and $X^{(j)}$.

For any two processes X and Y, the *covariation* between X and Y is defined as

$$\langle X, Y \rangle_t = \lim_{mesh(\tau) \to 0} \sum_{i=0}^{N-1} (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i}).$$

Notice that the quantity $\langle Y, Y \rangle_t$ corresponds to the quadratic variation and is usually written $\langle Y \rangle_t$. If X and Y are both Itô processes with respect to the same Brownian motion,

$$dX_t = \mu_t dt + \sigma_t dW_t, \qquad dY_t = \gamma_t dt + \rho_t dW_t, \tag{7.13}$$

then the covariation of X and Y is

$$\begin{aligned} \langle X, Y \rangle_t &= \frac{1}{2} \left(\langle X + Y \rangle_t - \langle X \rangle_t - \langle Y \rangle_t \right), \\ &= \frac{1}{2} \int_0^t (\sigma_s + \rho_s)^2 - \sigma_s^2 - \rho_s^2 ds, \\ &= \int_0^t \sigma_s \rho_s ds. \end{aligned}$$

$d \langle X, Y \rangle_t$ is computed according to the conventions:

$$\begin{aligned} dX_t \, dY_t &= d \langle X, Y \rangle_t , \\ (dt)^2 = 0, \quad dt \, dW_t &= dW_t \, dt = 0, \quad d \langle W \rangle_t = dt. \end{aligned}$$

The two variable case is the simplest way to illustrate Itô's formula in higher dimensions. To justify its form consider the Taylor expansion of a function $f(x, y)$, which maps $\mathbb{R}^2 \to \mathbb{R}$. Here $p = 1$ and $n = 2$.

$$\begin{aligned} f(x, y) &= f(x_0, y_0) + \frac{\partial f}{\partial x}(x - x_0) + \frac{\partial f}{\partial y}(y - y_0) + \\ &\quad \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(x - x_0)^2 + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(y - y_0)^2 + \frac{\partial^2 f}{\partial y \partial x}(x - x_0)(y - y_0). \end{aligned}$$

This suggests

$$df(X_t, Y_t) = \frac{\partial f}{\partial x} dX_t + \frac{\partial f}{\partial y} dY_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(dY_t)^2 + \frac{\partial^2 f}{\partial y \partial x} dX_t dY_t.$$

Substituting

$$\begin{aligned} (dX_t)^2 &= d\langle X\rangle_t = \sigma_t^2 dt, \\ (dY_t)^2 &= \rho_t^2 dt, \\ \text{and} \quad dX_t dY_t &= d\langle X, Y\rangle_t = \sigma_t \rho_t dt, \end{aligned}$$

gives an informal justification of the following result.

Corollary 13 (Itô's formula for functions of two variables). *If $f(x, y)$ is a twice differentiable function of two variables, and Z_t is a process defined as $Z_t = f(X_t, Y_t)$ for any two Itô processes X_t and Y_t from* (7.13)*, then Z_t follows the stochastic differential*

$$dZ_t = \frac{\partial f}{\partial x} dX_t + \frac{\partial f}{\partial y} dY_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma_t^2 dt + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}\rho_t^2 dt + \frac{\partial^2 f}{\partial y \partial x}\sigma_t \rho_t dt. \tag{7.14}$$

In the special case where one of the parameters is deterministic Itô's formula for two variables simplifies.

Example 19. *Find the stochastic differential of $Z_t = f(t, X_t)$.*

By Itô's formula (7.14) *in Corollary 13*

$$dZ_t = \frac{\partial f}{\partial x} dX_t + \frac{\partial f}{\partial t} dt + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma_t^2 dt + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial^2 f}{\partial t \partial x} dX_t dt.$$

Both terms involving $(dt)^2$ and $dX_t dt$ are negligible, and so we are left with

$$dZ_t = \frac{\partial f}{\partial x} dX_t + \frac{\partial f}{\partial t} dt + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma_t^2 dt.$$

Integration by parts

In classical calculus there exists a formula which allows integration of functions which have a product form. For differentiable functions u and v it states that

$$\int u'(x)v(x)dx = u(x)v(x) - \int v'(x)u(x)dx. \tag{7.15}$$

The integration-by-parts formula (7.15) is an easy consequence of the product rule for differentiation

$$\frac{d}{dx}(u\,v) = u'(x)\,v(x) + v'(x)\,u(x). \tag{7.16}$$

There is an analogous result in stochastic calculus. For any two Itô processes X and Y,

$$d(X_t Y_t) = X_t\,dY_t + Y_t\,dX_t + d\,\langle X, Y\rangle_t\,, \tag{7.17}$$

which doesn't look too dissimilar to Equation (7.16). Equation (7.17) is easy to justify using Itô's formula for functions of two variables; apply (7.14) to the function $f(x, y) = xy$.

Example 20 (The stochastic differential for $X_t = t\,W_t$). *By the product rule* (7.17)

$$dX_t = t\,dW_t + W_t\,dt + d\,\langle t, W\rangle_t\,,$$

By convention $d\,\langle t, W\rangle_t = dt.dW_t = 0$, *thus*

$$dX_t = t\,dW_t + W_t\,dt.$$

As an alternative method for reaching this result we could apply Itô's formula with $X_t = g(t, W_t)$ *where* $g(t, x) = t\,x$.

Example 21. *Suppose* Y_t *has the stochastic differential equation*

$$dY_t = \frac{1}{2}Y_t\,dt + Y_t\,dW_t, \quad Y_0 = 1,$$

and let $X_t = t\,W_t$. *What is the stochastic differential for* $X_t\,Y_t$*?*

Use the product rule:

$$d(X_t\,Y_t) = X_t\,dY_t + Y_t\,dX_t + d\,\langle X, Y\rangle_t\,.$$

The first two terms may be simplified by substituting the rules we know for dX_t *and* dY_t. *For the third term recall the convention* $d\,\langle X, Y\rangle_t = dX_t\,dY_t$ *and substitute for* dX_t *and* dY_t *to get*

$$d\,\langle X, Y\rangle_t = (t\,dW_t + W_t\,dt)\left(\frac{1}{2}Y_t\,dt + Y_t\,dW_t,\right).$$

All terms involving $dt\,dW_t$ *or* $(dt)^2$ *are taken to be zero, and* $(dW_t)^2$ *can be replaced by* dt, *this leaves us with* $d\,\langle X, Y\rangle_t = t\,Y_t\,dt$.

7.1.4 Processes with jumps

Brownian motion is good for modelling small, continuously chaotic, changes in a process. The Poisson process, in contrast, is good for modelling random shocks to the system. Shocks that occur infrequently but which are large in comparison to the Brownian noise, are well modelled by a Poisson process.

For this subsection let X_t denote the aptly named *jump diffusion* process

$$X_t = X_0 + \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s + J_t. \tag{7.18}$$

The J_t term represents the inclusion of a Markov jump process which has been added to the diffusion (7.1).

The paths of J_t are step functions characterised by the sequence of jump times and corresponding jump sizes. Since neither of these things are known in advance, a probabilistic construction is the best one can do. The process at some time t is at X_t. With some probability, say $\lambda(X_t)\delta t + o(\delta t)$, in the next instant $(t, t+\delta t]$, J_t will make a jump. The probability of two or more jumps in an interval is $o(\delta t)$. Given that a jump has occurred, the random jump size is distributed according to some measure $\bar{\Pi}(X_t, \cdot)$. Between jumps J_t is constant and X_t behaves like the diffusion process (7.1).

The calculus of jump processes is widely known on a superficial level. It is common for authors to simply quote the rules without giving a formal justification. Part of the reason for this neglect is that the underlying measure theory is quite difficult, and in practitioner often only need to end rules.

We shall briefly discuss the underlying probabilistic basis, before describing the calculus. Our motivation is to achieve a level of competence which will allow us to be not so overwhelmed by the more difficult texts such as Jacod & Shiryaev (1987). This deeper insight should allow us to not only understand how the calculus rules arise, but to modify them for processes which don't satisfy the standard assumptions. You are encouraged to develop an interest and seek the details which have been skipped in the following sub-subsection, even though it will not be assessed.

The stochastic basis

The approach that we shall follow for the uses random measures (Jacod & Shiryaev 1987). The theory behind this approach is much less widely known than the rest of the stochastic calculus in this chapter. The random measure approach provides an unambiguous, tractable description of the stochastic integral defining a jump process.

Use the measurable space $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ for time, and $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ for space. Here $\mathbb{R}^+ = [0, \infty)$, and $\mathcal{B}(A)$ denotes the Borel sets of A. For our purposes, an integer valued *random measure* on $(\mathbb{R}^+ \times \mathbb{R}, \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{B}(\mathbb{R}))$ is a family of functions $N = (N(\omega; t, y))$ such that

1. $N(\{0\} \times \mathbb{R}) = 0$, almost surely;
2. $N(\{t\} \times \mathbb{R}) = 0$ or 1, depending on whether a jump occurred at time t or not;
3. and for any $A = T \times Y \in \mathcal{B}(\mathbb{R}^+) \otimes \mathcal{B}(\mathbb{R})$, $N(A)$ takes values in $\mathbb{N} = \{0, 1, 2, \dots\}$.

For a fixed set of jump types $Y \in \mathcal{B}(\mathbb{R})$, $(N(t, Y); t \geq 0)$ may be thought of as a random process which realises the number of jumps in $[0, t]$ taking X from X_s to $X_s + q(\omega; s, X_s)$ with $q(\omega; s, X_s) \in Y$. This might be the set of all jumps which cause some portfolio to gain $\$q$ million, for example. More likely, Y could be the set of jumps which cause the portfolio to change in value by more than q percent. To determine whether a jump belongs to such a set, there needs to be information about the jump diffusion X at time t. It is Markovian though, since the described pure jump process is independent of $\mathcal{F}_s$, $s < t$.

The main result which allows a tractable description of the jump process using random measures is summarised in Theorem 1.8 of Jacod & Shiryaev (1987), and also in Theorem 9.14 of Klebaner (1998). There exists a predictable measure, say N^p, such that for every measurable random process q on $\Omega \times \mathbb{R}^+ \times \mathbb{R}$

$$\int_0^t \int_{\mathbb{R}} q_s \, N(\omega; ds, dy) - \int_0^t \int_{\mathbb{R}} q_s \, N^p(\omega; ds, dy),$$

is a local martingale. N^p is called the *compensator* of N. Furthermore, there exists a predictable process A and a kernel K such that the function N^p may be written

$$N^p(\omega; dt, dy) = dA_t(\omega) K(\omega; t, dy),$$

which is short for

$$\int_0^t \int_{\mathbb{R}} q_s \, N^p(\omega; ds, dy) = \int_0^t \int_{\mathbb{R}} q_s \, K(\omega; s, dy) dA_s(\omega).$$

This *disintegration* is not unique. The compensator process is predictable and of finite variation and therefore can play a role in the semimartingale decomposition.

In the case of integer-valued random measures there is a set of stopping times $(\tau_i; i \geq 1)$ (random times at which the process jumps) which allows us to write

$$J_t = \int_{[0,t]\times\mathbb{R}} q(\omega; s, X_s) N(\omega; ds, dy) = \sum_i q(\omega; \tau_i, X_{\tau_i}) 1_{\{\tau_i < t\}},$$

where 1 is an indicator function

$$1_A = \begin{cases} 1 & \text{if } A \text{ holds,} \\ 0 & \text{otherwise.} \end{cases}$$

It will be useful to know how to calculate the quadratic variation of any semimartingale. The decompositions of Section 6.1 will be useful in this task. Recall that any semimartingale can be decomposed into a local martingale M_t and a finite variation process A_t,

$$S_t = S_0 + M_t + A_t, \quad M_0 = A_0 = 0.$$

Furthermore, any local martingale M may be decomposed into

$$M_t = M_t^c + M_t^d,$$

where M^c is a continuous local martingale, and M^d is a purely discontinuous local martingale.

The continuous part of the jump diffusion process X_t defined by Equation (7.18) is

$$X_t^c = \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s,$$

and the purely discontinuous part is

$$X_t^d = J_t = \sum_{s \le t} (X_s - X_{s^-}).$$

Neither of these are necessarily local martingales, we may need to find the compensators to make further progress.

Here are some facts that will be used to prove Theorem 14.

- recapping the definition of a purely discontinuous local martingale M^d: for any continuous semimartingale X^c we have $\left\langle M^d, X^c \right\rangle_t = 0$;
- for any two purely discontinuous semimartingales X^d and Y^d,

$$\left\langle X^d, Y^d \right\rangle_t = \lim_{mesh(\tau_n) \to 0} \sum_{i=0}^{n-1} (X_{t_{i+1}}^d - X_{t_i}^d)(Y_{t_{i+1}}^d - Y_{t_i}^d) = \sum_{s \le t} (X_s - X_{s^-})(Y_s - Y_{s^-}).$$

  We shall often denote $X_s - X_{s^-}$ by ΔX_s.

- for any continuous process of finite variation A_t^c and any other semimartingale we have $\langle A^c, X \rangle_t = 0$.

The calculus of jumps

The following theorem is the generalisation of Theorem 11 from Itô processes to any semimartingale. The proof is similar to that found in Jacod & Shiryaev (1987).

Theorem 14 (Itô's formula for any semimartingale). *Let X be a semimartingale, and $f \in C^2(\mathbb{R})$. Then $f(X)$ is also a semimartingale and is governed by the differential*

$$\begin{aligned} f(X_t) =& f(X_0) + \int_0^t f'(X_{s^-}) dX_s + \int_0^t \frac{1}{2} f''(X_{s^-}) d \langle X^c \rangle_s \\ &+ \sum_{s \le t} \left[f(X_s) - f(X_{s^-}) - f'(X_{s^-}) \Delta X_s \right], \end{aligned} \tag{7.19}$$

where $\langle X^c \rangle_t$ denotes the quadratic variation of the continuous part of X.

Proof. We shall prove the result first for functions f which are polynomials.
The result is trivially true when $f(x) = x$. If we can prove the formula also holds for $f(x) = xg(x)$ by assuming (7.19) holds for g, then by induction (7.19) must hold for any polynomial.
To simplify the notation, associate to any $g \in C^2(\mathbb{R})$ $\hat{g}$ defined as

$$\hat{g}(x, y) = g(x) - g(y) - g'(y)(x - y).$$

By the product rule, then the inductive assumption

$$\begin{aligned} f(X_t) &= f(X_0) + \int_0^t X_{s^-} dg(X_s) + \int_0^t g(X_{s^-}) dX_s + \langle X, g(X) \rangle_t, \\ &= f(X_0) + \int_0^t X_{s^-} g'(X_s) dX_s + \int_0^t \frac{1}{2} X_{s^-} g''(X_{s^-}) d\langle X^c \rangle_s \qquad (7.20) \\ &\quad + \sum_{s \le t} X_{s^-} \hat{g}(X_s, X_{s^-}) + \int_0^t g(X_{s^-}) dX_s + \langle X, g(X) \rangle_t. \end{aligned}$$

Computing $\langle X, g(X) \rangle_t$ directly using (7.19), and noting that the second and third terms in (7.19) are processes of finite variation; the second being continuous, the third being purely discontinuous.

$$\begin{aligned} \langle X, g(X) \rangle_t &= \int_0^t g'(X_{s^-}) d\langle X \rangle_s + 0 + \sum_{s \le t} \Delta X_s \Delta \hat{g}(X_s, X_{s^-}), \\ &= \int_0^t g'(X_{s^-}) d\left(\langle X^c \rangle_s + \sum_{s \le t} (\Delta X_s)^2 \right) \\ &\quad + \sum_{s \le t} \Delta X_s [g(X_s) - g(X_{s^-}) - g'(X_{s^-}) \Delta X_s], \\ &= \int_0^t g'(X_{s^-}) d\langle X^c \rangle_s + \sum_{s \le t} \Delta X_s (g(X_s) - g(X_{s^-})). \qquad (7.21) \end{aligned}$$

Combining equations (7.20) and (7.21) verifies that f satisfies Equation (7.19).

Since f is C^2 we can always find a polynomial g which will match f and its partial derivatives of first and second order. An application of Lebesgue's Dominated Convergence Theorem shows that we can approximate each of $\int_0^t f'(X_{s^-}) dX_s$, $\int_0^t f''(X_{s^-}) d\langle X^c \rangle_s$, and $\sum_{s \le t} \hat{f}(X_s, X_{s^-})$ arbitrarily close using a polynomial g. $\square$

In the particular case of the jump diffusion (7.18), Itô's formula (7.19) may be evaluated further. Observing that

$$\int_0^t f'(X_{s^-})dX_s = \int_0^t \mu_t f'(X_{s^-})ds + \int_0^t \sigma_t f'(X_{s^-})dW_s + \sum_{s\le t} f'(X_{s^-})\Delta X_s,$$

and that

$$\int_0^t \frac{1}{2} f''(X_{s^-})d\left\langle X^c\right\rangle_s = \int_0^t \frac{1}{2}\sigma_s^2 f''(X_{s^-})ds,$$

enables us to write Itô's formula for the jump diffusion as

$$\begin{aligned} f(X_t) = f(X_0) + \int_0^t \mu_s f'(X_{s^-}) + \frac{1}{2}\sigma_s^2 f''(X_{s^-})ds + \int_0^t \sigma_s f'(X_{s^-})dW_s \\ + \sum_{s\le t}\left(f(X_s) - f(X_{s^-})\right). \end{aligned}$$

One can easily identify the continuous and purely discontinuous parts of this process. For many applications, we may wish to calculate the various moments of $f(X_t)$. If we can identify the compensator process which will make the purely discontinuous part of $f(X_t)$ a martingale, then evaluation of $\mathbb{E}(f(X_t))$ will be greatly simplified.

Recall that $\lambda(x)$ denotes the rate at which jumps occur when the process X_t is at x. Also remember that the distribution $\bar{\Pi}(X_t,\cdot)$ specifies the jump sizes at the times of the jumps

$$\Pr\left(J_{\tau_{i+1}} - J_{\tau_i} \le y \middle| X_{\tau_{i+1}^-}\right) = \int_{-\infty}^{y} \bar{\Pi}(X_{\tau_{i+1}^-}, dy).$$

The quantity

$$m(X_{\tau_{i+1}^-}) = \mathbb{E}_{\bar{\Pi}}\left(J_{\tau_{i+1}} - J_{\tau_i} \middle| X_{\tau_{i+1}^-}\right),$$

is the expected size of the $(i+1)$th jump, given the position of the process X just before the jump.

The compensator A_t of the jump process is given by

$$\begin{aligned} A_t = \int_0^t \int_{\mathbb{R}} q_s\, N^p(\omega; ds, dy) &= \int_0^t \lambda(X_s) \int_{\mathbb{R}} q_s\, \bar{\Pi}(X_s, dy)ds, \\ &= \int_0^t \lambda(X_s) m(X_s) ds. \end{aligned}$$

It follows that the jump process can be written as

$$\begin{aligned} J_t &= A_t + M_t, \\ &= \int_0^t \lambda(X_s) m(X_s) ds + M_t, \end{aligned}$$

where $M_t = J_t - A_t$ is a purely discontinuous local martingale. Thus we have the stochastic differential of the jump diffusion

$$X_t = X_0 + \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s + \int_{[0,t]\times\mathbb{R}} q(s, X_s) N(ds, dy),$$

given as

$$dX_t = (\mu_t + \lambda(X_t) m(X_t)) \, dt + \sigma_t dW_t + dM_t.$$

It follows that $A_t^f = \int_0^t \lambda(X_s) m^f(X_s) ds$, where

$$m^f(X_{\tau_{i+1}^-}) = \mathbb{E}_{\bar{\Pi}} \left(f(J_{\tau_{i+1}}) - f(J_{\tau_i}) \middle| X_{\tau_{i+1}^-} \right),$$

Is the corresponding compensator for the new process $f(X_t)$, chosen such that $M_t^f = \sum_{s \le t} \Delta f(X_t) - A_t^f$ is a martingale. The expected value of $f(X_t)$ is

$$\mathbb{E}(f(X_t)) = \mathbb{E}(f(X_0)) + \mathbb{E}\left[\int_0^t f'(X_{s^-})(\mu_s + \lambda(X_{s^-}) m^f(X_{s^-})) + \frac{1}{2}\sigma_s^2 f''(X_{s^-}) dt \right].$$

The quadratic variation of M_t is obtained in the following manner

$$\begin{aligned} \langle M \rangle_t &= \langle J - A \rangle_t, \\ &= \langle J \rangle_t, \quad \text{since } A_t \text{ is of finite variation} \\ &= \lim_{mesh \to 0} \sum_{i=1}^{n-1} (J_{t_{i+1}} - J_{t_i})^2, \\ &= \sum_i q_{\tau_i}^2 \, 1_{\{\tau_i < t\}}. \end{aligned}$$

Thus the process $\langle M \rangle_t$ is a pure jump process generated by the sequence (τ_i, q_{τ_i}). It's compensator is

$$\int_0^t \lambda(X_s)\nu(X_s)ds,$$

where

$$\nu(X_{\tau_{i+1}^-}) = \mathbb{E}_{\bar{\Pi}}\left((J_{\tau_{i+1}} - J_{\tau_i})^2 \middle| X_{\tau_{i+1}^-}\right).$$

For this reason the convention $d\langle M \rangle_t = \lambda(X_s)\nu(X_s)ds$ is often used.

7.1.5 Girsanov's Theorem

So far in this section we haven't mentioned the underlying probability measures behind our stochastic processes. We have developed some basic tools for manipulating stochastic equations, but they are manipulation of functions of Brownian motion, not a manipulation of measure. This apparent neglect is somewhat surprising given the remarks of Chapter 6 with respect to the important role the theory of martingales has in option pricing. Actually, we haven't ignored the importance of measure, just suppressed the dependence on it in our notation.

It is often the case that a stochastic process may be converted to a martingale by changing the underlying probability measure. Girsanov's Theorem tells us the conditions required for such a measure to exist and the form that the Radon-Nikodym derivative takes for this change of measure. Change of measure is useful for risk-neutral pricing of financial derivatives, solving stochastic differential equations, and calculating statistics of interest such as stopping time distributions.

Given a drifting Brownian motion the Cameron-Martin-Girsanov idea is to define a random variable Λ as

$$\Lambda = \exp\left(-\int_0^T \gamma_t dW_t - \frac{1}{2}\int_0^T \gamma_t^2 dt\right),$$

where γ_t corresponds to the drift process; in the case above $\gamma_t = \gamma = \frac{\mu - r + \frac{1}{2}\sigma^2}{\sigma}$. The following properties will then hold (check these)

- $\Lambda \geq 0$;
- $\mathbb{E}_{\mathbb{P}}(\Lambda) = 1$;
- the measure $\mathbb{Q}$ defined as

$$\mathbb{Q}(A) = \int_A \Lambda d\mathbb{P}, \tag{7.22}$$

  is a probability measure, and Λ is the Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}$;
- $\mathbb{Q}$ defined in (7.22) is the risk-neutral probability measure.

Theorem 15 (Girsanov's Theorem). *Let W_t be a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose γ_t to be a process adapted to the accompanying filtration $\mathcal{F}_t$. Define*

$$\tilde{W}_t = W_t + \int_0^t \gamma_s ds,$$

and

$$\Lambda = \exp\left(-\int_0^T \gamma_u dW_u - \frac{1}{2}\int_0^T \gamma_u^2 du\right),$$

and define a new probability measure $\mathbb{Q}$ by

$$\mathbb{Q}(A) = \int_A \Lambda d\mathbb{P},$$

then, provided $\mathbb{E}_{\mathbb{P}}\left(\exp(\frac{1}{2}\int_0^T \gamma_t^2 dt)\right) < \infty$, under $\mathbb{Q}$ the process $\tilde{W}_t$ is a Brownian motion. The random variable Λ is the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$.

Corollary 16 (Converse to Girsanov's Theorem). *If W_t is a $\mathbb{P}$-Brownian motion, and $\mathbb{Q}$ is a measure equivalent to $\mathbb{P}$, then there exists a γ_t such that*

$$\tilde{W}_t = W_t + \int_0^t \gamma_s ds,$$

is a $\mathbb{Q}$-Brownian motion.

The effect of changing measure using Girsanov's Theorem is to change the mean. The variance does not change. In consequence

- martingales may be destroyed or created; but
- volatilities, quadratic variation, and covariations are unaffected.

Example 22. *Suppose that X is a stochastic process which satisfies the stochastic differential equation*

$$dX_t = \sigma_t dW_t + \mu_t dt,$$

with W_t a $\mathbb{P}$-Brownian motion. Is there a measure $\mathbb{Q}$ such that

$$dX_t = \sigma d\tilde{W}_t + \nu_t dt,$$

where $\tilde{W}_t$ is a $\mathbb{Q}$-Brownian motion?

Rewrite dX_t as

$$dX_t = \sigma_t \left(dW_t + \left(\frac{\mu_t - \nu_t}{\sigma_t} \right) dt \right) + \nu_t dt,$$

then provided $\gamma_t = \frac{\mu_t - \nu_t}{\sigma_t}$ satisfies $\mathbb{E}_{\mathbb{P}} \left(\exp(\frac{1}{2} \int_0^T \gamma_t^2 dt) \right) < \infty$ by Girsanov's Theorem we have

$$\tilde{W}_t = W_t + \int_0^t \frac{\mu_s - \nu_s}{\sigma_s} ds,$$

is a $\mathbb{Q}$-Brownian motion. Furthermore, under $\mathbb{Q}$

$$dX_t = \sigma_t d\tilde{W}_t + \nu_t dt,$$

If in the above example ν_t was chosen to be identically 0, then the drift would have been eliminated. Thus, we might find the solution to the stochastic differential equation $dX_t = \sigma_t dW_t + \mu_t dt$ by looking for a solution to $dX_t = \sigma_t d\tilde{W}_t$ and applying the change of measure. In this way, Girsanov's Theorem provides us with yet another tool for solving stochastic differential equations.

7.2 Solving SDEs

We know have some of the tools which allow us to find the stochastic differential equation of some process given its definition. Conversely, given the stochastic differential $dX_t = a(t, X_t)dt + b(t, X_t)dW_t$, we can sometimes find an expression for X_t. In this section we present the solution to a few special SDE's.

Definition 2. *A* strong solution *to a stochastic differential equation is a function* $X_t = F(t, (W_s, s < t))$ *which satisfies the integral equation*

$$X_t = \int_0^t a(t, X_t)ds + \int_0^t b(t, X_t)dW_s.$$

The white noise process is defined as $\xi_t = \frac{dW_t}{dt}$. Since W_t is nowhere differentiable, this is not a derivative in the usual sense. A more sensible definition might read: if σ_t is the intensity of the noise at a some point in space and time, then

$$\int_0^t \sigma_t \xi_t dt = \int_0^t \sigma_t dW_t.$$

In other words, whenever we see $\xi_t dt$ we can replace it by dW_t and vice-versa.

Example 23 (Growth with uncertain interest). *Recall, B_t denotes the value \$1 grows to after time t. If invested with a continuously compounding interest risk-free rate r, then B_t satisfies the ordinary differential equation $\frac{dB_t}{dt} = rB_t$.*

Now suppose instead that the interest rate is modelled as $r+\sigma\xi_t$. The ODE becomes an SDE: $\frac{dB_t}{dt} = (r + \sigma\xi_t)B_t$, or in the standard form

$$dB_t = rB_t dt + \sigma B_t \xi_t dt.$$

Noticing $\xi_t dt$ can be replaced by dW_t yields

$$dB_t = rB_t dt + \sigma B_t dW_t,$$

which one might recognise as the SDE of geometric Brownian motion,

$$B_t = \exp\left((r - \frac{\sigma^2}{2})t + \sigma W_t\right).$$

The solution to some SDE's can be found by making an educated guess. If we know a few common forms and their solutions then given a new SDE we may be able to guess it's solution modulo some unknown constants, and then check using Itô's formula.

Lemma 17. *The process* $U_t = \exp(Y_t - \frac{1}{2}\langle Y\rangle_t)$ *satisfies the SDE*

$$dU_t = U_t dY_t.$$

Proof. Put $U_t = e^{V_t}$ then by Itô

$$dU_t = U_t dV_t + \frac{1}{2}U_t d\langle V\rangle_t,$$

but $V_t = Y_t - \frac{1}{2}\langle Y\rangle_t$, so $dV_t = dY_t - \frac{1}{2}d\langle Y\rangle_t$. The quadratic variation process $\langle Y\rangle_t$ is deterministic ($d\langle Y\rangle_t = \sigma_t^2 dt$ say), and therefore contributes nothing to the quadratic variation $\langle V\rangle_t$. In other words $d\langle V\rangle_t = d\langle Y\rangle_t$. The stochastic differential for U_t must be

$$dU_t = U_t dY_t - \frac{1}{2}U_t d\langle Y\rangle_t + \frac{1}{2}U_t d\langle Y\rangle_t.$$

□

We have shown that

$$U_t = e^{Y_t - \frac{1}{2}\langle Y\rangle_t} \implies dU_t = U_t dY_t,$$

and so, whenever we see a stochastic differential of a similar form to $dU_t = U_t dY_t$, a good guess might be to look for a solution of the form $U_t = e^{Y_t - \frac{1}{2}\langle Y\rangle_t}$.

Example 24. *Suppose we wish to solve the SDE we derived in our uncertain interest rate model (Example 23)*

$$dB_t = rB_t dt + \sigma B_t dW_t.$$

Put $dR_t = rdt + \sigma dW_t$, *then* $dB_t = B_t dR_t$. *Both of these equations are easy to solve:*

$$\begin{aligned} R_t &= rt + \sigma W_t, \quad \text{with } \langle R\rangle_t = \sigma^2 t, \\ B_t &= e^{R_t - \frac{1}{2}\langle R\rangle_t}. \end{aligned}$$

Substituting the known values for R_t and $\langle R \rangle_t$ suggests a solution

$$B_t = \exp\left((r - \frac{1}{2}\sigma^2)t + \sigma W_t\right),$$

which should be checked with Itô to ensure B_t satisfies the SDE

The next example introduces us to a process which has been suggested as an alternative to pure Brownian motion as the basis for stock price models (Hall, Mathews & Platen 1996).

Example 25 (Ornstein-Uhlenbeck process). *The Wiener process (Brownian motion) is good for modelling motion of a particle over long periods, however if we look closely it is a bad approximation to local movement of a particle since the paths are nowhere differentiable. According to Newton's laws of physics only particles of zero mass can follow such paths. The* Ornstein-Uhlenbeck process *assumes the velocity (rather than the position) of a particle undergoes Brownian motion. The ramifications are that the velocity paths are continuous and the sample paths of the process itself are almost surely continuously differentiable (smooth) functions of time.*

The Ornstein-Uhlenbeck process follows the Langevin equation

$$dX_t = cX_t dt + \sigma dW_t. \tag{7.23}$$

This process is one of the standard example given in most time-series courses: imagine we are observing a continuous-time process at discrete intervals with $dt = 1$, Equation (7.23) might be equivalently represented

$$\begin{aligned} X_{t+1} - X_t &= cX_t + \sigma(W_{t+1} - W_t), \\ \text{or} \quad X_{t+1} &= (c+1)X_t + \sigma Z_t, \end{aligned}$$

where $Z_t \sim N(0,1)$, for $t = 0, 1, 2, \ldots$.

In order to solve (7.23) consider the substitution $Y_t = e^{-ct}X_t$. This process is of product form, a deterministic part e^{-ct} times the random process X_t. The covariation between the two factors of the product is zero since one of the factors is purely

deterministic. By the product rule

$$\begin{aligned} dY_t &= e^{-ct}dX_t + X_t d(e^{-ct}) + 0, \\ &= e^{-ct}dX_t - cX_t e^{-ct}dt, \\ &= e^{-ct}(cX_t dt + \sigma dW_t) - cX_t e^{-ct}dt, \\ &= \sigma e^{-ct}dW_t, \end{aligned}$$

which as an integral equation reads

$$Y_t = Y_0 + \int_0^t \sigma e^{-cs}dW_s.$$

Substituting back to get X_t and noting that $X_0 = Y_0$

$$\begin{aligned} e^{-ct}X_t &= X_0 + \int_0^t \sigma e^{-cs}dW_s, \\ X_t &= e^{ct}X_0 + e^{ct}\int_0^t \sigma e^{-cs}dW_s. \end{aligned}$$

7.2.1 Solving linear SDEs

Stochastic differential equations for which there exist strong analytical solutions are a rare occurrence. In this section we consider a class of stochastic differential equations for which a unique strong solution can always be found. The class being that of the *linear stochastic differential equation.* We shall solve the general linear SDE and in doing so learn a general method by which every equation of this class may be solved.

The general linear stochastic differential equation takes the form

$$dX_t = (\alpha_t + \beta_t X_t)dt + (\gamma_t + \delta_t X_t)dW_t. \tag{7.24}$$

It is linear in the sense that the parameters $a(t,x) = \alpha_t + \beta_t x$ and $b(t,x) = \gamma_t + \delta_t x$ are linear in x. The functions $a(t,x)$ and $b(t,x)$ must satisfy the conditions on integrands (as stated in Section 10) for the differential dX_t to make sense.

We have seen a few linear stochastic differential equations already:

- Brownian motion with drift $dX_t = \alpha dt + \gamma dW_t$;

- Geometric Brownian motion $dX_t = \beta X_t dt + \delta X_t dW_t$;
- Ornstein-Uhlenbeck process $dX_t = \beta X_t dt + \gamma dW_t$;
  and soon we shall meet the
- Vasicek interest rate model $dX_t = (\alpha + \beta X_t)dt + \gamma dW_t$.

To find a solution in the general case with all non-identically-zero parameters, look for a solution of the form

$$\begin{aligned} X_t &= U_t V_t, \\ \text{where} \quad dU_t &= \beta_t U_t dt + \delta_t U_t dW_t, \quad U_0 = 1, \\ \text{and} \quad dV_t &= a_t dt + b_t dW_t, \quad V_0 = X_0. \end{aligned} \tag{7.25}$$

The parameters a_t and b_t are chosen such that Equation (7.25) actually holds. One way to determine a_t and b_t is to apply the product rule to X_t and equate the coefficients with those of Equation (7.24). By the product rule

$$\begin{aligned} dX_t &= d(U_t V_t), \\ &= U_t dV_t + V_t dU_t + \langle U, V \rangle_t, \\ &= U_t(a_t dt + b_t dW_t) + V_t(\beta_t U_t dt + \delta_t U_t dW_t) + U_t \delta_t b_t dt. \end{aligned}$$

Pattern matching with Equation (7.24) leads to the system

$$\begin{aligned} a_t U_t + \beta_t U_t V_t + U_t \delta_t b_t &= \alpha_t + \beta_t X_t, \\ b_t U_t + V_t \delta_t U_t &= \gamma_t + \delta_t X_t. \end{aligned}$$

These equations are satisfied if a_t and b_t are chosen such that

$$b_t U_t = \gamma_t, \quad \text{and } a_t U_t = \alpha_t - \delta_t \gamma_t.$$

Thus

$$V_t = V_0 + \int_0^t \frac{\alpha_s - \delta_s \gamma_s}{U_s} ds + \int_0^t \frac{\gamma_s}{U_s} dW_s.$$

Now we need to find an expression for U_t. The differential dU_t may be rewritten as

$$dU_t = U_t dY_t, \quad \text{where } dY_t = \beta_t dt + \delta_t dW_t.$$

Lemma 17 tells us that U_t has a solution in terms of Y_t given by

$$U_t = \exp(Y_t - \frac{1}{2}\langle Y\rangle_t).$$

After substituting $Y_t = \int_0^t \beta_s ds + \int_0^t \delta_s dW_s$, and $\langle Y\rangle_t = \int_0^t \delta_s^2 ds$, U_t becomes

$$U_t = \exp\left(\int_0^t (\beta_s - \frac{1}{2}\delta_s^2)ds + \int_0^t \delta_s dW_s\right).$$

Putting the two parts U_t and V_t together we find a solution to the general linear stochastic differential equation:

Lemma 18. *The general linear stochastic differential equation*

$$dX_t = (\alpha_t + \beta_t X_t)dt + (\gamma_t + \delta_t X_t)dW_t,$$

has a solution given by

$$X_t = X_0 + U_t \int_0^t \frac{\alpha_s - \delta_s\gamma_s}{U_s} ds + U_t \int_0^t \frac{\gamma_s}{U_s} dW_s,$$

where U_t is given by

$$U_t = \exp\left(\int_0^t (\beta_s - \frac{1}{2}\delta_s^2)ds + \int_0^t \delta_s dW_s\right),$$

and $U_0 = 1$.

Example 26 (The Vasicek interest rate model). *This is one of the standard models for describing the time value of money. The interest rate for borrowing and lending is not assumed to be a constant, but a random function R_t of time t. In the* Vasicek model *it is given by the linear stochastic differential equation*

$$dR_t = c(\mu - R_t)dt + \sigma dW_t. \tag{7.26}$$

This is a linear stochastic differential equation, the same kind as referred to by Lemma 18, and so we could simply read off the solution with

$$\alpha_t = c\mu, \quad \beta_t = -c, \quad \gamma_t = 0, \quad \text{and } \delta_t = \sigma,$$

or we could solve the equation using the method used to justify the Lemma.

Look for a solution of the form $R_t = U_t V_t$ where

$$\begin{aligned} dU_t &= -cU_t dt, \quad U_0 = 1, \\ \text{and} \quad dV_t &= a_t dt + b_t dW_t, \quad V_0 = R_0. \end{aligned} \tag{7.27}$$

Now Equation (7.27) is an ordinary differential equation resulting in the purely deterministic process

$$U_t = e^{-ct}.$$

Before we can write an expression for V_t we need to find a_t and b_t. If we compute dR_t in terms of dV_t (and in turn a_t and b_t) then by comparison with Equation (7.26) we may form a solvable system for the unknowns. By the product rule

$$\begin{aligned} dR_t &= U_t dV_t + V_t dU_t + d\langle V, U\rangle_t, \\ &= U_t a_t dt + U_t b_t dW_t - cV_t U_t dt. \end{aligned} \tag{7.28}$$

The covariation term $d\langle V, U\rangle_t$ makes no contribution because the process U_t is deterministic. Also note that V_t is being chosen such that $R_t = U_t V_t$, and therefore the term $-cV_t U_t dt$ is $-cR_t dt$. Equating the coefficients of dt and dW_t in Equation (7.28) with those of (7.26) yields

$$\begin{aligned} U_t a_t - cV_t U_t = c(\mu - R_t) \quad &\Longrightarrow \quad a_t = c\mu e^{ct}, \\ \sigma = U_t b_t \quad &\Longrightarrow \quad b_t = \sigma e^{ct}. \end{aligned}$$

Thus

$$\begin{aligned} V_t &= V_0 + \int_0^t c\mu e^{cs} ds + \int_0^t \sigma e^{cs} dW_s, \\ &= V_0 + \mu(e^{ct} - 1) + \int_0^t \sigma e^{cs} dW_s. \end{aligned}$$

And so

$$R_t = R_0 e^{-ct} + \mu(1 - e^{-ct}) + e^{-ct} \int_0^t \sigma e^{cs} dW_s.$$

7.2.2 Solving SDEs by Itô's formula

It is sometimes possible to reformulate a stochastic differential equation into a solvable system of partial differential equations.

To set the scene, imagine we are given and asked to solve

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t, \tag{7.29}$$

we can speculate that the solution X_t of this equation follows a function $f(t, W_t)$ taking time t as a parameter and driven by Brownian motion W_t. If the function $f(t, x)$ has continuous derivatives of the second order then we can apply Itô's formula and calculate

$$\begin{aligned} dX_t &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dW_t + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dW_t)^2 + \frac{\partial^2 f}{\partial t \partial x}dtdW_t, \\ &= \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\right)dt + \frac{\partial f}{\partial x}dW_t, \end{aligned}$$

where we've ignored all terms involving either $(dt)^2$ or $dtdW_t$. Equating coefficients with those of (7.29) leads to the system of partial differential equations

$$\begin{aligned} \mu(t, x) &= \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}, \\ \sigma(t, x) &= \frac{\partial f}{\partial x}. \end{aligned}$$

If we can solve this system then we have found a function $f(t, x)$ such that the process $X_t = f(t, W_t)$ satisfies the stochastic differential (7.29).

Example 27. *Suppose we are asked to solve*

$$dX_t = 1dt + 2W_t dW_t.$$

If X_t has the form $f(t, W_t)$, then

$$dX_t = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\right) dt + \frac{\partial f}{\partial x} dW_t,$$

which leads to the system of PDEs:

$$\begin{aligned} 1 &= \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}, \\ 2x &= \frac{\partial f}{\partial x}. \end{aligned}$$

Differentiating the second equation with respect to x reveals $2 = \frac{\partial^2 f}{\partial x^2}$, and substituting this into the first PDE yields $1 = \frac{\partial f}{\partial t} + \frac{1}{2}2$ or $\frac{\partial f}{\partial t} = 0$. So f is constant in t and $\frac{\partial f}{\partial x} = 2x$. Thus $f(t, x) = x^2 + c$ for some constant c. Using the initial condition $f(0, W_0) = X_0$ to determine c we have $X_t = X_0 + W_t^2$.

Example 28. *Suppose we are asked to find a process X_t which follows the stochastic differential*

$$dX_t = \mu X_t dt + \sigma X_t dW_t.$$

If $X_t = f(t, W_t)$, then

$$dX_t = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\right) dt + \frac{\partial f}{\partial x} dW_t,$$

which leads to the system of PDEs:

$$\begin{aligned} \mu f &= \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}, \\ \sigma f &= \frac{\partial f}{\partial x} \quad \Longrightarrow \sigma\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial x^2}. \end{aligned}$$

The first PDE becomes

$$\begin{aligned} \mu f &= \frac{\partial f}{\partial t} + \frac{1}{2}\sigma\frac{\partial f}{\partial x}, \\ &= \frac{\partial f}{\partial t} + \frac{1}{2}\sigma^2 f. \end{aligned}$$

After simplifying the system becomes

$$(\mu - \frac{1}{2}\sigma^2)f = \frac{\partial f}{\partial t}, \quad \sigma f = \frac{\partial f}{\partial x},$$

which may be solved using separation of variables. It is easy to check that

$$f(t,x) = Ae^{(\mu - \frac{1}{2}\sigma^2)t + \sigma x}$$

satisfies the system of PDEs. And that the process X_t defined as

$$X_t = X_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}$$

satisfies the original stochastic differential equation.

7.2.3 Solving SDEs by Stratonovich calculus

Analogous to the Itô stochastic differential equation is the *Stratonovich stochastic differential equation*

$$X_t = X_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s) \circ dW_s,$$

or in the equivalent notation

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t) \circ dW_t.$$

For the Stratonovich integral to be defined there is a set of criteria which the integrand must satisfy. These conditions are similar to those for the definition of the Itô integral. For our purposes it is enough to know that for some integrand Y_t adapted to the filtration of X_t, $\int_0^t Y_s dX_s$ exists implies $\int_0^t Y_s \circ dX_s$ is also defined. And furthermore, if the Stratonovich integral of Y_s with respect to X_s does exist then it satisfies the *transformation formula*:

$$\int_0^t Y_s \circ dX_s = \int_0^t Y_s dX_s + \frac{1}{2}\langle Y, X\rangle_t .$$

The transformation formula is a rule which allows us to convert from Stratonovich to Itô calculus and vice-versa.

In Stratonovich calculus classical rules such as the chain rule and product rule hold exactly. The proofs are quite simple. Start with Itô's formula or the Itô product rule and apply the transformation formula.

Theorem 19 (Integration by parts: the Stratonovich product rule). *For any two continuous processes X_t and Y_t adapted to the same Brownian filtration*

$$\circ d(X_t Y_t) = X_t \circ dY_t + Y_t \circ dX_t.$$

Proof. Left as an exercise. □

Theorem 20 (Change of variables: the Stratonovich chain rule). *For f three times continuously differentiable*

$$\circ d(f(X_t)) = f'(X_t) \circ dX_t.$$

Proof. First note that the Itô and Stratonovich integrals coincide when the integrand is deterministic (see Remark 7.1.3). In particular

$$\int_0^t 1 \circ d(f(X_s)) = \int_0^t 1 d(f(X_s)), \quad \text{or} \quad \circ\, d(f(X_s)) = d(f(X_s)).$$

Now by Itô's formula

$$d(f(X_t)) = f'(X_t) dX_t + \frac{1}{2} f''(X_t) d\langle X \rangle_t .$$

Put $Y_t = f'(X_t)$ and apply the transformation formula

$$Y_t dX_t = Y_t \circ dX_t - \frac{1}{2} d\langle X, Y \rangle_t ,$$

to get

$$d(f(X_t)) = f'(X_t) \circ dX_t - \frac{1}{2} d\langle X, Y \rangle_t + \frac{1}{2} f''(X_t) d\langle X \rangle_t .$$

The result is proven if we can show $f''(X_t) d\langle X \rangle_t - d\langle X, Y \rangle_t = 0$. Applying Itô's formula once again

$$dY_t = f''(X_t) dX_t + \frac{1}{2} f'''(X_t) d\langle X \rangle_t ,$$

and multiplying through by dX_t we have

$$dX_t dY_t = d\langle X, Y\rangle_t = f''(X_t)d\langle X\rangle_t + 0.$$

The second term is taken to be negligible since it contains $dX_t d\langle X\rangle_t$. □

Example 29. *Consider the Itô stochastic differential equation*

$$dX_t = \frac{1}{2}f(X_t)f'(X_t)dt + f(X_t)dW_t. \tag{7.30}$$

Applying the transformation formula to $f(X_t)dW_t$, this equation may be rewritten as an equivalent Stratonovich stochastic differential equation

$$dX_t = \frac{1}{2}f(X_t)f'(X_t)dt + f(X_t)\circ dW_t - \frac{1}{2}d\langle Y, W\rangle_t,$$

where $Y_t = f(X_t)$. Now $d\langle Y, W\rangle_t = dY_t dW_t$ and $dY_t = d(f(X_t)) = f'(X_t)\circ dX_t$, and so

$$\begin{aligned} dX_t &= \frac{1}{2}f(X_t)f'(X_t)dt + f(X_t)\circ dW_t - \frac{1}{2}f'(X_t)d\langle X, W\rangle_t, \\ &= \frac{1}{2}f(X_t)f'(X_t)dt + f(X_t)\circ dW_t - \frac{1}{2}f'(X_t)f(X_t)dt. \end{aligned}$$

The Stratonovich stochastic differential equation equivalent to (7.30) *is given by*

$$dX_t = f(X_t)\circ dW_t.$$

We try to solve this equation using the classical rules of calculus. That is to say, we attempt to solve the ordinary differential equation

$$dx = f(x)dw(t),$$

using separation of variables

$$\int_{x(0)}^{x(t)} \frac{1}{f(x)}dx = \int_0^t dw(t).$$

Hopefully one can solve this equation and replace $x(t)$ by X_t and $w(t)$ by W_t to get a solution to (7.30).

Example 30. *A slightly more complicated Itô stochastic differential equation is given by*

$$dX_t = (qf(X_t) + \frac{1}{2}f(X_t)f'(X_t))dt + f(X_t)dW_t, \tag{7.31}$$

where q is a constant. Applying the transformation formula to $f(X_t)dW_t$ as in the previous example we get

$$dX_t = qf(X_t)dt + f(X_t) \circ dW_t, \tag{7.32}$$

which we aim to solve via the equation

$$dx(t) = qf(x(t))dt + f(x(t))dw(t),$$

where $w(t)$ is a differentiable function. Separation of variables leads to

$$\int_{x(0)}^{x(t)} \frac{1}{f(x)}dx = qt + w(t) - w(0).$$

If the left-hand side can be evaluated to say

$$\int_{x(0)}^{x(t)} \frac{1}{f(x)}dx = g(x(t)) - g(x(0)),$$

then

$$X_t = g^{-1}(g(X_0) + qt + W_t - W - 0),$$

satisfies the Stratonovich stochastic differential Equation (7.32) *and hence it is also a solution to the Itô stochastic differential* (7.31).

7.2.4 Solving SDEs using Girsanov's Theorem

One of the applications of Girsanov's Theorem lies in solving certain nonlinear SDEs. Such SDEs are unable to be handled by the methods mentioned so far.

Consider the class of SDEs which fit the form

$$dX_t = f(t, X_t)dt + c(t)X_t dW_t, \tag{7.33}$$

where $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $c : \mathbb{R} \to \mathbb{R}$ are continuous functions, and W is a standard Brownian motion under some measure $\mathbb{P}$. Our use for Girsanov's Theorem is to completely remove the drift $f(t, X_t)dt$. Under the new measure, $\mathbb{Q}$, dX is much easier to solve. Then, provided we can determine exactly how much the $\mathbb{P}$-Brownian's drift was changed, the solution (under $\mathbb{Q}$) can be converted back to give the solution under $\mathbb{P}$. The final step needs to be qualified by "provided", since it requires us to solve a differential equation.

Suppose that we choose a measure $\mathbb{Q}$ under which the process $\tilde{W}$ is a standard Brownian motion and $\tilde{W} = W - \int_0^t \gamma_s ds$ for some adapted γ_t. Under this new measure, replacing dW_t by $d\tilde{W}_t + \gamma_t dt$, Equation (7.33) becomes

$$dX_t = (f(t, X_t) + c(t)X_t\gamma_t)\, dt + c(t)X_t d\tilde{W}_t. \tag{7.34}$$

Now (if possible) we should choose γ_t such that $f(t, X_t) + c(t)X_t\gamma_t$ is identically zero. Doing so reduces (7.34) to $dX_t = c(t)X_t d\tilde{W}_t$, thus

$$X_t = X_0 \exp\left(\int_0^t c(s)d\tilde{W}_s - \frac{1}{2}\int_0^t c(s)^2 ds\right).$$

Now changing back from $\mathbb{Q}$ to $\mathbb{P}$

$$\begin{aligned} X_t &= X_0 \exp\left(\int_0^t c(s)(dW_s - \gamma_s ds) - \frac{1}{2}\int_0^t c(s)^2 ds\right), \\ &= X_0 \exp\left(\int_0^t c(s)dW_s - \int_0^t c(s)\gamma_s + \frac{1}{2}c(s)^2 ds\right), \end{aligned} \tag{7.35}$$

which, once γ_t has been determined, will be a solution to the problem (7.33).

We still need to solve for γ_t in the equation $f(t, X_t) + c(t)X_t\gamma_t = 0$. An ODE is extracted using a clever change of variables. Put

$$Y_t = \exp\left(-\int_0^t c(s)\gamma_s ds\right),$$

then the equation we need to solve is

$$\begin{aligned} f(t, X_t) &+ c(t)X_t\gamma_t = 0, \\ f(t, F_t^{-1}Y_t) &+ c(t)F_t^{-1}Y_t\gamma_t = 0, \end{aligned}$$

where $F_t^{-1} = \exp\left(-\int_0^t c(s)dW_s + \frac{1}{2}\int_0^t c(s)^2 ds\right)$. So

$$Y_t = \frac{-F_t}{c(t)\gamma_t} f(t, F_t^{-1}Y_t), \tag{7.36}$$

but $\log(Y_t) = -\int_0^t c(s)\gamma_s ds$, which can be written $\frac{1}{Y_t}\frac{dY_t}{dt} = -c(t)\gamma_t$. Multiplying this identity on both sides of Equation (7.36) yields

$$\frac{dY_t}{dt} = F_t\, f(t, F_t^{-1}Y_t),$$

which is a deterministic differential equation for Y_t. Solving this equation will immediately allow us to write down the solution to (7.33).

Example 31. *Solve the SDE $dX_t = r\,dt + \alpha X_t\, dW_t$.*
Under the change of variables $W_t = \tilde{W}_t + \int_0^t \gamma_s ds$ the problem becomes that of solving

$$\begin{cases} dX_t &= \alpha X_t d\tilde{W}_t; \\ 0 &= r + \alpha X_t \gamma_t, \quad \forall t. \end{cases}$$

Now $dX_t = \alpha X_t d\tilde{W}_t$ implies

$$\begin{aligned} X_t &= \exp\left(\alpha\tilde{W}_t - \frac{1}{2}\alpha^2 t\right), \\ &= \exp\left(\alpha W_t - \frac{1}{2}\alpha^2 t\right)\exp\left(-\alpha\int_0^t \gamma_s ds\right), \\ &= G_t Y_t, \quad \textit{say.} \end{aligned}$$

If $Y_t = \exp\left(-\alpha\int_0^t \gamma_s ds\right)$, then $\frac{dY_t}{dt} = -\alpha\gamma_t Y_t$. Thus the problem $0 = r + \alpha X_t\gamma_t$ reduces to

$$\begin{aligned} r + \alpha G_t Y_t \gamma_t &= 0, \\ r &= G_t \frac{dY_t}{dt}. \end{aligned}$$

Thus

$$X_t = X_0 + r\exp\left(\alpha W_t - \frac{1}{2}\alpha^2 t\right)\int_0^t \exp\left(-\alpha W_t + \frac{1}{2}\alpha^2 t\right) ds.$$

7.3 Exercises

1. Let $\{W_t\}_{t\geq 0}$ denote a standard Brownian motion. Calculate the two integrals

   (a) $\lim_{mesh(\tau_n)\to 0} \sum_{j=0}^{n-1} W_{t_{j+1}}(W_{t_{j+1}} - W_{t_j})$,

   (b) $\int_0^T W_s \circ dW_s := \lim_{mesh(\tau_n)\to 0} \sum_{j=0}^{n-1} \frac{1}{2}(W_{t_{j+1}} + W_{t_j})(W_{t_{j+1}} - W_{t_j})$,

   This is the *Stratonovich integral* of W_t with respect to itself.

2. Apply Equation (7.14) to the function $f(x, y) = xy$ to obtain the product rule.

3. Using Itô's formula deduce that $W_t^2 - t$, $W_t^3 - 3tW_t$, and $W_t^4 - 6tW_t^2 + 3t^2$ are all martingales.

4. Find $f(x)$ such that $f(W_t + t)$ is a martingale.

5. Use Itô's formula to write down the stochastic differential equations for the following

   (a) $t^2 W_t$;

   (b) $\exp(\sigma W_t - \frac{1}{2}\sigma^2 t)$;

   (c) $f(t, X_t, Y_t)$.

6. Suppose S_T is lognormally distributed with $\log(S_T) \sim N(\mu T, \sigma^2 T)$. Use a change of measure to calculate $\mathbb{E}\left((S_T - K)^+\right)$ for constant K.

7. Solve $dX_t = \mu dt + (\gamma + \sigma X_t)dW_t$.

8. Solve the Langevin equation for the Ornstein-Uhlenbeck process.

9. The *Brownian bridge* from a to b follows the linear SDE

$$dY_t = \frac{b - Y_t}{1 - t}dt + dW_t; \quad 0 \leq t < 1, Y_0 = a.$$

Verify that

$$Y_t = a(1-t) + bt + (1-t)\int_0^t \frac{1}{1-s} dW_s; \quad 0 \le t < 1.$$

10. Solve $dX_t = X_t^3 dt + X_t^2 dW_t$.

11. Verify that $X_t = \sin(W_t)$ solves $dX_t = -\frac{1}{2}X_t dt + \sqrt{1-X_t^2}dW_t$, with $X_0 = 0$, for all times up until W leaves $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

12. Solve $dX_t = -\frac{1}{1+t}X_t dt + \frac{1}{1+t}dW_t$.

13. By using the Itô -Stratonovich transformation formula solve

$$dX_t = \frac{1}{2}n\int_0^t X_s^{2n-1} ds + \int_0^t X_s^n dW_s.$$

14. Redo Example 30 for the particular case $f(x) = x + 1$. Check that X_t is given by $-1 + e^{qt+W_t}$.

15. Check the answer quoted in Example 31 by solving the SDE using the method of Section 7.2.1.

16. Solve $dX_t = \frac{1}{X_t}dt + \alpha X_t dW_t$.

17. Apply the solution method of Section 7.2.4 to study the family

$$dX_t = X_t^\beta dt + \alpha X_t dW_t,$$

where α and β are constants.