CHAPTER FOURTEEN

Continuous-Time Financial Mathematics

The pursuit of mathematics is a divine madness of the human spirit.
Alfred North Whitehead (1861–1947), *Science and the Modern World*

This chapter introduces the mathematics behind continuous-time models. This approach to finance was initiated by Merton [290]. Formidable as the mathematics seems to be, it can be made accessible at some expense of rigor and generality. The theory will be applied to a few fundamental problems in finance.

14.1 Stochastic Integrals

From now on, we use $W \equiv \{ W(t), t \geq 0 \}$ to denote the Wiener process. The goal here is to develop stochastic integrals of X from a class of stochastic processes with respect to the Wiener process:

$$\mathrm{I}_t(X) \equiv \int_0^t X \, dW, \quad t \geq 0.$$

We saw in Subsection 13.3.4 that classical calculus cannot be applied to Brownian motion. One reason is that its sample path, regarded as a function, has unbounded total variation. $\mathrm{I}_t(X)$ is a random variable called the **stochastic integral** of X with respect to W. The stochastic process $\{ \mathrm{I}_t(X), t \geq 0 \}$ is denoted here by $\int X \, dW$. Typical requirements for X in financial applications are (1) $\mathrm{Prob}[\int_0^t X^2(s) \, ds < \infty] = 1$ for all $t \geq 0$ or the stronger $\int_0^t E[X^2(s)] \, ds < \infty$ and (2) that the information set at time t includes the history of X and W up to that point in time but nothing about the evolution of X or W after t (**nonanticipating**, so to speak). The future therefore cannot influence the present, and $\{ X(s), 0 \leq s \leq t \}$ is independent of $\{ W(t+u) - W(t), u > 0 \}$.

The **Ito integral** is a theory of stochastic integration. As with calculus, it starts with step functions. A stochastic process $\{X(t)\}$ is **simple** if there exist $0 = t_0 < t_1 < \cdots$ such that

$$X(t) = X(t_{k-1}) \quad \text{for } t \in [\, t_{k-1}, t_k), \quad k = 1, 2, \ldots$$

https://doi.org/10.1017/CBO9780511546839.015 Published online by Cambridge University Press

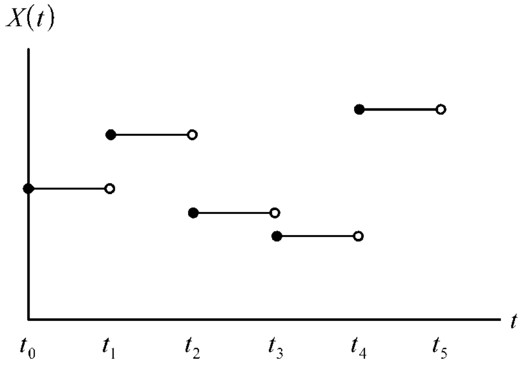

Figure 14.1: Simple stochastic process.

for any realization (see Fig. 14.1). The Ito integral of a simple process is defined as

$$\mathrm{I}_t(X) \equiv \sum_{k=0}^{n-1} X(t_k)[\, W(t_{k+1}) - W(t_k)\,], \tag{14.1}$$

where $t_n = t$. Note that the integrand X is evaluated at t_k, not t_{k+1}.

The natural step to follow is to define the Ito integral of more general processes as a limiting random variable of the Ito integral of simple stochastic processes. Indeed, for a general stochastic process $X = \{\, X(t), t \geq 0\,\}$, there exists a random variable $\mathrm{I}_t(X)$, unique almost certainly, such that $\mathrm{I}_t(X_n)$ converges in probability to $\mathrm{I}_t(X)$ for each sequence of simple stochastic processes $X_1, X_2, \ldots,$ such that X_n converges in probability to X. In particular, if X is continuous with probability one, then $\mathrm{I}_t(X_n)$ converges in probability to $\mathrm{I}_t(X)$ as $\delta_n \equiv \max_{1 \leq k \leq n}(t_k - t_{k-1})$ goes to zero, written as

$$\int_0^t X\, dW = \text{st-lim}_{\delta_n \to 0} \sum_{k=0}^{n-1} X(t_k)[\, W(t_{k+1}) - W(t_k)\,]. \tag{14.2}$$

It is a fundamental fact that $\int X dW$ is continuous almost surely [419, 566]. The following theorem says the Ito integral is a martingale (see Exercise 13.2.6 for its discrete analog), and a corollary is the mean-value formula $E[\int_a^b X dW\,] = 0$.

THEOREM 14.1.1 *The Ito integral $\int X\, dW$ is a martingale.*

Let us inspect Eq. (14.2) more closely. It says the following simple stochastic process $\{\, \widehat{X}(t)\,\}$ can be used in place of X to approximate the stochastic integral $\int_0^t X\, dW$:

$$\widehat{X}(s) \equiv X(t_{k-1}) \quad \text{for } s \in [\, t_{k-1}, t_k), \ \ k = 1, 2, \ldots, n.$$

The key here is the nonanticipating feature of $\widehat{X}$; that is, the information up to time s,

$$\{\, \widehat{X}(t), W(t), 0 \leq t \leq s\,\},$$

cannot determine the future evolution of either X or W. Had we defined the stochastic integral as $\sum_{k=0}^{n-1} X(t_{k+1})[\, W(t_{k+1}) - W(t_k)\,]$, we would have been using the following different simple stochastic process in the approximation,

$$\widehat{Y}(s) \equiv X(t_k) \quad \text{for } s \in [\, t_{k-1}, t_k), \ \ k = 1, 2, \ldots, n,$$

which clearly anticipates the future evolution of X. See Fig. 14.2 for illustration.

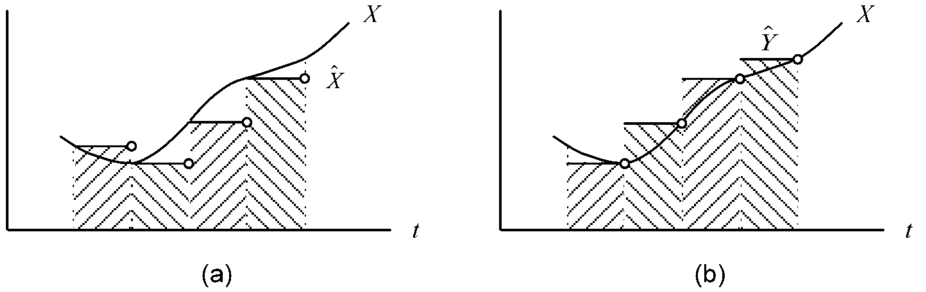

Figure 14.2: Stochastic integration. The simple process $\widehat{X}$ in (a) does not anticipate X, whereas the simple process $\widehat{X}$ in (b) does.

For example, $\int W\,dW$ can be approximated as follows:

$$\begin{aligned}
&\sum_{k=0}^{n-1} W(t_k)[\,W(t_{k+1}) - W(t_k)\,]\\
&\quad= \sum_{k=0}^{n-1} \frac{W(t_{k+1})^2 - W(t_k)^2}{2} - \sum_{k=0}^{n-1} \frac{[\,W(t_{k+1}) - W(t_k)\,]^2}{2}\\
&\quad= \frac{W(t)^2}{2} - \sum_{k=0}^{n-1} \frac{[\,W(t_{k+1}) - W(t_k)\,]^2}{2}.
\end{aligned}$$

Because the second term above converges to $t/2$ by Eq. (13.14),

$$\int_0^t W\,dW = \frac{W(t)^2}{2} - \frac{t}{2}. \tag{14.3}$$

One might have expected $\int_0^t W\,dW = W(t)^2/2$ from calculus. Hence the extra $t/2$ term may come as a surprise. It can be traced to the infinite total variation of Brownian motion. Another way to see the mistake of $\int_0^t W\,dW = W(t)^2/2$ is through Theorem 14.1.1: $W(t)^2/2$ is not a martingale (see Exercise 14.1.3), but $[\,W(t)^2 - t\,]/2$ is (see Exercise 13.3.5).

➤ **Exercise 14.1.1** Prove Theorem 14.1.1 for simple stochastic processes.

➤ **Exercise 14.1.2** Verify that using the following simple stochastic process,

$$Y(s) \equiv W(t_k) \quad \text{for } s \in [\,t_{k-1}, t_k),\ \ k = 1, 2, \ldots, n,$$

to approximate W results in $\int_0^t W\,dW = (W(t)^2 + t)/2$.

Comment 14.1.2 The different results in Exercise 14.1.2 and Eq. (14.3) show the importance of picking the intermediate point for stochastic integrals (here, right end point vs. left end point). The simple stochastic process in Exercise 14.1.2 anticipates the future evolution of W. In general, the following simple stochastic process,

$$Z(s) \equiv W((1-a)\,t_{k-1} + a t_k) \quad \text{for } s \in [\,t_{k-1}, t_k),\ \ k = 1, 2, \ldots, n,$$

gives rise to $\int_0^t W\,dW = W(t)^2/2 + (a - 1/2)\,t$. The Ito integral corresponds to the choice $a = 0$. Standard calculus rules apply when $a = 1/2$, which gives rise to the **Stratonovich stochastic integral**.

➤ **Exercise 14.1.3** Verify that $W(t)^2/2$ is not a martingale.

➤ **Exercise 14.1.4** Prove that $E[\int_0^t W\,dW] = 0$.

➤ **Exercise 14.1.5** Prove that stochastic integration reduces to the usual Riemann–Stieltjes form for constant processes.

14.2 Ito Processes

The stochastic process $X = \{X_t, t \geq 0\}$ that solves

$$X_t = X_0 + \int_0^t a(X_s, s)\,ds + \int_0^t b(X_s, s)\,dW_s, \quad t \geq 0$$

is called an **Ito process**. Here, X_0 is a scalar starting point and $\{a(X_t, t) : t \geq 0\}$ and $\{b(X_t, t) : t \geq 0\}$ are stochastic processes satisfying certain regularity conditions. The terms $a(X_t, t)$ and $b(X_t, t)$ are the **drift** and the **diffusion**, respectively. A shorthand that is due to Langevin's work in 1904 is the following **stochastic differential equation** for the **Ito differential** dX_t,

$$dX_t = a(X_t, t)\,dt + b(X_t, t)\,dW_t, \tag{14.4}$$

or simply $dX_t = a_t\,dt + b_t\,dW_t$ [30, 386]. This is Brownian motion with an instantaneous drift of a_t and an instantaneous variance of b_t^2. In particular, X is a martingale if the drift a_t is zero by Theorem 14.1.1. Recall that dW is normally distributed with mean zero and variance dt. A form equivalent to Eq. (14.4) is the so-called **Langevin equation**:

$$dX_t = a_t\,dt + b_t\sqrt{dt}\;\xi, \tag{14.5}$$

where $\xi \sim N(0, 1)$. This formulation makes it easy to derive Monte Carlo simulation algorithms. Although $dt \ll \sqrt{dt}$, the deterministic term a_t still matters because the random variable ξ makes sure the fluctuation term b_t over successive intervals tends to cancel each other out.

There are regularity conditions that guarantee the existence and the uniqueness of solution for stochastic differential equations [30, 373, 566]. The solution to a stochastic differential equation is also called a **diffusion process**.

14.2.1 Discrete Approximations

The following finite-difference approximation follows naturally from Eq. (14.5):

$$\widehat{X}(t_{n+1}) = \widehat{X}(t_n) + a(\widehat{X}(t_n), t_n)\,\Delta t + b(\widehat{X}(t_n), t_n)\,\Delta W(t_n), \tag{14.6}$$

where $t_n \equiv n\Delta t$. This method is called the **Euler method** or the **Euler–Maruyama method** [556]. Under mild conditions, $\widehat{X}(t_n)$ indeed converges to $X(t_n)$ [572]. Note that $\Delta W(t_n)$ should be interpreted as $W(t_{n+1}) - W(t_n)$ instead of $W(t_n) - W(t_{n-1})$ because a and b are required to be nonanticipating. With the drift a and the diffusion b determined at time t_n, $\widehat{X}$ is expected to be $\widehat{X}(t_n) + a(\widehat{X}(t_n), t_n)\,\Delta t$ at time t_{n+1}. However, the new information $\Delta W(t_n)$, which is unpredictable given the information available at time t_n, dislodges $\widehat{X}$ from its expected position by adding $b(\widehat{X}(t_n), t_n)\,\Delta W(t_n)$. This procedure then repeats itself at $\widehat{X}(t_{n+1})$. See Fig. 14.3 for an illustration.

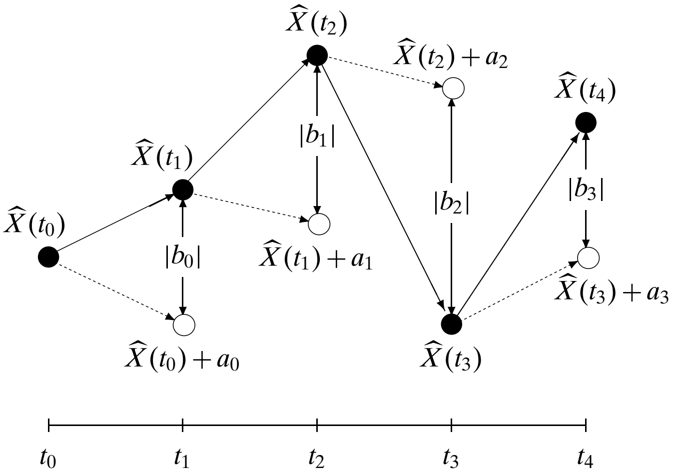

Figure 14.3: Dynamics of Ito process. The filled circles track the process, whereas the unfilled circles are the expected positions. In the plot, $a_i \equiv a\,(\widehat{X}(t_i), t_i)\,\Delta t$, the expected changes, and $b_i \equiv b(\widehat{X}(t_i), t_i)\,\Delta W(t_i)$, the random disturbances. Note that $\widehat{X}(t_{i+1}) = \widehat{X}(t_i) + a_i + b_i$.

The more advanced **Mil'shtein scheme** adds the $bb'[(\Delta W)^2 - \Delta t]/2$ term to Euler's method to provide better approximations [668, 669]. For geometric Brownian motion, for example, Euler's scheme yields

$$\widehat{X}(t_{n+1}) = \widehat{X}(t_n) + \mu\widehat{X}(t_n)\,\Delta t + \sigma\widehat{X}(t_n)\,\Delta W(t_n),$$

whereas Mil'shtein's scheme adds $\sigma^2\widehat{X}(t_n)\{\,[\,\Delta W(t_n)\,]^2 - \Delta t\,\}/2$ to the above.

Under fairly loose regularity conditions, approximation (14.6) can be replaced with

$$\widehat{X}(t_{n+1}) = \widehat{X}(t_n) + a(\widehat{X}(t_n), t_n)\,\Delta t + b(\widehat{X}(t_n), t_n)\sqrt{\Delta t}\,Y(t_n),$$

where $Y(t_0), Y(t_1), \ldots,$ are independent and identically distributed with zero mean and unit variance. This general result is guaranteed by Donsker's theorem [17]. The simpler discrete approximation scheme uses Bernoulli random variables instead:

$$\widehat{X}(t_{n+1}) = \widehat{X}(t_n) + a[\,\widehat{X}(t_n), t_n\,]\,\Delta t + b[\,\widehat{X}(t_n), t_n\,]\sqrt{\Delta t}\,\xi, \tag{14.7}$$

where $\text{Prob}[\,\xi = 1\,] = \text{Prob}[\,\xi = -1\,] = 1/2$. Note that $E[\,\xi\,] = 0$ and $\text{Var}[\,\xi\,] = 1$. This clearly defines a binomial model. As Δt goes to zero, $\widehat{X}$ converges to X [294, 434].

14.2.2 Trading and the Ito Integral

Consider an Ito process $d\boldsymbol{S}_t = \mu_t\,dt + \sigma_t\,dW_t$, where $\boldsymbol{S}_t$ is the vector of security prices at time t. Let ϕ_t be a trading strategy denoting the quantity of each type of security held at time t. Clearly the stochastic process $\phi_t\boldsymbol{S}_t$ is the value of the portfolio ϕ_t at time t. Then $\phi_t\,d\boldsymbol{S}_t \equiv \phi_t(\mu_t\,dt + \sigma_t\,dW_t)$ represents the change in the value from security price changes occurring at time t. The equivalent Ito integral,

$$G_T(\phi) \equiv \int_0^T \phi_t\,d\boldsymbol{S}_t = \int_0^T \phi_t\mu_t\,dt + \int_0^T \phi_t\sigma_t\,dW_t,$$

https://doi.org/10.1017/CBO9780511546839.015 Published online by Cambridge University Press

measures the gains realized by the trading strategy over the period $[0, T]$. A strategy is self-financing if

$$\phi_t \boldsymbol{S}_t = \phi_0 \boldsymbol{S}_0 + G_t(\phi) \tag{14.8}$$

for all $0 \leq t < T$. In other words, the investment at any time equals the initial investment plus the total capital gains up to that time.

Discrete-time models can clarify the above concepts. Let $t_0 < t_1 < \cdots < t_n$ denote the trading points. As before, $\boldsymbol{S}_k$ is the price vector at time t_k, and the vector ϕ_k denotes the quantity of each security held during $[t_k, t_{k+1})$. Thus $\phi_k \boldsymbol{S}_k$ stands for the value of portfolio ϕ_k right after its establishment at time t_k, and $\phi_k \boldsymbol{S}_{k+1}$ stands for the value of ϕ_k at time t_{k+1} *before* any transactions are made to set up the next portfolio ϕ_{k+1}. The nonanticipation requirement of the Ito integral means that ϕ_k must be established before $\boldsymbol{S}_{k+1}$ is known. The quantity $\phi_k \Delta \boldsymbol{S}_k \equiv \phi_k(\boldsymbol{S}_{k+1} - \boldsymbol{S}_k)$ represents the capital gains between times t_k and t_{k+1}, and the summation $G(n) \equiv \sum_{k=0}^{n-1} \phi_k \Delta \boldsymbol{S}_k$ is the total capital gains through time t_n. Note the similarity of this summation to the Ito integral of simple processes (14.1). A trading strategy is self-financing if the investment at any time is financed completely by the investment in the previous period, i.e.,

$$\phi_k \boldsymbol{S}_k = \phi_{k-1} \boldsymbol{S}_k$$

for all $0 < k \leq n$. The preceding condition and condition (14.8) are equivalent (see Exercise 14.2.1).

When an Ito process $dX_t = a_t\, dt + b_t\, dW_t$ is Markovian, the future evolution of X depends solely on its current value. The nonanticipating requirement further says that a_t and b_t cannot embody future values of dW. The Ito process is hence ideal for modeling asset price dynamics under the weak form of efficient markets.

➤ **Exercise 14.2.1** Prove that the self-financing definition $\phi_k \boldsymbol{S}_k = \phi_{k-1} \boldsymbol{S}_k$ implies the alternative condition $\phi_k \boldsymbol{S}_k = \phi_0 \boldsymbol{S}_0 + G_k$ for $0 < k \leq n$, and vice versa.

14.2.3 Ito's Lemma

The central tool in stochastic differential equations is **Ito's lemma**, which basically says that a smooth function of an Ito process is itself an Ito process.

THEOREM 14.2.1 *Suppose that $f: R \to R$ is twice continuously differentiable*[1] *and that $dX = a_t\, dt + b_t\, dW$. Then $f(X)$ is the Ito process*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\, a_s\, ds + \int_0^t f'(X_s)\, b_s\, dW + \frac{1}{2} \int_0^t f''(X_s)\, b_s^2\, ds$$

for $t \geq 0$.

In differential form, Ito's lemma becomes

$$df(X) = f'(X)\, a\, dt + f'(X)\, b\, dW + \frac{1}{2} f''(X)\, b^2\, dt. \tag{14.9}$$

Compared with calculus, the interesting part of Eq. (14.9) is the third term on the right-hand side. This can be traced to the positive quadratic variation of Brownian paths, making $(dW)^2$ nonnegligible. A convenient formulation of Ito's lemma suitable for generalization to higher dimensions is

$$df(X) = f'(X)\, dX + \frac{1}{2} f''(X)(dX)^2. \tag{14.10}$$

Here, we are supposed to multiply out $(dX)^2 = (a\,dt + b\,dW)^2$ symbolically according to the following multiplication table:

$\times$	dW	dt
dW	dt	0
dt	0	0

Note that the $(dW)^2 = dt$ entry is justified by Eq. (13.15). This form is easy to remember because of its similarity to Taylor expansion.

THEOREM 14.2.2 (Higher-Dimensional Ito's Lemma) *Let $W_1, W_2, \ldots, W_n$ be independent Wiener processes and let $X \equiv (X_1, X_2, \ldots, X_m)$ be a vector process. Suppose that $f : R^m \to R$ is twice continuously differentiable and X_i is an Ito process with $dX_i = a_i\,dt + \sum_{j=1}^{n} b_{ij}\,dW_j$. Then $df(X)$ is an Ito process with the differential*

$$df(X) = \sum_{i=1}^{m} f_i(X)\,dX_i + \frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{m} f_{ik}(X)\,dX_i\,dX_k,$$

where $f_i \equiv \partial f/\partial x_i$ and $f_{ik} \equiv \partial^2 f/\partial x_i \partial x_k$.

The multiplication table for Theorem 14.2.2 is

$\times$	dW_i	dt
dW_k	$\delta_{ik}\,dt$	0
dt	0	0

in which

$$\delta_{ik} = \begin{cases} 1, & \text{if } i = k \\ 0, & \text{otherwise} \end{cases}.$$

In applying the higher-dimensional Ito's lemma, usually one of the variables, say X_1, is the time variable t and $dX_1 = dt$. In this case, $b_{1j} = 0$ for all j and $a_1 = 1$. An alternative formulation of Ito's lemma incorporates the interdependence of the variables $X_1, X_2, \ldots, X_m$ into that between the Wiener processes.

THEOREM 14.2.3 *Let $W_1, W_2, \ldots, W_m$ be Wiener processes and let $X \equiv (X_1, X_2, \ldots, X_m)$ be a vector process. Suppose that $f : R^m \to R$ is twice continuously differentiable and X_i is an Ito process with $dX_i = a_i\,dt + b_i\,dW_i$. Then $df(X)$ is the following Ito process,*

$$df(X) = \sum_{i=1}^{m} f_i(X)\,dX_i + \frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{m} f_{ik}(X)\,dX_i\,dX_k,$$

with the following multiplication table:

$\times$	dW_i	dt
dW_k	$\rho_{ik}\,dt$	0
dt	0	0

Here, ρ_{ik} denotes the correlation between dW_i and dW_k.

In Theorem 14.2.3 the correlation between $dW_i = \sqrt{dt}\,\xi_i$ and $dW_k = \sqrt{dt}\,\xi_k$ refers to that between the normally distributed random variables ξ_i and ξ_k.

➤ **Exercise 14.2.2** Prove Eq. (14.3) by using Ito's formula.

14.3 Applications

This section presents applications of the Ito process, some of which will be useful later.

EXAMPLE 14.3.1 A (μ, σ) Brownian motion is $\mu\, dt + \sigma\, dW$ by Ito's lemma and Eq. (13.11).

EXAMPLE 14.3.2 Consider the Ito process $dX = \mu(t)\, dt + \sigma(t)\, dW$. It is identical to Brownian motion except that the drift $\mu(t)$ and diffusion $\sigma(t)$ are no longer constants. Again,

$$X(t) \sim N\left(X(0) + \int_0^t \mu(s)\, ds,\ \int_0^t \sigma^2(s)\, ds\right)$$

is normally distributed.

EXAMPLE 14.3.3 Consider the geometric Brownian motion process $Y(t) \equiv e^{X(t)}$, where $X(t)$ is a (μ, σ) Brownian motion. Ito's formula (14.9) implies that

$$\frac{dY}{Y} = (\mu + \sigma^2/2)\, dt + \sigma\, dW.$$

The instantaneous rate of return is $\mu + \sigma^2/2$, not μ.

EXAMPLE 14.3.4 Consider the Ito process $U \equiv YZ$ with $dY = a\, dt + b\, dW$ and $dZ = f\, dt + g\, dW$. Processes Y and Z share the Wiener process W. Ito's lemma (Theorem 14.2.2) can be used to show that $dU = Z\, dY + Y\, dZ + dY\, dZ$, which equals

$$Z\, dY + Y\, dZ + (a\, dt + b\, dW)(f\, dt + g\, dW) = Z\, dY + Y\, dZ + bg\, dt.$$

If either $b \equiv 0$ or $g \equiv 0$, then integration by parts holds.

EXAMPLE 14.3.5 Consider the Ito process $U \equiv YZ$, where $dY/Y = a\, dt + b\, dW_y$ and $dZ/Z = f\, dt + g\, dW_z$. The correlation between W_y and W_z is ρ. Apply Ito's lemma (Theorem 14.2.3):

$$\begin{aligned} dU &= Z\, dY + Y\, dZ + dY\, dZ \\ &= ZY(a\, dt + b\, dW_y) + YZ(f\, dt + g\, dW_z) \\ &\quad + YZ(a\, dt + b\, dW_y)(f\, dt + g\, dW_z) \\ &= U(a + f + bg\rho)\, dt + Ub\, dW_y + Ug\, dW_z. \end{aligned}$$

Note that dU/U has volatility $\sqrt{b^2 + 2bg\rho + g^2}$ by formula (6.9). The product of two (or more) correlated geometric Brownian motion processes thus remains geometric Brownian motion. This result has applications in **correlation options**, whose value depends on multiple assets. As

$$\begin{aligned} Y &= \exp[\,(a - b^2/2)\, dt + b\, dW_y], \\ Z &= \exp[\,(f - g^2/2)\, dt + g\, dW_z], \\ U &= \exp[\{a + f - (b^2 + g^2)/2\}\, dt + b\, dW_y + g\, dW_z\,], \end{aligned}$$

https://doi.org/10.1017/CBO9780511546839.015 Published online by Cambridge University Press

$\ln U$ is Brownian motion with a mean equal to the sum of the means of $\ln Y$ and $\ln Z$. This holds even if Y and Z are correlated. Finally, $\ln Y$ and $\ln Z$ have correlation ρ.

EXAMPLE 14.3.6 Suppose that S follows $dS/S = \mu\, dt + \sigma\, dW$. Then $F(S, t) \equiv Se^{y(T-t)}$ follows

$$\frac{dF}{F} = (\mu - y)\, dt + \sigma\, dW$$

by Ito's lemma. This result has applications in forward and futures contracts.

➤ **Exercise 14.3.1** Assume that $dX/X = \mu\, dt + \sigma\, dW$. (1) Prove that $\ln X$ follows $d(\ln X) = (\mu - \sigma^2/2)\, dt + \sigma\, dW$. (2) Derive the probability distribution of $\ln(X(t)/X(0))$.

➤ **Exercise 14.3.2** Let X follow the geometric Brownian motion process $dX/X = \mu\, dt + \sigma\, dW$. Show that $R \equiv \ln X + \sigma^2 t/2$ follows $dR = \mu\, dt + \sigma\, dW$.

➤ **Exercise 14.3.3** (1) What is the stochastic differential equation for the process W^n? (2) Show that

$$\int_s^t W^n\, dW = \frac{W(t)^{n+1} - W(s)^{n+1}}{n+1} - \frac{n}{2} \int_s^t W^{n-1}\, dt.$$

(Hint: Use Eqs. (13.15) and (13.16) or apply Ito's lemma.)

➤ **Exercise 14.3.4** Consider the Ito process $U \equiv (Y + Z)/2$, where $dY/Y = a\, dt + b\, dW$ and $dZ/Z = f\, dt + g\, dW$. Processes Y and Z share the Wiener process W. Derive the stochastic differential equation for dU/U.

➤ **Exercise 14.3.5** Redo Example 14.3.4 except that $dY = a\, dt + b\, dW_y$ and $dZ = f\, dt + g\, dW_z$, where dW_y and dW_z have correlation ρ.

➤ **Exercise 14.3.6** Verify that $U \equiv Y/Z$ follows $dU/U = (a - f - bg\rho)\, dt + b\, dW_y - g\, dW_z$, where Y and Z are drawn from Example 14.3.5.

➤ **Exercise 14.3.7** Given $dY/Y = \mu\, dt + \sigma\, dW$ and $dX/X = r\, dt$, derive the stochastic differential equation for $F \equiv X/Y$.

14.3.1 The Ornstein–Uhlenbeck Process

The **Ornstein–Uhlenbeck process** has the stochastic differential equation

$$dX = -\kappa X\, dt + \sigma\, dW, \tag{14.11}$$

where $\kappa, \sigma \geq 0$ (see Fig. 14.4). It is known that

$$E[\, X(t)\,] = e^{-\kappa(t-t_0)}\, E[\, x_0\,],$$

$$\mathrm{Var}[\, X(t)\,] = \frac{\sigma^2}{2\kappa}\left[1 - e^{-2\kappa(t-t_0)}\right] + e^{-2\kappa(t-t_0)}\, \mathrm{Var}[\, x_0\,],$$

$$\mathrm{Cov}[\, X(s), X(t)\,] = \frac{\sigma^2}{2\kappa}\, e^{-\kappa(t-s)}\left[1 - e^{-2\kappa(s-t_0)}\right] + e^{-\kappa(t+s-2t_0)}\, \mathrm{Var}[\, x_0\,]$$

for $t_0 \leq s \leq t$ and $X(t_0) = x_0$. In fact, $X(t)$ is normally distributed if x_0 is a constant or normally distributed [30]; X is said to be a **normal process**. Of course, $E[\, x_0\,] = x_0$ and $\mathrm{Var}[\, x_0\,] = 0$ if x_0 is a constant. When $x_0 \sim N(0, \frac{\sigma^2}{2\kappa})$, it is easy to see that X is

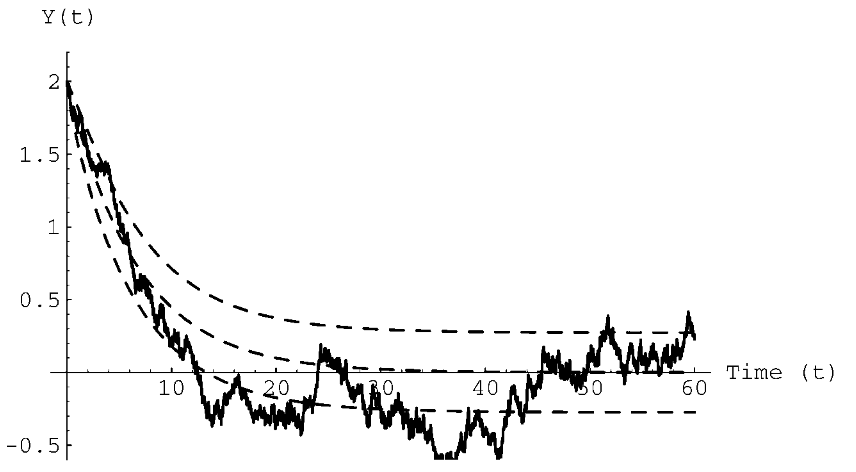

Figure 14.4: Sample path of Ornstein–Uhlenbeck process. Shown is a sample path of the Ornstein–Uhlenbeck process $dY = -0.15\,Y\,dt + 0.15\,dW$, starting at $Y(0) = 2$. The envelope is for one standard deviation $\sqrt{[(0.15)^2/0.3](1 - e^{-0.3t})}$ around the mean $2e^{-0.15t}$. In contrast to Brownian motion, which diverges to infinite values (see Fig. 13.2), the Ornstein–Uhlenbeck process converges to a stationary distribution.

stationary. The Ornstein–Uhlenbeck process describes the velocity of a tiny particle through a fluid in thermal equilibrium – in short, Brownian motion in nature [386].

The Ornstein–Uhlenbeck process has the following **mean-reversion** property. When $X > 0$, the dX term tends to be negative, pulling X toward zero, whereas if $X < 0$, the dX term tends to be positive, pulling X toward zero again.

EXAMPLE 14.3.7 Suppose that X is an Ornstein–Uhlenbeck process. Ito's lemma says that $V \equiv X^2$ has the differential

$$\begin{aligned} dV &= 2X\,dX + (dX)^2 = 2\sqrt{V}\,(-\kappa\sqrt{V}\,dt + \sigma\,dW) + \sigma^2\,dt \\ &= (-2\kappa V + \sigma^2)\,dt + 2\sigma\sqrt{V}\,dW, \end{aligned}$$

a **square-root process**.

Consider the following process, also called the Ornstein–Uhlenbeck process:

$$dX = \kappa(\mu - X)\,dt + \sigma\,dW, \tag{14.12}$$

where $\sigma \geq 0$. Given $X(t_0) = x_0$, a constant, it is known that

$$E[\,X(t)\,] = \mu + (x_0 - \mu)\,e^{-\kappa(t-t_0)} \tag{14.13}$$

$$\mathrm{Var}[\,X(t)\,] = \frac{\sigma^2}{2\kappa}\left[1 - e^{-2\kappa(t-t_0)}\right] \tag{14.14}$$

for $t_0 \leq t$ [855]. Because the mean and the standard deviation are roughly μ and $\sigma/\sqrt{2\kappa}$, respectively, for large t, the probability of X's being negative is extremely unlikely in any finite time interval when $\mu > 0$ is large relative to $\sigma/\sqrt{2\kappa}$ (say $\mu > 4\sigma/\sqrt{2\kappa}$). Process (14.12) has the salient mean-reverting feature that X tends to move toward μ, making it useful for modeling term structure [855], stock price volatility [823], and stock price return [613].

➤ **Exercise 14.3.8** Let $X(t)$ be the Ornstein–Uhlenbeck process in Eq. (14.11). Show that the differential for $Y(t) \equiv X(t)\, e^{\kappa t}$ is $dY = \sigma e^{\kappa t}\, dW$. (This implies that $Y(t)$, hence $X(t)$ as well, is normally distributed.)

➤ **Exercise 14.3.9** Justify the claim in Eq. (13.13) by showing that $Y(t) \equiv e^{-t}\, W(e^{2t})$ is the Ornstein–Uhlenbeck process $dY = -Y\, dt + \sqrt{2}\, dW$. (Hint: Consider $Y(t+dt) - Y(t)$.)

➤ **Exercise 14.3.10** Consider the following processes:

$$\begin{aligned} dS &= \mu S\, dt + \sigma S\, dW_1, \\ d\sigma &= \beta(\overline{\sigma} - \sigma)\, dt + \gamma\, dW_2, \end{aligned}$$

where dW_1 and dW_2 are Wiener processes with correlation ρ. Let $H(S, \sigma, \tau)$ be a function of S, σ, and τ. Derive its stochastic differential equation. (This process models stock price with a correlated stochastic volatility, which follows a mean-reverting process.)

➤ **Exercise 14.3.11** Show that the transition probability density function p of $dX = -(1/2)\, X dt + dW$ satisfies the backward equation

$$\frac{\partial p}{\partial s} = -\frac{1}{2}\frac{\partial^2 p}{\partial x^2} + \frac{1}{2}\, x\, \frac{\partial p}{\partial x}.$$

(Hint: $X(t) \sim N(xe^{-(t-s)/2}, 1 - e^{-(t-s)})$ when $X(s) = x$.)

14.3.2 The Square-Root Process

The square-root process has the stochastic differential equation

$$dX = \kappa(\mu - X)\, dt + \sigma\sqrt{X}\, dW,$$

where $\kappa, \sigma \geq 0$ and the initial value of X is a nonnegative constant. See Fig. 14.5 for an illustration. Like the Ornstein–Uhlenbeck process, the square-root process

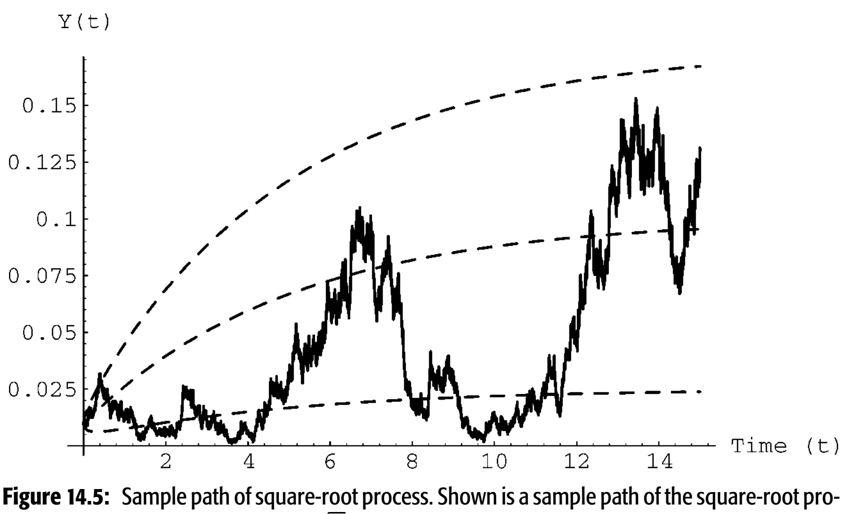

Figure 14.5: Sample path of square-root process. Shown is a sample path of the square-root process $dY = 0.2(0.1 - Y)\, dt + 0.15\sqrt{Y}\, dW$ with the initial condition $Y(0) = 0.01$. The envelope is for one standard deviation around the mean, which is $0.01\, e^{-0.2t} + 0.1\,(1 - e^{-0.2t})$.

possesses mean reversion in that X tends to move toward μ, but the volatility is proportional to $\sqrt{X}$ instead of a constant. When X hits zero and $\mu \geq 0$, the probability is one that it will not move below zero; in other words, zero is a **reflecting boundary**. Hence, the square-root process is a good candidate for modeling interest rate movements [234]. The Ornstein–Uhlenbeck process, in contrast, allows negative interest rates. The two processes are related (see Example 14.3.7).

Feller (1906–1970) showed that the random variable $2cX(t)$ follows the noncentral chi-square distribution

$$\chi\left(\frac{4\kappa\mu}{\sigma^2}, 2cX(0)\, e^{-\kappa t}\right),$$

where $c \equiv (2\kappa/\sigma^2)(1-e^{-\kappa t})^{-1}$ [234, 341]. Given $X(0) = x_0$, a constant, it can be proved that

$$E[\, X(t)\,] = x_0 e^{-\kappa t} + \mu(1-e^{-\kappa t}),$$
$$\mathrm{Var}[\, X(t)\,] = x_0\, \frac{\sigma^2}{\kappa}\,(e^{-\kappa t} - e^{-2\kappa t}) + \mu\, \frac{\sigma^2}{2\kappa}(1-e^{-\kappa t})^2$$

for $t \geq 0$.

14.4 Financial Applications

14.4.1 Transactions Costs

Transactions costs are a fact of life, never zero however negligible. Under the **proportional transactions cost model**, it is impossible to trade continuously. Intuitively, this is because the transactions cost per trade is proportional to $|dW|$, and $\int_0^T |dW| = \infty$ almost surely by Eq. (13.17). As a consequence, a continuous trader would be bankrupt with probability one [660]. Even stronger claims can be made. For instance, the cheapest trading strategy to dominate the value of European call at maturity is the covered call [814].

➤ **Exercise 14.4.1** Argue that an investor who has information about the entire future value of the Brownian motion's driving the stock price will have infinite wealth at any given horizon date. In other words, market fluctuations can be exploited.

14.4.2 Stochastic Interest Rate Models

Merton originated the following methodology to term structure modeling [493]. Suppose that the short rate r follows a Markov process $dr = \mu(r, t)\, dt + \sigma(r, t)\, dW$. Let $P(r, t, T)$ denote the price at time t of a zero-coupon bond that pays \$1 at time T. Its stochastic process must also be Markovian. Write its dynamics as

$$\frac{dP}{P} = \mu_p\, dt + \sigma_p\, dW$$

so that the expected instantaneous rate of return on a $(T-t)$-year zero-coupon bond is μ_p and the instantaneous variance is σ_p^2. Surely $P(r, T, T) = 1$ for any T. By Ito's lemma (Theorem 14.2.2),

$$\begin{aligned} dP &= \frac{\partial P}{\partial T}\, dT + \frac{\partial P}{\partial r}\, dr + \frac{1}{2}\, \frac{\partial^2 P}{\partial r^2}\, (dr)^2 \\ &= -\frac{\partial P}{\partial T}\, dt + \frac{\partial P}{\partial r}\, [\, \mu(r,t)\, dt + \sigma(r,t)\, dW\,] \\ &\quad + \frac{1}{2}\, \frac{\partial^2 P}{\partial r^2}\, [\, \mu(r,t)\, dt + \sigma(r,t)\, dW\,]^2 \\ &= \left[-\frac{\partial P}{\partial T} + \mu(r,t)\, \frac{\partial P}{\partial r} + \frac{\sigma(r,t)^2}{2}\, \frac{\partial^2 P}{\partial r^2} \right] dt + \sigma(r,t)\, \frac{\partial P}{\partial r}\, dW, \end{aligned}$$

where $dt = -dT$ in the second equality. Hence

$$-\frac{\partial P}{\partial T} + \mu(r,t)\, \frac{\partial P}{\partial r} + \frac{\sigma(r,t)^2}{2}\, \frac{\partial^2 P}{\partial r^2} = P\mu_p, \quad \sigma(r,t)\, \frac{\partial P}{\partial r} = P\sigma_p. \tag{14.15}$$

Models with the short rate as the only explanatory variable are called **short rate models**.

The Merton Model

Suppose we assume the local expectations theory, which means that μ_p equals the prevailing short rate $r(t)$ for all T, and we assume that μ and σ are constants. Then the partial differential equations (14.15) yield the following solution:

$$P(r,t,T) = \exp\left[-r(T-t) - \frac{\mu(T-t)^2}{2} + \frac{\sigma^2(T-t)^3}{6} \right]. \tag{14.16}$$

This model is due to Merton [660]. We make a few observations. First, $\sigma_p = -\sigma(T-t)$, which says sensibly that bonds with longer maturity are more volatile. The dynamics of P is $dP/P = r\, dt - \sigma(T-t)\, dW$. Now, P has no upper limits as T becomes large, which does not square with the reality. This happens because of negative rates in the model.

➤ **Exercise 14.4.2** Negative interest rates imply arbitrage profits for riskless bonds. Why?

Duration under Parallel Shifts

Consider duration with respect to parallel shifts in the spot rate curve. For convenience, assume that $t = 0$. Parallel shift means $S(r + \Delta r, T) = S(r, T) + \Delta r$ for any Δr; so $\partial S(r,T)/\partial r = 1$. This implies $S(r,T) = r + g(T)$ for some function g with $g(0) = 0$ because $S(r, 0) = r$. Consequently, $P(r,T) = e^{-[r+g(T)]\,T}$. Substitute this identity into the left-hand part of Eqs. (14.15) and assume the local expectations theory to obtain

$$g'(T) + \frac{g(T)}{T} = \mu(r) - \frac{\sigma(r)^2}{2}\, T.$$

As the left-hand side is independent of r, so must the right-hand side be. Because this must hold for all T, both $\mu(r)$ and $\sigma(r)$ must be constants, i.e., the Merton model. As mentioned before, this model is flawed, so must duration be, as such [496].

➤ **Exercise 14.4.3** Suppose that the current spot rate curve is flat. Under the assumption that only parallel shifts are allowed, what can we say about the parameters μ and σ governing the short rate process?

Immunization under Parallel Shifts Revisited

A duration-matched portfolio under parallel shifts in the spot rate curve begets arbitrage opportunities in that the portfolio value exceeds the liability for any instantaneous rate changes. This was shown in Subsection 5.8.2. However, this seeming inconsistency with equilibrium disappears if changes in portfolio value *through* time are taken into account. Indeed, for certain interest rate models, a liability immunized by a duration-matched portfolio exceeds the minimum portfolio value at any given time in the future. Thus the claimed arbitrage profit evaporates because the portfolio value does not always cover the liability.

We illustrate this point with the Merton model $dr = \mu\, dt + \sigma\, dW$, which results from parallel shifts in the spot rate curve and the local expectations theory. To immunize a \$1 liability that is due at time s, a two-bond portfolio is constructed now with maturity dates t_1 and t_2, where $t_1 < s < t_2$. Each bond is a zero-coupon bond with \$1 par value. The portfolio matches the present value of the liability today, and its value relative to the PV of the liability is minimum among all such portfolios (review Subsection 4.2.2). Consider any future time t such that $t < t_1$. With $A(t)$ denoting the portfolio value and $L(t)$ the liability value at time t, it can be shown that the asset/liability ratio $A(t)/L(t)$ is a convex function of the prevailing interest rate and $A(t) < L(t)$ (see Exercise 14.4.4). This conclusion holds for other interest rate models [52].

➤ **Exercise 14.4.4** (1) Prove that $A(t)/L(t)$ is a convex function of the prevailing interest rate. (2) Then verify $A(t) < L(t)$.

14.4.3 Modeling Stock Prices

The most popular stochastic model for stock prices has been geometric Brownian motion $dS/S = \mu\, dt + \sigma\, dW$. This model best describes an equilibrium in which expectations about future returns have settled down [660].

From the discrete-time analog $\Delta S/S = \mu\, \Delta t + \sigma \sqrt{\Delta t}\, \xi$, where $\xi \sim N(0, 1)$, we know that $\Delta S/S \sim N(\mu\, \Delta t, \sigma^2 \Delta t)$. The percentage return for the next Δt time hence has mean $\mu\, \Delta t$ and variance $\sigma^2 \Delta t$. In other words, the percentage return per unit time has mean μ and variance σ^2. For this reason, μ is called the expected instantaneous rate of return and σ^2 the instantaneous variance of the rate of return. If there is no uncertainty about the stock price, i.e., $\sigma = 0$, then $S(t) = S(0)\, e^{\mu t}$.

Comment 14.4.1 It may seem strange that the rate of return is μ instead of $\mu - \sigma^2/2$. Example 14.3.3 says that $S(t)/S(0) = e^{X(t)}$, where $X(t)$ is a $(\mu - \sigma^2/2, \sigma)$ Brownian motion and the continuously compounded rate of return over the time period $[\,0, T\,]$ is

$$\frac{\ln[\, S(T)/S(0)\,]}{T} = \frac{X(T) - X(0)}{T} \sim N\left(\mu - \frac{\sigma^2}{2}, \sigma^2\right). \tag{14.17}$$

The expected continuously compounded rate of return is then $\mu - \sigma^2/2$! Well, they refer to alternative definitions of rates of return. Unless stated otherwise, it is the *former* (instantaneous rate of return μ) that we have in mind from now on. It should be pointed out that the μ used in the BOPM in Eq. (9.13) referred to the latter rate of return. In summary,

$$\frac{E[\{S(\Delta t) - S(0)\}/S(0)]}{\Delta t} \to \mu,$$

$$\frac{\ln E[S(T)/S(0)]}{T} = \mu,$$

$$\frac{E[\ln(S(T)/S(0))]}{T} = \mu - \frac{\sigma^2}{2}.$$

(See Comment 9.3.2, Lemma 9.3.1, Example 14.3.3, and Exercises 13.3.8 and 14.4.5.)

➤ **Exercise 14.4.5** Prove that $E[S(T)] = S(0)\, e^{\mu T}$.

➤ **Exercise 14.4.6** Suppose the stock price follows the geometric Brownian motion process $dS/S = \sigma\, dW$. Example 14.3.3 says that $S(t)/S(0) = e^{X(t)}$, where $X(t)$ is a $(-\sigma^2/2, \sigma)$ Brownian motion. In other words, the stock is expected to have a negative growth rate. Explain why the growth rate is not zero.

➤ **Exercise 14.4.7** Show that the simple rate of return $[S(t)/S(0)] - 1$ has mean $e^{\mu t} - 1$ and variance $e^{2\mu t}(e^{\sigma^2 t} - 1)$.

➤ **Exercise 14.4.8** Assume that the volatility σ is stochastic but driven by an independent Wiener process. Suppose that the average variance over the time period $[0, T]$ as defined by $\widehat{\sigma^2} \equiv \frac{1}{T}\int_0^T \sigma^2(t)\, dt$ is given. Argue that

$$\ln \frac{S(T)}{S(0)} \sim N(\mu T - (\widehat{\sigma^2}\ T/2), \widehat{\sigma^2}\ T).$$

(Thus $S(T)$ seen from time zero remains lognormally distributed.)

➤ **Exercise 14.4.9** Justify using $\Delta S/(S\sqrt{\Delta t})$ to estimate volatility.

➤ **Exercise 14.4.10** What are the shortcomings of modeling the stock price dynamics by $dS = \mu\, dt + \sigma\, dW$ with constant μ and σ?

Continuous-Time Limit of the Binomial Model

What is the Ito process for the stock's rate of return in a risk-neutral economy to which the binomial model in Section 9.2 converges? The continuously compounded rate of return of the stock price over a period of length τ is a sum of the following n independent identically distributed random variables:

$$X_i = \begin{cases} \ln u & \text{with probability } p \\ \ln d & \text{with probability } 1 - p \end{cases},$$

where $u \equiv e^{\sigma\sqrt{\tau/n}}$, $d \equiv e^{-\sigma\sqrt{\tau/n}}$, and $p \equiv (e^{r\tau/n} - d)/(u - d)$. The rate of return is hence the random walk $\sum_{i=1}^n X_i$. It is straightforward to verify that

$$E\left[\sum_{i=1}^n X_i\right] \to \left(r - \frac{\sigma^2}{2}\right)\tau, \qquad \mathrm{Var}\left[\sum_{i=1}^n X_i\right] \to \sigma^2\tau. \tag{14.18}$$

The continuously compounded rate of return thus converges to a $(r - \sigma^2/2, \sigma)$ Brownian motion, and the stock price follows $dS/S = r\, dt + \sigma\, dW$ in a risk-neutral economy. Indeed, the discount process $\{Z(t) \equiv e^{-rt} S(t), t \geq 0\}$ is a martingale under this risk-neutral probability measure as $dZ/Z = \sigma\, dW$ is a driftless Ito process by Exercise 14.3.6.

➤ **Exercise 14.4.11** Verify Eqs. (14.18). (Hint: $e^x \approx 1 + x + x^2/2$.)

➤ **Exercise 14.4.12** From the above discussions, $E[\, X_i\,] \to (r - \sigma^2/2)\,\Delta t$ and $\sqrt{\mathrm{Var}[\, X_i\,]} \to \sigma\sqrt{\Delta t}$. Now $X_{i+1} = \ln(S_{i+1}/S_i)$, where $S_i \equiv S_0 e^{X_1 + X_2 + \cdots + X_i}$ is the stock price at time i. Hence

$$X_{i+1} = \ln\left(1 + \frac{S_{i+1} - S_i}{S_i}\right) \approx \frac{S_{i+1} - S_i}{S_i} \equiv \frac{\Delta S_i}{S_i}.$$

The above finding suggests that $dS/S = (r - \sigma^2/2)\, dt + \sigma\, dW$, contradicting the above! Find the hole in the argument and correct it.

➤ **Exercise 14.4.13** The continuously compounded rate of return $X \equiv \ln S$ follows $dX = (r - \sigma^2/2)\, dt + \sigma\, dW$ in a risk-neutral economy. Use this fact to show that

$$u = \exp[\,(r - \sigma^2/2)\ \Delta t + \sigma\sqrt{\Delta t}\,], \quad d = \exp[\,(r - \sigma^2/2)\ \Delta t - \sigma\sqrt{\Delta t}\ ]$$

under the alternative binomial model in which an up move occurs with probability 1/2.

Additional Reading

We followed [541, 543, 763, 764] in the exposition of stochastic processes and [30, 289, 419, 544] in the discussions of stochastic integrals. The Ito integral is due to Ito (1915–) [500]. Ito's lemma is also due to Ito under the strong influence of Bachelier's thesis [501, 512]. Rigorous proofs of Ito's lemma can be found in [30, 419], and informal ones can be found in [470, 492, 660]. *Mathematica* programs for carrying out some of the manipulations are listed in [854]. Consult [556, 557, 558, 774] for numerical solutions of stochastic differential equations. See [613] for the multivariate Ornstein–Uhlenbeck process. Other useful references include [822] (diffusion), [549] (Ito integral), [280, 761] (stochastic processes), [211, 364, 776] (stochastic differential equations), [261] (stochastic convergence), [542] (stochastic optimization in trading), [70] (nonprobabilistic treatment of continuous-time treading), [102, 181] (distribution-free competitive trading), and [112, 115, 262, 274, 681] (transactions costs).

NOTE

1. This means that all first- and second-order partial derivatives exist and are continuous.

> A proof is that which convinces a reasonable man; a rigorous proof is that which convinces an unreasonable man.
>
> Mark Kac (1914–1984)