# 9

# Change of Numéraire

## 9.1 Introduction

A *numéraire* is the unit of account in which other assets are denominated. One usually takes the numéraire to be the currency of a country. One might change the numéraire by changing to the currency of another country. As this example suggests, in some applications one must change the numéraire in which one works because of finance considerations. We shall see that sometimes it is convenient to change the numéraire because of modeling considerations as well. A model can be complicated or simple, depending on the choice of the numéraire for the model.

In this chapter, we will work within the multidimensional market model of Section 5.4. In particular, our model will be driven by a $d$-dimensional Brownian motion $W(t) = (W_1(t), \dots, W_d(t))$, $0 \le t \le T$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In particular, $W_1, \dots, W_d$ are independent Brownian motions. The filtration $\mathcal{F}(t)$, $0 \le t \le T$, is the one generated by this vector of Brownian motions. There is an adapted interest rate process $R(t)$, $0 \le t \le T$. This can be used to create a money market account whose price per share at time $t$ is

$$M(t) = e^{\int_0^t R(u)du}.$$

This is the capital an agent would have if the agent invested one unit of currency in the money market account at time zero and continuously rolled over the capital at the short-term interest rate. We also define the discount process

$$D(t) = e^{-\int_0^t R(u)du} = \frac{1}{M(t)}.$$

There are $m$ primary assets in the model of this chapter, and their prices satisfy equation (5.4.6), which we repeat here:

$$dS_i(t) = \alpha_i(t) S_i(t)\,dt + S_i(t) \sum_{j=1}^{d} \sigma_{ij}(t)\,dW_j(t), \quad i = 1, \dots, m. \tag{9.1.1}$$

We assume there is a unique risk-neutral measure $\widetilde{\mathbb{P}}$ (i.e., there is a unique $d$-dimensional process $\Theta(t) = (\Theta_1(t), \dots, \Theta_d(t))$ satisfying the market price of risk equations (5.4.18)). The risk-neutral measure is constructed using the multidimensional Girsanov Theorem 5.4.1. Under $\widetilde{\mathbb{P}}$, the Brownian motions

$$\widetilde{W}_j(t) = W_j(t) + \int_0^t \Theta_j(u)\,du, \quad j = 1, \dots, d,$$

are independent of one another. According to the Second Fundamental Theorem of Asset Pricing, Theorem 5.4.9, the market is complete; every derivative security can be hedged by trading in the primary assets and the money market account.

Under $\widetilde{\mathbb{P}}$, the discounted asset prices $D(t)S_i(t)$ are martingales, and so the discounted value of every portfolio process is also a martingale. The risk-neutral measure $\widetilde{\mathbb{P}}$ is thus associated with the money market account price $M(t)$ in the following way. If we were to denominate the $i$th asset in terms of the money market account, its price would be $S_i(t)/M(t) = D(t)S_i(t)$. In other words, at time $t$, the $i$th asset is worth $D(t)S_i(t)$ shares of the money market account. This process, the value of the $i$th asset denominated in shares of the money market account, is a martingale under $\widetilde{\mathbb{P}}$. We say the measure $\widetilde{\mathbb{P}}$ is *risk-neutral for the money market account numéraire.*

When we change the numéraire, denominating the $i$th asset in some other unit of account, it is no longer a martingale under $\widetilde{\mathbb{P}}$. When we change the numéraire, we need to also change the risk-neutral measure in order to maintain risk neutrality. The details and some applications of this idea are developed in this chapter.

## 9.2 Numéraire

In principle, we can take any positively priced asset as a *numéraire* and denominate all other assets in terms of the chosen numéraire. Associated with each numéraire, we shall have a risk-neutral measure. When making this association, we shall take only non-dividend-paying assets as numéraires. In particular, we regard $\widetilde{\mathbb{P}}$ as the risk-neutral measure associated with the domestic money market account, not the domestic currency. Currency pays a dividend because it can be invested in the money market. In contrast, in our model, a share of the money market account increases in value without paying a dividend.

The numéraires we consider in this chapter are:

- Domestic money market account. We denote the associated risk-neutral measure by $\widetilde{\mathbb{P}}$. It is the one discussed in Section 9.1.
- Foreign money market account. We denote the associated risk-neutral measure by $\widetilde{\mathbb{P}}^f$. It is constructed in Section 9.3 below.

- A zero-coupon bond maturing at time $T$. We denote the associated risk-neutral measure by $\widetilde{\mathbb{P}}^T$. It is called the *T-forward measure* and is used in Section 9.4.

The asset we take as numéraire could be one of the primary assets given by (9.1.1) or it could be a derivative asset. Regardless of which asset we take, it has the stochastic representation provided by the following theorem.

**Theorem 9.2.1 (Stochastic representation of assets).** *Let $N$ be a strictly positive price process for a non-dividend-paying asset, either primary or derivative, in the multidimensional market model of Section 9.1. Then there exists a* vector volatility process

$$\nu(t) = (\nu_1(t), \dots, \nu_d(t))$$

*such that*

$$dN(t) = R(t)N(t)\,dt + N(t)\nu(t) \cdot d\widetilde{W}(t). \tag{9.2.1}$$

*This equation is equivalent to each of the equations*

$$d\big(D(t)N(t)\big) = D(t)N(t)\nu(t) \cdot d\widetilde{W}(t), \tag{9.2.2}$$

$$D(t)N(t) = N(0)\exp\left\{\int_0^t \nu(u) \cdot d\widetilde{W}(u) - \frac{1}{2}\int_0^t \|\nu(u)\|^2\,du\right\}, \tag{9.2.3}$$

$$N(t) = N(0)\exp\left\{\int_0^t \nu(u) \cdot d\widetilde{W}(u) + \int_0^t \left(R(u) - \frac{1}{2}\|\nu(u)\|^2\right)du\right\}. \tag{9.2.4}$$

*In other words, under the risk-neutral measure,* every asset has a mean return equal to the interest rate. *The realized risk-neutral return for assets is characterized solely by their volatility vector processes (because initial conditions have no effect on return).*

PROOF: Under the risk-neutral measure $\widetilde{\mathbb{P}}$, the discounted price process $D(t)N(t)$ must be a martingale. The risk-neutral measure is constructed to enforce this condition for primary assets, and it is a consequence of the risk-neutral pricing formula for derivative assets. According to the Martingale Representation Theorem, Theorem 5.4.2,

$$d\left(D(t)N(t)\right) = \sum_{j=1}^{d} \widetilde{\Gamma}_j(t)\,d\widetilde{W}_j(t) = \widetilde{\Gamma}(t) \cdot d\widetilde{W}(t)$$

for some adapted $d$-dimensional process $\widetilde{\Gamma}(t) = (\widetilde{\Gamma}_1(t), \dots, \widetilde{\Gamma}_d(t))$. Because $N(t)$ is strictly positive, we can define the vector $\nu(t) = (\nu_1(t), \dots, \nu_d(t))$ by

$$\nu_j(t) = \frac{\widetilde{\Gamma}_j(t)}{D(t)N(t)}.$$

Then

$$d\left(D(t)N(t)\right) = D(t)N(t)\nu(t)\cdot d\widetilde{W}(t),$$

which is (9.2.2).

The solution to (9.2.2) is (9.2.3), as we now show. Define

$$\begin{aligned} X(t) &= \int_0^t \nu(u)\cdot d\widetilde{W}(u) - \frac{1}{2}\int_0^t \|\nu(u)\|^2\,du \\ &= \sum_{j=1}^d \int_0^t \nu_j(u)\,d\widetilde{W}_j(u) - \frac{1}{2}\sum_{j=1}^d \int_0^t \nu_j^2(u)\,du, \end{aligned}$$

so that

$$\begin{aligned} dX(t) &= \nu(t)\cdot d\widetilde{W}(t) - \frac{1}{2}\|\nu(t)\|^2\,dt \\ &= \sum_{j=1}^d \nu_j(t)\,d\widetilde{W}_j(t) - \frac{1}{2}\sum_{j=1}^d \nu_j^2(t)\,dt. \end{aligned}$$

Then

$$dX(t)\,dX(t) = \sum_{j=1}^d \nu_j^2(t)\,dt = \|\nu(t)\|^2\,dt.$$

Let $f(x) = N(0)e^x$, and compute

$$\begin{aligned} df(X(t)) &= f'(X(t))\,dX(t) + \frac{1}{2}f''(X(t))\,dX(t)\,dX(t) \\ &= f(X(t))\nu(t)\cdot d\widetilde{W}(t). \end{aligned}$$

We see that $f(X(t))$ solves (9.2.2), $f(X(t))$ has the desired initial condition $f(X(0)) = N(0)$, and $f(X(t))$ is the right-hand side of (9.2.3).

From (9.2.3), we have immediately that (9.2.4) holds. Applying the Itô-Doeblin formula to (9.2.4), we obtain (9.2.1). □

According to the multidimensional Girsanov Theorem, Theorem 5.4.1, we can use the volatility vector of $N(t)$ to change the measure. Define

$$\widetilde{W}_j^{(N)}(t) = -\int_0^t \nu_j(u)\,du + \widetilde{W}_j(t),\ \ j = 1,\ldots,d, \tag{9.2.5}$$

and a new probability measure

$$\widetilde{\mathbb{P}}^{(N)}(A) = \frac{1}{N(0)}\int_A D(T)N(T)\,d\widetilde{\mathbb{P}} \text{ for all } A\in\mathcal{F}. \tag{9.2.6}$$

We see from (9.2.3) that $\frac{D(T)N(T)}{N(0)}$ is the random variable $Z(T)$ appearing in (5.4.1) of the multidimensional Girsanov Theorem if we replace $\Theta_j(t)$ by

$-\nu_j(t)$ for $j = 1, \ldots, m$. Here we are using the probability measure $\widetilde{\mathbb{P}}$ in place of $\mathbb{P}$ in Theorem 5.4.1 and using the $d$-dimensional Brownian motion $(\widetilde{W}_1(t), \ldots, \widetilde{W}_d(t))$ under $\widetilde{\mathbb{P}}$ in place of the $d$-dimensional Brownian motion $(W_1(t), \ldots, W_d(t))$ under $\mathbb{P}$.

With these replacements, Theorem 5.4.1 implies that, under $\widetilde{\mathbb{P}}^{(N)}$, the process $\widetilde{W}^{(N)}(t) = (\widetilde{W}_1^{(N)}(t), \ldots, \widetilde{W}_d^{(N)}(t))$ is a $d$-dimensional Brownian motion. In particular, under $\widetilde{\mathbb{P}}^{(N)}$, the Brownian motions $\widetilde{W}_1^{(N)}, \ldots \widetilde{W}_d^{(N)}$ are independent of one another. The expected value of an arbitrary random variable $X$ under $\widetilde{\mathbb{P}}^{(N)}$ can be computed by the formula

$$\widetilde{\mathbb{E}}^{(N)} X = \frac{1}{N(0)} \widetilde{\mathbb{E}}\big[X D(T) N(T)\big]. \tag{9.2.7}$$

More generally,

$$\frac{D(t)N(t)}{N(0)} = \widetilde{\mathbb{E}}\left[\frac{D(T)N(T)}{N(0)}\middle| \mathcal{F}(t)\right], \quad 0 \leq t \leq T,$$

is the Radon-Nikodým derivative process $Z(t)$ in the Theorem 5.4.1, and Lemma 5.2.2 implies that for $0 \leq s \leq t \leq T$ and $Y$ an $\mathcal{F}(t)$-measurable random variable,

$$\widetilde{\mathbb{E}}^{(N)}[Y | \mathcal{F}(s)] = \frac{1}{D(s)N(s)} \widetilde{\mathbb{E}}\big[Y D(t) N(t) \big| \mathcal{F}(s)\big]. \tag{9.2.8}$$

**Theorem 9.2.2 (Change of risk-neutral measure).** *Let $S(t)$ and $N(t)$ be the prices of two assets denominated in a common currency, and let $\sigma(t) = (\sigma_1(t), \ldots, \sigma_d(t))$ and $\nu(t) = (\nu_1(t), \ldots, \nu_d(t))$ denote their respective volatility vector processes:*

$$d\big(D(t)S(t)\big) = D(t)S(t)\sigma(t) \cdot d\widetilde{W}(t), \quad d\big(D(t)N(t)\big) = D(t)N(t)\nu(t) \cdot d\widetilde{W}(t).$$

*Take $N(t)$ as the numéraire, so the price of $S(t)$ becomes $S^{(N)}(t) = \frac{S(t)}{N(t)}$. Under the measure $\widetilde{\mathbb{P}}^{(N)}$, the process $S^{(N)}(t)$ is a martingale. Moreover,*

$$dS^{(N)}(t) = S^{(N)}(t)\,[\sigma(t) - \nu(t)] \cdot d\widetilde{W}^{(N)}(t). \tag{9.2.9}$$

*Remark 9.2.3.* Equation (9.2.9) says that the volatility vector of $S^{(N)}(t)$ is the difference of the volatility vectors of $S(t)$ and $N(t)$. In particular, after the change of numéraire, the price of the numéraire becomes identically 1,

$$N^{(N)}(t) = \frac{N(t)}{N(t)} = 1,$$

and this has zero volatility vector:

$$dN^{(N)}(t) = N^{(N)}(t)[\nu(t) - \nu(t)] \cdot d\widetilde{W}^{(N)}(t) = 0.$$

We are not saying that volatilities subtract when we change the numéraire. We are saying that volatility *vectors* subtract. The process $N(t)$ in Theorem 9.2.2 has the stochastic differential representation (9.2.1), which we may rewrite as

$$dN(t) = R(t)N(t)\,dt + \|\nu(t)\|N(t)dB^N(t), \tag{9.2.10}$$

where

$$B^N(t) = \int_0^t \sum_{j=1}^d \frac{\nu_j(u)}{\|\nu(u)\|}\,d\widetilde{W}_u(t).$$

According to Lévy's Theorem, Theorem 4.6.4, $B^N(t)$ is a one-dimensional Brownian motion. From (9.2.10), we see that the volatility (not the volatility vector) of $N(t)$ is $\|\nu(t)\|$. Similarly, the volatility of $S(t)$ in Theorem 9.2.2 is $\|\sigma(t)\|$. Application of the same argument to equation (9.2.9) shows that the volatility of $S^{(N)}(t)$ is $\|\sigma(t)-\nu(t)\|$. This is not the difference of the volatilities $\|\sigma(t)\| - \|\nu(t)\|$ unless the volatility vector $\sigma(t)$ is a positive multiple of the volatility vector $\nu(t)$.

*Remark 9.2.4.* If we take the money market account as the numéraire in Theorem 9.2.2 (i.e., $N(t) = M(t) = \frac{1}{D(t)}$), then we have $d\big(D(t)N(t)\big) = 0$. The volatility vector for the money market account is $\nu(t) = 0$, and the volatility vector for an asset $S^{(N)}(t)$ denominated in units of money market account is the same as the volatility vector of the asset denominated in units of currency. Discounting an asset using the money market account does not affect its volatility vector.

*Remark 9.2.5.* Theorem 9.2.2 is a special case of a more general result. Whenever $M_1(t)$ and $M_2(t)$ are martingales under a measure $\mathbb{P}$, $M_2(0) = 1$, and $M_2(t)$ takes only positive values, then $M_1(t)/M_2(t)$ is a martingale under the measure $\mathbb{P}^{(M_2)}$ defined by

$$\mathbb{P}^{(M_2)}(A) = \int_A M_2(T)\,d\mathbb{P}.$$

See Exercise 9.1.

PROOF OF THEOREM 9.2.2: We have

$$D(t)S(t) = S(0)\exp\left\{\int_0^t \sigma(u)\cdot d\widetilde{W}(u) - \frac{1}{2}\int_0^t \|\sigma(u)\|^2\,du\right\},$$

$$D(t)N(t) = N(0)\exp\left\{\int_0^t \nu(u)\cdot d\widetilde{W}(u) - \frac{1}{2}\int_0^t \|\nu(u)\|^2\,du\right\},$$

and hence

$$S^{(N)}(t) = \frac{S(0)}{N(0)}\exp\left\{\int_0^t \big(\sigma(u)-\nu(u)\big)\cdot d\widetilde{W}(u) - \frac{1}{2}\int_0^t \left(\|\sigma(u)\|^2 - \|\nu(u)\|^2\right)du\right\}.$$

To apply the Itô-Doeblin formula to this, we first define

$$X(t) = \int_0^t \big(\sigma(u) - \nu(u)\big) \cdot d\widetilde{W}(u) - \frac{1}{2}\int_0^t \left(\|\sigma(u)\|^2 - \|\nu(u)\|^2\right) du,$$

so that

$$\begin{aligned}
dX(t) &= \big(\sigma(t) - \nu(t)\big) \cdot d\widetilde{W}(t) - \frac{1}{2}\left(\|\sigma(t)\|^2 - \|\nu(t)\|^2\right) dt \\
&= \sum_{j=1}^{d} \big(\sigma_j(t) - \nu_j(t)\big)\, d\widetilde{W}_j(t) - \frac{1}{2}\sum_{j=1}^{d}\left(\sigma_j^2(t) - \nu_j^2(t)\right) dt, \\
dX(t)\, dX(t) &= \sum_{j=1}^{d} \big(\sigma_j(t) - \nu_j(t)\big)^2 dt \\
&= \sum_{j=1}^{d} \left(\sigma_j^2(t) - 2\sigma_j(t)\nu_j(t) + \nu_j^2(t)\right) dt \\
&= \|\sigma(t)\|^2\, dt - 2\sigma(t)\cdot\nu(t)\, dt + \|\nu(t)\|^2\, dt.
\end{aligned}$$

With $f(x) = \frac{S(0)}{N(0)}e^x$, we have $S^{(N)}(t) = f(X(t))$ and

$$\begin{aligned}
dS^{(N)}(t) &= df(X(t)) \\
&= f'(X)\, dX + \frac{1}{2} f''(X)\, dX\, dX \\
&= S^{(N)}\bigg[(\sigma - \nu)\cdot d\widetilde{W} - \frac{1}{2}\|\sigma\|^2\, dt + \frac{1}{2}\|\nu\|^2\, dt \\
&\qquad\qquad + \frac{1}{2}\|\sigma\|^2\, dt - \sigma\cdot\nu\, dt + \frac{1}{2}\|\nu\|^2\, dt\bigg] \\
&= S^{(N)}\left[(\sigma - \nu)\cdot d\widetilde{W} - \nu\cdot(\sigma - \nu)\, dt\right] \\
&= S^{(N)}(\sigma - \nu)\cdot(-\nu\, dt + d\widetilde{W}) \\
&= S^{(N)}(\sigma - \nu)\cdot d\widetilde{W}^{(N)}.
\end{aligned}$$

Since $\widetilde{W}^{(N)}(t)$ is a $d$-dimensional Brownian motion under $\widetilde{\mathbb{P}}^{(N)}$, the process $S^{(N)}(t)$ is a martingale under this measure. □

## 9.3 Foreign and Domestic Risk-Neutral Measures

### 9.3.1 The Basic Processes

We now apply the ideas of the previous section to a market with two currencies, which we call foreign and domestic. This model is driven by

$$W(t) = (W_1(t), W_2(t)),$$

a two-dimensional Brownian motion on some $(\Omega, \mathcal{F}, \mathbb{P})$. In particular, we are assuming that $W_1$ and $W_2$ are independent under $\mathbb{P}$. We begin with a stock whose price in domestic currency, $S(t)$, satisfies

$$dS(t) = \alpha(t)S(t)\,dt + \sigma_1(t)S(t)\,dW_1(t). \tag{9.3.1}$$

There is a domestic interest rate $R(t)$, which leads to a domestic money market account price and domestic discount process

$$M(t) = e^{\int_0^t R(u)du}, \quad D(t) = e^{-\int_0^t R(u)du}.$$

There is also a foreign interest rate $R^f(t)$, which leads to a foreign money market account price and foreign discount process

$$M^f(t) = e^{\int_0^t R^f(u)du}, \quad D^f(t) = e^{-\int_0^t R^f(u)du}.$$

Finally, there is an exchange rate $Q(t)$, which gives units of domestic currency per unit of foreign currency. We assume this satisfies

$$dQ(t) = \gamma(t)Q(t)\,dt + \sigma_2(t)Q(t)\left[\rho(t)\,dW_1(t) + \sqrt{1-\rho^2(t)}\,dW_2(t)\right]. \tag{9.3.2}$$

We define

$$W_3(t) = \int_0^t \rho(u)\,dW_1(u) + \int_0^t \sqrt{1-\rho^2(u)}\,dW_2(u). \tag{9.3.3}$$

By Lévy's Theorem, Theorem 4.6.4, $W_3(t)$ is a Brownian motion under $\widetilde{\mathbb{P}}$. We may rewrite (9.3.2) as

$$dQ(t) = \gamma(t)Q(t)\,dt + \sigma_2(t)Q(t)\,dW_3(t), \tag{9.3.4}$$

from which we see that $Q(t)$ has volatility $\sigma_2(t)$.

We assume $R(t)$, $R^f(t)$, $\sigma_1(t)$, $\sigma_2(t)$, and $\rho(t)$ are processes adapted to the filtration $\mathcal{F}(t)$ generated by the two-dimensional Brownian motion $W(t) = (W_1(t), W_2(t))$, and

$$\sigma_1(t) > 0, \quad \sigma_2(t) > 0, \quad -1 < \rho(t) < 1$$

for all $t$ almost surely. Because

$$\frac{dS(t)}{S(t)} \cdot \frac{dQ(t)}{Q(t)} = \rho(t)\sigma_1(t)\sigma_2(t)\,dt,$$

the process $\rho(t)$ is the instantaneous correlation between relative changes in $S(t)$ and $Q(t)$.

### 9.3.2 Domestic Risk-Neutral Measure

There are three assets that can be traded: the domestic money market account, the stock, and the foreign money market account. We shall price each of these in domestic currency and discount at the domestic interest rate. The result is the price of each of them in units of the domestic money market account. Under the domestic risk-neutral measure, all three assets priced in units of the domestic money market account must be martingales. We use this observation to find the domestic risk-neutral measure.

We note that the first asset, the domestic money market account, when priced in units of the domestic money market, has constant price 1. This is always a martingale, regardless of the measure being used.

The second asset, the stock, in units of the domestic money market account has price $D(t)S(t)$, and this satisfies the stochastic differential equation

$$d\big(D(t)S(t)\big) = D(t)S(t)\big[\big(\alpha(t) - R(t)\big)\,dt + \sigma_1(t)\,dW_1(t)\big]. \tag{9.3.5}$$

We would like to construct a process

$$\widetilde{W}_1(t) = \int_0^t \Theta_1(u)\,du + W_1(t)$$

that permits us to rewrite (9.3.5) as

$$d\big(D(t)S(t)\big) = \sigma_1(t)D(t)S(t)\,d\widetilde{W}_1(t). \tag{9.3.6}$$

Equating the right-hand sides of (9.3.5) and (9.3.6), we see that $\Theta_1(t)$ must be chosen to satisfy the *first market price of risk equation*

$$\sigma_1(t)\Theta_1(t) = \alpha(t) - R(t). \tag{9.3.7}$$

The third asset available in the domestic market is the following. One can invest in the foreign money market account and convert that investment to domestic currency. The value of the foreign money market account in domestic currency is $M^f(t)Q(t)$, and its discounted value is $D(t)M^f(t)Q(t)$. The differential of this price is

$$\begin{aligned} d\big(D(t)M^f(t)Q(t)\big) &= D(t)M^f(t)Q(t)\big[\big(R^f(t) - R(t) + \gamma(t)\big)\,dt \\ &\quad + \sigma_2(t)\rho(t)\,dW_1(t) + \sigma_2(t)\sqrt{1-\rho^2(t)}\,dW_2(t)\big]. \end{aligned} \tag{9.3.8}$$

One can derive this using the fact that

$$d\left(M^f(t)\right) = R^f(t)M^f(t)\,dt,$$

using Itô's product rule to compute

$$\begin{aligned} d\big(M^f(t)Q(t)\big) &= M^f(t)Q(t)\big[\big(R^f(t) + \gamma(t)\big)\,dt \\ &\quad + \sigma_2(t)\rho(t)\,dW_1(t) + \sigma_2(t)\sqrt{1-\rho^2(t)}\,dW_2(t)\big], \end{aligned}$$

and then using Itô's product rule again on $D(t) \cdot M^f(t)Q(t)$ to obtain (9.3.8). The mean rate of change of $Q(t)$ is $\gamma(t)$. When we inflate this at the foreign interest rate and discount it at the domestic interest rate, (9.3.8) shows that the mean rate of return changes to $R^f(t) - R(t) + \gamma(t)$. The volatility terms are unchanged.

In addition to the process $\widetilde{W}_1(t)$, we would like to construct a process

$$\widetilde{W}_2(t) = \int_0^t \Theta_2(u)\, du + W_2(t)$$

so that (9.3.8) can be written as

$$d\big(D(t)M^f(t)Q(t)\big) \\ = D(t)M^f(t)Q(t)\big[\sigma_2(t)\rho(t)\, d\widetilde{W}_1(t) + \sigma_2(t)\sqrt{1-\rho^2(t)}\, d\widetilde{W}_2(t)\big]. \quad (9.3.9)$$

Equating the right-hand sides of (9.3.9) and (9.3.8), we obtain the *second market price of risk equation*

$$\sigma_2(t)\rho(t)\Theta_1(t) + \sigma_2(t)\sqrt{1-\rho^2(t)}\,\Theta_2(t) = R^f(t) - R(t) + \gamma(t). \quad (9.3.10)$$

The market price of risk equations (9.3.7) and (9.3.10) determine processes $\Theta_1(t)$ and $\Theta_2(t)$. We can solve explicitly for these processes by first solving (9.3.7) for $\Theta_1(t)$, substituting this into (9.3.10), and then solving (9.3.10) for $\Theta_2(t)$. The conditions $\sigma_1(t) > 0$, $\sigma_2(t) > 0$, and $-1 < \rho(t) < 1$ are needed to do this.

The particular formulas for $\Theta_1(t)$ and $\Theta_2(t)$ are irrelevant. What matters is that the market price of risk equations have one and only one solution, and so there is a unique risk-neutral measure $\widetilde{\mathbb{P}}$ given by the multidimensional Girsanov Theorem. Under this measure, $\widetilde{W}(t) = (\widetilde{W}_1(t), \widetilde{W}_2(t))$ is a two-dimensional Brownian motion and the processes 1, $D(t)S(t)$, and $D(t)M^f(t)Q(t)$ are martingales. In the spirit of (9.3.3), we may also define

$$\widetilde{W}_3(t) = \int_0^t \rho(u)\, d\widetilde{W}_1(u) + \int_0^t \sqrt{1-\rho^2(u)}\, d\widetilde{W}_2(t). \quad (9.3.11)$$

Then $\widetilde{W}_3(t)$ is a Brownian motion under $\widetilde{\mathbb{P}}$, and

$$d\widetilde{W}_1(t)\, d\widetilde{W}_3(t) = \rho(t)\, dt, \quad d\widetilde{W}_2(t)\, d\widetilde{W}_3(t) = \sqrt{1-\rho^2(t)}\, dt. \quad (9.3.12)$$

We can write the price processes 1, $D(t)S(t)$ and $D(t)M^f(t)Q(t)$ in undiscounted form by multiplying them by $M(t) = \frac{1}{D(t)}$ and using the formula $dM(t) = R(t)M(t)\, dt$ and Itô's product rule. This leads to the formulas

$$dM(t) = R(t)M(t)\, dt, \quad (9.3.13)$$

$$dS(t) = S(t)\big[R(t)\, dt + \sigma_1(t)\, d\widetilde{W}_1(t)\big], \quad (9.3.14)$$

$$d\big(M^f(t)Q(t)\big) = M^f(t)Q(t)\big[R(t)\, dt + \sigma_2(t)\rho(t)\, d\widetilde{W}_1(t) \\ + \sigma_2(t)\sqrt{1-\rho^2(t)}\, d\widetilde{W}_2(t)\big] \\ = M^f(t)Q(t)\big[R(t)\, dt + \sigma_2(t)\, d\widetilde{W}_3(t)\big]. \quad (9.3.15)$$

All these price processes have mean rate of return $R(t)$ under the domestic risk-neutral measure $\widetilde{\mathbb{P}}$. We constructed the domestic risk-neutral measure so this is the case.

We may multiply $M^f(t)Q(t)$ by $D^f(t)$ and use Itô's product rule again to obtain

$$\begin{aligned} dQ(t) &= Q(t)\big[\big(R(t) - R^f(t)\big)\,dt + \sigma_2(t)\rho(t)\,d\widetilde{W}_1(t) + \sigma_2(t)\sqrt{1-\rho^2(t)}\,d\widetilde{W}_2(t)\big] \\ &= Q(t)\big[\big(R(t) - R^f(t)\big)\,dt + \sigma_2(t)\,d\widetilde{W}_3(t)\big]. \end{aligned} \tag{9.3.16}$$

Under the domestic risk-neutral measure, the mean rate of change of the exchange rate is the difference between the domestic and foreign interest rates $R(t) - R^f(t)$. In particular, it is not $R(t)$, as would be the case for an asset. If one regards the exchange rate as an asset (i.e., hold a unit of foreign currency whose value is always $Q(t)$), then it is a *dividend-paying* asset. The unit of foreign currency can and should be invested in the foreign money market, and this pays out a continuous dividend at rate $R^f(t)$. If this dividend is reinvested in the foreign money market, then we get the asset in (9.3.15), which has mean rate of return $R(t)$; if the dividend is not reinvested, then the rate of return is reduced by $R^f(t)$ and we have (9.3.16) (cf. (5.5.6)).

It is important to note that (9.3.16) tells us about the mean rate of change of the exchange rate under the domestic risk-neutral measure. Under the actual probability measure $\mathbb{P}$, the mean rate of change of the exchange rate can be anything. There are no restrictions on the process $\gamma(t)$ in (9.3.2).

### 9.3.3 Foreign Risk-Neutral Measure

In this model, we have three assets: the domestic money market account, the stock, and the foreign money market account. We list these assets across the top of Figure 9.3.1, and down the side of the figure we list the four ways of denominating them.

In the previous subsection, we constructed the domestic risk-neutral measure $\widetilde{\mathbb{P}}$ under which the three entries in the second line of Figure 9.3.1 are martingales. In this subsection, we construct the foreign risk-neutral measure under which the entries in the fourth line are martingales. (We cannot make all the entries in the first line be martingales because every path of the process $M(t)$ is increasing, and thus this process is not a martingale under any measure. The same applies to the entries in the third line, which contains the increasing process $M^f(t)$.)

We observe that the fourth line in Figure 9.3.1 is obtained by dividing each entry of the second line by $D(t)M^f(t)Q(t)$. In other words, to find the foreign risk-neutral measure, we take the foreign money market account as the numéraire. Its value at time $t$, denominated in units of the domestic money market account, is $D(t)M^f(t)Q(t)$, and denominated in units of domestic currency, it is $M^f(t)Q(t)$. The differential of $M^f(t)Q(t)$ is given in (9.3.15), and from that formula we see that its volatility vector is

$$(\nu_1(t), \nu_2(t)) = \left(\sigma_2(t)\rho(t), \sigma_2(t)\sqrt{1-\rho^2(t)}\right),$$

the same as the volatility vector of $Q(t)$.

| | Domestic money market | Stock | Foreign money market |
|---|---|---|---|
| Domestic currency | $M(t)$ | $S(t)$ | $M^f(t)Q(t)$ |
| Domestic money market | $1$ | $D(t)S(t)$ | $D(t)M^f(t)Q(t)$ |
| Foreign currency | $M(t)/Q(t)$ | $S(t)/Q(t)$ | $M^f(t)$ |
| Foreign money market | $M(t)D^f(t)/Q(t)$ | $D^f(t)S(t)/Q(t)$ | $1$ |

**Fig. 9.3.1.** Prices under different numéraires.

According to Theorem 9.2.2, the risk-neutral measure associated with the numéraire $M^f(t)Q(t)$ is given by

$$\widetilde{\mathbb{P}}^f(A) = \frac{1}{Q(0)} \int_A D(T)M^f(T)Q(T)\, d\widetilde{\mathbb{P}} \text{ for all } A \in \mathcal{F}, \tag{9.3.17}$$

where we have used the fact that $D(0) = M^f(0) = 1$. Furthermore, the process $\widetilde{W}^f(t) = (\widetilde{W}_1^f(t), \widetilde{W}_2^f(t))$ given by

$$\widetilde{W}_1^f(t) = -\int_0^t \sigma_2(u)\rho(u)\, du + \widetilde{W}_1(t), \tag{9.3.18}$$

$$\widetilde{W}_2^f(t) = -\int_0^t \sigma_2(u)\sqrt{1-\rho^2(u)}\, du + \widetilde{W}_2(t), \tag{9.3.19}$$

is a two-dimensional Brownian motion under $\widetilde{\mathbb{P}}^f$. We call $\widetilde{\mathbb{P}}^f$ the *foreign risk-neutral measure*. Following (9.3.11), we may also define

$$\begin{aligned}
\widetilde{W}_3^f(t) &= \int_0^t \rho(u)\, d\widetilde{W}_1^f(u) + \int_0^t \sqrt{1-\rho^2(u)}\, d\widetilde{W}_2^f(t)\\
&= \int_0^t \rho(u)\big(-\sigma_2(u)\rho(u)\, du + d\widetilde{W}_1(u)\big)\\
&\qquad + \int_0^t \sqrt{1-\rho^2(u)}\big(-\sigma_2(u)\sqrt{1-\rho^2(u)} + d\widetilde{W}_2(u)\big)\\
&= -\int_0^t \sigma_2(u)\, du + \int_0^t \big(\rho(u)\, d\widetilde{W}_1(u) + \sqrt{1-\rho^2(u)}\, d\widetilde{W}_2(u)\big)\\
&= -\int_0^t \sigma_2(u)\, du + \widetilde{W}_3(t).
\end{aligned} \tag{9.3.20}$$

Then $\widetilde{W}_3^f(t)$ is a Brownian motion under $\widetilde{\mathbb{P}}^f$, and

$$d\widetilde{W}_1^f(t)\,d\widetilde{W}_3^f(t) = \rho(t)\,dt, \quad d\widetilde{W}_2^f(t)\,d\widetilde{W}_3^f(t) = \sqrt{1-\rho^2(t)}\,dt. \tag{9.3.21}$$

Instead of relying on Theorem 9.2.2, one can verify directly by Itô calculus that the first two entries in the last row of Figure 9.3.1 are martingales under $\widetilde{\mathbb{P}}^f$ (the third entry, 1, is obviously a martingale). One can verify by direct computation that

$$\begin{aligned} d\left(\frac{M(t)D^f(t)}{Q(t)}\right) &= \frac{M(t)D^f(t)}{Q(t)}\Big[-\sigma_2(t)\rho(t)\,d\widetilde{W}_1^f(t) \\ &\qquad -\sigma_2(t)\sqrt{1-\rho^2(t)}\,d\widetilde{W}_2^f(t)\Big] \\ &= -\frac{M(t)D^f(t)}{Q(t)}\sigma_2(t)\,d\widetilde{W}_3^f(t), \end{aligned} \tag{9.3.22}$$

$$\begin{aligned} d\left(\frac{D^f(t)S(t)}{Q(t)}\right) &= \frac{D^f(t)S(t)}{Q(t)}\Big[\big(\sigma_1(t)-\sigma_2(t)\rho(t)\big)\,d\widetilde{W}_1^f(t) \\ &\qquad -\sigma_2(t)\sqrt{1-\rho^2(t)}\,d\widetilde{W}_2^f(t)\Big] \\ &= \frac{D^f(t)S(t)}{Q(t)}\Big[\sigma_1(t)\,d\widetilde{W}_1^f(t)-\sigma_2(t)\,d\widetilde{W}_3^f(t)\Big], \end{aligned} \tag{9.3.23}$$

Because $\widetilde{W}_1^f(t)$, $\widetilde{W}_2^f(t)$, and $\widetilde{W}_3^f(t)$ are Brownian motions under $\widetilde{\mathbb{P}}^f$, the processes above are martingales under this measure. The Brownian motions $\widetilde{W}_1^f(t)$ and $\widetilde{W}_2^f(t)$ are independent under $\widetilde{\mathbb{P}}^f$, whereas $\widetilde{W}_3^f(t)$ has instantaneous correlations with $\widetilde{W}_1^f(t)$ and $\widetilde{W}_2^f(t)$ given by (9.3.21).

### 9.3.4 Siegel's Exchange Rate Paradox

In (9.3.16), we saw that under the domestic risk-neutral measure $\widetilde{\mathbb{P}}$, the mean rate of change for the exchange rate $Q(t)$ is $R(t)-R^f(t)$. From the foreign perspective, the exchange rate is $\frac{1}{Q(t)}$, and one should expect the mean rate of change of $\frac{1}{Q(t)}$ to be $R^f(t)-R(t)$. In other words, one might expect that if the average rate of change of the dollar against the euro is 5%, then the average rate of change of the euro against the dollar should be $-5\%$. This turns out not to be as straight forward as one might expect because of the convexity of the function $f(x)=\frac{1}{x}$.

For example, an exchange rate of 0.90 euros to the dollar would be 1.1111 dollars to the euro. If the dollar price of euro falls by 5%, then price of the euro would be only $0.95 \times 1.1111 = 1.0556$ dollars. This is an exchange rate of 0.9474 euros to the dollar. The change from 0.90 euros to the dollar to 0.9474 euros to the dollar is a 5.26% increase in the euro price of the dollar, not a 5% increase.

The convexity effect seen in the previous paragraph makes itself felt when we compute the differential of $\frac{1}{Q(t)}$. We take $f(x)=\frac{1}{x}$ so that $f'(x)=-\frac{1}{x^2}$ and $f''(x)=\frac{2}{x^3}$. Using (9.3.16), we obtain

$$\begin{aligned}
d\left(\frac{1}{Q}\right) &= df(Q)\\
&= f'(Q)\,dQ + \frac{1}{2}f''(Q)\,dQ\,dQ\\
&= \frac{1}{Q}\left[(R^f - R)\,dt - \sigma_2 d\widetilde{W}_3\right] + \frac{1}{Q}\sigma_2^2\,d\widetilde{W}_3\,d\widetilde{W}_3\\
&= \frac{1}{Q(t)}\left[(R^f - R + \sigma_2^2)\,dt - \sigma_2 d\widetilde{W}_3\right]. \qquad (9.3.24)
\end{aligned}$$

The mean rate of change under the domestic risk-neutral measure is $R^f(t) - R(t) + \sigma_2^2$, not $R^f(t) - R(t)$.

However, the asymmetry introduced by the convexity of $f(x) = \frac{1}{x}$ is resolved if we switch to the foreign risk-neutral measure, which is the appropriate one for derivative security pricing in the foreign currency. First recall the relationship (9.3.20)

$$d\widetilde{W}_3^f(t) = -\sigma_2(t)\,dt + d\widetilde{W}_3(t).$$

In terms of $\widetilde{W}_3^f(t)$, we may rewrite (9.3.24) as

$$d\left(\frac{1}{Q}\right) = \left(\frac{1}{Q}\right)\left[(R^f - R)\,dt - \sigma_2 d\widetilde{W}_3^f\right]. \qquad (9.3.25)$$

Under the foreign risk-neutral measure, the mean rate of change for $\frac{1}{Q}$ is $R^f - R$, as expected.

Under the actual probability measure $\mathbb{P}$, however, the asymmetry remains. When we begin with (9.3.4), which shows the mean rate of change of the exchange rate to be $\gamma(t)$ under $\mathbb{P}$ and is repeated below as (9.3.26), and then use the Itô-Doeblin formula as we did in (9.3.24), we obtain the formula (9.3.27) below:

$$dQ(t) = \gamma(t)Q(t)\,dt + \sigma_2(t)Q(t)\,dW_3(t), \qquad (9.3.26)$$

$$d\left(\frac{1}{Q(t)}\right) = \frac{1}{Q(t)}\big(-\gamma(t) + \sigma_2^2(t)\big)dt - \frac{1}{Q(t)}\sigma_2(t)\,dW_3(t). \qquad (9.3.27)$$

Both $Q$ and $\frac{1}{Q}$ have the same volatility. (A change of sign in the volatility does not affect volatility because Brownian motion is symmetric.) However, the mean rates of change of $Q$ and $\frac{1}{Q}$ are not negatives of one another.

### 9.3.5 Forward Exchange Rates

We assume in this subsection that the domestic and foreign interest rates are constant and denote these constants by $r$ and $r^f$, respectively. Recall that $Q$ is units of domestic currency per unit of foreign currency. The exchange rate from the *domestic* viewpoint is governed by the stochastic differential equation (9.3.16)

$$dQ(t) = Q(t)\left[(r - r^f)\,dt + \sigma_2(t)\rho(t)\,d\widetilde{W}_1(t) + \sigma_2(t)\sqrt{1-\rho^2(t)}\,d\widetilde{W}_2(t)\right].$$

Therefore

$$e^{-(r-r^f)t}Q(t)$$

is a martingale under $\widetilde{\mathbb{P}}$, the *domestic* risk-neutral measure.

At time zero, the (domestic currency) forward price $F$ for a unit of foreign currency, to be delivered at time $T$, is determined by the equation

$$\widetilde{\mathbb{E}}\left[e^{-rT}(Q(T) - F)\right] = 0.$$

The left-hand side is the risk-neutral pricing formula applied to the derivative security that pays $Q(T)$ in exchange for $F$ at time $T$. Setting this equal to zero determines the forward price. We may solve this equation for $F$ by observing that it implies

$$e^{-rT}F = \widetilde{\mathbb{E}}\left[e^{-rT}Q(T)\right] = e^{-r^fT}\widetilde{\mathbb{E}}\left[e^{-(r-r^f)T}Q(T)\right] = e^{-r^fT}Q(0),$$

which gives the $T$-forward (domestic per unit of foreign) exchange rate

$$F = e^{(r-r^f)T}Q(0).$$

The exchange rate from the foreign viewpoint is given by the stochastic differential equation (9.3.25)

$$d\left(\frac{1}{Q(t)}\right)$$
$$= \left(\frac{1}{Q(t)}\right)\left[(r^f - r)\,dt - \sigma_2(t)\rho(t)\,d\widetilde{W}_1^f(t) - \sigma_2(t)\sqrt{1-\rho^2(t)}\,d\widetilde{W}_2^f(t)\right].$$

Therefore,

$$e^{-(r^f-r)t}\frac{1}{Q(t)}$$

is a martingale under $\widetilde{\mathbb{P}}^f$, the *foreign* risk-neutral measure.

At time zero, the (foreign currency) forward price $F^f$ for a unit of domestic currency to be delivered at time $T$ is determined by the equation

$$\widetilde{\mathbb{E}}^f\left[e^{-r^fT}\left(\frac{1}{Q(T)} - F^f\right)\right] = 0.$$

The left-hand side is the risk-neutral pricing formula applied to the derivative security that pays $\frac{1}{Q(T)}$ in exchange for $F^f$ (both denominated in foreign currency) at time $T$. Setting this equal to zero determines the forward price. We may solve this equation for $F^f$ by observing that it implies

$$e^{-r^fT}F^f = \widetilde{\mathbb{E}}^f\left[e^{-r^fT}\frac{1}{Q(T)}\right] = e^{-rT}\widetilde{\mathbb{E}}^f\left[e^{-(r^f-r)T}\frac{1}{Q(T)}\right] = e^{-rT}\frac{1}{Q(0)},$$

which gives the $T$-forward (foreign per unit of domestic) exchange rate

$$F^f = e^{(r^f-r)T}\frac{1}{Q(0)} = \frac{1}{F}.$$

### 9.3.6 Garman-Kohlhagen Formula

In this section, we assume the domestic and foreign interest rates $r$ and $r^f$ and the volatility $\sigma_2$ are constant. Consider a call on a unit of foreign currency whose payoff in domestic currency is $(Q(T)-K)^+$. At time zero, the value of this is

$$\widetilde{\mathbb{E}}e^{-rT}(Q(T)-K)^+.$$

In this case, (9.3.16) becomes

$$dQ(t) = Q(t)\left[\left(r-r^f\right)dt + \sigma_2\, d\widetilde{W}_3(t)\right],$$

from which we conclude that

$$Q(T) = Q(0)\exp\left\{\sigma_2\widetilde{W}_3(T) + \left(r - r^f - \frac{1}{2}\sigma_2^2\right)T\right\}.$$

Define

$$Y = -\frac{\widetilde{W}_3(T)}{\sqrt{T}},$$

so $Y$ is a standard normal random variable under $\widetilde{\mathbb{P}}$. Then the price of the call is

$$\begin{aligned}
&\widetilde{\mathbb{E}}e^{-rT}\left(Q(T)-K\right)^+\\
&= \widetilde{\mathbb{E}}\left[e^{-rT}\left(Q(0)\exp\left\{-\sigma_2\sqrt{T}Y + \left(r - r^f - \frac{1}{2}\sigma_2^2\right)T\right\} - K\right)^+\right].
\end{aligned}$$

This expression is just like (5.5.10) with $\tau = T$, with $Q(0)$ in place of $x$, and with $r^f$ in place of the dividend rate $a$. According to (5.5.12), the call price is

$$\widetilde{\mathbb{E}}e^{-rT}\left(Q(T)-K\right)^+ = e^{-r^f T}Q(0)N(d_+) - e^{-rT}KN(d_-), \tag{9.3.28}$$

where

$$d_\pm = \frac{1}{\sigma_2\sqrt{T}}\left[\log\frac{Q(0)}{K} + \left(r - r^f \pm \frac{1}{2}\sigma_2^2\right)T\right]$$

and $N$ is the cumulative standard normal distribution function. Equation (9.3.28) is called the *Garman-Kohlhagen formula.*

### 9.3.7 Exchange Rate Put–Call Duality

In this subsection, we develop a relationship between a call on domestic currency, denominated in foreign currency, and a put on a foreign currency, denominated in the domestic currency.

Recall the numéraire $M^f(t)Q(t)$, which is the domestic price of the foreign money market account. The Radon-Nikodým derivative of the foreign risk-neutral measure with respect to the domestic risk-neutral measure is (see (9.3.17))

$$\frac{d\widetilde{\mathbb{P}}^f}{d\widetilde{\mathbb{P}}} = \frac{D(T)M^f(T)Q(T)}{Q(0)}.$$

Thus, for any random variable $X$,

$$\widetilde{\mathbb{E}}^f X = \widetilde{\mathbb{E}}\left[\frac{D(T)M^f(T)Q(T)}{Q(0)} X\right].$$

A call struck at $K$ on a unit of domestic currency denominated in the foreign currency pays off $\left(\frac{1}{Q(T)} - K\right)^+$ units of foreign currency at expiration time $T$. The foreign currency value of this at time zero, which is the foreign risk-neutral expected value of the discounted payoff, is

$$\begin{aligned}
&\widetilde{\mathbb{E}}^f\left[D^f(T)\left(\frac{1}{Q(T)} - K\right)^+\right]\\
&\quad = \widetilde{\mathbb{E}}\left[\frac{D(T)M^f(T)Q(T)}{Q(0)} \cdot D^f(T)\left(\frac{1}{Q(T)} - K\right)^+\right]\\
&\quad = \frac{1}{Q(0)}\widetilde{\mathbb{E}}\left[D(T)\big(1 - KQ(T)\big)^+\right]\\
&\quad = \frac{K}{Q(0)}\widetilde{\mathbb{E}}\left[D(T)\left(\frac{1}{K} - Q(T)\right)^+\right].
\end{aligned}$$

This is the time-zero value in domestic currency of $\frac{K}{Q(0)}$ puts on the foreign exchange rate. More specifically, a put struck at $\frac{1}{K}$ on a unit of foreign currency denominated in the domestic currency pays off $\left(\frac{1}{K} - Q(T)\right)^+$ units of domestic currency at expiration time $T$. The domestic currency value of this put at time zero, which is the domestic risk-neutral expected value of the discounted payoff, is

$$\widetilde{\mathbb{E}}\left[D(T)\left(\frac{1}{K} - Q(T)\right)^+\right].$$

The call we began with is worth $\frac{K}{Q(0)}$ of these puts.

The foreign currency price of the put struck at $\frac{1}{K}$ on a unit of foreign currency is

$$\frac{1}{Q(0)}\widetilde{\mathbb{E}}\left[D(T)\left(\frac{1}{K} - Q(T)\right)^+\right].$$

The call we began with has a value $K$ times this amount. When we denominate both the call and the put this way in foreign currency, we can then understand the final result. Indeed, we have seen that the option to exchange $K$ units of foreign currency for one unit of domestic currency (the call) is the same as $K$ options to exchange $\frac{1}{K}$ units of domestic currency for one unit of foreign currency (the put). Stated in this way, the result is almost obvious.

## 9.4 Forward Measures

Although there may be multiple Brownian motions driving the model of this section, in order to simplify the notation, we assume in this section that there is only one. It is not difficult to rederive the results presented here under the assumption that there are $d$ Brownian motions.

### 9.4.1 Forward Price

We recall the discussion of Section 5.6.1. Consider a zero-coupon bond that pays 1 unit of currency (all currency is domestic in this section) at maturity $T$. According to the risk-neutral pricing formula, the value of this bond at time $t \in [0, T]$ is

$$B(t,T) = \frac{1}{D(t)}\widetilde{\mathbb{E}}\big[D(T)\big|\mathcal{F}(t)\big]. \tag{9.4.1}$$

In particular, $B(T,T) = 1$.

Consider now an asset whose price denominated in currency is $S(t)$. A *forward contract* that delivers one share of this asset at time $T$ in exchange for $K$ has a time-$T$ payoff of $S(T) - K$. According to the risk-neutral pricing formula, the value of this contract at earlier times $t$ is

$$V(t) = \frac{1}{D(t)}\widetilde{\mathbb{E}}\big[D(T)(S(T)-K)\big|\mathcal{F}(t)\big].$$

Because $D(t)S(t)$ is a martingale under $\widetilde{\mathbb{P}}$, this reduces to

$$V(t) = S(t) - \frac{K}{D(t)}\widetilde{\mathbb{E}}\big[D(T)\big|\mathcal{F}(t)\big] = S(t) - KB(t,T). \tag{9.4.2}$$

The *$T$-forward price* $\mathrm{For}_S(t,T)$ at time $t \in [0,T]$ of an asset is the value of $K$ that causes the value of the forward contract in (9.4.2) to be zero:

$$\mathrm{For}_S(t,T) = \frac{S(t)}{B(t,T)}. \tag{9.4.3}$$

### 9.4.2 Zero-Coupon Bond as Numéraire

A zero-coupon bond is an asset, and therefore the discounted bond price $D(t)B(t,T)$ must be a martingale under the risk-neutral measure $\widetilde{\mathbb{P}}$. According to Theorem 9.2.1, there is a volatility process $\sigma^*(t,T)$ for the bond (a process in $t$; $T$ is fixed) such that

$$d\big(D(t)B(t,T)\big) = -\sigma^*(t,T)D(t)B(t,T)\,d\widetilde{W}(t). \tag{9.4.4}$$

In (9.4.4), we write $-\sigma^*(t,T)$ rather than $\sigma^*(t,T)$ in order to be consistent with the notation used in our discussion of the Heath-Jarrow-Morton model in Chapter 10. This has no effect on the distribution of the bond price process since we could just as well write (9.4.4) as

$$d\big(D(t)B(t,T)\big) = \sigma^*(t,T)D(t)B(t,T)\,d\big(-\widetilde{W}(t)\big),$$

and, just like $\widetilde{W}(t)$, the process $-\widetilde{W}(t)$ is a Brownian motion under $\widetilde{\mathbb{P}}$.

**Definition 9.4.1.** *Let $T$ be a fixed maturity date. We define the $T$-forward measure $\widetilde{\mathbb{P}}^T$ by*

$$\widetilde{\mathbb{P}}^T(A) = \frac{1}{B(0,T)}\int_A D(T)\,d\widetilde{\mathbb{P}} \text{ for all } A \in \mathcal{F}. \tag{9.4.5}$$

The $T$-forward measure corresponds to taking $N(t) = B(t,T)$ in (9.2.7) and (9.2.8). According to Theorem 9.2.2, the process

$$\widetilde{W}^T(t) = \int_0^t \sigma^*(u,T)\,du + \widetilde{W}(t)$$

is a Brownian motion under $\widetilde{\mathbb{P}}^T$. Furthermore, under the $T$-forward measure, all assets denominated in units of the zero-coupon bond maturing at time $T$ are martingale. In other words,

*$T$-forward prices are martingales under the $T$-forward measure $\widetilde{\mathbb{P}}^T$.*

Furthermore, the volatility vector of the $T$-forward price of an asset is the difference between the volatility vector of the asset and the volatility vector of the $T$-maturity zero-coupon bond (see Remark 9.2.3).

The reason to introduce the $T$-forward measure is that it often simplifies the risk-neutral pricing formula. According to that formula, the value at time $t$ of a contract that pays $V(T)$ at a later time $T$ is

$$V(t) = \frac{1}{D(t)}\widetilde{\mathbb{E}}\big[D(T)V(T)\big|\mathcal{F}(t)\big]. \tag{9.4.6}$$

The computation of the right-hand side of this formula requires that we know something about the dependence between the discount factor $D(T)$ and the payoff $V(T)$ of the derivative security. Especially when the derivative security depends on the interest rate, this can be difficult to model. However, according to (9.2.8) (with $t$ replacing $s$ and $T$ replacing $t$ in that formula), we have

$$\widetilde{\mathbb{E}}^T[V(T)|\mathcal{F}(t)] = \frac{1}{D(t)B(t,T)}\widetilde{\mathbb{E}}\big[D(T)V(T)\big|\mathcal{F}(t)\big] = \frac{1}{B(t,T)}V(t).$$

This gives us the simple formula

$$V(t) = B(t,T)\widetilde{\mathbb{E}}^T\big[V(T)\big|\mathcal{F}(t)\big]. \tag{9.4.7}$$

If we can find a simple model for the evolution of assets under the $T$-forward measure, we can use (9.4.7), in which we only need to estimate $V(T)$, instead of using (9.4.6), which requires us to estimate $D(T)V(T)$. We give an example of the power of this approach in the next subsection.

### 9.4.3 Option Pricing with a Random Interest Rate

The classical Black-Scholes-Merton option-pricing formula assumes a constant interest rate. For options on bonds and other interest-rate-dependent instruments, movements in the interest rate are critical. For these "fixed income" derivatives, the assumption of a constant interest rate is inappropriate.

In this section, we present a generalized Black-Scholes-Merton option-pricing formula that permits the interest rate to be random. The classical Black-Scholes-Merton assumption that the volatility of the underlying asset is constant is here replaced by the assumption that the volatility of the forward price of the underlying asset is constant. Because the forward price is a martingale under the forward measure, and $\widetilde{W}^T(t)$ is the Brownian motion used to drive asset prices under the forward measure, the assumption of constant volatility for the forward price is equivalent to the assumption

$$d\mathrm{For}_S(t,T) = \sigma \mathrm{For}_S(t,T)\, d\widetilde{W}^T(t), \tag{9.4.8}$$

where $\sigma$ is a constant. The bond maturity $T$ is chosen to coincide with the expiration time $T$ of the option.

**Theorem 9.4.2 (Black-Scholes-Merton option pricing with random interest rate).** *Let $S(t)$ be the price of an asset denominated in (domestic) currency, and assume the forward price of this asset satisfies (9.4.8) with a positive constant $\sigma$. The value at time $t \in [0,T]$ of a European call on this asset, expiring at time $T$ with strike price $K$, is*

$$V(t) = S(t)N(d_+(t)) - KB(t,T)N(d_-(t)), \tag{9.4.9}$$

*where the adapted processes $d_\pm(t)$ are given by*

$$d_\pm(t) = \frac{1}{\sigma\sqrt{T-t}}\left[\log\frac{\mathrm{For}_S(t,T)}{K} \pm \frac{1}{2}\sigma^2(T-t)\right]. \tag{9.4.10}$$

*Furthermore, a short position in the option can be hedged by holding $N(d_+(t))$ shares of the asset and shorting $KN(d_-(t))$ $T$-maturity zero-coupon bonds at each time $t$.*

*Remark 9.4.3.* If the interest rate is a constant $r$, then $B(t,T) = e^{-r(T-t)}$, $\mathrm{For}_S(t,T) = e^{r(T-t)}S(t)$, and this theorem reduces to the usual Black-Scholes-Merton formula and hedging strategy.

PROOF OF THEOREM 9.4.2: We prove formula (9.4.9) for $t = 0$. It is not difficult to modify the proof to account for general $t$.

We observe that $\mathrm{For}_S(0,T) = \frac{S(0)}{B(0,T)}$, and so the solution to (9.4.8) is

$$\mathrm{For}_S(t,T) = \frac{S(0)}{B(0,T)} \exp\left\{\sigma\widetilde{W}^T(t) - \frac{1}{2}\sigma^2 t\right\}. \tag{9.4.11}$$

For each $t$, this has a log-normal distribution under $\widetilde{\mathbb{P}}^T$, the measure under which $\widetilde{W}^T(t)$ is a Brownian motion.

We need one more change of measure. Suppose we take the asset price $S(t)$ to be the numéraire. In terms of this numéraire, the asset price is identically 1. The risk-neutral measure for this numéraire is given by

$$\widetilde{\mathbb{P}}^S(A) = \frac{1}{S(0)} \int_A D(T)S(T)\, d\widetilde{\mathbb{P}} \text{ for all } A \in \mathcal{F}.$$

Denominated in units of $S(t)$, the zero-coupon bond is

$$\frac{B(t,T)}{S(t)} = \frac{1}{\mathrm{For}_S(t,T)},\ 0 \le t \le T,$$

and, by Theorem 9.2.2, this is a martingale under $\widetilde{\mathbb{P}}^S$.

Indeed, we can compute the differential of $\frac{1}{\mathrm{For}_S(t,T)}$ using the Itô-Doeblin formula, the function $f(x) = \frac{1}{x}$, and (9.4.8). Since $f'(x) = -\frac{1}{x^2}$ and $f''(x) = \frac{2}{x^3}$, we have

$$\begin{aligned}
&d\left(\frac{1}{\mathrm{For}_S(t,T)}\right)\\
&= df(\mathrm{For}_S(t,T))\\
&= f'(\mathrm{For}_S(t,T))\, d\mathrm{For}_S(t,T) + \frac{1}{2} f''(\mathrm{For}_S(t,T))\, d\mathrm{For}_S(t,T)\, d\mathrm{For}_S(t,T)\\
&= -\frac{\sigma}{\mathrm{For}_S(t,T)}\, d\widetilde{W}^T(t) + \frac{\sigma^2}{\mathrm{For}_S(t,T)}\, dt\\
&= -\frac{\sigma}{\mathrm{For}_S(t,T)}\big( -\sigma\, dt + d\widetilde{W}^T\big).
\end{aligned} \tag{9.4.12}$$

Because we are guaranteed by Theorem 9.2.2 that $\frac{1}{\mathrm{For}_S(t,T)}$ is a martingale under $\widetilde{\mathbb{P}}^S$, we conclude that

$$\widetilde{W}^S(t) = -\sigma t + \widetilde{W}^T(t)$$

is a Brownian motion under $\widetilde{\mathbb{P}}^S$. We see also that $\frac{1}{\mathrm{For}_S(t,T)}$ has volatility $\sigma$. The solution to (9.4.12) is

$$\frac{1}{\mathrm{For}_S(t,T)} = \frac{B(0,T)}{S(0)} \exp\left\{-\sigma \widetilde{W}^S(t) - \frac{1}{2}\sigma^2 t\right\}. \tag{9.4.13}$$

For each $t$, this has a log-normal distribution under $\widetilde{\mathbb{P}}^S$, the measure under which $\widetilde{W}^S(t)$ is a Brownian motion.

At time zero, the value of a European call expiring at time $T$, according to the risk-neutral pricing formula, is

$$\begin{aligned}
V(0) &= \widetilde{\mathbb{E}}\left[D(T)(S(T)-K)^+\right]\\
&= \widetilde{\mathbb{E}}\left[D(T)S(T)\mathbb{I}_{\{S(T)>K\}}\right] - K\widetilde{\mathbb{E}}\left[D(T)\mathbb{I}_{\{S(T)>K\}}\right]\\
&= S(0)\widetilde{\mathbb{E}}\left[\frac{D(T)S(T)}{S(0)}\mathbb{I}_{\{S(T)>K\}}\right] - KB(0,T)\widetilde{\mathbb{E}}\left[\frac{D(T)}{B(0,T)}\mathbb{I}_{\{S(T)>K\}}\right]\\
&= S(0)\widetilde{\mathbb{P}}^S\{S(T)>K\} - KB(0,T)\widetilde{\mathbb{P}}^T\{S(T)>K\}\\
&= S(0)\widetilde{\mathbb{P}}^S\{\mathrm{For}_S(T,T)>K\} - KB(0,T)\widetilde{\mathbb{P}}^T\{\mathrm{For}_S(T,T)>K\}\\
&= S(0)\widetilde{\mathbb{P}}^S\left\{\frac{1}{\mathrm{For}_S(T,T)}<\frac{1}{K}\right\} - KB(0,T)\widetilde{\mathbb{P}}^T\{\mathrm{For}_S(T,T)>K\},
\end{aligned}$$

where in the next-to-last step we have used the fact that $\mathrm{For}_S(T,T) = S(T)$. Using the fact that $\widetilde{W}^S(T)$ is normal with mean zero and variance $T$ under $\widetilde{\mathbb{P}}^S$, we compute

$$\begin{aligned}
&\widetilde{\mathbb{P}}^S\left\{\frac{1}{\mathrm{For}_S(T,T)}<\frac{1}{K}\right\}\\
&\quad= \widetilde{\mathbb{P}}^S\left\{-\sigma\widetilde{W}^S(T) - \frac{1}{2}\sigma^2 T < \log\frac{S(0)}{KB(0,T)}\right\}\\
&\quad= \widetilde{\mathbb{P}}^S\left\{\frac{-\widetilde{W}^S(T)}{\sqrt{T}} < \frac{1}{\sigma\sqrt{T}}\left[\log\frac{S(0)}{KB(0,T)} + \frac{1}{2}\sigma^2 T\right]\right\}\\
&\quad= N(d_+(0)).
\end{aligned}$$

Using the fact that $\widetilde{W}^T(T)$ is normal with mean zero and variance $T$ under $\widetilde{\mathbb{P}}^T$, we obtain

$$\begin{aligned}
&\widetilde{\mathbb{P}}\{\mathrm{For}_S(T,T)>K\}\\
&\quad= \widetilde{\mathbb{P}}^T\left\{\sigma\widetilde{W}^T(T) - \frac{1}{2}\sigma^2 T > \log\frac{KB(0,T)}{S(0)}\right\}\\
&\quad= \widetilde{\mathbb{P}}^T\left\{\frac{\widetilde{W}^T(T)}{\sqrt{T}} > \frac{1}{\sigma\sqrt{T}}\log\left[\frac{KB(0,T)}{S(0)} + \frac{1}{2}\sigma^2 T\right]\right\}\\
&\quad= \widetilde{\mathbb{P}}^T\left\{-\frac{\widetilde{W}^T(T)}{\sqrt{T}} < \frac{1}{\sigma\sqrt{T}}\left[\log\frac{S(0)}{KB(0,T)} - \frac{1}{2}\sigma^2 T\right]\right\}\\
&\quad= N(d_-(0)).
\end{aligned}$$

This completes the proof of (9.4.9), at least for the case $t = 0$.

We now consider the hedge suggested by formula (9.4.9). It is easier to do this when we take the zero-coupon bond as the numéraire rather than when we use currency. Dividing (9.4.9) by $B(t,T)$, we obtain

$$\frac{V(t)}{B(t,T)} = \mathrm{For}_S(t,T)N(d_+(t)) - KN(d_-(t)). \tag{9.4.14}$$

This gives us the option price denominated in zero-coupon bonds. Suppose we hedge a short position in the option by holding $N(d_+(t))$ shares of the asset and shorting $KN(d_-(t))$ zero-coupon bonds at each time $t$. The value of this portfolio, denominated in units of zero-coupon bond, agrees with (9.4.14). To be sure this short option hedge works, however, we must verify that the portfolio just described is *self-financing*. In other words, we must be sure we do not need to infuse cash in order to maintain the positions just described. (A discussion related to this, passing from discrete to continuous time, is provided in Exercise 4.10 of Chapter 4.) The capital gains differential associated with this portfolio, again denominated in units of zero-coupon bond, is

$$N(d_+(t))\, d\mathrm{For}_S(t,T).$$

(When measuring wealth in units of zero-coupon bond, there is no capital gain from movements in the bond price.) The differential of the portfolio, according to Itô's formula, is

$$\begin{aligned} d\left(\frac{V(t)}{B(t,T)}\right) &= N(d_+(t))\, d\mathrm{For}_S(t,T) + \mathrm{For}_S(t,T)\, dN(d_+(t)) \\ &\quad + d\mathrm{For}_S(t,T)\, dN(d_+(t)) - K\, dN(d_-(t)). \end{aligned} \tag{9.4.15}$$

In order for the portfolio to be self-financing, we must have

$$\mathrm{For}_S(t,T)\, dN(d_+(t)) + d\mathrm{For}_S(t,T)\, dN(d_+(t)) - K\, dN(d_-(t)) = 0, \tag{9.4.16}$$

so that the change of value in the portfolio is entirely due to capital gains. The verification of (9.4.16) is Exercise 9.6. □

## 9.5 Summary

This chapter discusses the fact that when we change the units of account, the so-called *numéraire*, we must change the risk-neutral measure. Fortunately, the Radon-Nikodým derivative process needed to effect this change of measure is simple; it is the numéraire itself, discounted in order to be a martingale and normalized by its initial condition in order to have expected value 1. This is the content of Theorem 9.2.2.

In this chapter, we apply the change-of-numéraire idea in two cases: foreign exchange models and option pricing in the presence of a random interest rate. It was also used in the discussion of Asian options in Section 7.5.

In the context of foreign exchange models, we show that the mean rate of change of the exchange rate is the difference between the interest rates in the two economies *under the risk-neutral measure for the economy in which the exchange rate is being considered.* We show that one can derive other expected symmetries (e.g., the forward exchange rate in one currency is the reciprocal of the foreign exchange rate in the other currency), provided one is careful to use the appropriate risk-neutral measures.

When the interest rate is random, the classical Black-Scholes-Merton option-pricing formula does not apply. However, if one is willing to assume that the $T$-forward price of the underlying asset has constant volatility, then the price of a call expiring at time $T$ has a simple formula and a simple hedging strategy (Theorem 9.4.2). This fact is exploited to build *LIBOR models* in Section 10.4.

## 9.6 Notes

The model of foreign and domestic markets presented in this chapter is a simplification of one in Musiela and Rutkowski [126]. The model in [126], drawn from Amin and Jarrow [2], permits foreign and domestic interest rates to be random. The Garman-Kohlhagen formula of Subsection 9.3.6 is taken from Garman and Kohlhagen [68]. The option to exchange one risky asset for another, of which Subsection 9.3.7 is a special case, was studied by Margrabe [117].

Theorem 9.4.2, option pricing with a random interest rate, is taken from Geman, El Karoui, and Rochet [70]. It traces back at least to Geman [69] and Jamshidian [94], who observed that the forward price of an asset is its price when denominated in the numéraire of the zero-coupon bond maturing at the delivery date. Even earlier, Merton [122] proposed hedging European options by using a bond maturing on the option expiration date.

## 9.7 Exercises

**Exercise 9.1.** This exercise provides an alternate proof of the main assertion of Theorem 9.2.2.

(i) Use Lemma 5.2.2 to prove Remark 9.2.5.
(ii) Let $S(t)$ and $N(t)$ be prices of two assets, denominated in a common currency, and assume $N(t)$ is always strictly positive. Let $\widetilde{\mathbb{P}}$ be the risk-neutral measure under which the discounted asset prices $D(t)S(t)$ and $D(t)N(t)$ are martingales. Apply Remark 9.2.5 to show that $S^{(N)}(t) = \frac{S(t)}{N(t)}$ is a martingale under $\widetilde{\mathbb{P}}^{(N)}$ defined by (9.2.6).

**Exercise 9.2 (Portfolios under change of numéraire).** Consider two assets with prices $S(t)$ and $N(t)$ given by

$$S(t) = S(0)\exp\left\{\sigma\widetilde{W}(t) + \left(r - \frac{1}{2}\sigma^2\right)t\right\},$$
$$N(t) = N(0)\exp\left\{\nu\widetilde{W}(t) + \left(r - \frac{1}{2}\nu^2\right)t\right\},$$

where $\widetilde{W}(t)$ is a one-dimensional Brownian motion under the risk-neutral measure $\widetilde{\mathbb{P}}$ and the volatilities $\sigma > 0$ and $\nu > 0$ are constant, as is the interest rate $r$. We define a third asset, the money market account, whose price per share at time $t$ is $M(t) = e^{rt}$.

Let us now denominate prices in terms of the numéraire $N$, so that the redenominated first asset price is

$$\widehat{S}(t) = \frac{S(t)}{N(t)}$$

and the redenominated money market account price is

$$\widehat{M}(t) = \frac{M(t)}{N(t)}.$$

According to Theorem 9.2.2, $d\widehat{S}(t) = (\sigma-\nu)\widehat{S}(t)\,d\widehat{W}(t)$, where $\widehat{W}(t) = \widetilde{W}(t) - \nu t$.

(i) Compute the differential of $\frac{1}{N(t)}$.
(ii) Compute the differential of $\widehat{M}(t)$, expressing it in terms of $d\widehat{W}(t)$.

Consider a portfolio that at each time $t$ holds $\Delta(t)$ shares of the first asset and finances this by investing in or borrowing from the money market. According to the usual formula, the differential of the value $X(t)$ of this portfolio is

$$dX(t) = \Delta(t)\,dS(t) + r\big(X(t) - \Delta(t)S(t)\big)\,dt.$$

We define

$$\Gamma(t) = \frac{X(t) - \Delta(t)S(t)}{M(t)}$$

to be the number of shares of money market account held by this portfolio at time $t$ and can then rewrite the differential of $X(t)$ as

$$dX(t) = \Delta(t)\,dS(t) + \Gamma(t)\,dM(t). \tag{9.7.1}$$

Note also that by the definition of $\Gamma(t)$, we have

$$X(t) = \Delta(t)S(t) + \Gamma(t)M(t). \tag{9.7.2}$$

We redenominate the portfolio value, defining

$$\widehat{X}(t) = \frac{X(t)}{N(t)}, \tag{9.7.3}$$

so that (dividing (9.7.2) by $N(t)$) we have

$$\widehat{X}(t) = \Delta(t)\widehat{S}(t) + \Gamma(t)\widehat{M}(t). \tag{9.7.4}$$

(iii) Use stochastic calculus to show that

$$d\widehat{X}(t) = \Delta(t)\, d\widehat{S}(t) + \Gamma(t)\, d\widehat{M}(t).$$

This equation is the counterpart in the new numéraire of equation (9.7.1) and says that the change in $\widehat{X}(t)$ is solely due to changes in the prices of the assets held by the portfolio. (Hint: Start from equation (9.7.3) and use (9.7.1) and (9.7.4) along the way.)

**Exercise 9.3 (Change in volatility caused by change of numéraire).** Let $S(t)$ and $N(t)$ be the prices of two assets, denominated in a common currency, and let $\sigma$ and $\nu$ denote their volatilities, which we assume are constant. We assume also that the interest rate $r$ is constant. Then

$$dS(t) = rS(t)\, dt + \sigma S(t)\, d\widetilde{W}_1(t),$$
$$dN(t) = rN(t)\, dt + \nu N(t)\, d\widetilde{W}_3(t),$$

where $\widetilde{W}_1(t)$ and $\widetilde{W}_3(t)$ are Brownian motions under the risk-neutral measure $\widetilde{\mathbb{P}}$. We assume these Brownian motions are correlated, with $d\widetilde{W}_1(t)\, d\widetilde{W}_3(t) = \rho\, dt$ for some constant $\rho$.

(i) Show that $S^{(N)}(t) = \frac{S(t)}{N(t)}$ has volatility $\gamma = \sqrt{\sigma^2 - 2\rho\sigma\nu + \nu^2}$. In other words, show that there exists a Brownian motion $\widetilde{W}_4$ under $\widetilde{\mathbb{P}}$ such that

$$\frac{dS^{(N)}(t)}{S^{(N)}(t)} = (\text{Something})\, dt + \gamma\, d\widetilde{W}_4(t).$$

(ii) Show how to construct a Brownian motion $\widetilde{W}_2(t)$ under $\widetilde{\mathbb{P}}$ that is independent of $\widetilde{W}_1(t)$ such that $dN(t)$ may be written as

$$dN(t) = rN(t)\, dt + \nu N(t)\big[\rho\, d\widetilde{W}_1(t) + \sqrt{1-\rho^2}\, d\widetilde{W}_2(t)\big].$$

(iii) Using Theorem 9.2.2, determine the volatility vector of $S^{(N)}(t)$. In other words, find a vector $(v_1, v_2)$ such that

$$dS^{(N)}(t) = S^{(N)}(t)\big[v_1\, d\widetilde{W}_1^{(N)}(t) + v_2\, d\widetilde{W}_2^{(N)}(t)\big],$$

where $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ are independent Brownian motions under $\widetilde{\mathbb{P}}^{(N)}$. Show that

$$\sqrt{v_1^2 + v_2^2} = \sqrt{\sigma^2 - 2\rho\sigma\nu + \nu^2}.$$

**Exercise 9.4.** From the differential formulas (9.3.14) and (9.3.15) for the stock and discounted exchange rate in terms of the Brownian motions under the domestic risk-neutral measure, derive the differential formulas (9.3.22) and (9.3.23) for the redenominated money market account and stock discounted at the foreign interest rate and written in terms of the Brownian motions under the *foreign* risk-neutral measure.

**Exercise 9.5 (Quanto option).** A *quanto option* pays off in one currency the price in another currency of an underlying asset without taking the currency conversion into account. For example, a quanto call on a British asset struck at \$25 would pay \$5 if the price of the asset upon expiration of the option is £30. To compute the payoff of the option, the price 30 is treated as if it were dollars, even though it is pounds sterling.

In this problem we consider a quanto option in the foreign exchange model of Section 9.3. We take the domestic and foreign interest rates to be constants $r$ and $r^f$, respectively, and we assume that $\sigma_1 > 0$, $\sigma_2 > 0$, and $\rho \in (-1, 1)$ are likewise constant.

(i) From (9.3.14), show that

$$S(t) = S(0) \exp\left\{\sigma_1 \widetilde{W}_1(t) + \left(r - \frac{1}{2}\sigma_1^2\right) t\right\}.$$

(ii) From (9.3.16), show that

$$Q(t) = Q(0) \exp\left\{\sigma_2 \rho \widetilde{W}_1(t) + \sigma_2 \sqrt{1-\rho^2}\, \widetilde{W}_2(t) + \left(r - r^f - \frac{1}{2}\sigma_2^2\right) t\right\}.$$

(iii) Show that

$$\frac{S(t)}{Q(t)} = \frac{S(0)}{Q(0)} \exp\left\{\sigma_4 \widetilde{W}_4(t) + \left(r - a - \frac{1}{2}\sigma_4^2\right) t\right\},$$

where

$$\sigma_4 = \sqrt{\sigma_1^2 - 2\rho\sigma_1\sigma_2 + \sigma_2^2},$$
$$a = r - r^f + \rho\sigma_1\sigma_2 - \sigma_2^2,$$

and

$$\widetilde{W}_4(t) = \frac{\sigma_1 - \sigma_2\rho}{\sigma_4} \widetilde{W}_1(t) - \frac{\sigma_2\sqrt{1-\rho^2}}{\sigma_4} \widetilde{W}_2(t)$$

is a Brownian motion.

(iv) Consider a quanto call that pays off

$$\left(\frac{S(T)}{Q(T)} - K\right)^+$$

units of domestic currency at time $T$. (Note that $\frac{S(T)}{Q(T)}$ is denominated in units of foreign currency, but in this payoff it is treated as if it is a number of units of domestic currency.) Show that if at time $t \in [0, T]$ we have $\frac{S(t)}{Q(t)} = x$, then the price of the quanto call at this time is

$$q(t, x) = x e^{-a\tau} N(d_+(\tau, x)) - e^{-r\tau} K N(d_-(\tau, x)),$$

where $\tau = T - t$ and

$$d_{\pm}(\tau, x) = \frac{1}{\sigma_4\sqrt{\tau}}\left[\log\frac{x}{K} + \left(r - a \pm \frac{1}{2}\sigma_4^2\right)\tau\right].$$

(Hint: Argue that this is a case of formula (5.5.12).)

**Exercise 9.6.** Verify equation (9.4.16),

$$\mathrm{For}_S(t,T)\, dN(d_+(t)) + d\mathrm{For}_S(t,T)\, dN(d_+(t)) - K\, dN(d_-(t)) = 0,$$

in the following steps.

(i) Use (9.4.10) to show that

$$d_-(t) = d_+(t) - \sigma\sqrt{T-t}.$$

(ii) Use (9.4.10) to show that

$$d_+^2(t) - d_-^2(t) = 2\log\frac{\mathrm{For}_S(t,T)}{K}.$$

(iii) Use (ii) to show that

$$\mathrm{For}_S(t,T)e^{-d_+^2(t)/2} - Ke^{-d_-^2(t)/2} = 0.$$

(iv) Use (9.4.8) and the Itô-Doeblin formula to show that

$$dd_+(t) = \frac{1}{2\sigma(T-t)^{3/2}}\log\frac{\mathrm{For}_S(t,T)}{K}\, dt - \frac{3\sigma}{4\sqrt{T-t}}\, dt + \frac{1}{\sqrt{T-t}}\, d\widetilde{W}$$

(v) Use (i) to show that

$$dd_-(t) = dd_+(t) + \frac{\sigma}{2\sqrt{T-t}}\, dt.$$

(vi) Use (iv) and (v) to show that

$$dd_+(t)\, dd_+(t) = dd_-(t)\, dd_-(t) = \frac{dt}{T-t}.$$

(vii) Use the Itô-Doeblin formula to show that

$$dN(d_+(t)) = \frac{1}{\sqrt{2\pi}}e^{-d_+^2(t)/2}\, dd_+(t) - \frac{d_+(t)}{2(T-t)\sqrt{2\pi}}e^{-d_+^2(t)/2}\, dt.$$

(viii) Use the Itô-Doeblin formula, (v), (i), and (vi) to show that

$$dN(d_-(t)) = \frac{1}{\sqrt{2\pi}}e^{-d_-^2(t)/2}\, dd_+(t) + \frac{\sigma}{\sqrt{2\pi(T-t)}}e^{-d_-^2(t)/2}\, dt - \frac{d_+(t)}{2(T-t)\sqrt{2\pi}}e^{-d_-^2(t)/2}\, dt.$$

(ix) Use (9.4.8), (vii), and (iv) to show that

$$d\mathrm{For}_S(t,T)\, dN(d_+(t)) = \frac{\sigma\mathrm{For}_S(t,T)}{\sqrt{2\pi(T-t)}}e^{-d_+^2(t)/2}\, dt.$$

(x) Now prove (9.4.16).

# 10

# Term-Structure Models

## 10.1 Introduction

Real markets do not have a single interest rate. Instead, they have bonds of different maturities, some paying coupons and others not paying coupons. From these bonds, *yields* to different maturities can be implied. More specifically, let $0 = T_0 < T_1 < T_2 < \cdots < T_n$ be a given set of dates, and let $B(0, T_j)$ denote the price at time zero of a zero-coupon bond paying 1 at maturity $T_j$. Consider a coupon-paying bond that makes fixed payments $C_1, C_2, \ldots, C_j$ at dates $T_1, T_2, \ldots, T_j$, respectively. Each of the numbers $C_1, C_2, \ldots, C_{j-1}$ represents a coupon (interest payment), and $C_j$ represents the interest plus principal paid at the maturity $T_j$ of the bond. The price of this bond at time zero can be decomposed as

$$\sum_{j=i}^{j} C_i B(0, T_i). \tag{10.1.1}$$

On the other hand, if one is given the price of a coupon-paying bond of each maturity $T_1, T_2, \ldots, T_n$, then using (10.1.1) one can solve recursively for $B(0, T_1), \ldots, B(0, T_n)$ by first observing that $B(0, T_1)$ is the price of the $T_1$-maturity bond divided by the payment it will make at $T_1$, then using this value of $B(0, T_1)$ and the price of the $T_2$-maturity bond to solve for $B(0, T_2)$, and continuing in this manner. This method of determining zero-coupon bond prices from coupon-paying bond prices is called *bootstrapping.*

In any event, from market data one can ultimately determine prices of zero-coupon bonds for a number of different maturity dates. Each of these bonds has a *yield* specific to its maturity, where yield is defined to be the constant continuously compounding interest rate over the lifetime of the bond that is consistent with its price:

$$\text{price of zero-coupon bond} = \text{face value} \times e^{-\text{yield} \times \text{time to maturity}}.$$

The *face value* of a zero-coupon bond is the amount it promises to pay upon maturity. The formula above implies that capital equal to the price of the bond, invested at a continuously compounded interest rate equal to the yield, would, over the lifetime of the bond, result in a final payment of the face value. In this chapter, we shall normalize zero-coupon bonds by taking the face value to be 1.

In summary, instead of having a single interest rate, real markets have a *yield curve*, which one can regard either as a function of finitely many yields plotted versus their corresponding maturities or more often as a function of a nonnegative real variable (time) obtained by interpolation from the finitely many maturity–yield pairs provided by the market. The interest rate (sometimes called the *short rate*) is an idealization corresponding to the shortest-maturity yield or perhaps the overnight rate offered by the government, depending on the particular application.

We assume throughout this chapter that the bonds have no risk of default. One generally regards U.S. government bonds to be nondefaultable.

Models for interest rates have already appeared in this text, most notably in Section 6.5, where the partial differential equation satisfied by zero-coupon bonds in a one-factor short-rate model was developed and the Hull-White and Cox-Ingersoll-Ross models were given as examples. In Section 10.2 of this chapter, we extend these models to permit finitely many factors. These are Markov models in which the state of the model at each time is a multidimensional vector.

Unlike the models for equities considered heretofore and the Heath-Jarrow-Morton model considered later, the multifactor models in Section 10.2 do not immediately provide a mechanism for evolution of the prices of tradeable assets. In the earlier models, we assume an evolution of the price of a primary asset or the prices of multiple primary assets under the actual measure and then determine the market prices of risk that enable us to switch to a risk-neutral measure. In the multifactor models of Section 10.2, we begin with the evolution of abstract "factors," and from these the interest rate is obtained. But the interest rate is not the price of an asset, and we cannot infer a market price of risk from the interest rate alone. If we also had prices of some primary assets, say zero-coupon bonds, we could determine market prices of risk. However, in the models of Section 10.2, the only way to get prices of zero-coupon bonds is to use the risk-neutral pricing formula, and this cannot be done until we have a risk-neutral measure. Therefore, we build these models under the risk-neutral measure from the outset. Zero-coupon bond prices are given by the risk-neutral pricing formula, which implies that discounted zero-coupon bond prices are martingales under the risk-neutral measure. This implies in turn that no arbitrage can be achieved by trading in the zero-coupon bonds and the money market. After these models are built, they are calibrated to market prices for zero-coupon bonds and probably also some fixed income derivatives. The actual probability measure and the market prices of risk never enter the picture.

In contrast to the models of Section 10.2, the *Heath-Jarrow-Morton (HJM) model* takes its state at each time to be the forward curve at that time. The forward rate $f(t,T)$, which is the state of the HJM model, is the instantaneous rate that can be locked in at time $t$ for borrowing at time $T \geq t$. For fixed $t$, one calls the function $T \mapsto f(t,T)$, defined for $T \geq t$, the *forward rate curve.* The HJM model provides a mechanism for evolving this curve (a "curve" in the variable $T$) forward in time (the variable $t$). The forward rate curve can be deduced from the zero-coupon bond prices, and the zero-coupon bond prices can be deduced from the forward rate curve. Because zero-coupon bond prices are given directly by the HJM model rather than indirectly by the risk-neutral pricing formula, one needs to be careful that the model does not generate prices that admit arbitrage. Hence, HJM is more than a model because it provides a necessary and sufficient condition for a model driven by Brownian motion to be free of arbitrage. Every Brownian-motion-driven model must satisfy the HJM no-arbitrage condition, and to illustrate that point we provide Exercise 10.10 to verify that the Hull-White and Cox-Ingersoll-Ross models satisfy this condition.

For practical applications, it would be convenient to build a model where the forward rate had a log-normal distribution. Unfortunately, this is not possible. However, if one instead models the *simple interest rate* $L(t,T)$ that one can lock in at time $t$ for borrowing over the interval $T$ to $T+\delta$, where $\delta$ is a positive constant, this problem can be overcome. We call $L(t,T)$ *forward LIBOR (London interbank offered rate).* The constant $\delta$ is typically 0.25 (*three-month LIBOR*) or 0.50 (*six-month LIBOR*). The model that takes forward LIBOR as its state is often called the *forward LIBOR model*, the *market model*, or the *Brace-Gatarek-Musiela (BGM) model.* It is presented in Section 10.4.

## 10.2 Affine-Yield Models

The one-factor Cox-Ingersoll-Ross (CIR) and Hull-White models appearing in Section 6.5 are called *affine-yield models* because in these models the yield for zero-coupon bond prices is an affine (linear plus constant) function of the interest rate. In this section, we develop the two-factor, constant-coefficient versions of these models. (The constant-coefficient version of the Hull-White model is the Vasicek model.) Models with three or more factors can be developed along the lines of the two-factor models of this section.

It turns out that there are essentially three different two-factor affine-yield models, one in which both factors have constant diffusion terms (and hence are Gaussian processes, taking negative values with positive probability), one in which both factors appear under the square root in diffusion terms (and hence must be nonnegative at all times), and one in which only one factor appears under the square root in the diffusion terms (and only this factor is nonnegative at all times, whereas the other factor can become negative). We shall call these the *two-factor Vasicek*, the *two-factor CIR*, and the *two-factor mixed* term-structure models, respectively. For each of these types of models, there is a *canonical model* (i.e., a simplest way of writing the model).

Two-factor affine yield-models appearing in the literature, which often seem to be more complicated than the canonical models of this section, can always be obtained from one of the three canonical models by changing variables. It is desirable when calibrating a model to first change the variables to put the model into a form having the minimum number of parameters; otherwise, the calibration can be confounded by the fact that multiple sets of parameters yield the same result. The canonical models presented here have the minimmu number of parameters.

### 10.2.1 Two-Factor Vasicek Model

For the two-factor Vasicek model, we let the factors $X_1(t)$ and $X_2(t)$ be given by the system of stochastic differential equations

$$dX_1(t) = \big(a_1 - b_{11}X_1(t) - b_{12}X_2(t)\big)\,dt + \sigma_1\,d\widetilde{B}_1(t), \tag{10.2.1}$$

$$dX_1(t) = \big(a_2 - b_{21}X_1(t) - b_{22}X_2(t)\big)\,dt + \sigma_2\,d\widetilde{B}_2(t), \tag{10.2.2}$$

where the processes $\widetilde{B}_1(t)$ and $\widetilde{B}_2(t)$ are Brownian motions under a risk-neutral measure $\widetilde{\mathbb{P}}$ with constant correlation $\nu \in (-1,1)$ (i.e., $d\widetilde{B}_1(t)\,d\widetilde{B}_2(t) = \nu\,dt$). The constants $\sigma_1$ and $\sigma_2$ are assumed to be strictly positive. We further assume that the matrix

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

has strictly positive eigenvalues $\lambda_1$ and $\lambda_2$. The positivity of these eigenvalues causes the factors $X_1(t)$ and $X_2(t)$, as well as the canonical factors $Y_1(t)$ and $Y_2(t)$ defined below, to be mean-reverting. Finally, we assume the interest rate is an affine function of the factors,

$$R(t) = \epsilon_0 + \epsilon_1 X_1(t) + \epsilon_2 X_2(t), \tag{10.2.3}$$

where $\epsilon_0$, $\epsilon_1$, and $\epsilon_2$ are constants. This is the most general two-factor Vasicek model.

#### Canonical Form

As presented above, the two-factor Vasicek model is "overparametrized" (i.e., different choices of the parameters $a_i$, $b_{ij}$, $\sigma_i$, and $\epsilon_i$ can lead to the same distribution for the process $R(t)$). To eliminate this overparametrization, we reduce the model (10.2.1)–(10.2.3) to the *canonical two-factor Vasicek model*

$$dY_1(t) = -\lambda_1 Y_1(t)\,dt + d\widetilde{W}_1(t), \tag{10.2.4}$$

$$dY_2(t) = -\lambda_{21} Y_1(t)\,dt - \lambda_2 Y_2(t)\,dt + d\widetilde{W}_2(t), \tag{10.2.5}$$

$$R(t) = \delta_0 + \delta_1 Y_1(t) + \delta_2 Y_2(t), \tag{10.2.6}$$

where $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ are independent Brownian motions.

The canonical two-factor Vasicek model has six parameters:

$$\lambda_1 > 0, \; \lambda_2 > 0, \; \lambda_{21} \; \delta_0, \; \delta_1, \; \delta_2.$$

The parameters are used to calibrate the model. In practice, one often permits some of these parameters to be time-varying but nonrandom in order to make the model fit the initial yield curve; see Exercise 10.3.

To achieve this reduction, we first transform $B$ to its Jordan canonical form by choosing a nonsingular matrix

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

such that

$$K = PBP^{-1} = \begin{bmatrix} \lambda_1 & 0 \\ \kappa & \lambda_2 \end{bmatrix}.$$

If $\lambda_1 \neq \lambda_2$, then the columns of $P^{-1}$ are eigenvectors of $B$ and $\kappa = 0$ (i.e., $K$ is diagonal). If $\lambda_1 = \lambda_2$, then $\kappa$ might be zero, but it can also happen that $\kappa \neq 0$, in which case we may choose $P$ so that $\kappa = 1$. Using the notation

$$X(t) = \begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix}, \; A = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}, \; \Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}, \; \widetilde{B}(t) = \begin{bmatrix} \widetilde{B}_1(t) \\ \widetilde{B}_2(t) \end{bmatrix},$$

we may rewrite (10.2.1) and (10.2.2) in vector notation:

$$dX(t) = A\,dt - BX(t)\,dt + \Sigma\,d\widetilde{B}(t).$$

Multiplying both sides by $P$ and defining $\overline{X}(t) = PX(t)$, we obtain

$$d\overline{X}(t) = PA\,dt - K\overline{X}(t)\,dt + P\Sigma\,d\widetilde{B}(t),$$

which can be written componentwise as

$$\begin{aligned} d\overline{X}_1(t) &= (p_{11}a_1 + p_{12}a_2)\,dt - \lambda_1\overline{X}_1(t)\,dt \\ &\qquad + p_{11}\sigma_1\,d\widetilde{B}_1(t) + p_{12}\sigma_2\,d\widetilde{B}_2(t), && (10.2.7) \\ d\overline{X}_2(t) &= (p_{21}a_1 + p_{22}a_2)\,dt - \kappa\overline{X}_1(t)\,dt - \lambda_2\overline{X}_2(t)\,dt \\ &\qquad + p_{21}\sigma_1\,d\widetilde{B}_1(t) + p_{22}\sigma_2\,d\widetilde{B}_2(t). && (10.2.8) \end{aligned}$$

**Lemma 10.2.1.** *Under our assumptions that $\sigma_1 > 0$, $\sigma_2 > 0$, $-1 < \nu < 1$, and $P$ is nonsingular, we have*

$$\gamma_i = p_{i1}^2\sigma_1^2 + 2\nu p_{i1}p_{i2}\sigma_1\sigma_2 + p_{i2}^2\sigma_2^2, \quad i = 1, 2, \tag{10.2.9}$$

*are strictly positive, and*

$$\rho = \frac{1}{\sqrt{\gamma_1\gamma_2}}(p_{11}p_{21}\sigma_1^2 + \nu(p_{11}p_{22} + p_{12}p_{21})\sigma_1\sigma_2 + p_{12}p_{22}\sigma_2^2) \tag{10.2.10}$$

*is in $(-1, 1)$.*

PROOF: Because $\nu \in (-1,1)$, the matrix

$$N = \begin{bmatrix} 1 & \nu \\ \nu & 1 \end{bmatrix}$$

has a matrix square root. Indeed, one such square root is

$$\sqrt{N} = \begin{bmatrix} a & \sqrt{1-a^2} \\ \sqrt{1-a^2} & a \end{bmatrix},$$

where $a = \mathrm{sign}(\nu)\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{1-\nu^2}}$. Verification of this uses the equation

$$\begin{aligned} 2a\sqrt{1-a^2} &= 2\mathrm{sign}(\nu)\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{1-\nu^2}} \cdot \sqrt{\frac{1}{2} - \frac{1}{2}\sqrt{1-\nu^2}} \\ &= 2\mathrm{sign}(\nu)\sqrt{\frac{1}{4} - \frac{1}{4}(1-\nu^2)} \\ &= 2\mathrm{sign}(\nu) \cdot \frac{1}{2}|\nu| \ = \ \nu. \end{aligned}$$

The matrices $\sqrt{N}$, $\Sigma$, and $P^{\mathrm{tr}}$ are nonsingular, which implies nonsingularity of the matrix

$$\sqrt{N}\Sigma P^{\mathrm{tr}} = \begin{bmatrix} p_{11}\sigma_1 a + p_{12}\sigma_2\sqrt{1-a^2} & p_{21}\sigma_1 a + p_{22}\sigma_2\sqrt{1-a^2} \\ p_{11}\sigma_1\sqrt{1-a^2} + p_{12}\sigma_2 a & p_{21}\sigma_1\sqrt{1-a^2} + p_{22}\sigma_2 a \end{bmatrix}.$$

Let $c_1$ be the first column of this matrix and $c_2$ the second column. Because of the nonsingularity of $\sqrt{N}\Sigma P^{\mathrm{tr}}$, these vectors are linearly independent, and hence neither of them is the zero vector,

Therefore,

$$\gamma_i = \|c_i\|^2 > 0, \quad i = 1, 2.$$

For linearly independent vectors, the Cauchy-Schwarz inequality implies

$$-\|c_1\|\,\|c_2\| < c_1 \cdot c_2 < \|c_1\|\,\|c_2\|.$$

This is equivalent to $-1 < \rho < 1$. □

We define

$$\overline{B}_i(t) = \frac{1}{\sqrt{\gamma_i}}\big(p_{i1}\sigma_1\widetilde{B}_1(t) + p_{i2}\sigma_2\widetilde{B}_2(t)\big), \quad i = 1, 2.$$

The processes $\overline{B}_1(t)$ and $\overline{B}_2(t)$ are continuous martingales starting at zero. Furthermore,

$$d\overline{B}_1(t)\,d\overline{B}_1(t) = d\overline{B}_2(t)\,d\overline{B}_2(t) = dt.$$

According to Lévy's Theorem, Theorem 4.6.4, $\overline{B}_1(t)$ and $\overline{B}_2(t)$ are Brownian motions. Furthermore,

$$d\overline{B}_1(t)\,d\overline{B}_2(t) = \rho\,dt,$$

where $\rho$ is defined by (10.2.10). We may rewrite (10.2.7) and (10.2.8) as

$$d\overline{X}_1(t) = (p_{11}a_1 + p_{12}a_2)\,dt - \lambda_1\overline{X}_1(t)\,dt + \sqrt{\gamma_1}\,d\overline{B}_1(t), \tag{10.2.11}$$
$$d\overline{X}_2(t) = (p_{21}a_1 + p_{22}a_2)\,dt - \kappa\overline{X}_1(t)\,dt - \lambda_2\overline{X}_2(t)\,dt + \sqrt{\gamma_2}\,d\overline{B}_2(t). \tag{10.2.12}$$

Setting

$$\widehat{X}_1(t) = \frac{1}{\sqrt{\gamma_1}}\left(\overline{X}_1(t) - \frac{p_{11}a_1 + p_{12}a_2}{\lambda_1}\right),$$
$$\widehat{X}_2(t) = \frac{1}{\sqrt{\gamma_2}}\left(\overline{X}_2(t) + \frac{\kappa(p_{11}a_1 + p_{12}a_2)}{\lambda_1\lambda_2} - \frac{p_{21}a_1 + p_{22}a_2}{\lambda_2}\right),$$

we may further rewrite (10.2.11) and (10.2.12) as

$$d\widehat{X}_1(t) = -\lambda_1\widehat{X}_1(t)\,dt + d\overline{B}_1(t), \tag{10.2.13}$$
$$d\widehat{X}_2(t) = -\kappa\sqrt{\frac{\gamma_1}{\gamma_2}}\,\widehat{X}_1(t)\,dt - \lambda_2\widehat{X}_2(t)\,dt + d\overline{B}_2(t). \tag{10.2.14}$$

As the last step, we define

$$\widetilde{W}_1(t) = \overline{B}_1(t), \quad \widetilde{W}_2(t) = \frac{1}{\sqrt{1-\rho^2}}\big[-\rho\overline{B}_1(t) + \overline{B}_2(t)\big].$$

Both $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ are continuous martingales, and it is easily verified that

$$d\widetilde{W}_1(t)\,d\widetilde{W}_1(t) = dt, \quad d\widetilde{W}_1(t)\,d\widetilde{W}_2(t) = 0, \quad d\widetilde{W}_2(t)\,d\widetilde{W}_2(t) = dt.$$

According to Lévy's Theorem, Theorem 4.6.4, $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ are independent Brownian motions. Setting

$$Y_1(t) = \widehat{X}_1(t), \quad Y_2(t) = \frac{-\rho\widehat{X}_1(t) + \widehat{X}_2(t)}{\sqrt{1-\rho^2}},$$

we have

$$\begin{aligned}
dY_2(t) &= \frac{1}{\sqrt{1-\rho^2}}\big[-\rho\,d\widehat{X}_1(t) + d\widehat{X}_2(t)\big]\\
&= \frac{1}{\sqrt{1-\rho^2}}\left[\left(\rho\lambda_1 - \kappa\sqrt{\frac{\gamma_1}{\gamma_2}}\right)\widehat{X}_1(t) - \lambda_2\widehat{X}_2(t)\right]dt\\
&\quad + \frac{1}{\sqrt{1-\rho^2}}\big[-\rho\,d\overline{B}_1(t) + d\overline{B}_2(t)\big]\\
&= \frac{1}{\sqrt{1-\rho^2}}\left(\rho\lambda_1 - \rho\lambda_2 - \kappa\sqrt{\frac{\gamma_1}{\gamma_2}}\right)Y_1(t)\,dt - \lambda_2 Y_2(t)\,dt + d\widetilde{W}_2(t).
\end{aligned}$$

We may thus rewrite (10.2.13) and (10.2.14) as

$$dY_1(t) = -\lambda_1 Y_1(t)\,dt + d\widetilde{W}_1(t),$$
$$dY_2(t) = -\lambda_{21} Y_1(t)\,dt - \lambda_2 Y_2(t)\,dt + d\widetilde{W}_2(t),$$

where

$$\lambda_{21} = \frac{1}{\sqrt{1-\rho^2}}\left(-\rho\lambda_1 + \rho\lambda_2 + \kappa\sqrt{\frac{\gamma_1}{\gamma_2}}\right).$$

These are the canonical equations (10.2.4) and (10.2.5).

To obtain (10.2.6), we trace back through the changes of variables:

$$\begin{aligned}
Y_1(t) &= \widehat{X}_1(t)\\
&= \frac{1}{\sqrt{\gamma_1}}\left(\overline{X}_1(t) - \frac{p_{11}a_1 + p_{12}a_2}{\lambda_1}\right)\\
&= \frac{1}{\sqrt{\gamma_1}}\left(p_{11}X_1(t) + p_{12}X_2(t) - \frac{p_{11}a_1 + p_{12}a_2}{\lambda_1}\right),\\
Y_2(t) &= \frac{1}{\sqrt{1-\rho^2}}\big(-\rho\widehat{X}_1(t) + \widehat{X}_2(t)\big)\\
&= -\frac{\rho}{\sqrt{\gamma_1(1-\rho^2)}}\left(\overline{X}_1(t) - \frac{p_{11}a_1 + p_{12}a_2}{\lambda_1}\right)\\
&\quad + \frac{1}{\sqrt{\gamma_2(1-\rho^2)}}\left(\overline{X}_2(t) + \frac{\kappa(p_{11}a_1 + p_{12}a_2)}{\lambda_1\lambda_2} - \frac{p_{21}a_1 + p_{22}a_2}{\lambda_2}\right)\\
&= -\frac{\rho}{\sqrt{\gamma_1(1-\rho^2)}}\left(p_{11}X_1(t) + p_{12}X_2(t) - \frac{p_{11}a_1 + p_{12}a_2}{\lambda_1}\right)\\
&\quad + \frac{1}{\sqrt{\gamma_2(1-\rho^2)}}\left(p_{21}X_1(t) + p_{22}X_2(t) + \frac{\kappa(p_{11}a_1 + p_{12}a_2)}{\lambda_1\lambda_2}\right.\\
&\qquad\qquad \left. - \frac{p_{21}a_1 + p_{22}a_2}{\lambda_2}\right).
\end{aligned}$$

In vector notation,

$$Y(t) = \Gamma\big(PX(t) + V\big), \tag{10.2.15}$$

where

$$Y(t) = \begin{bmatrix} Y_1(t) \\ Y_2(t) \end{bmatrix}, \quad \Gamma = \begin{bmatrix} \frac{1}{\sqrt{\gamma_1}} & 0 \\ -\frac{\rho}{\sqrt{\gamma_1(1-\rho^2)}} & \frac{1}{\sqrt{\gamma_2(1-\rho^2}} \end{bmatrix},$$

$$V = \begin{bmatrix} -\frac{p_{11}a_1 + p_{12}a_2}{\lambda_1} \\ \frac{\kappa(p_{11}a_1 + p_{12}a_2)}{\lambda_1\lambda_2} - \frac{p_{21}a_1 + p_{22}a_2}{\lambda_2} \end{bmatrix}.$$

We solve (10.2.15) for $X(t)$:

$$X(t) = P^{-1}\big(\Gamma^{-1}Y(t) - V\big).$$

Therefore,

$$\begin{aligned} R(t) &= \epsilon_0 + [\epsilon_1\ \epsilon_2]X(t) \\ &= \epsilon_0 + [\epsilon_1\ \epsilon_2]P^{-1}\Gamma^{-1}Y(t) - [\epsilon_1\ \epsilon_2]P^{-1}V \\ &= \delta_0 + [\delta_1\ \delta_2]Y(t), \end{aligned}$$

where

$$\delta_0 = \epsilon_0 - [\epsilon_1\ \epsilon_2]P^{-1}V, \quad [\delta_1\ \delta_2] = [\epsilon_1\ \epsilon_2]P^{-1}\Gamma^{-1}.$$

We have obtained (10.2.6).

## Bond Prices

We derive the formula for zero-coupon bond prices in the canonical two-factor Vasicek model. According to the risk-neutral pricing formula, the price at time $t$ of a zero-coupon bond paying 1 at a later time $T$ is

$$B(t,T) = \widetilde{\mathbb{E}}\left[e^{-\int_t^T R(u)du}\middle| \mathcal{F}(t)\right], \quad 0 \le t \le T.$$

Because $R(t)$ given by (10.2.6) is a function of the factors $Y_1(t)$ and $Y_2(t)$, and the solution of the system of stochastic differential equations (10.2.4) and (10.2.5) is Markov, there must be some function $f(t, y_1, y_2)$ such that

$$B(t,T) = f\big(t, Y_1(t), Y_2(t)\big). \tag{10.2.16}$$

The discount factor $D(t) = e^{-\int_0^t R(u)du}$ satisfies $dD(t) = -R(t)D(t)\,dt$ (see (5.2.18)). Iterated conditioning implies that the discounted bond price $D(t)B(t,T)$ is a martingale under $\widetilde{\mathbb{P}}$. Therefore, the differential of $D(t)B(t,T)$ has $dt$ term zero. We compute this differential:

$$\begin{aligned} &d\big(D(t)B(t,T)\big) \\ &= d\Big(D(t)f\big(t, Y_1(t), Y_2(t)\big)\Big) \\ &= -R(t)D(t)f\big(t, Y_1(t), Y_2(t)\big)\,dt + D(t)\,df\big(t, Y_1(t), Y_2(t)\big) \\ &= D\Bigg[-Rf\,dt + f_t\,dt + f_{y_1}\,dY_1 + f_{y_2}\,dY_2 \\ &\qquad + \frac{1}{2}f_{y_1y_1}\,dY_1\,dY_1 + f_{y_1y_2}\,dY_1\,dY_2 + \frac{1}{2}f_{y_2y_2}\,dY_2\,dY_2\Bigg]. \end{aligned} \tag{10.2.17}$$

We use equations (10.2.4)–(10.2.6) to take the next step:

$$
\begin{aligned}
&d(D(t)B(t,T))\\
&= D\Big[-(\delta_0+\delta_1 Y_1+\delta_2 Y_2)f + f_t - \lambda_1 Y_1 f_{y_1} - \lambda_{21} Y_1 f_{y_2}\\
&\qquad -\lambda_2 Y_2 f_{y_2} + \frac{1}{2} f_{y_1 y_1} + \frac{1}{2} f_{y_2 y_2}\Big]\,dt + D\left[f_{y_1}\,d\widetilde{W}_1 + f_{y_2}\,d\widetilde{W}_2\right].
\end{aligned}
$$

Setting the $dt$ term equal to zero, we obtain the partial differential equation

$$
\begin{aligned}
&-(\delta_0+\delta_1 y_1+\delta_2 y_2)f(t,y_1,y_2) + f_t(t,y_1,y_2)\\
&\quad -\lambda_1 y_1 f_{y_1}(t,y_1,y_2) - \lambda_{21} y_1 f_{y_2}(t,y_1,y_2) - \lambda_2 y_2 f_{y_2}(t,y_1,y_2)\\
&\qquad +\frac{1}{2} f_{y_1 y_1}(t,y_1,y_2) + \frac{1}{2} f_{y_2 y_2}(t,y_1,y_2) = 0 \qquad (10.2.18)
\end{aligned}
$$

for all $t \in [0,T)$ and all $y_1 \in \mathbb{R}$, $y_2 \in \mathbb{R}$. We have also the terminal condition

$$
f(T,y_1,y_2) = 1 \text{ for all } y_1 \in \mathbb{R},\ y_2 \in \mathbb{R}. \qquad (10.2.19)
$$

To solve this equation, we seek a solution of the affine-yield form

$$
f(t,y_1,y_2) = e^{-y_1 C_1(T-t) - y_2 C_2(T-t) - A(T-t)} \qquad (10.2.20)
$$

for some functions $C_1(\tau)$, $C_2(\tau)$, and $A(\tau)$. Here we define $\tau = T - t$ to be the *relative maturity* (i.e., the time until maturity). So long as the model parameters do not depend on $t$, zero-coupon bond prices will depend on $t$ and $T$ only through $\tau$. The terminal condition (10.2.19) implies that

$$
C_1(0) = C_2(0) = A(0) = 0. \qquad (10.2.21)
$$

We compute derivatives, where $'$ denotes differentiation with respect to $\tau$. We use the fact $\frac{d}{dt}C_i(\tau) = C_i'(\tau)\cdot\frac{d}{dt}\tau = -C_i'(\tau)$, $i=1,2$, and the similar equation $\frac{d}{dt}A(\tau) = -A'(\tau)$ to obtain

$$
\begin{aligned}
f_t &= [y_1 C_1' + y_2 C_2' + A']f, \quad & f_{y_1} &= -C_1 f, \quad & f_{y_2} &= -C_2 f,\\
f_{y_1 y_1} &= C_1^2 f, & f_{y_1 y_2} &= C_1 C_2 f, & f_{y_2 y_2} &= C_2^2 f.
\end{aligned}
$$

Equation (10.2.18) becomes

$$
\begin{aligned}
&\Big[(C_1' + \lambda_1 C_1 + \lambda_{21} C_2 - \delta_1)y_1 + (C_2' + \lambda_2 C_2 - \delta_2)y_2\\
&\qquad +\Big(A' + \frac{1}{2}C_1^2 + \frac{1}{2}C_2^2 - \delta_0\Big)\Big]f = 0. \qquad (10.2.22)
\end{aligned}
$$

Because (10.2.22) must hold for all $y_1$ and $y_2$, the term $C_1' + \lambda C_1 + \lambda_{21}C_2 - \delta_1$ multiplying $y_1$ must be zero. If it were not, and (10.2.22) held for one value of $y_1$, then a change in the value of $y_1$ would cause the equation to be violated. Similarly, the term $C_2' + \lambda_2 C_2 - \delta_2$ multiplying $y_2$ must be zero, and consequently the remaining term $A' + \frac{1}{2}C_1^2 + \frac{1}{2}C_2^2 - \delta_0$ must also be zero. This gives us a system of three ordinary differential equations:

$$C_1'(\tau) = -\lambda_1 C_1(\tau) - \lambda_{21} C_2(\tau) + \delta_1, \tag{10.2.23}$$
$$C_2'(\tau) = -\lambda_2 C_2(\tau) + \delta_2, \tag{10.2.24}$$
$$A'(\tau) = -\frac{1}{2} C_1^2(\tau) - \frac{1}{2} C_2^2(\tau) + \delta_0. \tag{10.2.25}$$

The solution of (10.2.24) satisfying the initial condition $C_2(0) = 0$ (see (10.2.21)) is

$$C_2(\tau) = \frac{\delta_2}{\lambda_2}\left(1 - e^{-\lambda_2 \tau}\right). \tag{10.2.26}$$

We substitute this into (10.2.23) and solve using the initial condition $C_1(0) = 0$. In particular, (10.2.23) implies

$$\begin{aligned}
\frac{d}{d\tau}\left(e^{\lambda_1 \tau} C_1(\tau)\right) &= e^{\lambda_1 \tau}\left(\lambda_1 C_1(\tau) + C_1'(\tau)\right) \\
&= e^{\lambda_1 \tau}\left( - \lambda_{21} C_2(\tau) + \delta_1\right) \\
&= e^{\lambda_1 \tau}\left(-\frac{\lambda_{21}\delta_2}{\lambda_2}\left(1 - e^{-\lambda_2 \tau}\right) + \delta_1\right).
\end{aligned}$$

If $\lambda_1 \neq \lambda_2$, integration from 0 to $\tau$ yields

$$C_1(\tau) = \frac{1}{\lambda_1}\left(\delta_1 - \frac{\lambda_{21}\delta_2}{\lambda_2}\right)\left(1 - e^{-\lambda_1 \tau}\right) + \frac{\lambda_{21}\delta_2}{\lambda_2(\lambda_1 - \lambda_2)}\left(e^{-\lambda_2 \tau} - e^{-\lambda_1 \tau}\right). \tag{10.2.27}$$

If $\lambda_1 = \lambda_2$, we obtain instead

$$C_1(\tau) = \frac{1}{\lambda_1}\left(\delta_1 - \frac{\lambda_{21}\delta_2}{\lambda_1}\right)\left(1 - e^{-\lambda_1 \tau}\right) + \frac{\lambda_{21}\delta_2}{\lambda_1} \tau e^{-\lambda_1 \tau}. \tag{10.2.28}$$

Finally, (10.2.25) and the initial condition $A(0) = 0$ imply

$$A(\tau) = \int_0^\tau \left[-\frac{1}{2} C_1^2(u) - \frac{1}{2} C_2^2(u) + \delta_0\right] du, \tag{10.2.29}$$

and this can be obtained in closed form by a lengthy but straightforward computation.

### Short Rate and Long Rate

We fix a positive relative maturity $\bar{\tau}$ (say, thirty years) and call the yield at time $t$ on the zero-coupon bond with relative maturity $\bar{\tau}$ (i.e., the bond maturing at date $t+\bar{\tau}$) the *long rate* $L(t)$. Once we have a model for evolution of the short rate $R(t)$ under the risk-neutral measure, then for each $t \geq 0$ the price of the $(t + \bar{\tau})$-maturity zero-coupon bond is determined by the risk-neutral pricing formula, and hence the short-rate model alone determines the long rate. We cannot therefore write down an arbitrary stochastic differential equation for the long rate. Nonetheless, in any affine-yield model, the long rate satisfies some stochastic differential equation, and we can work out this equation.

Consider the canonical two-factor Vasicek model. As we have seen in the previous discussion, zero-coupon bond prices in this model are of the form

$$B(t,T) = e^{-Y_1(t)C_1(T-t)-Y_2(t)C_2(T-t)-A(T-t)},$$

where $C_1(\tau)$, $C_2(\tau)$, and $A(\tau)$ are given by (10.2.26)–(10.2.29). Thus, the long rate at time $t$ is

$$L(t) = -\frac{1}{\overline{\tau}}\log B(t,t+\overline{\tau}) = \frac{1}{\overline{\tau}}\big[C_1(\overline{\tau})Y_1(t) + C_2(\overline{\tau})Y_2(t) + A(\overline{\tau})\big], \quad (10.2.30)$$

which is an affine function of the canonical factors $Y_1(t)$ and $Y_2(t)$ at time $t$. Because the canonical factors do not have an economic interpretation, we may wish to use $R(t)$ and $L(t)$ as the model factors. We now show how to do this, obtaining a two-factor Vasicek model of the form (10.2.1), (10.2.2), and (10.2.3), where $X_1(t)$ is replaced by $R(t)$ and $X_2(t)$ is replaced by $L(t)$.

We begin by writing the formulas (10.2.6) and (10.2.30) in vector notation:

$$\begin{bmatrix} R(t) \\ L(t) \end{bmatrix} = \begin{bmatrix} \delta_1 & \delta_2 \\ \frac{1}{\overline{\tau}}C_1(\overline{\tau}) & \frac{1}{\overline{\tau}}C_2(\overline{\tau}) \end{bmatrix} \begin{bmatrix} Y_1(t) \\ Y_2(t) \end{bmatrix} + \begin{bmatrix} \delta_0 \\ \frac{1}{\overline{\tau}}A(\overline{\tau}) \end{bmatrix}. \quad (10.2.31)$$

We wish to solve this system for $(Y_1(t), Y_2(t))$.

**Lemma 10.2.2.** *The matrix*

$$D = \begin{bmatrix} \delta_1 & \delta_2 \\ \frac{1}{\overline{\tau}}C_1(\overline{\tau}) & \frac{1}{\overline{\tau}}C_2(\overline{\tau}) \end{bmatrix}$$

*is nonsingular if and only if* $\delta_2 \neq 0$ *and*

$$(\lambda_1 - \lambda_2)\delta_1 + \lambda_{21}\delta_2 \neq 0. \quad (10.2.32)$$

.

PROOF: Consider the function $f(x) = 1 - e^{-x} - xe^{-x}$, for which $f(0) = 0$ and $f'(x) = xe^{-x} > 0$ for all $x > 0$. We have $f(x) > 0$ for all $x > 0$. Define $h(x) = \frac{1}{x}(1 - e^{-x})$. Since $h'(x) = -x^{-2}f(x)$, which is strictly negative for all $x > 0$, $h(x)$ is strictly decreasing on $(0, \infty)$.

To examine the nonsingularity of $D$, we consider first the case $\lambda_1 \neq \lambda_2$. In this case, (10.2.26) and (10.2.27) imply

$$\begin{aligned} \det(D) &= \frac{1}{\overline{\tau}}\big[\delta_1 C_2(\overline{\tau}) - \delta_2 C_1(\overline{\tau})\big] \\ &= \frac{\delta_1\delta_2}{\lambda_2\overline{\tau}}\left(1 - e^{-\lambda_2\overline{\tau}}\right) - \frac{\delta_1\delta_2}{\lambda_1\overline{\tau}}\left(1 - e^{-\lambda_1\overline{\tau}}\right) \\ &\quad + \frac{\lambda_{21}\delta_2^2}{(\lambda_1 - \lambda_2)\lambda_1\lambda_2\overline{\tau}}\big[(\lambda_1 - \lambda_2)(1 - e^{-\lambda_1\overline{\tau}}) - \lambda_1 e^{\lambda_2\overline{\tau}} + \lambda_1 e^{\lambda_1\overline{\tau}}\big] \end{aligned}$$

$$
\begin{aligned}
&= \delta_1\delta_2\left[\frac{1}{\lambda_2\bar{\tau}}\left(1-e^{-\lambda_2\bar{\tau}}\right)-\frac{1}{\lambda_1\bar{\tau}}\left(1-e^{-\lambda_1\bar{\tau}}\right)\right]\\
&\quad+\frac{\lambda_{21}\delta_2^2}{(\lambda_1-\lambda_2)\lambda_1\lambda_2\bar{\tau}}\left[\lambda_1\left(1-e^{-\lambda_2\bar{\tau}}\right)-\lambda_2\left(1-e^{-\lambda_1\bar{\tau}}\right)\right]\\
&= \delta_2\left(\delta_1+\frac{\lambda_{21}\delta_2}{(\lambda_1-\lambda_2)}\right)\left[\frac{1}{\lambda_2\bar{\tau}}\left(1-e^{-\lambda_2\bar{\tau}}\right)-\frac{1}{\lambda_1\bar{\tau}}\left(1-e^{-\lambda_1\bar{\tau}}\right)\right]\\
&= \delta_2\left(\delta_1+\frac{\lambda_{21}\delta_2}{(\lambda_1-\lambda_2)}\right)\left[h(\lambda_2\bar{\tau})-h(\lambda_1\bar{\tau})\right].
\end{aligned}
$$

Because $\lambda_1 \neq \lambda_2$ and $h$ is strictly decreasing, $h(\lambda_2\bar{\tau}) \neq h(\lambda_2\bar{\tau})$. The determinant of $D$ is nonzero if and only if $\delta_2 \neq 0$ and (10.2.32) holds.

Next consider the case $\lambda_1 = \lambda_2$. In this case, (10.2.26) and (10.2.28) imply

$$
\begin{aligned}
\det(D) &= \frac{1}{\bar{\tau}}\left[\delta_1 C_2(\bar{\tau})-\delta_2 C_1(\bar{\tau})\right]\\
&= \frac{\delta_1\delta_2}{\lambda_1\bar{\tau}}\left(1-e^{-\lambda_1\bar{\tau}}\right)-\frac{\delta_2}{\lambda_1\bar{\tau}}\left(\delta_1-\frac{\lambda_{21}\delta_2}{\lambda_1}\right)\left(1-e^{-\lambda_1\bar{\tau}}\right)-\frac{\lambda_{21}\delta_2^2}{\lambda_1}e^{-\lambda_1\bar{\tau}}\\
&= \frac{\lambda_{21}\delta_2^2}{\lambda_1^2\bar{\tau}}f(\lambda_1\bar{\tau}).
\end{aligned}
$$

Because $\lambda_1\bar{\tau}$ is positive, $f(\lambda_1\bar{\tau})$ is not zero. In this case, (10.2.32) is equivalent to $\delta_2 \neq 0$ and $\lambda_{21} \neq 0$. The determinant of $D$ is nonzero if and only if (10.2.32) holds (in which case $\delta_2 \neq 0$ and $\lambda_{21} \neq 0$). □

Under the assumptions of Lemma 10.2.2, we can invert (10.2.31) to obtain

$$
\begin{bmatrix} Y_1(t)\\ Y_2(t)\end{bmatrix} = \begin{bmatrix} \delta_1 & \delta_2\\ \frac{1}{\bar{\tau}}C_1(\bar{\tau}) & \frac{1}{\bar{\tau}}C_2(\bar{\tau})\end{bmatrix}^{-1}\left(\begin{bmatrix} R(t)\\ L(t)\end{bmatrix}-\begin{bmatrix}\delta_0\\ \frac{1}{\bar{\tau}}A(\bar{\tau})\end{bmatrix}\right). \tag{10.2.33}
$$

We can compute the differential in (10.2.31) using (10.2.4) and (10.2.5). This leads to a formula in which $Y_1(t)$ and $Y_2(t)$ appear on the right-hand side, but we can then use (10.2.33) to rewrite the right-hand side in terms of $R(t)$, $L(t)$. These steps result in the equation

$$
\begin{aligned}
&\begin{bmatrix} dR(t)\\ dL(t)\end{bmatrix}\\
&= \begin{bmatrix} \delta_1 & \delta_2\\ \frac{1}{\bar{\tau}}C_1(\bar{\tau}) & \frac{1}{\bar{\tau}}C_2(\bar{\tau})\end{bmatrix}\begin{bmatrix} dY_1(t)\\ dY_2(t)\end{bmatrix}\\
&= \begin{bmatrix} \delta_1 & \delta_2\\ \frac{1}{\bar{\tau}}C_1(\bar{\tau}) & \frac{1}{\bar{\tau}}C_2(\bar{\tau})\end{bmatrix}\left(-\begin{bmatrix} \lambda_1 & 0\\ \lambda_{21} & \lambda_2\end{bmatrix}\begin{bmatrix} Y_1(t)\\ Y_2(t)\end{bmatrix}dt+\begin{bmatrix} d\widetilde{W}_1(t)\\ d\widetilde{W}_2(t)\end{bmatrix}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \begin{bmatrix} \delta_1 & \delta_2 \\ \frac{1}{\overline{\tau}} C_1(\overline{\tau}) & \frac{1}{\overline{\tau}} C_2(\overline{\tau}) \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ \lambda_{21} & \lambda_2 \end{bmatrix} \begin{bmatrix} \delta_1 & \delta_2 \\ \frac{1}{\overline{\tau}} C_1(\overline{\tau}) & \frac{1}{\overline{\tau}} C_2(\overline{\tau}) \end{bmatrix}^{-1} \begin{bmatrix} \delta_0 \\ \frac{1}{\overline{\tau}} A(\overline{\tau}) \end{bmatrix} dt \\
&\quad - \begin{bmatrix} \delta_1 & \delta_2 \\ \frac{1}{\overline{\tau}} C_1(\overline{\tau}) & \frac{1}{\overline{\tau}} C_2(\overline{\tau}) \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ \lambda_{21} & \lambda_2 \end{bmatrix} \begin{bmatrix} \delta_1 & \delta_2 \\ \frac{1}{\overline{\tau}} C_1(\overline{\tau}) & \frac{1}{\overline{\tau}} C_2(\overline{\tau}) \end{bmatrix}^{-1} \begin{bmatrix} R(t) \\ L(t) \end{bmatrix} dt \\
&\quad + \begin{bmatrix} \delta_1 & \delta_2 \\ \frac{1}{\overline{\tau}} C_1(\overline{\tau}) & \frac{1}{\overline{\tau}} C_2(\overline{\tau}) \end{bmatrix} \begin{bmatrix} d\widetilde{W}_1(t) \\ d\widetilde{W}_2(t) \end{bmatrix}.
\end{aligned}
$$

This is the vector notation for a pair of equations of the form (10.2.1) and (10.2.2) for a two-factor Vasicek model for the short rate $R(t)$ and the long rate $L(t)$. The parameters $a_1$ and $a_2$ appearing in (10.2.1) and (10.2.2) are given by

$$
\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \delta_1 & \delta_2 \\ \frac{1}{\overline{\tau}} C_1(\overline{\tau}) & \frac{1}{\overline{\tau}} C_2(\overline{\tau}) \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ \lambda_{21} & \lambda_2 \end{bmatrix} \begin{bmatrix} \delta_1 & \delta_2 \\ \frac{1}{\overline{\tau}} C_1(\overline{\tau}) & \frac{1}{\overline{\tau}} C_2(\overline{\tau}) \end{bmatrix}^{-1} \begin{bmatrix} \delta_0 \\ \frac{1}{\overline{\tau}} A(\overline{\tau}) \end{bmatrix}.
$$

The matrix $B$ is

$$
\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} = \begin{bmatrix} \delta_1 & \delta_2 \\ \frac{1}{\overline{\tau}} C_1(\overline{\tau}) & \frac{1}{\overline{\tau}} C_2(\overline{\tau}) \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ \lambda_{21} & \lambda_2 \end{bmatrix} \begin{bmatrix} \delta_1 & \delta_2 \\ \frac{1}{\overline{\tau}} C_1(\overline{\tau}) & \frac{1}{\overline{\tau}} C_2(\overline{\tau}) \end{bmatrix}^{-1},
$$

and the eigenvalues of $B$ are $\lambda_1 > 0$, $\lambda_2 > 0$. With

$$
\sigma_1 = \sqrt{\delta_1^2 + \delta_2^2}, \quad \sigma_2 = \frac{1}{\overline{\tau}} \sqrt{C_1^2(\overline{\tau}) + C_2^2(\overline{\tau})},
$$

the processes

$$
\begin{aligned}
\widetilde{B}_1(t) &= \frac{1}{\sigma_1} \big( \delta_1 \widetilde{W}_1(t) + \delta_2 \widetilde{W}_2(t) \big), \\
\widetilde{B}_2(t) &= \frac{1}{\sigma_2 \overline{\tau}} \big( C_1(\overline{\tau}) \widetilde{W}_1(t) + C_2(\overline{\tau}) \widetilde{W}_2(t) \big),
\end{aligned}
$$

are the Brownian motions appearing in (10.2.1) and (10.2.2). Finally, equation (10.2.3) takes the form

$$
R(t) = 0 + 1 \cdot R(t) + 0 \cdot L(t)
$$

(i.e., $\epsilon_0 = \epsilon_2 = 0$, $\epsilon_1 = 1$).

## Gaussian Factor Processes

The canonical two-factor Vasicek model in vector notation is

$$
dY(t) = -\Lambda Y(t) + d\widetilde{W}(t), \tag{10.2.34}
$$

where

$$Y(t) = \begin{bmatrix} Y_1(t) \\ Y_2(t) \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ \lambda_{21} & \lambda_2 \end{bmatrix}, \quad \widetilde{W}(t) = \begin{bmatrix} \widetilde{W}_1(t) \\ \widetilde{W}_2(t) \end{bmatrix}.$$

Recall that $\lambda_1 > 0$, $\lambda_2 > 0$. There is a closed-form solution to this matrix differential equation. To derive this solution, we first form the matrix exponential $e^{\Lambda t}$ defined by

$$e^{\Lambda t} = \sum_{n=0}^{\infty} \frac{1}{n!} (\Lambda t)^n,$$

where $(\Lambda t)^0 = I$, the $2 \times 2$ identity matrix.

**Lemma 10.2.3.** *If $\lambda_1 \neq \lambda_2$, then*

$$e^{\Lambda t} = \begin{bmatrix} e^{\lambda_1 t} & 0 \\ \frac{\lambda_{21}}{\lambda_1 - \lambda_2}\left(e^{\lambda_1 t} - e^{\lambda_2 t}\right) & e^{\lambda_2 t} \end{bmatrix}. \tag{10.2.35}$$

*If $\lambda_1 = \lambda_2$, then*

$$e^{\Lambda t} = \begin{bmatrix} e^{\lambda_1 t} & 0 \\ \lambda_{21} t e^{\lambda_1 t} & e^{\lambda_1 t} \end{bmatrix}. \tag{10.2.36}$$

*In either case,*

$$\frac{d}{dt} e^{\Lambda t} = \Lambda e^{\Lambda t} = e^{\Lambda t} \Lambda, \tag{10.2.37}$$

*where the derivative is defined componentwise, and*

$$e^{-\Lambda t} = \left(e^{\Lambda t}\right)^{-1}, \tag{10.2.38}$$

*where $e^{-\Lambda t}$ is obtained by replacing $\lambda_1$, $\lambda_2$, and $\lambda_{21}$ in the formula for $e^{\Lambda t}$ by $-\lambda_1$, $-\lambda_2$, and $-\lambda_{21}$, respectively.*

PROOF: We consider first the case $\lambda_1 \neq \lambda_2$. We claim that in this case

$$(\Lambda t)^n = \begin{bmatrix} (\lambda_1 t)^n & 0 \\ \lambda_{21} t^n \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2} & (\lambda_2 t)^n, \end{bmatrix}, \quad n = 0, 1, \dots. \tag{10.2.39}$$

This equation holds for the base case $n = 0$: $(\Lambda t)^0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. We show by mathematical induction that the equation holds in general. Assume (10.2.39) is true for some value of $n$. Then

$$\begin{aligned}
(\Lambda t)^{n+1} &= (\Lambda t)(\Lambda t)^n \\
&= \begin{bmatrix} \lambda_1 t & 0 \\ \lambda_{21} t & \lambda_2 t \end{bmatrix} \begin{bmatrix} (\lambda_1 t)^n & 0 \\ \lambda_{21} t^n \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2} & (\lambda_2 t)^n \end{bmatrix} \\
&= \begin{bmatrix} (\lambda_1 t)^{n+1} & 0 \\ \lambda_{21} t^{n+1} \left(\lambda_1^n + \lambda_2 \frac{\lambda_1^n - \lambda_2^n}{\lambda_1 - \lambda_2}\right) & (\lambda_2 t)^{n+1} \end{bmatrix} \\
&= \begin{bmatrix} (\lambda_1 t)^{n+1} & 0 \\ \lambda_{21} t^{n+1} \frac{\lambda_1^{n+1} - \lambda_2^{n+1}}{\lambda_1 - \lambda_2} & (\lambda_2 t)^{n+1} \end{bmatrix},
\end{aligned}$$

which is (10.2.39) with $n$ replaced by $n+1$. Having thus established (10.2.39) for all values of $n$, we have

$$\begin{aligned}
e^{\Lambda t} &= \sum_{n=0}^{\infty} \frac{1}{n!}(\Lambda t)^n \\
&= \begin{bmatrix} \sum_{n=0}^{\infty} \frac{1}{n!}(\lambda_1 t)^n & 0 \\ \frac{\lambda_{21}}{\lambda_1-\lambda_2}\left(\sum_{n=0}^{\infty} \frac{1}{n!}(\lambda_1 t)^n - \sum_{n=0}^{\infty} \frac{1}{n!}(\lambda_2 t)^n\right) & \sum_{n=0}^{\infty} \frac{1}{n!}(\lambda_2 t)^n \end{bmatrix} \\
&= \begin{bmatrix} e^{\lambda_1 t} & 0 \\ \frac{\lambda_{21}}{\lambda_1-\lambda_2}\left(e^{\lambda_1 t} - e^{\lambda_2 t}\right) & e^{\lambda_2 t} \end{bmatrix}.
\end{aligned}$$

This is (10.2.35).

We next consider the case $\lambda_1 = \lambda_2$. We claim in this case that

$$(\Lambda t)^n = \begin{bmatrix} (\lambda_1 t)^n & 0 \\ n\lambda_{21}\lambda_1^{n-1}t^n & (\lambda_1 t)^n \end{bmatrix}, \quad n = 1, 2, \ldots. \tag{10.2.40}$$

This equation holds for the base case $n = 0$. We again use mathematical induction to establish the equation for all $n$. Assume (10.2.40) holds for some value of $n$. Then

$$\begin{aligned}
(\Lambda t)^{n+1} &= (\Lambda t)(\Lambda t)^n \\
&= \begin{bmatrix} \lambda_1 t & 0 \\ \lambda_{21} t & \lambda_1 t \end{bmatrix} \begin{bmatrix} (\lambda_1 t)^n & 0 \\ n\lambda_{21}\lambda_1^{n-1}t^n & (\lambda_1 t)^n \end{bmatrix} \\
&= \begin{bmatrix} (\lambda_1 t)^{n+1} & 0 \\ \left(\lambda_{21}\lambda_1^n + n\lambda_{21}\lambda_1^n\right)t^{n+1} & (\lambda_1 t)^{n+1} \end{bmatrix} \\
&= \begin{bmatrix} (\lambda_1 t)^{n+1} & 0 \\ (n+1)\lambda_{21}\lambda_1^n t^{n+1} & (\lambda_1 t)^{n+1} \end{bmatrix},
\end{aligned}$$

which is (10.2.40) with $n$ replaced by $n+1$. Having thus established (10.2.40) for all values of $n$, we have

$$e^{\Lambda t} = \sum_{n=0}^{\infty} \frac{1}{n!}(\Lambda t)^n = \begin{bmatrix} \sum_{n=0}^{\infty} \frac{1}{n!}(\lambda_1 t)^n & 0 \\ \lambda_{21}\sum_{n=0}^{\infty} \frac{n}{n!}\lambda_1^{n-1}t^n & \sum_{n=0}^{\infty} \frac{1}{n!}(\lambda_1 t)^n \end{bmatrix}. \tag{10.2.41}$$

But

$$\lambda_{21}\sum_{n=0}^{\infty} \frac{n}{n!}\lambda_1^{n-1}t^n = \lambda_{21}\frac{d}{d\lambda_1}\sum_{n=0}^{\infty} \frac{1}{n!}(\lambda_1 t)^n = \lambda_{21}\frac{d}{d\lambda_1}e^{\lambda_1 t} = \lambda_{21} t e^{\lambda_1 t}.$$

Substituting this into (10.2.41), we obtain (10.2.36).

When $\lambda_1 \neq \lambda_2$, we have

$$\frac{d}{dt}e^{\Lambda t} = \begin{bmatrix} \lambda_1 e^{\lambda_1 t} & 0 \\ \frac{\lambda_{21}}{\lambda_1-\lambda_2}\left(\lambda_1 e^{\lambda_1 t} - \lambda_2 e^{\lambda_2 t}\right) & \lambda_2 e^{\lambda_2 t} \end{bmatrix}$$

and

$$e^{-\Lambda t} = \begin{bmatrix} e^{-\lambda_1 t} & 0 \\ \frac{\lambda_{21}}{\lambda_1-\lambda_2}\left(e^{-\lambda_1 t} - e^{-\lambda_2 t}\right) & e^{-\lambda_2 t} \end{bmatrix}.$$

When $\lambda_1 = \lambda_2$,

$$\frac{d}{dt}e^{\Lambda t} = \begin{bmatrix} \lambda_1 e^{\lambda_1 t} & 0 \\ \lambda_{21}(1+\lambda_1 t)e^{\lambda_1 t} & \lambda_1 e^{\lambda_1 t} \end{bmatrix}$$

and

$$e^{-\Lambda t} = \begin{bmatrix} e^{-\lambda_1 t} & 0 \\ -\lambda_{21} t e^{-\lambda_1 t} & e^{-\lambda_1 t} \end{bmatrix}.$$

The verification of (10.2.37) and (10.2.38) can be done by straightforward matrix multiplications. □

We use (10.2.34) to compute

$$d\left(e^{\Lambda t}Y(t)\right) = e^{\Lambda t}\left(\Lambda Y(t)\,dt + dY(t)\right) = e^{\Lambda t}\,d\widetilde{W}(t).$$

Integration from 0 to $t$ yields

$$e^{\Lambda t}Y(t) = Y(0) + \int_0^t e^{\Lambda u}\,d\widetilde{W}(u).$$

We solve for

$$\begin{aligned} Y(t) &= e^{-\Lambda t}Y(0) + e^{-\Lambda t}\int_0^t e^{\Lambda u}\,d\widetilde{W}(u) \\ &= e^{-\Lambda t}Y(0) + \int_0^t e^{-\Lambda(t-u)}\,d\widetilde{W}(u). \end{aligned} \tag{10.2.42}$$

If $\lambda_1 \neq \lambda_2$, equation (10.2.42) may be written componentwise as

$$Y_1(t) = e^{-\lambda_1 t}Y_1(0) + \int_0^t e^{-\lambda_1(t-u)}\,d\widetilde{W}_1(u), \tag{10.2.43}$$

$$\begin{aligned} Y_2(t) = {} & \frac{\lambda_{21}}{\lambda_1-\lambda_2}\left(e^{-\lambda_1 t} - e^{-\lambda_2 t}\right)Y_1(0) + e^{-\lambda_2 t}Y_2(0) \\ & + \frac{\lambda_{21}}{\lambda_1-\lambda_2}\int_0^t \left(e^{-\lambda_1(t-u)} - e^{-\lambda_2(t-u)}\right) d\widetilde{W}_1(u) \\ & + \int_0^t e^{-\lambda_2(t-u)}\,d\widetilde{W}_2(u). \end{aligned} \tag{10.2.44}$$

If $\lambda_1 = \lambda_2$, then the componentwise form of (10.2.42) is

$$Y_1(t) = e^{-\lambda_1 t}Y_1(0) + \int_0^t e^{-\lambda_1(t-u)}\,d\widetilde{W}_1(u), \tag{10.2.45}$$

$$\begin{aligned} Y_2(t) = {} & -\lambda_{21} t e^{-\lambda_1 t}Y_1(0) + e^{-\lambda_1 t}Y_2(0) \\ & -\lambda_{21}\int_0^t (t-u)e^{-\lambda_1(t-u)}\,d\widetilde{W}_1(u) + \int_0^t e^{-\lambda_1(t-u)}\,d\widetilde{W}_2(u). \end{aligned} \tag{10.2.46}$$

Being nonrandom quantities plus Itô integrals of nonrandom integrands, the processes $Y_1(t)$ and $Y_2(t)$ are Gaussian, and so $R(t) = \delta_0 + \delta_1 Y_1(t) + \delta_2 Y_2(t)$ is normally distributed. The statistics of $Y_1(t)$ and $Y_2(t)$ are provided in Exercise 10.1.

### 10.2.2 Two-Factor CIR Model

In the two-factor Vasicek model, the canonical factors $Y_1(t)$ and $Y_2(t)$ are jointly normally distributed. Because these factors are driven by independent Brownian motions, they are not perfectly correlated and hence, for all $t > 0$,

$$R(t) = \delta_0 + \delta_1 Y_1(t) + \delta_2 Y_2(t) \tag{10.2.47}$$

is a normal random variable with positive variance except in the degenerate case $\delta_1 = \delta_2 = 0$. In particular, for each $t > 0$, there is a positive probability that $R(t)$ is strictly negative.

In the *two-factor Cox-Ingersoll-Ross model (CIR)* of this subsection, both factors are guaranteed to be nonnegative at all times almost surely. We again define the interest rate by (10.2.47) but now assume that

$$\delta_0 \geq 0, \quad \delta_1 > 0, \quad \delta_2 > 0. \tag{10.2.48}$$

We take the initial interest rate $R(0)$ to be nonnegative, and then we have $R(t) \geq 0$ for all $t \geq 0$ almost surely.

The evolution of the factor processes in the canonical two-factor CIR model is given by

$$dY_1(t) = \big(\mu_1 - \lambda_{11} Y_1(t) - \lambda_{12} Y_2(t)\big)\, dt + \sqrt{Y_1(t)}\, d\widetilde{W}_1(t), \tag{10.2.49}$$

$$dY_2(t) = \big(\mu_2 - \lambda_{21} Y_1(t) - \lambda_{22} Y_2(t)\big)\, dt + \sqrt{Y_2(t)}\, d\widetilde{W}_2(t). \tag{10.2.50}$$

In addition to (10.2.48), we assume

$$\mu_1 \geq 0,\ \mu_2 \geq 0,\ \lambda_{11} > 0,\ \lambda_{22} > 0,\ \lambda_{12} \leq 0,\ \lambda_{21} \leq 0. \tag{10.2.51}$$

These conditions guarantee that although the drift term $\mu_1 - \lambda_{11} Y_1(t) - \lambda_{12} Y_2(t)$ can be negative, it is nonnegative whenever $Y_1(t) = 0$ and $Y_2(t) \geq 0$. Similarly, the drift term $\mu_2 - \lambda_{21} Y_1(t) - \lambda_{22} Y_2(t)$ is nonnegative whenever $Y_2(t) = 0$ and $Y_1(t) \geq 0$. Starting with $Y_1(0) \geq 0$ and $Y_2(0) \geq 0$, we have $Y_1(t) \geq 0$ and $Y_2(t) \geq 0$ for all $t \geq 0$ almost surely.

The Brownian motions $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ in (10.2.49) and (10.2.50) are assumed to be independent. We do not need this assumption to guarantee nonnegativity of $Y_1(t)$ and $Y_2(t)$ but rather to obtain the affine-yield result below; see Remark 10.2.4.

**Bond Prices**

We derive the formula for zero-coupon bond prices in the canonical two-factor CIR model. As in the two-factor Vasicek model, the price at time $t$ of a zero-coupon bond maturing at a later time $T$ must be of the form

$$B(t,T) = f\big(t, Y_1(t), Y_2(t)\big)$$

for some function $f(t, y_1, y_2)$. The discounted bond price has differential

$$\begin{aligned}
&d\big(D(t)B(t,T)\big)\\
&= d\Big(D(t)f\big(t,Y_1(t),Y_2(t)\big)\Big)\\
&= -R(t)D(t)f\big(t,Y_1(t),Y_2(t)\big)\,dt + D(t)\,df\big(t,Y_1(t),Y_2(t)\big)\\
&= D\Big[-Rf\,dt + f_t\,dt + f_{y_1}\,dY_1 + f_{y_2}\,dY_2\\
&\qquad + \frac{1}{2}f_{y_1y_1}\,dY_1\,dY_1 + f_{y_1y_2}\,dY_1\,dY_2 + \frac{1}{2}f_{y_2y_2}\,dY_2\,dY_2\Big]\\
&= D\Bigg[-(\delta_0+\delta_1Y_1+\delta_2Y_2)f + f_t + (\mu_1 - \lambda_{11}Y_1 - \lambda_{12}Y_2)f_{y_1}\\
&\qquad + (\mu_2 - \lambda_{21}Y_1 - \lambda_{22}Y_2)f_{y_2} + \frac{1}{2}Y_1f_{y_1y_1} + \frac{1}{2}Y_2f_{y_2y_2}\Bigg]\,dt\\
&\qquad + D\Big[\sqrt{Y_1}\,f_{y_1}\,d\widetilde{W}_1 + \sqrt{Y_2}\,f_{y_2}\,d\widetilde{W}_2\Big].
\end{aligned}$$

Setting the $dt$ term equal to zero, we obtain the partial differential equation

$$\begin{aligned}
&-(\delta_0+\delta_1y_1+\delta_2y_2)f(t,y_1,y_2) + f_t(t,y_1,y_2)\\
&\quad + (\mu_1 - \lambda_{11}y_1 - \lambda_{12}y_2)f_{y_1}(t,y_1,y_2) + (\mu_2 - \lambda_{21}y_1 - \lambda_{22}y_2)f_{y_2}(t,y_1,y_2)\\
&\qquad + \frac{1}{2}y_1f_{y_1y_1}(t,y_1,y_2) + \frac{1}{2}y_2f_{y_2y_2}(t,y_1,y_2) = 0 \qquad (10.2.52)
\end{aligned}$$

for all $t \in [0,T)$ and all $y_1 \geq 0$, $y_2 \geq 0$. To solve this equation, we seek a solution of the affine-yield form

$$f(t,y_1,y_2) = e^{-y_1C_1(T-t)-y_2C_2(T-t)-A(T-t)} \qquad (10.2.53)$$

for some functions $C_1(\tau)$, $C_2(\tau)$, and $A(\tau)$, where $\tau = T - t$. The terminal condition

$$f\big(T, Y_1(T), Y_2(T)\big) = B(T,T) = 1$$

implies

$$C_1(0) = C_2(0) = A(0) = 0. \qquad (10.2.54)$$

With $'$ denoting differentiation with respect to $\tau$, we have $\frac{d}{dt}C_i(\tau) = -C_i'(\tau)$, $i = 1, 2$, $\frac{d}{dt}A(\tau) = -A'(\tau)$, and (10.2.52) becomes

$$\Bigg[\left(C_1' + \lambda_{11}C_1 + \lambda_{21}C_2 + \frac{1}{2}C_1^2 - \delta_1\right)y_1$$
$$+\left(C_2' + \lambda_{12}C_1 + \lambda_{22}C_2 + \frac{1}{2}C_2^2 - \delta_2\right)y_2$$
$$+ (A' - \mu_1 C_1 - \mu_2 C_2 - \delta_0)\Bigg] f = 0. \quad (10.2.55)$$

Because (10.2.55) must hold for all $y_1 \geq 0$ and $y_2 \geq 0$, the term $C_1' + \lambda_{11}C_1 + \lambda_{21}C_2 + \frac{1}{2}C_1^2 - \delta_1$ multiplying $y_1$ must be zero. Similarly, the term $C_2' + \lambda_{12}C_1 + \lambda_{22}C_2 + \frac{1}{2}C_2^2 - \delta_2$ multiplying $y_2$ must be zero, and consequently the remaining term $A' - \mu_1 C_2 - \mu_2 C_2 - \delta_0$ must also be zero. This gives us a system of three ordinary differential equations:

$$C_1'(\tau) = -\lambda_{11}C_1(\tau) - \lambda_{21}C_2(\tau) - \frac{1}{2}C_1^2(\tau) + \delta_1, \quad (10.2.56)$$
$$C_2'(\tau) = -\lambda_{12}C_1(\tau) - \lambda_{22}C_2(\tau) - \frac{1}{2}C_2^2(\tau) + \delta_2, \quad (10.2.57)$$
$$A'(\tau) = \mu_1 C_1(\tau) + \mu_2 C_2(\tau) + \delta_0. \quad (10.2.58)$$

The solution to these equations satisfying the initial condition (10.2.54) can be found numerically. Solving this system of ordinary differential equations numerically is simpler than solving the partial differential equation (10.2.52).

*Remark 10.2.4.* We note that if the Brownian motions $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ in (10.2.49) and (10.2.50) were correlated with some correlation coefficient $\rho \neq 0$, then the partial differential equation (10.2.52) would have the additional term $\rho\sqrt{y_1 y_2}\, f_{y_1 y_2}$ on the left-hand side. This term would ruin the argument that led to the system of ordinary differential equations (10.2.56)–(10.2.58). For this reason, we assume at the outset that these Brownian motions are independent.

### 10.2.3 Mixed Model

Both factors in the two-factor CIR model are always nonnegative. In the two-factor Vasicek model, both factors can become negative. In the two-factor mixed model, one of the factors is always nonnegative and the other can become negative.

The *canonical two-factor mixed model* is

$$dY_1(t) = \big(\mu - \lambda_1 Y_1(t)\big)\, dt + \sqrt{Y_1(t)}\, d\widetilde{W}_1(t), \quad (10.2.59)$$
$$dY_2(t) = -\lambda_2 Y_2(t)\, dt + \sigma_{21}\sqrt{Y_1(t)}\, d\widetilde{W}_1(t)$$
$$+ \sqrt{\alpha + \beta Y_1(t)}\, d\widetilde{W}_2(t). \quad (10.2.60)$$

We assume $\mu \geq 0$, $\lambda_1 > 0$, $\lambda_2 > 0$, $\alpha \geq 0$, $\beta \geq 0$, and $\sigma_{21} \in \mathbb{R}$. The Brownian motions $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ are independent. We assume $Y_1(0) \geq 0$, and we have $Y_1(t) \geq 0$ for all $t \geq 0$ almost surely. On the other hand, even if $Y_2(t)$ is positive, $Y_2(t)$ can take negative values for $t > 0$. The interest rate is defined by

$$R(t) = \delta_0 + \delta_1 Y_1(t) + \delta_2 Y_2(t). \tag{10.2.61}$$

In this model, zero-coupon bond prices have the affine-yield form

$$B(t,T) = e^{-Y_1(t)C_1(T-t) - Y_2(t)C_2(T-t) - A(T-t)}. \tag{10.2.62}$$

Just as in the two-factor Vasicek model and the two-factor CIR model, the functions $C_1(\tau)$, $C_2(\tau)$, and $A(\tau)$ must satisfy the terminal condition

$$C_1(0) = C_2(0) = A(0). \tag{10.2.63}$$

Exercise 10.2 derives the system of ordinary differential equations that determine the functions $C_1(\tau)$, $C_2(\tau)$, and $A(\tau)$.

## 10.3 Heath-Jarrow-Morton Model

The Heath-Jarrow-Morton (HJM) model of this section evolves the whole yield curve forward in time. There are several possible ways to represent the yield curve, and the one chosen by the HJM model is in terms of the *forward rates* that can be locked in at one time for borrowing at a later time. In this section, we first discuss forward rates, then write down the HJM model for their evolution, discuss how to guarantee that the resulting model does not admit arbitrage, and conclude with a procedure for calibrating the HJM model.

### 10.3.1 Forward Rates

Let us fix a time horizon $\overline{T}$ (say 50 years). All bonds in the following discussion will mature at or before time $\overline{T}$. For $0 \le t \le T \le \overline{T}$, as before, we denote by $B(t,T)$ the price at time $t$ of a zero-coupon bond maturing at time $T$ and having face value 1. We assume this bond bears no risk of default. We assume further that, for every $t$ and $T$ satisfying $0 \le t \le T \le \overline{T}$, the bond price $B(t,T)$ is defined. If the interest rate is strictly positive between times $t$ and $T$, then $B(t,T)$ must be strictly less than one whenever $t < T$. This is the situation to keep in mind, although some implementations of the HJM model violate it.

At time $t$, we can engage in *forward investing* at the later time $T$ by setting up the following portfolio. Here $\delta$ is a small positive number.

- Take a short position of size 1 in $T$-maturity bonds. This generates income $B(t,T)$.
- Take a long position of size $\frac{B(t,T)}{B(t,T+\delta)}$ in $(T+\delta)$-maturity bonds. This costs $B(t,T)$.

The net cost of setting up this portfolio at time $t$ is zero. At the later time $T$, holding this portfolio requires that we pay 1 to cover the short position in the $T$-maturity bond. At the still later time $T+\delta$, we receive $\frac{B(t,T)}{B(t,T+\delta)}$ from the long position in the $T+\delta$-maturity bond. In other words, we have invested 1 at time $T$ and received more than 1 at time $T+\delta$. The yield that explains the surplus received at time $T+\delta$ is

$$\frac{1}{\delta}\log\frac{B(t,T)}{B(t,T+\delta)} = -\frac{\log B(t,T+\delta)-\log B(t,T)}{\delta}. \tag{10.3.1}$$

This is the continuously compounding rate of interest that, applied to the 1 invested at time $T$, would return $\frac{B(t,T)}{B(t,T+\delta)}$ at time $T+\delta$. If the bond $B(t,T+\delta)$ with the longer time to maturity has the smaller price, as it would if the interest rate is strictly positive, then the quotient $\frac{B(t,T)}{B(t,T+\delta)}$ is strictly greater than 1 and the yield is strictly positive.

Note that the yield in (10.3.1) is $\mathcal{F}(t)$-measurable. Although it is an interest rate for investing at time $T$, it can be "locked in" at the earlier time $t$. In fact, if someone were to propose any other interest rate for investing (or borrowing) at time $T$ that is set at the earlier time $t$, then by accepting this interest rate and setting up the portfolio described above or its opposite, one could create an arbitrage.

We define the *forward rate at time t for investing at time T* to be

$$\begin{aligned} f(t,T) &= -\lim_{\delta\downarrow 0}\frac{\log B(t,T+\delta)-\log B(t,T)}{\delta} \\ &= -\frac{\partial}{\partial T}\log B(t,T). \end{aligned} \tag{10.3.2}$$

This is the limit of the yield in (10.3.1) as $\delta \downarrow 0$ and can thus be regarded as the *instantaneous interest rate* at time $T$ that can be locked in at the earlier time $t$.

If we know $f(t,T)$ for all values of $0 \le t \le T \le \overline{T}$, we can recover $B(t,T)$ for all values of $0 \le t \le T \le \overline{T}$ by the formula

$$\int_t^T f(t,v)\,dv = -\big[\log B(t,T) - \log B(t,t)\big] = -\log B(t,T),$$

where we have used the fact that $B(t,t)=1$. Therefore,

$$B(t,T) = \exp\left\{-\int_t^T f(t,v)\,dv\right\}, \quad 0 \le t \le T \le \overline{T}. \tag{10.3.3}$$

From bond prices, we can determine forward rates from (10.3.2). From forward rates, we can determine bond prices from (10.3.3). Therefore, at least theoretically, it does not appear to matter whether we build a model for forward rates or for bond prices. In fact, the no-arbitrage condition works out to have a simple form when we model forward rates. From a practical point of view, forward rates are a more difficult object to determine from market data because the differentiation in (10.3.2) is sensitive to small changes in the bond prices. On the other hand, once we have forward rates, bond prices are easy to determine because the integration in (10.3.3) is not sensitive to small changes in the forward rates.

The interest rate at time $t$ is

$$R(t) = f(t,t). \tag{10.3.4}$$

This is the instantaneous rate we can lock in at time $t$ for borrowing at time $t$.

### 10.3.2 Dynamics of Forward Rates and Bond Prices

Assume that $f(0,T)$, $0 \le T \le \overline{T}$, is known at time 0. We call this the *initial forward rate curve*. In the HJM model, the forward rate at later times $t$ for investing at still later times $T$ is given by

$$f(t,T) = f(0,T) + \int_0^t \alpha(u,T)\,du + \int_0^t \sigma(u,T)\,dW(u). \tag{10.3.5}$$

We may write this in differential form as

$$df(t,T) = \alpha(t,T)\,dt + \sigma(t,T)\,dW(t), \quad 0 \le t \le T. \tag{10.3.6}$$

Here and elsewhere in this section, $d$ indicates the differential with respect to the variable $t$; the variable $T$ is being held constant in (10.3.6).

Here the process $W(u)$ is a Brownian motion under the actual measure $\mathbb{P}$. In particular, $\alpha(t,T)$ is the drift of $f(t,T)$ under the actual measure. The processes $\alpha(t,T)$ and $\sigma(t,T)$ may be random. For each fixed $T$, they are adapted processes in the $t$ variable. To simplify the notation, we assume that the forward rate is driven by a single Brownian motion. The case when the forward rate is driven by multiple Brownian motions is addressed in Exercise 10.9.

From (10.3.6), we can work out the dynamics of the bond prices given by (10.3.3). Note first that because $-\int_t^T f(t,v)\,dv$ has a $t$-variable in two places, its differential has two terms. Indeed,

$$d\left(-\int_t^T f(t,v)\,dv\right) = f(t,t)\,dt - \int_t^T df(t,v)\,dv.$$

The first term on the right-hand side is the result of taking the differential with respect to the lower limit of integration $t$. The fact that this is the *lower* limit produces a minus sign, which cancels the minus sign on the left-hand side. The other term is the result of taking the differential with respect to the $t$ under the integral sign. Using (10.3.4) and (10.3.6), we see that

$$d\left(-\int_t^T f(t,v)\,dv\right) = R(t)\,dt - \int_t^T \left[\alpha(t,v)\,dt + \sigma(t,v)\,dW(t)\right]dv.$$

We next reverse the order of the integration (see Exercise 10.8), writing

$$\int_t^T \alpha(t,v)\,dt\,dv = \int_t^T \alpha(t,v)\,dv\,dt \;=\; \alpha^*(t,T)\,dt, \tag{10.3.7}$$

$$\int_t^T \sigma(t,v)\,dW(t)\,dv = \int_t^T \sigma(t,v)\,dv\,dW(t) \;=\; \sigma^*(t,T)\,dW(t), \tag{10.3.8}$$

where

$$\alpha^*(t,T) = \int_t^T \alpha(t,v)\,dv, \quad \sigma^*(t,T) = \int_t^T \sigma(t,v)\,dv. \tag{10.3.9}$$

In conclusion, we have

$$d\left(-\int_t^T f(t,v)\,dv\right) = R(t)\,dt - \alpha^*(t,T)\,dt - \sigma^*(t,T)\,dW(t). \tag{10.3.10}$$

Let $g(x) = e^x$, so that $g'(x) = e^x$ and $g''(x) = e^x$. According to (10.3.3),

$$B(t,T) = g\left(-\int_t^T f(t,v)\,dv\right).$$

The Itô-Doeblin formula implies

$$\begin{aligned}
dB(t,T) &= g'\left(-\int_t^T f(t,v)\,dv\right) d\left(-\int_t^T f(t,v)\,dv\right)\\
&\quad +\frac{1}{2}g''\left(-\int_t^T f(t,v)\,dv\right)\left[d\left(-\int_t^T f(t,v)\,dv\right)\right]^2\\
&= B(t,T)\left[R(t)\,dt - \alpha^*(t,T)\,dt - \sigma^*(t,T)\,dW(t)\right]\\
&\quad +\frac{1}{2}B(t,T)\left(\sigma^*(t,T)\right)^2 dt\\
&= B(t,T)\left[R(t) - \alpha^*(t,T) + \frac{1}{2}\left(\sigma^*(t,T)\right)^2\right] dt\\
&\quad -\sigma^*(t,T)B(t,T)\,dW(t).
\end{aligned} \tag{10.3.11}$$

### 10.3.3 No-Arbitrage Condition

The HJM model has a zero-coupon bond with maturity $T$ for every $T \in [0,\overline{T}]$. We need to make sure there is no opportunity for arbitrage by trading in these bonds. The First Fundamental Theorem of Asset Pricing, Theorem 5.4.7, says that, in order to guarantee this, we should seek a probability measure $\widetilde{\mathbb{P}}$ under which each discounted bond price

$$D(t)B(t,T) = \exp\left\{-\int_0^t R(u)\,du\right\} B(t,T), \quad 0 \le t \le T,$$

is a martingale. Because $dD(t) = -R(t)D(t)\,dt$, we have the differential

$$\begin{aligned} &d\big(D(t)B(t,T)\big) \\ &= -R(t)D(t)B(t,T)\,dt + D(t)\,dB(t,T) \\ &= D(t)B(t,T)\left[\left(-\alpha^*(t,T) + \frac{1}{2}(\sigma^*(t,T))^2\right)dt - \sigma^*(t,T)\,dW(t)\right]. \quad (10.3.12) \end{aligned}$$

We want to write the term in square brackets as

$$-\sigma^*(t,T)[\Theta(t)\,dt + dW(t)],$$

and we can then use Girsanov's Theorem, Theorem 5.2.3, to change to a probability measure $\widetilde{\mathbb{P}}$ under which

$$\widetilde{W}(t) = \int_0^t \Theta(u)\,du + W(t) \tag{10.3.13}$$

is a Brownian motion. Using this Brownian motion, we may rewrite (10.3.12) as

$$d\big(D(t)B(t,T)\big) = -D(t)B(t,T)\sigma^*(t,T)\,d\widetilde{W}(t). \tag{10.3.14}$$

It would then follow that $D(t)B(t,T)$ is a martingale under $\widetilde{\mathbb{P}}$ (i.e., $\widetilde{\mathbb{P}}$ would be risk-neutral).

For the program above to work, we must solve the equation

$$\begin{aligned} &\left[\left(-\alpha^*(t,T) + \frac{1}{2}(\sigma^*(t,T))^2\right)dt - \sigma^*(t,T)\,dW(t)\right] \\ &\qquad\qquad = -\sigma^*(t,T)\big[\Theta(t)\,dt + dW(t)\big] \end{aligned}$$

for $\Theta(t)$. In other words, we must find a process $\Theta(t)$ satisfying

$$-\alpha^*(t,T) + \frac{1}{2}\big(\sigma^*(t,T)\big)^2 = -\sigma^*(t,T)\Theta(t). \tag{10.3.15}$$

Actually, (10.3.15) represents infinitely many equations, one for each maturity $T \in (0,\overline{T}]$. These are the *market price of risk equations*, and we have one such equation for each bond (maturity). However, there is only one process $\Theta(t)$. This process is the *market price of risk*, and we have as many such processes as there are sources of uncertainty. In this case, there is only one Brownian motion driving the model.

To solve (10.3.15), we recall from (10.3.9) that

$$\frac{\partial}{\partial T}\alpha^*(t,T) = \alpha(t,T), \quad \frac{\partial}{\partial T}\sigma^*(t,T) = \sigma(t,T).$$

Differentiating (10.3.15) with respect to $T$, we obtain

$$-\alpha(t,T) + \sigma^*(t,T)\sigma(t,T) = -\sigma(t,T)\Theta(t)$$

or, equivalently,

$$\alpha(t,T) = \sigma(t,T)\big[\sigma^*(t,T) + \Theta(t)\big]. \tag{10.3.16}$$

**Theorem 10.3.1 (Heath-Jarrow-Morton no-arbitrage condition).** *A term-structure model for zero-coupon bond prices of all maturities in $(0,\overline{T}]$ and driven by a single Brownian motion does not admit arbitrage if there exists a process $\Theta(t)$ such that (10.3.16) holds for all $0 \le t \le T \le \overline{T}$. Here $\alpha(t,T)$ and $\sigma(t,T)$ are the drift and diffusion, respectively, of the forward rate (i.e., the processes satisfying (10.3.6)), $\sigma^*(t,T) = \int_t^T \sigma(t,v)\,dv$, and $\Theta(t)$ is the market price of risk.*

PROOF: It remains only to check that if $\Theta(t)$ solves (10.3.16), then it also satisfies (10.3.15), for then we can use Girsanov's Theorem as described above to construct a risk-neutral measure. The existence of a risk-neutral measure guarantees the absence of arbitrage.

Suppose $\Theta(t)$ solves (10.3.16). We rewrite this equation, replacing $T$ by $v$:

$$\alpha(t,v) = \sigma(t,v)\big[\sigma^*(t,v) + \Theta(t)\big].$$

Integrating with respect to $v$ from $v=t$ to $v=T$, we obtain

$$\alpha^*(t,v)\Big|_{v=t}^{v=T} = \frac{1}{2}\big(\sigma^*(t,v)\big)^2\Big|_{v=t}^{v=T} + \sigma^*(t,v)\Theta(t)\Big|_{v=t}^{v=T}.$$

But because $\alpha^*(t,t) = \sigma^*(t,t) = 0$, this reduces to

$$\alpha^*(t,T) = \frac{1}{2}\big(\sigma^*(t,T)\big)^2 + \sigma^*(t,T)\Theta(t),$$

which is (10.3.15). □

So long as $\sigma(t,T)$ is nonzero, we can solve (10.3.16) for $\Theta(t)$:

$$\Theta(t) = \frac{\alpha(t,T)}{\sigma(t,T)} - \sigma^*(t,T), \quad 0 \le t \le T. \tag{10.3.17}$$

This shows that $\Theta(t)$ is unique, and hence the risk-neutral measure is unique. In this case, the Second Fundamental Theorem of Asset Pricing, Theorem 5.4.9, guarantees that the model is complete (i.e., all interest rate derivatives can be hedged by trading in zero-coupon bonds).

### 10.3.4 HJM Under Risk-Neutral Measure

We began with the formula (10.3.5) for the evolution of the forward rate, and the driving process $W(u)$ appearing in (10.3.5) is a Brownian motion under the actual measure $\mathbb{P}$. Assuming the model satisfies the HJM no-arbitrage condition (10.3.16), we may rewrite (10.3.5) as

$$\begin{aligned} df(t,T) &= \alpha(t,T)\,dt + \sigma(t,T)\,dW(t) \\ &= \sigma(t,T)\sigma^*(t,T)\,dt + \sigma(t,T)\Big[\Theta(t) + dW(t)\Big] \\ &= \sigma(t,T)\sigma^*(t,T)\,dt + \sigma(t,T)\,d\widetilde{W}(t), \end{aligned}$$

where $\widetilde{W}(t)$ is given by (10.3.13). To conclude that there is no arbitrage, we need the drift of the forward rate under the risk-neutral measure to be $\sigma(t,T)\sigma^*(t,T)$. We saw in the proof of Theorem 10.3.1 that the no-arbitrage condition (10.3.16) implies (10.3.15), and using (10.3.15) we may rewrite the differential of the discounted bond price (10.3.12) as

$$\begin{aligned} d\big(D(t)B(t,T)\big) &= -\sigma^*(t,T)D(t)B(t,T)\big[\Theta(t)\,dt + dW(t)\big] \\ &= -\sigma^*(t,T)D(t)B(t,T)\,d\widetilde{W}(t). \end{aligned}$$

Because $d\frac{1}{D(t)} = \frac{R(t)}{D(t)}\,dt$, the differential of the undiscounted bond price is

$$\begin{aligned} dB(t,T) &= d\left(\frac{1}{D(t)}\cdot D(t)B(t,T)\right) \\ &= \frac{R(t)}{D(t)}D(t)B(t,T)\,dt - \sigma^*(t,T)\frac{1}{D(t)}D(t)B(t,T)\,d\widetilde{W}(t) \\ &= R(t)B(t,T)\,dt - \sigma^*(t,T)B(t,T)\,d\widetilde{W}(t). \end{aligned}$$

The following theorem summarizes this discussion.

**Theorem 10.3.2 (Term-structure evolution under risk-neutral measure).** *In a term-structure model satisfying the HJM no-arbitrage condition of Theorem 10.3.1, the forward rates evolve according to the equation*

$$df(t,T) = \sigma(t,T)\sigma^*(t,T)\,dt + \sigma(t,T)\,d\widetilde{W}(t), \tag{10.3.18}$$

*and the zero-coupon bond prices evolve according to the equation*

$$dB(t,T) = R(t)B(t,T)\,dt - \sigma^*(t,T)B(t,T)\,d\widetilde{W}(t), \tag{10.3.19}$$

*where $\widetilde{W}(t)$ is a Brownian motion under a risk-neutral measure $\widetilde{\mathbb{P}}$. Here $\sigma^*(t) = \int_t^T \sigma(t,v)\,dv$ and $R(t) = f(t,t)$ is the interest rate. The discounted bond prices satisfy*

$$d\big(D(t)B(t,T)\big) = -\sigma^*(t,T)D(t)B(t,T)\,d\widetilde{W}(t), \tag{10.3.20}$$

*where $D(t) = e^{-\int_0^t R(u)du}$ is the discount process. The solution to the stochastic differential equation (10.3.19) is*

$$
\begin{aligned}
&B(t,T)\\
&= B(0,T)\exp\left\{\int_0^t R(u)du - \int_0^t \sigma^*(u,T)\,d\widetilde{W}(u) - \frac{1}{2}\int_0^T (\sigma^*(u,T))^2\,du\right\}\\
&= \frac{B(0,T)}{D(t)}\exp\left\{-\int_0^t \sigma^*(u,T)\,d\widetilde{W}(u) - \frac{1}{2}\int_0^T \left(\sigma^*(u,T)\right)^2 du\right\}. \quad (10.3.21)
\end{aligned}
$$

### 10.3.5 Relation to Affine-Yield Models

Every term-structure model driven by Brownian motion is an HJM model. In any such model, there are forward rates. The drift and diffusion of the forward rates must satisfy the conditions of Theorem 10.3.1 in order for a risk-neutral measure to exist, which rules out arbitrage. Under these conditions, the formulas of Theorem 10.3.2 describe the evolution of the forward rates and bonds under the risk-neutral measure.

We illustrate this with the one-factor Hull-White and Cox-Ingersoll-Ross (CIR) models of Examples 6.5.1 and 6.5.2. For both these models, the interest rate dynamics are of the form

$$dR(t) = \beta\big(t, R(t)\big)\,dt + \gamma\big(t, R(t)\big)\,d\widetilde{W}(t),$$

where $\widetilde{W}(t)$ is a Brownian motion under a risk-neutral probability measure $\widetilde{\mathbb{P}}$. In the case of the Hull-White model,

$$\beta(t,r) = a(t) - b(t)r, \quad \gamma(t,r) = \sigma(t),$$

for some nonrandom positive functions $a(t)$, $b(t)$, and $\sigma(t)$. For the CIR model,

$$\beta(t,r) = a - br, \quad \gamma(t,r) = \sigma\sqrt{r}, \tag{10.3.22}$$

for some positive constants $a$, $b$, and $\sigma$. The zero-coupon bond prices are of the form

$$B(t,T) = e^{-R(t)C(t,T) - A(t,T)}, \tag{10.3.23}$$

where $C(t,T)$ and $A(t,T)$ are nonrandom functions. In the case of the Hull-White model, $C(t,T)$ and $A(t,T)$ are given by (6.5.10) and (6.5.11), which we repeat here:

$$C(t,T) = \int_t^T e^{-\int_t^s b(v)dv}\,ds, \tag{10.3.24}$$

$$A(t,T) = \int_t^T \left(a(s)C(s,T) - \frac{1}{2}\sigma^2(s)C^2(s,T)\right) ds. \tag{10.3.25}$$

In the case of the CIR model, $C(t,T)$ and $A(t,T)$ are given by (6.5.16) and (6.5.17). According to (10.3.2), the forward rates are

$$f(t,T) = -\frac{\partial}{\partial T}\log B(t,T) = R(t)\frac{\partial}{\partial T}C(t,T) + \frac{\partial}{\partial T}A(t,T).$$

With $C'(t,T)$ and $A'(t,T)$ denoting derivatives with respect to $t$, we have the forward rate differential

$$\begin{aligned}
df(t,T) &= \frac{\partial}{\partial T}C(t,T)\,dR(t) + R(t)\frac{\partial}{\partial T}C'(t,T)\,dt + \frac{\partial}{\partial T}A'(t,T)\,dt \\
&= \left[\frac{\partial}{\partial T}C(t,T)\beta\big(t,R(t)\big) + R(t)\frac{\partial}{\partial T}C'(t,T) + \frac{\partial}{\partial T}A'(t,T)\right]dt \\
&\quad + \frac{\partial}{\partial T}C(t,T)\gamma(t,R(t))\,d\widetilde{W}(t).
\end{aligned}$$

This is an HJM model with

$$\sigma(t,T) = \frac{\partial}{\partial T}C(t,T)\gamma\big(t,R(t)\big). \tag{10.3.26}$$

Since we are working under the risk-neutral measure, Theorem 10.3.2 implies that the drift term should be $\sigma(t,T)\sigma^*(t,T) = \sigma(t,T)\int_t^T\sigma(t,v)\,dv$. In other words, for these affine-yield models, the HJM no-arbitrage condition becomes

$$\begin{aligned}
&\frac{\partial}{\partial T}C(t,T)\beta\big(t,R(t)\big) + R(t)\frac{\partial}{\partial T}C'(t,T) + \frac{\partial}{\partial T}A'(t,T) \\
&= \left(\frac{\partial}{\partial T}C(t,T)\right)\gamma\big(t,R(t)\big)\int_t^T\frac{\partial}{\partial v}C(t,v)\gamma\big(t,R(t)\big)\,dv \\
&= \left(\frac{\partial}{\partial T}C(t,T)\right)\gamma\big(t,R(t)\big)\big[C(t,T)-C(t,t)\big]\gamma\big(t,R(t)\big) \\
&= \left(\frac{\partial}{\partial T}C(t,T)\right)C(t,T)\gamma^2\big(t,R(t)\big).
\end{aligned} \tag{10.3.27}$$

We verify (10.3.27) for the *Vasicek model*, which is the Hull-White model with constant $a$, $b$, and $\sigma$, and we leave the verification for the Hull-White and CIR models as Exercise 10.10. For the Vasicek model, (10.3.24) and (10.3.25) reduce to

$$\begin{aligned}
C(t,T) &= \frac{1}{b}\left(1-e^{-b(T-t)}\right), \\
A'(t,T) &= -aC(t,T) + \frac{1}{2}\sigma^2C^2(t,T),
\end{aligned}$$

and hence

$$\begin{aligned}
\frac{\partial}{\partial T}C(t,T) &= e^{-b(T-t)}, \\
\frac{\partial}{\partial T}A'(t,T) &= \left(\frac{\sigma^2}{b}-a\right)e^{-b(T-t)} - \frac{\sigma^2}{a}e^{-2b(T-t)}.
\end{aligned}$$

Therefore,

$$\sigma(t,T) = \sigma e^{-b(T-t)},$$

and

$$\sigma^*(t,T) = \int_t^T \sigma(t,u)\,du = \sigma \int_t^T e^{-b(T-u)}\,du = \frac{\sigma}{b}\left(1 - e^{-b(T-t)}\right).$$

It follows that

$$\begin{aligned}
&\frac{\partial}{\partial T}C(t,T)\beta(t,R(t)) + R(t)\frac{\partial}{\partial T}C'(t,T) + \frac{\partial}{\partial T}A'(t,T)\\
&= e^{-b(T-t)}(a - bR(t)) + R(t)be^{-b(T-t)} + \left(\frac{\sigma^2}{b} - a\right)e^{-b(T-t)}\\
&\qquad -\frac{\sigma^2}{b}e^{-2b(t-t)}\\
&= \frac{\sigma^2}{b}\left(e^{-b(T-t)} - e^{-2b(T-t)}\right)\\
&= \sigma(t,T)\sigma^*(t,T),
\end{aligned}$$

as expected.

### 10.3.6 Implementation of HJM

To implement an HJM model, we need to know $\sigma(t,T)$ for $0 \le t \le T \le \overline{T}$. We can use historical data to estimate this because the same diffusion process $\sigma(t,T)$ appears in both the stochastic differential equation (10.3.6) driven by the Brownian motion $W(t)$ under the actual probability measure $\mathbb{P}$ and in the stochastic differential equation (10.3.18) driven by the Brownian motion $\widetilde{W}(t)$ under the risk-neutral measure $\widetilde{\mathbb{P}}$. Once we have $\sigma(t,T)$, we can compute $\sigma^*(t,T) = \int_t^T \sigma(t,v)\,dv$. This plus the initial forward curve $f(0,T)$, $0 \le T \le \overline{T}$, permits us to determine all the terms appearing in the formulas in Theorem 10.3.2. In particular, we use the initial forward curve to compute

$$R(t) = f(t,t) = f(0,t) + \int_0^t \sigma(u,t)\sigma^*(u,t)\,du + \int_0^t \sigma(u,t)\,d\widetilde{W}(u). \quad (10.3.28)$$

Since all expectations required for pricing interest rate derivatives are computed under $\widetilde{\mathbb{P}}$, we need only the formulas in Theorem 10.3.2; the market price of risk $\Theta(t)$ and the drift of the forward rate $\alpha(t,T)$ in (10.3.6) are irrelevant to derivative pricing. They are relevant, however, if we want to estimate nondiffusion terms from historical data (e.g., the probability of credit class migration for defaultable bonds) or we want to compute a quantity such as Value-at-Risk that requires use of the actual measure.

Assume for the moment that $\sigma(t,T)$ is of the form

$$\sigma(t,T) = \widetilde{\sigma}(T-t)\,\min\{M, f(t,T)\} \quad (10.3.29)$$

for some nonrandom function $\widetilde{\sigma}(\tau)$, $\tau \geq 0$, and some positive constant $M$. In (10.3.29), we need to have the capped forward rate $\min\{M, f(t,T)\}$ on the right-hand side rather than the forward rate $f(t,T)$ itself to prevent explosion of the forward rate. This is discussed in more detail in Subsection 10.4.1. One consequence of this fact is that forward rates (recall we are working here with *continuously compounding* forward rates; see (10.3.2)) cannot be log-normal. This is a statement about forward rates, not about the HJM model. Section 10.4 discusses how to overcome this feature of continuously compounding forward rates by building a model for simple forward rates.

We choose $\widetilde{\sigma}(T-t)$ to match historical data. The forward rate evolves according to the continuous-time model

$$df(t,T) = \alpha(t,T)\,dt + \widetilde{\sigma}(T-t)\min\{M, f(t,T)\}\,dW(t).$$

Suppose we have observed this forward rate at times $t_1 < t_2 < \cdots < t_J < 0$ in the past, and the forward rate we observed at those times was for the relative maturities $\tau_1 < \tau_2 < \cdots < \tau_K$ (i.e., we have observed $f(t_j, t_j + \tau_k)$ for $j = 1, \dots, J$ and $k = 1, \dots, K$). Suppose further that for some small positive $\delta$ we have also observed $f(t_j + \delta, t_j + \tau_k)$. We assume that $\delta$ is sufficiently small that $t_j + \delta < t_{j+1}$ for $j = 1, \dots, J-1$ and $t_J + \delta \leq 0$. According to our model,

$$\begin{aligned}&f(t_j+\delta, t_j+\tau_k) - f(t_j, t_j+\tau_k)\\ &\approx \delta\alpha(t_j, t_j+\tau_k) + \widetilde{\sigma}(\tau_k)\min\{M, f(t_j, t_j+\tau_k)\}\big(W(t_j+\delta) - W(t_j)\big).\end{aligned}$$

We identify $\widetilde{\sigma}$ by defining

$$D_{j,k} = \frac{f(t_j+\delta, t_j+\tau_k) - f(t_j, t_j+\tau_k)}{\sqrt{\delta}\,\min\{M, f(t_j, t_j+\tau_k)\}} \tag{10.3.30}$$

and observing that

$$D_{j,k} \approx \frac{\sqrt{\delta}\,\alpha(t_j, t_j+\tau_k)}{\min\{M, f(t_j, t_j+\tau_k)\}} + \widetilde{\sigma}(\tau_k)\frac{W(t_j+\delta) - W(t_j)}{\sqrt{\delta}}.$$

The first term on the right-hand side is small relative to the second term because the first term contains the factor $\sqrt{\delta}$. We define

$$X_j = \frac{W(t_j+\delta) - W(t_j)}{\sqrt{\delta}}, \quad j = 1, \dots, J, \tag{10.3.31}$$

the expression appearing in the second term, which is a standard normal random variable. We conclude that

$$D_{j,k} \approx \widetilde{\sigma}(\tau_k) X_j. \tag{10.3.32}$$

Observe that not only are $X_1, \dots, X_J$ standard normal random variables but are also independent of one another. The approximation (10.3.32) permits us to regard $D_{1k}, D_{2k}, \dots, D_{Jk}$ as independent observations taken at times $t_1, t_2, \ldots t_J$ on forward rates, all with the same relative maturity $\tau_k$. We compute the empirical covariance

$$C_{k_1,k_2} = \frac{1}{J}\sum_{j=1}^{J} D_{j,k_1} D_{j,k_2}.$$

The theoretical covariance, computed from the right-hand side of (10.3.32), is

$$\mathbb{E}\big[\widetilde{\sigma}(\tau_{k_1})\widetilde{\sigma}(\tau_{k_2})X_j^2\big] = \widetilde{\sigma}(\tau_{k_1})\widetilde{\sigma}(\tau_{k_2}).$$

Ideally, we would find $\widetilde{\sigma}(\tau_1), \widetilde{\sigma}(\tau_2), \ldots \widetilde{\sigma}(\tau_K)$ so that

$$C_{k_1 k_2} = \widetilde{\sigma}(\tau_{k_1})\widetilde{\sigma}(\tau_{k_2}), \quad k_1, k_2 = 1, 2, \ldots, K. \tag{10.3.33}$$

However, we have $K^2$ equations and only $K$ unknowns. (Actually, for different values of $k_1$ and $k_2$, the equations $C_{k_1,k_2} = \widetilde{\sigma}(\tau_{k_1})\widetilde{\sigma}(\tau_{k_2})$ and $C_{k_2,k_1} = \widetilde{\sigma}(\tau_{k_2})\widetilde{\sigma}(\tau_{k_1})$ are the same. By eliminating these duplicates, one can reduce the system to $\frac{1}{2}K(K+1)$ equations, but this is still more than the number of unknowns if $K \geq 2$.)

To determine a best choice of $\widetilde{\sigma}(\tau_1), \widetilde{\sigma}(\tau_2), \ldots \widetilde{\sigma}(\tau_K)$, we use *principal components analysis.* Set

$$D = \begin{bmatrix} D_{1,1} & D_{1,2} & \cdots & D_{1,K} \\ D_{2,1} & D_{2,2} & \cdots & D_{2,K} \\ \vdots & \vdots & & \vdots \\ D_{J,1} & D_{J,2} & \cdots & D_{J,K} \end{bmatrix}.$$

The $J$ rows of $D$ correspond to observation times, and the $K$ columns correspond to relative maturities. Then

$$C = \begin{bmatrix} C_{1,1} & C_{1,2} & \cdots & C_{1,K} \\ C_{2,1} & C_{2,2} & \cdots & C_{2,K} \\ \vdots & \vdots & & \vdots \\ C_{K,1} & C_{K,2} & \cdots & C_{K,K} \end{bmatrix} = \frac{1}{J} D^{\mathrm{tr}} D$$

is symmetric and positive semidefinite. Every symmetric, positive semidefinite matrix has a principal component decomposition

$$C = \lambda_1 e_1 e_1^{\mathrm{tr}} + \lambda_2 e_2 e_2^{\mathrm{tr}} + \cdots + \lambda_K e_K e_K^{\mathrm{tr}},$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_K \geq 0$ are the eigenvalues of $C$ and the column vectors $e_1, e_2, \ldots e_K$ are the orthogonal eigenvectors, all normalized to have length one. We want to write

$$C = \begin{bmatrix} \widetilde{\sigma}(\tau_1) \\ \widetilde{\sigma}(\tau_2) \\ \vdots \\ \widetilde{\sigma}(\tau_K) \end{bmatrix} \Big[\widetilde{\sigma}(\tau_1), \widetilde{\sigma}(\tau_2), \ldots, \widetilde{\sigma}(\tau_K)\Big].$$

However, this cannot be done exactly. The best approximation is

$$\begin{bmatrix} \widetilde{\sigma}(\tau_1) \\ \widetilde{\sigma}(\tau_2) \\ \vdots \\ \widetilde{\sigma}(\tau_K) \end{bmatrix} = \sqrt{\lambda_1}\, e_1.$$

To get a better approximation to $C$, we can introduce more Brownian motions into the equation driving the forward rates (see Exercise 10.9). Each of these has its own $\widetilde{\sigma}$ vector, and these can be chosen to be $\sqrt{\lambda_2}\, e_2$, $\sqrt{\lambda_3}\, e_3$, etc.

So far we have used only historical data. In the final step of the calibration, we introduce a nonrandom function $s(t)$ into the forward rate evolution under the risk-neutral measure, writing

$$df(t,T) = \sigma(t,T)\sigma^*(t,T)\,dt + s(t)\widetilde{\sigma}(T-t)\min\{M, f(t,T)\}\,d\widetilde{W}(t). \quad (10.3.34)$$

This is our final model. We use the values of $\widetilde{\sigma}(T-t)$ estimated from historical data under the assumption $s(t) \equiv 1$. We then allow the possibility that $s(t)$ is different from 1. We have $\sigma(t,T) = s(t)\widetilde{\sigma}(T-t)\min\{M, f(t,T)\}$. Therefore,

$$\sigma^*(t,T) = \int_t^T \sigma(t,v)\,dv = s(t)\int_t^T \widetilde{\sigma}(v-t)\min\{M, f(t,v)\}\,dv. \quad (10.3.35)$$

We substitute this function into (10.3.34) and evolve the forward rate. Even with this last-minute introduction of $s(t)$ into the model, the model is free of arbitrage when $\sigma^*(t,T)$ in (10.3.34) is defined by (10.3.35). Typically, one assumes that $s(t)$ is piecewise constant, and the values of these constants are free parameters that can be used to get the model to agree with market prices. Recalibrations of the model affect $s(t)$ only.

## 10.4 Forward LIBOR Model

In this section, we present the *forward LIBOR model*, which leads to the *Black caplet formula*. This requires us to build a model for *LIBOR* (London interbank offered rate) and use the *forward measures* introduced in Section 9.4. We begin by explaining why the continuously compounding forward rates of Section 10.3 are inadequate for the purposes of this section.

### 10.4.1 The Problem with Forward Rates

We have seen in Theorem 10.3.2 that in an arbitrage-free term-structure model, forward rates must evolve according to (10.3.18),

$$df(t,T) = \sigma(t,T)\sigma^*(t,T)\,dt + \sigma(t,T)\,d\widetilde{W}(t), \quad (10.3.18)$$

where $\widetilde{W}$ is a Brownian motion under a risk-neutral measure $\widetilde{\mathbb{P}}$. In order to adapt the Black-Scholes formula for equity options to fixed income markets, and thereby obtain the Black caplet formula (see Theorem 10.4.2 below), it would be desirable to build a model in which forward rates are log-normal under a risk-neutral measure. To do that, we should set $\sigma(t,T) = \sigma f(t,T)$ in (10.3.18), where $\sigma$ is a positive constant. However, we would then have

$$\sigma^*(T,t) = \int_t^T \sigma(t,v)\,dv = \sigma \int_t^T f(t,v)\,dv,$$

and the $dt$ term in (10.3.18) would be

$$\sigma^2 f(t,T) \int_t^T f(t,v)\,dv. \tag{10.4.1}$$

Heath, Jarrow, and Morton [83] show that this drift term causes forward rates to explode. For $T$ near $t$, the $dt$ term (10.4.1) is approximately equal to $\sigma^2(T-t)f^2(t,T)$, and the square of the forward rate creates the problem. With the drift term (10.4.1), equation (10.3.18) is similar to the deterministic ordinary differential equation

$$f'(t) = \sigma^2 f^2(t) \tag{10.4.2}$$

with a positive initial condition $f(0)$. The solution to (10.4.2) is

$$f(t) = \frac{f(0)}{1-\sigma^2 f(0)t},$$

as can easily be verified by computing

$$f'(t) = \frac{\sigma^2 f^2(0)}{\big(1-\sigma^2 f(0)t\big)^2}.$$

The function $f(t)$ explodes at time $t = \frac{1}{\sigma^2 f(0)}$. In fact, when the drift function (10.4.1) is used in (10.3.18), then (10.3.18) is worse than (10.4.2) because the randomness in (10.3.18) causes some paths to explode immediately no matter what initial condition is given. This difficulty with continuously compounding forward rates causes us to introduce *forward LIBOR*.

### 10.4.2 LIBOR and Forward LIBOR

Let $0 \le t \le T$ and $\delta > 0$ be given. We recall the discussion in Subsection 10.3.1 of how at time $t$ one can lock in an interest rate for investing over the interval $[T, T+\delta]$ by taking a short position of size 1 in a $T$-maturity zero-coupon bond and a long position of size $\frac{B(t,T)}{B(t,T+\delta)}$ in $(T+\delta)$-maturity zero-coupon bonds. This position can be created at zero cost at time $t$, it calls for "investment" of 1 at time $T$ to cover the short position, and it "repays" $\frac{B(t,T)}{B(t,T+\delta)}$ at time $T+\delta$. The continuously compounding interest rate that would explain this repayment on the investment of 1 over the time interval $[T, T+\delta]$ is given by (10.3.1). In this section, we study the *simple* interest rate that would explain this repayment, and this interest rate $L(t,T)$ is determined by the equation

$$\text{investment} \times \big(1 + \text{duration of investment} \times \text{interest rate}\big) = \text{repayment},$$

or in symbols:

$$1 + \delta L(t,T) = \frac{B(t,T)}{B(t,T+\delta)}. \tag{10.4.3}$$

We solve this equation for $L(t,T)$:

$$L(t,T) = \frac{B(t,T) - B(t,T+\delta)}{\delta B(t,T+\delta)}. \tag{10.4.4}$$

When $0 \le t < T$, we call $L(t,T)$ *forward LIBOR*. When $t = T$, we call it *spot LIBOR*, or simply *LIBOR*, set at time $T$. The positive number $\delta$ is called the *tenor* of the LIBOR, and it is usually either 0.25 years or 0.50 years.

### 10.4.3 Pricing a Backset LIBOR Contract

An *interest rate swap* is an agreement between two parties A and B that A will make fixed interest rate payments on some "notional amount" to B at regularly spaced dates and B will make variable interest rate payments on the same notional amount on these same dates. The variable rate is often *backset LIBOR*, defined on one payment date to be the LIBOR set on the previous payment date. The no-arbitrage price of a payment of backset LIBOR on a notional amount of 1 is given by the following theorem.

**Theorem 10.4.1 (Price of backset LIBOR).** *Let $0 \le t \le T$ and $\delta > 0$ be given. The no-arbitrage price at time $t$ of a contract that pays $L(T,T)$ at time $T+\delta$ is*

$$S(t) = \begin{cases} B(t,T+\delta)L(t,T), & 0 \le t \le T, \\ B(t,T+\delta)L(T,T), & T \le t \le T+\delta. \end{cases} \tag{10.4.5}$$

PROOF: There are two cases to consider. In the first case, $T \le t \le T+\delta$, LIBOR has been set at $L(T,T)$ and is known at time $t$. The value at time $t$ of a contract that pays 1 at time $T+\delta$ is $B(t,T+\delta)$, so the value at time $t$ of a contract that pays $L(T,T)$ at time $T+\delta$ is $B(t,T+\delta)L(T,T)$.

In the second case, $0 \le t \le T$, we note from (10.4.4) that

$$B(t,T+\delta)L(t,T) = \frac{1}{\delta}\big[B(t,T) - B(t,T+\delta)\big].$$

We must show that the right-hand side is the value at time $t$ of the backset LIBOR contract. To do this, suppose at time $t$ we have $\frac{1}{\delta}\big[B(t,T) - B(t,T+\delta)\big]$, and we use this capital to set up a portfolio that is:

- long $\frac{1}{\delta}$ bonds maturing at $T$;
- short $\frac{1}{\delta}$ bonds maturing at $T+\delta$.

At time $T$, we receive $\frac{1}{\delta}$ from the long position and use it to buy $\frac{1}{\delta B(T,T+\delta)}$ bonds maturing at time $T+\delta$, so that we now have a position of $\frac{1}{\delta B(T,T+\delta)} - \frac{1}{\delta}$ in $(T+\delta)$-maturity bonds. At time $T+\delta$, this portfolio pays

$$\frac{1}{\delta B(T,T+\delta)} - \frac{1}{\delta} = \frac{B(T,T) - B(T,T+\delta)}{\delta B(T,T+\delta)} = L(T,T).$$

We conclude that the capital $\frac{1}{\delta}\big[B(t,T) - B(t,T+\delta)\big]$ we used at time $t$ to set up the portfolio must be the value at time $t$ of the payment $L(T,T)$ at time $T+\delta$. □

We have proved Theorem 10.4.1 by a no-arbitrage argument. One can also obtain (10.4.5) from the risk-neutral pricing formula. For the case $t=0$, this is Exercise 10.12.

### 10.4.4 Black Caplet Formula

A common fixed income derivative security is an *interest rate cap*, a contract that pays the difference between a variable interest rate applied to a principal and a fixed interest rate (a *cap*) applied to the same principal whenever the variable interest rate exceeds the fixed rate. More specifically, let the *tenor* $\delta$, the *principal* (also called the *notional amount*) $P$, and the *cap* $K$ be fixed positive numbers. An interest rate cap pays $\big(\delta P L(\delta j, \delta j) - K\big)^+$ at time $\delta(j+1)$ for $j = 0, \dots, n$. To determine the price at time zero of the cap, it suffices to price one of the payments, a so-called *interest rate caplet*, and then sum these prices over the payments. We show here how to do this and obtain the *Black caplet formula*. We also note that each of these payments is of the form $\delta P(L(\delta j, \delta j) - K')^+$, where $K' = \frac{K}{\delta P}$. Thus, it suffices to determine the time-zero price of the payment $(L(T,T) - K)^+$ at time $T+\delta$ for an arbitrary $T$ and $K > 0$.

Consider the contract that pays $L(T,T)$ at time $T+\delta$ whose price $S(t)$ at earlier times is given by Theorem 10.4.1. Suppose we use the zero-coupon bond $B(t,T+\delta)$ as the numéraire. In terms of this numéraire, the price of the contract paying backset LIBOR is

$$\frac{S(t)}{B(t,T+\delta)} = \begin{cases} L(t,T), & 0 \le t \le T, \\ L(T,T), & T \le T \le T+\delta. \end{cases} \tag{10.4.6}$$

Recalling Definition 5.6.1 and Theorem 5.6.2, at least for $0 \le t \le T$, we see that forward LIBOR $L(t,T)$ is the $(T+\delta)$-forward price of the contract paying backset LIBOR $L(T,T)$ at time $T+\delta$.

If we build a term-structure model driven by a single Brownian motion under the actual probability measure $\mathbb{P}$ and satisfying the Heath-Jarrow-Morton no-arbitrage condition of Theorem 10.3.1, then there is a Brownian motion $\widetilde{W}(t)$ under a risk-neutral probability measure $\widetilde{\mathbb{P}}$ such that forward rates are given by (10.3.18) and bond prices by (10.3.19). Theorem 9.2.2 implies that the risk-neutral measure corresponding to numéraire $B(t, T+\delta)$ is given by

$$\widetilde{\mathbb{P}}^{T+\delta}(A) = \frac{1}{B(0, T+\delta)} \int_A D(T+\delta)\, d\widetilde{\mathbb{P}} \text{ for all } A \in \mathcal{F} \tag{10.4.7}$$

and

$$\widetilde{W}^{T+\delta}(t) = \int_0^t \sigma^*(u, T+\delta)\, du + \widetilde{W}(t) \tag{10.4.8}$$

is a Brownian motion under $\widetilde{\mathbb{P}}^{T+\delta}$. We call $\widetilde{\mathbb{P}}^{T+\delta}$ the $(T+\delta)$-forward measure.

Theorem 9.2.2 implies that $\frac{S(t)}{B(t,T+\delta)}$ is a martingale under $\widetilde{\mathbb{P}}^{T+\delta}$. (See the discussion in Subsections 9.4.1 and 9.4.2.) According to the Martingale Representation Theorem (see Corollary 5.3.2), there must exist some process $\gamma(t, T)$, a process in $t \in [0, T]$ for each fixed $T$, such that

$$dL(t, T) = \gamma(t, T) L(t, T)\, d\widetilde{W}^{T+\delta}(t), \; 0 \le t \le T. \tag{10.4.9}$$

We relate this process to the zero-coupon bond volatilities in Subsection 10.4.5. The point of (10.4.9) is that there is no $dt$ term, which was the term causing the problem with forward rates in Subsection 10.4.1. The $dt$ term has been removed by changing to the $(T+\delta)$-forward measure, under which $L(t, T)$ is a martingale.

The forward LIBOR model is constructed so that $\gamma(t, T)$, defined for $0 \le t \le T \le \overline{T}$, is nonrandom. When $\gamma(t, T)$ is nonrandom, forward LIBOR $L(t, T)$ will be log-normal under the forward measure $\widetilde{\mathbb{P}}^{T+\delta}$. This leads to the following pricing result.

**Theorem 10.4.2 (Black caplet formula).** *Consider a* caplet *that pays $\big(L(T, T) - K\big)^+$ at time $T+\delta$, where $K$ is some nonnegative constant. Assume forward LIBOR is given by (10.4.9) and $\gamma(t, T)$ is nonrandom. Then the price of the caplet at time zero is*

$$B(0, T+\delta)\big[L(0, T)N(d_+) - KN(d_-)\big], \tag{10.4.10}$$

*where*

$$d_\pm = \frac{1}{\sqrt{\int_0^T \gamma^2(t, T)\, dt}} \left[\log \frac{L(0, T)}{K} \pm \frac{1}{2} \int_0^T \gamma^2(t, T)\, dt\right]. \tag{10.4.11}$$

PROOF: According to the risk-neutral pricing formula, the price of the caplet at time zero is the discounted risk-neutral (under $\widetilde{\mathbb{P}}$) expected value of the payoff, which is

$$\begin{aligned}
&\widetilde{\mathbb{E}}\left[D(T+\delta)\left(L(T,T)-K\right)^+\right]\\
&= B(0,T+\delta)\widetilde{\mathbb{E}}\left[\frac{D(T+\delta)}{B(0,T+\delta)}\left(L(T,T)-K\right)^+\right]\\
&= B(0,T+\delta)\widetilde{\mathbb{E}}^{T+\delta}\left(L(T,T)-K\right)^+. \qquad (10.4.12)
\end{aligned}$$

The solution to the stochastic differential equation (10.4.9) is

$$L(T,T) = L(0,T)\exp\left\{\int_0^T \gamma(t,T)\,d\widetilde{W}^{T+\delta}(t) - \frac{1}{2}\int_0^t \gamma^2(t,T)\,dt\right\}.$$

Let us define $\overline{\gamma}(T) = \sqrt{\frac{1}{T}\int_0^T \gamma^2(t,T)\,dt}$. According to Example 4.7.3, the Itô integral $\int_0^T \gamma(t,T)\,d\widetilde{W}^{T+\delta}(t)$ is a normal random variable under $\widetilde{\mathbb{P}}^{T+\delta}$ with mean zero and variance $\overline{\gamma}^2(T)T$; we may thus write it as $-\overline{\gamma}(T)\sqrt{T}\,X$, where $X = -\frac{1}{\overline{\gamma}(T)\sqrt{T}}\int_0^T \gamma(t,T)\,d\widetilde{W}^{T+\delta}(t)$ is a standard normal random variable under $\widetilde{\mathbb{P}}^{T+\delta}$. In this notation,

$$L(T,T) = L(0,T)e^{-\overline{\gamma}(T)\sqrt{T}\,X - \frac{1}{2}\overline{\gamma}^2(T)T},$$

and

$$\widetilde{\mathbb{E}}^{T+\delta}\left(L(T,T)-K\right)^+ = \widetilde{\mathbb{E}}^{T+\delta}\left[\left(L(0,T)e^{-\overline{\gamma}(T)\sqrt{T}\,X - \frac{1}{2}\overline{\gamma}^2(T)T} - K\right)^+\right].$$

This is the same computation as in (5.2.35), which led to (5.2.36). Therefore,

$$\begin{aligned}
\widetilde{\mathbb{E}}^{T+\delta}(L(T,T)-K)^+ &= BS(T, L(0,T); K, 0, \overline{\gamma}(T))\\
&= L(0,T)N(d_+) - KN(d_-),
\end{aligned}$$

and the risk-neutral price of the caplet (10.4.12) is (10.4.10). □

### 10.4.5 Forward LIBOR and Zero-Coupon Bond Volatilities

Recall that forward LIBOR is determined by the equation (10.4.3), which we can rewrite as

$$L(t,T) + \frac{1}{\delta} = \frac{B(t,T)}{\delta B(t,T+\delta)}.$$

We work out the evolution of $L(t,T)$ under the forward measure $\widetilde{\mathbb{P}}^{T+\delta}$. According to Theorem 10.3.2,

$$\begin{aligned}
&D(t)B(t,T)\\
&= B(0,T)\exp\left\{-\int_0^t \sigma^*(u,T)\,d\widetilde{W}(u) - \frac{1}{2}\int_0^t (\sigma^*(u,T))^2\,du\right\},\\
&D(t)B(t,T+\delta)\\
&= B(0,T+\delta)\exp\left\{-\int_0^t \sigma^*(u,T+\delta)\,d\widetilde{W}(u) - \frac{1}{2}\int_0^t (\sigma^*(u,T+\delta))^2\,du\right\}.
\end{aligned}$$

This implies

$$\begin{aligned}
L(t,T) &+ \frac{1}{\delta} \\
&= \frac{B(t,T)}{\delta B(t,T+\delta)} \\
&= \frac{B(0,T)}{\delta B(0,T+\delta)} \exp\left\{ \int_0^t \left[\sigma^*(u,T+\delta) - \sigma^*(u,T)\right] d\widetilde{W}(u) \right. \\
&\qquad \left. + \frac{1}{2}\int_0^t \left[(\sigma^*(u,T+\delta))^2 - (\sigma^*(u,T))^2\right] du \right\}.
\end{aligned}$$

The Itô-Doeblin formula implies

$$\begin{aligned}
dL&(t,T) \\
&= \left(L(t,T) + \frac{1}{\delta}\right)\Bigg\{\left[\sigma^*(t,T+\delta) - \sigma^*(t,T)\right] d\widetilde{W}(t) \\
&\qquad + \frac{1}{2}\left[(\sigma^*(t,T+\delta))^2 - (\sigma^*(t,T))^2\right] dt \\
&\qquad + \frac{1}{2}\left[\sigma^*(t,T+\delta) - \sigma^*(t,T)\right]^2 d\widetilde{W}(t)\, d\widetilde{W}(t)\Bigg\} \\
&= \left(L(t,T) + \frac{1}{\delta}\right)\Bigg\{\left[\sigma^*(t,T+\delta) - \sigma^*(t,T)\right] d\widetilde{W}(t) \\
&\qquad + \frac{1}{2}\Big[(\sigma^*(t,T+\delta))^2 - (\sigma^*(t,T))^2 + (\sigma^*(t,T+\delta))^2 \\
&\qquad\qquad - 2\sigma^*(t,T+\delta)\sigma^*(t,T) + (\sigma^*(t,T))^2\Big] dt\Bigg\} \\
&= \left(L(t,T) + \frac{1}{\delta}\right)\Big\{\left[\sigma^*(t,T+\delta) - \sigma^*(t,T)\right] d\widetilde{W}(t) \\
&\qquad + \left[(\sigma^*(t,T+\delta))^2 - \sigma^*(t,T+\delta)\sigma^*(t,T)\right] dt\Big\} \\
&= \left(L(t,T) + \frac{1}{\delta}\right)\left[\sigma^*(t,T+\delta) - \sigma^*(t,T)\right]\left[\sigma^*(t,T+\delta)dt + d\widetilde{W}(t)\right].
\end{aligned}$$

From (10.4.8), we have

$$d\widetilde{W}^{T+\delta}(t) = \sigma^*(t,T+\delta)\,dt + d\widetilde{W}(t). \tag{10.4.13}$$

Therefore,

$$dL(t,T) = \frac{1}{\delta}\big(1 + \delta L(t,T)\big)\left[\sigma^*(t,T+\delta) - \sigma^*(t,T)\right] d\widetilde{W}^{T+\delta}(t). \tag{10.4.14}$$

Comparing this with (10.4.9), we conclude that the forward LIBOR volatility $\gamma(t,T)$ of (10.4.9) and thc $(T+\delta)$- and $T$-maturity zero-coupon bond volatilities $\sigma^*(t,T+\delta)$ and $\sigma^*(t,T)$ are related by the formula

$$\gamma(t,T) = \frac{1+\delta L(t,T)}{\delta L(t,T)}\big[\sigma^*(t,T+\delta) - \sigma^*(t,T)\big]. \tag{10.4.15}$$

### 10.4.6 A Forward LIBOR Term-Structure Model

The Black caplet formula of Theorem 10.4.2 is used to calibrate the forward LIBOR model. However, this calibration does not determine all the parameters needed to have a full term-structure model. In this section, we discuss the calibration and display some of the choices left open by it. We begin by collecting the equations appearing earlier in this section that we need for this subsection:

$$1+\delta L(t,T) = \frac{B(t,T)}{B(t,T+\delta)}, \quad 0 \le t \le T \le \overline{T} - \delta, \tag{10.4.3}$$

$$\widetilde{\mathbb{P}}^{T+\delta}(A) = \frac{1}{B(0,T+\delta)} \int_A D(T+\delta)\, d\widetilde{\mathbb{P}} \text{ for all } A \in \mathcal{F},\ 0 \le T \le \overline{T} - \delta, \tag{10.4.7}$$

$$d\widetilde{W}^{T+\delta}(t) = \sigma^*(t,T+\delta)\,dt + d\widetilde{W}(t), \quad 0 \le t \le T \le \overline{T} - \delta, \tag{10.4.8}$$

$$dL(t,T) = \gamma(t,T)L(t,T)\,d\widetilde{W}^{T+\delta}(t), \quad 0 \le t \le T \le \overline{T} - \delta, \tag{10.4.9}$$

$$\gamma(t,T) = \frac{1+\delta L(t,T)}{\delta L(t,T)}\big[\sigma^*(t,T+\delta) - \sigma^*(t,T)\big], \quad 0 \le t \le T \le \overline{T} - \delta. \tag{10.4.15}$$

Suppose now, at time zero, that market data allow us to determine caplet prices for maturity dates $T_j = j\delta$ for $j = 1, \dots, n$. We can then imply the volatilities $\overline{\gamma}(T_j)$, $j = 1, \dots, n$, appearing in the proof of Theorem 10.4.2. We wish to build a term structure model consistent with these data. We begin by setting $\overline{T}$ in the equations above equal to $(n+1)\delta$.

- We choose nonrandom nonnegative functions

$$\gamma(t,T_j), \quad 0 \le t \le T_j,\ j = 1, \dots, n,$$

  so that $\sqrt{\frac{1}{T_j}\int_0^{T_j} \gamma^2(t,T_j)\,dt} = \overline{\gamma}(T_j)$.

For example, we could take $\gamma(t,T_j) = \overline{\gamma}(T_j)$ for $0 \le t \le T_j$.

With these volatility functions $\gamma(t,T_j)$, we can evolve forward LIBORs by equation (10.4.9), at least for $T = T_j$, $j = 1, \dots, n$, and the forward LIBORs we obtain will agree with the market cap prices. However, (10.4.9) with $T = T_j$ gives us a formula for forward LIBOR $L(t,T_j)$ in terms of the forward Brownian motion $\widetilde{W}^{T_{j+1}}(t)$, and these are different for different values of $j$. Before we use (10.4.9) to evolve forward LIBORs, we must determine the relationship among these different equations.

### Construction of Forward LIBOR Processes

Observe from (10.4.8) that

$$d\widetilde{W}^{T_j}(t) = \sigma^*(t, T_j)\,dt + d\widetilde{W}(t), \quad 0 \le t \le T_j.$$

Similarly,

$$d\widetilde{W}^{T_{j+1}}(t) = \sigma^*(t, T_{j+1})\,dt + d\widetilde{W}(t), \quad 0 \le t \le T_{j+1}.$$

Subtracting these equations, we obtain

$$\begin{aligned} d\widetilde{W}^{T_j}(t) &= \left[\sigma^*(t, T_j) - \sigma^*(t, T_{j+1})\right] dt + d\widetilde{W}^{T_{j+1}}(t) \\ &= -\frac{\delta\gamma(t, T_j)L(t, T_j)}{1 + \delta L(t, T_j)}\,dt + d\widetilde{W}^{T_{j+1}}(t), \quad 0 \le t \le T_j, \end{aligned} \tag{10.4.16}$$

where we have used (10.4.15) for the second equality. Setting $j = n$ in (10.4.16), we have

$$d\widetilde{W}^{T_n}(t) = -\frac{\delta\gamma(t, T_n)L(t, T_n)}{1 + \delta L(t, T_n)}\,dt + d\widetilde{W}^{T_{n+1}}(t), \quad 0 \le t \le T_n. \tag{10.4.17}$$

Setting $j = n - 1$ in (10.4.16) and using (10.4.17), we obtain

$$\begin{aligned} d\widetilde{W}^{T_{n-1}}(t) &= -\frac{\delta\gamma(t, T_{n-1})L(t, T_{n-1})}{1 + \delta L(t, T_{n-1})}\,dt + d\widetilde{W}^{T_n}(t) \\ &= -\frac{\delta\gamma(t, T_{n-1})L(t, T_{n-1})}{1 + \delta L(t, T_{n-1})}\,dt - \frac{\delta\gamma(t, T_n)L(t, T_n)}{1 + \delta L(t, T_n)}\,dt + d\widetilde{W}^{T_{n+1}}(t), \\ &\qquad 0 \le t \le T_{n-1}. \end{aligned}$$

Repeating this process, we conclude that

$$d\widetilde{W}^{T_{j+1}}(t) = -\sum_{i=j+1}^{n} \frac{\delta\gamma(t, T_i)L(t, T_i)}{1 + \delta L(t, T_i)}\,dt + d\widetilde{W}^{T_{n+1}}(t), \quad 0 \le t \le T_{j+1}. \tag{10.4.18}$$

Equation (10.4.18) holds for $j = 0, \ldots, n$, provided we interpret $\sum_{i=n+1}^{n}$ to be zero.

We return to (10.4.9), using (10.4.18) to write

$$dL(t, T_j) = \gamma(t, T_j)L(t, T_j)\left[-\sum_{i=j+1}^{n} \frac{\delta\gamma(t, T_i)L(t, T_i)}{1 + \delta L(t, T_i)}\,dt + d\widetilde{W}^{T_{n+1}}(t)\right],$$
$$0 \le t \le T_j, \; j = 1, \ldots, n. \tag{10.4.19}$$

Now we have a single Brownian motion driving all $n$ equations. Thus, to construct the forward LIBOR model, we choose a Brownian motion, which we call $\widetilde{W}^{T_{n+1}}(t)$, $0 \leq t \leq T_{n+1}$, under a probability measure we call $\widetilde{\mathbb{P}}^{T_{n+1}}$. That is, we start with a probability space $(\Omega, \mathcal{F}, \widetilde{\mathbb{P}}^{T_{n+1}})$ on which is defined a Brownian motion $\widetilde{W}^{T_{n+1}}(t)$, $0 \leq t \leq T_{n+1}$. We assume the initial forward LIBORs $L(0, T_j)$, $j = 1, \ldots, n+1$, are known from market data. With these initial conditions, (10.4.19) generates the forward LIBOR processes $L(t, T_j)$, $0 \leq t \leq T_j$, generating first $L(t, T_n)$, which has no drift in (10.4.19), then using $L(t, T_n)$ in the differential equation for $L(t, T_{n-1})$ to generate that process, then using $L(t, T_n)$ and $L(t, T_{n-1})$ in the differential equation for $L(t, T_{n-2})$ to generate that process, and so on. Implicit in this computation is a dependence among these different forward LIBOR processes.

### Construction of $T_j$-Maturity Discounted Bond Prices

We construct the volatility $\sigma^*(t, T_j)$ for the zero-coupon bond maturing at $T_j$, $j = 1, \ldots, n+1$. The forward LIBOR model has a tenor $\delta > 0$, and while it puts constraints on the cumulative effect of processes between set points $T_j$, it does not provide fine detail about what happens between set points. In particular, we are free to choose the bond volatilities $\sigma^*(t, T_j)$ for $T_{j-1} \leq t < T_j$. The only constraint is that

$$\lim_{t \uparrow T_j} \sigma^*(t, T_j) = \sigma^*(T_j, T_j) = 0. \tag{10.4.20}$$

This constraint is present because the bond price $B(t, T_j)$ converges to 1 as $t \uparrow T_j$, and so the volatility must vanish. This is also apparent in the second formula in (10.3.9).

- For each $j = 1, \ldots, n+1$, choose $\sigma^*(t, T_j)$ for $T_{j-1} \leq t < T_j$ so that (10.4.20) is satisfied.

We show that this determines $\sigma(t, T_j)$ for all values of $t \in [0, T_j)$. (Again, we know from the outset that $\sigma(T_j, T_j) = 0$; that does not need to be chosen or determined.)

First of all, the initial choice of $\sigma^*(t, T_1)$ determines this function for all relevant values of $t$, namely, for $0 \leq t < T_1$. From (10.4.15), we have

$$\sigma^*(t, T_2) = \sigma^*(t, T_1) + \frac{\delta\gamma(t, T_1)L(t, T_1)}{1 + \delta L(t, T_1)},$$

and since $\sigma^*(t, T_1)$ has been chosen for $0 \leq t < T_1$, the function $\sigma(t, T_2)$ is determined by this equation for $0 \leq t < T_1$. For $T_1 \leq t < T_2$, $\sigma(t, T_2)$ has already been chosen. Therefore, $\sigma^*(t, T_2)$ is determined for $0 \leq t < T_2$. From (10.4.15), we also have

$$\sigma^*(t, T_3) = \sigma^*(t, T_2) + \frac{\delta\gamma(t, T_2)L(t, T_2)}{1 + \delta L(t, T_2)},$$

and since $\sigma^*(t, T_2)$ has been determined for $0 \leq t < T_2$, the function $\sigma(t, T_3)$ is determined by this equation for $0 \leq t < T_2$. For $T_2 \leq t < T_3$, $\sigma(t, T_3)$

has already been chosen. Therefore, $\sigma^*(t,T_3)$ is determined for $0 \le t < T_3$. Continuing in this way, we determine $\sigma(t,T_j)$ for all $j = 1,\dots,n+1$ and $0 \le t < T_j$.

Using the bond volatilities $\sigma^*(t,T)$ and (10.4.8), we may write the zero-coupon bond price formula (10.3.19) of Theorem 10.3.2 as

$$\begin{aligned}
dB(t,T_j) &= R(t)B(t,T_j)\,dt - \sigma^*(t,T_j)B(t,T_j)\,d\widetilde{W}(t)\\
&= R(t)B(t,T_j)\,dt + \sigma^*(t,T_j)\sigma^*(t,T_{n+1})B(t,T_j)\,dt\\
&\qquad -\sigma^*(t,T_j)B(t,T_j)\,d\widetilde{W}^{T_{n+1}}(t).
\end{aligned}$$

However, we have not yet determined an interest rate process $R(t)$, and so we prefer to write this equation in discounted form. For $j = 1,\dots,n+1$,

$$\begin{aligned}
d\big(D(t)B(t,T_j)\big) &= \sigma^*(t,T_j)\sigma^*(t,T_{n+1})D(t)B(t,T_j)\,dt\\
&\qquad -\sigma^*(t,T_j)D(t)B(t,T_j)\,d\widetilde{W}^{T_{n+1}}(t),\ \ 0 \le t \le T_j. \quad (10.4.21)
\end{aligned}$$

The initial condition can be obtained from (10.4.3):

$$D(0)B(0,T_j) = B(0,T_j) = \prod_{i=0}^{j-1}\frac{B(0,T_{i+1})}{B(0,T_i)} = \prod_{i=0}^{j-1}\big(1+\delta L(0,T_i)\big)^{-1}. \quad (10.4.22)$$

This permits us to generate the discounted bond prices $D(t)B(t,T_j)$, $j = 1,\dots,n+1$. Indeed, the solution to (10.4.21) is

$$\begin{aligned}
D(t)B(t,T_j) = B(0,T_j)\exp\Bigg\{&-\int_0^t \sigma^*(u,T_j)\,d\widetilde{W}^{T_{n+1}}(u)\\
&-\int_0^t\left[\frac{1}{2}\big(\sigma^*(u,T_j)\big)^2 - \sigma^*(u,T_j)\sigma^*(u,T_{n+1})\right]du\Bigg\}. \quad (10.4.23)
\end{aligned}$$

*Remark 10.4.3.* Equation (10.4.23) does not determine the discount process $D(t)$ and the bond price $B(t,T_j)$ separately, except when $t = T_j$ for some $j$. In the case when $t = T_j$, we have $B(T_j,T_j) = 1$, so

$$\begin{aligned}
D(T_j) &= D(T_j)B(T_j,T_j)\\
&= B(0,T_j)\exp\Bigg\{-\int_0^{T_j}\sigma^*(u,T_j)\,d\widetilde{W}^{T_{n+1}}(u)\\
&\qquad -\int_0^{T_j}\left[\frac{1}{2}\big(\sigma^*(u,T_j)\big)^2 - \sigma^*(u,T_j)\sigma^*(u,T_{n+1})\right]du\Bigg\}.
\end{aligned} \quad (10.4.24)$$

In the special case when $j = n+1$, we obtain

$$D(T_{n+1}) = B(0, T_{n+1}) \exp\left\{ -\int_0^{T_{n+1}} \sigma^*(u, T_{n+1})\, d\widetilde{W}^{T_{n+1}}(u) + \frac{1}{2} \int_0^{T_{n+1}} \left(\sigma^*(u, T_{n+1})\right)^2 du \right\}. \quad (10.4.25)$$

### Risk-Neutral Measure

The risk-neutral measure $\widetilde{\mathbb{P}}$ is related to the forward measure $\widetilde{\mathbb{P}}^{T_{n+1}}$ by (10.4.7),

$$\widetilde{\mathbb{P}}^{T_{n+1}}(A) = \int_A \frac{D(T_{n+1})}{B(0, T_{n+1})}\, d\widetilde{\mathbb{P}} \text{ for all } A \in \mathcal{F},$$

or, equivalently,

$$\widetilde{\mathbb{P}}(A) = \int_A \frac{B(0, T_{n+1})}{D(T_{n+1})}\, d\widetilde{\mathbb{P}}^{T_{n+1}} \text{ for all } A \in \mathcal{F}. \quad (10.4.26)$$

Because we have begun with the measure $\widetilde{\mathbb{P}}^{T_{n+1}}$ rather than $\widetilde{\mathbb{P}}$, we use (10.4.26) to define $\widetilde{\mathbb{P}}$. According to (10.4.25),

$$\frac{B(0, T_{n+1})}{D(T_{n+1})} = \exp\left\{ \int_0^{T_{n+1}} \sigma^*(u, T_{n+1})\, d\widetilde{W}^{T_{n+1}}(u) - \frac{1}{2} \int_0^{T_{n+1}} \sigma^*(u, T_{n+1})\, du \right\}, \quad (10.4.27)$$

and so the terms appearing on the right-hand side of (10.4.26) are defined. The following theorem justifies calling $\widetilde{\mathbb{P}}$ the risk-neutral measure.

**Theorem 10.4.4.** *Under $\widetilde{\mathbb{P}}$ given by (10.4.26), the discounted zero-coupon bond prices given by (10.4.21) and (10.4.22), or equivalently by (10.4.23), are martingales.*

PROOF: With

$$\widetilde{W}(t) = \widetilde{W}^{T_{n+1}}(t) - \int_0^t \sigma^*(u, T_{n+1})\, du, \quad 0 \le t \le T_{n+1},$$

(10.4.21) may be written as

$$d\big(D(t)B(t, T_j)\big) = -\sigma^*(t, T_j) D(t) B(t, T_j)\, d\widetilde{W}(t). \quad (10.4.28)$$

It suffices to show that $\widetilde{W}(t)$ is a Brownian motion under $\widetilde{\mathbb{P}}$ defined by (10.4.26). According to Girsanov's Theorem, Theorem 5.2.3, with $\Theta(u) = -\sigma^*(u, T_{n+1})$, since $\widetilde{W}^{T_{n+1}}(t)$ is a Brownian motion under $\widetilde{\mathbb{P}}^{T_{n+1}}$, then $\widetilde{W}(t)$ is a Brownian motion under a measure $\widehat{\mathbb{P}}$ defined by

$$\widehat{\mathbb{P}}(A) = \int_A Z(T_{n+1})\, d\widetilde{\mathbb{P}}^{T_{n+1}} \text{ for all } A \in \mathcal{F},$$

where

$$Z(T_{n+1}) = \exp\left\{ -\int_0^{T_{n+1}} \Theta(u)\, d\widetilde{W}^{T_{n+1}}(u) - \frac{1}{2}\int_0^{T_{n+1}} \Theta^2(u)\, du \right\}.$$

From (10.4.27), we see that $Z(T_{n+1}) = \frac{B(0,T_{n+1})}{D(T_{n+1})}$, so $\widehat{\mathbb{P}} = \widetilde{\mathbb{P}}$. □

*Remark 10.4.5.* In order to complete the determination of a full term-structure model with bond prices for all maturities $T$, a discount process, and forward rates, it is necessary to choose $\gamma(t,T)$ for $0 \le t \le T$ and $T \in (0, T_{n+1}) \setminus \{T_1, \dots, T_n\}$ and to also make some choices in order to determine bond volatility $\sigma^*(t,T)$ for $0 \le t \le T$ and $T \in (0, T_{n+1}) \setminus \{T_1, \dots, T_n\}$. This can be done, and thus the forward LIBOR model is consistent with a full term-structure model. However, the model obtained by exercising these choices arbitrarily is not a reliable vehicle for pricing instruments that depend on these choices.

## 10.5 Summary

We have presented three types of term-structure models: finite-factor Markov models for the short rate, the Heath-Jarrow-Morton model, and the forward LIBOR model.

There are many finite-factor short-rate models. For all of them, one writes down a stochastic differential equation or system of stochastic differential equations for the "factors", and then provides a formula for the interest rate as a function of these factors. One then uses the risk-neutral pricing formula to obtain prices of bonds and fixed income derivatives. In particular, these models begin under the risk-neutral measure, for otherwise there is no way to infer prices of assets from the factor processes and the interest rate.

Affine-yield models belong to the class of finite-factor short-rate models, and we have presented the two-factor affine-yield models. In these models, the interest rate is given by an equation of the form

$$R(t) = \delta_0 + \delta_1 Y_1(t) + \delta_2 Y_2(t), \tag{10.2.6}$$

where $\delta_0$, $\delta_1$, and $\delta_2$ are either constants (as in the text) or nonrandom functions of time (as in Exercise 10.3), and $Y_1(t)$ and $Y_2(t)$ are the factor processes. When regarded as a two-dimensional process, $\big(Y_1(t), Y_2(t)\big)$ is Markov, and hence bond prices and the prices of interest rate derivatives are functions of these processes. These functions can be determined by solving partial differential equations with boundary conditions depending on the particular instrument being priced. For the boundary condition associated with zero-coupon bonds, the partial differential equations reduce to a system of ordinary differential equations, which permits rapid calibration of the models.

For the two-factor affine-yield models, the price at time $t$ of a zero-coupon bond maturing at a later time $T$ and paying 1 upon maturity is of the form

$$B(t,T) = e^{-Y_1(t)C_1(t,T)-Y_2(t)C_2(t,T)-A(t,T)}. \tag{10.5.1}$$

The nonrandom functions $C_1(t,T)$, $C_2(t,T)$, and $A(t,T)$ are given by a system of ordinary differential equations in the $t$ variable and the boundary condition

$$C_1(T,T) = C_2(T,T) = A(T,T) = 0.$$

When the model coefficients, both $\delta_0$, $\delta_1$, and $\delta_2$ in (10.2.6) and the coefficients in the differential equations satisfied by the factor processes, are constant, the functions $C_1(t,T)$, $C_2(t,T)$, and $A(t,T)$ depend on $t$ and $T$ only through their difference $\tau = T - t$.

The affine-yield models are calibrated by choosing the coefficients in (10.2.6) and/or in the stochastic differential equations for the factor processes. To introduce more variables for the calibration, it is customary to take the coefficients to be nonrandom, often piecewise constant, functions of time. It is helpful before beginning the calibration to make sure that the models are written in their most parsimonious form so that one cannot obtain the same model statistics from two different sets of parameter choices. The canonical forms presented here are "most parsimonious" in this sense.

There are three canonical two-factor affine-yield models, which we call the *two-factor Vasicek model*, the *two-factor Cox-Ingersoll-Ross model*, and the *two-factor mixed model*. In the first of these, both factors can become negative. In the second, both factors are guaranteed to be nonnegative. In the third, one factor is guaranteed to be nonnegative and the other can become negative. All three of these models are driven by independent Brownian motions $\widetilde{W}_1(t)$, $\widetilde{W}_2(t)$ under a risk-neutral measure $\widetilde{\mathbb{P}}$.

The canonical two-factor Vasicek model is

$$dY_1(t) = -\lambda_1 Y_1(t)\,dt + d\widetilde{W}_1(t), \tag{10.2.4}$$

$$dY_2(t) = -\lambda_{21} Y_1(t)\,dt - \lambda_2 Y_2(t)\,dt + d\widetilde{W}_1(t), \tag{10.2.5}$$

where $\lambda_1 > 0$ and $\lambda_2 > 0$. These factors are Gaussian processes, and their statistics and the statistics of the resulting interest rate $R(t)$ can be determined (Exercise 10.2). The functions $C_1(T-t)$, $C_2(T-t)$, and $A(T-t)$ in (10.5.1) are determined by the system of ordinary differential equations (10.2.23)–(10.2.25), and the solution to this system is given by (10.2.26)–(10.2.29). The canonical two-factor Cox-Ingersoll-Ross model is

$$dY_1(t) = \big(\mu_1 - \lambda_{11}Y_1(t) - \lambda_{12}Y_2(t)\big)\,dt + \sqrt{Y_1(t)}\,d\widetilde{W}_1(t), \tag{10.2.49}$$

$$dY_2(t) = \big(\mu_2 - \lambda_{21}Y_1(t) - \lambda_{22}Y_2(t)\big)\,dt + \sqrt{Y_1(t)}\,d\widetilde{W}_2(t), \tag{10.2.50}$$

where $\mu_1 \geq 0$, $\mu_2 \geq 0$, $\lambda_{11} > 0$, $\lambda_{22} > 0$, $\lambda_{12} \leq 0$, and $\lambda_{21} \leq 0$. The system of ordinary differential equations (10.2.56)–(10.2.58) determines the functions

$C_1(T-t)$, $C_2(T-t)$, and $A(T-t)$ in (10.5.1). The canonical two-factor mixed model is

$$dY_1(t) = \big(\mu - \lambda_1 Y_1(t)\big)\,dt + \sqrt{Y_1(t)}\,d\widetilde{W}_1(t), \tag{10.2.59}$$
$$dY_2(t) = -\lambda_2 Y_2(t)\,dt + \sigma_{21}\sqrt{Y_1(t)}\,d\widetilde{W}_1(t) + \sqrt{\alpha + \beta Y_1(t)}\,d\widetilde{W}_2(t), \tag{10.2.60}$$

where $\mu \geq 0$, $\lambda_1 > 0$, $\lambda_2 > 0$, $\alpha \geq 0$, and $\beta \geq 0$. The functions $C_1(T-t)$, $C_2(T-t)$, and $A(T-t)$ in (10.5.1) are determined by the system of differential equations (10.7.4)–(10.7.6). When the model coefficients depend on time, the differential equations in all three cases are modified by replacing the constant coefficients by time-varying coefficients and replacing $C_i'$ in these equations (which is the derivative of $C_i$ with respect to $\tau = T - t$) by $-\frac{\partial}{\partial t}C_i(t,T)$ and making the similar replacement for $A'$.

The Heath-Jarrow-Morton (HJM) model evolves the whole yield curve forward in time rather than a finite set of factors. The yield curve is an infinite-dimensional object. Note, however, that the HJM model is driven by finitely many Brownian motions (in fact, by one Brownian motion in Section 10.3 but by multiple Brownian motions in Exercise 10.9). As a result, the HJM model is "finite-dimensional" in the sense that not every possible yield curve can be obtained from the model.

The yield curve in the HJM model is characterized by *forward rates*. The forward rate $f(t,T)$ is the instantaneous interest rate that can be locked in at time $t$ for borrowing at a later time $T$. The HJM model begins under the actual probability measure $\mathbb{P}$ and derives a condition on the drift $\alpha(t,T)$ and diffusion $\sigma(t,T)$ of $f(t,T)$ that guarantees the existence of a risk-neutral measure $\widetilde{\mathbb{P}}$ and hence guarantees the absence of arbitrage. This condition is that there must exist a *market price of risk process* $\Theta(t)$ that does not depend on $T$ and that satisfies

$$\alpha(t,T) = \sigma(t,T)\big[\sigma^*(t,T) + \Theta(t)\big], \quad 0 \leq t \leq T; \tag{10.3.16}$$

see Theorem 10.3.1. Although this condition was developed within the HJM model, one would not encounter in practice an arbitrage-free term-structure model driven by a single Brownian motion and not satisfying this condition. For term-structure models driven by multiple Brownian motions, the analogous condition appears in Exercise 10.9(i).

In terms of the Brownian motion $\widetilde{W}(t)$ under the risk-neutral measure, bond prices in the HJM model satisfy

$$dB(t,T) = R(t)B(t,T)\,dt - \sigma^*(t,T)B(t,T)\,d\widetilde{W}(t),$$

where $\sigma^*(t,T) = \int_t^T \sigma(t,v)\,dv$; see Theorem 10.3.2. A calibration procedure for the HJM model is provided in Subsection 10.3.6.

In contrast to the continuously compounding forward rate $f(t,T)$, which is the basis of the HJM model, forward LIBOR $L(t,T)$ is the simple interest rate that can be locked in at time $t$ for borrowing at a later time $T$ over the interval $[T, T+\delta]$. Here $\delta$ is a positive constant, and although not indicated by the notation, $L(t,T)$ depends on the choice of this constant.

Section 10.4 introduces a model for forward LIBOR. One can build this model so that forward LIBOR $L(t,T)$ is log-normal under the forward measure $\widetilde{\mathbb{P}}^{T+\delta}$, and this permits a mathematically rigorous derivation of the *Black caplet formula*. This formula is similar to the Black-Scholes-Merton formula for equities but used in fixed income markets in which the essence of the market is that the interest rate is random, in contrast to the Black-Scholes-Merton assumption.

## 10.6 Notes

The Vasicek model appears in [154] and the Cox-Ingersoll-Ross model in [41]. Hull and White generalized the Vasicek model in [88]. The general concept of multifactor affine-yield models is developed in Duffie and Kan [57], [58]. The reduction of affine-yield models to canonical versions is due to Dai and Singleton [44]. A sampling of other articles related to affine-yield models includes Ait-Sahalia [1], Balduzzi, Das, Foresi, and Sundaram [7], Chen [29], Chen and Scott [30], [31], [32], Collin-Dufresne and Goldstein [38], [39], Duffee [55], and Piazzesi [132]. Maghsoodi [116] provides a detailed study of the one-dimensional CIR equation when the parameters are time-varying.

Although affine-yield models have simple bond price formulas, the prices for fixed income derivatives are more complicated. However, numerical solution of partial differential equations can be avoided by Fourier transform analysis; see, Duffie, Pan, and Singleton [59].

Some other common short rate models are those of Black, Derman, and Toy [15], Black and Karasinski [16], and Longstaff and Schwartz [111]. An empirical comparison of various short rate models is provided by Chan et al. [28].

Ho and Lee [85] developed a discrete-time model for the evolution of the yield curve. The continuous-time limit of the Ho-Lee model is a constant-diffusion forward rate. In particular, the interest rate behaves like that in a Vasicek model and can become negative.

An arbitrage-free framework for the evolution of the yield curve in continuous time was developed by Heath, Jarrow, and Morton [83]. Related papers are [81] and [82]. The HJM framework presented in this chapter is general, but it can be specialized to obtain a Markov implementation; see Brace and Musiela [20], Cheyette [34], and Hunt, Kennedy, and Pelsser [90]. Filipović [66] examines the issue of making the yield curves generated by the HJM model consistent with the scheme used to generate the initial yield curve. Jara [96] considers an HJM-type model but for interest rate futures rather than forward rates. The advantage is that the drift term causing the explosion discussed in Subsection 10.4.1 does not appear in such a model.

The switch from continuously compounding forward rates to simple forward rates in order to remove the explosion problem described in Section 10.4.1 was proposed by Sandmann and Sondermann [146], [147]. The use of a log-normal simple interest rate to price caps and floors was worked out by Miltersen, Sandmann, and Sondermann [125]. This idea was embedded in a full forward LIBOR term-structure model by Brace, Gątarek, and Musiela [19]. This was the first full term-structure model consistent with the heuristic formula provided by Black [13] in 1976 and in common use since then.

Recently, a variation on forward LIBOR models has been developed for swaps markets; see Jamshidian [95] and the three books cited below. Term-structure models with jumps have been studied by Björk, Kabanov, and Runggaldier [12], Das [46], Das and Foresi [47], Glasserman and Kou [73], Glasserman and Merener [74], and Shirakawa [149].

Three recent books by authors with practical experience in term-structure modeling are Pelsser [131], Brigo and Mercurio [21], and Rebonato [137]. Pelsser's text [131] is succinct but comprehensive, Brigo and Mercurio's text [21] contains considerably more detail, and Rebonato's book [137] is devoted to forward LIBOR models.

## 10.7 Exercises

**Exercise 10.1 (Statistics in the two-factor Vasicek model).** According to Example 4.7.3, $Y_1(t)$ and $Y_2(t)$ in (10.2.43)–(10.2.46) are Gaussian processes.

(i) Show that

$$\widetilde{\mathbb{E}}Y_1(t) = e^{-\lambda_1 t}Y_1(0), \tag{10.7.1}$$

that when $\lambda_1 \neq \lambda_2$, then

$$\widetilde{\mathbb{E}}Y_2(t) = \frac{\lambda_{21}}{\lambda_1 - \lambda_2}(e^{-\lambda_1 t} - e^{-\lambda_2 t})Y_1(0) + e^{-\lambda_2 t}Y_2(0), \tag{10.7.2}$$

and when $\lambda_1 = \lambda_2$, then

$$\widetilde{\mathbb{E}}Y_2(t) = -\lambda_{21}te^{-\lambda_1 t}Y_1(0) + e^{-\lambda_1 t}Y_2(0). \tag{10.7.3}$$

We can write

$$Y_1(t) - \widetilde{\mathbb{E}}Y_1(t) = e^{-\lambda_1 t}I_1(t),$$

when $\lambda_1 \neq \lambda_2$,

$$Y_2(t) - \mathbb{E}Y_2(t) = \frac{\lambda_{21}}{\lambda_1 - \lambda_2}\big(e^{-\lambda_1 t}I_1(t) - e^{-\lambda_2 t}I_2(t)\big) - e^{-\lambda_2 t}I_3(t),$$

and when $\lambda_1 = \lambda_2$,

$$Y_2(t) - \widetilde{\mathbb{E}}Y_2(t) = -\lambda_{21}te^{-\lambda_1 t}I_1(t) + \lambda_{21}e^{-\lambda_1 t}I_4(t) + e^{-\lambda_1 t}I_3(t),$$

where the Itô integrals

$$I_1(t) = \int_0^t e^{\lambda_1 u}\, d\widetilde{W}_1(u),\ I_2(t) = \int_0^t e^{\lambda_2 u}\, d\widetilde{W}_1(u),$$
$$I_3(t) = \int_0^t e^{\lambda_2 u}\, d\widetilde{W}_2(u),\ I_4(t) = \int_0^t u e^{\lambda_1 u}\, d\widetilde{W}_1(u),$$

all have expectation zero under the risk-neutral measure $\widetilde{\mathbb{P}}$. Consequently, we can determine the variances of $Y_1(t)$ and $Y_2(t)$ and the covariance of $Y_1(t)$ and $Y_2(t)$ under the risk-neutral measure from the variances and covariances of $I_j(t)$ and $I_k(t)$. For example, if $\lambda_1 = \lambda_2$, then

$$\begin{aligned}
&\operatorname{Var}\big(Y_1(t)\big)\\
&= e^{-2\lambda_1 t}\widetilde{\mathbb{E}} I_1^2(t),\\
&\operatorname{Var}\big(Y_2(t)\big)\\
&= \lambda_{21}^2 t^2 e^{-2\lambda_1 t}\widetilde{\mathbb{E}} I_1^2(t) + \lambda_{21}^2 e^{-2\lambda_1 t}\widetilde{\mathbb{E}} I_4^2(t) + e^{-2\lambda_1 t}\widetilde{\mathbb{E}} I_3^2(t)\\
&\quad -2\lambda_{21}^2 t e^{-2\lambda_1 t}\widetilde{\mathbb{E}}\big[I_1(t)I_4(t)\big] - 2\lambda_{21} t e^{-2\lambda_1 t}\widetilde{\mathbb{E}}\big[I_1(t)I_3(t)\big]\\
&\quad +2\lambda_{21} e^{-2\lambda_1 t}\widetilde{\mathbb{E}}\big[I_4(t)I_3(t)\big],\\
&\operatorname{Cov}\big(Y_1(t), Y_2(t)\big)\\
&= -\lambda_{21} t e^{-2\lambda_1 t}\widetilde{\mathbb{E}} I_1^2(t) + \lambda_{21} e^{-2\lambda_1 t}\widetilde{\mathbb{E}}\big[I_1(t)I_4(t)\big] + e^{-2\lambda_1 t}\widetilde{\mathbb{E}}\big[I_1(t)I_3(t)\big],
\end{aligned}$$

where the variances and covariance above are under the risk-neutral measure $\widetilde{\mathbb{P}}$.

(ii) Compute the five terms

$$\widetilde{\mathbb{E}} I_1^2(t),\ \widetilde{\mathbb{E}}\big[I_1(t)I_2(t)\big],\ \widetilde{\mathbb{E}}\big[I_1(t)I_3(t)\big],\ \widetilde{\mathbb{E}}\big[I_1(t)I_4(t)\big],\ \widetilde{\mathbb{E}}\big[I_4^2(t)\big].$$

The five other terms, which you are not being asked to compute, are

$$\begin{aligned}
\mathbb{E} I_2^2(t) &= \frac{1}{2\lambda_2}\big(e^{2\lambda_2 t} - 1\big),\\
\mathbb{E}\big[I_2(t)I_3(t)\big] &= 0,\\
\mathbb{E}\big[I_2(t)I_4(t)\big] &= \frac{t}{\lambda_1+\lambda_2} e^{(\lambda_1+\lambda_2)t} + \frac{1}{(\lambda_1+\lambda_2)^2}\big(1 - e^{(\lambda_1+\lambda_2)t}\big),\\
\mathbb{E} I_3^2(t) &= \frac{1}{\lambda_2}\big(e^{2\lambda_2 t} - 1\big),\\
\mathbb{E}\big[I_3(t)I_4(t)\big] &= 0.
\end{aligned}$$

(iii) Some derivative securities involve *time spread* (i.e., they depend on the interest rate at two different times). In such cases, we are interested in the joint statistics of the factor processes at different times. These are still jointly normal and depend on the statistics of the Itô integrals $I_j$ at different times. Compute $\widetilde{\mathbb{E}}\big[I_1(s)I_2(t)\big]$, where $0 \le s < t$. (Hint: Fix $s \ge 0$ and define

$$J_1(t) = \int_0^t e^{\lambda_1 u}\mathbb{I}_{\{u\le s\}}\, d\widetilde{W}_1(u),$$

where $\mathbb{I}_{\{u\le s\}}$ is the function of $u$ that is 1 if $u \le s$ and 0 if $u > s$. Note that $J_1(t) = I_1(s)$ when $t \ge s$.)

**Exercise 10.2 (Ordinary differential equations for the mixed affine-yield model).** In the mixed model of Subsection 10.2.3, as in the two-factor Vasicek model and the two-factor Cox-Ingersoll-Ross model, zero-coupon bond prices have the affine-yield form

$$f(t, y_1, y_2) = e^{-y_1 C_1(T-t) - y_2 C_2(T-t) - A(T-t)},$$

where $C_1(0) = C_2(0) = A(0) = 0$.

(i) Find the partial differential equation satisfied by $f(t, y_1, y_2)$.
(ii) Show that $C_1$, $C_2$, and $A$ satisfy the system of ordinary differential equations

$$C_1' = -\lambda_1 C_1 - \frac{1}{2}C_1^2 - \sigma_{21}C_1C_2 - (1+\beta)C_2^2 + \delta_1, \tag{10.7.4}$$

$$C_2' = -\lambda_2 C_2 + \delta_2, \tag{10.7.5}$$

$$A' = \mu C_1 - \frac{1}{2}\alpha C_2^2 + \delta_0. \tag{10.7.6}$$

**Exercise 10.3 (Calibration of the two-factor Vasicek model).** Consider the canonical two-factor Vasicek model (10.2.4), (10.2.5), but replace the interest rate equation (10.2.6) by

$$R(t) = \delta_0(t) + \delta_1 Y_1(t) + \delta_2 Y_2(t), \tag{10.7.7}$$

where $\delta_1$ and $\delta_2$ are constant but $\delta_0(t)$ is a nonrandom function of time. Assume that for each $T$ there is a zero-coupon bond maturing at time $T$. The price of this bond at time $t \in [0, T]$ is

$$B(t,T) = \widetilde{\mathbb{E}}\left[e^{-\int_t^T R(u)du}\,\Big|\,\mathcal{F}(t)\right].$$

Because the pair of processes $\big(Y_1(t), Y_2(t)\big)$ is Markov, there must exist some function $f(t, T, y_1, y_2)$ such that $B(t,T) = f\big(t, T, Y_1(t), Y_2(t)\big)$. (We indicate the dependence of $f$ on the maturity $T$ because, unlike in Subsection 10.2.1, here we shall consider more than one value of $T$.)

(i) The function $f(t, T, y_1, y_2)$ is of the affine-yield form

$$f(t, T, y_1, y_2) = e^{-y_1 C_1(t,T) - y_2 C_2(t,T) - A(t,T)}. \tag{10.7.8}$$

Holding $T$ fixed, derive a system of ordinary differential equations for $\frac{d}{dt}C_1(t,T)$, $\frac{d}{dt}C_2(t,T)$, and $\frac{d}{dt}A(t,T)$.

(ii) Using the terminal conditions $C_1(T,T) = C_2(T,T) = 0$, solve the equations in (i) for $C_1(t,T)$ and $C_2(t,T)$. (As in Subsection 10.2.1, the functions $C_1$ and $C_2$ depend on $t$ and $T$ only through the difference $\tau = T - t$; however, the function $A$ discussed in part (iii) below depends on $t$ and $T$ separately.)

(iii) Using the terminal condition $A(T,T) = 0$, write a formula for $A(t,T)$ as an integral involving $C_1(u,T)$, $C_2(u,T)$, and $\delta_0(u)$. You do not need to evaluate this integral.

(iv) Assume that the model parameters $\lambda_1 > 0$ $\lambda_2 > 0$, $\lambda_{21}$, $\delta_1$, and $\delta_2$ and the initial conditions $Y_1(0)$ and $Y_2(0)$ are given. We wish to choose a *function* $\delta_0$ so that the zero-coupon bond prices given by the model match the bond prices given by the market at the initial time zero. In other words, we want to choose a function $\delta(T)$, $T \geq 0$, so that

$$f\big(0,T,Y_1(0),Y_2(0)\big) = B(0,T), \quad T \geq 0.$$

In this part of the exercise, we regard both $t$ and $T$ as variables and use the notation $\frac{\partial}{\partial t}$ to indicate the derivative with respect to $t$ when $T$ is held fixed and the notation $\frac{\partial}{\partial T}$ to indicate the derivative with respect to $T$ when $t$ is held fixed. Give a formula for $\delta_0(T)$ in terms of $\frac{\partial}{\partial T}\log B(0,T)$ and the model parameters. (Hint: Compute $\frac{\partial}{\partial T}A(0,T)$ in two ways, using (10.7.8) and also using the formula obtained in (iii). Because $C_i(t,T)$ depends only on $t$ and $T$ through $\tau = T - t$, there are functions $\overline{C}_i(\tau)$ such that $\overline{C}_i(\tau) = \overline{C}_i(T-t) = C_i(t,T)$, $i = 1,2$. Then

$$\frac{\partial}{\partial t}C_i(t,T) = -\overline{C}_i'(\tau), \quad \frac{\partial}{\partial T}C_i(t,T) = \overline{C}_i'(\tau),$$

where $'$ denotes differentiation with respect to $\tau$. This shows that

$$\frac{\partial}{\partial T}C_i(t,T) = -\frac{\partial}{\partial t}C_i(t,T), \; i = 1,2, \tag{10.7.9}$$

a fact that you will need.)

**Exercise 10.4.** Hull and White [89] propose the two-factor model

$$dU(t) = -\lambda_1 U(t)\,dt + \sigma_1\,d\widetilde{B}_2(t), \tag{10.7.10}$$

$$dR(t) = \big[\theta(t) + U(t) - \lambda_2 R(t)\big]\,dt + \sigma_2\,d\widetilde{B}_1(t), \tag{10.7.11}$$

where $\lambda_1$, $\lambda_2$, $\sigma_1$, and $\sigma_2$ are positive constants, $\theta(t)$ is a nonrandom function, and $\widetilde{B}_1(t)$ and $\widetilde{B}_2(t)$ are correlated Brownian motions with $d\widetilde{B}_1(t)\,d\widetilde{B}_2(t) = \rho\,dt$ for some $\rho \in (-1,1)$. In this exercise, we discuss how to reduce this to the two-factor Vasicek model of Subsection 10.2.1, except that, instead of (10.2.6), the interest rate is given by (10.7.7), in which $\delta_0(t)$ is a nonrandom function of time.

(i) Define

$$X(t) = \begin{bmatrix} U(t) \\ R(t) \end{bmatrix}, \quad K = \begin{bmatrix} \lambda_1 & 0 \\ -1 & \lambda_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$

$$\Theta(t) = \begin{bmatrix} 0 \\ \theta(t) \end{bmatrix}, \quad \widetilde{B}(t) = \begin{bmatrix} \widetilde{B}_1(t) \\ \widetilde{B}_2(t) \end{bmatrix},$$

so that (10.7.10) and (10.7.11) can be written in vector notation as

$$dX(t) = \Theta(t)\,dt - KX(t)\,dt + \Sigma\,d\widetilde{B}(t). \tag{10.7.12}$$

Now set

$$\widehat{X}(t) = X(t) - e^{-Kt}\int_0^t e^{Ku}\Theta(u)\,du.$$

Show that

$$d\widehat{X}(t) = -K\widehat{X}(t)\,dt + \Sigma d\widetilde{B}(t). \tag{10.7.13}$$

(ii) With

$$C = \begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ -\frac{\rho}{\sigma_1\sqrt{1-\rho^2}} & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{bmatrix},$$

define $Y(t) = C\widehat{X}(t)$, $\widetilde{W}(t) = C\Sigma\widetilde{B}(t)$. Show that the components of $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ are independent Brownian motions and

$$dY(t) = -\Lambda Y(t) + d\widetilde{W}(t), \tag{10.7.14}$$

where

$$\Lambda = CKC^{-1} = \begin{bmatrix} \lambda_1 & 0 \\ \frac{\rho\sigma_2(\lambda_2-\lambda_1)-\sigma_1}{\sigma_2\sqrt{1-\rho^2}} & \lambda_2. \end{bmatrix}.$$

Equation (10.7.14) is the vector form of the canonical two-factor Vasicek equations (10.2.4) and (10.2.5).

(iii) Obtain a formula for $R(t)$ of the form (10.7.7). What are $\delta_0(t)$, $\delta_1$, and $\delta_2$?

**Exercise 10.5 (Correlation between long rate and short rate in the one-factor Vasicek model).** The one-factor Vasicek model is the one-factor Hull-White model of Example 6.5.1 with constant parameters,

$$dR(t) = \big(a - bR(t)\big)\,dt + \sigma\,d\widetilde{W}(t), \tag{10.7.15}$$

where $a$, $b$, and $\sigma$ are positive constants and $\widetilde{W}(t)$ is a one-dimensional Brownian motion. In this model, the price at time $t \in [0,T]$ of the zero-coupon bond maturing at time $T$ is

$$B(t,T) = e^{-C(t,T)R(t)-A(t,T)},$$

where $C(t,T)$ and $A(t,T)$ are given by (6.5.10) and (6.5.11):

$$\begin{aligned}
C(t,T) &= \int_t^T e^{-\int_t^s b\,dv}\,ds \;=\; \frac{1}{b}\big(1-e^{-b(T-t)}\big),\\
A(t,T) &= \int_t^T \left(aC(s,T) - \frac{1}{2}\sigma^2 C^2(s,T)\right) ds\\
&= \frac{2ab-\sigma^2}{2b^2}(T-t) + \frac{\sigma^2-ab}{b^3}\big(1-e^{-b(T-t)}\big) - \frac{\sigma^2}{4b^3}\big(1-e^{-2b(T-t)}\big).
\end{aligned}$$

In the spirit of the discussion of the short rate and the long rate in Subsection 10.2.1, we fix a positive relative maturity $\overline{\tau}$ and define the long rate $L(t)$ at time $t$ by (10.2.30):

$$L(t) = -\frac{1}{\overline{\tau}}\log B(t, t+\overline{\tau}).$$

Show that changes in $L(t)$ and $R(t)$ are perfectly correlated (i.e., for any $0 \le t_1 < t_2$, the correlation coefficient between $L(t_2)-L(t_1)$ and $R(t_2)-R(t_1)$ is one). This characteristic of one-factor models caused the development of models with more than one factor.

**Exercise 10.6 (Degenerate two-factor Vasicek model).** In the discussion of short rates and long rates in the two-factor Vasicek model of Subsection 10.2.1, we made the assumptions that $\delta_2 \neq 0$ and $(\lambda_1-\lambda_2)\delta_1 + \lambda_{21}\delta_2 \neq 0$ (see Lemma 10.2.2). In this exercise, we show that if either of these conditions is violated, the two-factor Vasicek model reduces to a one-factor model, for which long rates and short rates are perfectly correlated (see Exercise 10.5).

(i) Show that if $\delta_2 = 0$ (and $\delta_0 > 0$, $\delta_1 > 0$), then the short rate $R(t)$ given by the system of equations (10.2.4)–(10.2.6) satisfies the one-dimensional stochastic differential equation

$$dR(t) = \big(a - bR(t)\big)\,dt + d\widetilde{W}_1(t). \tag{10.7.16}$$

Define $a$ and $b$ in terms of the parameters in (10.2.4)–(10.2.6).

(ii) Show that if $(\lambda_1-\lambda_2)\delta_1 + \lambda_{21}\delta_2 = 0$ (and $\delta_0 > 0$, $\delta_1^2+\delta_2^2 \neq 0$), then the short rate $R(t)$ given by the system of equations (10.2.4)–(10.2.6) satisfies the one-dimensional stochastic differential equation

$$dR(t) = \big(a - bR(t)\big)\,dt + \sigma\,d\widetilde{B}(t). \tag{10.7.17}$$

Define $a$ and $b$ in terms of the parameters in (10.2.4)–(10.2.6) and define the Brownian motion $\widetilde{B}(t)$ in terms of the independent Brownian motions $\widetilde{W}_1(t)$ and $\widetilde{W}_2(t)$ in (10.2.4) and (10.2.5).

**Exercise 10.7 (Forward measure in the two-factor Vasicek model).** Fix a maturity $T > 0$. In the two-factor Vasicek model of Subsection 10.2.1, consider the $T$-forward measure $\widetilde{\mathbb{P}}^T$ of Definition 9.4.1:

$$\widetilde{\mathbb{P}}^T(A) = \frac{1}{B(0,T)} \int_A D(T)\, d\widetilde{\mathbb{P}} \text{ for all } A \in \mathcal{F}.$$

(i) Show that the two-dimensional $\widetilde{\mathbb{P}}^T$-Brownian motions $\widetilde{W}_1^T(t)$, $\widetilde{W}_2^T(t)$ of (9.2.5) are

$$\widetilde{W}_j^T(t) = \int_0^t C_1(T-u)\, du + \widetilde{W}_j(t), \quad j = 1, 2, \tag{10.7.18}$$

where $C_1(\tau)$ and $C_2(\tau)$ are given by (10.2.26)–(10.2.28).

(ii) Consider a call option on a bond maturing at time $\overline{T} > T$. The call expires at time $T$ and has strike price $K$. Show that at time zero the risk-neutral price of this option is

$$B(0,T)\widetilde{\mathbb{E}}^T\left[\left(e^{-C_1(\overline{T}-T)Y_1(T)-C_2(\overline{T}-T)Y_2(T)-A(\overline{T}-T)} - K\right)^+\right]. \tag{10.7.19}$$

(iii) Show that, under the $T$-forward measure $\widetilde{\mathbb{P}}^T$, the term

$$X = -C_1(\overline{T}-T)Y_1(T) - C_2(\overline{T}-T)Y_1(T) - A(\overline{T}-T)$$

appearing in the exponent in (10.7.19) is normally distributed.

(iv) It is a straightforward but lengthy computation, like the computations in Exercise 10.1, to determine the mean and variance of the term $X$. Let us call its variance $\sigma^2$ and its mean $\mu - \frac{1}{2}\sigma^2$, so that we can write $X$ as

$$X = \mu - \frac{1}{2}\sigma^2 - \sigma Z,$$

where $Z$ is a standard normal random variable under $\widetilde{\mathbb{P}}^T$. Show that the call option price in (10.7.19) is

$$B(0,T)\big(e^{\mu} N(d_+) - K N(d_-)\big),$$

where

$$d_\pm = \frac{1}{\sigma}\left(\mu - \log K \pm \frac{1}{2}\sigma^2\right).$$

**Exercise 10.8 (Reversal of order of integration in forward rates).** The forward rate formula (10.3.5) with $v$ replacing $T$ states that

$$f(t,v) = f(0,v) + \int_0^t \alpha(u,v)\, du + \int_0^t \sigma(u,v)\, dW(u).$$

Therefore,

$$-\int_t^T f(t,v)\, dv = -\int_t^T \left[f(0,v) + \int_0^t \alpha(u,v)\, du + \int_0^t \sigma(u,v)\, dW(u)\right] dv. \tag{10.7.20}$$

(i) Define

$$\widehat{\alpha}(u,t,T) = \int_t^T \alpha(u,v)\,dv, \quad \widehat{\sigma}(u,t,T) = \int_t^T \sigma(u,v)\,dv.$$

Show that if we reverse the order of integration in (10.7.20), we obtain the equation

$$\begin{aligned} &-\int_t^T f(t,v)\,dv \\ &= -\int_t^T f(0,v)\,dv - \int_0^t \widehat{\alpha}(u,t,T)\,du - \int_0^t \widehat{\sigma}(u,t,T)\,dW(u). \end{aligned} \tag{10.7.21}$$

(In one case, this is a reversal of the order of two Riemann integrals, a step that uses only the theory of ordinary calculus. In the other case, the order of a Riemann and an Itô integral are being reversed. This step is justified in the appendix of [83]. You may assume without proof that this step is legitimate.)

(ii) Take the differential with respect to $t$ in (10.7.21), remembering to get two terms from each of the integrals $\int_0^t \widehat{\alpha}(u,t,T)\,du$ and $\int_0^t \widehat{\sigma}(u,t,T)\,dW(u)$ because one must differentiate with respect to each of the two $t$s appearing in these integrals.

(iii) Check that your formula in (ii) agrees with (10.3.10).

**Exercise 10.9 (Multifactor HJM model).** Suppose the Heath-Jarrow-Morton model is driven by a $d$-dimensional Brownian motion, so that $\sigma(t,T)$ is also a $d$-dimensional vector and the forward rate dynamics are given by

$$df(t,T) = \alpha(t,T)\,dt + \sum_{j=1}^{d} \sigma_j(t,T)\,dW_j(t).$$

(i) Show that (10.3.16) becomes

$$\alpha(t,T) = \sum_{j=1}^{d} \sigma_j(t,T)\left[\sigma_j^*(t,T) + \Theta_j(t)\right].$$

(ii) Suppose there is an adapted, $d$-dimensional process

$$\Theta(t) = \big(\Theta_1(t), \ldots, \Theta_d(t)\big)$$

satisfying this equation for all $0 \le t \le T \le \overline{T}$. Show that if there are maturities $T_1, \ldots, T_d$ such that the $d \times d$ matrix $\big(\sigma_j(t,T_i)\big)_{i,j}$ is nonsingular, then $\Theta(t)$ is unique.

**Exercise 10.10.** (i) Use the ordinary differential equations (6.5.8) and (6.5.9) satisfied by the functions $A(t,T)$ and $C(t,T)$ in the one-factor Hull-White model to show that this model satisfies the HJM no-arbitrage condition (10.3.27).

(ii) Use the ordinary differential equations (6.5.14) and (6.5.15) satisfied by the functions $A(t,T)$ and $C(t,T)$ in the one-factor Cox-Ingersoll-Ross model to show that this model satisfies the HJM no-arbitrage condition (10.3.27).

**Exercise 10.11.** Let $\delta > 0$ be given. Consider an interest rate swap paying a fixed interest rate $K$ and receiving backset LIBOR $L(T_{j-1}, T_{j-1})$ on a principal of 1 at each of the payment dates $T_j = \delta j$, $j = 1, 2, \ldots, n+1$. Show that the value of the swap is

$$\delta K \sum_{j=1}^{n+1} B(0,T_j) - \delta \sum_{j=1}^{n+1} B(0,T_j) L(0,T_{j-1}). \tag{10.7.22}$$

*Remark 10.7.1.* The *swap rate* is defined to be the value of $K$ that makes the initial value of the swap equal to zero. Thus, the swap rate is

$$K = \frac{\sum_{j=1}^{n+1} B(0,T_j) L(0,T_{j-1})}{\sum_{j=1}^{n+1} B(0,T_j)}. \tag{10.7.23}$$

**Exercise 10.12.** In the proof of Theorem 10.4.1, we showed by an arbitrage argument that the value at time 0 of a payment of backset LIBOR $L(T,T)$ at time $T+\delta$ is $B(0,T+\delta)L(0,T)$. The risk-neutral price of this payment, computed at time zero, is

$$\widetilde{\mathbb{E}}\left[D(T+\delta)L(T,T)\right].$$

Use the definitions

$$L(T,T) = \frac{1 - B(T,T+\delta)}{\delta B(T,T+\delta)},$$

$$B(0,T+\delta) = \widetilde{\mathbb{E}}\left[D(T+\delta)\right],$$

and the properties of conditional expectations to show that

$$\widetilde{\mathbb{E}}\left[D(T+\delta)L(T,T)\right] = B(0,T+\delta)L(0,T).$$

This page intentionally left blank

# 11

# Introduction to Jump Processes

## 11.1 Introduction

This chapter studies *jump-diffusion* processes. The "diffusion" part of the nomenclature refers to the fact that these processes can have a Brownian motion component or, more generally, an integral with respect to Brownian motion. In addition, the paths of these processes may have jumps. We consider in this chapter the special case when there are only finitely many jumps in each finite time interval.

One can also construct processes in which there are infinitely many jumps in a finite time interval, although for such processes it is necessarily the case that, for each positive threshold, only finitely many jumps can have a size exceeding the threshold in any finite time interval. The number exceeding the threshold can depend on the threshold and become arbitrarily large as the threshold approaches zero. Such processes are not considered here, although the theory provided here gives some idea of how such processes can be analyzed.

The fundamental pure jump process is the *Poisson process*, and this is presented in Section 11.2. All jumps of a Poisson process are of size one. A *compound Poisson process* is like a Poisson process, except that the jumps are of random size. Compound Poisson processes are the subject of Section 11.3.

In Section 11.4, we define a *jump process* to be the sum of a nonrandom initial condition, an Itô integral with respect to a Brownian motion $dW(t)$, a Riemann integral with respect to $dt$, and a *pure jump* process. A pure jump process begins at zero, has finitely many jumps in each finite time interval, and is constant between jumps. Section 11.4 defines stochastic integrals with respect to jump processes. These stochastic integrals are themselves jump processes. Section 11.4 also examines the quadratic variation of jump processes and their stochastic integrals.

In Section 11.5, we present the stochastic calculus for jump processes. The key result is the extension of the Itô-Doeblin formula to cover these processes.

In Section 11.6, we take up the matter of changing measures for Poisson processes and for compound Poisson processes. We conclude with a discussion of how to simultaneously change the measure for a Brownian motion and a compound Poisson process. The effect of this change is to adjust the drift of the Brownian motion and to adjust the *intensity* (average rate of jump arrival) and the distribution of the jump sizes for the compound Poisson process.

In Section 11.7, we apply this theory to the problem of pricing and partially hedging a European call in a jump-diffusion model.

## 11.2 Poisson Process

In the way that Brownian motion is the basic building block for continuous-path processes, the Poisson process serves as the starting point for jump processes. In this section, we construct the Poisson process and develop its basic properties.

### 11.2.1 Exponential Random Variables

Let $\tau$ be a random variable with density

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0, \\ 0, & t < 0, \end{cases} \tag{11.2.1}$$

where $\lambda$ is a positive constant. We say that $\tau$ has the *exponential distribution* or simply that $\tau$ is an *exponential random variable.*

The expected value of $\tau$ can be computed by an integration by parts:

$$\begin{aligned} \mathbb{E}\tau = \int_0^\infty t f(t)\, dt = \lambda \int_0^\infty t e^{-\lambda t}\, dt &= -te^{-\lambda t}\Big|_{t=0}^{t=\infty} + \int_0^\infty e^{-\lambda t}\, dt \\ &= 0 - \tfrac{1}{\lambda} e^{-\lambda t}\Big|_{t=0}^{t=\infty} = \tfrac{1}{\lambda}. \end{aligned}$$

For the cumulative distribution function, we have

$$F(t) = \mathbb{P}\{\tau \leq t\} = \int_0^t \lambda e^{-\lambda u}\, du = -e^{-\lambda u}\Big|_{u=0}^{u=t} = 1 - e^{-\lambda t}, \quad t \geq 0,$$

and hence

$$\mathbb{P}\{\tau > t\} = e^{-\lambda t}, \quad t \geq 0. \tag{11.2.2}$$

Suppose we are waiting for an event, such as default of a bond, and we know that the distribution of the time of this event is exponential with mean $\frac{1}{\lambda}$ (i.e., it has the density (11.2.1)). Suppose we have already waited $s$ time units, and we are interested in the probability that we will have to wait an additional $t$ time units (conditioned on knowing that the event has not occurred during the time interval $[0, s]$). This probability is

$$\begin{aligned}\mathbb{P}\{\tau > t+s \mid \tau > s\} &= \frac{\mathbb{P}\{\tau > t+s \text{ and } \tau > s\}}{\mathbb{P}\{\tau > s\}}\\ &= \frac{\mathbb{P}\{\tau > t+s\}}{\mathbb{P}\{\tau > s\}} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}.\end{aligned} \tag{11.2.3}$$

In other words, after waiting $s$ time units, the probability that we will have to wait an additional $t$ time units is the same as the probability of having to wait $t$ time units when we were starting at time 0. The fact that we have already waited $s$ time units does not change the distribution of the remaining time. This property for the exponential distribution is called *memorylessness.*

### 11.2.2 Construction of a Poisson Process

To construct a Poisson process, we begin with a sequence $\tau_1, \tau_2, \ldots$ of independent exponential random variables, all with the same mean $\frac{1}{\lambda}$. We will build a model in which an event, which we call a "jump," occurs from time to time. The first jump occurs at time $\tau_1$, the second occurs $\tau_2$ time units after the first, the third occurs $\tau_3$ time units after the second, etc. The $\tau_k$ random variables are called the *interarrival times.* The *arrival times* are

$$S_n = \sum_{k=1}^{n} \tau_k \tag{11.2.4}$$

(i.e., $S_n$ is the time of the $n$th jump). The *Poisson process* $N(t)$ counts the number of jumps that occur at or before time $t$. More precisely,

$$N(t) = \begin{cases} 0 \text{ if } 0 \le t < S_1, \\ 1 \text{ if } S_1 \le t < S_2, \\ \vdots \\ n \text{ if } S_n \le t < S_{n+1}, \\ \vdots \end{cases}$$

Note that at the jump times $N(t)$ is defined so that it is *right-continuous* (i.e., $N(t) = \lim_{s \downarrow t} N(s)$). We denote by $\mathcal{F}(t)$ the $\sigma$-algebra of information acquired by observing $N(s)$ for $0 \le s \le t$.

Because the expected time between jumps is $\frac{1}{\lambda}$, the jumps are arriving at an average rate of $\lambda$ per unit time. We say the Poisson process $N(t)$ has *intensity* $\lambda$. Figure 11.2.1 shows one path of a Poisson process.

### 11.2.3 Distribution of Poisson Process Increments

In order to determine the distribution of the increments of a Poisson process, we must first determine the distribution of the jump times $S_1, S_2, \ldots$.

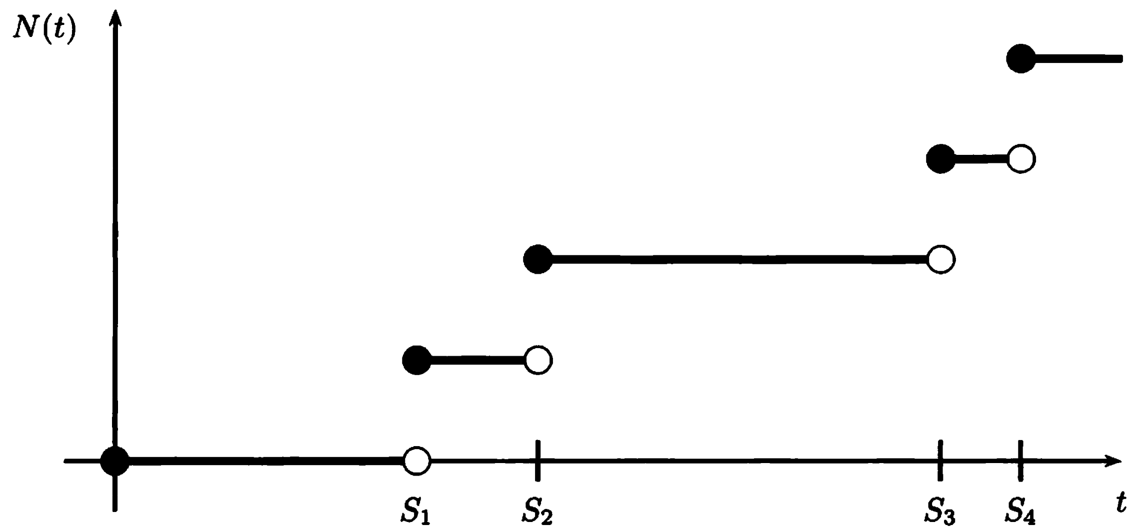


**Fig. 11.2.1.** One path of a Poisson process.

**Lemma 11.2.1.** *For $n \geq 1$, the random variable $S_n$ defined by (11.2.4) has the* gamma density

$$g_n(s) = \frac{(\lambda s)^{n-1}}{(n-1)!}\lambda e^{-\lambda s}, \quad s \geq 0. \tag{11.2.5}$$

PROOF: We prove (11.2.5) by induction on $n$. For $n = 1$, we have that $S_1 = \tau_1$ is exponential, and (11.2.5) becomes the exponential density

$$g_1(s) = \lambda e^{-\lambda s}, \quad s \geq 0.$$

(Recall that 0! is defined to be 1.) Having thus established the base case, let us assume that (11.2.5) holds for some value of $n$ and prove it for $n+1$. In other words, we assume $S_n$ has density $g_n(s)$ given in (11.2.5) and we want to compute the density of $S_{n+1} = S_n + \tau_{n+1}$. Since $S_n$ and $\tau_{n+1}$ are independent, the density of $S_{n+1}$ can be computed by the convolution

$$\begin{aligned}
\int_0^s g_n(v) f(s-v)\, dv &= \int_0^s \frac{(\lambda v)^{n-1}}{(n-1)!}\lambda e^{-\lambda v} \cdot \lambda e^{-\lambda(s-v)}\, dv \\
&= \frac{\lambda^{n+1} e^{-\lambda s}}{(n-1)!} \int_0^s v^{n-1}\, ds \;=\; \frac{\lambda^{n+1} e^{-\lambda s}}{n!}\, v^n \Big|_{v=0}^{v=s} \\
&= \frac{(\lambda s)^n}{n!}\lambda e^{-\lambda s} \;=\; g_{n+1}(s).
\end{aligned}$$

This completes the induction step and proves the lemma. □

**Lemma 11.2.2.** *The Poisson process $N(t)$ with intensity $\lambda$ has the distribution*

$$\mathbb{P}\{N(t) = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, \ldots. \tag{11.2.6}$$

PROOF: For $k \geq 1$, we have $N(t) \geq k$ if and only if there are at least $k$ jumps by time $t$ (i.e., if and only if $S_k$, the time of the $k$th jump, is less than or equal to $t$). Therefore,

$$\mathbb{P}\{N(t) \geq k\} = \mathbb{P}\{S_k \leq t\} = \int_0^t \frac{(\lambda s)^{k-1}}{(k-1)!} \lambda e^{-\lambda s}\, ds.$$

Similarly,

$$\mathbb{P}\{N(t) \geq k+1\} = \mathbb{P}\{S_{k+1} \leq t\} = \int_0^t \frac{(\lambda s)^k}{k!} \lambda e^{-\lambda s}\, ds.$$

We integrate this last expression by parts to obtain

$$\begin{aligned}\mathbb{P}\{N(t) \geq k+1\} &= -\frac{(\lambda s)^k}{k!} e^{-\lambda s}\Big|_{s=0}^{s=t} + \int_0^t \frac{(\lambda s)^{k-1}}{(k-1)!} \lambda e^{-\lambda s}\, ds \\ &= -\frac{(\lambda t)^k}{k!} e^{-\lambda t} + \mathbb{P}\{N(t) \geq k\}.\end{aligned}$$

This implies that for $k \geq 1$,

$$\mathbb{P}\{N(t) = k\} = \mathbb{P}\{N(t) \geq k\} - \mathbb{P}\{N(t) \geq k+1\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

For $k = 0$, we have from (11.2.2)

$$\mathbb{P}\{N(t) = 0\} = \mathbb{P}\{S_1 > t\} = \mathbb{P}\{\tau_1 > t\} = e^{-\lambda t},$$

which is (11.2.6) with $k = 0$. □

Suppose we observe the Poisson process up to time $s$ and then want to know the distribution of $N(t+s) - N(s)$, conditioned on knowing what has happened up to and including time $s$. It turns out that the information about what has happened up to and including time $s$ is irrelevant. This is a consequence of the memorylessness of exponential random variables (see (11.2.3)). Because $N(t+s) - N(s)$ is the number of jumps in the time interval $(s, t+s]$, in order to compute the distribution of $N(t+s) - N(s)$, we are interested in the time of the next jump after $s$. At time $s$, we know the time since the last jump, but the time between $s$ and the next jump does not depend on this. Indeed, the time between $s$ and the first jump after $s$ has an exponential distribution with mean $\frac{1}{\lambda}$, independent of everything that has happened up to time $s$. The time between that jump and the one after it is also exponentially distributed with mean $\frac{1}{\lambda}$, independent of everything that has happened up to time $s$. The same applies for all subsequent jumps. Consequently, $N(t+s) - N(s)$ is independent of $\mathcal{F}(s)$. Furthermore, the distribution of $N(t+s) - N(s)$ is the same as the distribution of $N(t)$. In both cases, one is simply counting the number of jumps that occur in a time interval of length $t$, and the jumps are independent and exponentially distributed with mean $\frac{1}{\lambda}$. When a process has the property that the distribution of the increment depends only on the difference between the two time points, the increments are said to be *stationary*. Both the Poisson process and Brownian motion have stationary independent increments.

**Theorem 11.2.3.** *Let $N(t)$ be a Poisson process with intensity $\lambda > 0$, and let $0 = t_0 < t_1 < \cdots < t_n$ be given. Then the increments*

$$N(t_1) - N(t_0),\ N(t_2) - N(t_1), \dots, N(t_n) - N(t_{n-1})$$

*are stationary and independent, and*

$$\mathbb{P}\{N(t_{j+1}) - N(t_j) = k\} = \frac{\lambda^k (t_{j+1} - t_j)^k}{k!} e^{-\lambda(t_{j+1} - t_j)}, \quad k = 0, 1, \dots. \tag{11.2.7}$$

OUTLINE OF PROOF: Let $\mathcal{F}(t)$ be the $\sigma$-algebra of information acquired by observing $N(s)$ for $0 \le s \le t$. As we just discussed, $N(t_n) - N(t_{n-1})$ is independent of $\mathcal{F}(t_{n-1})$ and has the same distribution as $N(t_n - t_{n-1})$, which by Lemma 11.2.2 is the distribution given by (11.2.7) with $j = n - 1$. Since the other increments $N(t_1) - N(t_0), \dots, N(t_{n-1}) - N(t_{n-2})$ are $\mathcal{F}(t_{n-1})$-measurable, these increments are independent of $N(t_n) - N(t_{n-1})$. We now repeat the argument for the next-to-last increment $N(t_{n-1}) - N(t_{n-2})$, then the increment before that, etc. □

### 11.2.4 Mean and Variance of Poisson Increments

Let $0 \le s < t$ be given. According to Theorem 11.2.3, the Poisson increment $N(t) - N(s)$ has distribution

$$\mathbb{P}\{N(t) - N(s) = k\} = \frac{\lambda^k (t - s)^k}{k!} e^{-\lambda(t-s)}, \quad k = 0, 1, \dots. \tag{11.2.8}$$

Recall the exponential power series, which we shall use in the three different forms given below:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} = \sum_{k=2}^{\infty} \frac{x^{k-2}}{(k-2)!}.$$

We note first of all from this that

$$\sum_{k=0}^{\infty} \mathbb{P}\{N(t) - N(s) = k\} = e^{-\lambda(t-s)} \sum_{k=0}^{\infty} \frac{\lambda^k (t-s)^k}{k!} = e^{-\lambda(t-s)} \cdot e^{\lambda(t-s)} = 1,$$

as we would expect. We next compute the expected increment

$$\begin{aligned}
\mathbb{E}\big[N(t)-N(s)\big] &= \sum_{k=0}^{\infty} k\frac{\lambda^k(t-s)^k}{k!}e^{-\lambda(t-s)} \\
&= \lambda(t-s)e^{-\lambda(t-s)}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}(t-s)^{k-1}}{(k-1)!} \\
&= \lambda(t-s)\cdot e^{-\lambda(t-s)}\cdot e^{\lambda(t-s)} \\
&= \lambda(t-s). \qquad (11.2.9)
\end{aligned}$$

This is consistent with our observation at the end of Subsection 11.2.2 that jumps are arriving at an average rate of $\lambda$ per unit time. Therefore, the average number of jumps between times $s$ and $t$ is $\mathbb{E}\big[N(t)-N(s)\big]=\lambda(t-s)$.

Finally, we compute the second moment of the increment

$$\begin{aligned}
\mathbb{E}\big[\big(N(t)-N(s)\big)^2\big] &= \sum_{k=0}^{\infty} k^2\frac{\lambda^k(t-s)^k}{k!}e^{-\lambda(t-s)} \\
&= e^{-\lambda(t-s)}\sum_{k=1}^{\infty}(k-1+1)\frac{\lambda^k(t-s)^k}{(k-1)!} \\
&= e^{-\lambda(t-s)}\sum_{k=2}^{\infty}\frac{\lambda^k(t-s)^k}{(k-2)!} + e^{-\lambda(t-s)}\sum_{k=1}^{\infty}\frac{\lambda^k(t-s)^k}{(k-1)!} \\
&= \lambda^2(t-s)^2e^{-\lambda(t-s)}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}(t-s)^{k-2}}{(k-2)!} \\
&\quad +\lambda(t-s)e^{-\lambda(t-s)}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}(t-s)^{k-1}}{(k-1)!} \\
&= \lambda^2(t-s)^2+\lambda(t-s).
\end{aligned}$$

This implies

$$\begin{aligned}
\mathrm{Var}\big[N(t)-N(s)\big] &= \mathbb{E}\big[\big(N(t)-N(s)\big)^2\big]-\big(\mathbb{E}\big[N(t)-N(s)\big]\big)^2 \\
&= \lambda^2(t-s)^2+\lambda(t-s)-\lambda^2(t-s)^2 \\
&= \lambda(t-s); \qquad (11.2.10)
\end{aligned}$$

the variance is the same as the mean.

### 11.2.5 Martingale Property

**Theorem 11.2.4.** *Let $N(t)$ be a Poisson process with intensity $\lambda$. We define the* compensated Poisson process *(see Figure 11.2.2)*

$$M(t)=N(t)-\lambda t.$$

*Then $M(t)$ is a martingale.*

PROOF: Let $0 \leq s < t$ be given. Because $N(t) - N(s)$ is independent of $\mathcal{F}(s)$ and has expected value $\lambda(t-s)$, we have

$$\begin{aligned}
\mathbb{E}\big[M(t)\big|\mathcal{F}(s)\big] &= \mathbb{E}\big[M(t) - M(s)\big|\mathcal{F}(s)\big] + \mathbb{E}\big[M(s)\big|\mathcal{F}(s)\big] \\
&= \mathbb{E}\big[N(t) - N(s) - \lambda(t-s)\big|\mathcal{F}(s)\big] + M(s) \\
&= \mathbb{E}\big[N(t) - N(s)\big] - \lambda(t-s) + M(s) \\
&= M(s).
\end{aligned}$$

□

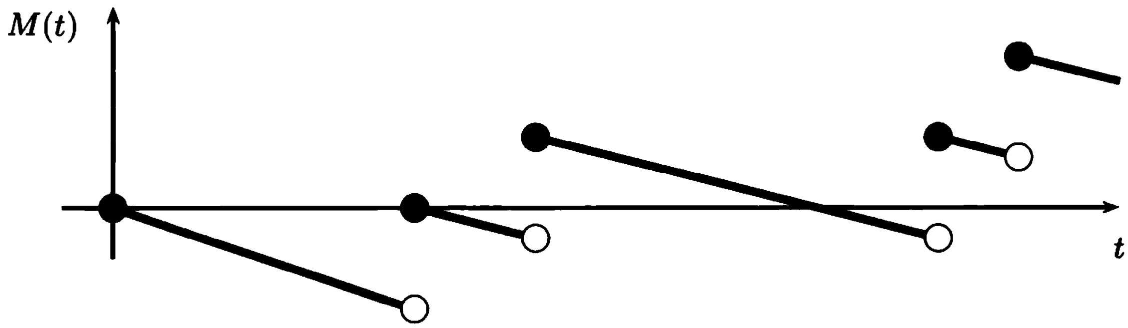


**Fig. 11.2.2.** One path of a compensated Poisson process.

## 11.3 Compound Poisson Process

When a Poisson process or a compensated Poisson process jumps, it jumps up one unit. For models of financial markets, we need to allow the jump size to be random. We introduce random jump sizes in this section.

### 11.3.1 Construction of a Compound Poisson Process

Let $N(t)$ be a Poisson process with intensity $\lambda$, and let $Y_1, Y_2, \ldots$ be a sequence of identically distributed random variables with mean $\beta = \mathbb{E}Y_i$. We assume the random variables $Y_1, Y_2, \ldots$ are independent of one another and also independent of the Poisson process $N(t)$. We define the *compound Poisson process*

$$Q(t) = \sum_{i=1}^{N(t)} Y_i, \quad t \geq 0. \tag{11.3.1}$$

The jumps in $Q(t)$ occur at the same times as the jumps in $N(t)$, but whereas the jumps in $N(t)$ are always of size 1, the jumps in $Q(t)$ are of random size. The first jump is of size $Y_1$, the second of size $Y_2$, etc. Figure 11.3.1 shows one path of a compound Poisson process.

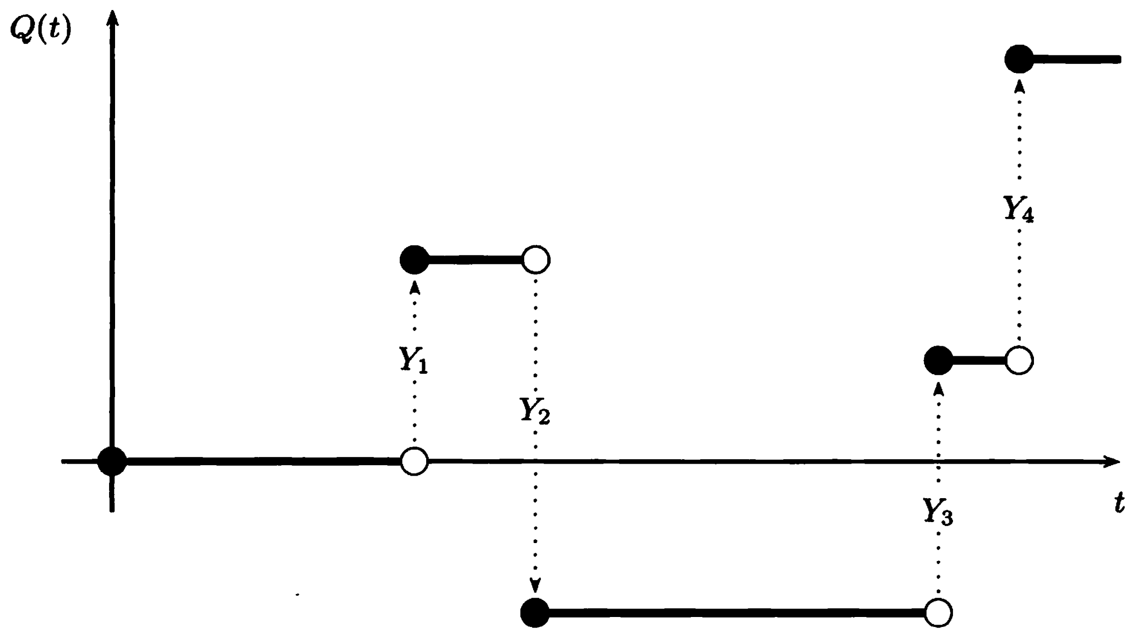


**Fig. 11.3.1.** One path of a compound Poisson process.

Like the simple Poisson process $N(t)$, the increments of the compound Poisson process $Q(t)$ are independent. In particular, for $0 \leq s < t$,

$$Q(s) = \sum_{i=1}^{N(s)} Y_i,$$

which sums up the first $N(s)$ jumps, and

$$Q(t) - Q(s) = \sum_{i=N(s)+1}^{N(t)} Y_i,$$

which sums up jumps $N(s)+1$ to $N(t)$, are independent. Moreover, $Q(t)-Q(s)$ has the same distribution as $Q(t-s)$ because $N(t) - N(s)$ has the same distribution as $N(t-s)$.

The mean of the compound Poisson process is

$$\begin{aligned}\mathbb{E}Q(t) &= \sum_{k=0}^{\infty} \mathbb{E}\Big[\sum_{i=1}^{k} Y_i \Big| N(t) = k\Big] \mathbb{P}\{N(t) = k\} \\ &= \sum_{k=0}^{\infty} \beta k \frac{(\lambda t)^k}{k!} e^{-\lambda t} \;=\; \beta \lambda t e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^{k-1}}{(k-1)!} \;=\; \beta \lambda t.\end{aligned}$$

On average, there are $\lambda t$ jumps in the time interval $[0, t]$, the average jump size is $\beta$, and the number of jumps is independent of the size of the jumps. Hence, $\mathbb{E}Q(t)$ is the product $\beta \lambda t$.

**Theorem 11.3.1.** *Let $Q(t)$ be the compound Poisson process defined above. Then the* compensated compound Poisson process

$$Q(t) - \beta\lambda t$$

*is a martingale.*

PROOF: Let $0 \leq s < t$ be given. Because the increment $Q(t) - Q(s)$ is independent of $\mathcal{F}(s)$ and has mean $\beta\lambda(t-s)$, we have

$$\begin{aligned}\mathbb{E}\big[Q(t) - \beta\lambda t \big| \mathcal{F}(s)\big] &= \mathbb{E}\big[Q(t) - Q(s) \big| \mathcal{F}(s)\big] + Q(s) - \beta\lambda t \\ &= \beta\lambda(t-s) + Q(s) - \beta\lambda t \\ &= Q(s) - \beta\lambda s. \end{aligned}$$

□

Just like a Poisson process, a compound Poisson process has stationary independent increments. We give the precise statement below.

**Theorem 11.3.2.** *Let $Q(t)$ be a compound Poisson process and let $0 = t_0 < t_1 < \cdots < t_n$ be given. The increments*

$$Q(t_1) - Q(t_0), Q(t_2) - Q(t_1), \ldots, Q(t_n) - Q(t_{n-1}),$$

*are independent and stationary. In particular, the distribution of $Q(t_j) - Q(t_{j-1})$ is the same as the distribution of $Q(t_j - t_{j-1})$.*

### 11.3.2 Moment-Generating Function

In Theorem 11.3.2, we did not write an explicit formula for the distribution of $Q(t_j - t_{j-1})$ because the formula for the density or probability mass function of this random variable is quite complicated. However, the formula for its moment-generating function is simple. For this reason, we use moment generating functions rather than densities or probability mass functions in much of what follows.

Let $Q(t)$ be the compound Poisson process defined by (11.3.1). Denote the moment-generating function of the random variable $Y_i$ by

$$\varphi_Y(u) = \mathbb{E}e^{uY_i}.$$

This does not depend on the index $i$ because $Y_1, Y_2, \ldots$ all have the same distribution. The moment generating function for the compound Poisson process $Q(t)$ is

$$\begin{aligned}\varphi_{Q(t)}(u) &= \mathbb{E}e^{uQ(t)} \\ &= \mathbb{E}\exp\left\{u\sum_{i=1}^{N(t)} Y_i\right\} \\ &= \mathbb{P}\{N(t)=0\} + \sum_{k=1}^{\infty}\mathbb{E}\exp\left\{u\sum_{i=1}^{k} Y_i \Big| N(t)=k\right\}\mathbb{P}\{N(t)=k\}\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{P}\{N(t)=0\} + \sum_{k=1}^{\infty} \mathbb{E}\exp\Big\{u\sum_{i=1}^{k} Y_i\Big\}\mathbb{P}\{N(t)=k\} \\
&= e^{-\lambda t} + \sum_{k=1}^{\infty} \mathbb{E}e^{uY_1}\mathbb{E}e^{uY_2}\cdots\mathbb{E}e^{uY_k}\frac{(\lambda t)^k}{k!}e^{-\lambda t} \\
&= e^{-\lambda t} + e^{-\lambda t}\sum_{k=1}^{\infty}\frac{\big(\varphi_Y(u)\lambda t\big)^k}{k!} \\
&= e^{-\lambda t}\sum_{k=0}^{\infty}\frac{\big(\varphi_Y(u)\lambda t\big)^k}{k!} \\
&= \exp\big\{\lambda t\big(\varphi_Y(u)-1\big)\big\}. \qquad (11.3.2)
\end{aligned}
$$

If the random variables $Y_i$ are not really random but rather always take the constant value $y$, then the compound Poisson process $Q(t)$ is actually $yN(t)$ and $\varphi_Y(u) = e^{uy}$. It follows that $y$ times a Poisson process has the moment-generating function

$$
\varphi_{yN(t)}(u) = \mathbb{E}e^{uyN(t)} = \exp\{\lambda t(e^{uy}-1)\}. \qquad (11.3.3)
$$

When $y=1$, we have the Poisson process, whose moment-generating function is thus

$$
\varphi_{N(t)}(u) = \mathbb{E}e^{uN(t)} = \exp\{\lambda t(e^{u}-1)\}. \qquad (11.3.4)
$$

Finally, consider the case when $Y_i$ takes one of finitely many possible nonzero values $y_1, y_2, \dots, y_M$, with $p(y_m) = \mathbb{P}\{Y_i = y_m\}$ so that $p(y_m) > 0$ for every $m$ and $\sum_{m=1}^{M} p(y_m) = 1$. Then $\varphi_Y(u) = \sum_{m=1}^{M} p(y_m)e^{uy_m}$. It follows from (11.3.2) that

$$
\begin{aligned}
\varphi_{Q(t)}(u) &= \exp\Big\{\lambda t\Big(\sum_{m=1}^{M} p(y_m)e^{uy_m} - 1\Big)\Big\} \\
&= \exp\Big\{\lambda t\sum_{m=1}^{M} p(y_m)\big(e^{uy_m}-1\big)\Big\} \\
&= \exp\{\lambda p(y_1)t(e^{uy_1}-1)\}\exp\{\lambda p(y_2)t(e^{uy_2}-1)\}\cdots \\
&\qquad \cdots\exp\{\lambda p(y_M)t(e^{uy_M}-1)\}. \qquad (11.3.5)
\end{aligned}
$$

This last expression is the product of the moment generating-functions for $M$ scaled Poisson processes, the $m$th process having intensity $\lambda p(y_m)$ and jump size $y_m$ (see (11.3.3)). This observation leads to the following theorem.

**Theorem 11.3.3 (Decomposition of a compound Poisson process).** *Let $y_1, y_2, \dots, y_M$ be a finite set of nonzero numbers, and let $p(y_1), p(y_2), \dots, p(y_M)$ be positive numbers that sum to 1. Let $\lambda > 0$ be given, and let $\overline{N}_1(t), \overline{N}_2(t), \dots, \overline{N}_M(t)$ be independent Poisson processes, each $\overline{N}_m(t)$ having intensity $\lambda p(y_m)$. Define*

$$\overline{Q}(t) = \sum_{m=1}^{M} y_m \overline{N}_m(t), \quad t \geq 0. \tag{11.3.6}$$

*Then $\overline{Q}(t)$ is a compound Poisson process. In particular, if $\overline{Y}_1$ is the size of the first jump of $\overline{Q}(t)$, $\overline{Y}_2$ is the size of the second jump, etc., and*

$$\overline{N}(t) = \sum_{m=1}^{M} \overline{N}_m(t), \quad t \geq 0,$$

*is the total number of jumps on the time interval $(0,t]$, then $\overline{N}(t)$ is a Poisson process with intensity $\lambda$, the random variables $\overline{Y}_1, \overline{Y}_2, \ldots$ are independent with $\mathbb{P}\{\overline{Y}_i = y_m\} = p(y_m)$ for $m = 1, \ldots, M$, the random variables $\overline{Y}_1, \ldots, \overline{Y}_M$ are independent of $\overline{N}(t)$, and*

$$\overline{Q}(t) = \sum_{i=0}^{\overline{N}(t)} \overline{Y}_i, \quad t \geq 0.$$

OUTLINE OF PROOF: According to (11.3.3), for each $m$, the characteristic function of $y_m \overline{N}_m(t)$ is

$$\varphi_{y_m \overline{N}_m(t)}(u) = \exp\{\lambda p(y_m) t (e^{uy_m} - 1)\}.$$

With $\overline{Q}(t)$ defined by (11.3.6), we use the fact that $\overline{N}_1(t), \overline{N}_2(t), \ldots, \overline{N}_M(t)$ are independent of one another to write

$$\begin{aligned}
\varphi_{\overline{Q}(t)}(u) &= \mathbb{E} \exp\left\{u \sum_{m=1}^{M} y_m \overline{N}_m(t)\right\} \\
&= \mathbb{E} e^{uy_1 \overline{N}_1(t)} \mathbb{E} e^{uy_2 \overline{N}_2(t)} \ldots \mathbb{E} e^{uy_M \overline{N}_M(t)} \\
&= \varphi_{y_1 \overline{N}_1(t)}(u) \varphi_{y_2 \overline{N}_2(t)}(u) \cdots \varphi_{y_M \overline{N}_M(t)}(u) \\
&= \exp\{\lambda p(y_1) t (e^{uy_1} - 1)\} \exp\{\lambda p(y_2) t (e^{uy_2} - 1)\} \cdots \\
&\qquad \cdots \exp\{\lambda p(y_M) t (e^{uy_M} - 1)\},
\end{aligned}$$

which is the right-hand side of (11.3.5). It follows that the random variable $\overline{Q}(t)$ of (11.3.6) has the same distribution as the random variable $Q(t)$ appearing on the left-hand side of (11.3.5). With a bit more work, one can show that the distribution of the whole path of $\overline{Q}$ defined by (11.3.6) agrees with the distribution of the whole path of the process $Q$ appearing on the left-hand side of (11.3.5).

Recall that the process $Q$ appearing on the left-hand side of (11.3.5) is the compound Poisson process defined by (11.3.1). For this process $N(t)$, the total number of jumps by time $t$ is Poisson with intensity $\lambda$, and the sizes of the jumps, $Y_1, Y_2, \ldots$, are identically distributed random variables, independent of one another and independent of $N(t)$, and with $\mathbb{P}\{Y_i = y_m\} = p(y_m)$ for

$m = 1, \ldots, M$. Because the processes $Q$ and $\overline{Q}$ have the same distribution, these statements must also be true for the total number of jumps and the sizes of the jumps of the process $\overline{Q}$ of (11.3.6), which is what the theorem asserts. □

The substance of Theorem 11.3.3 is that there are two equivalent ways of regarding a compound Poisson process that has only finitely many possible jump sizes. It can be thought of as a single Poisson process in which the size-one jumps are replaced by jumps of random size. Alternatively, it can be regarded as a sum of independent Poisson processes in each of which the size-one jumps are replaced by jumps of a fixed size. We restate Theorem 11.3.3 in a way designed to make this more clear.

**Corollary 11.3.4.** *Let $y_1, \ldots, y_M$ be a finite set of nonzero numbers, and let $p(y_1), \ldots, p(y_M)$ be positive numbers that sum to 1. Let $Y_1, Y_2, \ldots$ be a sequence of independent, identically distributed random variables with $\mathbb{P}\{Y_i = y_m\} = p(y_m)$, $m = 1, \ldots, M$. Let $N(t)$ be a Poisson process and define the compound Poisson process*

$$Q(t) = \sum_{i=1}^{N(t)} Y_i.$$

*For $m = 1, \ldots, M$, let $N_m(t)$ denote the number of jumps in $Q$ of size $y_m$ up to and including time $t$. Then*

$$N(t) = \sum_{m=1}^{M} N_m(t) \text{ and } Q(t) = \sum_{m=1}^{M} y_m N_m(t).$$

*The processes $N_1, \ldots, N_M$ defined this way are independent Poisson processes, and each $N_m$ has intensity $\lambda p(y_m)$.*

## 11.4 Jump Processes and Their Integrals

In this section, we introduce the stochastic integral when the integrator is a process with jumps, and we develop properties of this integral. We shall have a Brownian motion and Poisson and compound Poisson processes. There will always be a single filtratoin associated with all of them, in the sense of the following definition.

**Definition 11.4.1.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathcal{F}(t)$, $t \geq 0$, be a filtration on this space. We say that a Brownian motion $W$ is a* Brownian motion relative to this filtration *if $W(t)$ is $\mathcal{F}(t)$-measurable for every $t$ and for every $u > t$ the increment $W(u) - W(t)$ is independent of $\mathcal{F}(t)$. Similarly, we say that a Poisson process $N$ is a* Poisson process relative to this filtration *if $N(t)$ is $\mathcal{F}(t)$-measurable for every $t$ and for every $u > t$ the increment*

*$N(u) - N(t)$ is independent of $\mathcal{F}(t)$. Finally, we say that a compound Poisson process $Q$ is a* compound Poisson process relaative to this filtration *if $Q(t)$ is $\mathcal{F}(t)$-measurable for every $t$ and for every $u > t$ the increment $Q(u) - Q(t)$ is independent of $\mathcal{F}(t)$.*

### 11.4.1 Jump Processes

We wish to define the stochastic integral

$$\int_0^t \Phi(s)\, dX(s),$$

where the integrator $X$ can have jumps. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which is given a filtration $\mathcal{F}(t)$, $t \geq 0$. All processes will be adapted to this filtration. Furthermore, the integrators we consider in this section will be right-continuous and of the form

$$X(t) = X(0) + I(t) + R(t) + J(t). \tag{11.4.1}$$

In (11.4.1), $X(0)$ is a nonrandom initial condition. The process

$$I(t) = \int_0^t \Gamma(s)\, dW(s) \tag{11.4.2}$$

is an *Itô integral* of an adapted process $\Gamma(s)$ with respect to a Brownian motion relative to the filtration. We shall call $I(t)$ the *Itô integral part of $X$*. The process $R(t)$ in (11.4.1) is a *Riemann integral*[1]

$$R(t) = \int_0^t \Theta(s)\, ds \tag{11.4.3}$$

for some adapted process $\Theta(t)$. We shall call $R(t)$ the *Riemann integral part of $X$*. The *continuous part* of $X(t)$ is defined to be

$$X^c(t) = X(0) + I(t) + R(t) = X(0) + \int_0^t \Gamma(s)\, dW(s) + \int_0^t \Theta(s)\, ds.$$

The quadratic variation of this process is

$$[X^c, X^c](t) = \int_0^t \Gamma^2(s)\, ds,$$

an equation that we write in differential form as

[1] One usually takes this to be a Lebesgue integral with respect to $dt$, but for all the cases we consider, the Riemann integral is defined and agrees with the Lebesgue integral.

$$dX^c(t)\,dX^c(t) = \Gamma^2(t)\,dt.$$

Finally, in (11.4.1), $J(t)$ is an adapted, right-continuous *pure jump process* with $J(0) = 0$. By *right-continuous*, we mean that $J(t) = \lim_{s\downarrow t} J(s)$ for all $t \geq 0$. The *left-continuous* version of such a process will be denoted $J(t-)$. In other words, if $J$ has a jump at time $t$, then $J(t)$ is the value of $J$ immediately after the jump, and $J(t-)$ is its value immediately before the jump. We assume that $J$ does not jump at time zero, has only finitely many jumps on each finite time interval $(0, T]$, and is constant between jumps. The constancy between jumps is what justifies calling $J(t)$ a *pure* jump process. A Poisson process and a compound Poisson process have this property. A compensated Poisson process does not because it decreases between jumps. We shall call $J(t)$ the *pure jump part of* $X$.

**Definition 11.4.2.** *A process $X(t)$ of the form (11.4.1), with Itô integral part $I(t)$, Riemann integral part $R(t)$, and pure jump part $J(t)$ as described above, will be called a* jump process. *The* continuous part *of this process is $X^c(t) = X(0) + I(t) + R(t)$.*

A jump process in this book is not the most general possible because we permit only finitely many jumps in finite time. For many applications, these processes are sufficient. Furthermore, the stochastic calculus for these processes gives a good indication of how the stochastic calculus works for the more general case.

A jump process $X(t)$ is right-continuous and adapted. Because both $I(t)$ and $R(t)$ are continuous, the left-continuous version of $X(t)$ is

$$X(t-) = X(0) + I(t) + R(t) + J(t-).$$

The jump size of $X$ at time $t$ is denoted

$$\Delta X(t) = X(t) - X(t-).$$

If $X$ is continuous at $t$, then $\Delta X(t) = 0$. If $X$ has a jump at time $t$, then $\Delta X(t)$ is the size of this jump, which is also $\Delta J(t) = J(t) - J(t-)$, the size of the jump in $J$. Whenever $X(0-)$ appears in the formulas below, we mean it to be $X(0)$. In particular, $\Delta X(0) = 0$; there is no jump at time zero.

**Definition 11.4.3.** *Let $X(t)$ be a jump process of the form (11.4.1)–(11.4.3) and let $\Phi(s)$ be an adapted process. The* stochastic integral *of $\Phi$ with respect to $X$ is defined to be*

$$\int_0^t \Phi(s)\,dX(s) = \int_0^t \Phi(s)\Gamma(s)\,dW(s) + \int_0^t \Phi(s)\Theta(s)\,ds + \sum_{0<s\leq t} \Phi(s)\Delta J(s). \tag{11.4.4}$$

*In differential notation,*

$$\begin{aligned}\Phi(t)dX(t) &= \Phi(t)\,dI(t) + \Phi(t)\,dR(t) + \Phi(t)\,dJ(t)\\ &= \Phi(t)\,dX^c(t) + \Phi(t)\,dJ(t),\end{aligned}$$

*where*

$$\Phi(t)\,dI(t) = \Phi(t)\Gamma(t)\,dW(t),\ \ \Phi(t)\,dR(t) = \Phi(t)\Theta(t)\,dt,$$
$$\Phi(t)\,dX^c(t) = \Phi(t)\Gamma(t)\,dW(t) + \Phi(t)\Theta(t)\,dt.$$

*Example 11.4.4.* Let $X(t) = M(t) = N(t) - \lambda t$, where $N(t)$ is a Poisson process with intensity $\lambda$ so that $M(t)$ is the compensated Poisson process of Theorem 11.2.4. In the terminology of Definition 11.4.2, $I(t) = 0$, $X^c(t) = R(t) = -\lambda t$, and $J(t) = N(t)$. Let $\Phi(s) = \Delta N(s)$ (i.e., $\Phi(s)$ is 1 if $N$ has a jump at time $s$, and $\Phi(s)$ is zero otherwise). For $s \in [0, t]$, $\Phi(s)$ is zero except for finitely many values of $s$, and thus

$$\int_0^t \Phi(s)\,dX^c(s) = \int_0^t \Phi(s)\,dR(s) = -\lambda \int_0^t \Phi(s)\,ds = 0.$$

However,

$$\int_0^t \Phi(s)\,dN(s) = \sum_{0<s\leq t} \left(\Delta N(s)\right)^2 = N(t).$$

Therefore,

$$\int_0^t \Phi(s)\,dM(s) = -\lambda \int_0^t \Phi(s)\,ds + \int_0^t \Phi(s)\,dN(s) = N(t). \qquad (11.4.5)$$

□

For Brownian motion $W(t)$, we defined the stochastic integral

$$I(t) = \int_0^t \Gamma(s)\,dW(s)$$

in a way that caused $I(t)$ to be a martingale. To define the stochastic integral, we approximated the integrand $\Gamma(s)$ by simple integrands $\Gamma_n(s)$, wrote down a formula for

$$I_n(t) = \int_0^t \Gamma_n(s)\,dW(s),$$

and verified that, for each $n$, $I_n(t)$ is a martingale. We defined $I(t)$ as the limit of $I_n(t)$ as $n \to \infty$ and, because it is the limit of martingales, $I(t)$ also is a martingale. The only conditions we needed on $\Gamma(s)$ for this construction were that it be adapted and that it satisfy the technical condition $\mathbb{E}\int_0^t \Gamma^2(s)\,ds < \infty$ for every $t > 0$.

This construction makes sense for finance because we ultimately replace $\Gamma(s)$ by a position in an asset and replace $W(s)$ by the price of that asset. If the asset price is a martingale (i.e., it is pure volatility with no underlying trend), then the gain we make from investing in the asset should also be a martingale. The stochastic integral is this gain.

In the context of processes that can jump, we still want the stochastic integral with respect to a martingale to be a martingale. However, we see in Example 11.4.4 that this is not always the case. The integrator $M(t)$ in that example is a martingale (see Theorem 11.2.4), but the integral $N(t)$ in (11.4.5) is not because it goes up but cannot go down.

An agent who invests in the compensated Poisson process $M(t)$ by choosing his position according to the formula $\Phi(s) = \Delta N(s)$ has created an arbitrage. To do this, he is holding a zero position at all times except the jump times of $N(s)$, which are also the jump times of $M(s)$, at which times he holds a position one. Because the jumps in $M(s)$ are always up and our investor holds a long position at exactly the jump times, he will reap the upside gain from all these jumps and have no possibility of loss.

In reality, the portfolio process $\Phi(s) = \Delta N(s)$ cannot be implemented because investors must take positions before jumps occur. No one without insider information can arrange consistently to take a position exactly at the jump times. However, $\Phi(s)$ depends only on the path of the underlying process $M$ up to and including at time $s$ and does not depend on the future of the path. That is the definition of adapted we used when constructing stochastic integrals with respect to Brownian motion. Here we see that it is not enough to require the integrand to be adapted. A mathematically convenient way of formulating the extra condition is to insist that our integrands be *left-continuous*. That rules out $\Phi(s) = \Delta N(s)$. In the time interval between jumps, this process is zero, and a left-continuous process that is zero between jumps must also be zero at the jump times.

We give the following theorem without proof.

**Theorem 11.4.5.** *Assume that the jump process $X(s)$ of (11.4.1)–(11.4.3) is a martingale, the integrand $\Phi(s)$ is* left-continuous *and adapted, and*

$$\mathbb{E}\int_0^t \Gamma^2(s)\Phi^2(s)\,ds < \infty \text{ for all } t \geq 0.$$

*Then the stochastic integral $\int_0^t \Phi(s)\,dX(s)$ is also a martingale.*

The mathematical literature on integration with respect to jump processes gives a slightly more general version of Theorem 11.4.5 in which the integrand is required only to be *predictable*. Roughly speaking, such processes are those that can be gotten as the limit of left-continuous processes. We shall not need this more general concept.

Note that although we require the integrand $\Phi(s)$ to be left-continuous in Theorem 11.4.5, the integrator $X(t)$ is always taken to be right-continuous, and so the integral $\int_0^t \Phi(s)\,dX(s)$ will be right-continuous in the upper limit of integration $t$. The integral jumps whenever $X$ jumps and $\Phi$ is simultaneously not zero. The value of the integral at time $t$ includes the jump at time $t$ if there is a jump; see (11.4.4).

*Example 11.4.6.* Let $N(t)$ be a Poisson process with intensity $\lambda$, let $M(t) = N(t) - \lambda t$ be the compensated Poisson process, and let

$$\Phi(s) = \mathbb{I}_{[0,S_1]}(s)$$

be 1 up to and including the time of the first jump and zero thereafter. Note that $\Phi$ is left-continuous. We have

$$\begin{aligned}\int_0^t \Phi(s)\,dM(s) &= \begin{cases} -\lambda t, & 0 \le t < S_1, \\ 1 - \lambda S_1, & t \ge S_1 \end{cases} \\ &= \mathbb{I}_{[S_1,\infty)}(t) - \lambda(t \wedge S_1). \end{aligned} \tag{11.4.6}$$

The notation $t \wedge S_1$ in (11.4.6) denotes the minimum of $t$ and $S_1$. See Figure 11.4.1.

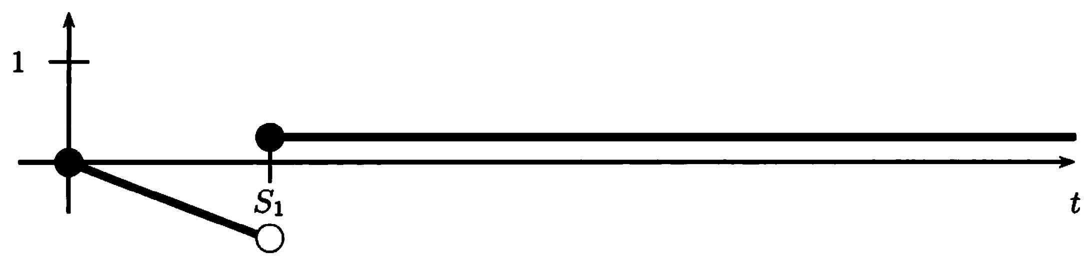


**Fig. 11.4.1.** $\mathbb{I}_{[S_1,\infty)}(t) - \lambda(t \wedge S_1)$.

We verify the martingale property for the process $\mathbb{I}_{[S_1,\infty)}(t) - \lambda(t \wedge S_1)$ by direct computation. For $0 \le s < t$, we have

$$\mathbb{E}\big[\mathbb{I}_{[S_1,\infty)}(t) - \lambda(t \wedge S_1)\big|\mathcal{F}(s)\big] = \mathbb{P}\{S_1 \le t|\mathcal{F}(s)\} - \lambda\mathbb{E}\big[t \wedge S_1\big|\mathcal{F}(s)\big]. \tag{11.4.7}$$

If $S_1 \le s$, then at time $s$ we know the value of $S_1$ and the conditional expectations above give us the random variables being estimated. In particular, the right-hand side of (11.4.7) is $1 - \lambda S_1 = \mathbb{I}_{[S_1,\infty)}(s) - \lambda(s \wedge S_1)$, and the martingale property is satisfied. On the other hand, if $S_1 > s$, then

$$\mathbb{P}\{S_1 \le t|\mathcal{F}(s)\} = 1 - \mathbb{P}\{S_1 > t|S_1 > s\} = 1 - e^{-\lambda(t-s)}, \tag{11.4.8}$$

where we have used the fact that $S_1$ is exponentially distributed and used the memorylessness (11.2.3) of exponential random variables. In fact, the memorylessness says that, conditioned on $S_1 > s$, the density of $S_1$ is

$$-\frac{\partial}{\partial u}\mathbb{P}\{S_1 > u|S_1 > s\} = -\frac{\partial}{\partial u}e^{-\lambda(u-s)} = \lambda e^{-\lambda(u-s)}, \quad u > s.$$

It follows that, when $S_1 > s$,

$$\begin{aligned}
\lambda\mathbb{E}\big[t\wedge S_1\big|\mathcal{F}(s)\big] &= \lambda\mathbb{E}\big[t\wedge S_1\big|S_1>s\big]\\
&= \lambda^2\int_s^\infty (t\wedge u)e^{-\lambda(u-s)}\,du\\
&= \lambda^2\int_s^t ue^{-\lambda(u-s)}du+\lambda^2\int_t^\infty te^{-\lambda(u-s)}du\\
&= -\lambda ue^{-\lambda(u-s)}\Big|_{u=s}^{u=t}+\lambda\int_s^t e^{-\lambda(u-s)}du-\lambda te^{-\lambda(u-s)}\Big|_{u=t}^{u=\infty}\\
&= \lambda s-\lambda te^{-\lambda(t-s)}-e^{-\lambda(u-s)}\Big|_{u=s}^{u=t}+\lambda te^{-\lambda(t-s)}\\
&= \lambda s-e^{-\lambda(t-s)}+1. \qquad (11.4.9)
\end{aligned}$$

Subtracting (11.4.9) from (11.4.8), we obtain in the case $S_1>s$ that

$$\mathbb{E}\big[\mathbb{I}_{[S_1,\infty)}(t)-\lambda(t\wedge S_1)\big|\mathcal{F}(s)\big]=-\lambda s=\mathbb{I}_{[S_1,\infty)}(s)-\lambda(s\wedge S_1).$$

This completes the verification of the martingale property for the stochastic integral in (11.4.6).

Note that if we had taken the integrand in (11.4.6) to be $\mathbb{I}_{[0,S_1)}(t)$, which is right-continuous rather than left-continuous at $S_1$, then we would have gotten

$$\int_0^t \mathbb{I}_{[0,S_1)}(u)dM(u)=-\lambda(t\wedge S_1). \qquad (11.4.10)$$

According to (11.4.9) with $s=0$,

$$\mathbb{E}\big[-\lambda(t\wedge S_1)\big]=e^{-\lambda t}-1,$$

which is strictly decreasing in $t$. Consequently, the integral (11.4.10) obtained from the right-continuous integrand $\mathbb{I}_{[0,S_1)}(t)$ is not a martingale. □

### 11.4.2 Quadratic Variation

In order to write down the Itô-Doeblin formula for processes with jumps, we need to discuss *quadratic variation.* Let $X(t)$ be a jump process. To compute the quadratic variation of $X$ on $[0,T]$, we choose $0=t_0<t_1<t_2<\cdots<t_n=T$, denote the set of these times by $\Pi=\{t_0,t_1,\ldots,t_n\}$, denote the length of the longest subinterval by $\|\Pi\|=\max_j(t_{j+1}-t_j)$, and define

$$Q_\Pi(X)=\sum_{j=0}^{n-1}\big(X(t_{j+1})-X(t_j)\big)^2.$$

The *quadratic variation* of $X$ on $[0,T]$ is defined to be

$$[X,X](T)=\lim_{\|\Pi\|\to 0}Q_\Pi(X),$$

where of course as $\|\Pi\| \to 0$ the number of points in $\Pi$ must approach infinity.

In general, $[X,X](T)$ can be random (i.e., can depend on the path of $X$). However, in the case of Brownian motion, we know that $[W,W](T) = T$ does not depend on the path. In the case of an Itô integral $I(T) = \int_0^T \Gamma(s)dW(s)$ with respect to Brownian motion, $[I,I](T) = \int_0^T \Gamma^2(s)ds$ can depend on the path because $\Gamma(s)$ can depend on the path.

We will also need the concept of *cross variation*. Let $X_1(t)$ and $X_2(t)$ be jump processes. We define

$$C_\Pi(X_1,X_2) = \sum_{j=0}^{n-1} \big(X_1(t_{j+1}) - X_1(t_j)\big)\big(X_2(t_{j+1}) - X_2(t_j)\big)$$

and

$$[X_1,X_2](T) = \lim_{\|\Pi\|\to 0} C_\Pi(X_1,X_2).$$

**Theorem 11.4.7.** *Let $X_1(t) = X_1(0) + I_1(t) + R_1(t) + J_1(t)$ be a jump process, where $I_1(t) = \int_0^t \Gamma_1(s)\,dW(s)$, $R_1(t) = \int_0^t \Theta_1(s)\,ds$, and $J_1(t)$ is a right-continuous pure jump process. Then $X_1^c(t) = X_1(0) + I_1(t) + R_1(t)$ and*

$$[X_1,X_1](T) = [X_1^c,X_1^c](T) + [J_1,J_1](T) = \int_0^T \Gamma_1^2(s)\,ds + \sum_{0<s\le T} \big(\Delta J_1(s)\big)^2. \tag{11.4.11}$$

*Let $X_2(t) = X_2(0) + I_2(t) + R_2(t) + J_2(t)$ be another jump process, where $I_2(t) = \int_0^t \Gamma_2(s)\,dW(s)$, $R_2(t) = \int_0^t \Theta_2(s)\,ds$, and $J_2(t)$ is a right-continuous pure jump process. Then $X_2^c(t) = X_2(0) + I_2(t) + R_2(t)$, and*

$$\begin{aligned}[X_1,X_2](T) &= [X_1^c,X_2^c](T) + [J_1,J_2](T)\\ &= \int_0^T \Gamma_1(s)\Gamma_2(s)\,ds + \sum_{0<s\le T} \Delta J_1(s)\Delta J_2(s). \end{aligned} \tag{11.4.12}$$

PROOF: We only need to prove (11.4.12) since (11.4.11) is the special case of (11.4.12) in which $X_2 = X_1$. We have

$$\begin{aligned}C_\Pi(X_1,X_2) &= \sum_{j=0}^{n-1} \big(X_1(t_{j+1}) - X_1(t_j)\big)\big(X_2(t_{j+1}) - X_2(t_j)\big)\\ &= \sum_{j=0}^{n-1} \big(X_1^c(t_{j+1}) - X^c(t_j) + J_1(t_{j+1}) - J_1(t_j)\big)\\ &\qquad\qquad \times\big(X_2^c(t_{j+1}) - X_2^c(t_j) + J_2(t_{j+1}) - J_2(t_j)\big)\end{aligned}$$

$$= \sum_{j=0}^{n-1} \big(X_1^c(t_{j+1}) - X_1^c(t_j)\big)\big(X_2^c(t_{j+1}) - X_2^c(t_j)\big)$$
$$+ \sum_{j=0}^{n-1} \big(X_1^c(t_{j+1}) - X_1^c(t_j)\big)\big(J_2(t_{j+1}) - J_2(t_j)\big)$$
$$+ \sum_{j=0}^{n-1} \big(J_1(t_{j+1}) - J_1(t_j)\big)\big(X_2^c(t_{j+1}) - X_2^c(t_j)\big)$$
$$+ \sum_{j=0}^{n-1} \big(J_1(t_{j+1}) - J_1(t_j)\big)\big(J_2(t_{j+1}) - J_2(t_j)\big). \quad (11.4.13)$$

We know from the theory of continuous processes that

$$\lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \big(X_1^c(t_{j+1}) - X_1^c(t_j)\big)\big(X_2^c(t_{j+1}) - X_2^c(t_j)\big) = \big[X_1^c, X_2^c\big](T)$$
$$= \int_0^T \Gamma_1(s)\Gamma_2(s)\,ds.$$

We shall show that the second and third terms appearing on the right-hand side of (11.4.13) have limit zero as $\|\Pi\| \to 0$, and the fourth term has limit

$$[J_1, J_2](T) = \sum_{0<s\le T} \Delta J_1(s)\Delta J_2(s).$$

We consider the second term on the right-hand side of (11.4.13):

$$\left|\sum_{j=0}^{n-1} \big(X_1^c(t_{j+1}) - X_1^c(t_j)\big)\big(J_2(t_{j+1}) - J_2(t_j)\big)\right|$$
$$\le \max_{0\le j\le n-1} \big|X_1^c(t_{j+1}) - X_1^c(t_j)\big| \cdot \sum_{j=0}^{n-1} \big|J_2(t_{j+1}) - J_2(t_j)\big|$$
$$\le \max_{0\le j\le n-1} \big|X_1^c(t_{j+1}) - X_1^c(t_j)\big| \cdot \sum_{0<s\le T} |\Delta J_2(s)|.$$

As $\|\Pi\| \to 0$, the factor $\max_{0\le j\le n-1} \big|X_1^c(t_{j+1}) - X_1^c(t_j)\big|$ has limit zero, whereas $\sum_{0<s\le T} |\Delta J_2(s)|$ is a finite number not depending on $\Pi$. Hence, the second term on the right-hand side of (11.4.13) has limit zero as $\|\Pi\| \to 0$. Similarly, the third term on the right-hand side of (11.4.13) has limit zero.

Let us fix an arbitrary $\omega \in \Omega$, which fixes the paths of these processes, and choose the time points in $\Pi$ so close together that there is at most one jump of $J_1$ in each interval $(t_j, t_{j+1}]$, at most one jump of $J_2$ in each interval $(t_j, t_{j+1}]$, and if $J_1$ and $J_2$ have a jump in the same interval, then these jumps are simultaneous. Let $A_1$ denote the set of indices $j$ for which $(t_j, t_{j+1}]$ contains a jump of $J_1$, and let $A_2$ denote the set of indices $j$ for which $(t_j, t_{j+1}]$ contains a jump of $J_2$. The fourth term on the right-hand side of (11.4.13) is

$$\begin{aligned}\sum_{j=0}^{n-1} &\big(J_1(t_{j+1}) - J_1(t_j)\big)\big(J_2(t_{j+1}) - J_2(t_j)\big)\\ &= \sum_{j\in A_1\cap A_2} \big(J_1(t_{j+1}) - J_1(t_j)\big)\big(J_2(t_{j+1}) - J_2(t_j)\big)\\ &= \sum_{0<s\le t} \Delta J_1(s)\Delta J_2(s).\end{aligned}$$

This completes the proof. □

*Remark 11.4.8.* In differential notation, equation (11.4.12) of Theorem 11.4.7 says that if

$$X_1(t) = X_1(0) + X_1^c(t) + J_1(t), \quad X_2(t) = X_2(0) + X_2^c(t) + J_2(t),$$

then

$$dX_1(t)\,dX_2(t) = dX_1^c(t)\,dX_2^c(t) + dJ_1(t)\,dJ_2(t).$$

In particular,

$$dX_1^c(t)\,dJ_2(t) = dX_2^c(t)\,dJ_1(t) = 0;$$

the cross variation between a continuous process and a pure jump process is zero. It follows that the cross variation between a Brownian motion and a Poisson process is zero.

More generally, the cross variation between two processes is zero if one of them is continuous and the other has no Itô integral part. In order to get a nonzero cross variation, both processes must have a $dW$ term or the processes must have simultaneous jumps. This means that the cross variation between a Brownian motion and a compensated Poisson process is also zero. We state this last fact as a corollary. □

**Corollary 11.4.9.** *Let $W(t)$ be a Brownian motion and $M(t) = N(t) - \lambda t$ be a compensated Poisson process relative to the same filtration $\mathcal{F}(t)$ (Definition 11.4.1). Then*

$$[W, M](t) = 0, \quad t \ge 0.$$

PROOF: In Theorem 11.4.7, take $I_1(t) = W(t)$, $R_1(t) = J_1(t) = 0$ and take $I_2(t) = 0$, $R_2(t) = -\lambda t$, and $J_2(t) = N(t)$. □

We shall see in Corollary 11.5.3 that the equation $[W, M](t) = 0$ implies that $W$ and $M$ are independent, and hence $W$ and $N$ are independent. *A Brownian motion and a Poisson process relative to the same filtration must be independent.*

**Corollary 11.4.10.** *For $i = 1, 2$, let $X_i(t)$ be an adapted, right-continuous jump process. In other words, $X_i(t) = X_i(0) + I_i(t) + R_i(t) + J_i(t)$, where $I_i(t) = \int_0^t \Gamma_i(s)\,dW(s)$, $R_i(t) = \int_0^t \Theta_i(s)\,ds$, and $J_i(t)$ is a pure jump process. Let $\widetilde{X}_i(0)$ be a constant, let $\Phi_i(s)$ be an adapted process, and set*

$$\widetilde{X}_i(t) = \widetilde{X}_i(0) + \int_0^t \Phi_i(s)\,dX_i(s).$$

*By definition,*

$$\widetilde{X}_i(t) = \widetilde{X}_i(0) + \widetilde{I}_i(t) + \widetilde{R}_i(t) + \widetilde{J}_i(t),$$

*where*

$$\widetilde{I}_i(t) = \textstyle\int_0^t \Phi_i(s)\Gamma_i(s)\,dW(s), \quad \widetilde{R}_i(t) = \int_0^t \Phi_i(s)\Theta_i(s)\,ds,$$
$$\widetilde{J}_i(t) = \textstyle\sum_{0<s\le t} \Phi_i(s)\Delta J_i(s).$$

*Note that $\widetilde{X}_i(t)$ is a jump process with continuous part $\widetilde{X}_i^c(t) = \widetilde{X}_i(0) + \widetilde{I}_i(t) + \widetilde{R}_i(t)$ and pure jump part $\widetilde{J}_i(t)$. We have*

$$\begin{aligned}
&[\widetilde{X}_1, \widetilde{X}_2](t)\\
&= [\widetilde{X}_1^c, \widetilde{X}_2^c](t) + [\widetilde{J}_1, \widetilde{J}_2](t)\\
&= \int_0^t \Phi_1(s)\Phi_2(s)\Gamma_1(s)\Gamma_2(s)\,ds + \sum_{0<s\le t} \Phi_1(s)\Phi_2(s)\Delta J_1(s)\Delta J_2(s)\\
&= \int_0^t \Phi_1(s)\Phi_2(s)\,d[X_1, X_2](s).
\end{aligned}$$

*Remark 11.4.11.* Corollary 11.4.10 may be rewritten using differential notation. The corollary says that if

$$d\widetilde{X}_1(t) = \Phi_1(t)\,dX_1(t) \text{ and } d\widetilde{X}_2(t) = \Phi_2(t)\,dX_2(t),$$

then

$$d\widetilde{X}_1(t)\,d\widetilde{X}_2(t) = \Phi_1(t)\Phi_2(t)\,dX_1(t)\,dX_2(t).$$

## 11.5 Stochastic Calculus for Jump Processes

### 11.5.1 Itô-Doeblin Formula for One Jump Process

For a continuous-path process, the Itô-Doeblin formula is the following. Let

$$X^c(t) = X^c(0) + \int_0^t \Gamma(s)\,dW(s) + \int_0^t \Theta(s)\,ds, \tag{11.5.1}$$

where $\Gamma(s)$ and $\Theta(s)$ are adapted processes. In differential notation, we write

$$dX^c(s) = \Gamma(s)\,dW(s) + \Theta(s)\,ds, \quad dX^c(s)\,dX^c(s) = \Gamma^2(s)\,ds.$$

Let $f(x)$ be a function whose first and second derivatives are defined and continuous. Then

$$\begin{aligned} df\big(X^c(s)\big) &= f'\big(X^c(s)\big)\,dX^c(s) + \frac{1}{2}f''\big(X^c(s)\big)\,dX^c(s)\,dX^c(s) \\ &= f'\big(X^c(s)\big)\Gamma(s)\,dW(s) + f'\big(X^c(s)\big)\Theta(s)\,ds \\ &\quad + \frac{1}{2}f''\big(X^c(s)\big)\Gamma^2(s)\,ds. \end{aligned} \tag{11.5.2}$$

We write this in integral form as

$$\begin{aligned} f\big(X^c(t)\big) = f\big(X^c(0)\big) &+ \int_0^t f'\big(X^c(s)\big)\Gamma(s)\,dW(s) + \int_0^t f'\big(X^c(s)\big)\Theta(s)\,ds \\ &+ \frac{1}{2}\int_0^t f''\big(X^c(s)\big)\Gamma^2(s)\,ds. \end{aligned}$$

We now add a right-continuous pure jump term $J$ into (11.5.1), setting

$$X(t) = X(0) + I(t) + R(t) + J(t),$$

where $I(t) = \int_0^t \Gamma(s)\,dW(s)$ and $R(t) = \int_0^t \Theta(s)\,ds$. As usual, we denote by $X^c(t) = X(0) + I(t) + R(t)$ the continuous part of $X(t)$. Between jumps of $J$, the analogue of (11.5.2) holds:

$$\begin{aligned} df\big(X(s)\big) &= f'\big(X(s)\big)\,dX(s) + \frac{1}{2}f''\big(X(s)\big)\,dX(s)\,dX(s) \\ &= f'\big(X(s)\big)\Gamma(s)\,dW(s) + f'\big(X(s)\big)\Theta(s)\,ds \\ &\quad + \frac{1}{2}f''\big(X(s)\big)\Gamma^2(s)\,ds \\ &= f'\big(X(s)\big)\,dX^c(s) + \frac{1}{2}f''\big(X(s)\big)\,dX^c(s)\,dX^c(s). \end{aligned} \tag{11.5.3}$$

When there is a jump in $X$ from $X(s-)$ to $X(s)$, there is typically also a jump in $f(X)$ from $f\big(X(s-)\big)$ to $f\big(X(s)\big)$. When we integrate both sides of (11.5.3) from 0 to $t$, we must add in all the jumps that occur between these two times. This leads to the following theorem.

**Theorem 11.5.1 (Itô-Doeblin formula for one jump process).** *Let $X(t)$ be a jump process and $f(x)$ a function for which $f'(x)$ and $f''(x)$ are defined and continuous. Then*

$$\begin{aligned} f\big(X(t)\big) = f\big(X(0)\big) &+ \int_0^t f'\big(X(s)\big)\,dX^c(s) + \frac{1}{2}\int_0^t f''\big(X(s)\big)\,dX^c(s)\,dX^c(s) \\ &+ \sum_{0<s\le t}\big[f\big(X(s)\big) - f\big(X(s-)\big)\big]. \end{aligned} \tag{11.5.4}$$

PROOF: Fix $\omega \in \Omega$, which fixes the path of $X$, and let $0 < \tau_1 < \tau_2 < \cdots < \tau_{n-1} < t$ be the jump times in $[0, t)$ of this path of the process $X$. We set $\tau_0 = 0$, which is not a jump time, and $\tau_n = t$, which may or may not be a jump time. Whenever $u < v$ are both in the same interval $(\tau_j, \tau_{j+1})$, there is no jump between times $u$ and $v$, and the Itô-Doeblin formula (11.5.3) for continuous processes applies. We thus have

$$f(X(v)) - f(X(u)) = \int_u^v f'(X(s))\,dX^c(s) + \frac{1}{2}\int_u^v f''(X(s))\,dX^c(s)\,dX^c(s).$$

Letting $u \downarrow \tau_j$ and $v \uparrow \tau_{j+1}$ and using the right-continuity of $X$, we conclude that

$$\begin{aligned} &f(X(\tau_{j+1}-)) - f(X(\tau_j)) \\ &= \int_{\tau_j}^{\tau_{j+1}} f'(X(s))\,dX^c(s) + \frac{1}{2}\int_{\tau_j}^{\tau_{j+1}} f''(X(s))\,dX^c(s)\,dX^c(s). \end{aligned} \tag{11.5.5}$$

(Note here that

$$\lim_{v \uparrow \tau_{j+1}} \int_u^v f'(X(s))\,dX^c(s) = \int_v^{\tau_{j+1}} f'(X(s))\,dX^c(s),$$

but this is not the case if we replace $dX^c(s)$ by $dX(s)$ in this equation. If we made this replacement, the jump in $X$ at time $\tau_{j+1}$ would appear on the right-hand side of the equation but not on the left-hand side. It is for this reason that we integrate with respect to $dX^c(s)$ in (11.5.5).) We now add the jump in $f(X)$ at time $\tau_{j+1}$ into (11.5.5), obtaining thereby

$$\begin{aligned} &f(X(\tau_{j+1})) - f(X(\tau_j)) \\ &= \int_{\tau_j}^{\tau_{j+1}} f'(X(s))\,dX^c(s) + \frac{1}{2}\int_{\tau_j}^{t_{j+1}} f''(X(s))\,dX^c(s)\,dX^c(s) \\ &\qquad + f(X(\tau_{j+1})) - f(X(\tau_{j+1}-)). \end{aligned}$$

Summing over $j = 0, \ldots, n-1$, we obtain

$$\begin{aligned} &f(X(t)) - f(X(0)) \\ &= \sum_{j=0}^{n-1} \left[f(X(\tau_{j+1})) - f(X(\tau_j))\right] \\ &= \int_0^t f'(X(s))\,dX^c(s) + \frac{1}{2}\int_0^t f''(X(s))\,dX^c(s)\,dX^c(s) \\ &\qquad + \sum_{j=0}^{n-1} \left[f(X(\tau_{j+1})) - f(X(\tau_{j+1}-))\right], \end{aligned}$$

which is (11.5.4). Note in this connection that if there is no jump at $\tau_n = t$, then the last term in the sum on the right-hand side, $f\big(X(\tau_n)\big) - f\big(X(\tau_n-)\big)$, is zero. □

It is not always possible to rewrite (11.5.4) in differential form because it is not always possible to find a differential form for the sum of jumps. We provide one case in which this can be done in the next example.

*Example 11.5.2 (Geometric Poisson process).* Consider the geometric Poisson process

$$S(t) = S(0)\exp\big\{N(t)\log(\sigma+1) - \lambda\sigma t\big\} = S(0)e^{-\lambda\sigma t}(\sigma+1)^{N(t)}, \quad (11.5.6)$$

where $\sigma > -1$ is a constant. If $\sigma > 0$, this process jumps up and moves down between jumps; if $-1 < \sigma < 0$, it jumps down and moves up between jumps. We show that the process is a martingale.

We may write $S(t) = S(0)f\big(X(t)\big)$, where $f(x) = e^x$ and

$$X(t) = N(t)\log(\sigma+1) - \lambda\sigma t$$

has continuous part $X^c(t) = -\lambda\sigma t$ and pure jump part $J(t) = N(t)\log(\sigma+1)$. According to the Itô-Doeblin formula for jump processes,

$$\begin{aligned} S(t) &= f\big(X(t)\big) \\ &= f\big(X(0)\big) - \lambda\sigma\int_0^t f'\big(X(u)\big)\,du + \sum_{0<u\le t}\big[f\big(X(u)\big) - f\big(X(u-)\big)\big] \\ &= S(0) - \lambda\sigma\int_0^t S(u)\,du + \sum_{0<u\le t}\big[S(u) - S(u-)\big]. \end{aligned} \quad (11.5.7)$$

If there is a jump at time $u$, then $S(u) = (\sigma+1)S(u-)$. Therefore,

$$S(u) - S(u-) = \sigma S(u-) \quad (11.5.8)$$

whenever there is a jump at time $u$, and of course $S(u) - S(u-) = 0$ if there is no jump at time $u$. In either case, we have

$$S(u) - S(u-) = \sigma S(u-)\Delta N(u).$$

This observation permits us to rewrite the sum on the right-hand side of (11.5.7) as

$$\sum_{0<u\le t}\big[S(u) - S(u-)\big] = \sum_{0<u\le t}\sigma S(u-)\Delta N(u) = \sigma\int_0^t S(u-)\,dN(u).$$

It does not matter whether we write the Riemann integral on the right-hand side of (11.5.7) as $\int_0^t S(u)\,du$ or as $\int_0^t S(u-)\,du$. The integrands in these two integrals differ at only finitely many times, and when we integrate with respect to $du$, these differences do not matter. Therefore, we may rewrite (11.5.7) as

$$\begin{aligned} S(t) &= S(0) - \lambda\sigma \int_0^t S(u-)\,du + \sigma \int_0^t S(u-)\,dN(u) \\ &= S(0) + \sigma \int_0^t S(u-)\,dM(u), \end{aligned}$$

where $M$ is the compensated Poisson process $M(u) = N(u) - \lambda u$, which is a martingale. Because the integrand $S(u-)$ is left-continuous, Theorem 11.4.5 guarantees that $S(t)$ is a martingale.

In this case, the Itô-Doeblin formula (11.5.7) has a differential form, namely,

$$dS(t) = \sigma S(t-)\,dM(t) = -\lambda\sigma S(t)\,dt + \sigma S(t-)\,dN(t). \tag{11.5.9}$$

We were able to obtain this differential form because in (11.5.8) we were able to write the jump in $f(X)$ (i.e., the jump in $S$) at time $u$ in terms of $f\big(X(u-)\big)$ (i.e., in terms of $S(u-)$). □

**Corollary 11.5.3.** *Let $W(t)$ be a Brownian motion and let $N(t)$ be a Poisson process with intensity $\lambda > 0$, both defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and relative to the same filtration $\mathcal{F}(t)$, $t \geq 0$. Then the processes $W(t)$ and $N(t)$ are independent.*

KEY STEP IN PROOF: Let $u_1$ and $u_2$ be fixed real numbers and define

$$Y(t) = \exp\left\{u_1 W(t) + u_2 N(t) - \frac{1}{2}u_1^2 t - \lambda\big(e^{u_2} - 1\big)t\right\}.$$

We use the Itô-Doeblin formula to show that $Y$ is a martingale.

To do this, we define

$$X(s) = u_1 W(s) + u_2 N(s) - \frac{1}{2}u_1^2 s - \lambda\big(e^{u_2} - 1\big)s$$

and $f(x) = e^x$, so that $Y(s) = f\big(X(s)\big)$. The process $X(s)$ has Itô integral part $I(s) = u_1 W(s)$, Riemann integral part $R(s) = -\frac{1}{2}u_1^2 s - \lambda\big(e^{u_2} - 1\big)s$, and pure jump part $J(s) = u_2 N(s)$. In particular,

$$dX^c(s) = u_1\,dW(s) - \frac{1}{2}u_1^2\,ds - \lambda\big(e^{u_2} - 1\big)\,ds, \quad dX^c(s)\,dX^c(s) = u_1^2\,ds.$$

We next observe that if $Y$ has a jump at time $s$, then

$$Y(s) = \exp\left\{u_1 W(s) + u_2(N(s-) + 1) - \frac{1}{2}u_1^2 s - \lambda\big(e^{u_2} - 1\big)s\right\} = Y(s-)e^{u_2}.$$

Therefore,

$$Y(s) - Y(s-) = \big(e^{u_2} - 1\big)Y(s-)\Delta N(s).$$

According to the Itô-Doeblin formula for jump processes,

$$\begin{aligned}
Y(t) &= f\big(X(t)\big)\\
&= f\big(X(0)\big) + \int_0^t f'\big(X(s)\big)\,dX^c(s) + \frac{1}{2}\int_0^t f''\big(X(s)\big)\,dX^c(s)\,dX^c(s)\\
&\quad + \sum_{0<s\le t}\big[f\big(X(s)\big) - f\big(X(s-)\big)\big]\\
&= 1 + u_1\int_0^t Y(s)\,dW(s) - \frac{1}{2}u_1^2\int_0^t Y(s)\,ds - \lambda\big(e^{u_2}-1\big)\int_0^t Y(s)\,ds\\
&\quad + \frac{1}{2}u_1^2\int_0^t Y(s)\,ds + \sum_{0<s\le t}\big[Y(s) - Y(s-)\big]\\
&= 1 + u_1\int_0^t Y(s)\,dW(s) - \lambda\big(e^{u_2}-1\big)\int_0^t Y(s-)\,ds\\
&\quad + \big(e^{u_2}-1\big)\int_0^t Y(s-)\,dN(s)\\
&= 1 + u_1\int_0^t Y(s)\,dW(s) + \big(e^{u_2}-1\big)\int_0^t Y(s-)\,dM(s), \qquad (11.5.10)
\end{aligned}$$

where $M(s) = N(s) - \lambda s$ is a compensated Poisson process. Here we have used the fact that because $Y$ has only finitely many jumps, $\int_0^t Y(s)\,ds = \int_0^t Y(s-)\,ds$. The Itô integral $\int_0^t Y(s)\,dW(s)$ in the last line of (11.5.10) is a martingale, and the integral of the left-continuous process $Y(s-)$ with respect to the martingale $M(s)$ is also. Therefore, $Y$ is a martingale.

Because $Y(0) = 1$ and $Y$ is a martingale, we have $\mathbb{E}Y(t) = 1$ for all $t$. In other words,

$$\mathbb{E}\exp\Big\{u_1W(t) + u_2N(t) - \frac{1}{2}u_1^2t - \lambda\big(e^{u_2}-1\big)t\Big\} = 1 \text{ for all } t \ge 0.$$

We have obtained the joint moment-generating function formula

$$\mathbb{E}e^{u_1W(t)+u_2N(t)} = \exp\Big\{\frac{1}{2}u_1^2t\Big\}\cdot\exp\big\{\lambda t\big(e^{u_2}-1\big)\big\}.$$

This is the product of the moment-generating function $\mathbb{E}e^{u_1W(t)} = \exp\big\{\frac{1}{2}u_1^2t\big\}$ for $W(t)$ (see Exercise 1.6(i)) and the moment-generating function $\mathbb{E}e^{u_2N(t)} = \exp\big\{\lambda t\big(e^{u_2}-1\big)\big\}$ for $N(t)$ (see (11.3.4)). Since the joint moment-generating function factors into the product of moment-generating functions, the random variables $W(t)$ and $N(t)$ are independent.

The corollary asserts more than the independence between $N(t)$ and $W(t)$ for fixed $t$, saying that the *processes* $N$ and $W$ are independent (i.e., anything depending only on the path of $W$ is independent of anything depending only on the path of $N$). For example, the corollary asserts that $\max_{0\le s\le t} W(s)$ is independent of $\int_0^t N(s)\,ds$. The first step in the proof of this statement is the one just given, which shows that the *random variables* $W(t)$ and $N(t)$ are independent of each fixed $t$. The next step, which we omit, is to show that for any finite set of times $0 \le t_1 < t_2 < \cdots < t_n$, the *vector* of random variables $(W(t_1), W(t_2), \ldots, W(t_n))$ is independent of the *vector* of random variables $(N(t_1), N(t_2), \ldots, N(t_n))$. The assertion of the corollary follows from this. □

### 11.5.2 Itô-Doeblin Formula for Multiple Jump Processes

There is a multidimensional version of the Itô-Doeblin formula for jump processes. We give the two-dimensional version. The formula for higher dimensions follows the same pattern.

**Theorem 11.5.4 (Two-dimensional Itô-Doeblin formula for processes with jumps).** *Let $X_1(t)$ and $X_2(t)$ be jump processes, and let $f(t,x_1,x_2)$ be a function whose first and second partial derivatives appearing in the following formula are defined and are continuous. Then*

$$\begin{aligned}
&f\big(t,X_1(t),X_2(t)\big)\\
&= f\big(0,X_1(0),X_2(0)\big) + \int_0^t f_t\big(s,X_1(s),X_2(s)\big)\,ds\\
&\quad+\int_0^t f_{x_1}\big(s,X_1(s),X_2(s)\big)\,dX_1^c(s) + \int_0^t f_{x_2}\big(s,X_1(s),X_2(s)\big)\,dX_2^c(s)\\
&\quad+\frac{1}{2}\int_0^t f_{x_1,x_1}\big(s,X_1(s),X_2(s)\big)\,dX_1^c(s)\,dX_1^c(s)\\
&\quad+\int_0^t f_{x_1,x_2}\big(s,X_1(s),X_2(s)\big)\,dX_1^c(s)\,dX_2^c(s)\\
&\quad+\frac{1}{2}\int_0^t f_{x_2,x_2}\big(s,X_1(s),X_2(s)\big)\,dX_2^c(s)\,dX_2^c(s)\\
&\quad+\sum_{0<s\le t}\Big[f\big(s,X_1(s),X_2(s)\big) - f\big(s,X_1(s-),X_2(s-)\big)\Big].
\end{aligned}$$

**Corollary 11.5.5 (Itô's product rule for jump processes).** *Let $X_1(t)$ and $X_2(t)$ be jump processes. Then*

$$\begin{aligned}
X_1(t)X_2(t) &= X_1(0)X_2(0) + \int_0^t X_2(s)\,dX_1^c(s) + \int_0^t X_1(s)\,dX_2^c(s)\\
&\quad+[X_1^c,X_2^c](t) + \sum_{0<s\le t}\big[X_1(s)X_2(s) - X_1(s-)X_2(s-)\big]\\
&= X_1(0)X_2(0) + \int_0^t X_2(s-)\,dX_1(s) + \int_0^t X_1(s-)\,dX_2(s)\\
&\quad+[X_1,X_2](t). \qquad (11.5.11)
\end{aligned}$$

PROOF: Take $f(x_1, x_2) = x_1 x_2$ so that

$$f_{x_1} = x_2, \; f_{x_2} = x_1, \; f_{x_1 x_1} = 0, \; f_{x_1 x_2} = 1, \; f_{x_2 x_2} = 0.$$

The two-dimensional Itô-Doeblin formula implies

$$\begin{aligned} X_1(t)X_2(t) = X_1(0)X_2(0) &+ \int_0^t X_2(s)\, dX_1^c(s) + \int_0^t X_1(s)\, dX_2^c(s) \\ &+ \int_0^t 1\, dX_1^c(s)\, dX_2^c(s) + \sum_{0<s\leq t} \Big[X_1(s)X_2(s) - X_1(s-)X_2(s-)\Big]. \end{aligned} \tag{11.5.12}$$

The notation $\int_0^t 1\, dX_1^c(s)\, dX_2^c(s)$ in (11.5.12) means $[X_1^c, X_2^c](t)$ (see Remark 11.4.8). This establishes the first equality in (11.5.11).

To obtain the second equality, we denote by $J_1(t) = X_1(t) - X_1^c(t)$ and $J_2(t) = X_2(t) - X_2^c(t)$ the pure jump parts of $X_1(t)$ and $X_2(t)$, respectively, and begin with the last line of (11.5.11), using (11.4.12) to compute

$$\begin{aligned} &X_1(0)X_2(0) + \int_0^t X_2(s-)\, dX_1(s) + \int_0^t X_1(s-)\, dX_2(s) + [X_1, X_2](t) \\ &= X_1(0)X_2(0) + \int_0^t X_2(s-)\, dX_1^c(s) + \int_0^t X_2(s-)\, dJ_1(s) \\ &\quad + \int_0^t X_1(s-)\, dX_2^c(s) + \int_0^t X_1(s-)\, dJ_2(s) \\ &\quad + [X_1^c, X_2^c](t) + \sum_{0<s\leq t} \Delta J_1(s) \Delta J_2(s) \\ &= X_1(0)X_2(0) + \int_0^t X_2(s)\, dX_1^c(s) + \int_0^t X_1(s)\, dX_2^c(s) + [X_1^c, X_2^c](t) \\ &\quad + \sum_{0<s\leq t} \big[X_2(s-)\Delta X_1(s) + X_1(s-)\Delta X_2(s) + \Delta X_1(s)\Delta X_2(s)\big]. \end{aligned} \tag{11.5.13}$$

We have also used the fact that the jumps in $X_i(t)$ are the same as the jumps in $J_i(t)$. It remains to show that this last sum is the same as the sum

$$\sum_{0<s\leq t} \big[X_1(s)X_2(s) - X_1(s-)X_2(s-)\big]$$

in the second line of (11.5.11). We expand the typical term in the sum in the second line of (11.5.11):

$$\begin{aligned}
&X_1(s)X_2(s) - X_1(s-)X_2(s-)\\
&= \big(X_1(s-) + \Delta X_1(s)\big)\big(X_2(s-) + \Delta X_2(s)\big) - X_1(s-)X_2(s-)\\
&= X_1(s-)X_2(s-) + X_1(s-)\Delta X_2(s) + \Delta X_1(s)X_2(s-) + \Delta X_1(s)\Delta X_2(s)\\
&\qquad -X_1(s-)X_2(s-)\\
&= X_1(s-)\Delta X_2(s) + \Delta X_1(s)X_2(s-) + \Delta X_1(s)\Delta X_2(s).
\end{aligned}$$

This is the typical term in the sum appearing at the end of (11.5.13). □

For stochastic calculus without jumps, Girsanov's Theorem tells us how to change the measure using the Radon-Nikodým derivative process

$$Z(t) = \exp\left\{ - \int_0^t \Gamma(s)\,dW(s) - \frac{1}{2}\int_0^t \Gamma^2(s)\,ds\right\}.$$

This process satisfies the stochastic differential equation

$$dZ(t) = -\Gamma(t)Z(t)\,dW(t) = Z(t)\,dX^c(t),$$

where $X^c(t) = -\int_0^t \Gamma(s)\,dW(s)$ and $[X^c, X^c](t) = \int_0^t \Gamma^2(s)\,ds$. We may rewrite $Z(t)$ as

$$Z(t) = \exp\left\{X^c(t) - \frac{1}{2}[X^c, X^c](t)\right\}. \tag{11.5.14}$$

In stochastic calculus for processes with jumps, the analogous stochastic differential equation is

$$dZ^X(t) = Z^X(t-)\,dX(t), \tag{11.5.15}$$

where the integrator $X$ is now allowed to have jumps. The solution to (11.5.15) is like (11.5.14), except now, whenever there is a jump in $X$, (11.5.15) says there is a jump in $Z^X$ of size

$$\Delta Z^X(s) = Z^X(s-)\Delta X(s).$$

Therefore,

$$Z^X(s) = Z^X(s-) + \Delta Z^X(s) = Z^X(s-)\big(1 + \Delta X(s)\big).$$

The following corollary presents the result.

**Corollary 11.5.6.** *Let $X(t)$ be a jump process. The* Doleans-Dade exponential *of $X$ is defined to be the process*

$$Z^X(t) = \exp\left\{X^c(t) - \frac{1}{2}[X^c, X^c](t)\right\} \prod_{0<s\le t} \big(1 + \Delta X(s)\big).$$

*This process is the solution to the stochastic differential equation (11.5.15) with initial condition $Z^X(0) = 1$, which in integral form is*

$$Z^X(t) = 1 + \int_0^t Z^X(s-)\,dX(s). \tag{11.5.16}$$

PROOF: We may write $X(t)$ as $X(t) = X^c(t) + J(t)$, where

$$X^c(t) = \int_0^t \Gamma(s)\,dW(s) + \int_0^t \Theta(s)\,ds \tag{11.5.17}$$

is the continuous part of $X$ and $J(t)$ is the pure jump part. We define

$$\begin{aligned} Y(t) &= \exp\left\{\int_0^t \Gamma(s)\,dW(s) + \int_0^t \Theta(s)\,ds - \frac{1}{2}\int_0^t \Gamma^2(s)\,ds\right\} \\ &= \exp\left\{X^c(t) - \frac{1}{2}[X^c, X^c](t)\right\}. \end{aligned} \tag{11.5.18}$$

From the Itô-Doeblin formula for continuous processes, we know that

$$dY(t) = Y(t)\,dX^c(t) = Y(t-)\,dX^c(t). \tag{11.5.19}$$

We next define $K(t) = 1$ for $t$ between 0 and the time of the first jump of $X$, and we set

$$K(t) = \prod_{0<s\le t} \big(1 + \Delta X(s)\big) \tag{11.5.20}$$

for $t$ greater than or equal to the first jump time of $X$. The process $K(t)$ is a pure jump process, and $Z^X(t) = Y(t)K(t)$. If $X$ has a jump at time $t$, then $K(t) = K(t-)\big(1 + \Delta X(t)\big)$. Therefore,

$$\Delta K(t) = K(t) - K(t-) = K(t-)\Delta X(t). \tag{11.5.21}$$

Because $Y(t)$ is continuous and $K(t)$ is a pure jump process, $[Y, K](t) = 0$. We now use Itô's product rule for jump processes to obtain

$$\begin{aligned} Z^X(t) &= Y(t)K(t) \\ &= Y(0) + \int_0^t K(s-)\,dY(s) + \int_0^t Y(s-)\,dK(s) \\ &= 1 + \int_0^t Y(s-)K(s-)\,dX^c(s) + \sum_{0<s\le t} Y(s-)K(s-)\Delta X(s) \\ &= 1 + \int_0^t Y(s-)K(s-)\,dX(s) \\ &= 1 + \int_0^t Z^X(s-)\,dX(s). \end{aligned} \tag{11.5.22}$$

This is (11.5.16). □

## 11.6 Change of Measure

Just as we can use Girsanov's Theorem to change the measure so that a Brownian motion with drift becomes a Brownian motion without drift, we can change the measure for Poisson processes and compound Poisson processes. For a Poisson process, the change of measure affects the intensity. For a compound Poisson process, the change of measure can affect both the intensity and the distribution of the jump sizes. We treat these two situations in the next two subsections, and in the third subsection we also include a Brownian motion component in the process under consideration.

### 11.6.1 Change of Measure for a Poisson Process

Let $N(t)$ be a Poisson process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ relative to a filtration $\mathcal{F}(t)$, $t \geq 0$. We denote the intensity of $N(t)$ by $\lambda$, a positive constant (i.e., $\mathbb{E}N(t) = \lambda t$). The compensated Poisson process $M(t) = N(t) - \lambda t$ is a martingale under $\mathbb{P}$ (Theorem 11.2.4). Let $\tilde{\lambda}$ be a positive number. We define

$$Z(t) = e^{(\lambda - \tilde{\lambda})t} \left( \frac{\tilde{\lambda}}{\lambda} \right)^{N(t)}. \tag{11.6.1}$$

We fix a time $T > 0$ and will use $Z(T)$ to change to a new measure $\widetilde{\mathbb{P}}$ under which $N(t)$, $0 \leq t \leq T$, has intensity $\tilde{\lambda}$ rather than $\lambda$. It is clear that $Z(T) > 0$ almost surely. In order to use $Z(T)$ to change the measure, we also need to verify that $\mathbb{E}Z(T) = 1$.

**Lemma 11.6.1.** *The process $Z(t)$ of (11.6.1) satisfies*

$$dZ(t) = \frac{\tilde{\lambda} - \lambda}{\lambda} Z(t-)\, dM(t). \tag{11.6.2}$$

*In particular, $Z(t)$ is a martingale under $\mathbb{P}$ and $\mathbb{E}Z(t) = 1$ for all $t$.*

PROOF: Define $X(t) = \frac{\tilde{\lambda} - \lambda}{\lambda} M(t)$, which is a martingale with continuous part $X^c(t) = (\lambda - \tilde{\lambda})t$ and pure jump part $J(t) = \frac{\tilde{\lambda} - \lambda}{\lambda} N(t)$. Then $[X^c, X^c](t) = 0$, and if there is a jump at time $t$, then $\Delta X(t) = \frac{\tilde{\lambda} - \lambda}{\lambda}$, so

$$1 + \Delta X(t) = \frac{\tilde{\lambda}}{\lambda}.$$

Therefore, the process in (11.6.1) may be written as

$$Z(t) = \exp\left\{ X^c(t) - \frac{1}{2}[X^c, X^c](t) \right\} \prod_{0 < s \leq t} (1 + \Delta X(s)).$$

We see from this formula that $Z(t)$ is the Doleans-Dade exponential $Z^X(t)$ of Corollary 11.5.6. In particular,

$$Z(t) = 1 + \int_0^t Z(s-)\, dX(s).$$

Since $X$ is a martingale and $Z(s-)$ is left-continuous, $Z(t)$ is a martingale. Because $Z(t)$ is a martingale and $Z(0) = 1$, we know that $\mathbb{E}Z(t) = 1$ for all $t \geq 0$. □

We may now fix a positive time $T$ and use $Z(T)$ to change the measure. We define

$$\widetilde{\mathbb{P}}(A) = \int_A Z(T)\, d\mathbb{P} \quad \text{for all } A \in \mathcal{F}. \tag{11.6.3}$$

**Theorem 11.6.2 (Change of Poisson intensity).** *Under the probability measure $\widetilde{\mathbb{P}}$, the process $N(t)$, $0 \leq t \leq T$, is Poisson with intensity $\tilde{\lambda}$.*

KEY STEP IN PROOF: We compute the moment-generating function of $N(t)$ under $\widetilde{\mathbb{P}}$. For $0 \leq t \leq T$, we can change the $\widetilde{\mathbb{E}}$ expectation of $e^{uN(t)}$ to the $\mathbb{E}$ expectation by using $Z(t)$ as the Radon-Nikodým derivative rather than $Z(T)$ (see Lemma 5.2.1). Using the formula for $Z(t)$ and the moment-generating function formula (11.3.4), we obtain

$$\begin{aligned}
\mathbb{E}\left[e^{uN(t)}Z(t)\right] &= e^{(\lambda-\tilde{\lambda})t}\,\mathbb{E}\left[e^{uN(t)}\left(\frac{\tilde{\lambda}}{\lambda}\right)^{N(t)}\right] \\
&= e^{(\lambda-\tilde{\lambda})t}\,\mathbb{E}\left[\exp\left\{\left(u + \log\frac{\tilde{\lambda}}{\lambda}\right)N(t)\right\}\right] \\
&= e^{(\lambda-\tilde{\lambda})t}\exp\left\{\lambda t\left(e^{u+\log(\tilde{\lambda}/\lambda)} - 1\right)\right\} \\
&= \exp\left\{\tilde{\lambda}t(e^u - 1)\right\},
\end{aligned}$$

which is the moment generating function for a Poisson process with intensity $\tilde{\lambda}$ (see again (11.3.4)). □

*Example 11.6.3.* Consider a stock modeled as a geometric Poisson process

$$S(t) = S(0)\exp\{\alpha t + N(t)\log(\sigma+1) - \lambda\sigma t\} = S(0)e^{(\alpha-\lambda\sigma)t}(\sigma+1)^{N(t)},$$

where $\sigma > -1$, $\sigma \neq 0$, and $N(t)$ is a Poisson process with intensity $\lambda$ under the actual probability measure $\mathbb{P}$. We saw in Example 11.5.2 that $e^{-\alpha t}S(t)$ is a martingale under $\mathbb{P}$, and hence $S(t)$ has mean rate of return $\alpha$. Indeed, in place of (11.5.9), we now have

$$dS(t) = \alpha S(t)\,dt + \sigma S(t-)\,dM(t), \tag{11.6.4}$$

where $M(t)$ is the compensated Poisson process $M(t) = N(t) - \lambda t$. We would like to change to a probability measure $\widetilde{\mathbb{P}}$ under which

$$dS(t) = rS(t)\,dt + \sigma S(t-)\,d\widetilde{M}(t), \tag{11.6.5}$$

where $r$ is the interest rate, $N(t)$ is a Poisson process with intensity $\tilde{\lambda}$ under $\widetilde{\mathbb{P}}$, and $\widetilde{M}(t) = N(t) - \tilde{\lambda}t$ is a compensated Poisson process under $\widetilde{\mathbb{P}}$. Then, under $\widetilde{\mathbb{P}}$, the geometric Poisson process would have mean rate of return equal to the interest rate, and $\widetilde{\mathbb{P}}$ would be the risk-neutral measure.

To accomplish this, we note that the "$dt$" term in (11.6.4) is

$$(\alpha - \lambda\sigma)S(t)\,dt \tag{11.6.6}$$

(recall that $dM(t) = dN(t) - \lambda\,dt$) and the "$dt$" term in (11.6.5) is

$$(r - \tilde{\lambda}\sigma)S(t)\,dt. \tag{11.6.7}$$

(Here again we are using the fact that $S(t-)\,dt$ and $S(t)\,dt$ have the same integrals, and we can thus use them interchangeably.) We set (11.6.6) and (11.6.7) equal and solve for

$$\tilde{\lambda} = \lambda - \frac{\alpha - r}{\sigma}.$$

We then change to the risk-neutral measure by formula (11.6.3) with $Z(T)$ defined by (11.6.1).

To make the change of measure, we must have $\tilde{\lambda} > 0$, which is equivalent to

$$\lambda > \frac{\alpha - r}{\sigma}. \tag{11.6.8}$$

If condition (11.6.8) does not hold, then there is no risk-neutral measure and hence there must be an arbitrage. Indeed, if $\sigma > 0$ and (11.6.8) fails, then

$$S(t) \geq S(0)e^{rt}(\sigma + 1)^{N(t)} \geq S(0)e^{rt},$$

and borrowing at the interest rate $r$ to invest in the stock is an arbitrage. If $-1 < \sigma < 0$, the inequalities are reversed and the arbitrage consists of shorting the stock to invest in the money market account. □

### 11.6.2 Change of Measure for a Compound Poisson Process

Let $N(t)$ be a Poisson process with intensity $\lambda$, and let $Y_1, Y_2, \ldots$ be a sequence of identically distributed random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We assume the random variables $Y_1, Y_2, \ldots$ are independent of one another and also independent of the Poisson process $N(t)$. We define the *compound Poisson process*

$$Q(t) = \sum_{i=1}^{N(t)} Y_i. \tag{11.6.9}$$

Note for future reference that if $N$ jumps at time $t$, then $Q$ jumps at time $t$ and

$$\Delta Q(t) = Y_{N(t)}. \tag{11.6.10}$$

Our goal is to change the measure so that the intensity of $N(t)$ and the distribution of the jump sizes $Y_1, Y_2, \ldots$ both change. We first consider the case when the jump-size random variables have a discrete distribution (i.e., each $Y_i$ takes one of finitely many possible nonzero values $y_1, y_2, \ldots, y_M$). Let $p(y_m)$ denote the probability that a jump is of size $y_m$:

$$p(y_m) = \mathbb{P}\{Y_i = y_m\}, \ m = 1, \ldots, M.$$

This does not depend on $i$ since $Y_1, Y_2, \ldots$ are identically distributed. We assume that $p(y_m) > 0$ for every $m$ and, of course, that $\sum_{m=1}^{M} p(y_m) = 1$.

Let $N_m(t)$ denote the number of jumps in $Q(t)$ of size $y_m$ up to and including time $t$, so that

$$N(t) = \sum_{m=1}^{M} N_m(t) \text{ and } Q(t) = \sum_{m=1}^{M} y_m N_m(t).$$

According to Corollary 11.3.4, $N_1, \ldots, N_M$ are independent Poisson processes and each $N_m$ has intensity $\lambda_m = \lambda p(y_m)$.

Let $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_M$ be given positive numbers, and set

$$Z_m(t) = e^{(\lambda_m - \tilde{\lambda}_m)t} \left( \frac{\tilde{\lambda}_m}{\lambda_m} \right)^{N_m(t)} \quad \text{and} \quad Z(t) = \prod_{m=1}^{M} Z_m(t). \tag{11.6.11}$$

**Lemma 11.6.4.** *The process $Z(t)$ of (11.6.11) is a martingale. In particular, $\mathbb{E}Z(t) = 1$ for all $t$.*

PROOF: From Lemma 11.6.1, we have

$$dZ_m(t) = \frac{\tilde{\lambda}_m - \lambda_m}{\lambda_m} Z_m(t-)\, dM_m(t), \tag{11.6.12}$$

where

$$M_m(t) = N_m(t) - \lambda_m\, dt.$$

Because the integrand in (11.6.12) is left-continuous and the compensated Poisson process is a martingale, the process $Z_m$ is a martingale (Theorem 11.4.5).

For $m \neq n$, the Poisson processes $N_m$ and $N_n$ have no simultaneous jumps, and hence $[Z_m, Z_n] = 0$. Itô's product rule (Corollary 11.5.5) implies that

$$d\big(Z_1(t)Z_2(t)\big) = Z_2(t-)\, dZ_1(t) + Z_1(t-)\, dZ_2(t). \tag{11.6.13}$$

Because both $Z_1$ and $Z_2$ are martingales and the integrands in (11.6.13) are left-continuous, the process $Z_1 Z_2$ is a martingale. Because $Z_1 Z_2$ has no jumps simultaneous with the jumps of $Z_3$, Itô's product rule further implies

$$d\big(Z_1(t)Z_2(t)Z_3(t)\big) = Z_3(t-)\, d\big(Z_1(t)Z_2(t)\big) + \big(Z_1(t-)Z_2(t-)\big)\, dZ_3(t).$$

Once again, the integrators are martingales and the integrands are left-continuous. Therefore, $Z_1 Z_2 Z_3$ is a martingale. Continuing this process, we eventually conclude that $Z(t) = Z_1(t) Z_2(t) \cdots Z_m(t)$ is a martingale. □

Fix $T > 0$. Because $Z(T) > 0$ almost surely and $\mathbb{E}Z(T) = 1$, we can use $Z(T)$ to change the measure, defining

$$\widetilde{\mathbb{P}}(A) = \int_A Z(T)\, dP \text{ for all } Z \in \mathcal{F}.$$

**Theorem 11.6.5 (Change of compound Poisson intensity and jump distribution for finitely many jump sizes).** *Under $\widetilde{\mathbb{P}}$, $Q(t)$ is a compound Poisson process with intensity $\tilde{\lambda} = \sum_{m=1}^{M} \tilde{\lambda}_m$, and $Y_1, Y_2, \ldots$ are independent, identically distributed random variables with*

$$\widetilde{\mathbb{P}}\{Y_i = y_m\} = \tilde{p}(y_m) = \frac{\tilde{\lambda}_m}{\tilde{\lambda}}. \tag{11.6.14}$$

KEY STEP IN PROOF: We use the independence of $N_1, \ldots, N_M$ under $\mathbb{P}$ to compute the moment-generating function of $Q(t)$ under $\widetilde{\mathbb{P}}$. For $0 \leq t \leq T$, Lemma 5.2.1 and the moment-generating function formula (11.3.4) imply

$$\begin{aligned}
\widetilde{\mathbb{E}}\left[e^{uQ(t)}\right] &= \mathbb{E}\left[e^{uQ(t)} Z(t)\right] \\
&= \mathbb{E}\left[\exp\left\{u \sum_{m=1}^{M} y_m N_m(t)\right\} \cdot \prod_{m=1}^{M} e^{(\lambda_m - \tilde{\lambda}_m)t} \left(\frac{\tilde{\lambda}_m}{\lambda_m}\right)^{N_m(t)}\right] \\
&= \prod_{m=1}^{M} \exp\{(\lambda_m - \tilde{\lambda}_m)t\} \cdot \mathbb{E}\left[\exp\left\{\left(u y_m + \log \frac{\tilde{\lambda}_m}{\lambda_m}\right) N_m(t)\right\}\right] \\
&= \prod_{m=1}^{M} \exp\{(\lambda_m - \tilde{\lambda}_m)t\} \exp\left\{\lambda_m t \left(e^{u y_m + \log(\tilde{\lambda}_m/\lambda_m)} - 1\right)\right\} \\
&= \prod_{m=1}^{M} \exp\left\{(\lambda_m - \tilde{\lambda}_m)t + \tilde{\lambda}_m t e^{u y_m} - \lambda_m t\right\} \\
&= \prod_{m=1}^{M} \exp\left\{\tilde{\lambda}_m t (e^{u y_m} - 1)\right\} \\
&= \prod_{m=1}^{M} \exp\left\{\tilde{\lambda} t \tilde{p}(y_m) e^{u y_m} - \tilde{\lambda}_m t\right\} \\
&= \exp\left\{\tilde{\lambda} t \left(\sum_{m=1}^{M} \tilde{p}(y_m) e^{u y_m} - 1\right)\right\}.
\end{aligned}$$

According to (11.3.5), this is the moment-generating function for a compound Poisson process with intensity $\tilde{\lambda}$ and jump-size distribution (11.6.14). □

The Radon-Nikodým derivative process $Z(t)$ of (11.6.11) may be written as

$$Z(t) = \exp\left\{\sum_{m=1}^{M}(\lambda_m - \tilde{\lambda}_m)t\right\} \cdot \prod_{m=1}^{M}\left(\frac{\tilde{\lambda}\tilde{p}(y_m)}{\lambda p(y_m)}\right)^{N_m(t)} = e^{(\lambda-\tilde{\lambda})t}\prod_{i=1}^{N(t)}\frac{\tilde{\lambda}\tilde{p}(Y_i)}{\lambda p(Y_i)}.$$

This suggests that if $Y_1, Y_2, \dots$ are not discrete but instead have a common density $f(y)$, then we could change the measure so that $Q(t)$ has intensity $\tilde{\lambda}$ and $Y_1, Y_2, \dots$ have a different density $\widetilde{f}(y)$ by using the Radon-Nikodým derivative process

$$Z(t) = e^{(\lambda-\tilde{\lambda})t}\prod_{i=1}^{N(t)}\frac{\tilde{\lambda}\widetilde{f}(Y_i)}{\lambda f(Y_i)}. \tag{11.6.15}$$

This is in fact the case, although the proof, given below, is harder than the one just given for the case of a discrete jump-size distribution.

To avoid division by zero in (11.6.15), we assume that $\widetilde{f}(y) = 0$ whenever $f(y) = 0$. This means that if a certain set of jump sizes has probability zero under $\mathbb{P}$, then it will also have probability zero under $\widetilde{\mathbb{P}}$ considered in Theorem 11.6.7 below.

**Lemma 11.6.6.** *The process $Z(t)$ of (11.6.15) is a martingale. In particular, $\mathbb{E}Z(t) = 1$ for all $t \geq 0$.*

PROOF: We define the pure jump process

$$J(t) = \prod_{i=1}^{N(t)}\frac{\tilde{\lambda}\widetilde{f}(Y_i)}{\lambda f(Y_i)}. \tag{11.6.16}$$

At the jump times of $Q$, which are also the jump times of $N$ and $J$, we have (recall (11.6.10))

$$J(t) = J(t-)\frac{\tilde{\lambda}\widetilde{f}(Y_{N(t)})}{\lambda f(Y_{N(t)})} = J(t-)\frac{\tilde{\lambda}\widetilde{f}(\Delta Q(t))}{\lambda f(\Delta Q(t))},$$

and hence

$$\Delta J(t) = J(t) - J(t-) = \left[\frac{\tilde{\lambda}\widetilde{f}(\Delta Q(t))}{\lambda f(\Delta Q(t))} - 1\right]J(t-) \tag{11.6.17}$$

at the jump times of $Q$.

We define the compound Poisson process

$$H(t) = \sum_{i=1}^{N(t)}\frac{\tilde{\lambda}\widetilde{f}(Y_i)}{\lambda f(Y_i)} \tag{11.6.18}$$

for which

$$\Delta H(t) = \frac{\tilde{\lambda}\tilde{f}\big(\Delta Q(t)\big)}{\lambda f\big(\Delta Q(t)\big)}. \tag{11.6.19}$$

Because

$$\mathbb{E}\left[\frac{\tilde{\lambda}\tilde{f}(Y_i)}{\lambda f(Y_i)}\right] = \frac{\tilde{\lambda}}{\lambda}\int_{-\infty}^{\infty} \frac{\tilde{f}(y)}{f(y)} f(y)\, dy = \frac{\tilde{\lambda}}{\lambda}\int_{-\infty}^{\infty} \tilde{f}(y)\, dy = \frac{\tilde{\lambda}}{\lambda},$$

the compensated compound Poisson process $H(t) - \tilde{\lambda}t$ is a martingale (Theorem 11.3.1 with $\beta = \frac{\tilde{\lambda}}{\lambda}$). We may rewrite (11.6.17) as

$$\Delta J(t) = J(t-)\Delta H(t) - J(t-)\Delta N(t), \tag{11.6.20}$$

and because all these terms are zero if there is no jump at $t$, this equation holds at all times $t$, not just at the jump times of $Q$. Because $J$, $H$, and $N$ are all pure jump processes, we may also write (11.6.20) as

$$dJ(t) = J(t-)\, dH(t) - J(t-)\, dN(t).$$

Because $J(t)$ is a pure jump process and $e^{(\lambda-\tilde{\lambda})t}$ is continuous, the cross variation between these two processes is zero. Therefore, Itô's product rule for jump processes (Corollary 11.5.5) implies that $Z(t) = e^{(\lambda-\tilde{\lambda})t}J(t)$ may be written as

$$\begin{aligned}
Z(t) &= Z(0) + \int_0^t J(s-)(\lambda - \tilde{\lambda})e^{(\lambda-\tilde{\lambda})s}\, ds + \int_0^t e^{(\lambda-\tilde{\lambda})s}\, dJ(s)\\
&= 1 + \int_0^t e^{(\lambda-\tilde{\lambda})s}J(s-)(\lambda - \tilde{\lambda})\, ds + \int_0^t e^{(\lambda-\tilde{\lambda})s}J(s-)\, dH(s)\\
&\qquad - \int_0^t e^{(\lambda-\tilde{\lambda})s}J(s-)\, dN(s)\\
&= 1 + \int_0^t e^{(\lambda-\tilde{\lambda})s}J(s-)\, d\big(H(s) - \tilde{\lambda}s\big) - \int_0^t e^{(\lambda-\tilde{\lambda})s}J(s-)\, d\big(N(s) - \lambda s\big)\\
&= 1 + \int_0^t Z(s-)\, d\big(H(s) - \tilde{\lambda}s\big) - \int_0^t Z(s-)\, d\big(N(s) - \lambda s\big).
\end{aligned} \tag{11.6.21}$$

Theorem 11.4.5 implies that $Z(t)$ is a martingale. Since $Z(t)$ is a martingale and $Z(0) = 1$, we have $\mathbb{E}Z(t) = 1$ for all $t$. □

For future reference, we rewrite (11.6.21) in the differential form

$$dZ(t) = Z(t-)\, d\big(H(t) - \tilde{\lambda}t\big) - Z(t-)\, d\big(N(t) - \lambda t\big).$$

This equation implies

$$\Delta Z(t) = Z(t-)\Delta H(t) - Z(t-)\Delta N(t). \tag{11.6.22}$$

Fix a positive $T$ and define

$$\widetilde{\mathbb{P}}(A) = \int_A Z(T)\,d\mathbb{P} \quad \text{for all } A \in \mathcal{F}. \tag{11.6.23}$$

**Theorem 11.6.7 (Change of compound Poisson intensity and jump distribution for a continuum of jump sizes).** *Under the probability measure $\widetilde{\mathbb{P}}$, the process $Q(t)$, $0 \le t \le T$, of (11.6.9) is a compound Poisson process with intensity $\tilde{\lambda}$. Furthermore, the jumps in $Q(t)$ are independent and identically distributed with density $\widetilde{f}(y)$.*

KEY STEP IN PROOF: We need to show that, under $\widetilde{\mathbb{P}}$, the process $Q(t)$ has the moment-generating function corresponding to a compound Poisson process with intensity $\tilde{\lambda}$ and jump density $\widetilde{f}(y)$. In other words, we must show that (see (11.3.2))

$$\widetilde{\mathbb{E}} e^{uQ(t)} = \exp\left\{\tilde{\lambda} t\big(\widetilde{\varphi}_Y(u) - 1\big)\right\}, \tag{11.6.24}$$

where

$$\widetilde{\varphi}_Y(u) = \int_{-\infty}^{\infty} e^{uy}\widetilde{f}(y)\,dy. \tag{11.6.25}$$

We define

$$X(t) = \exp\left\{uQ(t) - \tilde{\lambda} t\big(\widetilde{\varphi}_Y(u) - 1\big)\right\}$$

and show that $X(t)Z(t)$ is a martingale under $\mathbb{P}$. At jump times of $Q$,

$$X(t) = X(t-)e^{u\Delta Q(t)},$$

and hence

$$\Delta X(t) = X(t) - X(t-) = X(t-)\left(e^{u\Delta Q(t)} - 1\right). \tag{11.6.26}$$

We introduce the compound Poisson process

$$V(t) = \sum_{i=1}^{N(t)} e^{uY_i}\frac{\tilde{\lambda}\widetilde{f}(Y_i)}{\lambda f(Y_i)}.$$

Because

$$\mathbb{E}\left[e^{uY_i}\frac{\tilde{\lambda}\widetilde{f}(Y_i)}{\lambda f(Y_i)}\right] = \frac{\tilde{\lambda}}{\lambda}\int_{-\infty}^{\infty} e^{uy}\frac{\widetilde{f}(y)}{f(y)}f(y)\,dy = \frac{\tilde{\lambda}}{\lambda}\widetilde{\varphi}_Y(u),$$

the compensated compound Poisson process $V(t) - \tilde{\lambda} t\widetilde{\varphi}_Y(u)$ is a martingale (see Theorem 11.3.1 with $\beta = \frac{\tilde{\lambda}}{\lambda}\widetilde{\varphi}_Y(u)$). At jump times of $Q$,

$$\Delta V(t) = e^{u\Delta Q(t)}\frac{\tilde{\lambda}\widetilde{f}(\Delta Q(t))}{\lambda f(\Delta Q(t))} = e^{u\Delta Q(t)}\Delta H(t), \tag{11.6.27}$$

where $H(t)$, defined by (11.6.18), satisfies (11.6.19) at jump times of $Q$.

Because $X(t)$ and $Z(t)$ have no Itô integral components, (11.6.26), (11.6.22), and (11.6.27) imply

$$\begin{aligned}
[X,Z](t) &= \sum_{0<s\le t} \Delta X(s)\Delta Z(s)\\
&= \sum_{0<s\le t} X(s-)Z(s-)\big(e^{u\Delta Q(s)}-1\big)\Delta H(s)\\
&\quad - \sum_{0<s\le t} X(s-)Z(s-)\big(e^{u\Delta Q(s)}-1\big)\Delta N(s)\\
&= \sum_{0<s\le t} X(s-)Z(s-)\Delta V(s) - \sum_{0<s\le t} X(s-)Z(s-)\Delta H(s)\\
&\quad - \sum_{0<s\le t} X(s-)Z(s-)\big(e^{u\Delta Q(s)}-1\big). \qquad (11.6.28)
\end{aligned}$$

We have omitted $\Delta N(s)$ in the last term because it is always either 1 or 0, and when it is zero, $e^{u\Delta Q(s)}-1$ is also zero. In other words,

$$\big(e^{u\Delta Q(s)}-1\big)\Delta N(s) = \big(e^{u\Delta Q(s)}-1\big).$$

We use Itô's product rule for jump processes to write

$$X(t)Z(t) = 1 + \int_0^t X(s-)\,dZ(s) + \int_0^t Z(s-)\,dX(s) + [X,Z](t).$$

We show that the right-hand side is a martingale under $\mathbb{P}$. The integral $\int_0^t X(s-)\,dZ(s)$ is a martingale because the integrand is left-continuous and $Z$ is a martingale. We examine the two other terms, using (11.6.26) and (11.6.28):

$$\begin{aligned}
&\int_0^t Z(s-)\,dX(s) + [X,Z](t)\\
&= \int_0^t Z(s-)\,dX^c(s) + \sum_{0<s\le t} Z(s-)\Delta X(s) + [X,Z](t)\\
&= -\tilde{\lambda}\big(\widetilde{\varphi}_Y(u)-1\big)\int_0^t X(s-)Z(s-)ds + \sum_{0<s\le t} X(s-)Z(s-)\big(e^{u\Delta Q(s)}-1\big)\\
&\quad + \sum_{0<s\le t} X(s-)Z(s-)\Delta V(s) - \sum_{0<s\le t} X(s-)Z(s-)\Delta H(s)\\
&\quad - \sum_{0<s\le t} X(s-)Z(s-)\big(e^{u\Delta Q(s)}-1\big)\\
&= \int_0^t X(s-)Z(s-)\,d\big(V(s)-\tilde{\lambda}s\widetilde{\varphi}_Y(u)\,ds\big) - \int_0^t X(s-)Z(s-)\,d\big(H(s)-\tilde{\lambda}s\big).
\end{aligned}$$

This is a martingale because the processes $V(t)-\tilde{\lambda}t\widetilde{\varphi}_Y(u)$ and $H(t)-\tilde{\lambda}t$ are martingales and the integrands are left-continuous.

We can now prove (11.6.24). Using Lemma 5.2.1, we may write

$$\widetilde{\mathbb{E}}\big[e^{uQ(t)}\big] = \mathbb{E}\big[e^{uQ(t)}Z(t)\big]. \tag{11.6.29}$$

But the martingale $X(t)Z(t)$ has constant expectation 1, which implies

$$\begin{aligned} 1 &= \mathbb{E}\big[X(t)Z(t)\big] \\ &= \exp\big\{-\tilde{\lambda}t\big(\widetilde{\varphi}_Y(u)-1\big)\big\}\cdot\mathbb{E}\big[e^{uQ(t)}Z(t)\big]. \end{aligned} \tag{11.6.30}$$

Combining (11.6.29) and (11.6.30), we obtain (11.6.24). □

### 11.6.3 Change of Measure for a Compound Poisson Process and a Brownian Motion

Suppose now that we have a probability space $(\Omega,\mathcal{F},\mathbb{P})$ on which is defined a Brownian motion $W(t)$. Suppose that on this same probability space there is defined a compound Poisson process

$$Q(t) = \sum_{i=1}^{N(t)} Y_i$$

as in (11.3.1) with intensity $\lambda$ and jumps having density function $f(y)$. We assume there is a single filtration $\mathcal{F}(t)$, $t \geq 0$, for both the Brownian motion and the compound Poisson process. In this case, the Brownian motion and compound Poisson process must be independent. (See Corollary 11.4.9 for the case of a Brownian motion and a Poisson process. The case of a Brownian motion and a compound Poisson process is Exercise 11.6.)

Let $\tilde{\lambda}$ be a positive number, let $\widetilde{f}(y)$ be another density function with the property that $\widetilde{f}(y)=0$ whenever $f(y)=0$, and let $\Theta(t)$ be an adapted process. We define

$$Z_1(t) = \exp\left\{-\int_0^t \Theta(u)\,dW(u) - \frac{1}{2}\int_0^t \Theta^2(u)\,du\right\}, \tag{11.6.31}$$

$$Z_2(t) = e^{(\lambda-\tilde{\lambda})t}\prod_{i=1}^{N(t)}\frac{\tilde{\lambda}\widetilde{f}(Y_i)}{\lambda f(Y_i)}, \tag{11.6.32}$$

$$Z(t) = Z_1(t)Z_2(t). \tag{11.6.33}$$

**Lemma 11.6.8.** *The process $Z(t)$ of (11.6.33) is a martingale. In particular, $\mathbb{E}Z(t)=1$ for all $t \geq 0$.*

PROOF: We know from stochastic calculus for continuous processes that $Z_1(t)$ is a martingale and from Lemma 11.6.6 that $Z_2(t)$ is a martingale. Since $Z_1(t)$ is continuous and $Z_2(t)$ has no Itô integral part, $[Z_1,Z_2](t)=0$. Itô's product rule for jump processes thus implies

$$Z_1(t)Z_2(t) = Z_1(0)Z_2(0) + \int_0^t Z_1(s-)\,dZ_2(s) + \int_0^t Z_2(s-)\,dZ_1(s), \quad (11.6.34)$$

and both integrals are martingales because of Theorem 11.4.5. This implies that $Z(t)$ is a martingale, and because $Z(0) = 1$, we have $\mathbb{E}Z(t) = 1$ for all $t \geq 0$. □

Fix a positive $T$ and define $\widetilde{\mathbb{P}}(A) = \int_A Z(T)\,d\mathbb{P}$ for all $A \in \mathcal{F}$. We have the following.

**Theorem 11.6.9.** *Under the probability measure $\widetilde{\mathbb{P}}$, the process*

$$\widetilde{W}(t) = W(t) + \int_0^t \Theta(s)\,ds$$

*is a Brownian motion, $Q(t)$ is a compound Poisson process with intensity $\tilde{\lambda}$ and independent, identically distributed jump sizes having density $\widetilde{f}(y)$, and the processes $\widetilde{W}(t)$ and $Q(t)$ are independent.*

The key step in the proof of Theorem 11.6.9 is to show that $\widetilde{W}(t)$ and $Q(t)$ have the correct joint moment-generating function under $\widetilde{\mathbb{P}}$. In other words, we must show

$$\widetilde{\mathbb{E}}\left[e^{u_1\widetilde{W}(t)+u_2Q(t)}\right] = \exp\left\{\frac{1}{2}u_1^2 t\right\} \cdot \exp\left\{\tilde{\lambda}t\big(\widetilde{\varphi}_Y(u_2) - 1\big)\right\}, \quad (11.6.35)$$

where $\widetilde{\varphi}_Y(u_2)$ is given by (11.6.25). Since $e^{\frac{1}{2}u_1^2 t}$ is the moment-generating function for a normal random variable with mean zero and variance $t$, $\exp\left\{\tilde{\lambda}t\big(\widetilde{\varphi}_Y(u_2) - 1\big)\right\}$ is the moment-generating function for a compound Poisson process with intensity $\tilde{\lambda}$ and jump density $\widetilde{f}(y)$, and since the joint moment-generating function factors into the product of these two moment-generating functions, we would then know that $\widetilde{W}(t)$ and $Q(t)$ have the right distributions under $\widetilde{\mathbb{P}}$ and are independent.

If the process $\Theta(t)$ is independent of the process $Q(t)$, then $Z_1$ is independent of $Q$ and we can obtain (11.6.35) from the following independence-based computation:

$$\begin{aligned}\widetilde{\mathbb{E}}\left[e^{u_1\widetilde{W}(t)+u_2Q(t)}\right] &= \mathbb{E}\left[e^{u_1\widetilde{W}(t)}Z_1(t)\cdot e^{u_2Q(t)}Z_2(t)\right]\\ &= \mathbb{E}\left[e^{u_1\widetilde{W}(t)}Z_1(t)\right]\cdot\mathbb{E}\left[e^{u_2Q(t)}Z_2(t)\right].\end{aligned}$$

Girsanov's Theorem from stochastic calculus for continuous processes implies

$$\mathbb{E}\left[e^{u_1\widetilde{W}(t)}Z_1(t)\right] = \exp\left\{\frac{1}{2}u_1^2 t\right\},$$

and (11.6.30) implies

$$\mathbb{E}\left[e^{u_2Q(t)}Z_2(t)\right] = \exp\left\{\tilde{\lambda}t\big(\widetilde{\varphi}_Y(u_2) - 1\big)\right\}.$$

Equation (11.6.35) follows.

The surprising fact is that (11.6.35) and hence the conclusion of Theorem 11.6.9 hold even if $\Theta(t)$ is allowed to depend on $Q(t)$. Indeed, we could have $\Theta(t)$ *equal to* $Q(t)$. We give the proof of this fact.

PROOF OF (11.6.35): We define

$$X_1(t) = \exp\left\{u_1\widetilde{W}(t) - \frac{1}{2}u_1^2 t\right\},$$
$$X_2(t) = \exp\left\{u_2 Q(t) - \tilde{\lambda}t\big(\widetilde{\varphi}_Y(u_2) - 1\big)\right\},$$

and show below that $X_1(t)Z_1(t)$, $X_2(t)Z_2(t)$, and $X_1(t)Z_1(t)X_2(t)Z_2(t)$ are martingales under $\mathbb{P}$.

The Itô-Doeblin formula for continuous processes implies

$$\begin{aligned} dX_1(t) &= X_1(t)\Big(u_1\,d\widetilde{W}(t) - \frac{1}{2}u_1^2\,dt\Big) + \frac{1}{2}u_1^2 X_1(t)\,dt \\ &= u_1 X_1(t)\,d\widetilde{W}(t) \\ &= u_1 X_1(t)\,dW(t) + u_1\Theta(t)X_1(t)\,dt. \end{aligned}$$

The Itô-Doeblin formula also implies

$$dZ_1(t) = -\Theta(t)Z_1(t)\,dW(t).$$

Itô's product rule yields

$$\begin{aligned} d\big(X_1(t)Z_1(t)\big) &= X_1(t)\,dZ_1(t) + Z_1(t)\,dX_1(t) + dX_1(t)\,dZ_1(t) \\ &= -\Theta(t)X_1(t)Z_1(t)\,dW(t) + u_1X_1(t)Z_1(t)\,dW(t) \\ &\quad + u_1\Theta(t)X_1(t)Z_1(t)\,dt - u_1\Theta(t)X_1(t)Z_1(t)\,dt \\ &= \big(u_1 - \Theta(t)\big)X_1(t)Z_1(t)\,dW(t). \end{aligned}$$

Because its differential has no $dt$ term, $X_1(t)Z_1(t)$ is a martingale.

We showed in the proof of Theorem 11.6.7 that $X_2(t)Z_2(t)$ is a martingale.

Finally, because $X_1(t)Z_1(t)$ is continuous and $X_2(t)Z_2(t)$ has no Itô integral part, $[X_1Z_1, X_2Z_2](t) = 0$. Therefore, Itô's product rule implies

$$\begin{aligned} X_1(t)Z_1(t)X_2(t)Z_2(t) = 1 &+ \int_0^t X_1(s-)Z_1(s-)\,d\big(X_2(s)Z_2(s)\big) \\ &+ \int_0^t X_2(s-)Z_2(s-)\,d\big(X_1(s)Z_1(s)\big), \end{aligned}$$

and Theorem 11.4.5 implies that $X_1(t)Z_1(t)X_2(t)Z_2(t)$ is a martingale. It follows that

$$\mathbb{E}\big[X_1(t)Z_1(t)X_2(t)Z_2(t)\big] = 1;$$

this gives us (11.6.35). □

Suppose a compound Poisson process $Q(t)$ has jumps $Y_1, Y_2, \dots$ that take only finitely many nonzero values $y_1, y_2, \dots, y_M$, with $p(y_m) = \mathbb{P}\{Y_i = y_m\}$ so that $p(y_m) > 0$ and $\sum_{m=1}^M p_m = 1$. Let $\tilde{\lambda}$ be a positive constant and let $\tilde{p}(y_1), \dots, \tilde{p}(y_M)$ be positive numbers that sum to 1. In place of (11.6.32), we now define

$$Z_2(t) = e^{(\lambda-\tilde{\lambda})t} \prod_{i=1}^{N(t)} \frac{\tilde{\lambda}\tilde{p}(Y_i)}{\lambda p(Y_i)}$$

and then define $Z(t)$ by (11.6.33). Lemma 11.6.8 still applies and permits us to define the probability measure $\widetilde{\mathbb{P}}$ by the formula $\widetilde{\mathbb{P}}(A) = \int_A Z(T)\, d\mathbb{P}$ for all $Z \in \mathcal{F}$. A straightforward modification of the proof of Theorem 11.6.9 gives the following result.

**Theorem 11.6.10.** *Under the probability measure $\widetilde{\mathbb{P}}$, the process*

$$\widetilde{W}(t) = W(t) + \int_0^t \Theta(s)\, ds$$

*is a Brownian motion, $Q(t)$ is a compound Poisson process with intensity $\tilde{\lambda}$ and independent, identically distributed jump sizes satisfying $\widetilde{\mathbb{P}}\{Y_i = y_m\} = \tilde{p}(y_m)$ for all $i$ and $m = 1, \dots, M$, and the processes $\widetilde{W}(t)$ and $Q(t)$ are independent.*

## 11.7 Pricing a European Call in a Jump Model

In this section, we consider the problem of pricing a European call when the underlying asset is a jump process. We work out the details for two cases: (1) the underlying asset is driven by a single Poisson process, and (2) the underlying asset is driven by a Brownian motion and a compound Poisson process. The market is complete in the first case and incomplete in the second. We discuss the nature of the incompleteness in the second case.

### 11.7.1 Asset Driven by a Poisson Process

We return to Example 11.6.3, in which the underlying asset price is given by

$$\begin{aligned} S(t) &= S(0) \exp\left\{\alpha t + N(t) \log(\sigma + 1) - \lambda\sigma t\right\} \\ &= S(0) e^{(\alpha - \lambda\sigma)t} (\sigma + 1)^{N(t)}, \end{aligned} \tag{11.7.1}$$

for which the differential is

$$dS(t) = \alpha S(t)\, dt + \sigma S(t-)\, dM(t).$$

In this model, $N(t)$ is a Poisson process with intensity $\lambda > 0$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and $M(t) = N(t) - \lambda t$ is the compensated Poisson process.

We fix a positive time $T$ and wish to price a European call whose payoff at time $T$ is

$$V(T) = \big(S(T) - K\big)^+.$$

We saw in Example 11.6.3 that we must assume $\lambda > \frac{\alpha - r}{\sigma}$ in order to rule out arbitrage. Under this assumption,

$$\tilde{\lambda} = \lambda - \frac{\alpha - r}{\sigma}$$

is positive, and there is a risk-neutral measure given by

$$\widetilde{\mathbb{P}}(A) = \int_A Z(T)\, d\mathbb{P} \text{ for all } A \in \mathcal{F},$$

where $Z(t) = e^{(\lambda - \tilde{\lambda})t}\left(\frac{\tilde{\lambda}}{\lambda}\right)^{N(t)}$. This risk-neutral measure is in fact unique; see Remark 11.7.2 below.

Under the risk-neutral measure, the compensated Poisson process $\widetilde{M}(t) = N(t) - \tilde{\lambda}t$ is a martingale, and

$$dS(t) = rS(t)\, dt + \sigma S(t-)\, d\widetilde{M}(t) \tag{11.7.2}$$

or, equivalently,

$$d\big(e^{-rt}S(t)\big) = \sigma e^{-rt} S(t-)\, d\widetilde{M}(t).$$

The discounted asset price is a martingale under $\widetilde{\mathbb{P}}$. In terms of $\tilde{\lambda}$, we may rewrite the second line in (11.7.1) as

$$S(t) = S(0)e^{(r - \tilde{\lambda}\sigma)t}(\sigma + 1)^{N(t)}.$$

For $0 \le t \le T$, let $V(t)$ denote the risk-neutral price of a European call paying $V(T) = \big(S(T) - K\big)^+$ at time $T$. The discounted call price is a martingale under the risk-neutral measure. In other words, the call price $V(t)$ satisfies

$$e^{-rt}V(t) = \widetilde{\mathbb{E}}\big[e^{-rT}V(T)\big|\mathcal{F}(t)\big] = \widetilde{\mathbb{E}}\big[e^{-rT}\big(S(T) - K\big)^+\big|\mathcal{F}(t)\big].$$

We have

$$\begin{aligned} S(T) &= S(0)e^{(r-\tilde{\lambda}\sigma)t}(\sigma + 1)^{N(t)} \cdot e^{(r-\tilde{\lambda}\sigma)(T-t)}(\sigma + 1)^{N(T)-N(t)} \\ &= S(t) \cdot e^{(r-\tilde{\lambda}\sigma)(T-t)}(\sigma + 1)^{N(T)-N(t)}. \end{aligned}$$

It follows that

$$\begin{aligned} V(t) &= \widetilde{\mathbb{E}}\big[e^{-r(T-t)}\big(S(T) - K\big)^+\big|\mathcal{F}(t)\big] \\ &= \widetilde{\mathbb{E}}\Big[e^{-r(T-t)}\Big(S(t)e^{(r-\tilde{\lambda}\sigma)(T-t)}(\sigma + 1)^{N(T)-N(t)} - K\Big)^+\Big|\mathcal{F}(t)\Big]. \end{aligned}$$

The random variable $S(t)$ is $\mathcal{F}(t)$-measurable, whereas

$$e^{(r-\tilde{\lambda}\sigma)(T-t)}(\sigma+1)^{N(T)-N(t)}$$

is independent of $\mathcal{F}(t)$. According to the Independence Lemma, Lemma 2.3.4,

$$V(t) = c\big(t, S(t)\big),$$

where

$$\begin{aligned} c(t,x) &= \widetilde{\mathbb{E}}\Big[e^{-r(T-t)}\Big(xe^{(r-\tilde{\lambda}\sigma)(T-t)}(\sigma+1)^{N(T)-N(t)} - K\Big)^+\Big] \\ &= \sum_{j=0}^{\infty} e^{-r(T-t)}\Big(xe^{(r-\tilde{\lambda}\sigma)(T-t)}(\sigma+1)^j - K\Big)^+ \frac{\tilde{\lambda}^j(T-t)^j}{j!}e^{-\tilde{\lambda}(T-t)} \\ &= \sum_{j=0}^{\infty} \Big(xe^{-\tilde{\lambda}\sigma(T-t)}(\sigma+1)^j - Ke^{-r(T-t)}\Big)^+ \frac{\tilde{\lambda}^j(T-t)^j}{j!}e^{-\tilde{\lambda}(T-t)}. \end{aligned} \tag{11.7.3}$$

From this formula, the risk-neutral price of the call $c(t,x)$ can be computed. The $j=0$ term in (11.7.3) is

$$\big(xe^{-\tilde{\lambda}\sigma(T-t)} - Ke^{-r(T-t)}\big)^+ e^{-\tilde{\lambda}(T-t)}.$$

When $t=T$, this term is $(x-K)^+$, and it is the only nonzero term in the sum in (11.7.3) when $t=T$. Therefore, the function $c$ satisfies the terminal condition

$$c(T,x) = (x-K)^+ \text{ for all } x \geq 0. \tag{11.7.4}$$

We next derive the "partial differential equation" that $c(t,x)$ must satisfy. The usual iterated conditioning argument shows that

$$e^{-rt}c\big(t,S(t)\big) = e^{-rt}V(t) = \widetilde{\mathbb{E}}\big[e^{-rT}\big(S(T)-K\big)^+\big|\mathcal{F}(t)\big]$$

is a martingale under $\widetilde{\mathbb{P}}$. Therefore, we compute $d\big(e^{-rt}c(t,S(t))\big)$ and set the "$dt$" term equal to zero. The stochastic differential equation (11.7.2) may be rewritten as

$$dS(t) = (r-\tilde{\lambda}\sigma)S(t)\,dt + \sigma S(t-)\,dN(t), \tag{11.7.5}$$

which shows that the continuous part of the stock price satisfies

$$dS^c(t) = (r-\tilde{\lambda}\sigma)S(t)\,dt.$$

On the other hand, if the stock price jumps at time $t$, then

$$\varDelta S(t) = S(t) - S(t-) = \sigma S(t-), \quad S(t) = (\sigma+1)S(t-).$$

The Itô-Doeblin formula implies

$$
\begin{aligned}
&e^{-rt}c(t,S(t))\\
&\quad= c(0,S(0)) + \int_0^t e^{-ru}\big[-rc(u,S(u))\,du + c_t(u,S(u))\,du\\
&\qquad\qquad\qquad +c_x(u,S(u))\,dS^c(u)\big]\\
&\qquad + \sum_{0<u\le t} e^{-ru}\big[c(u,S(u)) - c(u,S(u-))\big]\\
&\quad= c(0,S(0)) + \int_0^t e^{-ru}\big[-rc(u,S(u)) + c_t(u,S(u))\\
&\qquad\qquad\qquad +(r-\tilde\lambda\sigma)S(u)c_x(u,S(u))\big]\,du\\
&\qquad + \int_0^t e^{-ru}\big[c(u,(\sigma+1)S(u-)) - c(u,S(u-))\big]\,dN(u)\\
&\quad= c(0,S(0)) + \int_0^t e^{-ru}\big[-rc(u,S(u)) + c_t(u,S(u))\\
&\qquad\qquad\qquad +(r-\tilde\lambda\sigma)S(u)c_x(u,S(u))\big]\,du\\
&\qquad + \int_0^t e^{-ru}\big[c(u,(\sigma+1)S(u-)) - c(u,S(u-))\big]\tilde\lambda\,du\\
&\qquad + \int_0^t e^{-ru}\big[c(u,(\sigma+1)S(u-)) - c(u,S(u-))\big]\,d\widetilde{M}(u).
\end{aligned}
$$

However, the integral

$$\int_0^t e^{-ru}\big[c(u,(\sigma+1)S(u-)) - c(u,S(u-))\big]\tilde\lambda\,du$$

is the same as the integral

$$\int_0^t e^{-ru}\big[c(u,(\sigma+1)S(u)) - c(u,S(u))\big]\tilde\lambda\,du.$$

We have shown that

$$
\begin{aligned}
&e^{-rt}c(t,S(t))\\
&\quad= c(0,S(0))\\
&\qquad + \int_0^t e^{-ru}\big[-rc(u,S(u)) + c_t(u,S(u)) + (r-\tilde\lambda\sigma)S(u)c_x(u,S(u))\\
&\qquad\qquad\qquad +\tilde\lambda\big(c(u,(\sigma+1)S(u)) - c(u,S(u))\big)\big]\,du\\
&\qquad + \int_0^t e^{-ru}\big[c(u,(\sigma+1)S(u-)) - c(u,S(u-))\big]\,d\widetilde{M}(u). \qquad (11.7.6)
\end{aligned}
$$

The last integral is a martingale because the integrator $\widetilde{M}(u)$ is a martingale and the integrand is left-continuous. Because left-hand side of (11.7.6), $e^{-rt}c(t,S(t))$, is also a martingale we can then solve for

$$c(0,S(0)) + \int_0^t e^{-ru}\big[-rc(u,S(u)) + c_t(u,S(u)) + (r-\tilde{\lambda}\sigma)S(u)c_x(u,S(u)) \\ +\tilde{\lambda}\big(c(u,(\sigma+1)S(u)) - c(u,S(u))\big)\big]\,du$$

and see that it is the difference of two martingales and hence is itself a martingale. This can only happen if the integrand is zero:

$$-rc(t,S(t)) + c_t(t,S(t)) + (r-\tilde{\lambda}\sigma)S(t)c_x(t,S(t)) \\ +\tilde{\lambda}\big(c(t,(\sigma+1)S(t)) - c(t,S(t))\big) = 0. \quad (11.7.7)$$

The way we have in the past argued for (11.7.7) using (11.7.6) (see the discussion preceding Theorem 6.4.3) is by first taking the differential in (11.7.6) to obtain

$$d\big(e^{-rt}c(t,S(t))\big) \\ = e^{-rt}\big[-rc(t,S(t)) + c_t(t,S(t)) + (r-\tilde{\lambda}\sigma)S(t)c_x(t,S(t)) \\ +\tilde{\lambda}\big(c(t,(\sigma+1)S(t)) - c(t,S(t))\big)\big]\,dt \\ +e^{-rt}\big[c(t,(\sigma+1)S(t-)) - c(t,S(t-))\big]\,d\widetilde{M}(t)$$

and then setting the $dt$ term equal to zero. This still works, provided we make sure the non-$dt$ term has a martingale integrator, and if this integrator has jumps, then the integrand for this martingale is left-continuous. In particular, we also have

$$d\big(e^{-rt}c(t,S(t))\big) \\ = e^{-rt}\big[-rc(t,S(t)) + c_t(t,S(t)) + (r-\tilde{\lambda}\sigma)S(t)c_x(t,S(t))\big]\,dt \\ +e^{-rt}\big[c(t,(\sigma+1)S(t-)) - c(t,S(t-))\big]\,dN(t), \quad (11.7.8)$$

but setting the "$dt$" term

$$e^{-rt}\big[-rc(t,S(t)) + c_t(t,S(t)) + (r-\tilde{\lambda}\sigma)S(t)c_x(t,S(t))\big]\,dt$$

in this expression equal to zero gives an incorrect result because the non-$dt$ term has integrator $dN(t)$ and $N(t)$ is not a martingale.

We conclude by replacing the stock price process $S(t)$ in (11.7.7) by a dummy variable $x$. This gives the equation

$$-rc(t,x)+c_t(t,x)+(r-\tilde{\lambda}\sigma)xc_x(t,x)+\tilde{\lambda}\big(c(t,(\sigma+1)x)-c(t,x)\big) = 0, \quad (11.7.9)$$

which must hold for $0 \le t < T$ and $x \ge 0$. This is sometimes called a *differential-difference* equation because it involves $c$ at two different values of the stock price, namely $x$ and $(\sigma+1)x$. The function $c(t,x)$ defined by (11.7.3) satisfies this equation because, by its construction, $e^{-rt}c(t,S(t))$ is a martingale under $\widetilde{\mathbb{P}}$.

Returning to (11.7.6) and using equation (11.7.9), we see that for $0 \le t \le T$,

$$\begin{aligned} & e^{-rt}c(t,S(t)) \\ & = c(0,S(0)) + \int_0^t e^{-ru}\big[c(u,(\sigma+1)S(u-)) - c(u,S(u-))\big]\,d\widetilde{M}(u). \quad (11.7.10) \end{aligned}$$

In particular,

$$\begin{aligned} & e^{-rT}\big(S(T)-K\big)^+ \\ & = e^{-rT}c(T,S(T)) \\ & = c(0,S(0)) + \int_0^T e^{-ru}\big[c(u,(\sigma+1)S(u-)) - c(u,S(u-))\big]\,d\widetilde{M}(u). \quad (11.7.11) \end{aligned}$$

We use this observation to construct the hedge for a short position in the call.

Suppose we sell the call at time zero in exchange for initial capital $X(0) = c(0,S(0))$. We want to invest in the stock and money market account so that $X(t) = c(t,S(t))$ for all $t$ or, equivalently,

$$e^{-rt}X(t) = e^{-rt}c(t,S(t)) \text{ for all } t \in [0,T].$$

To accomplish this, we match differentials. From (11.7.10), we see that the differential of $e^{-rt}c(t,S(t))$ is

$$d\big(e^{-rt}c(t,S(t))\big) = e^{-rt}\big[c(t,(\sigma+1)S(t-)) - c(t,S(t-))\big]\,d\widetilde{M}(t). \quad (11.7.12)$$

The differential of the value $X(t)$ of a portfolio that at each time $t$ holds $\Gamma(t)$ shares of stock (we use $\Gamma(t)$ rather than $\Delta(t)$ to denote the number of shares of stock held in the hedging portfolio to avoid confusion with the use of $\Delta$ as the size of the jump in a process) is

$$dX(t) = \Gamma(t-)\,dS(t) + r\big[X(t) - \Gamma(t)S(t)\big]\,dt.$$

Therefore,

$$\begin{aligned} d\big(e^{-rt}X(t)\big) &= e^{-rt}[-rX(t)\,dt + dX(t)] \\ &= e^{-rt}[\Gamma(t-)\,dS(t) - r\Gamma(t)S(t)\,dt] \\ &= e^{-rt}\sigma\Gamma(t-)S(t-)\,d\widetilde{M}(t), \quad (11.7.13) \end{aligned}$$

where we have used (11.7.2) in the last step. We are interested in determining the value of $\Gamma(t-)$, the position held just before any jump that may occur at time $t$. Comparing (11.7.12) and (11.7.13), we conclude that we should take

$$\Gamma(t-) = \frac{c(t,(\sigma+1)S(t-)) - c(t,S(t-))}{\sigma S(t-)}. \quad (11.7.14)$$

This is the hedging position we should hold at all times, whether they are jump times or not. More specifically, if we define

$$\Gamma(t) = \frac{c(t,(\sigma+1)S(t)) - c(t,S(t))}{\sigma S(t)} \text{ for all } t \in [0,T], \tag{11.7.15}$$

then (11.7.14) will also hold and integration of (11.7.13) yields

$$\begin{aligned} &e^{-rt}X(t) \\ &= X(0) + \int_0^t e^{-ru}\big[c(u,(\sigma+1)S(u-)) - c(u,S(u-))\big]\, d\widetilde{M}(u). \end{aligned} \tag{11.7.16}$$

Comparison of (11.7.10) with (11.7.16) shows that $X(t) = c(t, S(t))$ for all $t$. In particular, (11.7.11) shows that $X(T) = \big(S(T) - K\big)^+$; the short position in the European call has been hedged.

*Remark 11.7.1 (Sanity check).* To convince ourselves that the hedge (11.7.15) really works, we consider separately the cases when the stock jumps at time $t$ and when the stock does not jump at time $t$. In the event of a jump, the change in the option price is $c(t,(\sigma+1)S(t-)) - c(t,S(t-))$. The change in the hedging portfolio value is

$$\Gamma(t-)\big(S(t) - S(t-)\big) = \Gamma(t-)\sigma S(t-) = c(t,(\sigma+1)S(t-)) - c(t,S(t-)),$$

which agrees with the change in the option price.

On the other hand, if the stock price does not jump at time $t$, then the stock price follows equation (11.7.5) without the $dN(t)$ term at time $t$:

$$dS(t) = (r - \tilde{\lambda}\sigma)S(t)\,dt.$$

At this time, (11.7.8) shows that the discounted option price has the differential

$$\begin{aligned} &d\big(e^{-rt}c(t,S(t))\big) \\ &= e^{-rt}\big[-rc(t,S(t)) + c_t(t,S(t)) + (r-\tilde{\lambda}\sigma)S(t)c_x(t,S(t))\big]\,dt \\ &= -e^{-rt}\tilde{\lambda}\big[c(t,(\sigma+1)S(t)) - c(t,S(t))\big]\,dt, \end{aligned}$$

where we have used the differential-difference equation (11.7.9) to obtain the second equality. The differential of the discounted portfolio value at this time is (from (11.7.13) without the $dN(t)$ term implicit in $d\widetilde{M}(t)$)

$$\begin{aligned} d\big(e^{-rt}X(t)\big) &= e^{-rt}\sigma\Gamma(t)S(t)(-\tilde{\lambda}\,dt) \\ &= -e^{-rt}\tilde{\lambda}\big[c(t,(\sigma+1)S(t)) - c(t,S(t))\big]\,dt. \end{aligned}$$

Once again, the discounted portfolio value tracks the discounted option price.
□

*Remark 11.7.2 (Completeness).* In this subsection, we have constructed the price and hedge for a European call on a stock driven by a single Poisson process. It is clear from the analysis that this same argument would work for an arbitrary European derivative security with payoff $h(S(T))$ at time $T$ written on a stock modeled this way. One could simply replace the call payoff by the function $h$ in equation (11.7.3). The differential-difference equation (11.7.9) would still apply, although now with terminal condition $c(T,x) = h(x)$ replacing (11.7.4), and the hedging formula (11.7.15) would still be correct.

The model is complete and the risk-neutral measure is unique if and only if every derivative security can be hedged (Second Fundamental Theorem of Asset Pricing, Theorem 5.4.9). "Every" derivative security means also those derivative securities that are path-dependent. We have not considered path-dependent derivative securities in this subsection, but one can show that they also can be hedged, and thus the model is complete.

### 11.7.2 Asset Driven by a Brownian Motion and a Compound Poisson Process

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which is defined a Brownian motion $W(t)$, $0 \le t \le T$, and $M$ independent Poisson processes $N_1(t), \dots, N_M(t)$, $0 \le t \le T$. Let $\mathcal{F}(t)$, $0 \le t \le T$, be the filtration generated by the Brownian motion and the $M$ Poisson processes.

Let $\lambda_m > 0$ be the intensity of the $m$th Poisson process and let $-1 < y_1 < \cdots < y_M$ be nonzero numbers. Set

$$N(t) = \sum_{m=1}^{M} N_m(t), \quad Q(t) = \sum_{m=1}^{M} y_m N_m(t).$$

Then $N$ is a Poisson process with intensity $\lambda = \sum_{m=1}^{M} \lambda_m$ and $Q$ is a compound Poisson process. Let $Y_i$ denote the size of the $i$th jump of $Q$. Then the $Y_i$ random variables take values in the set $\{y_1, \dots, y_M\}$, and $Q(t)$ can be written as

$$Q(t) = \sum_{i=1}^{N(t)} Y_i.$$

Define

$$p(y_m) = \frac{\lambda_m}{\lambda}.$$

The random variables $Y_1, Y_2, \dots$ are independent and identically distributed, with $\mathbb{P}\{Y_i = y_m\} = p(y_m)$. These assertions all follow from Theorem 11.3.3.

Set

$$\beta = \mathbb{E}Y_i = \sum_{m=1}^{M} y_m p(y_m) = \frac{1}{\lambda} \sum_{m=1}^{M} \lambda_m y_m. \tag{11.7.17}$$

According to Theorem 11.3.1,

$$Q(t) - \beta\lambda t = Q(t) - t\sum_{m=1}^{M} \lambda_m y_m$$

is a martingale.

In this subsection, the stock price will be modeled by the stochastic differential equation

$$\begin{aligned} dS(t) &= \alpha S(t)\,dt + \sigma S(t)\,dW(t) + S(t-)d\big(Q(t) - \beta\lambda t\big) \\ &= (\alpha - \beta\lambda)S(t)\,dt + \sigma S(t)\,dW(t) + S(t-)\,dQ(t). \end{aligned} \tag{11.7.18}$$

Under the original probability measure $\mathbb{P}$, the mean rate of return on the stock is $\alpha$. The assumption that $y_i > -1$ for $i = 1, \ldots, M$ guarantees that although the stock price can jump down, it cannot jump from a positive to a negative value or to zero. We begin with a positive initial stock price $S(0)$, and the stock price is positive at all subsequent times; see (11.7.19) below. If $S(0) = 0$, then $S(t) = 0$ for all $t$.

**Theorem 11.7.3.** *The solution to (11.7.18) is*

$$S(t) = S(0)\exp\left\{\sigma W(t) + \left(\alpha - \beta\lambda - \frac{1}{2}\sigma^2\right)t\right\}\prod_{i=1}^{N(t)}(Y_i + 1). \tag{11.7.19}$$

PROOF: We show that $S(t)$ defined by the right-hand side of (11.7.19) satisfies the stochastic differential equation (11.7.18). Toward this end, define the continuous stochastic process

$$X(t) = S(0)\exp\left\{\sigma W(t) + \left(\alpha - \beta\lambda - \frac{1}{2}\sigma^2\right)t\right\}$$

and the pure jump process

$$J(t) = \prod_{i=1}^{N(t)}(Y_i + 1).$$

Then $S(t) = X(t)J(t)$. We show that $S(t) = X(t)J(t)$ is a solution to the stochastic differential equation (11.7.18).

The Itô-Doeblin formula for a continuous process says that

$$dX(t) = (\alpha - \beta\lambda)X(t)\,dt + \sigma X(t)\,dW(t). \tag{11.7.20}$$

At the time of the $i$th jump, $J(t) = J(t-)(Y_i + 1)$ and hence

$$\Delta J(t) = J(t) - J(t-) = J(t-)Y_i = J(t-)\Delta Q(t).$$

The equation $\Delta J(t) = J(t-)\Delta Q(t)$ also holds at nonjump times, with both sides equal to zero. Therefore,

$$dJ(t) = J(t-)\,dQ(t). \tag{11.7.21}$$

Itô's product rule for jump processes implies that

$$S(t) = X(t)J(t) = S(0) + \int_0^t X(s-)\,dJ(s) + \int_0^t J(s)\,dX(s) + [X,J](t). \tag{11.7.22}$$

Since $J$ is a pure jump process and $X$ is continuous, $[X,J](t) = 0$. Substituting (11.7.20) and (11.7.21) into (11.7.22), we obtain

$$\begin{aligned} S(t) &= X(t)J(t) \\ &= S(0) + \int_0^t X(s-)J(s-)\,dQ(s) + (\alpha - \beta\lambda)\int_0^t J(s)X(s)\,ds \\ &\quad +\sigma\int_0^t J(s)X(s)\,dW(s), \end{aligned}$$

which in differential form is

$$\begin{aligned} dS(t) &= d\big(X(t)J(t)\big) \\ &= X(t-)J(t-)\,dQ(t) + (\alpha - \beta\lambda)J(t)X(t)\,dt + \sigma J(t)X(t)\,dW(t) \\ &= S(t-)\,dQ(t) + (\alpha - \beta\lambda)S(t)\,dt + \sigma S(t)\,dW(t). \end{aligned}$$

This is (11.7.18). □

We now undertake to construct a risk-neutral measure. Let $\theta$ be a constant and let $\tilde{\lambda}_1, \dots, \tilde{\lambda}_M$ be positive constants.[2] Define

$$\begin{aligned} Z_0(t) &= \exp\left\{ -\theta W(t) - \frac{1}{2}\theta^2 t \right\}, \\ Z_m(t) &= e^{(\lambda_m - \tilde{\lambda}_m)t} \left( \frac{\tilde{\lambda}_m}{\lambda_m} \right)^{N_m(t)}, \quad m = 1, \dots, M, \\ Z(t) &= Z_0(t) \prod_{m=1}^M Z_m(t), \\ \widetilde{\mathbb{P}}(A) &= \int_A Z(T)\,d\mathbb{P} \text{ for all } A \in \mathcal{F}. \end{aligned}$$

The following assertions follow from Theorem 11.6.10 and Corollary 11.3.4. Independence under $\widetilde{\mathbb{P}}$ between $\widetilde{W}$ and each of the Poisson processes $N_m$, asserted in (iii) below, follows from Corollary 11.5.3. Under the probability measure $\widetilde{\mathbb{P}}$,

[2] One could create more risk-neutral measures than we consider here by letting $\theta$ and $\tilde{\lambda}_1, \dots, \tilde{\lambda}_M$ be adapted stochastic processes.

(i) the process
$$\widetilde{W}(t) = W(t) + \theta t \tag{11.7.23}$$
is a Brownian motion,
(ii) each $N_m$ is a Poisson process with intensity $\tilde{\lambda}_m$, and
(iii) $\widetilde{W}$ and $N_1, \dots, N_m$ are independent of one another.

Define
$$\tilde{\lambda} = \sum_{m=1}^{M} \tilde{\lambda}_m, \quad \tilde{p}(y_m) = \frac{\tilde{\lambda}_m}{\tilde{\lambda}}.$$
Under $\widetilde{\mathbb{P}}$, the process $N(t) = \sum_{m=1}^{M} N_m(t)$ is Poisson with intensity $\tilde{\lambda}$, the jump-size random variables $Y_1, Y_2, \dots$ are independent and identically distributed with $\widetilde{\mathbb{P}}\{Y_i = y_m\} = \tilde{p}(y_m)$, and $Q(t) - \tilde{\beta}\tilde{\lambda}t$ is a martingale, where
$$\tilde{\beta} = \widetilde{\mathbb{E}}Y_i = \sum_{m=1}^{M} y_m \tilde{p}(y_m) = \frac{1}{\tilde{\lambda}} \sum_{m=1}^{M} \tilde{\lambda}_m y_m.$$

The probability measure $\widetilde{\mathbb{P}}$ is risk-neutral if and only if the mean rate of return of the stock under $\widetilde{\mathbb{P}}$ is the interest rate $r$. In other words, $\widetilde{\mathbb{P}}$ is risk-neutral if and only if
$$\begin{aligned} dS(t) &= (\alpha - \beta\lambda)S(t)\,dt + \sigma S(t)\,dW(t) + S(t-)\,dQ(t) \\ &= rS(t)\,dt + \sigma S(t)\,d\widetilde{W}(t) + S(t-)d(Q(t) - \tilde{\beta}\tilde{\lambda}t). \end{aligned} \tag{11.7.24}$$
This is equivalent to the equation
$$\alpha - \beta\lambda = r + \sigma\theta - \tilde{\beta}\tilde{\lambda}, \tag{11.7.25}$$
which is the *market price of risk equation* for this model. Recalling the definitions of $\beta$ and $\tilde{\beta}$, we may rewrite the market price of risk equation (11.7.25) as
$$\begin{aligned} \alpha - r &= \sigma\theta + \beta\lambda - \tilde{\beta}\tilde{\lambda} \\ &= \sigma\theta + \sum_{m=1}^{M} (\lambda_m - \tilde{\lambda}_m) y_m. \end{aligned} \tag{11.7.26}$$
Because there is one equation and $M + 1$ unknowns, $\theta, \lambda_1, \dots, \lambda_M$, there are multiple risk-neutral measures.

Extra stocks would help determine a unique risk-neutral measure. We illustrate this point by taking $M = 2$ in the following example.

*Example 11.7.4 (Three stocks and two Poisson processes).* With one Brownian motion $W$ and two independent Poisson processes $N_1$ and $N_2$, define three compound Poisson processes

$$Q_i(t) = y_{i,1}N_1(t) + y_{i,2}N_2(t), \ i = 1, 2, 3,$$

where $y_{i,m} > -1$ for $i = 1, 2, 3$ and $m = 1, 2$. Set

$$\beta_i = \frac{1}{\lambda}(\lambda_1 y_{i,1} + \lambda_2 y_{i,2}), \ i = 1, 2, 3,$$

where $\lambda_1$ and $\lambda_2$ are the intensities of $N_1$ and $N_2$, respectively, under the original measure $\mathbb{P}$. For $i = 1, 2, 3$, we have a stock process modeled by

$$dS_i(t) = (\alpha_i - \beta_i\lambda)S_i(t)\,dt + \sigma_i S_i(t)\,dW(t) + S_i(t-)\,dQ_i(t).$$

In this model, there is a market price of risk equation analogous to (11.7.26) for each stock. The market price of risk equations are

$$\begin{aligned}
\alpha_1 - r &= \sigma_1\theta + (\lambda_1 - \tilde{\lambda}_1)y_{1,1} + (\lambda_2 - \tilde{\lambda}_2)y_{1,2},\\
\alpha_2 - r &= \sigma_2\theta + (\lambda_1 - \tilde{\lambda}_1)y_{2,1} + (\lambda_2 - \tilde{\lambda}_2)y_{2,2},\\
\alpha_3 - r &= \sigma_3\theta + (\lambda_1 - \tilde{\lambda}_1)y_{3,1} + (\lambda_2 - \tilde{\lambda}_2)y_{3,2}.
\end{aligned}$$

These are three equations in the three unknowns $\theta$, $\tilde{\lambda}_1$, and $\tilde{\lambda}_2$. If they have a unique solution, then there is a unique risk-neutral measure. In that case, the market would be complete and free of arbitrage. □

We return to the discussion of the model with a single stock given by (11.7.18) and (11.7.19). Let us choose some $\theta$ and $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_M$ satisfying the market price of risk equations (11.7.26). Then, in the notation of (11.7.24), we have

$$\begin{aligned}
dS(t) &= rS(t) + \sigma S(t)\,d\widetilde{W}(t) + S(t-)d(Q(t) - \tilde{\beta}\tilde{\lambda}t)\\
&= (r - \tilde{\beta}\tilde{\lambda})\,dt + \sigma S(t)\,d\widetilde{W}(t) + S(t-)dQ(t).
\end{aligned} \tag{11.7.27}$$

This is like equation (11.7.18), and just as (11.7.19) is the solution to (11.7.18), the solution to (11.7.27) is

$$S(t) = S(0)\exp\left\{\sigma\widetilde{W}(t) + \left(r - \tilde{\beta}\tilde{\lambda} - \frac{1}{2}\sigma^2\right)t\right\}\prod_{i=1}^{N(t)}(Y_i + 1). \tag{11.7.28}$$

Indeed, it is a straightforward matter to use (11.7.25) to verify that (11.7.19) and (11.7.28) are in fact the same equation. We have not changed the stock price process; we have changed only its distribution.

We compute the risk-neutral price of a call on the stock with price process given by (11.7.28). Because $\theta$ does not appear explicitly in (11.7.28), it will not appear in our pricing formula. However,

$$\tilde{\beta}\tilde{\lambda} = \sum_{m=1}^{M}\tilde{\lambda}_m y_m$$

will appear in this formula, and we can choose the risk-neutral intensities $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_M$ to be any positive constants and subsequently choose $\theta$ so that the market price of risk equation (11.7.25) is satisfied. We assume for the remainder of this section that some choice has been made. Our pricing formula will depend on the choice. It is common to use these free parameters to calibrate the model to market data.

For the next step, we need some notation. Define

$$\kappa(\tau, x) = xN(d_+(\tau, x)) - Ke^{-r\tau}N(d_-(\tau, x)), \tag{11.7.29}$$

where

$$d_\pm(\tau, x) = \frac{1}{\sigma\sqrt{\tau}}\left[\log\frac{x}{K} + \left(r \pm \frac{1}{2}\sigma^2\right)\tau\right]$$

and

$$N(y) = \frac{1}{\sqrt{2\pi}}\int_{\infty}^{y} e^{-\frac{1}{2}z^2}\,dz$$

is the cumulative standard normal distribution function. In other words, $\kappa(\tau, x)$ is the standard Black-Scholes-Merton call price on a geometric Brownian motion with volatility $\sigma$ when the current stock price is $x$, the expiration date is $\tau$ time units in the future, the interest rate is $r$, and the strike price is $K$. We have

$$\kappa(\tau, x) = \widetilde{\mathbb{E}}\left[e^{-r\tau}\left(x\exp\left\{-\sigma\sqrt{\tau}Y + \left(r - \frac{1}{2}\sigma^2\right)\tau\right\} - K\right)^+\right],$$

where $Y$ is a standard normal random variable under $\widetilde{\mathbb{P}}$; see Subsection 5.2.5.

**Theorem 11.7.5.** *For $0 \le t < T$, the risk-neutral price of a call,*

$$V(t) = \widetilde{\mathbb{E}}\left[e^{-r(T-t)}(S(T) - K)^+ \big| \mathcal{F}(t)\right],$$

*is given by $V(t) = c(t, S(t))$, where*

$$c(t, x) = \sum_{j=0}^{\infty} e^{-\tilde{\lambda}(T-t)}\frac{\tilde{\lambda}^j(T-t)^j}{j!}\,\widetilde{\mathbb{E}}\,\kappa\left(T-t, xe^{-\tilde{\beta}\tilde{\lambda}(T-t)}\prod_{i=1}^{j}(Y_i+1)\right). \tag{11.7.30}$$

PROOF: Let $t \in [0, T)$ be given and define $\tau = T - t$. From (11.7.28), we see that

$$S(T) = S(t)\exp\left\{\sigma(\widetilde{W}(T) - \widetilde{W}(t)) + \left(r - \tilde{\beta}\tilde{\lambda} - \frac{1}{2}\sigma^2\right)\tau\right\}\prod_{i=N(t)+1}^{N(T)}(Y_i + 1). \tag{11.7.31}$$

The term $S(t)$ is $\mathcal{F}(t)$-measurable, and the other term appearing on the right-hand side of (11.7.31) is independent of $\mathcal{F}(t)$. Therefore, the Independence Lemma, Lemma 2.3.4, implies that

$$V(t) = \widetilde{\mathbb{E}}\big[e^{-r\tau}(S(T)-K)^{+}\big|\mathcal{F}(t)\big] = c(t, S(t)),$$

where

$$\begin{aligned}
&c(t,x)\\
&= \widetilde{\mathbb{E}}\Bigg[e^{-r\tau}\Bigg(x\exp\Big\{\sigma(\widetilde{W}(T)-\widetilde{W}(t)) + \Big(r-\tilde{\beta}\tilde{\lambda}-\frac{1}{2}\sigma^2\Big)\tau\Big\}\\
&\qquad\qquad\qquad \times \prod_{i=N(t)+1}^{N(T)}(Y_i+1) - K\Bigg)^{+}\Bigg]\\
&= \widetilde{\mathbb{E}}\Bigg[\widetilde{\mathbb{E}}\Bigg[e^{-r\tau}\Bigg(x\exp\Big\{\sigma(\widetilde{W}(T)-\widetilde{W}(t)) + \Big(r-\tilde{\beta}\tilde{\lambda}-\frac{1}{2}\sigma^2\Big)\tau\Big\}\\
&\qquad\qquad\qquad \times \prod_{i=N(t)+1}^{N(T)}(Y_i+1) - K\Bigg)^{+}\Bigg|\sigma\Bigg(\prod_{i=N(t)+1}^{N(T)}(Y_i+1)\Bigg)\Bigg]\Bigg]\\
&= \widetilde{\mathbb{E}}\Bigg[\widetilde{\mathbb{E}}\Bigg[e^{-r\tau}\Bigg(xe^{-\tilde{\beta}\tilde{\lambda}}\exp\Big\{-\sigma\sqrt{\tau}Y + \Big(r-\frac{1}{2}\sigma^2\Big)\tau\Big\}\\
&\qquad\qquad\qquad \times \prod_{i=N(t)+1}^{N(T)}(Y_i+1) - K\Bigg)^{+}\Bigg|\sigma\Bigg(\prod_{i=N(t)+1}^{N(T)}(Y_i+1)\Bigg)\Bigg]\Bigg],
\end{aligned}$$

where

$$Y = -\frac{\widetilde{W}(T)-\widetilde{W}(t)}{\sqrt{\tau}}$$

is a standard normal random variable under $\widetilde{\mathbb{P}}$, and where the conditioning $\sigma$-algebra $\sigma\left(\prod_{i=N(t)+1}^{N(T)}(Y_i+1)\right)$ is the one generated by the random variable $\prod_{i=N(t)+1}^{N(T)}(Y_i+1)$. Because $\prod_{i=N(t)+1}^{N(T)}(Y_i+1)$ is $\sigma\Big(\prod_{i=N(t)+1}^{N(T)}(Y_i+1)\Big)$-measurable and $Y$ is independent of $\sigma\Big(\prod_{i=N(t)+1}^{N(T)}(Y_i+1)\Big)$, we may use the Independence Lemma, Lemma 2.3.4, again to obtain

$$\begin{aligned}
&\widetilde{\mathbb{E}}\Bigg[e^{-r\tau}\Bigg(xe^{-\tilde{\beta}\tilde{\lambda}}\exp\Big\{-\sigma\sqrt{\tau}Y + \Big(r-\frac{1}{2}\sigma^2\Big)\tau\Big\}\\
&\qquad\qquad\qquad \times \prod_{i=N(t)+1}^{N(T)}(Y_i+1) - K\Bigg)^{+}\Bigg|\sigma\Bigg(\prod_{i=N(t)+1}^{N(T)}(Y_i+1)\Bigg)\Bigg]\\
&\qquad\qquad\qquad\qquad = \kappa\Bigg(\tau, xe^{-\tilde{\beta}\tilde{\lambda}\tau}\prod_{i=N(t)+1}^{N(T)}(Y_i+1)\Bigg).
\end{aligned}$$

It follows that

$$c(t,x) = \mathbb{E}\,\kappa\left(\tau, xe^{-\tilde{\beta}\tilde{\lambda}\tau}\prod_{i=N(t)+1}^{N(T)}(Y_i+1)\right). \tag{11.7.32}$$

To see that (11.7.32) agrees with (11.7.30), we note that conditioned on $N(T) - N(t) = j$, the random variable $\prod_{i=N(t)+1}^{N(T)}(Y_i+1)$ has the same distribution as $\prod_{i=1}^{j}(Y_i+1)$. Furthermore,

$$\mathbb{P}\{N(T) - N(t) = j\} = e^{-\tilde{\lambda}\tau}\frac{\tilde{\lambda}^j\tau^j}{j!}. \qquad \square$$

*Remark 11.7.6 (Continuous jump distribution).* Suppose the jump sizes $Y_i$ have a density $f(y)$ rather than a probability mass function $p(y_1), \dots, p(y_m)$, and this density is strictly positive on a set $B \subset (-1,\infty)$ and zero elsewhere. In this case, we replace (11.7.17) by the formula

$$\beta = \mathbb{E}Y_i = \int_{-1}^{\infty} yf(y)\,dy.$$

For the risk-neutral measure, we can choose $\theta$, $\tilde{\lambda} > 0$ and any density $\widetilde{f}(y)$ that is strictly positive on $B$ and zero elsewhere so that the market price of risk equation (see (11.7.26))

$$\alpha - r = \sigma\theta + \beta\lambda - \tilde{\beta}\tilde{\lambda}$$

is satisfied, where now

$$\tilde{\beta} = \widetilde{\mathbb{E}}Y_i = \int_{-1}^{\infty} y\widetilde{f}(y)\,dy.$$

Under these conditions, Theorem 11.7.5 still holds. $\square$

We return to the model with discrete jump sizes. The following theorem provides the differential-difference equation satisfied by the call price.

**Theorem 11.7.7.** *The call price $c(t,x)$ of (11.7.30) satisfies the equation*

$$-rc(t,x) + c_t(t,x) + (r - \tilde{\beta}\tilde{\lambda})xc_x(t,x) + \frac{1}{2}\sigma^2x^2c_{xx}(t,x)$$
$$+\tilde{\lambda}\Big[\sum_{m=1}^{M}\tilde{p}(y_m)c(t,(y_m+1)x) - c(t,x)\Big] = 0,\ 0 \le t < T,\ x \ge 0, \tag{11.7.33}$$

*and the terminal condition*

$$c(T,x) = (x-K)^+,\ x \ge 0.$$

PROOF: From (11.7.27), we see that the continuous part of the stock price satisfies $dS^c(t) = (r - \tilde{\beta}\tilde{\lambda})S(t)\,dt + \sigma S(t)\,d\widetilde{W}(t)$. Therefore, the Itô-Doeblin formula implies

$$\begin{aligned}
&e^{-rt}c(t,S(t)) - c(0,S(0))\\
&= \int_0^t e^{-ru}\Big[-rc(u,S(u)) + c_t(u,S(u)) + (r-\tilde{\beta}\tilde{\lambda})S(u)c_x(u,S(u))\\
&\qquad +\frac{1}{2}\sigma^2 S^2(u)c_{xx}(u,S(u))\Big]\,du + \int_0^t e^{-ru}\sigma S(u)c_x(u,S(u))\,d\widetilde{W}(u)\\
&\quad + \sum_{0<u\le t} e^{-ru}\big[c(u,S(u)) - c(u,S(u-))\big]. \qquad (11.7.34)
\end{aligned}$$

We examine the last term in (11.7.34). If $u$ is a jump time of the $m$th Poisson process $N_m$, the stock price satisfies $S(u) = (y_m+1)S(u-)$. Therefore,

$$\begin{aligned}
&\sum_{0<u\le t} e^{-ru}\big[c(u,S(u)) - c(u,S(u-))\big]\\
&= \sum_{m=1}^M \sum_{0<u\le t} e^{-ru}\big[c(u,(y_m+1)S(u-)) - c(u,S(u-))\big]\Delta N_m(u)\\
&= \sum_{m=1}^M \int_0^t e^{-ru}\big[c(u,(y_m+1)S(u-)) - c(u,S(u-))\big]\,d(N_m(u)-\tilde{\lambda}_m u)\\
&\qquad + \int_0^t e^{-ru}\Big[\sum_{m=1}^M \frac{\tilde{\lambda}_m}{\tilde{\lambda}} c(u,(y_m+1)S(u)) - c(u,S(u))\Big]\tilde{\lambda}\,du\\
&= \sum_{m=1}^M \int_0^t e^{-ru}\big[c(u,(y_m+1)S(u-)) - c(u,S(u-))\big]\,d(N_m(u)-\tilde{\lambda}_m u)\\
&\qquad + \int_0^t e^{-ru}\tilde{\lambda}\sum_{m=1}^M \big[\tilde{p}(y_m)c(u,(y_m+1)S(u)) - c(u,S(u))\big]\Big\}\,du.
\end{aligned}$$

Substituting this into (11.7.34) and taking differentials, we obtain

$$\begin{aligned}
&d\big(e^{-rt}c(t,S(t))\big)\\
&= e^{-rt}\Big\{-rc(t,S(t)) + c_t(t,S(t)) + (r-\tilde{\beta}\tilde{\lambda})S(t)c_x(t,S(t))\\
&\qquad +\frac{1}{2}\sigma^2 S^2(t)c_{xx}(t,S(t))\\
&\qquad +\tilde{\lambda}\sum_{m=1}^M \big[\tilde{p}(y_m)c(t,(y_m+1)S(t)) - c(t,S(t))\big]\Big\}\,dt\\
&\quad +e^{-rt}\sigma S(t)c_x(t,S(t))\,d\widetilde{W}(t)\\
&\quad +\sum_{m=1}^M e^{-rt}\big[c(t,(y_m+1)S(t-)) - c(t,S(t-))\big]\,d(N_m(t)-\tilde{\lambda}_m t). \qquad (11.7.35)
\end{aligned}$$

The integrators $N_m(t) - \tilde{\lambda}_m t$ in the last term are martingales under $\widetilde{\mathbb{P}}$, and the integrands $e^{-rt}[c(t,(y_m+1)S(t-)) - c(t,S(t-))]$ are left-continuous. Therefore, the integral of this term is a martingale. Likewise, the integral of the next-to-last term $e^{-rt}c_x(t,S(t))\,d\widetilde{W}(t)$ is a martingale. Since the discounted option price appearing on the left-hand side of (11.7.35) is also a martingale, the remaining term in (11.7.35) is a martingale as well. Because the remaining term is a $dt$ term, it must be zero. Replacing the price process $S(t)$ by the dummy variable $x$ in the integrand of this term, we obtain (11.7.33). □

**Corollary 11.7.8.** *The call price $c(t,x)$ of (11.7.30) satisfies*

$$
\begin{aligned}
&d\big(e^{-rt}c(t,S(t))\big)\\
&= e^{-rt}\sigma S(t)c_x(t,S(t))\,d\widetilde{W}(t)\\
&\quad+\sum_{m=1}^{M} e^{-rt}\big[c(t,(y_m+1)S(t-)) - c(t,S(t-))\big]\,d(N_m(t)-\tilde{\lambda}_m t)\\
&= e^{-rt}\sigma S(t)c_x(t,S(t))\,d\widetilde{W}(t)\\
&\quad+e^{-rt}\big[c(t,S(t)) - c(t,S(t-))\big]\,dN(t)\\
&\quad-e^{-rt}\tilde{\lambda}\left[\sum_{m=1}^{M}\tilde{p}(y_m)c\big(t,(y+1)S(t-)\big) - c\big(t,S(t-)\big)\right]dt. \qquad (11.7.36)
\end{aligned}
$$

PROOF: We use (11.7.33) to cancel the $dt$ term in (11.7.35) and obtain the first equality in (11.7.36). For the second equality, recall that $N(t) = \sum_{m=1}^{M} N_m(t)$, $\tilde{\lambda} = \sum_{m=1}^{M}\tilde{\lambda}_m$, and $\tilde{\lambda}\tilde{p}(y_m) = \tilde{\lambda}_m$. □

*Remark 11.7.9 (Continuous jump distribution).* There are modifications of Theorem 11.7.7 and Corollary 11.7.8 for the case when the jump sizes $Y_i$ have a density $\tilde{f}(y)$ under the risk-neutral measure $\widetilde{\mathbb{P}}$. In (11.7.33), the term $\sum_{m=1}^{M}\tilde{p}(y_m)c(t,(y_m+1)x)$ would be replaced by $\int_{-1}^{\infty} c(t,(y+1)x)\tilde{f}(y)\,dy$. In (11.7.36), we would use the second formula for $d\big(e^{-rt}c(t,S(t))\big)$, which is written in terms of the total number of jumps (i.e., in terms of the Poisson process $N(t) = \sum_{n=1}^{M} N_m(t)$) rather than in terms of the individual Poisson processes $N_m$, and replace $\sum_{m}^{M}\tilde{p}(y_m)c\big(t,(y_m+1)S(t-)\big)$ by $\int_{-1}^{\infty} c\big(t,y+1)S(t-)\big)\tilde{f}(y)\,dy$ □

Finally, we think about hedging a short position in the European call whose discounted price satisfies (11.7.36). Suppose we begin with a short call position and a hedging portfolio whose initial capital is $X(0) = c(0,S(0))$. We compare the differential of the discounted call price with the differential of the discounted value of the hedging portfolio. If $\Gamma(t)$ shares of stock are held by the hedging portfolio at each time $t$, then

$$dX(t) = \Gamma(t-)\,dS(t) + r\big[X(t) - \Gamma(t)S(t)\big]\,dt$$

and

$$\begin{aligned}
d\big(e^{-rt}X(t)\big) &= e^{-rt}\big[-rX(t)\,dt + dX(t)\big]\\
&= e^{-rt}\big[\Gamma(t-)\,dS(t) - r\Gamma(t)S(t)\,dt\big]\\
&= e^{-rt}\big[\Gamma(t)\sigma S(t)\,d\widetilde{W}(t) + \Gamma(t-)S(t-)\,d(Q(t)-\tilde{\beta}\lambda t)\big]\\
&= e^{-rt}\Big[\Gamma(t)\sigma S(t)\,d\widetilde{W}(t)\\
&\qquad +\Gamma(t-)S(t-)\sum_{m=1}^{M} y_m\,(dN_m(t)-\tilde{\lambda}_m\,dt)\Big], \qquad (11.7.37)
\end{aligned}$$

where we have used (11.7.27). It is natural to try the "delta-hedging" strategy

$$\Gamma(t) = c_x(t, S(t)).$$

This equates the $d\widetilde{W}(t)$ terms in (11.7.36) and (11.7.37) (i.e., it provides a perfect hedge against the risk introduced by the Brownian motion).

However, the delta hedge leaves us with

$$\begin{aligned}
&d\big[e^{-rt}c(t,S(t)) - e^{-rt}X(t)\big]\\
&= \sum_{m=1}^{M} e^{-rt}\big[c(t,(y_m+1)S(t-)) - c(t,S(t-)) - y_mS(t-)c_x(t,S(t-))\big]\\
&\qquad \times (dN_m(t) - \tilde{\lambda}_m\,dt). \qquad (11.7.38)
\end{aligned}$$

The function $c(t,x)$ is strictly convex in $x$. This is a consequence of the strict convexity of the function $\kappa(\tau,x)$ of (11.7.29) and equation (11.7.30). From strict convexity, we have

$$c(t,x_2) - c(t,x_1) > (x_2 - x_1)c_x(t,x_1)$$

for all $x_1 \geq 0$, $x_2 \geq 0$ such that $x_1 \neq x_2$. Therefore,

$$c(t,(y_m+1)S(t-)) - c(t,S(t-)) > y_mS(t-)c_x(t,S(t-)), \qquad (11.7.39)$$

the strict inequality being a consequence of the assumption that each $y_m$ is greater than $-1$ and different from 0. It follows from (11.7.39) and (11.7.38) that between jumps

$$d\big[e^{-rt}c(t,S(t)) - e^{-rt}X(t)\big] < 0.$$

Between jumps, the hedging portfolio outperforms the option. However, at jump times, the option outperforms the hedging portfolio.

Because both $e^{-rt}c(t,S(t))$ and $e^{-rt}X(t)$ are martingales under $\widetilde{\mathbb{P}}$, so is their difference. Furthermore, at the initial time, the difference is $c(0,S(0)) - X(0) = 0$. Therefore, the expected value of the difference is always zero:

$$\widetilde{\mathbb{E}}\big[e^{-rt}c(t,S(t))\big] = \widetilde{\mathbb{E}}\big[e^{-rt}X(t)\big], \; 0 \leq t \leq T.$$

"On average," the delta-hedging formula hedges the option, where the average is computed under the risk-neutral measure we have chosen. This provides some justification for choosing $\tilde{\lambda}_m = \lambda_m$, so that, at least as far as the jumps are concerned, the average under the risk-neutral measure we are using is also the average under the actual probability measure.

*Remark 11.7.10 (Continuous jump distribution).* When the risk-neutral distribution of the jumps $Y_i$ has density $\widetilde{f}(y)$, (11.7.38) becomes

$$\begin{aligned} d\big[e^{-rt}c(t,S(t)) - e^{-rt}X(t)\big] \\ = e^{-rt}\big[c(t,S(t)) - c(t,S(t-)) - (S(t)-S(t-))c_x(t,S(t-))\big]\,dN(t) \\ -e^{-rt}\tilde{\lambda}\int_{-1}^{\infty}\big[c(t,(y+1)S(t-)) - c(t,S(t-)) \\ -yS(t-)c_x(t,S(t-))\big]\widetilde{f}(y)\,dy\,dt. \end{aligned} \tag{11.7.40}$$

Equation (11.7.40) can be interpreted just as (11.7.38) was. Because

$$c(t,(y+1)S(t-)) - c(t,S(t-)) - yS(t-)c_x(t,S(t-)) > 0$$

for all $y > -1$, $y \neq 0$, between jumps

$$d\big[e^{-rt}c(t,S(t)) - e^{-rt}X(t)\big] < 0,$$

the hedging portfolio outperforms the option. At jump times, the option outperforms the hedging portfolio because

$$c(t,S(t)) - c(t,S(t-)) - (S(t)-S(t-))c_x(t,S(t-)) > 0.$$

On "average," where the average is computed under the risk-neutral measure we have chosen, these two effects cancel one another.

## 11.8 Summary

The fundamental pure jump process is the *Poisson process.* Like Brownian motion, the Poisson process is Markov, but unlike Brownian motion, it is not a martingale. The Possion process only jumps up, and between jumps it is constant. To obtain a martingale, one must subtract away the mean of the Poisson process to obtain a *compensated Poisson process* (Theorem 11.2.4).

All jumps of a Poisson process are of size one. A *compound Poisson process* is like a Poisson process, except that the jumps are of random size. Like the Poisson process, a compound Poisson process is Markov (Exercise 11.7), and although it is generally not a martingale, one can obtain a martingale by subtracting away its mean (Theorem 11.3.1). A compound Poisson process that has only finitely many, say $M$, possible jump sizes can be decomposed into a sum of $M$ independent scaled Poisson processes (Theorem 11.3.3 and Corollary 11.3.4).

A *jump process* has four components: an initial condition, an Itô integral, a Riemann integral, and a pure jump process. The sum of the first three constitute the *continuous part* of the jump process. Stochastic integrals and stochastic calculus for the continuous part of a jump process were treated in Chapter 4. In this chapter, the pure jump part is a right-continuous process that has finitely many jumps in each finite time interval and is constant between jumps. Stochastic integrals with respect to such processes are straightforward. The quadratic variation of such a process over a time interval is the sum of the squares of the jumps within that time interval, and the quadratic variation of a (nonpure) jump process is the quadratic variation of the continuous part plus the quadratic variation of the pure jump part. These observations lead to a version of the Itô-Doeblin formula for jump processes (Theorems 11.5.1 and 11.5.4). One of the consequences of these theorems is that a Brownian motion and a Poisson process relative to the same filtration must be independent (Corollary 11.5.3) and that two Poisson processes are independent if and only if they have no simultaneous jumps (Exercises 11.4 and 11.5).

If we integrate an adapted process with respect to a jump process that is a martingale, the resulting stochastic integral can fail to be a martingale. However, if the integrand is left-continuous, then the stochastic integral will be a martingale.

For compound Poisson processes, one can change the measure in order to obtain an arbitrary positive *intensity* (average rate of jump arrival) and an arbitrary distribution of jump sizes, subject to the condition that every jump size that was impossible before the change of measure is still impossible after the change of measure. This provides a great deal of freedom when constructing risk-neutral measures. In particular, if there are $M$ possible jump sizes, there are $M-1$ degrees of freedom in the assignment of probabilities to these jump sizes (the probabilities must sum to one, and thus there are not $M$ degrees of freedom). In order to have a complete market, there must be a money market account and as many nonredundant securities as there are sources of uncertainty. Each possible jump size counts as a source of uncertainty. If there is no Brownian motion and only one possible jump size, a single security in addition to the money market account will make the model complete (Section 11.7.1). If there are two possible jump sizes and an additional source of uncertainty due to a Brownian motion, three securities in addition to the money market account are required (Example 11.7.4). If there are infinitely many possible jump sizes, infinitely many securities would be required to make the model complete.

As the discussion above suggests, jump-diffusion models are generally incomplete and there are typically multiple risk-neutral measures in such models. The practice is to consider a parametrized class of such measures and then calibrate the model to market prices to determine values for the parameters. One can then apply the risk-neutral pricing formula to price derivative securities, but this formula can no longer be justified by a hedging argument. It is instead an elaborate interpolation procedure by which prices of nontraded securities are computed based on prices of traded ones. One can use this formula to examine the effectiveness of various hedging techniques. This is done for the delta-hedging rule in Subsection 11.7.2 following Remark 11.7.9.

## 11.9 Notes

A text on Poisson and compound Possion processes, but that does not include the ideas of change of measure, is Ross [141]. The easiest place to read about stochastic calculus for processes with jumps is Protter [133].

In Section 11.7, we consider a European call in two models, one in which the driving process for the underlying asset is a single Poisson process and the other in which the underlying asset is driven by a Brownian motion and multiple Poisson processes. In both these models, there are only finitely many jump sizes, but the analogous results for models with a continuous jump distribution are presented in Remarks 11.7.6, 11.7.9, and 11.7.10. Such a model was first treated by Merton [123], who considered the case in which one plus the jump size has a log-normal distribution. Some of the more recent works on option pricing in models with jumps are Brockhaus et al. [23], Elliott and Kopp [63], Madan, Carr, and Chang [113], Madan and Milne [114], Madan and Seneta [115], Mercurio and Runggaldier [120], and Overhaus et al [130]. Term-structure models with jumps are treated by Björk, Kabanov and Runggaldier [12], Das [46], Das and Foresi [47], Glasserman and Kou [73], and Glasserman and Merener [74].

## 11.10 Exercises

**Exercise 11.1.** Let $M(t)$ be the compensated Poisson process of Theorem 11.2.4.

(i) Show that $M^2(t)$ is a submartingale.
(ii) Show that $M^2(t) - \lambda t$ is a martingale.

**Exercise 11.2.** Suppose we have observed a Poisson process up to time $s$, have seen that $N(s) = k$, and are interested in the value of $N(s+t)$ for small positive $t$. Show that

$$\begin{aligned}
\mathbb{P}\{N(s+t) = k \big| N(s) = k\} &= 1 - \lambda t + O(t^2), \\
\mathbb{P}\{N(s+t) = k+1 \big| N(s) = k\} &= \lambda t + O(t^2), \\
\mathbb{P}\{N(s+t) \geq k+2 \big| N(s) = k\} &= O(t^2),
\end{aligned}$$

where $O(t^2)$ is used to denote terms involving $t^2$ and higher powers of $t$.

**Exercise 11.3 (Geometric Poisson process).** Let $N(t)$ be a Poisson process with intensity $\lambda > 0$, and let $S(0) > 0$ and $\sigma > -1$ be given. Using Theorem 11.2.3 rather than the Itô-Doeblin formula for jump processes, show that

$$S(t) = \exp\{N(t)\log(\sigma + 1) - \lambda\sigma t\} = (\sigma + 1)^{N(t)} e^{-\lambda\sigma t}$$

is a martingale.

**Exercise 11.4.** Suppose $N_1(t)$ and $N_2(t)$ are Poisson processes with intensities $\lambda_1$ and $\lambda_2$, respectively, both defined on the same probability space $(\Omega, \mathcal{F}, P)$ and relative to the same filtration $\mathcal{F}(t)$, $t \geq 0$. Show that almost surely $N_1(t)$ and $N_2(t)$ can have no simultaneous jump. (Hint: Define the compensated Poisson processes $M_1(t) = N_1(t) - \lambda_1 t$ and $M_2(t) = N_2(t) - \lambda_2 t$, which like $N_1$ and $N_2$ are independent. Use Itô's product rule for jump processes to compute $M_1(t)M_2(t)$ and take expectations.)

**Exercise 11.5.** Suppose $N_1(t)$ and $N_2(t)$ are Poisson processes defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ relative to the same filtration $\mathcal{F}(t)$, $t \geq 0$. Assume that almost surely $N_1(t)$ and $N_2(t)$ have no simultaneous jump. Show that, for each fixed $t$, the random variables $N_1(t)$ and $N_2(t)$ are independent. (Hint: Adapt the proof of Corollary 11.5.3.) (In fact, the whole path of $N_1$ is independent of the whole path of $N_2$, although you are not being asked to prove this stronger statement.)

**Exercise 11.6.** Let $W(t)$ be a Brownian motion and let $Q(t)$ be a compound Poisson process, both defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and relative to the same filtration $\mathcal{F}(t)$, $t \geq 0$. Show that, for each $t$, the random variables $W(t)$ and $Q(t)$ are independent. (In fact, the whole path of $W$ is independent of the whole path of $Q$, although you are not being asked to prove this stronger statement.)

**Exercise 11.7.** Use Theorem 11.3.2 to prove that a compound Poisson process is Markov. In other words, show that, whenever we are given two times $0 \leq t \leq T$ and a function $h(x)$, there is another function $g(t, x)$ such that

$$\mathbb{E}\big[h(Q(T))\big|\mathcal{F}(t)\big] = g(t, Q(t)).$$

# A

# Advanced Topics in Probability Theory

This appendix to Chapter 1 examines more deeply some of the topics touched upon in that chapter. It is intended for readers who desire a fuller explanation. The material in this appendix is not used in the text.

## A.1 Countable Additivity

It is tempting to believe that the finite-additivity condition (1.1.5) can be used to obtain the countable-additivity condition (1.1.2). However, the right-hand side of (1.1.5) is a finite sum, whereas the right-hand side of (1.1.2) is an infinite sum. An infinite sum is not really a sum at all but rather a limit of finite sums:

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \lim_{N \to \infty} \sum_{n=1}^{N} \mathbb{P}(A_n). \tag{A.1.1}$$

Because of this fact, there is no way to get condition (1.1.2) from condition (1.1.5), and so we build the stronger condition (1.1.2) into the definition of probability space.

In fact, condition (1.1.2) is so strong that it is not possible to define $\mathbb{P}(A)$ for every subset $A$ of an uncountably infinite sample space $\Omega$ so that (1.1.2) holds. Because of this, we content ourselves with defining $P(A)$ for every set $A$ in a $\sigma$-algebra $\mathcal{F}$ that contains all the sets we will need for our analysis but omits some of the pathological sets that a determined mathematician can construct.

There are two other consequences of (1.1.2) that we often use implicitly, and these are provided by the next theorem.

**Theorem A.1.1.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $A_1, A_2, A_3, \ldots$ be a sequence of sets in $\mathcal{F}$.*

*(i) If $A_1 \subset A_2 \subset A_3 \subset \ldots$, then*

$$\mathbb{P}(\cup_{k=1}^{\infty} A_k) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

*(ii) If* $A_1 \supset A_2 \supset A_3 \supset \dots$, *then*

$$\mathbb{P}(\cap_{k=1}^{\infty} A_k) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

PROOF: In the first case, we define

$$B_1 = A_1,\ B_2 = A_2 \setminus A_1,\ B_3 = A_3 \setminus A_2, \dots,$$

where $A_{k+1} \setminus A_k = A_{k+1} \cap A_k^c$. Then $B_1, B_2, B_3, \dots$ are disjoint sets, and

$$A_n = \bigcup_{k=1}^{n} A_k = \bigcup_{k=1}^{n} B_k, \quad \bigcup_{k=1}^{\infty} A_k = \bigcup_{k=1}^{\infty} B_k.$$

Condition (1.1.2) used to justify the second equality below and (1.1.5) used to justify the fourth imply

$$\begin{aligned}\mathbb{P}(\cup_{k=1}^{\infty} A_k) = \mathbb{P}(\cup_{k=1}^{\infty} B_k) &= \sum_{k=1}^{\infty} \mathbb{P}(B_k) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{P}(B_k) \\ &= \lim_{n\to\infty} \mathbb{P}(\cup_{k=1}^{n} B_k) = \lim_{n\to\infty} \mathbb{P}(A_n).\end{aligned}$$

This concludes the proof of (i).

Let us now assume $A_1 \supset A_2 \supset A_3 \supset \dots$. We define $C_k = A_k^c$, so that $C_1 \subset C_2 \subset C_3 \subset \dots$ and $\cap_{k=1}^{\infty} A_k = (\cup_{k=1} C_k)^c$. Then (1.1.6) and (i) imply

$$\begin{aligned}\mathbb{P}(\cap_{k=1}^{\infty} A_k) = 1 - \mathbb{P}(\cup_{k=1}^{\infty} C_k) &= 1 - \lim_{n\to\infty} \mathbb{P}(C_n) \\ = \lim_{n\to\infty} (1 - \mathbb{P}(C_n)) &= \lim_{n\to\infty} \mathbb{P}(A_n).\end{aligned}$$

Thus we have (ii). □

Property (i) of Theorem A.1.1 was used in (1.2.6) at the step

$$\lim_{n\to\infty} \mathbb{P}\{-n \le X \le n\} = \mathbb{P}\{X \in \mathbb{R}\}.$$

Property (ii) of this theorem was used in (1.2.4). Property (ii) can also be used in the following example.

*Example A.1.2.* We continue Example 1.1.3, the uniform measure on $[0,1]$. Recall the $\sigma$-algebra $\mathcal{B}[0,1]$ of Borel subsets of $[0,1]$, obtained by beginning with the closed intervals and adding all other sets necessary in order to have a $\sigma$-algebra. A complicated but instructive example of a set in $\mathcal{B}[0,1]$ is the *Cantor set*, which we now construct. We also compute its probability, where the probability measure $\mathbb{P}$ we use is the uniform measure, assigning a probability to each interval $[a,b] \subset [0,1]$ equal to its length $b-a$.

From the interval $[0,1]$, remove the middle third (i.e., the open interval $(\frac{1}{3},\frac{2}{3})$). The remaining set is

$$C_1 = \left[0,\frac{1}{3}\right] \cup \left[\frac{2}{3},1\right],$$

which has two pieces, each with probability $\frac{1}{3}$, and the whole set $C_1$ has probability $\frac{2}{3}$. From each of the two pieces of $C_1$, remove the middle third (i.e., remove the open intervals $(\frac{1}{9},\frac{2}{9})$ and $(\frac{7}{9},\frac{8}{9})$). The remaining set is

$$C_2 = \left[0,\frac{1}{9}\right] \cup \left[\frac{2}{9},\frac{1}{3}\right] \cup \left[\frac{2}{3},\frac{7}{9}\right] \cup \left[\frac{8}{9},1\right],$$

which has four pieces, each with probability $\frac{1}{9}$, and the whole set $C_2$ has probability $\frac{4}{9}$. See Figure A.1.1.

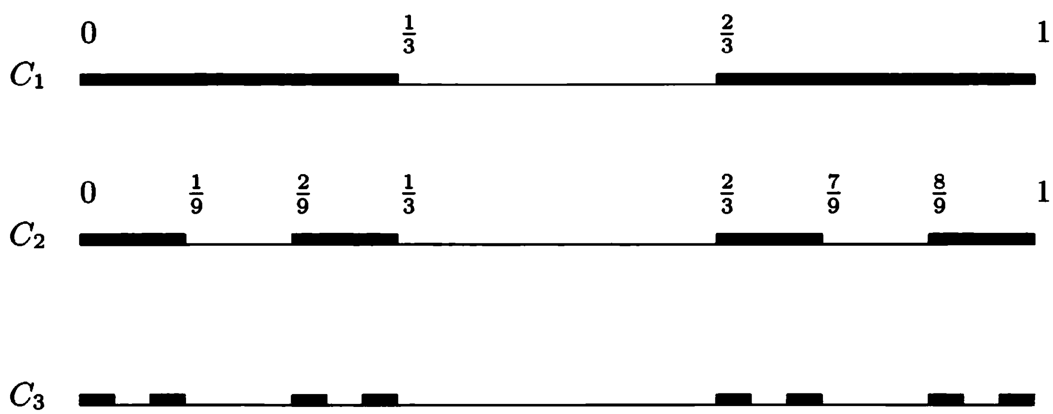


**Fig. A.1.1.** Constructing the Cantor set.

Continue this process so at stage $k$ we have a set $C_k$ that has $2^k$ pieces, each with probability $\frac{1}{3^k}$, and the whole set $C_k$ has probability $\left(\frac{2}{3}\right)^k$. The Cantor set is defined to be $C = \cap_{k=1}^{\infty} C_k$. From Theorem A.1.1(ii), we see that

$$\mathbb{P}(C) = \lim_{k\to\infty} \mathbb{P}(C_k) = \lim_{k\to\infty} \left(\frac{2}{3}\right)^k = 0.$$

Despite the fact that it has zero probability, the Cantor set has infinitely many points. It certainly contains the points $0, \frac{1}{3}, \frac{2}{3}, 1, \frac{1}{9}, \frac{2}{9}, \frac{7}{9}, \frac{8}{9}, \frac{1}{27}, \frac{2}{27}, \dots$, which are the endpoints of the intervals appearing at the successive stages, because these are never removed. This is a countably infinite set of points. In fact, the Cantor set has uncountably many points. To see this, assume that all the points in the Cantor set can be listed in a sequence $x_1, x_2, x_3, \dots$. Let $K_1$ denote the piece of $C_1$, either $\left[0,\frac{1}{3}\right]$ or $\left[\frac{2}{3},1\right]$, that does not contain $x_1$. Let $K_2$ be a piece of $K_1 \cap C_2$ that does not contain $x_2$. For example, if $K_1 = \left[0,\frac{1}{3}\right]$

and $x_2 \in \left[\frac{2}{9}, \frac{1}{3}\right]$, we take $K_2 = \left[0, \frac{1}{9}\right]$. If $x_2 \notin K_1$, it does not matter whether we take $K_2 = \left[0, \frac{1}{9}\right]$ or $K_2 = \left[\frac{2}{9}, \frac{1}{3}\right]$. Next let $K_3$ be a piece of $K_2 \cap C_3$ that does not contain $x_3$. Continue this process. Then

$$K_1 \supset K_2 \supset K_3 \supset \dots, \tag{A.1.2}$$

and $x_1 \notin K_1$, $x_2 \notin K_2$, $x_3 \notin K_3, \dots$. In particular, $\cap_{n=1}^{\infty} K_n$ does not contain any point in the sequence $x_1, x_2, x_3, \dots$. But the intersection of a sequence of nonempty closed intervals that are "nested" as described by (A.1.2) must contain something, and so there is a point $y$ satisfying $y \in \cap_{n=1}^{\infty} K_n$. But $\cap_{n=1}^{\infty} K_n \subset C$, and so the point $y$ is in the Cantor set but not on the list $x_1, x_2, x_3, \dots$. This shows that the list cannot include every point in the Cantor set. The set of all points in the Cantor set cannot be listed in a sequence, which means that the Cantor set is uncountably infinite. □

## A.2 Generating $\sigma$-algebras

We often have some collection $\mathcal{C}$ of subsets of a sample space $\Omega$ and want to put in all other sets necessary in order to have a $\sigma$-algebra. We did this in Example 1.1.3 when we constructed the $\sigma$-algebra $\mathcal{B}[0,1]$ and again in Example 1.1.4 when we constructed $\mathcal{F}_\infty$. In the former case, $\mathcal{C}$ was the collection of all closed intervals $[a, b] \subset [0, 1]$; in the latter case, $\mathcal{C}$ was the collection of all subsets of $\Omega_\infty$ that could be described in terms of finitely many coin tosses.

In general, when we begin with a collection $\mathcal{C}$ of subsets of $\Omega$ and put in all other sets necessary in order to have a $\sigma$-algebra, the resulting $\sigma$-algebra is called the *$\sigma$-algebra generated by $\mathcal{C}$* and is denoted by $\sigma(\mathcal{C})$. The description just given of $\sigma(\mathcal{C})$ is not mathematically precise because it is difficult to determine how and whether the process of "putting in all other sets necessary in order to have a $\sigma$-algebra" terminates. We provide a precise mathematical definition at the end of this discussion.

The precise definition of $\sigma(\mathcal{C})$ works from the outside in rather than the inside out. In particular, we define $\sigma(\mathcal{C})$ to be the "smallest" $\sigma$-algebra containing all the sets in $\mathcal{C}$ in the following sense. Put in $\sigma(\mathcal{C})$ every set that is in every $\sigma$-algebra that is "bigger" than $\mathcal{C}$ (i.e., that contains all the sets in $\mathcal{C}$). There is at least one $\sigma$-algebra containing all the sets in $\mathcal{C}$, the $\sigma$-algebra of all subsets of $\Omega$. If this is the only $\sigma$-algebra bigger than $\mathcal{C}$, then we put every subset of $\Omega$ into $\sigma(\mathcal{C})$ and we are done. If there are other $\sigma$-algebras bigger than $\mathcal{C}$, then we put into $\sigma(\mathcal{C})$ only those sets that are in *every* such $\sigma$-algebra. We note the following items.

(i) The empty set $\emptyset$ is in $\sigma(\mathcal{C})$ because it is in every $\sigma$-algebra bigger than $\mathcal{C}$.
(ii) If $A \in \sigma(\mathcal{C})$, then $A$ is in every $\sigma$-algebra bigger than $\mathcal{C}$. Therefore, $A^c$ is in every such $\sigma$-algebra, which implies that $A^c$ is in $\sigma(\mathcal{C})$.
(iii) If $A_1, A_2, A_3, \dots$ is a sequence of sets in $\sigma(\mathcal{C})$, then this sequence is in every $\sigma$-algebra bigger than $\mathcal{C}$, and so the union $\cup_{n=1}^{\infty} A_n$ is also in every such $\sigma$-algebra. This shows that the union is in $\sigma(\mathcal{C})$.

(iv) By definition, every set in $\mathcal{C}$ is in every $\sigma$-algebra bigger than $\mathcal{C}$ and so is in $\sigma(\mathcal{C})$.
(v) Suppose $\mathcal{G}$ is a $\sigma$-algebra bigger than $\mathcal{C}$. By definition, every set in $\sigma(\mathcal{C})$ is also in $\mathcal{G}$.

Properties (i)–(iii) show that $\sigma(\mathcal{C})$ is a $\sigma$-algebra. Property (iv) shows that $\sigma(\mathcal{C})$ contains all the sets in $\mathcal{C}$. Property (v) shows that $\sigma(\mathcal{C})$ is the "smallest" $\sigma$-algebra containing all the sets in $\mathcal{C}$.

**Definition A.2.1** *Let $\mathcal{C}$ be a collection of subsets of a nonempty set $\Omega$. The* $\sigma$-algebra generated by $\mathcal{C}$, *denoted $\sigma(\mathcal{C})$, is the collection of sets that belong to all $\sigma$-algebras bigger than $\mathcal{C}$ (i.e., all $\sigma$-algebras containing all the sets in $\mathcal{C}$).*

## A.3 Random Variable with Neither Density nor Probability Mass Function

Using the notation of Example 1.2.5, let us define

$$Y = \sum_{n=1}^{\infty} \frac{2Y_n}{3^n}.$$

If $Y_1 = 0$, which happens with probability $\frac{1}{2}$, then $0 \le Y \le \frac{1}{3}$. If $Y_1 = 1$, which also happens with probability $\frac{1}{2}$, then $\frac{2}{3} \le Y \le 1$. If $Y_1 = 0$ and $Y_2 = 0$, which happens with probability $\frac{1}{4}$, then $0 \le Y \le \frac{1}{9}$. If $Y_1 = 0$ and $Y_2 = 1$, which also happens with probability $\frac{1}{4}$, then $\frac{2}{9} \le Y \le \frac{1}{3}$. This pattern continues. Indeed, when we consider the first $n$ tosses we see that the random variable $Y$ takes values in the set $C_n$ defined in Example A.1.2, and hence $Y$ can only take values in the Cantor set $C = \cap_{n=1}^{\infty} C_n$.

We first argue that $Y$ cannot have a density. If it did, then the density $f$ would have to be zero except on the set $C$. But $C$ has zero Lebesgue measure, and so $f$ is almost everywhere zero and $\int_0^1 f(x)\,dx = 0$ (i.e., the function $f$ would not integrate to one, as is required of a density).

We next argue that $Y$ cannot have a probability mass function. If it did, then for some number $x \in C$ we would have $\mathbb{P}(Y = x) > 0$. But $x$ has a unique base-three expansion

$$x = \sum_{n=1}^{\infty} \frac{x_n}{3^n},$$

where each $x_n$ is either 0, 1, or 2 unless $x$ is of the form $\frac{k}{3^n}$ for some positive integers $k$ and $n$. In the latter case, $x$ has two base-three expansions. For example, $\frac{7}{9}$ can be written as both

$$\frac{7}{9} = \frac{2}{3} + \frac{1}{9} + \frac{0}{27} + \frac{0}{81} + \frac{0}{243} + \dots$$

and
$$\frac{7}{9} = \frac{2}{3} + \frac{0}{9} + \frac{1}{27} + \frac{1}{81} + \frac{1}{243} + \cdots.$$
In either case, there are at most two choices of $\omega \in \Omega_\infty$ for which $Y(\omega) = x$. In other words, the set $\{\omega \in \Omega; Y(\omega) = x\}$ has either one or two elements. The probability of a set with one element is zero, and the probability of a set with two elements is $0 + 0 = 0$. Hence $\mathbb{P}\{Y = x\} = 0$.

The cumulative distribution function $F(x) = \mathbb{P}\{Y \le x\}$ satisfies (see Figure A.3.1 for a partial rendition of $F(x)$)

$$\begin{aligned}
F(0) = 0,\ F(1) = 1,\ \ F(x) &= \frac{1}{2} \text{ for } \frac{1}{3} \le x \le \frac{2}{3},\\
F(x) = \frac{1}{4} \text{ for } \frac{1}{9} \le x \le \frac{2}{9},\ \ F(x) &= \frac{3}{4} \text{ for } \frac{7}{9} \le x \le \frac{8}{9},\\
F(x) = \frac{1}{8} \text{ for } \frac{1}{27} \le x \le \frac{2}{27},\ \ F(x) &= \frac{3}{8} \text{ for } \frac{7}{27} \le x \le \frac{8}{27},\\
F(x) = \frac{5}{8} \text{ for } \frac{19}{27} \le x \le \frac{20}{27},\ \ F(x) &= \frac{7}{8} \text{ for } \frac{25}{27} \le x \le \frac{26}{27},\\
&\ \ \vdots
\end{aligned}$$

and, because $\mathbb{P}\{Y = x\} = 0$ for every $x$, $F$ is continuous. Furthermore, $F'(x) = 0$ for every $x \in [0,1] \setminus C$, which is almost every $x \in [0,1]$. A nonconstant continuous function whose derivative is almost everywhere zero is said to be *singularly continuous*.

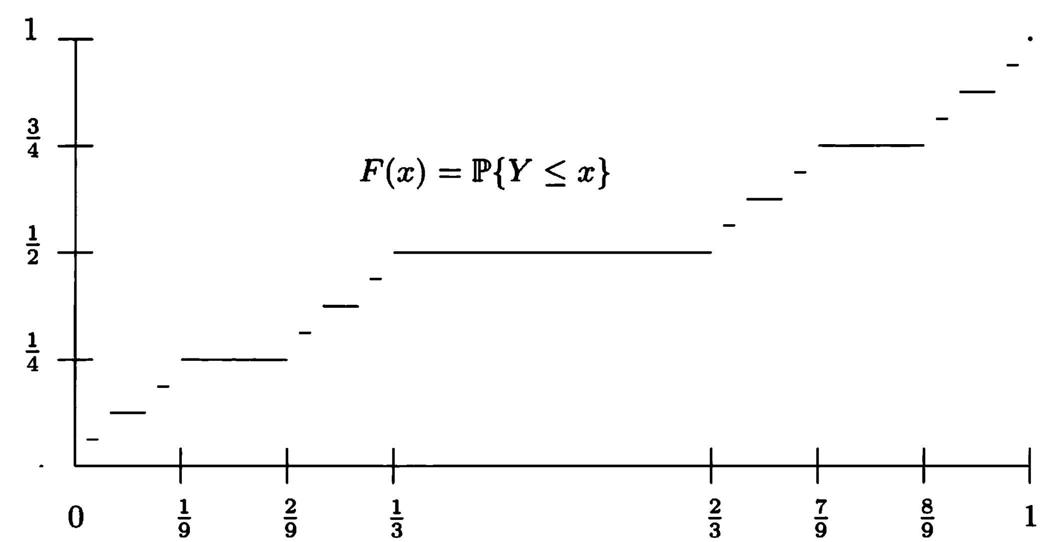


**Fig. A.3.1.** A singularly continuous function.

# B

# Existence of Conditional Expectations

This appendix uses the Radon-Nikodým Theorem, Theorem 1.6.7, to establish the existence of the conditional expectation of a random variable $X$ with respect to a $\sigma$-algebra $\mathcal{G}$. Here we treat the case when $X$ is nonnegative *and* integrable. If $X$ is only integrable, one can decompose it in the usual way as $X = X^+ - X^-$, the difference of nonnegative integrable random variables, and then apply Theorem B.1 below to $X^+$ and $X^-$ separately. If $X$ is only nonnegative, one can write it as the limit of a nondecreasing sequence of nonnegative integrable random variables and use the Monotone Convergence Theorem, Theorem 1.4.5, to extend Theorem B.1 below to cover this case.

**Theorem B.1.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{G}$ be a sub-$\sigma$-algebra of $\mathcal{F}$, and let $X$ be an integrable nonnegative random variable. Then there exists a $\mathcal{G}$-measurable random variable $Y$ such that*

$$\int_A Y(\omega)\, d\mathbb{P}(\omega) = \int_A X(\omega)\, d\mathbb{P}(\omega) \text{ for every } A \in \mathcal{G}. \tag{B.1}$$

In light of Definition 2.3.1, the random variable $Y$ in the theorem above is the conditional expectation $\mathbb{E}[X|\mathcal{G}]$.

PROOF OF THEOREM B.1: We define a probability measure by

$$\widetilde{\mathbb{P}}(A) = \int_A \frac{X(\omega)+1}{\mathbb{E}[X+1]}\, d\mathbb{P}(\omega) \text{ for every } A \in \mathcal{F}.$$

Because the integrand $\frac{X+1}{\mathbb{E}[X+1]}$ is strictly positive almost surely and has expectation 1, $\mathbb{P}$ and $\widetilde{\mathbb{P}}$ are equivalent probability measures (see Theorem 1.6.1 and the comment following Definition 1.6.3).

The probabilities $\mathbb{P}(A)$ and $\widetilde{\mathbb{P}}(A)$ are defined for every subset $A$ of $\Omega$ that is in $\mathcal{F}$. We define two equivalent probability measures on the smaller $\sigma$-algebra $\mathcal{G}$. The first is simply $\mathbb{P}$ restricted to $\mathcal{G}$ (i.e., we define $\mathbb{Q}(A) = \mathbb{P}(A)$ for every

$A \in \mathcal{G}$, and we leave $\mathbb{Q}(A)$ undefined for $A \notin \mathcal{G}$). The second is $\widetilde{\mathbb{P}}$ restricted to $\mathcal{G}$ (i.e., we define $\widetilde{\mathbb{Q}}(A) = \widetilde{\mathbb{P}}(A)$ for every $A \in \mathcal{G}$, and we leave $\widetilde{\mathbb{Q}}(A)$ undefined for $A \notin \mathcal{G}$). We now have two probability spaces, $(\Omega, \mathcal{G}, \mathbb{Q})$ and $(\Omega, \mathcal{G}, \widetilde{\mathbb{Q}})$, which differ only by their probability measures $\mathbb{Q}$ and $\widetilde{\mathbb{Q}}$. Moreover, $\mathbb{Q}$ and $\widetilde{\mathbb{Q}}$ are equivalent. The Radon-Nikodým Theorem, Theorem 1.6.7, implies the existence of a random variable $Z$ such that

$$\widetilde{\mathbb{Q}}(A) = \int_A Z(\omega)\, d\mathbb{Q}(\omega) \text{ for every } A \in \mathcal{G}.$$

However, since we are now working on probability spaces with $\sigma$-algebra $\mathcal{G}$, the random variable $Z$ whose existence is guaranteed by the Radon-Nikodým Theorem will be $\mathcal{G}$-measurable rather than $\mathcal{F}$-measurable. (Recall from Definition 1.2.1 that every random variable is measurable with respect to the $\sigma$-algebra in the space on which it is defined.)

Since $\widetilde{\mathbb{Q}}$ and $\mathbb{Q}$ agree with $\widetilde{\mathbb{P}}$ and $\mathbb{P}$ on $\mathcal{G}$, we may rewrite the formula above as

$$\widetilde{\mathbb{P}}(A) = \int_A Z(\omega)\, d\mathbb{P}(\omega) \text{ for every } A \in \mathcal{G}$$

or, equivalently,

$$\int_A \frac{X(\omega)+1}{\mathbb{E}[X+1]}\, d\mathbb{P}(\omega) = \int_A Z(\omega)\, d\mathbb{P}(\omega) \text{ for every } A \in \mathcal{G}.$$

Multiplication by $\mathbb{E}[X+1]$ leads to the equation

$$\int_A X(\omega)\, d\mathbb{P}(\omega) + \int_A 1\, d\mathbb{P}(\omega) = \int_A \mathbb{E}[X+1] Z(\omega)\, d\mathbb{P}(\omega) \text{ for every } A \in \mathcal{G}.$$

We conclude that

$$\int_A X(\omega)\, d\mathbb{P}(\omega) = \int_A \big(\mathbb{E}[X+1] Z(\omega) - 1\big)\, d\mathbb{P}(\omega) \text{ for every } A \in \mathcal{G}.$$

Taking $Y(\omega) = \mathbb{E}[X+1]Z(\omega) - 1$, we have (B.1). Because $Z$ is $\mathcal{G}$-measurable and $\mathbb{E}[X+1]$ is constant, $Y$ is also $\mathcal{G}$-measurable. $\square$

# C

# Completion of the Proof of the Second Fundamental Theorem of Asset Pricing

This appendix provides a lemma that is the last step in the proof of the Second Fundamental Theorem of Asset Pricing, Theorem 5.4.9 of Chapter 5.

**Lemma C.1** *Let $A$ be an $m \times d$-dimensional matrix, $b$ an $m$-dimensional vector, and $c$ a $d$-dimensional vector. If the equation*

$$Ax = b \tag{C.1}$$

*has a unique solution $x_0$, a $d$-dimensional vector, then the equation*

$$A^{\mathrm{tr}} y = c \tag{C.2}$$

*has at least one solution $y_0$, an $m$-dimensional vector. (Here, $A^{\mathrm{tr}}$ denotes the transpose of the matrix $A$.)*

PROOF: We regard $A$ as a mapping from $\mathbb{R}^d$ to $\mathbb{R}^m$ and define the *kernel of* $A$ to be

$$K(A) = \{x \in \mathbb{R}^d : Ax = 0\}.$$

If $x_0$ solves (C.1) and $x \in K(A)$, then $x_0 + x$ also solves (C.1). Thus, the assumption of a unique solution to (C.1) implies that $K(A)$ contains only the $d$-dimensional zero vector.

The rank of $A$ is defined to be the number of linearly independent columns of $A$. Because $K(A)$ contains only the $d$-dimensional zero vector, the rank must be $d$. Otherwise, we could find a linear combination of these columns that would be the $m$-dimensional zero vector, and the coefficients in this linear combination would give us a non-zero vector in $K(A)$. But any matrix and its transpose have the same rank, and so the rank of $A^{\mathrm{tr}}$ is $d$ as well. The rank of a matrix is also the dimension of its range space. The range space of $A^{\mathrm{tr}}$ is

$$R(A^{\mathrm{tr}}) = \{z \in \mathbb{R}^d : z = A^{\mathrm{tr}} y \text{ for some } y \in \mathbb{R}^m\}.$$

Because the dimension of this space is $d$ and it is a subspace of $\mathbb{R}^d$, it must in fact be equal to $\mathbb{R}^d$. In other words, for every $z \in \mathbb{R}^d$, there is some $y \in \mathbb{R}^m$ such that $z = A^{\mathrm{tr}} y$. Hence, (C.2) has a solution $y_0 \in \mathbb{R}^m$. □

This page intentionally left blank