Math 6810 (Probability)

Spring 2013

Lecture notes

Pieter Allaart
University of North Texas

January 30, 2013

Text: *Introduction to Stochastic Calculus with Applications*, by Fima C. Klebaner (3rd edition), Imperial College Press.

Other recommended reading: (Do not purchase these books before consulting with your instructor!)

1. *Real Analysis* by H. L. Royden (3rd edition), Prentice Hall.
2. *Probability and Measure* by Patrick Billingsley (3rd edition), Wiley.
3. *Probability with Martingales* by David Williams, Cambridge University Press.
4. *Stochastic Calculus for Finance I and II* by Steven E. Shreve, Springer.
5. *Brownian Motion and Stochastic Calculus* by Ioannis Karatzas and Steven E. Shreve, Springer. (Warning: this requires stamina, but is one of the few texts that is complete and mathematically rigorous)

Chapter 4

The stochastic integral

Our goal in this chapter is to define the *stochastic integral*

$$\int_0^T X(t)\, dW(t), \tag{4.1}$$

where $(W(t) : t \geq 0)$ is a Brownian motion and $(X(t) : 0 \leq t \leq T)$ a stochastic process adapted to the natural filtration $(\mathcal{F}_t)_t$ of $(W(t))_t$. First, we review the Riemann-Stieltjes integral.

Literature abbreviation: "Kleb" = "Klebaner"; "K&S" = "Karatzas and Shreve".

4.1 The Riemann-Stieltjes integral

Let f and g be real-valued functions on an interval $[a, b]$. For a partition $\mathcal{P} = \{a = t_0 < t_1 < \cdots < t_n = b\}$ and a choice of points $\xi_i \in [t_{i-1}, t_i]$, $i = 1, \ldots, n$, define the sum

$$S(\mathcal{P}, f, g) = \sum_{i=1}^{n} f(\xi_i)\big(g(t_i) - g(t_{i-1})\big). \tag{4.2}$$

Put $\|\mathcal{P}\| = \max_{1 \leq i \leq n}(t_i - t_{i-1})$. If $S(\mathcal{P}, f, g)$ converges to some finite limit L as $\|\mathcal{P}\| \to 0$ (regardless the precise choice of the ξ_i and the t_i), we call L the *Riemann-Stieltjes integral* of f with respect to g over $[a, b]$, and write

$$L = \lim_{\|\mathcal{P}\| \to 0} S(\mathcal{P}, f, g) = \int_a^b f(t)\, dg(t). \tag{4.3}$$

Note that the choice $g(t) = t$ gives the Riemann integral.

Theorem 4.1. *If f is continuous and g is of bounded variation, then the Riemann-Stieltjes integral $\int_a^b f(t)\, dg(t)$ exists.*

Unfortunately, if g is of not of bounded variation then $\int_a^b f(t)\,dg(t)$ need not exist in general; Theorem 1.14 of Klebaner makes this precise. In particular, the integral (4.1) is not defined as a Riemann-Stieltjes integral, since Brownian motion is not of bounded variation. In this chapter, we shall use a different approach to give meaning to expressions such as (4.1).

4.2 The Itô integral

From now on, $(W(t) : t \geq 0)$ represents a one-dimensional Brownian motion, and $(\mathcal{F}_t)_t$ the natural filtration of $W(t)$, unless otherwise specified. We first define the stochastic integral for a simple process $X(t)$.

4.2.1 Itô integral of a simple process

Fix $T > 0$, and let $(X(t) : 0 \leq t \leq T)$ be a left-continuous, adapted simple process. That is, $X(t)$ is of the form

$$X(t) = \xi_0 \,\mathrm{I}_{[0,t_1]}(t) + \sum_{i=1}^{n-1} \xi_i \,\mathrm{I}_{(t_i,t_{i+1}]}(t), \qquad 0 \leq t \leq T, \tag{4.4}$$

where $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$, and $\xi_0, \ldots, \xi_{n-1}$ are random variables such that $\mathrm{E}(\xi_i^2) < \infty$ and ξ_i is $\mathcal{F}_{t_i}$-measurable for $i = 1, \ldots, n$. In what follows, we shall always assume all adapted simple processes to be left-continuous without further mention.

Definition 4.2. For a simple, adapted process $X(t)$ as above, the *stochastic integral* (or *Itô integral*) $\int_0^T X(t)\,dW(t)$ is defined by

$$\int_0^T X(t)\,dW(t) = \sum_{i=1}^{n} \xi_i \big(W(t_{i+1}) - W(t_i)\big). \tag{4.5}$$

(Compare with the martingale transforms of Chapter 2! Recall that $W(t)$ is a martingale. Basically, a martingale transform is a "discrete stochastic integral".)

It is easy to see that the above definition of $\int_0^T X(t)\,dW(t)$ does not depend on the representation of $X(t)$ as a simple process.

Remark 4.3. The stochastic integral has the following interpretation in finance: suppose $W(t)$ represents the price of an asset (such as a share of stock) at time t. Think about $t_0, t_1, \ldots, t_{n-1}$ as *trading dates*, and about $X(t)$ as the position (number of shares) taken in the asset at time t. Naturally, the number of shares held between time t_i and t_{i+1} must be decided at time t_i, based only on the history of the asset price up to time t_i. Thus, ξ_i should be $\mathcal{F}_{t_i}$-measurable. Further, the stochastic integral $\int_0^T X(t)\,dW(t)$ defined above can be interpreted as the total net gain over the time interval $[0, T]$ from the trading strategy $X(t)$.

Note that if the ξ_i are non-random, then the stochastic integral $\int_0^T X(t)\,dW(t)$ is a Gaussian (= normally distributed) random variable with mean zero. If the ξ_i are random, $\int_0^T X(t)\,dW(t)$ is in general not Gaussian.

For an arbitrary subinterval $(a,b]$ of $[0,T]$, we can define

$$\int_a^b X(t)\,dW(t) := \int_0^T \mathrm{I}_{(a,b]}(t)X(t)\,dW(t), \tag{4.6}$$

which is well defined since if $X(t)$ is a simple adapted process, then so it $\mathrm{I}_{(a,b]}(t)X(t)$.

Theorem 4.4. *Let $X(t)$ and $Y(t)$ be adapted simple processes. The Itô integral has the following properties:*

(i) (Linearity) *If c and d are constants, then*

$$\int_0^T (cX(t)+dY(t))\,dW(t) = c\int_0^T X(t)\,dW(t) + d\int_0^T Y(t)\,dW(t).$$

(ii) (Zero mean property)

$$\mathrm{E}\int_0^T X(t)\,dW(t) = 0.$$

(iii) (Isometry)

$$\mathrm{E}\left(\int_0^T X(t)\,dW(t)\right)^2 = \int_0^T \mathrm{E}(X^2(t))\,dt. \tag{4.7}$$

Proof. See Kleb, p. 94. □

4.2.2 Itô integral of an adapted process

Now let $(X(t) : 0 \le t \le T$ be a measurable adapted process. We assume throughout that

$$\int_0^T \mathrm{E}(X^2(t))\,dt < \infty. \tag{4.8}$$

In other words, $X \in L^2([0,T]\times\Omega)$.

Proposition 4.5. *If $(X(t) : 0 \le t \le T)$ is measurable adapted satisfying* (4.8)*, then there is a sequence of adapted simple processes $X_n = (X_n(t) : 0 \le t \le T), n \in \mathbb{N}$ such that*

$$\lim_{n\to\infty}\int_0^T \mathrm{E}\big(X_n(t)-X(t)\big)^2\,dt = 0. \tag{4.9}$$

In other words, $X_n \to X$ in $L^2([0,T]\times\Omega)$.

Proof. If $X(t)$ is continuous, we can simply put

$$X_n(t) = X(0)\,\mathrm{I}_{\{0\}}(t) + \sum_{k=0}^{2^n-1} X(kT/2^n)\,\mathrm{I}_{(kT/2^n,(k+1)T/2^n]}(t), \qquad n \in \mathbb{N}, \tag{4.10}$$

and the BCT gives (4.9). For not necessarily continuous $X(t)$, a more sophisticated argument is needed; see K&S, Lemma 3.2.4 and Proposition 3.2.6. □

Proposition 4.6. *Let $\{X_n\}$ be a sequence of adapted simple processes satisfying* (4.9), *and let*

$$I_n := \int_0^T X_n(t)\,dW(t).$$

Then $\{I_n\}$ is a Cauchy sequence in $L^2(\Omega)$, and hence converges in L^2 to a limit I.

Proof. See Kleb, pp. 97-98. □

Definition 4.7. The limit I in Proposition 4.6 is called the *Itô integral* of $X(t)$ over $[0,T]$, denoted

$$I = \int_0^T X(t)\,dW(t).$$

Note that the Itô integral is not defined pathwise; that is, we do not have a definition of $\int_0^T X(t,\omega)\,dW(t,\omega)$ for each $\omega \in \Omega$. Instead, the Itô integral is defined as a limit in L^2 of a sequence of pathwise defined Itô integrals of simple processes.

Theorem 4.8. *Let $X(t)$ and $Y(t)$ be measurable adapted processes in $L^2([0,T]\times\Omega)$. The Itô integral has the following properties:*

(i) (Linearity) *If c and d are constants, then*

$$\int_0^T (cX(t) + dY(t))\,dW(t) = c\int_0^T X(t)\,dW(t) + d\int_0^T Y(t)\,dW(t).$$

(ii) (Zero mean property)

$$\mathrm{E}\int_0^T X(t)\,dW(t) = 0.$$

(iii) (Isometry)

$$\mathrm{E}\left(\int_0^T X(t)\,dW(t)\right)^2 = \int_0^T \mathrm{E}(X^2(t))\,dt. \tag{4.11}$$

Proof. Exercise. (Hint: use the corresponding properties for simple processes.) □

Remark 4.9. The stochastic integral does not possess the monotonicity property: $X(t) \le Y(t)$ does not imply $\int_0^T X(t)\,dW(t) \le \int_0^T Y(t)\,dW(t)$. (For example: $\int_0^T 1\,dW(t) = W(T) - W(0) = W(T)$, which is negative with probability 1/2.)

Example 4.10 (Kleb, Example 4.2)**.** Using the quadratic variation of Brownian motion and a sequence of partitions of $[0, T]$ it can be calculated directly that

$$\int_0^T W(t)\,dW(t) = \frac{1}{2}W^2(T) - \frac{1}{2}T.$$

Recall that for a continuously differentiable function $f : [0, T] \to \mathbb{R}$ with $f(0) = 0$,

$$\int_0^T f(t)\,df(t) = \int_0^T f(t)f'(t)\,dt = \frac{1}{2}(f^2(T) - f^2(0)) = \frac{1}{2}f^2(T).$$

Compared to this, the Itô integral $\int_0^T W(t)\,dW(t)$ has an extra term. This is a first indication that the Itô integral does not behave like integrals in ordinary calculus. We will need to develop an alternative to the chain rule (Itô's rule), in order to correctly compute stochastic integrals.

Example 4.11. Kleb, p. 99 gives examples of processes $X(t)$ which do or do not satisfy the condition (4.8). Note that an adapted process $X(t)$ does not have to be a function of $W(t)$, but may depend on the entire path $W(s) : 0 \le s \le t$. However, $X(t) = f(W(t))$ gives the simplest examples.

Theorem 4.12 (Product of two Itô integrals)**.** *Let $X(t)$ and $Y(t)$ be measurable adapted processes in $L^2([0, T] \times \Omega)$. Then*

$$\mathrm{E}\left(\int_0^T X(t)\,dW(t) \int_0^T Y(t)\,dW(t)\right) = \int_0^T \mathrm{E}\left(X(t)Y(t)\right)dt.$$

Proof. Let $I_1 = \int_0^T X(t)\,dW(t)$ and $I_2 = \int_0^T Y(t)\,dW(t)$. Now write

$$I_1 I_2 = \frac{1}{2}\left[(I_1 + I_2)^2 - I_1^2 - I_2^2\right],$$

and use the isometry property (4.11) three times. □

4.2.3 The Itô integral process

We will now treat the upper limit of the stochastic integral as variable, and consider properties of the stochastic process

$$I(t) := \int_0^t X(u)\,dW(u), \tag{4.12}$$

for a measurable adapted process $X(t)$ satisfying (4.8).

Theorem 4.13. *The process $(I(t) : 0 \le t \le T)$ is a continuous, square-integrable martingale.*

Proof. We already know that $I(t)$ is square-integrable. To prove the other two properties, suppose first that $X(t)$ is a simple process, defined by (4.4). Then continuity of $I(t)$ follows directly from continuity of $W(t)$. To show that $I(t)$ is a martingale, we must show that

$$\mathrm{E}\left[\int_s^t X(u)\,dW(u)|\mathcal{F}_s\right] = 0 \quad \text{a.s., for all } 0 \le s < t \le T. \tag{4.13}$$

By adding extra partition points if necessary, we may assume that $s = t_j$ and $t = t_k$ for some $j < k$. Then

$$\int_s^t X(u)\,dW(u) = \sum_{i=j}^{k-1} \xi_i \big(W(t_{i+1} - W(t_i)\big),$$

so that

$$\mathrm{E}\left[\int_s^t X(u)\,dW(u)|\mathcal{F}_s\right] = \sum_{i=j}^{k-1} \mathrm{E}\left[\xi_i\big(W(t_{i+1} - W(t_i)\big)|\mathcal{F}_{t_j}\right].$$

But by the tower law, for each $i \ge j$,

$$\begin{aligned}\mathrm{E}\left[\xi_i\big(W(t_{i+1} - W(t_i)\big)|\mathcal{F}_{t_j}\right] &= \mathrm{E}\left[\mathrm{E}\left[\xi_i\big(W(t_{i+1} - W(t_i)\big)|\mathcal{F}_{t_i}\right]|\mathcal{F}_{t_j}\right] \\ &= \mathrm{E}\left[\xi_i\,\mathrm{E}\left(W(t_{i+1}) - W(t_i)\right)|\mathcal{F}_{t_j}\right] \\ &= 0,\end{aligned}$$

where the second equality follows since ξ_i is $\mathcal{F}_{t_i}$-measurable and $W(t_{i+1}) - W(t_i)$ is independent of $\mathcal{F}_{t_i}$; and the last equality follows because $\mathrm{E}\big(W(t_{i+1}) - W(t_i)\big) = 0$. Thus, we have (4.13).

Now let $X(t)$ be an arbitrary measurable adapted process. Let $X_n(t)$, $n \in \mathbb{N}$ be a sequence of adapted simple processes converging in mean square to $X(t)$. To check (4.13), we must show that

$$\mathrm{E}\left[\left(\int_s^t X(u)\,dW(u)\right)\mathrm{I}_A\right] = 0, \qquad \forall\, A \in \mathcal{F}_s. \tag{4.14}$$

Since X_n satisfies (4.13), (4.14) holds with X_n in place of X. Now we use the Schwartz inequality to write

$$\begin{aligned}&\left|\mathrm{E}\left[\left(\int_s^t X(u)\,dW(u)\right)\mathrm{I}_A\right] - \mathrm{E}\left[\left(\int_s^t X_n(u)\,dW(u)\right)\mathrm{I}_A\right]\right| \\ &\quad= \left|\mathrm{E}\left[\left(\int_s^t \{X(u) - X_n(u)\}\,dW(u)\right)\mathrm{I}_A\right]\right| \\ &\quad\le \left(\mathrm{E}\,\mathrm{I}_A^2\right)^{1/2}\left[\mathrm{E}\left(\int_s^t \{X(u) - X_n(u)\}\,dW(u)\right)^2\right]^{1/2}.\end{aligned}$$

This last expression tends to 0, since the first factor is bounded by 1, and by the Itô isometry

$$\mathrm{E}\left(\int_s^t \{X(u) - X_n(u)\}\, dW(u)\right)^2 = \int_s^t \mathrm{E}\left(X(u) - X_n(u)\right)^2 du \to 0.$$

From the above development, we obtain (4.14). Thus, $I(t)$ is a martingale.

Again by Itô isometry,

$$\mathrm{E}\left(\int_s^t X(u)\, dW(u)\right)^2 = \int_s^t \mathrm{E}\left(X^2(u)\right) du,$$

from which we get, by DCT (bound the last integral by the finite number $\int_0^T \mathrm{E}\left(X^2(u)\right) du$),

$$\lim_{t\to t_0} \mathrm{E}\left(I(t) - I(t_0)\right)^2 = 0$$

for each t_0. We say $I(t)$ is *continuous in mean square*. From our definition of the stochastic integral it does not follow logically that $I(t)$ is pathwise continuous. But Proposition 4.6 and the definition that follows can be modified to guarantee this. It can be shown that the space $\mathcal{M}_2^c$ of continuous, square-integrable martingales on $[0, T]$ is a complete metric space with the L^2 norm

$$\|Y\| := \sqrt{\mathrm{E}(Y_T^2)}.$$

(We omit the technical details here! But see K&S, Proposition 1.5.23.) If we put $I_n(t) = \int_0^T X_n(t)\, dW(t)$, then the processes $I_n(t)$ are in $\mathcal{M}_2^c$, and they form a Cauchy sequence in that space by the argument of the proof of Proposition 4.6. Thus, the limit $I(t)$, now a stochastic process on $[0, T]$, lies in $\mathcal{M}_2^c$ as well. □

We next consider the quadratic variation of the process $I(t)$. Before stating the theorem we prove a small lemma.

Lemma 4.14. *Let $\{Y_n\}$ be a sequence of r.v.'s that converges in L^1 to a r.v. Y, and let Z be a nonnegative r.v. with finite expectation, which is independent of Y_n and Y. Then $ZY_n \to_P ZY$.*

Proof. We have, for any $\varepsilon > 0$,

$$\mathrm{P}(|ZY_n - ZY| \geq \varepsilon) \leq \varepsilon^{-1}\,\mathrm{E}\,|ZY_n - ZY| = \varepsilon^{-1}E(Z)E|Y_n - Y| \to 0.$$

□

Theorem 4.15 (Quadratic variation of the Itô integral)**.** *The quadratic variation of the Itô integral $\int_0^t X(u)\, dW(u)$ is*

$$[I, I](t) = \int_0^t X^2(u)\, du, \tag{4.15}$$

where $[I, I](t)$ *is understood as the limit in probability of*

$$\sum_{i=1}^{n} [I(t_i) - I(t_{i-1})]^2$$

for partitions $0 = t_0 < t_1 < \cdots < t_n = t$, *as* $\max_i(t_i - t_{i-1}) \to 0$.

Proof. Let $X(t)$ be a simple process as in (4.4). Consider an interval $[t_j, t_{j+1}]$, and partition it by

$$t_j = s_0 < s_1 < \cdots < s_m = t_{j+1}.$$

Then

$$\sum_{i=0}^{m-1} [I(s_{i+1}) - I(s_i)]^2 = \xi_j^2 \sum_{i=0}^{m-1} \big(W(s_{i+1}) - W(s_i)\big)^2. \tag{4.16}$$

As $m \to \infty$ and $\max_i(s_{i+1} - s_i) \to 0$, the summation on the right converges in L^2 (and hence in L^1) to the quadratic variation of Brownian motion over the interval $[t_j, t_{j+1}]$, which is $t_{j+1} - t_j$. Since $\mathrm{E}(\xi_j^2) < \infty$, Lemma 4.14 implies that

$$\sum_{i=0}^{m-1} [I(s_{i+1}) - I(s_i)]^2 \to_P \xi_j^2 (t_{j+1} - t_j) = \int_{t_j}^{t_{j+1}} X^2(u)\, du.$$

Summing over j gives the desired result.

If $X(t)$ is a general adapted process, it is difficult to prove (4.15) directly. (Approximation by simple processes doesn't seem to work very well.) Instead we use the characterization (stated somewhat imprecisely without proof, but see K&S, Section 1.5) of the quadratic variation of a square integrable martingale $M(t)$ as the unique increasing process $A(t)$, null at zero, such that $M^2(t) - A(t)$ is a martingale. Thus, (4.15) will follow once we check that

$$I^2(t) - \int_0^t X^2(u)\, du, \qquad 0 \le t \le T$$

is a martingale. It's not hard to prove the following conditional version of the isometry property:

$$\mathrm{E}\left[\big(I(t) - I(s)\big)^2 | \mathcal{F}_s\right] = \mathrm{E}\left[\int_s^t X^2(u)\, du | \mathcal{F}_s\right].$$

(The argument is similar to that in the proof of Theorem 4.13: First prove it for a simple process $X(t)$, then use an approximation argument.) By the usual trick (remember?),

$$\mathrm{E}\left[I^2(t) - I^2(s) | \mathcal{F}_s\right] = \mathrm{E}\left[\big(I(t) - I(s)\big)^2 | \mathcal{F}_s\right].$$

From the last two equations, the martingale property of $I^2(t) - \int_0^t X^2(u)\, du$ follows immediately, and from it, we get (4.15). □

Remark 4.16. The relationship

$$I(t) = \int_0^t X(u)\,dW(u) \tag{4.17}$$

is often written in differential notation as

$$dI(t) = X(t)dW(t). \tag{4.18}$$

Note, however, that the differential $dW(t)$ in itself does not carry any meaning, and (4.18) only makes sense as an alternative representation of (4.17). The advantage of differential notation is that it is both more compact and more heuristic.

Remark 4.17. The definition of the Itô integral $\int_0^T X(t)\,dW(t)$ can be extended to processes $X(t)$ which, instead of the condition (4.8), merely satisfy

$$\int_0^T X^2(t)\,dt < \infty \qquad \text{a.s.} \tag{4.19}$$

The details can be found in K&S, pp. 146-147. In a nutshell, $X(t)$ can be approximated by simple processes which converge to $X(t)$ in probability (rather than L^2), and the resulting Itô integrals converge in probability to a finite limit, which we call the Itô integral of $X(t)$. In this setting, however, the Itô integral typically fails to have an expectation and variance. Hence, the Itô integral process is not a martingale and does not satisfy the isometry property (4.11). It does, however, have quadratic variation given by (4.15).

Note that (4.19) is automatically satisfied if the integrand $X(t)$ is continuous in t, since a continuous function on a compact interval is bounded. Thus, for any continuous function f, the integral

$$\int_0^T f(W(t))\,dW(t)$$

is well defined.

4.3 Itô's formula

Just like the Fundamental Theorem of Calculus provides a mechanical way for computing ordinary definite integrals, Itô's formula does that (to some extent) for stochastic integrals. It will also give rise to an integration by parts formula for Itô integrals. But it has another, even more important use: Because the Itô integral process is a martingale, Itô's formula tells us how functions of Brownian motion (or later, of more general processes) must be compensated in order to obtain a martingale. This basic idea leads to the so-called "equivalent martingale measure" or "risk-neutral measure" in the theory of option pricing, that we will develop in a later chapter.

Theorem 4.18 (Itô's formula for Brownian motion)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be twice continuously differentiable. Then for any $t \leq T$,*

$$f(W(t)) = f(0) + \int_0^t f'(W(s))\,dW(s) + \frac{1}{2}\int_0^t f''(W(s))\,ds. \tag{4.20}$$

In differential form:

$$df(W(t)) = f'(W(t))dW(t) + \frac{1}{2}f''(W(t))dt. \tag{4.21}$$

Proof. See Kleb, Theorem 4.13 and Theorem 4.14. Note that the stochastic integral in (4.20) is well defined in view of Remark 4.17. □

The most remarkable feature of the formula (4.20) is the last term. If $u : [0, T] \to \mathbb{R}$ is differentiable, then we have

$$f(u(t)) = f(u(0)) + \int_0^t f(u(s))\,du(s).$$

The extra term in Itô's formula is caused by the nondifferentiability of $W(t)$, and it is what sets stochastic calculus apart from ordinary differential calculus. The "dt" at the end derives from the quadratic variation of Brownian motion.

Now let f be a continuously differentiable function, and F an antiderivative of f. Then (4.20) gives

$$\int_0^t f(W(s))\,dW(s) = F(W(t)) - F(0) - \frac{1}{2}\int_0^t f'(W(s))\,ds. \tag{4.22}$$

Applying this to $f(x) = x$ gives

$$\int_0^T W(s)\,dW(s) = \frac{1}{2}W^2(t) - \frac{1}{2}t,$$

which we derived earlier directly from the definition of the Itô integral. (**Exercise:** Write this in differential form.)

Example 4.19. Let $f(x) = e^x$. Then (4.20) gives

$$e^{W(t)} = 1 + \int_0^t e^{W(s)}\,dW(s) + \frac{1}{2}\int_0^t e^{W(s)}\,ds. \tag{4.23}$$

Since $\int_0^t \mathrm{E}\left(e^{2W(s)}\right)ds < \infty$, this shows that the process

$$e^{W(t)} - \frac{1}{2}\int_0^t e^{W(s)}\,ds, \qquad 0 \leq t \leq T$$

is a martingale with mean 1.

In differential form, (4.23) is

$$de^{W(t)} = e^{W(t)}dW(t) + \frac{1}{2}e^{W(t)}dt, \tag{4.24}$$

or, with $Y(t) = e^{W(t)}$,

$$dY(t) = Y(t)dW(t) + \frac{1}{2}Y(t)dt. \tag{4.25}$$

This last equation is called a *stochastic differential equation* (SDE), with $Y(t)$ considered as the unknown process. More on SDEs later.

4.4 Interlude: quadratic covariation

Definition 4.20. The *quadratic covariation*, or *cross variation*, of two functions f and g on $[0, t]$ is defined by

$$[f, g](t) := \lim_{\delta \to 0} \sum_{i=0}^{n-1} \big(f(t_{i+1}) - f(t_i)\big)\big(g(t_{i+1}) - g(t_i)\big) \tag{4.26}$$

for partitions $0 = t_0 < t_1 < \cdots < t_n = t$, where $\delta = \max_i(t_{i+1} - t_i)$, provided the limit exists.

Exercise 4.21. a) Show that $[\cdot, \cdot]$ is bilinear; that is,

$$[a_1f_1 + a_2f_2, b_1g_1 + b_2g_2] = a_1b_1[f_1, g_1] + a_1b_2[f_1, g_2] + a_2b_1[f_2, g_1] + a_2b_2[f_2, g_2].$$

b) Verify the *polarization identity*:

$$[f, g](t) = \frac{1}{2}\big([f + g, f + g](t) - [f, f](t) - [g, g](t)\big).$$

(Hint: start with $[f + g, f + g](t)$, and expand.)

Proposition 4.22. *If f is continuous and g is of bounded variation, then* $[f, g](t) = 0$.

Proof. The proof is essentially the same as that for the quadratic variation of a continuous function of bounded variation; see Kleb, Theorem 1.10. □

For stochastic processes $X(t)$ and $Y(t)$, $[X, Y](t)$ is defined by (4.26), where the limit is understood as limit in probability.

Now let $X_1(t)$ and $X_2(t)$ be measurable, adapted processes and $I_i(t) = \int_0^t X_i(u)\, dW(u)$, $i = 1, 2$. Then one easily calculates using the polarization identity and (4.15),

$$[I_1, I_2](t) = \int_0^t X_1(u)X_2(u)\, du. \tag{4.27}$$

It can be shown that, if $M(t)$ and $N(t)$ are continuous square-integrable martingales, then $[M,N](t)$ is the unique adapted process $A(t)$ of bounded variation such that $M(t)N(t) - A(t)$ is martingale. Thus, for the two stochastic integrals above,

$$I_1(t)I_2(t) - \int_0^t X_1(u)X_2(u)\,du, \qquad 0 \le t \le T$$

is a martingale. We shall prove this fact later using a two-dimensional Itô formula.

4.5 Itô processes

Definition 4.23. An *Itô process* is a process of the form

$$Y(t) = Y(0) + \int_0^t \mu(s)\,ds + \int_0^t \sigma(s)\,dW(s), \qquad 0 \le t \le T, \tag{4.28}$$

where $\mu(t)$ and $\sigma(t)$ are adapted processes such that

$$\int_0^T |\mu(t)|\,dt < \infty \qquad \text{and} \qquad \int_0^T \sigma^2(t)\,dt < \infty \qquad \text{a.s.} \tag{4.29}$$

In stochastic differential form,

$$dY(t) = \mu(t)dt + \sigma(t)dW(t). \tag{4.30}$$

(Of course, (4.30) has meaning only as a rewriting of (4.28).) The process $Y(t)$ is a special case of a *diffusion process*, and $\sigma(t)$ is called the *diffusion coefficient* of $Y(t)$.

Theorem 4.24 (Quadratic variation of an Itô process)**.** *If $Y(t)$ is an Itô process as above, then the quadratic variation of $Y(t)$ exists and equals*

$$[Y,Y](t) = \int_0^t \sigma^2(s)\,ds. \tag{4.31}$$

Proof. Write $Y(t) = Y(0) + B(t) + I(t)$, where $B(t) = \int_0^t \mu(s)\,ds$ and $I(t) = \int_0^t \sigma(s)\,dW(s)$. Observe that $B(t)$ is absolutely continuous and hence of bounded variation. Thus, by Proposition 4.22, $[B,I](t) = 0$ and likewise, $[B,B](t) = 0$. The result therefore follows from (4.15) and bilinearity of $[\cdot,\cdot]$. □

Remark 4.25. The process $B(t)$ in the above proof has zero quadratic variation, but that does not mean it is nonrandom. However, $B(t)$ is less volatile than $I(t)$. At time t we can not with certainty predict the value of $B(t+\Delta t)$, but we have a good estimate:

$$B(t+\Delta t) \approx B(t) + \Delta\mu(t).$$

In finance applications, $B(t)$ is like the value of a money market account at a variable interest rate. It is random, but not very volatile. By contrast, an estimate for $I(t+\Delta t)$ is

$$I(t+\Delta t) \approx I(t) + \sigma(t)\big(W(t+\Delta t) - W(t)\big).$$

Since $W(t+\Delta t) - W(t)$ is independent of the information available at time t, there is more uncertainty about the value of $I(t+\Delta t)$. Thus, $I(t)$ is more like the value of a stock portfolio. It has more volatility.

A trivial modification of the proof of Theorem 4.24 gives:

Theorem 4.26 (Covariation of two Itô processes)**.** *Let $X_1(t)$ and $X_2(t)$ be Itô processes driven by the* same *Brownian motion $W(t)$. That is,*

$$X_i(t) = X_i(0) + \int_0^t \mu_i(s)\,ds + \int_0^t \sigma_i(s)\,dW(s), \qquad 0 \le t \le T, \qquad i = 1, 2.$$

Then

$$[X_1, X_2](t) = \int_0^t \sigma_1(s)\sigma_2(s)\,ds. \tag{4.32}$$

Exercise 4.27. Let f be a C^2 function. Use Itô's formula to write $f(W(t))$ as an Itô process, and compute $[f(W), W]$ and $[f(W), f(W)]$.

For applications in finance, it is necessary to define the stochastic integral with respect to an Itô process.

Definition 4.28 (Stochastic integral with respect to an Itô process)**.** Let $Y(t)$ be an Itô process given by (4.30). Let $H(t)$ be an adapted process satisfying

$$\int_0^t H^2(s)\sigma^2(s)\,ds < \infty \qquad \text{and} \qquad \int_0^t |H(s)\mu(s)|\,ds < \infty \qquad \text{a.s.}$$

Then we define

$$\int_0^t H(s)\,dY(s) := \int_0^t H(s)\mu(s)\,ds + \int_0^t H(s)\sigma(s)\,dW(s). \tag{4.33}$$

(This definition is suggested by the differential form (4.30): multiply both sides by $H(t)$, then formally integrate both sides.) In the above integral, we call $H(t)$ the *integrand* and $Y(t)$ the *integrator*. Itô processes provide a fairly large class of integrators for stochastic calculus – large enough for our purposes. But in more advanced stochastic calculus one considers an even larger class of integrators, called *semimartingales*, which also include processes with jumps such as the Poisson process. Defining stochastic integrals with respect to semimartingales is a considerable technical undertaking which we shall avoid.

Theorem 4.29 (Itô's formula for Itô processes)**.** *Let $X(t)$ be an Itô process with stochastic differential*

$$dX(t) = \mu(t)dt + \sigma(t)dW(t),$$

and let f be a C^2 function. Then

$$\begin{aligned} f(X(t)) &= f(X(0)) + \int_0^t f'(X(s))\,dX(s) + \frac{1}{2}\int_0^t f''(X(s))\sigma^2(s)\,ds \\ &= f(X(0)) + \int_0^t \left[f'(X(s))\mu(s) + \frac{1}{2}f''(X(s))\sigma^2(s)\right] ds + \int_0^t f'(X(s))\sigma(s)\,dW(s). \end{aligned}$$

In differential notation,

$$df(X(t)) = \left(f'(X(t))\mu(t) + \frac{1}{2}f''X(t)\sigma^2(t)\right) dt + f'(X(t))\sigma(t)dW(t). \tag{4.34}$$

(The proof is similar to that of Theorem 4.18, and is omitted.)

Example 4.30 (A simple SDE)**.** We find a positive process $X(t)$ satisfying the SDE

$$dX(t) = X(t)dW(t) + \frac{1}{2}X(t)dt. \tag{4.35}$$

Applying Itô's formula with $f(x) = \log x$, we calculate formally,

$$d\log X(t) = dW(t).$$

(Check!!) Thus, $\log X(t) = \log X(0) + W(t)$ and $X(t) = X(0)e^{W(t)}$. Since computation with stochastic differentials is not rigorous, we check this solution by applying Itô's formula (the one for Brownian motion) with $f(x) = e^x$. (Do this!)

Exercise 4.31. Klebaner, Exercise 4.8 (p. 121)

Remark 4.32. (*Note on manipulation with stochastic differentials.*) It is often convenient to use the heuristic notation

$$dX(t)dY(t) = d[X,Y](t). \tag{4.36}$$

In particular, we then have

$$(dX(t))^2 = d[X,X](t).$$

These equations should be taken as a definition of the left-hand sides. The right-hand sides have meaning through their integral representations. We can now formally manipulate with iterated differentials by using the 2×2 table:

$$(dW(t))^2 = dt, \qquad dW(t)dt = 0,$$

$$dt\,dW(t) = 0, \qquad (dt)^2 = 0.$$

The first expression comes of course from $(dW(t))^2 = d[W,W](t) = dt$.

4.6 Itô' formula for functions of two variables

Let $f(x, y)$ be a function with continuous partial derivatives f_x, f_y, f_{xx}, f_{xy} and f_{yy} up to order 2. The multivariable Taylor formula is

$$\begin{aligned} f(x,y) = f(x_0,y_0) &+ f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) \\ &+ \frac{1}{2}\left(f_{xx}(x_0,y_0)(x-x_0)^2 + f_{yy}(x_0,y_0)(y-y_0)^2\right) + f_{xy}(x_0,y_0)(x-x_0)(y-y_0) \\ &+ \text{ higher order terms} \end{aligned}$$

Using this one can prove, in a manner similar to the proof of Theorem 4.18:

Theorem 4.33 (Two-dimensional Itô formula)**.** *Let $f(x, y)$ be a function with continuous partial derivatives up to order 2. Let X and Y be Itô processes given by*

$$\begin{aligned} dX(t) &= \mu_X(t)dt + \sigma_X(t)dW(t), \\ dY(t) &= \mu_Y(t)dt + \sigma_Y(t)dW(t). \end{aligned} \tag{4.37}$$

Then

$$\begin{aligned} df(X(t),Y(t)) &= f_x(X(t),Y(t))dX(t) + f_y(X(t),Y(t))dY(t) \\ &\quad + \frac{1}{2}f_{xx}(X(t),Y(t))d[X,X](t) + \frac{1}{2}f_{yy}(X(t),Y(t))d[Y,Y](t) \\ &\quad + f_{xy}(X(t),Y(t))d[X,Y](t) \\ &= f_x(X(t),Y(t))dX(t) + f_y(X(t),Y(t))dY(t) \\ &\quad + \frac{1}{2}f_{xx}(X(t),Y(t))\sigma_X^2(t)dt + \frac{1}{2}f_{yy}(X(t),Y(t))\sigma_Y^2(t)dt \\ &\quad + f_{xy}(X(t),Y(t))\sigma_X(t)\sigma_Y(t)dt. \end{aligned}$$

(Exercise: write this in integral form!)

(Again, we use differential notation to keep the formula more compact, but it is the integral form that has a clearly defined meaning.)

One of the main applications of the two-dimensional Itô formula is to derive a product rule and integration by parts formula for stochastic calculus. To do this, take $f(x, y) = xy$. Then $f_x = y$, $f_y = x$, $f_{xx} = f_{yy} = 0$ and $f_{xy} = 1$. Thus, we obtain the **Itô product rule**

$$d(X(t)Y(t)) = X(t)dY(t) + Y(t)dX(t) + d[X,Y](t).$$

Writing this in integral form and rearranging gives the **integration by parts formula**

$$\int_0^t X(s)\,dY(s) = X(t)Y(t) - X(0)Y(0) - \int_0^t Y(s)\,dX(s) - [X,Y](t).$$

Compared to the usual product rule and integration by parts formula, there is an extra term given by the quadratic covariation of X and Y. Note that if X and Y were differentiable, their covariation would be zero and we'd get the usual formulas back.

Exercise 4.34. Let $X(t) = \sin(W(t))$. Compute $d(X(t)W(t))$ two ways:

a) using the one-dimensional Itô formula

b) by writing $X(t)$ as an Itô process and using the Itô product rule.

Check that the answers match.

The two-dimensional Itô formula has another important special case:

Corollary 4.35. *Let $f(x,t)$ be twice continuously differentiable in x and continuously differentiable in t. Let X be an Itô process given by* (4.37)*. Then*

$$df(X(t),t) = f_x(X(t),t)dX(t) + f_t(X(t),t)dt + \frac{1}{2}\sigma_X^2(t)f_{xx}(X(t),t)dt. \tag{4.38}$$

(Exercise: write this in integral form.)

Proof. This follows from Theorem 4.33 by putting $Y(t) = t$ (why is this an Itô process?), and noting that in this case, $[X,Y](t) = [Y,Y](t) = 0$. □

Example 4.36. Let $X(t) = \exp\{\sigma W(t) - \sigma^2 t/2\}$. Taking $f(x,t) = \exp(\sigma x - \sigma^2 t/2)$, we obtain

$$\begin{aligned} dX(t) &= f_x(W(t),t)dW(t) + f_t(W(t),t)dt + \frac{1}{2}f_{xx}(W(t),t)dt \\ &= \sigma f(W(t),t)dW(t) - \frac{1}{2}\sigma^2 f(W(t),t)dt + \frac{1}{2}\sigma^2 f(W(t),t)dt \\ &= \sigma X(t)dW(t). \end{aligned}$$

In integral form:

$$X(t) = X(0) + \int_0^t \sigma X(s)\, dW(s).$$

This shows again that $X(t)$ is a (square-integrable) martingale, something we knew from Chapter 3!

Exercise 4.37. Show that $M(t) = e^{t/2}\sin(W(t))$ is a martingale.

4.7 Multidimensional Itô processes

See Kleb, Section 4.7. Also note (and check!) Remark 4.9 on correlated Brownian motions.

4.8 Exact solution of linear SDEs

A *stochastic differential equation* (SDE) is an equation of the form

$$dX(t) = \mu(X(t),t)dt + \sigma(X(t),t)dW(t). \tag{4.39}$$

Here the solution $X(t)$ is assumed to be an adapted process, and $\mu(x,t)$ and $\sigma(x,t)$ are real-valued (nonrandom) functions. By a *solution* of (4.39) we mean here what is called a *strong solution* in the literature, that is, a process $X(t)$ defined on the same probability space as $W(t)$ and adapted to the natural filtration of $W(t)$, such that the integrals below are well defined, and

$$X(t) = X(0) + \int_0^t \mu(X(s), s)ds + \int_0^t \sigma(X(s), s)dW(s).$$

SDEs that don't have strong solutions may have *weak solutions*, a concept we will not go into for the moment. As with ordinary differential equations (ODEs), there is substantial theory about existence and uniqueness of solutions of SDEs. But one case can be solved explicitly, namely the linear SDE

$$dX(t) = \big(\alpha(t) + \beta(t)X(t)\big)dt + \big(\gamma(t) + \delta(t)X(t)\big)dW(t). \tag{4.40}$$

In this section we will see how. Note first that in the most basic case,

$$dX(t) = \big(\alpha(t) + \beta(t)X(t)\big)dt + \sigma dW(t),$$

the linear SDE is simply a linear ODE perturbed by noise. In that case we would expect the solution $X(t)$ to look globally like the solution of the corresponding ODE, but with locally irregular behavior. One simple example of this is the *Langevin equation*

$$dX(t) = -\alpha X(t)dt + \sigma dW(t),$$

whose solution is the *Ornstein-Uhlenbeck process*, which models the motion of a particle suspended in liquid in the presence of friction. (We will see another example in the next section as a model for interest rates.)

4.8.1 Review of linear ODEs

For comparison, we first recall the solution of the first order linear ODE

$$y'(t) = a(t) + b(t)y(t). \tag{4.41}$$

First find an *integrating factor* $\mu(t)$ satisfying

$$\mu'(t) = -\mu(t)b(t). \tag{4.42}$$

This last differential equation can be solved by separating variables:

$$\frac{\mu'(t)}{\mu(t)} = -b(t) \qquad \Rightarrow \qquad \log\mu(t) = -\int b(t)dt \qquad \Rightarrow \qquad \mu(t) = \exp\left\{-\int_0^t b(s)ds\right\},$$

where we have chosen the solution with $\mu(0) = 1$. Now multiply both sides of (4.41) by $\mu(t)$ and rearrange to get, using (4.42) and the product rule,

$$(\mu(t)y(t))' = \mu(t)a(t).$$

Integrating both sides gives the general solution

$$y(t) = \frac{1}{\mu(t)} \left[y(0) + \int_0^t \mu(s)a(s)ds \right].$$

4.8.2 The stochastic exponential

Before solving the linear SDE in full generality, we first consider a special case.

Definition 4.38. Let $X(t)$ be an Itô process given by

$$dX(t) = \mu(t)dt + \sigma(t)dW(t). \tag{4.43}$$

The *stochastic exponential* of X, denoted $\mathcal{E}(X)$, is a process $U(t)$ satisfying

$$dU(t) = U(t)dX(t), \qquad U(0) = 1, \tag{4.44}$$

or in integral form,

$$U(t) = 1 + \int_0^t U(s)dX(s).$$

If X is of bounded variation, we simply have $U(t) = e^{X(t)}$. For a general Itô process, we have:

Theorem 4.39. *Let $X(t)$ satisfy* (4.43)*. Then* (4.44) *has a unique solution given by*

$$\begin{aligned} U(t) = \mathcal{E}(X)(t) &= \exp\left\{ X(t) - X(0) - \frac{1}{2}[X,X](t) \right\} \qquad (4.45) \\ &= \exp\left\{ \int_0^t \mu(s)ds + \int_0^t \sigma(s)dW(s) - \frac{1}{2}\int_0^t \sigma^2(s)ds \right\}. \end{aligned}$$

Proof. For a proof that the right hand side of (4.45) satisfies (4.44), see Kleb, Theorem 5.2 (p.129).

For uniqueness, suppose there is another solution $U_1(t)$. That is,

$$dU_1(t) = U_1(t)dX(t), \qquad U_1(0) = 1.$$

By formal manipulation with differentials,

$$d[U,U](t) = (dU(t))^2 = U^2(t)(dX(t))^2 = U^2(t)d[X,X](t) = U^2(t)\sigma^2(t)dt$$

and

$$d[U_1,U](t) = dU_1(t)dU(t) = U_1(t)U(t)(dX(t))^2 = U_1(t)U(t)\sigma^2(t)dt.$$

Now use the two-variable Itô rule (Theorem 4.33) with $f(x,y) = x/y$ to show that

$$d\left(\frac{U_1(t)}{U(t)}\right) = 0,$$

so $U_1(t)/U(t)$ is constant. Since $U_1(0) = U(0) = 1$, this means $U_1(t) = U(t)$ and the solution is unique. □

4.8.3 Solution of the general linear SDE

See Kleb, Section 5.3 (pp. 131-132).

4.8.4 Example: Brownian bridge

See Kleb, pp. 133-134. (Hint for checking the last equation in Example 5.11: use l'Hopital's rule!)

4.9 Examples: Financial models

4.9.1 Geometric Brownian motion as a stock price model

It has long been known that Brownian motion has the same kind of irregular behavior as stock price movements. Bachelier (1900) already proposed using Brownian motion as a model for stock prices. The main disadvantage, of course, is that Brownian motion can become negative, whereas real stock prices can't. A better model is the *geometric Brownian motion*

$$S(t) = S(0)\exp\{\sigma W(t) + \mu t\},$$

where $\sigma > 0$ and μ are real constants. We have already seen a special case of geometric Brownian motion in Example 4.36, where it was shown that for the choice $\mu = -\sigma^2/2$, $S(t)$ is a martingale.

More generally, let $\sigma(t)$ and $\alpha(t)$ be adapted processes, and define the Itô process

$$X(t) = \int_0^t \sigma(s)dW(s) + \int_0^t \left(\alpha(s) - \frac{1}{2}\sigma^2(s)\right) ds.$$

In differential notation:

$$dX(t) = \sigma(t)dW(t) + \left(\alpha(t) - \frac{1}{2}\sigma^2(t)\right) dt.$$

Basically, $X(t)$ is a Brownian motion that has been given a *drift* $\alpha(t) - \frac{1}{2}\sigma^2(t)$ and a *volatility* $\sigma(t)$. The larger $\sigma(t)$, the "wilder" the up- and down swings of the process $X(t)$. Now define

$$S(t) = S(0)e^{X(t)}, \qquad 0 \le t \le T,$$

where $S(0)$ is nonrandom and positive. By analogy with Example 4.36 we can show, using Itô's formula (Theorem 4.29), that $S(t)$ satisfies the SDE

$$dS(t) = \alpha(t)S(t)dt + \sigma(t)S(t)dW(t).$$

(Try it!) The financial interpretation is this: $S(t)$ is the price of a stock, $\alpha(t)$ is the *instantaneous mean rate of return* from investing in this stock, and $\sigma(t)$ is the *volatility*. In the special case $\alpha(t) \equiv 0$, the SDE simplifies to

$$dS(t) = \sigma(t)S(t)dW(t),$$

and formally integrating both sides gives

$$S(t) = S(0) + \int_0^t \sigma(s)S(s)dW(s).$$

It is not clear whether the Itô integral on the right is a martingale, since it seems hard to check the condition $\int_0^T \mathrm{E}[\sigma^2(t)S^2(t)]dt < \infty$. It can be shown, however, that if

$$\mathrm{E}\left[\exp\left\{\frac{1}{2}\int_0^T \sigma^2(s)ds\right\}\right] < \infty$$

(the *Novikov condition*), then $\int_0^t \sigma(s)S(s)dW(s)$ is indeed a martingale, and hence the price process $S(t)$ is a martingale. In the long run, we expect to exactly break even when investing in this stock. This corresponds with the mean rate of return being zero.

The above example includes *all* possible models of an asset price model that is always positive, has no jumps, and is driven by a single Brownian motion.

4.9.2 Two interest rate models

Example 4.40 (Vasicek interest rate model)**.** The Vasicek model for an interest rate process $R(t)$ is

$$dR(t) = \big(\alpha - \beta R(t)\big)dt + \sigma dW(t), \tag{4.46}$$

where α, β and σ are positive constants. Note that this is an example of a linear ODE perturbed by noise. Using the method of Subsection 4.8.3 we obtain the closed-form solution

$$R(t) = e^{-\beta t}R(0) + \frac{\alpha}{\beta}\big(1 - e^{-\beta t}\big) + \sigma e^{-\beta t}\int_0^t e^{\beta s}dW(s). \tag{4.47}$$

The first part of the solution, the nonrandom function

$$r(t) = e^{-\beta t}R(0) + \frac{\alpha}{\beta}\big(1 - e^{-\beta t}\big),$$

is the solution of the ODE

$$r'(t) = \alpha - \beta r(t).$$

Addition of the noise term $\sigma dW(t)$ in (4.46) has the effect of addition of the Itô integral $\sigma e^{-\beta t}\int_0^t e^{\beta s}dW(s)$ to the solution. The Itô integral has mean zero, and by Itô isometry,

$$\mathrm{E}\left(\int_0^t e^{\beta s}dW(s)\right)^2 = \int_0^t e^{2\beta s}ds = \frac{e^{2\beta t} - 1}{2\beta}.$$

Since the Itô integral is the only nonrandom part in the solution (4.47), it follows that $R(t)$ has mean and variance given by

$$\mathrm{E}(R(t)) = e^{-\beta t}R(0) + \frac{\alpha}{\beta}\big(1 - e^{-\beta t}\big), \tag{4.48}$$

$$\mathrm{Var}(R(t)) = \frac{\sigma^2}{2\beta}\big(1 - e^{-2\beta t}\big). \tag{4.49}$$

In fact, $R(t)$ is normally distributed with the above mean and variance, because an Itô integral of a nonrandom function is normally distributed.

The Vasicek model has the advantage that the interest rate is *mean-reverting*: When $R(t) > \alpha/\beta$, the drift term in (4.46) is negative, pushing $R(t)$ back toward α/β. Likewise, when $R(t) < \alpha/\beta$, the drift is positive, again pushing $R(t)$ back toward α/β. Note that $\lim_{t\to\infty} \mathrm{E}(R(t)) = \alpha/\beta$, and if $R(0) = \alpha/\beta$ then $\mathrm{E}(R(t)) = \alpha/\beta$ for all t.

The Vasicek model also has an undesirable feature: $R(t)$ can become negative with positive probability, no matter how the parameters α, β and σ are chosen. The next example does not have this problem, but the price to pay is that we do not have a closed-form solution.

Example 4.41 (Cox-Ingersoll-Ross (CIR) interest rate model)**.** The Cox-Ingersoll-Ross (CIR) model for the interest rate process is

$$dR(t) = \big(\alpha - \beta R(t)\big)dt + \sigma\sqrt{R(t)}dW(t), \tag{4.50}$$

where, as in the previous example, α, β and σ are positive constants. It can be shown that a solution of (4.50) exists, though it can not be expressed in closed form. The advantage of the CIR model is that $R(t)$ is always nonnegative. If $R(t)$ reaches zero, the diffusion coefficient $\sigma\sqrt{R(t)}$ becomes zero and the drift $\alpha - \beta R(t) = \alpha > 0$ drives the interest rate back into positive territory. Like the Vasicek model, the CIR model is mean-reverting.

We now illustrate how Itô's formula can be used to compute the mean and variance of $R(t)$, despite the lack of a closed-form expression for $R(t)$ itself. First, using Corollary 4.35 with $f(x,t) = e^{\beta t}x$, we calculate

$$d\big(e^{\beta t}R(t)\big) = \alpha e^{\beta t}dt + \sigma e^{\beta t}\sqrt{R(t)}dW(t),$$

or in integral form,

$$\begin{aligned} e^{\beta t}R(t) &= R(0) + \alpha\int_0^t e^{\beta s}ds + \sigma\int_0^t e^{\beta s}\sqrt{R(s)}dW(s) \\ &= R(0) + \frac{\alpha}{\beta}(e^{\beta t} - 1) + \sigma\int_0^t e^{\beta s}\sqrt{R(s)}dW(s). \end{aligned}$$

Unfortunately, here we do not know (a priori!) whether $\int_0^t \mathrm{E}[e^{2\beta s}R(s)] < \infty$, so we do not know if the Itô integral in the last equation has an expectation. It can be shown, however, that it does, and in that case its expectation must be zero. Thus, upon dividing by $e^{\beta t}$ we find that $\mathrm{E}(R(t))$ is given by (4.48), as in the Vasicek model.

To compute the variance of $R(t)$, apply Corollary 4.35 this time with $f(x,t) = e^{2\beta t}x^2$ to obtain

$$e^{2\beta t}R^2(t) = R^2(0) + (2\alpha + \sigma^2)\int_0^t e^{2\beta s}R(s)ds + 2\sigma\int_0^t e^{2\beta s}R^{3/2}(s)dW(s).$$

Again it can be shown that the Itô integral has zero mean, so we get

$$e^{2\beta t}\,\mathrm{E}(R^2(t)) = R^2(0) + (2\alpha + \sigma^2)\int_0^t e^{2\beta s}\,\mathrm{E}(R(s))ds.$$

Substituting the expression from (4.48) for $\mathrm{E}(R(s))$ and dividing by $e^{2\beta t}$ gives a (somewhat unwieldy) expression for $\mathrm{E}(R^2(t))$. The variance can then be computed to be

$$\begin{aligned}\mathrm{Var}(R(t)) &= \mathrm{E}\,R^2(t) - \big(\mathrm{E}\,R(t)\big)^2\\ &= \frac{\sigma^2}{\beta}R(0)e^{-\beta t}\big(1 - e^{\beta t}\big) + \frac{\alpha\sigma^2}{2\beta^2}\big(1 - e^{-\beta t}\big)^2.\end{aligned}$$

In particular,

$$\lim_{t\to\infty}\mathrm{Var}(R(t)) = \frac{\alpha\sigma^2}{2\beta^2}.$$

Thus, the oscillations of $R(t)$ about its mean do not get out of hand as t becomes large, a desirable feature for an interest rate process.