5

Interest Rate Derivative Securities

5.1 Introduction

This chapter is devoted to interest rate derivatives. Interest rate derivatives are financial products derived from interest rates. There are various interest rates that will be mentioned in this chapter. Here we first give the meaning of each interest rate and derive some relations among them.

An N-year zero-coupon yield or an N-year spot interest rate is the interest rate on an investment starting at time t and lasting for N years. The investment is a "pure" N-year investment with no intermediate payments. Assume that the interest is compounded continuously. In this case, suppose that at time t the N-year zero-coupon yield is $Y(t, t+N)$, then the investor will get

$$\mathrm{e}^{Y(t,t+N)N}$$

at the end of year N for each dollar invested. A zero-coupon yield curve is a curve showing the relation between $Y(t, t+N)$ and N.

A zero-coupon bond with a face value or a par value of one dollar is a contract whose holder will get one dollar at the maturity of the contract from its issuer. Let $Z(t;T)$ denote the money a person needs to pay in order to have the contract with maturity date T at time t. Then, between $Y(t,T)$ and $Z(t;T)$, there is the following relation

$$Z(t;T) = \mathrm{e}^{-Y(t,T)(T-t)}, \tag{5.1}$$

or

$$Y(t,T) = \frac{-\ln Z(t;T)}{T-t}.$$

Suppose $t \le T_1 \le T_2$. An interest rate determined at time t for a period $[T_1, T_2]$ and paid at time T_2 is called a forward interest rate. Let us denote this rate by $f(t, T_1, T_2)$ and again assume that the interest is compounded continuously. Among $f(t, T_1, T_2)$, $Z(t;T_1)$, and $Z(t;T_2)$, there is the following relation:

Y.-l. Zhu et al., *Derivative Securities and Difference Methods*, Springer Finance,
DOI 10.1007/978-1-4614-7306-0_5,

$$Z(t;T_1) = Z(t;T_2)\mathrm{e}^{f(t,T_1,T_2)(T_2-T_1)},$$

or

$$f(t,T_1,T_2) = \frac{1}{T_2-T_1}\ln\frac{Z(t;T_1)}{Z(t;T_2)}. \tag{5.2}$$

The reason is the following. If we borrow one dollar at time T_1, then we need to return $\mathrm{e}^{f(t,T_1,T_2)(T_2-T_1)}$ dollars at time T_2 according to the forward interest rate at time t. At time t, the values of one dollar at time T_1 and $\mathrm{e}^{f(t,T_1,T_2)(T_2-T_1)}$ dollars at time T_2 should be the same, otherwise there is an arbitrage opportunity.

An instantaneous forward interest rate $F(t,T_1)$ is the limit of $f(t,T_1,T_2)$ as $T_2 \to T_1$, written as

$$\begin{aligned} F(t,T_1) &= \lim_{T_2\to T_1} f(t,T_1,T_2) = \lim_{T_2\to T_1}\frac{-[\ln Z(t;T_2) - \ln Z(t;T_1)]}{T_2-T_1} \\ &= \frac{-1}{Z(t;T_1)}\frac{\partial Z(t;T_1)}{\partial T_1}. \end{aligned} \tag{5.3}$$

This gives

$$Z(t;T) = Z(t;t)\mathrm{e}^{-\int_t^T F(t,u)du} = \mathrm{e}^{-\int_t^T F(t,u)du}.$$

Furthermore, combining this expression for $Z(t,T)$ with the relation (5.1) yields

$$Y(t,T) = \frac{1}{T-t}\int_t^T F(t,u)du. \tag{5.4}$$

The limit of $Y(t,T)$ as $T \to t$ is called the instantaneous short rate (see [43]), the short-term interest rate, the short rate, or the spot rate (see [84]), denoted by $r(t)$, so

$$r(t) = \lim_{T\to t} Y(t,T) = Y(t,t).$$

Because from Eq. (5.4) we also have

$$\lim_{T\to t} Y(t,T) = \lim_{T\to t}\frac{1}{T-t}\int_t^T F(t,u)du = F(t,t),$$

we get

$$r(t) = Y(t,t) = F(t,t). \tag{5.5}$$

Clearly, if $Y(t,T)$ is equal to a constant r, then

$$Z(t;T) = \mathrm{e}^{-r(T-t)},$$

and

$$f(t,T_1,T_2) = F(t,T_1) = F(t,t) = Y(t,t) = r(t) = r.$$

In practice, the interest is often compounded discretely. If a loan of one dollar is required to pay at an interest rate $\bar{r}$ compounded m times per year, then the amount of each payment is

$$\frac{\bar{r}}{m}.$$

For an investment with an interest rate r compounded continuously, the interest payment for a period $\dfrac{1}{m}$ years is

$$\mathrm{e}^{r/m} - 1.$$

If

$$\mathrm{e}^{r/m} - 1 = \frac{\bar{r}}{m},$$

that is,

$$r = m \ln(1 + \bar{r}/m),$$

then the two investments are equivalent. Suppose that a forward interest rate at time t for the period $[T_1, T_1 + 1/m]$ is an interest rate compounded m times per year and we use $\bar{f}(t, T_1, T_1 + 1/m)$ to denote this forward interest rate. Let $f(t, T_1, T_1 + 1/m)$ be equivalent to the interest rate $\bar{f}(t, T_1, T_1 + 1/m)$. Then we have

$$f(t, T_1, T_1 + 1/m) = m \ln\left(1 + \frac{\bar{f}(t, T_1, T_1 + 1/m)}{m}\right)$$

and the relation (5.2) can be rewritten as

$$m \ln\left(1 + \frac{\bar{f}(t, T_1, T_1 + 1/m)}{m}\right) = m \ln\left(\frac{Z(t; T_1)}{Z(t; T_1 + 1/m)}\right)$$

or

$$\bar{f}(t, T_1, T_1 + 1/m) = m\left[\frac{Z(t; T_1)}{Z(t; T_1 + 1/m)} - 1\right]. \tag{5.6}$$

This is the counterpart of the relation (5.2) for an interest rate compounded m times per year. Actually, this relation can also be derived directly. For the formulae (5.1) and (5.3)–(5.5), we can also have their counterparts for interest rates compounded discretely.

As we know, the value of a bond is related to interest rates. There are many other financial contracts related to interest rates, which are signed between two parties, for example, a bank and a company. These are called interest rate derivatives. For an equity option, a typical life span is 9 months or less. In this case, the assumption of a short rate being a deterministic function of t, even a constant, is acceptable. If this is not the case, it may be necessary to consider a short rate as a random variable. For example, a life span of a bond may be 5 years, 10 years, even 30 years. Therefore, it is more realistic to deal with a short-term interest rate as a random variable. An interest rate cannot be traded on the market. In Chap. 2, we pointed out that there is a unknown function called the market price of risk for a short rate in the governing partial differential equation (PDE) for interest rate derivatives. Before using such an equation to price a derivative security, one has to find this function. From the mathematical point of view, to find a unknown function in the partial differential equation is to solve an inverse problem. This function in the PDE is determined by some data associated with solutions of the equation. The values of zero-coupon bonds with various maturity dates on the market or some other data can be taken as the data needed. Moreover, reducing the randomness of a zero-coupon bond curve to the randomness of the short rate is not a good approximation in many cases. Thus, describing the randomness of a zero-coupon bond curve by the randomness of several interest rates, namely, considering multi-factor models, is necessary.

Therefore, the rest of this chapter is organized as follows. In Sects. 5.2 and 5.3, the problem for a bond is formulated and for four special models, explicit solutions are derived. In Sect. 5.4, we discuss the inverse problem of determining the market price of risk and give a formulation of the inverse problem so that the determination of the unknown function can be reduced to solving such a problem. Then, we discuss bond options, swaps, swaptions, and so forth in Sect. 5.5. Section 5.6 is devoted to multi-factor interest rate models, especially, a three-factor model that can used in practice easily. Finally, two-factor convertible bonds are discussed in Sect. 5.7.

5.2 Bonds

A bond is a long-term contract under which the issuer promises to pay the holder a specified amount of money on a specified date. The specified amount is called the face value of the bond, which is denoted by Z in this chapter, and the specified date is named the maturity date T. Usually, the holder is also paid a specified amount at fixed times during the life of the contract. Such a specified amount is called a coupon. If there is no coupon payment, the bond is known as a zero-coupon bond. Clearly, the bondholder must pay a certain amount of money to the issuer when the bond is purchased. This amount is called the upfront premium. In this section, we will mainly derive the equations by which one can determine a fair value of the bond for any time t, including the upfront premium.

5.2.1 Bond Values for Deterministic Short Rates

Let r be the interest rate for the shortest possible deposit, which is called the short-term interest rate or, for short, the short rate in this book. For a short period, r may be assumed to be a constant. For a long period, for example, a few years, it is unreasonable to consider r as a constant. As a starting point, we assume that the short rate is a known function of t, i.e., $r = r(t)$. Let $V(t)$ stand for the value of a bond with coupon rate $k(t)$ at time t. Assume that the return rate of a bond during the time interval $[t, t + dt]$ be the risk-free short rate, so we have

$$dV + Zk(t)dt = r(t)V dt,$$

where $Zk(t)dt$ is the coupon payment the bondholder receives during the time interval. If the coupon is paid continuously, $k(t)$ is a continuous function of t. If it is paid at fixed times, $k(t)$ is a linear combination of Dirac delta functions, i.e., $k(t) = \sum_i k_i \delta(t - t_i)$, $t_i \le T$. The relation above can also be written as

$$dV - r(t)V dt = -Zk(t)dt.$$

Multiplying both sides of the equation by $\mathrm{e}^{\int_t^T r(\tau)d\tau}$, which is usually referred to as the integrating factor, yields

$$\mathrm{e}^{\int_t^T r(\tau)d\tau}\left[dV - r(t)V dt\right] = -Zk(t)\mathrm{e}^{\int_t^T r(\tau)d\tau}dt.$$

The left-hand side actually is $d\left(\mathrm{e}^{\int_t^T r(\tau)d\tau}V\right)$. Therefore, we have

$$\int_t^T d\left(\mathrm{e}^{\int_t^T r(\tau)d\tau}V\right) = V(T) - \mathrm{e}^{\int_t^T r(\tau)d\tau}V(t) = -Z\int_t^T k(\bar{\tau})\mathrm{e}^{\int_{\bar{\tau}}^T r(\tau)d\tau}d\bar{\tau}$$

and

$$\begin{aligned} V(t) &= \mathrm{e}^{-\int_t^T r(\tau)d\tau}\left[V(T) + Z\int_t^T k(\bar{\tau})\mathrm{e}^{\int_{\bar{\tau}}^T r(\tau)d\tau}d\bar{\tau}\right] \\ &= V(T)\mathrm{e}^{-\int_t^T r(\tau)d\tau}\left[1 + \int_t^T k(\bar{\tau})\mathrm{e}^{\int_{\bar{\tau}}^T r(\tau)d\tau}d\bar{\tau}\right], \end{aligned} \tag{5.7}$$

where we have used the condition $Z = V(T)$. For a zero-coupon bond, $k(t) = 0$ and

$$V(t) = V(T)\mathrm{e}^{-\int_t^T r(\tau)d\tau} = Z\mathrm{e}^{-\int_t^T r(\tau)d\tau}.$$

From the right-hand side, we see that the value of $V(t)$ depends on T. However, this dependence is suppressed in this expression. In order to express this dependence explicitly, the relation above can be rewritten as

$$V(t;T) = V(T;T)\mathrm{e}^{-\int_t^T r(\tau)d\tau}, \tag{5.8}$$

where $V(T;T) = Z$.

At time t, the values of zero-coupon bonds with various maturities can be obtained from the market, i.e., $V(t;T)$ with a fixed t and various T is observable. Suppose we have such a function. Differentiating the formula (5.8) with respect to T yields

$$\frac{\partial V\,(t;T)}{\partial T} = -V(T;T)\mathrm{e}^{-\int_t^T r(\tau)d\tau} r(T) = -V(t;T)r(T)$$

and

$$r(T) = \frac{-1}{V(t;T)} \frac{\partial V\,(t;T)}{\partial T}.$$

This means that the short rate at time T can be determined by the value and the slope of the function $V(t;T)$. It is clear that $r(T)$ does not depend on Z. Let $Z = 1$, then comparing the expression for $r(T)$ and the formula (5.3) yields

$$F(t,T) = r(T)$$

and

$$Z(t;T) = \mathrm{e}^{-\int_t^T r(u)du} \tag{5.9}$$

if $Z = 1$. Also for a zero-coupon bond,

$$\frac{V(t;T)}{V(T;T)} = Z(t;T).$$

Thus, from the relation (5.1) we have

$$Y(t,T) = \frac{-\ln Z(t;T)}{T-t} = \frac{-\ln\left(V(t;T)/V(T;T)\right)}{T-t}, \tag{5.10}$$

which is usually called the yield of a bond during the time interval $[t,T]$. A plot of Y against the time to maturity, $T-t$, is called the yield curve. The dependence of the yield on $T-t$ is called the term structure of interest rates. The historical data on bonds are usually given in the form of yields for various $T-t$.

5.2.2 Bond Equations for Random Short Rates

It will be more realistic to consider the short rate r as a random variable. Suppose

$$dr = u(r,t)dt + w(r,t)dX. \tag{5.11}$$

From Sect. 2.3, we know that the value of a bond as a short rate derivative, $V\,(r,t)$, satisfies Eq. (2.34) with only one random variable r:

$$\frac{\partial V}{\partial t} + \frac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + (u - \lambda w)\frac{\partial V}{\partial r} - rV + kZ = 0, \tag{5.12}$$

where kZ is the coupon payment and $\lambda = \lambda(r,t)$ is the market price of risk for r. For a bond the value at maturity date T is a constant Z, i.e.,

$$V(r,T) = Z. \tag{5.13}$$

If the short rate model satisfies the conditions (2.39) and (2.40), then no boundary condition is needed, i.e., Eq. (5.12) with the final condition (5.13) has a unique solution.

5.3 Some Explicit Solutions of Bond Equations

There exist many short rate models. Here, we discuss the following model (see [84]):

$$dr = [\bar{\mu}(t) - \bar{\gamma}(t)r]\, dt + \sqrt{\alpha(t)r - \beta(t)}dX, \tag{5.14}$$

where $\alpha(t)$, $\beta(t)$, $\bar{\gamma}(t)$, and $\bar{\mu}(t)$ are given functions of t. Several important models, for example, the Vasicek model (see [81]), the Cox–Ingersoll–Ross model (see [23]), the Ho–Lee model (see [41]), and the Hull–White model (see [44]) possess this form. For the models in the form (5.14), the determination of the value of a zero-coupon bond can be reduced to solving two ordinary differential equations. Sometimes we can find analytic solutions or the solution can be expressed in terms of integrals with known integrands. Such a solution is referred to as an explicit solution here.

If a short rate model is in the form (5.14) and we take

$$\lambda\,(r,t) = \bar{\lambda}(t)\sqrt{\alpha(t)r - \beta(t)}, \tag{5.15}$$

then Eq. (5.12) can be written as

$$\frac{\partial V}{\partial t} + \frac{1}{2}\left[\alpha(t)r - \beta(t)\right]\frac{\partial^2 V}{\partial r^2} + \left[\mu(t) - \gamma(t)r\right]\frac{\partial V}{\partial r} - rV = 0, \tag{5.16}$$

where

$$\mu(t) = \bar{\mu}(t) + \bar{\lambda}(t)\beta(t) \tag{5.17}$$

and

$$\gamma(t) = \bar{\gamma}(t) + \bar{\lambda}(t)\alpha(t). \tag{5.18}$$

Here, we let $k = 0$ because we are going to determine the value of a zero-coupon bond. Because the coefficients of $\dfrac{\partial^2 V}{\partial r^2}$ and $\dfrac{\partial V}{\partial r}$ are linear functions in r, the solution of Eq. (5.16) with the condition (5.13) has the following form

$$V(r,t) = Z\mathrm{e}^{A(t,T) - rB(t,T)} \tag{5.19}$$

with

$$A(T, T) = 0 \tag{5.20}$$

and

$$B(T, T) = 0. \tag{5.21}$$

In fact, because the conditions (5.20) and (5.21) hold, we have

$$V(r, T) = Z.$$

Substituting the expression (5.19) into Eq. (5.16) yields

$$\frac{dA}{dt} - r\frac{dB}{dt} + \frac{1}{2}\left[\alpha(t)r - \beta(t)\right]B^2 - \left[\mu(t) - \gamma(t)r\right]B - r = 0.$$

If the sum of the terms independent of r is equal to zero, i.e.,

$$\frac{dA}{dt} - \frac{1}{2}\beta(t)B^2 - \mu(t)B = 0$$

and the sum of all coefficients of r is equal to zero, i.e.,

$$-\frac{dB}{dt} + \frac{1}{2}\alpha B^2 + \gamma(t)B - 1 = 0,$$

then the expression (5.19) is a solution to a zero-coupon bond. These two equations above, which can be rewritten as

$$\frac{dA}{dt} = \frac{1}{2}\beta(t)B^2 + \mu(t)B \tag{5.22}$$

and

$$\frac{dB}{dt} = \frac{1}{2}\alpha(t)B^2 + \gamma(t)B - 1, \tag{5.23}$$

have unique solutions satisfying the conditions (5.20) and (5.21). Thus, it is true that Eq. (5.16) with the condition (5.13) has a solution in the form (5.19) satisfying the conditions (5.20) and (5.21), and the solution of the problem can be reduced to solving the two ordinary differential equations (5.22) and (5.23) with the conditions (5.20) and (5.21).

5.3.1 Analytic Solutions for the Vasicek and Cox–Ingersoll–Ross Models

If $\alpha, \beta, \gamma, \mu$ in Eqs. (5.22) and (5.23) are constant, then we can find analytic expressions for A and B. When A and B have such expressions, the expression (5.19) gives an analytic solution for a zero-coupon bond. In this case, Eq. (5.23) can be rewritten as

$$\frac{dB}{B^2 + \dfrac{2\gamma}{\alpha}B - \dfrac{2}{\alpha}} = \frac{\alpha}{2}dt. \tag{5.24}$$

Since

$$B^2 + \frac{2\gamma}{\alpha}B - \frac{2}{\alpha} = \left(B + \frac{\gamma - \psi}{\alpha}\right)\left(B + \frac{\gamma + \psi}{\alpha}\right),$$

where

$$\psi = \sqrt{\gamma^2 + 2\alpha}, \tag{5.25}$$

using the method of partial fraction decomposition, we can have

$$\frac{1}{B^2 + \frac{2\gamma}{\alpha}B - \frac{2}{\alpha}} = \frac{\frac{\alpha}{2\psi}}{B + \frac{\gamma - \psi}{\alpha}} - \frac{\frac{\alpha}{2\psi}}{B + \frac{\gamma + \psi}{\alpha}}.$$

Noticing this relation, we can easily find the solution to Eq. (5.24) by integrating both sides of the equation:

$$\begin{aligned}
&\int_{B(t,T)}^{B(T,T)} \frac{dB}{B^2 + \frac{2\gamma}{\alpha}B - \frac{2}{\alpha}} \\
&= \frac{\alpha}{2\psi}\left[\int_{B(t,T)}^{0} \frac{dB}{B + (\gamma - \psi)/\alpha} - \int_{B(t,T)}^{0} \frac{dB}{B + (\gamma + \psi)/\alpha}\right] \\
&= \frac{\alpha}{2\psi}\left[\ln \frac{(\gamma - \psi)/\alpha}{B + (\gamma - \psi)/\alpha} - \ln \frac{(\gamma + \psi)/\alpha}{B + (\gamma + \psi)/\alpha)}\right] \\
&= \frac{\alpha}{2}\int_t^T dt = \frac{\alpha}{2}(T - t).
\end{aligned}$$

From this we have

$$\frac{B + (\gamma + \psi)/\alpha}{B + (\gamma - \psi)/\alpha} = \frac{\gamma + \psi}{\gamma - \psi}\, \mathrm{e}^{\psi(T-t)}$$

or

$$\begin{aligned}
B &= \frac{\frac{\gamma + \psi}{\alpha}\, \mathrm{e}^{\psi(T-t)} - \frac{\gamma + \psi}{\alpha}}{1 - \frac{\gamma + \psi}{\gamma - \psi}\, \mathrm{e}^{\psi(T-t)}} \\
&= \frac{2\left[\mathrm{e}^{\psi(T-t)} - 1\right]}{(\gamma + \psi)\, \mathrm{e}^{\psi(T-t)} - (\gamma - \psi)},
\end{aligned} \tag{5.26}$$

where we have used the relation $\psi^2 - \gamma^2 = 2\alpha$. After we find B, from Eq. (5.22) we have

$$\begin{aligned}
\int_{A(t,T)}^{A(T,T)} dA &= A(T,T) - A(t,T) \\
&= \int_t^T \left(\frac{1}{2}\beta B^2 + \mu B\right) dt
\end{aligned}$$

or

$$A(t,T) = -\frac{1}{2}\beta \int_t^T B^2 dt - \mu \int_t^T B dt.$$

Using the relation (5.24), we can obtain the results of $\int_t^T B dt$ and $\int_t^T B^2 dt$ easily as follows:

$$\begin{aligned}
\int_t^T B dt &= \int_{B(t,T)}^0 \frac{2B/\alpha}{B^2 + 2\gamma B/\alpha - 2/\alpha} dB \\
&= \frac{2}{\alpha} \int_{B(t,T)}^0 \left[\frac{-(\gamma-\psi)/(2\psi)}{B+(\gamma-\psi)/\alpha} + \frac{(\gamma+\psi)/(2\psi)}{B+(\gamma+\psi)/\alpha} \right] dB \\
&= -\frac{\gamma-\psi}{\alpha\psi} \ln \frac{(\gamma-\psi)/\alpha}{B+(\gamma-\psi)/\alpha} + \frac{\gamma+\psi}{\alpha\psi} \ln \frac{(\gamma+\psi)/\alpha}{B+(\gamma+\psi)/\alpha}
\end{aligned}$$

and

$$\begin{aligned}
\int_t^T B^2 dt &= \frac{2}{\alpha} \int_{B(t,T)}^0 \frac{B^2}{B^2 + 2\gamma B/\alpha - 2/\alpha} dB \\
&= \frac{2}{\alpha} \int_{B(t,T)}^0 \left(1 - \frac{2\gamma B/\alpha}{B^2 + 2\gamma B/\alpha - 2/\alpha} + \frac{2/\alpha}{B^2 + 2\gamma B/\alpha - 2/\alpha} \right) dB \\
&= \frac{2}{\alpha} \int_{B(t,T)}^0 \left(1 + \frac{\gamma(\gamma-\psi)/(\alpha\psi)}{B+(\gamma-\psi)/\alpha} - \frac{\gamma(\gamma+\psi)/(\alpha\psi)}{B+(\gamma+\psi)/\alpha} \right. \\
&\quad \left. + \frac{1/\psi}{B+(\gamma-\psi)/\alpha} - \frac{1/\psi}{B+(\gamma+\psi)/\alpha} \right) dB \\
&= \frac{2}{\alpha} \left\{ -B + \left[\frac{\gamma(\gamma-\psi)}{\alpha\psi} + \frac{1}{\psi} \right] \ln \frac{(\gamma-\psi)/\alpha}{B+(\gamma-\psi)/\alpha} \right. \\
&\quad \left. - \left[\frac{\gamma(\gamma+\psi)}{\alpha\psi} + \frac{1}{\psi} \right] \ln \frac{(\gamma+\psi)/\alpha}{B+(\gamma+\psi)/\alpha} \right\} \\
&= \frac{2}{\alpha} \left\{ -B - \frac{\gamma-\psi}{(\gamma+\psi)\psi} \ln \frac{(\gamma-\psi)/\alpha}{B+(\gamma-\psi)/\alpha} \right. \\
&\quad \left. + \frac{(\gamma+\psi)}{(\gamma-\psi)\psi} \ln \frac{(\gamma+\psi)/\alpha}{B+(\gamma+\psi)/\alpha} \right\}.
\end{aligned}$$

Therefore

$$\begin{aligned}
A &= \frac{\beta}{\alpha} B + \left[\frac{\beta(\gamma-\psi)}{\alpha(\gamma+\psi)\psi} + \mu \frac{\gamma-\psi}{\alpha\psi} \right] \ln \frac{(\gamma-\psi)/\alpha}{B+(\gamma-\psi)/\alpha} \\
&\quad - \left[\frac{\beta(\gamma+\psi)}{\alpha(\gamma-\psi)\psi} + \mu \frac{\gamma+\psi}{\alpha\psi} \right] \ln \frac{(\gamma+\psi)/\alpha}{B+(\gamma+\psi)/\alpha} \qquad (5.27)
\end{aligned}$$

and

$$V(r,t) = Z\left[\frac{B+(\gamma-\psi)/\alpha}{(\gamma-\psi)\alpha}\right]^{\beta(\psi-\gamma)/\alpha(\gamma+\psi)\psi+\mu(\psi-\gamma)/\alpha\psi} \times\left[\frac{B+(\gamma+\psi)/\alpha}{(\gamma+\psi)/\alpha}\right]^{\beta(\gamma+\psi)/\alpha(\gamma-\psi)\psi+\mu(\gamma+\psi)/\alpha\psi} \mathrm{e}^{(\beta/\alpha-r)B}. \quad (5.28)$$

This is a solution of a zero-coupon bond suitable for all the models (5.14) with constant $\alpha, \beta, \bar{\gamma}$, and $\bar{\mu}$ as long as we choose the market price of risk in the form $\lambda(r,t) = \bar{\lambda}\sqrt{\alpha r - \beta}$. The parameters $\alpha, \beta, \bar{\gamma}$, and $\bar{\mu}$ can be obtained from the data on the short rate on the market. However $\bar{\lambda}$, a parameter in the expression of the market price of risk, cannot be determined from the data on the short rate and should be obtained from the other data on the market. For example, $\bar{\lambda}$ can be determined from the yield function on the market by the least squares method, i.e., by choosing $\bar{\lambda}$ so that $\int_t^T \left[Y(t,T;\bar{\lambda}) - \tilde{Y}(t-T)\right]^2 dT$ is minimized, or

$$\int_t^T \left[Y\left(t,T;\bar{\lambda}\right) - \tilde{Y}(t,T)\right] \frac{\partial Y\left(t,T;\bar{\lambda}\right)}{\partial \bar{\lambda}} dT = 0. \quad (5.29)$$

Here, $\tilde{Y}(t,T)$ is the yield function observed on the market, whereas according to the expressions (5.10) and (5.19), the function $Y(t,T;\bar{\lambda})$ is given by

$$Y\left(t,T;\bar{\lambda}\right) = \frac{rB(t,T;\bar{\lambda}) - A(t,T;\bar{\lambda})}{T-t}, \quad (5.30)$$

where $A(t,T;\bar{\lambda})$ and $B(t,T;\bar{\lambda})$ are given by the expressions (5.26) and (5.27), respectively, but the dependence of A and B on $\bar{\lambda}$ is expressed explicitly here. If the value of yield is only available discretely on the market, then we can find a $\bar{\lambda}$ such that $\sum_i \left[Y(t,T_i;\bar{\lambda}) - \tilde{Y}(t,T_i)\right]^2$ is minimized, or

$$\sum_i \left[Y(t,T_i;\bar{\lambda}) - \tilde{Y}(t,T_i)\right] \frac{\partial Y(t,T_i;\bar{\lambda})}{\partial \bar{\lambda}} = 0. \quad (5.31)$$

As soon as we have $\bar{\lambda}$, the (5.16) with constant α, β, γ, and μ can be used to determine the value of any other short rate derivative. Generally speaking, it is impossible to fit the entire yield curve by choosing only one parameter. This is a drawback of such a model.

For some special models, for example, the Vasicek model (see [81]) and the Cox–Ingersoll–Ross model (see [23]), the expression can be simplified. Let us do this for these two models.

The Vasicek model is in the form

$$dr = (\bar{\mu} - \bar{\gamma}r)\,dt + \sqrt{-\beta}dX, \quad \beta < 0, \quad \bar{\gamma} > 0.$$

Therefore, the expressions (5.26) and (5.27) with

$$\alpha = 0, \quad \mu = \bar{\mu} + \bar{\lambda}\beta$$

and

$$\gamma = \bar{\gamma}$$

give B and A for this model. In this case, the expression (5.26) becomes[1]

$$B = \frac{\mathrm{e}^{\gamma(T-t)} - 1}{\gamma\, \mathrm{e}^{\gamma(T-t)}} = \frac{1}{\gamma}(1 - \mathrm{e}^{-\gamma(T-t)}). \tag{5.32}$$

However, the expression (5.27) cannot be used for calculation because of $\alpha = 0$. In order to have an expression for A that can be used for calculation, we need to find the limit of the expression (5.27) as $\alpha \to 0$ or solve Eq. (5.22) with B given by the expression (5.32) directly. Let us solve Eq. (5.22) directly. Putting the expression (5.32) into Eq. (5.22), we have:

$$\begin{aligned}
&A(T,T) - A(t,T) = \int_{A(t,T)}^{A(T,T)} dA \\
&= \int_t^T \left[\frac{\beta}{2\gamma^2}\left(1 - \mathrm{e}^{-\gamma(T-t)}\right)^2 + \frac{\mu}{\gamma}\left(1 - \mathrm{e}^{-\gamma(T-t)}\right)\right] dt \\
&= \int_t^T \left[\frac{\beta}{2\gamma^2}\left(1 - \mathrm{e}^{-\gamma(T-t)}\right)\left(-\mathrm{e}^{-\gamma(T-t)}\right)\right. \\
&\quad \left. + \left(\frac{\beta}{2\gamma^2} + \frac{\mu}{\gamma}\right)\left(1 - \mathrm{e}^{-\gamma(T-t)}\right)\right] dt \\
&= \left[\frac{\beta}{4\gamma^3}\left(1 - \mathrm{e}^{-\gamma(T-t)}\right)^2 + \left(\frac{\beta}{2\gamma^2} + \frac{\mu}{\gamma}\right)\left(t - \frac{1}{\gamma}\mathrm{e}^{-\gamma(T-t)}\right)\right]\Bigg|_t^T \\
&= -\frac{\beta}{4\gamma^3}\left(1 - \mathrm{e}^{-\gamma(T-t)}\right)^2 + \left(\frac{\beta}{2\gamma^2} + \frac{\mu}{\gamma}\right)(T-t) \\
&\quad - \left(\frac{\beta}{2\gamma^2} + \frac{\mu}{\gamma}\right)\frac{1}{\gamma}\left(1 - \mathrm{e}^{-\gamma(T-t)}\right).
\end{aligned}$$

Because of $A(T,T) = 0$, we obtain

$$A = -\left(\frac{\beta}{2\gamma^2} + \frac{\mu}{\gamma}\right)(T-t) + \left(\frac{\beta}{2\gamma^2} + \frac{\mu}{\gamma}\right)B + \frac{\beta B^2}{4\gamma}. \tag{5.33}$$

Consequently

$$\begin{aligned}
V(r,t) &= Z\mathrm{e}^{-(\beta/2\gamma^2+\mu/\gamma)(T-t)+(\beta/2\gamma^2+\mu/\gamma-r)B+\beta B^2/4\gamma} \\
&= Z\mathrm{e}^{-(\beta/2\gamma^2+\mu/\gamma)(T-t)+(\beta/2\gamma^2+\mu/\gamma-r)(1-\mathrm{e}^{-\gamma(T-t)})/\gamma+\beta\left(1-\mathrm{e}^{-\gamma(T-t)}\right)^2/4\gamma^3}.
\end{aligned} \tag{5.34}$$

[1] This expression can also be obtained by integrating Eq. (5.23) with $\alpha = 0$ directly, and for the case $\alpha = 0$, this direct way of finding the solution is easier.

This is the value of a zero-coupon bond if the Vasicek model is adopted. As pointed out above, the solution (5.34) can also be obtained by finding the limit of the solution (5.28). This is left to readers as Problem 7.

Noticing that B does not depend on $\bar{\lambda}$ in this case, we have

$$\begin{aligned}
Y(t,T;\bar{\lambda}) &= \frac{rB\,(t,T) - A(t,T;\bar{\lambda})}{T-t} \\
&= \frac{\left(\frac{\beta}{2\gamma^2}+\frac{\mu}{\gamma}\right)(T-t) - \left(\frac{\beta}{2\gamma^2}+\frac{\mu}{\gamma}-r\right)B - \frac{\beta B^2}{4\gamma}}{T-t} \\
&= \frac{\left(\frac{\beta}{2\gamma^2}+\frac{\bar{\mu}}{\gamma}\right)(T-t) - \left(\frac{B}{2\gamma^2}+\frac{\bar{\mu}}{\gamma}-r\right)B - \frac{\beta B^2}{4\gamma}}{T-t} \\
&\quad + \frac{\bar{\lambda}\beta(T-t-B)}{\gamma\,(T-t)}
\end{aligned}$$

and

$$\frac{\partial Y}{\partial \bar{\lambda}}(t,T;\bar{\lambda}) = \frac{\beta(T-t-B)}{\gamma(T-t)}.$$

Hence, Eq. (5.29) becomes a linear equation for $\bar{\lambda}$. From the linear equation, we see that $\bar{\lambda}$ is given by

$$\frac{\displaystyle\int_t^T \left[\frac{\left(\frac{\beta}{2\gamma^2}+\frac{\bar{\mu}}{\gamma}\right)(T-t) - \left(\frac{B}{2\gamma^2}+\frac{\bar{\mu}}{\gamma}-r\right)B - \frac{\beta B^2}{4\gamma}}{T-t} - \tilde{Y}\right]\frac{T-t-B}{(T-t)}dT}{\displaystyle -\frac{\beta}{\gamma}\int_t^T \frac{(T-t-B)^2}{(T-t)^2}dT}. \tag{5.35}$$

Because only $\bar{\lambda}$ is chosen, the yield curve cannot be fitted entirely. Another problem of this model is that r may be negative.

In order to rectify this problem, Cox, Ingersoll, and Ross (see [23]) proposed another model:

$$dr = (\bar{\mu} - \bar{\gamma}r)dt + \sqrt{\alpha r}dX. \tag{5.36}$$

This is also in the form (5.14) and $\beta = 0$ here. In this case, the solution for a zero-coupon bond is

$$V(r,t) = Z\left[\frac{B+(\gamma-\psi)/\alpha}{(\gamma-\psi)/\alpha}\right]^{\mu(\psi-\gamma)/\alpha\psi}\left[\frac{B+(\gamma+\psi)/\alpha}{(\gamma+\psi)/\alpha}\right]^{\mu(\gamma+\psi)/\alpha\psi}\mathrm{e}^{-rB}.$$

Here, B is given by the expression (5.26), i.e.,

$$B = \frac{2\left(\mathrm{e}^{\psi(T-t)}-1\right)}{(\gamma+\psi)\,\mathrm{e}^{\psi(T-t)}-(\gamma-\psi)},$$

$$\mu = \bar{\mu}$$

and

$$\gamma = \bar{\gamma} + \bar{\lambda}\alpha,$$

where $\bar{\lambda}$ is a parameter in the expression (5.15) for the market price of risk. However, the solution can have another form. Because

$$\begin{aligned} A(T,T) - A(t,T) &= \int_{A(t,T)}^{A(T,T)} dA = \int_t^T \mu B dt \\ &= \mu \int_t^T \frac{2\left(\mathrm{e}^{\psi(T-t)}-1\right)}{(\gamma+\psi)\,\mathrm{e}^{\psi(T-t)}-(\gamma-\psi)} dt, \end{aligned}$$

noticing $A(T,T)=0$ and setting $\xi = \mathrm{e}^{\psi(T-t)}$, we have

$$\begin{aligned} &A(t,T) \\ &= -\mu \int_t^T \frac{2\left(\mathrm{e}^{\psi(T-t)}-1\right)}{(\gamma+\psi)\,\mathrm{e}^{\psi(T-t)}-(\gamma-\psi)} dt = \mu \int_\xi^1 \frac{2(\xi-1)d\xi}{[(\gamma+\psi)\xi-(\gamma-\psi)]\psi\xi} \\ &= \frac{2\mu}{\psi(\gamma+\psi)} \int_\xi^1 \left[\frac{-2\psi/(\gamma-\psi)}{\xi-(\gamma-\psi)/(\gamma+\psi)} + \frac{(\gamma+\psi)/(\gamma-\psi)}{\xi}\right] d\xi \\ &= \frac{-4\mu}{\gamma^2-\psi^2}\left[\ln(\xi-(\gamma-\psi)/(\gamma+\psi)) - (\gamma+\psi)\ln\xi/2\psi\right]\Big|_\xi^1 \\ &= \frac{2\mu}{\alpha}\left[\ln(1-(\gamma-\psi)/(\gamma+\psi)) - \ln(\xi-(\gamma-\psi)/(\gamma+\psi)) \right. \\ &\quad \left. +(\gamma+\psi)\ln\xi/2\psi\right] \\ &= \ln\left(\frac{2\psi\mathrm{e}^{(\gamma+\psi)(T-t)/2}}{(\gamma+\psi)\mathrm{e}^{\psi(T-t)}-(\gamma-\psi)}\right)^{2\mu/\alpha}. \end{aligned}$$

Therefore, we have a solution

$$V(r,t) = Z\left[\frac{2\psi\mathrm{e}^{(\gamma+\psi)(T-t)/2}}{(\gamma+\psi)\mathrm{e}^{\psi(T-t)}-(\gamma-\psi)}\right]^{2\mu/\alpha} \mathrm{e}^{-rB}, \qquad (5.37)$$

which is the form given in the paper by Cox, Ingersoll, and Ross. It can be proved that the two expressions are identical. This is left to readers to prove as Problem 9.

In this case

$$Y(t,T;\bar{\lambda}) = \frac{\dfrac{2\left(\mathrm{e}^{\psi(T-t)}-1\right)r}{(\gamma+\psi)\mathrm{e}^{\psi(T-t)}-(\gamma-\psi)} - \dfrac{2\mu}{\alpha}\ln\left(\dfrac{2\psi\mathrm{e}^{(\gamma+\psi)(T-t)/2}}{(\gamma+\psi)\mathrm{e}^{\psi(T-t)}-(\gamma-\psi)}\right)}{T-t},$$

where $\gamma = \bar{\gamma} + \bar{\lambda}\alpha$ and $\psi = \sqrt{\gamma^2 + 2\alpha}$, so the dependence of $Y(t, T; \bar{\lambda})$ on $\bar{\lambda}$ is quite complicated.

As we have stated, in order to use the partial differential equation (5.16) to price the value of other derivatives, we need to determine $\bar{\lambda}$ so that we can have $\gamma = \bar{\gamma} + \bar{\lambda}\alpha$. For example, we can obtain $\bar{\lambda}$ by solving Eq. (5.29). Because the dependence of $Y(t, T; \bar{\lambda})$ on $\bar{\lambda}$ in this case is quite complicated, Eq. (5.29) has to be solved numerically. Just like the Vasicek model, generally speaking, it is impossible to "build" the entire term structure of the short rate into a parameter $\bar{\lambda}$.

5.3.2 Explicit Solutions for the Ho–Lee and Hull–White Models

In order to fit the entire term structure of interest rates, it seems to be necessary to require $\bar{\lambda}$ to be dependent on t or r. If $\bar{\lambda}$ depends on t, then for some models in the form (5.14), the solution of a zero-coupon bond can explicitly be expressed by elemental functions and integrals with known integrands. We refer to such a solution as an explicit solution or a closed-form solution. The Ho–Lee model (see [41])

$$dr = \bar{\mu}(t)dt + \sqrt{-\beta}dX \tag{5.38}$$

and the Hull–White model (see [44])

$$dr = (\bar{\mu}(t) - \bar{\gamma}r)dt + \sqrt{-\beta}dX \tag{5.39}$$

are such models. We note that the Hull–White model is an extension of the Ho–Lee model and the Vasicek model. For the Hull–White model, $B(t, T)$ is the same as for the Vasicek model, given by the expression (5.32):

$$B(t, T) = \frac{1}{\gamma}(1 - \mathrm{e}^{-\gamma(T-t)}),$$

where

$$\gamma = \bar{\gamma}.$$

Let $\gamma \to 0$, we have

$$B(t, T) = T - t, \tag{5.40}$$

which is the expression of $B(t, T)$ for the Ho–Lee model. For both of them, Eq. (5.22) is in the form

$$\frac{dA}{dt} = \frac{1}{2}\beta B^2 + \mu(t)B,$$

where $\mu(t)$ is given by the expression (5.17):

$$\mu(t) = \bar{\mu}(t) + \bar{\lambda}(t)\beta.$$

Here, we assume that the market price of risk is $\lambda(r,t) = \bar{\lambda}(t)\sqrt{-\beta}$. From the ordinary differential equation above, we can find

$$A(t,T) = -\frac{1}{2}\beta \int_t^T B^2(\tau,T)d\tau - \int_t^T \mu(\tau)B(\tau,T)d\tau$$

and

$$V(r,t) = Z\mathrm{e}^{-\frac{1}{2}\beta\int_t^T B^2(\tau,T)d\tau - \int_t^T \mu(\tau)B(\tau,T)d\tau - rB(t,T)}. \tag{5.41}$$

Here, B is given by the expression (5.32) or the expression (5.40), depending on which model is used. Therefore, if $\bar{\lambda}$ is given, we can find $V(r,t)$ without any difficulties.

In practice, we need to find $\bar{\lambda}(t)$ from some data on the market, for example, a given yield function $Y(t,T)$. In order to do this, we rewrite the solution (5.41) as

$$\ln V(r,t) = \ln Z - \frac{1}{2}\beta \int_t^T B^2(\tau,T)d\tau - \int_t^T \mu(\tau)B(\tau,T)d\tau - rB(t,T)$$

or if we require that the solution (5.41) fits the yield function on the market, we furthermore have

$$\int_t^T \mu(\tau)B(\tau,T)d\tau = Y(t,T)(T-t) - \frac{1}{2}\beta \int_t^T B^2(\tau,T)d\tau - rB(t,T), \tag{5.42}$$

where we have used the definition of the yield (5.10). If we define

$$F_1(t,T) \equiv Y(t,T)(T-t) - \frac{1}{2}\beta \int_t^T B^2(\tau,T)d\tau - rB(t,T) \tag{5.43}$$

and substitute $(1-\mathrm{e}^{-\gamma(T-\tau)})/\gamma$ for B on the left-hand side of Eq. (5.42), it becomes

$$\frac{1}{\gamma}\int_t^T \mu(\tau)(1-\mathrm{e}^{-\gamma(T-\tau)})d\tau = F_1(t,T).$$

Differentiating both sides of this relation with respect to T twice yields

$$\mu(T) = \frac{\partial^2 F_1(t,T)}{\partial T^2} + \gamma\frac{\partial F_1(t,T)}{\partial T}, \tag{5.44}$$

which is the function $\mu(t)$ for the Hull–White model. After having the function $\mu(t)$, we can obtain $\bar{\lambda}(t)$ immediately by

$$\bar{\lambda}(t) = \frac{1}{\beta}\left[\mu(t) - \bar{\mu}(t)\right]$$

if we want. Therefore, for the Hull–White model, we can find a function for the market price of risk for r such that the entire term structure of interest rate can be fitted. For the Ho–Lee model, in order to do this, we can use the same formula with $\gamma = 0$, so in the expression (5.43) $B = T - t$. Because in these models the entire term structure of interest rate is built into the function $\bar{\lambda}(t)$, these two models are often referred to as no-arbitrage interest rate models. The difference between them is that the Hull–White model has the mean reversion property that an interest rate model should have, whereas the Ho–Lee model does not. However, even though the Hull-White model has the mean reversion property, r is still defined on $(-\infty, \infty)$ because the coefficient of dX in the model is a constant.

5.4 Inverse Problem on the Market Price of Risk

As we saw in Sect. 5.3, for some special interest rate models and some special function of the market price of risk, we can find an explicit solution for a zero-coupon bond and furthermore explicit expressions for the market price of risk for which the entire term structure of interest rate or the entire zero-coupon bond price curve is fitted. However, even though the model is in the form (5.14) and solving the partial differential equation (5.16) can be reduced to solving ordinary differential equations (5.22) and (5.23), we still may not be able to find an explicit expression for the market price of risk if $\alpha(t)$ really depends on t or even if α is a nonzero constant. In this case, the unknown function $\bar{\lambda}(t)$ appears in both Eqs. (5.22) and (5.23) and it may be necessary to use numerical methods.

Also, there are other models, for example, the Black–Derman–Toy model (see [9]):

$$d \ln r = \left[\bar{\mu}(t) - \frac{\sigma_r'(t)}{\sigma_r(t)} \ln r\right] dt + \sigma_r(t) dX$$

and the Black–Karasinski model (see [10]):

$$d \ln r = \left[\bar{\mu}(t) - \bar{\gamma}(t) \ln r\right] dt + \sigma_r(t) dX.$$

For these models, it might even be impossible to reduce solving a partial differential equation into solving two ordinary differential equations. In addition, it may be necessary to consider interest rate models (5.11):

$$dr = u(r,t)dt + w(r,t)dX$$

with more general functions $u(r,t)$ and $w(r,t)$. For example, a model might be more useful if $u(r,t)$ and $w(r,t)$ is determined from the data of the short rate on the market. Also in order for the model to be more realistic, the model should guarantee that the random variable r will be in a finite interval $[r_l, r_u]$ in the future if r is in the interval $[r_l, r_u]$ now. According to Sect. 2.4, if u and w satisfy

$$\begin{cases} u(r_l,t) - w(r_l,t)\dfrac{\partial}{\partial r}w(r_l,t) \geq 0, \\ w(r_l,t) = 0 \end{cases} \tag{5.45}$$

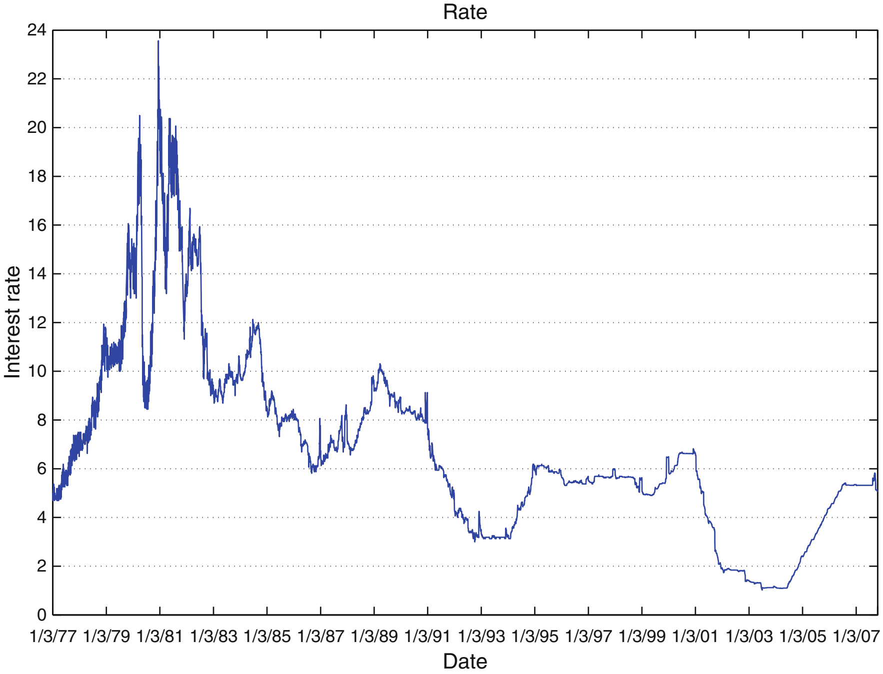

Fig. 5.1. One month LIBOR on US dollar during Jan 1977–Sep 2007

and

$$\begin{cases} u(r_u,t) - w(r_u,t)\dfrac{\partial}{\partial r}w(r_u,t) \le 0, \\ w(r_u,t) = 0, \end{cases} \tag{5.46}$$

then the random variable r is always in $[r_l, r_u]$. In what follows, we will describe a model having such properties.

The real data of the 1-month LIBOR (London Interbank Offer Rate) on U.S. dollar during January 1977–2010 is available and is shown as a curve in Fig. 5.1. From the data we know that the minimum interest rate $r_{\min}$ is 0.0022906 and the maximum interest rate $r_{\max}$ is 0.23562. Thus we assume that for the interval $[r_l, r_u]$ the lower bound r_l is 0.0 and the upper bound r_u is 0.24. From the data we can also determine the standard deviation of r for 40 values of r by statistics, which are shown as "$\circ$" in Fig. 5.2. Assuming

$$w(r) = (r - r_l)(r_u - r)(a_0 r^2 + b_0 r + c_0),$$

using the values of standard deviation of r obtained, and using the least squares method, we can find the values of a_0, b_0, and c_0, which are

$$a_0 = 4.1, \quad b_0 = -0.51, \quad c_0 = 0.0224.$$

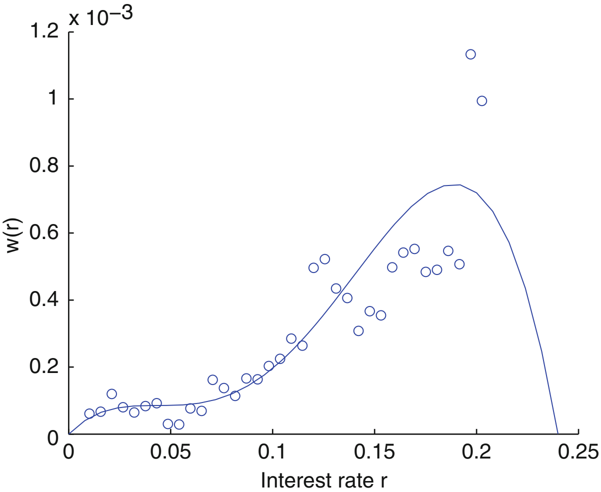

Fig. 5.2. $w(r)$ with $r_l = 0$ and $r_u = 0.24$

That is, the function $w(r)$ in the form above from the real data is:

$$w(r) = (r - r_l)(r_u - r)4.1r^2 - 0.51r + 0.0224$$

The curve of $w(r)$ is also given in Fig. 5.2. This function $w(r)$ satisfies the second conditions in the conditions (5.45) and (5.46). We can also find a function $u(r)$ satisfying the conditions (5.45) and (5.46), so r will be in $[r_l, r_u]$ if such a model is used. However here we do not discuss how to determine such a $u(r,t)$ from the real data. This is because, as we will see, we can choose $\lambda(r,t)$ so that $u(r,t)$ will not be used in order to do computation by using this model. If such a model is used, we have to solve PDE problems numerically in order to get market price of risk and the values of derivatives (for details, see the paper [72] by Shi). In what follows, we briefly discuss how to obtain the market price of risk numerically.

As pointed out in Sect. 5.2.2, if we use the model (5.11), then any interest rate derivative, $V(r,t)$, satisfies Eq. (5.12):

$$\frac{\partial V}{\partial t} + \frac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + (u - \lambda w)\frac{\partial V}{\partial r} - rV = 0, \qquad 0 \le t \le T,$$

where we assume that there is no coupon related to the derivative, so $kZ = 0$. This parabolic partial differential equation degenerates to a hyperbolic partial differential equation or an ordinary differential equation at $r = r_l$ and r_u when $w\ (r_l, t) = 0$ and $w(r_u, t) = 0$. Moreover, if the condition (5.45) holds, from Sect. 2.4.2, we see that no extra boundary condition at $r = r_l$ is needed in order to find a unique solution. Similarly, if the condition (5.46) holds, then no extra boundary condition at $r = r_u$ is needed. Consequently, the final value problem without any boundary conditions

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2}w^2 \dfrac{\partial^2 V}{\partial r^2} + (u - \lambda w)\dfrac{\partial V}{\partial r} - rV = 0, & r_l \le r \le r_u, \quad t \le T, \\ V(r,T) = f(r), & r_l \le r \le r_u \end{cases}$$

has a unique solution if the conditions (5.45) and (5.46) hold. As we have discussed, u and w can be determined from the historical data of the short rate on the market. However, in order to use this equation to price any derivatives, we need to know $\lambda(r,t)$. As soon as such a $\lambda(r,t)$ is determined, an interest rate model (5.11) becomes a no-arbitrage interest rate model. Thus, it is important in practice. Suppose $\lambda(r,t)$ is a function of t plus $u(r,t)/w(r)$, i.e., $\lambda(r,t) = \bar{\lambda}(t)+u(r,t)/w(r)$.[2] Then $\bar{\lambda}(t)$, as the solution of the following inverse problem, can be determined numerically by the term structure of interest rates or, equivalently, by the zero-coupon bond price curve. Suppose that $t = 0$ corresponds to today and today's short rate is r^*. Let $V(r,t;T^*)$ be the solution of the problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2}w^2 \dfrac{\partial^2 V}{\partial r^2} + \left[u - (\bar{\lambda}(t) + u/w)w\right] \dfrac{\partial V}{\partial r} - rV = 0, \\ \qquad\qquad\qquad\qquad r_l \le r \le r_u, \quad 0 \le t \le T^*, \\ V(r,T^*;T^*) = 1, \qquad r_l \le r \le r_u. \end{cases} \tag{5.47}$$

Here T^* is a parameter. We need to find a function $\bar{\lambda}(t)$ defined on $[0, T^*_{\max}]$ such that $V(r^*,0;T^*)$ is equal to the today's value of the zero-coupon bond maturing at time T^* and with a face value of one dollar for any $T^* \in [0, T^*_{\max}]$, where $T^*_{\max}$ in the longest maturity of zero-coupon bonds on the market.

In this problem, the value of $\bar{\lambda}(t)$ for $t \in [0, T^*_1]$ is determined by the value of zero-coupon bonds maturing at time $T^* \in [0, T^*_1]$. If $\bar{\lambda}(t)$ for $t \in [0, T^*_1]$ has been obtained and $T^*_2 > T^*_1$, then the value of $\bar{\lambda}(t)$ for $t \in [T^*_1, T^*_2]$ will be found by letting $V(r^*,0;T^*)$ be equal to the value of a zero-coupon bond maturing at time T^* for any $T^* \in [T^*_1, T^*_2]$. Therefore, the value of $\bar{\lambda}(t)$ at $t = T^*$ is determined by the value of a zero-coupon bond maturing at time T^* if the value of $\bar{\lambda}(t)$ for $t \in [0, T^*)$ has been obtained. In order to find the value of $\bar{\lambda}(T^*)$, we need to make a guess about it and solve the problem (5.47) from $t = T^*$ to $t = 0$ and then check if $V(r^*,0;T^*)$ is equal to the value of the zero-coupon bond maturing at time T^*. If T^* is 20 or 30 years, then the procedure of solving the problem (5.47) is quite long.

Actually the property of the function $\bar{\lambda}(t)$ has a close relation with the second derivative of the zero-coupon bond curve with respect to the maturity time T^* (see Sect. 10.1.1). If the zero-coupon bond curve is obtained by the cubic spline interpolation, then the second derivative is continuous, but the third derivative is discontinuous. The non-smoothness of $\bar{\lambda}(t)$ sometimes causes quite big oscillation of the solution of the bond equation if $T^*_{\max}$ is big.

[2] For such a choice of $\lambda(r,t)$, $u(r,t)-\lambda(r,t)w(r) = -\bar{\lambda}(t)w(r)$, so $u(r,t)$ disappears from the PDE. Thus we do not need $u(r,t)$ in order to solve the PDE.

5.5 Application of Bond Equations

The bond equation (5.12) can be applied to evaluating not only bonds but also bond options, options on bond futures contracts, swaps, caps, floors, collars, and even options on them. In what follows, we describe these applications.

5.5.1 Bond Options and Options on Bond Futures Contracts

A bond option is similar to an equity option except that the underlying asset is a bond. A bond depends on the interest r, and consequently, a bond option will also depend on r. Consider a T-year European option on a N-year bond. Suppose that the time today is zero. Then the bond should be issued on time T and will mature at time $T+N$. In what follows, let T_b denote $T+N$ and for simplicity, let the face value of the bond be equal to one. Thus, the bond price is the solution of the problem

$$\begin{cases} \dfrac{\partial V_b}{\partial t} + \dfrac{1}{2} w^2 \dfrac{\partial^2 V_b}{\partial r^2} + (u - \lambda w) \dfrac{\partial V_b}{\partial r} - r V_b + k = 0, \\ \qquad\qquad\qquad\qquad r_l \le r \le r_u, \quad T \le t \le T_b, \\ V_b(r, T_b; T_b) = 1, \quad r_l \le r \le r_u, \end{cases} \tag{5.48}$$

where we consider a coupon-bearing bond with a coupon payment $k(t)dt$ during a time period $[t, t+dt]$ and use V_b to represent the price of the bond. In practice the coupon is not paid continuously, the equation should be

$$\frac{\partial V_b}{\partial t} + \frac{1}{2} w^2 \frac{\partial^2 V_b}{\partial r^2} + (u - \lambda w) \frac{\partial V_b}{\partial r} - r V_b + \sum_i k_i \delta(t - t_i) = 0.$$

In this case V_b gives the quoted price (clean price). The price a purchaser needs to pay is the cash price (dirty price)—the clean price plus the accrued interest, which should be close to the price given by the model with a continuous coupon payment. Here, we assume that the conditions (5.45) and (5.46) hold, so at $r = r_l$ and $r = r_u$ the equation degenerates to a hyperbolic equation and does not require any boundary conditions. Every model can be modified locally, so the conditions (5.45) and (5.46) hold. Therefore, this assumption is realistic. We also assume that $\lambda(r,t)$ is known. A European call bond option is a contract whose holder has a right to purchase a bond at time T at a price E. Let $V(r,t)$ be the price of the option. Clearly, $V(r,t)$ should be the solution of the problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2} w^2 \dfrac{\partial^2 V}{\partial r^2} + (u - \lambda w) \dfrac{\partial V}{\partial r} - r V = 0, & r_l \le r \le r_u, \\ & t \le T, \\ V(r, T) = \max(V_b(r, T; T_b) - E, 0), & r_l \le r \le r_u. \end{cases} \tag{5.49}$$

For a European put bond option, the final condition is

$$V(r,T) = \max\left(E - V_b(r,T;T_b), 0\right).$$

For American call and put bond options, we need to require

$$V(r,t) \geq \max\left(V_b\left(r,t;t+N\right) - E, 0\right)$$

and

$$V(r,t) \geq \max(E - V_b(r,t;t+N), 0)$$

for $t \in [0,t]$, respectively. For example, if the option is on a 3-year bond, then $N = 3$. In this case, in order to determine the solution, we need to solve a problem involving free boundaries, and the constraint is a function of t. Therefore, this free-boundary problem is more complicated than that in equity option cases.

We can also determine the value of an option on a bond futures contract, which is denoted by $V(r,t)$ in what follows. Again, let T_b be the maturity date of the bond and T be the expiry of the option and the date the futures contract is initiated. Also, suppose that the futures contract is matured at time $T_f \in (T, T_b)$ and that the delivery price given in the option—the exercise price of the option is K. When $V(r,T)$ is given, we can obtain the value of the option today by solving a problem similar to the problem (5.49). How do we find $V(r,T)$?

Let $V_{b0}(r,t;T_f)$ be the value of the zero-coupon bond with maturity date T_f, which is the solution of the following problem

$$\begin{cases} \dfrac{\partial V_{b0}}{\partial t} + \dfrac{1}{2}w^2\dfrac{\partial^2 V_{b0}}{\partial r^2} + (u - \lambda w)\dfrac{\partial V_{b0}}{\partial r} - rV_{b0} = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad r_l \leq r \leq r_u, \quad t \leq T_f, \\ V_{b0}(r,T_f;T_f) = 1, \qquad\qquad\qquad r_l \leq r \leq r_u. \end{cases} \tag{5.50}$$

Then the value of the bond futures contract with a delivery price K given in the option can be expressed as

$$V_f(r,t;T_f) = V_b(r,t;T_b) - KV_{b0}(r,t;T_f) \tag{5.51}$$

for any $t \leq T_f$. Let K^* be the futures price for the futures contract with maturity date T_f at time T. K^* should be determined by the condition that the value of the futures contract is equal to zero when it is initiated at time T, i.e.,

$$V_f(r,T;T_f) = V_b(r,T;T_b) - K^*V_{b0}(r,T;T_f) = 0.$$

From this condition, we immediately know that the futures price K^* is equal to $V_b(r,T;T_b)/V_{b0}(r,T;T_f)$. If

$$K < K^* = V_b(r,T;T_b)/V_{b0}(r,T;T_f),$$

the holder of the option will exercise the option because the value of the bond futures contract

$$V_f(r, T; T_f) = V_b(r, T; T_b) - KV_{b0}(r, T; T_f) > 0.$$

Actually this is the value of the option for this case. If

$$K \geq K^* = V_b(r, T; T_b)/V_{b0}(r, T; T_f),$$

the value of the bond futures contract

$$V_f(r, T; T_f) = V_b(r, T; T_b) - KV_b(r, T; T_f) \leq 0,$$

and the holder will not exercise the option, which means $V(r, T) = 0$. Putting the two cases together, for $V(r, T)$ we have the following expression

$$V(r, T) = \max\left(V_b(r, T; T_b) - KV_{b0}(r, T; T_f), 0\right). \tag{5.52}$$

Therefore, we can first solve the problem (5.48) from T_b to T to get $V_b(r, T; T_b)$ and solve the problem (5.50) from T_f to T to get $V_{b0}(r, T; T_f)$, and then use the formula (5.52) in order to get $V(r, T)$. As soon as we have $V(r, T)$, we can solve the problem (5.49) with $V(r, T)$ as the final condition in order to find the price of the option on a bond futures contract today.

It is possible to consider V_b as a state variable and let the bond option price depend on V_b and t. For example, suppose

$$dV_b = \mu V_b dt + \sigma V_b dX,$$

where μ and σ is constant. In this case, we get the Black–Scholes equation with independent variables t and V_b, and use the Black–Scholes formulae to find the prices of European bond options. However, because the bond price must be equal to the face value at time T_b, which is often referred to as the pull-to-par phenomenon, a bond has different features from an equity, especially when $t \approx T_b$ (see Fig. 1.3). Therefore, even though a model in the form $dV_b = \mu V_b dt + \sigma V_b dX$ can describe the dynamics of an equity well, it could not state that of a bond. Consequently, the bond price obtained in this way is expected to have a large error, especially when $T \approx T_b$. If the model is in the form

$$dV_b = \alpha(t)(1 - V_b)dt + \sigma(t)V_b dX,$$

where $\alpha(t) \to \infty$ and $\sigma(t) \to 0$ as $t \to T_b$, then the result might be much better because such a model guarantees that V_b has a unique value one at time T_b. Of course, in this case it might be necessary to get solutions by numerical methods. Another problem of pricing a bond option in this way is to assume that the short rate is constant throughout the whole life of the option. If T is not small, it is not a good assumption.

Promising to pay an amount E at time T is equivalent to issuing a bond maturing at time T with a face value E. Thus, a right to pay E for a bond with a maturity date T_b at time T is the same right to exchange a bond of a face value E with a maturity date T for another bond with a maturity date T_b at time T. Therefore, a bond option can be understood as an exchange option that allows the holder to exchange a bond maturing at time T for another bond maturing at time T_b. If a bond option is dealt with in this way, it may be necessary to choose a model so that at least the random variable for the bond maturing at time T has the property of "pull-to-par."

5.5.2 Interest Rate Swaps and Swaptions

This subsection is devoted to plain vanilla interest rate swaps and options on such swaps—swaptions. As an example, let us look at the following N-year swap on a notional principal Q between a bank and a company.[3] In the swap, the bank and the company agree that during the next N years, the company will pay the bank the interest payment on the notional principal Q at a fixed rate $r_s(N)$ semiannually and in return, the bank will pay the company the interest payment on the same principal at a floating rate at the same times. Here, the floating rate in many interest rate swap agreements is the 6-month London Interbank Offer Rate (LIBOR) prevailing 6 months before the payment date. When the swap is initiated, both parties do not need to pay any money. Thus, the contract has no value at initiation. The fixed rate $r_s(N)$ is called the swap rate for an N year swap and determined through negotiation by the two parties. Clearly, the company wants $r_s(N)$ to be as small as possible, and the bank prefers a higher $r_s(N)$. What is the value of $r_s(N)$ both parties can accept? $r_s(N)$ should be a rate such that the value of the swap at initiation is zero. In order to know what equation $r_s(N)$ should satisfy, we need to find out how the value of the swap is related to r_s, where r_s denotes a swap rate that might not equal $r_s(N)$.

Suppose the swap is initiated at time T and today's time is $t^* \geq T$. The interest payments are exchanged semiannually at time

$$t_k = T + k/2,$$

$k = k^* + 1, k^* + 2, \cdots, 2N$, where k^* is the integer part of $2(t^* - T)$. Suppose today the price of the zero-coupon bond with a face value of one dollar and with maturity date t_k is $Z(t^*; t_k)$. In the swap given above, the company will pay cash $Qr_s/2$ at time $t_k, k = k^* + 1, k^* + 2, \cdots, 2N$. The present value of this cash flow is

$$\sum_{k=k^*+1}^{2N} \frac{Qr_s}{2} Z(t^*; t_k).$$

[3] A swap can also be between two companies.

At the same times, the bank will pay the company an amount of cash $\frac{Q}{2}\bar{f}(t_{k-1}, t_{k-1}, t_k)$ at time t_k, $k = k^*+1, k^*+2, \cdots, 2N$, where $\bar{f}(t_{k-1}, t_{k-1}, t_k)$ is the forward rate for the period $[t_{k-1}, t_k]$ determined at time t_{k-1} and we define $t_{k^*} = T + k^*/2$. Because $t_{k^*} \le t^*$, $\bar{f}(t_{k^*}, t_{k^*}, t_{k^*+1})$ is known today and the present value of the first payment is

$$\frac{Q}{2}Z(t^*; t_{k^*+1})\bar{f}(t_{k^*}, t_{k^*}, t_{k^*+1}).$$

What is the present value of the other payments? Suppose we deposit Q in the bank at time t_{k^*+1} for a period $[t_{k^*+1}, t_{k^*+2}]$ at a floating rate $f(t_{k^*+1}, t_{k^*+1}, t_{k^*+2})$. At time t_k, $k = k^*+2, k^*+3, \cdots, 2N-1$, we take the interest payment away and still leave Q in the bank for the next half year. In this way, we can generate a cash flow $\frac{Q}{2}\bar{f}(t_{k-1}, t_{k-1}, t_k)$ at time $t_k, k = k^*+2, k^*+3, \cdots, 2N-1$ and cash $\frac{Q}{2}\bar{f}(t_{2N-1}, t_{2N-1}, t_{2N}) + Q$ at time t_{2N}. Therefore, the value of the other payments is the difference between Q at time t_{k^*+1} and Q at time $t_{2N} = T + N$. Written mathematically, the present value of the other payments is

$$QZ(t^*; t_{k^*+1}) - QZ(t^*; T+N).$$

This result also can be obtained analytically. In fact, from the relation (5.6) we know that the forward interest rate compounded semiannually at time t_k during a period $[t_k, t_{k+1}]$ is

$$\bar{f}(t_k, t_k, t_{k+1}) = 2\left[\frac{Z(t_k; t_k)}{Z(t_k; t_{k+1})} - 1\right],$$

where $t_{k+1} = t_k + 1/2$. Therefore at time t^*, the value of the cash flow $\frac{Q}{2}$ $\bar{f}(t_k, t_k, t_{k+1})$ at time t_{k+1}, $k = k^*+1, k^*+2, \cdots, 2N-1$, is

$$\begin{aligned}
&\sum_{k=k^*+1}^{2N-1} \frac{Q}{2}\bar{f}(t_k, t_k, t_{k+1})Z(t^*; t_{k+1}) \\
&= Q\sum_{k=k^*+1}^{2N-1}\left[\frac{Z(t_k; t_k)}{Z(t_k; t_{k+1})} - 1\right]Z(t^*; t_{k+1}) \\
&= Q\sum_{k=k^*+1}^{2N-1}\frac{[Z(t_k; t_k) - Z(t_k; t_{k+1})]Z(t^*; t_k)}{Z(t_k; t_{k+1})Z(t^*; t_k)}Z(t^*; t_{k+1}) \\
&= Q\sum_{k=k^*+1}^{2N-1}[Z(t^*; t_k) - Z(t^*; t_{k+1})] \\
&= Q[Z(t^*; t_{k^*+1}) - Z(t^*; t_{2N})] \\
&= QZ(t^*; t_{k^*+1}) - QZ(t^*; T+N).
\end{aligned}$$

Let $V_s(t^*, r_s)$ be the present value of the swap to the company, which is the present value of the cash flow the company will receive minus the present value of the cash flow it will pay. From previous results, we arrive at

$$\begin{aligned}V_s(t^*;r_s) =& \frac{Q}{2}Z(t^*;t_{k^*+1})\bar{f}(t_{k^*},t_{k^*},t_{k^*+1}) + QZ(t^*;t_{k^*+1}) - QZ(t^*;T+N)\\ &- \sum_{k=k^*+1}^{2N} \frac{Qr_s}{2} Z(t^*;t_k)\\ =& QZ(t^*;t_{k^*+1})\left[1+\frac{1}{2}\bar{f}(t_{k^*},t_{k^*},t_{k^*+1})\right]\\ &- Q\left[Z(t^*;T+N) + \sum_{k=k^*+1}^{2N} \frac{r_s}{2} Z(t^*;t_k)\right]. \qquad (5.53)\end{aligned}$$

The expression $Q\left[Z(t^*;T+N) + \sum_{k=k^*+1}^{2N} \frac{r_s}{2} Z(t^*;t_k)\right]$ can be understood as the present value of a coupon-bearing bond, and the expression $QZ(t^*;t_{k^*+1}) \times \left[1+\frac{1}{2}\bar{f}(t_{k^*},t_{k^*},t_{k^*+1})\right]$ is the present value of another coupon-bearing bond. Therefore, a swap can be seen as a combination of a long position in one coupon-bearing bond with a short position in another coupon-bearing bond.

Here, we also need to point out that the values of a swap to two parties have the same magnitude but opposite signs. Thus, the value of the swap mentioned above to the bank is

$$\begin{aligned}&Q\left[Z(t^*;T_s+N) + \sum_{k=k^*+1}^{2N} \frac{r_s}{2} Z(t^*;t_k)\right]\\ &-QZ(t^*;t_{k^*+1})\left[1+\frac{1}{2}\bar{f}(t_{k^*},t_{k^*},t_{k^*+1})\right].\end{aligned}$$

In the case $t^* = T$, we have $k^* = 0, t_{k^*+1} = T + 1/2$ and

$$\bar{f}(T,T,T+1/2) = 2\left[\frac{1}{Z(T;T+1/2)} - 1\right],$$

that is,

$$Z(T;T+1/2)\left[1+\frac{1}{2}\bar{f}(T,T,T+1/2)\right] = 1, \qquad (5.54)$$

so we have

$$V_s(T;r_s) = Q\left[1 - Z(T;T+N) - \frac{r_s}{2}\sum_{k=1}^{2N} Z(T;T+k/2)\right]. \qquad (5.55)$$

As we have stated, when the swap is initiated, the value of the swap should be zero. Therefore, for the fixed rate in the contract we obtain

$$r_s(N) = 2\frac{1 - Z(T;T+N)}{\sum\limits_{k=1}^{2N} Z(T;T+k/2)}. \tag{5.56}$$

Therefore, between the swap rate for an N-year swap and $Z(T;T+k/2), k = 1, 2, \cdots, 2N$, there is a simple relation: $r_s(N)$ can be determined by $Z(T;T+k/2)$, $k = 1, 2, \cdots 2N$. This relation is true for $N = 1/2, 1, 3/2, \cdots$. Actually, $Z(T;T+k/2)$, $k = 1, 2, \cdots, 2N$, can also be obtained recursively by

$$Z(T;T+k/2) = \frac{1 - \dfrac{r_s(k/2)}{2}\sum\limits_{i=1}^{k-1} Z(T;T+i/2)}{1 + \dfrac{r_s(k/2)}{2}} \tag{5.57}$$

if $r_s(k/2), k = 1, 2, \cdots, 2N$ are given. Therefore, knowing $r_s(k/2)$ for different k is the same as knowing the yield curve.

As we have mentioned, a swap can be understood as the difference between two different coupon-bearing bonds. From the expression (5.53), we know that the face values of both bonds are Q. The expiration date of one bond is t_{k^*+1} and it pays a coupon $\dfrac{Q}{2}\bar{f}(t_{k^*}, t_{k^*}, t_{k^*+1})$ at $t = t_{k^*+1}$. Let V_i denote the value of this bond. The expiration date of the other bond is $T + N$, and it pays coupons $\dfrac{Qr_s}{2}$ semiannually starting at $t = t_{k^*+1}$. Let V_o represent the value of the other bond. The value of swap $V_s(t)$ is equal to $V_i - V_o$. Any bond can be priced by the bond equation. In fact, $V_i(r,t)$ is the solution of the problem

$$\begin{cases} \dfrac{\partial V_i}{\partial t} + \dfrac{1}{2}w^2\dfrac{\partial^2 V_i}{\partial r^2} + (u - \lambda w)\dfrac{\partial V_i}{\partial r} - rV_i = 0, & r_l \le r \le r_u,\ t^* \le t \le t_{k^*+1}, \\ V_i(r, t_{k^*+1}) = Q[1 + \bar{f}(t_{k^*}, t_{k^*}, t_{k^*+1})/2], & r_l \le r \le r_u \end{cases}$$

and $V_o(r,t)$ is the solution of the problem

$$\begin{cases} \dfrac{\partial V_o}{\partial t} + \dfrac{1}{2}w^2\dfrac{\partial^2 V_o}{\partial r^2} + (u - \lambda w)\dfrac{\partial V_o}{\partial r} - rV_o + \sum\limits_{k=k^*+1}^{2N} \dfrac{Qr_s}{2}\delta(t - t_k) = 0, \\ \qquad r_l \le r \le r_u,\ t^* \le t \le T + N, \\ V_o(r, T+N) = Q, \qquad r_l \le r \le r_u. \end{cases}$$

Let $r = r^*$ today and let $\lambda(r,t)$ be chosen so that $V(r^*, t^*; t_k) = Z(t^*; t_k)$, $k = k^* + 1, k^* + 2, \cdots, 2N$, where $V(r, t; t_k)$ is the solution of the following problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2}w^2\dfrac{\partial^2 V}{\partial r^2} + (u - \lambda w)\dfrac{\partial V}{\partial r} - rV = 0, \\ \qquad\qquad\qquad\qquad\qquad r_l \le r \le r_u,\ t^* \le t \le t_k, \\ V(r, t_k; t_k) = 1, \qquad\qquad r_l \le r \le r_u, \end{cases} \tag{5.58}$$

then

$$\begin{aligned} V_o(r^*, t^*; r_s) &= QV(r^*, t^*; t_{2N}) + \frac{Qr_s}{2}\sum_{k=k^*+1}^{2N} V(r^*, t^*; t_k) \\ &= Q\left[Z(t^*; t_{2N}) + \frac{r_s}{2}\sum_{k=k^*+1}^{2N} Z(t^*; t_k)\right] \end{aligned}$$

and

$$\begin{aligned} V_i(r^*, t^*) &= Q\left[1 + \bar{f}(t_{k^*}, t_{k^*}, t_{k^*+1})/2\right] V(r^*, t^*; t_{k^*+1}) \\ &= Q\left[1 + \bar{f}(t_{k^*}, t_{k^*}, t_{k^*+1})/2\right] Z(t^*; t_{k^*+1}). \end{aligned}$$

From these two expressions, we can see that

$$V_i(r^*, t^*) - V_o(r^*, t^*; r_s)$$

will have the same value as that given by the expression (5.53). When the bond equation is used, the value of the swap is not only given at $r = r^*$, and V_s is considered as a function of r and t, i.e., $V_s = V_s(r, t)$. The value of the swap is also dependent on the value of r_s. Therefore sometimes $V_s(r, t)$ is written as $V_s(r, t; r_s)$, where r_s is a parameter.

Indeed, in order to find $V_s(r, t)$, it is not necessary to find $V_i(r, t)$ and $V_o(r, t)$ separately; instead, we only need to solve

$$\begin{cases} \dfrac{\partial V_s}{\partial t} + \dfrac{1}{2}w^2\dfrac{\partial^2 V_s}{\partial r^2} + (u - \lambda w)\dfrac{\partial V_s}{\partial r} - rV_s - \displaystyle\sum_{k=k^*+1}^{2N} \dfrac{Qr_s}{2}\delta(t - t_k) \\ \qquad\qquad + Q\left[1 + \dfrac{\bar{f}(t_{k^*}, t_{k^*}, t_{k^*+1})}{2}\right]\delta(t - t_{k^*+1}) = 0, \\ \qquad\qquad\qquad\qquad r_l \le r \le r_u, \quad t^* \le t \le T + N, \\ V_s(r, T + N) = -Q, \quad r_l \le r \le r_u. \end{cases} \tag{5.59}$$

It is not difficult to show this conclusion, and we leave this proof to the reader as problem 15. Now we can find the value of a swap either using the formula (5.53) or solving the problem (5.59) and get the same answer. Many people will choose to calculate the value of the swap by the expression (5.53) because it is simple. Why do we need to consider problem (5.59)? It can provide some information on $\dfrac{\partial V_s(r^*, t)}{\partial r}$ and the bond equation will be useful when pricing a swaption by solving bond equations.

An option on a swap, or a swaption, is a contract to give the holder the right to enter into a certain interest rate swap by a certain time in the future. Consider a European swaption. Its holder has the right to choose if he should have an N-year swap at time T under which he will pay interest at a fixed rate r_{se} (the so-called exercise swap rate) and receive interest payment at a floating rate. Let r'_s be the N-year swap rate at time T, which can have infinitely many possible values. If $r_{se} < r'_s$, then the holder will choose to exercise the swaption because the value of a swap with a swap rate r'_s at time T is 0 and the value of a swap with a swap rate $r_{se} < r'_s$ should be positive, but the holder can enter into such a swap without paying any money. If $r_{se} > r'_s$, then the holder will choose not to exercise the option because the swap rate is lower on the market.

Such an option interests companies who plan to enter into a swap as a fixed rate payer because the swaption provides the companies with a guarantee that the fixed rate of interest they will pay on a loan will not exceed r_{se}.

According to the result (5.55), at time T, the values of the swaps with swap rates r'_s and r_{se} to the company are

$$V_s(T; r'_s) = Q\left[1 - Z(T; T+N) - \sum_{k=1}^{2N} \frac{r'_s}{2} Z(T; t_k)\right]$$

and

$$V_s(T; r_{se}) = Q\left[1 - Z(T; T+N) - \sum_{k=1}^{2N} \frac{r_{se}}{2} Z(T; t_k)\right]$$

respectively. If $r_{se} \le r'_s$, then the value of the swaption V at time T is

$$V(r'_s, T) = V_s(T; r_{se}) - V_s(T; r'_s) = Q\frac{r'_s - r_{se}}{2} \sum_{k=1}^{2N} Z(T; t_k);$$

while if $r_{se} > r'_s$, then $V(r'_s, T) = 0$. Consequently, the payoff of the swaption is

$$V(r'_s, T) = \frac{Q}{2} \sum_{k=1}^{2N} Z(T; t_k) \max(r'_s - r_{se}, 0). \tag{5.60}$$

Suppose that at time T, r'_s has a lognormal distribution with the following probability density function

$$G(r'_s) = \frac{1}{r'_s \sigma \sqrt{2\pi(T-t)}} \mathrm{e}^{-\left[\ln(r'_s/r_s) + \sigma^2(T-t)/2\right]^2 / 2\sigma^2(T-t)},$$

where r_s is the swap rate at time t. This model is often referred to as Black's model (see [8]). This probability density function is the probability density function (2.85) with $r - D_0 = 0$. Thus, the expectation of $\max(r'_s - r_{se}, 0)$ is $\mathrm{e}^{r(T-t)}$ times the price of a call option with $r - D_0 = 0$. That is

$$\mathrm{E}\left[\max(r_s' - r_{se}, 0)\right] = r_s N\left(\frac{\ln(r_s/r_{se}) + \sigma^2(T-t)/2}{\sigma\sqrt{T-t}}\right)$$
$$-r_{se} N\left(\frac{\ln(r_s/r_{se}) - \sigma^2(T-t)/2}{\sigma\sqrt{T-t}}\right).$$

Therefore at time t, the value of the payoff is

$$\begin{aligned}
&Z(t;T)\frac{Q}{2}\sum_{k=1}^{2N} Z(T;t_k)\\
&\quad\times\left[r_s N\left(\frac{\ln\frac{r_s}{r_{se}} + \frac{\sigma^2(T-t)}{2}}{\sigma\sqrt{T-t}}\right) - r_{se} N\left(\frac{\ln\frac{r_s}{r_{se}} - \frac{\sigma^2(T-t)}{2}}{\sigma\sqrt{T-t}}\right)\right]\\
&= \frac{Q}{2}\sum_{k=1}^{2N} Z(t;t_k)\\
&\quad\times\left[r_s N\left(\frac{\ln\frac{r_s}{r_{se}} + \frac{\sigma^2(T-t)}{2}}{\sigma\sqrt{T-t}}\right) - r_{se} N\left(\frac{\ln\frac{r_s}{r_{se}} - \frac{\sigma^2(T-t)}{2}}{\sigma\sqrt{T-t}}\right)\right],
\end{aligned}$$

where $Z(t;T)$ is the discounting factor between t and T and we have used the relation $Z(t;T)\ Z(T;t_k) = Z(t;t_k)$. European swaptions are frequently valued in this way. Obviously, it is an approximate method.

We may also evaluate the European swaption by solving bond equations. As is given by the formula (5.60), the payoff of the swaption is

$$\frac{Q}{2}\sum_{k=1}^{2N} Z(T;T+k/2)\max(r_s' - r_{se}, 0).$$

At time T, $r_s(N)$ is determined by the formula (5.56), i.e., r_s' is given by

$$2\frac{1 - Z(T;T+N)}{\sum\limits_{k=1}^{2N} Z(T;T+k/2)}.$$

Thus the payoff of the swaption can be rewritten as

$$\begin{aligned}
&\frac{Q}{2}\sum_{k=1}^{2N} Z(T;T+k/2)\max\left(2\frac{1 - Z(T;T+N)}{\sum\limits_{k=1}^{2N} Z(T;T+k/2)} - r_{se},\ 0\right)\\
&= Q\max\left(1 - Z(T;T+N) - \frac{r_{se}}{2}\sum_{k=1}^{2N} Z(T;T+k/2),\ 0\right).
\end{aligned}$$

Let $V_{so}(r,t)$ denote the value of the swaption. In this case the procedure of determining swaption price is divided into two steps. The first step is to determine the value of swap with r_{se} as the swap rate for all $r \in [r_l, r_u]$, $V_s(r,t;r_{se},T)$, where r_{se} and T are parameters and T denotes the starting date of the swap, and the second step is to obtain the payoff of swaption and to find the value of swaption. In order to get $V_s(r,t;r_{se},T)$, we need to solve the following problem

$$\begin{cases} \dfrac{\partial V_s}{\partial t} + \dfrac{1}{2}w^2\dfrac{\partial^2 V_s}{\partial r^2} + (u-\lambda w)\dfrac{\partial V_s}{\partial r} - rV_s \\ \qquad\qquad -\dfrac{Qr_{se}}{2}\displaystyle\sum_{k=1}^{2N}\delta(t-T-k/2) + Q\delta(t-T) = 0, \\ \qquad\qquad r_l \le r \le r_u, \quad T \le t \le T+N, \\ V_s(r,T+N) = -Q, \ r_l \le r \le r_u. \end{cases} \tag{5.61}$$

This problem actually is Eq. (5.59) with $k^* = 0$, $t^* = T$, and replacing

$$Q\left[1 + \frac{\bar{f}(t_{k^*}, t_{k^*}, t_{k^*+1})}{2}\right]\delta(t - t_{k^*+1})$$

by $Q\delta(t-T)$ because of the relation (5.54). After we obtain $V_s(r,T;r_{se},T)$, we need to get the payoff of the swaption

$$V_{so}(r,T) = \max\left(V_s(r,T;r_{se},T), 0\right)$$

and then to solve the following bond equation from $t=T$ to $t=0$ in order to find the value of swaption:

$$\begin{cases} \dfrac{\partial V_{so}}{\partial t} + \dfrac{1}{2}w^2\dfrac{\partial^2 V_{so}}{\partial r^2} + (u-\lambda w)\dfrac{\partial V_{so}}{\partial r} - rV_{so} = 0, & r_l \le r \le r_u, \\ & 0 \le t \le T, \\ V_{so}(r,T) = \max\left(V_s(r,T;r_{se},T), 0\right), & r_l \le r \le r_u. \end{cases} \tag{5.62}$$

The value of the swaption today is $V_{so}(r^*, 0)$, where r^* is the interest rate today. For American swaptions we can use a similar treatment to find their prices.

5.5.3 Interest Rate Caps, Floors, and Collars

An interest rate cap is a contract whose holder will receive some money from the issuer if the floating rate exceeds a certain level r_c, where r_c is called the cap rate. Therefore, interest rate caps can be used to provide insurance against the rate of interest on a floating-rate loan rising above a certain level.

For example, someone is going to get an N-year floating-rate loan with a principal amount Q at time t^*. The borrower will pay interest quarterly at time

$$t_k = t^* + k/4, \quad k = 1, 2, \cdots, 4N.$$

The rate is a 3-month LIBOR determined at time t_{k-1} for the period $[t_{k-1}, t_k]$, where we define $t_0 = t^*$. The LIBOR is a forward interest rate. According to the notation given in Sect. 5.1, $\bar{f}(t_{k-1}, t_{k-1}, t_k)$ stands for this rate. In what follows, we use the notation $\bar{f}_{k-1}$ instead of $\bar{f}(t_{k-1}, t_{k-1}, t_k)$ for brevity. The borrower is worrying that he will pay too much interest if the 3-month LIBOR becomes very high during the period $[t^*, t^* + N]$. Therefore, he is interested in such a cap: it starts from t^* and lasts N years, and at time t_k, the issuer of the cap will pay the holder an amount of cash $Q \max(\bar{f}_{k-1} - r_c, 0)/4$. Suppose he purchases this cap. Then when $\bar{f}_{k-1} < r_c$, he will pay interest payment on the loan $Q\bar{f}_{k-1}/4$ and receive zero from the issuer of the cap; whereas $\bar{f}_{k-1} > r_c$, his actual payment is $Qr_c/4$ because he receives $Q \max(\bar{f}_{k-1} - r_c, 0)/4$ from the cap. Hence the cap provides insurance against the interest rate on the floating-rate loan rising above an upper bound r_c.

How much should be paid in order to obtain such an insurance? The present value of the payment $Q \max(\bar{f}_{k-1} - r_c, 0)/4$ at time t_k is actually the value of a call option with expiry t_k. This call option is usually called the kth caplet. The LIBOR $\bar{f}_{k-1}$ is a forward rate determined at time t_{k-1} for the period $[t_{k-1}, t_k]$, so an amount $Q \max(\bar{f}_{k-1} - r_c, 0)/4$ at time t_k is equivalent to the amount

$$\frac{Q}{4(1 + \bar{f}_{k-1}/4)} \max(\bar{f}_{k-1} - r_c, 0) = \max\left(Q - Q\frac{1 + r_c/4}{1 + \bar{f}_{k-1}/4}, 0\right)$$

at time t_{k-1}. A loan with a face value $Q\,(1 + r_c/4)$ and maturity t_k is worth $Q(1 + r_c/4)/(1 + \bar{f}_{k-1}/4)$ at time t_{k-1} for any $\bar{f}_{k-1}$. Therefore, a caplet with a payoff $Q \max(\bar{f}_{k-1} - r_c, 0)/4$ at time t_k is equivalent to a put option with maturity t_{k-1} and a strike price Q on a zero-coupon bond with maturity t_k and a face value $Q\,(1 + r_c/4)$. At time $t^* (= t_0)$, the value of the first caplet is equal to a known value $\dfrac{Q}{4(1 + \bar{f}_0/4)} \max(\bar{f}_0 - r_c, 0)$. Usually, this value is excluded from the premium and there is no payment at time t_1 even if the LIBOR is greater than r_c. Thus, a cap comprises $4N - 1$ put options on zero-coupon bonds. Because a bond or an option on a bond can be seen as a derivative on the short rate r, their values can be calculated by the bond equation. Let the value of the bond with maturity t_k be $V_{bk}(r, t)$. Then, $V_{bk}(r, t)$ is the solution of the problem

$$\begin{cases} \dfrac{\partial V_{bk}}{\partial t} + \dfrac{1}{2}w^2\dfrac{\partial^2 V_{bk}}{\partial r^2} + (u - \lambda w)\dfrac{\partial V_{bk}}{\partial r} - rV_{bk} = 0, & r_l \le r \le r_u, \\ & t_{k-1} \le t \le t_k, \\ V_{bk}(r, t_k) = (1 + r_c/4)\,Q, & r_l \le r \le r_u, \end{cases} \tag{5.63}$$

where $k = 2, 3, \cdots, 4N$. Let $V_c(r,t)$ be the solution of the problem

$$\begin{cases} \dfrac{\partial V_c}{\partial t} + \dfrac{1}{2}w^2 \dfrac{\partial^2 V_c}{\partial r^2} + (u - \lambda w)\dfrac{\partial V_c}{\partial r} - rV_c \\ \qquad + \displaystyle\sum_{k=2}^{4N} \max(Q - V_{bk}(r, t_{k-1}), 0)\delta(t - t_{k-1}) = 0, \\ \qquad\qquad r_l \le r \le r_u, \quad t^* \le t \le t_{4N-1}, \\ V_c(r, t_{4N-1}) = 0, \qquad r_l \le r \le r_u. \end{cases} \tag{5.64}$$

Then, $V_c(r,t^*)$ gives the value of the cap — the total value of the $4N-1$ put options at time t^* and the premium of the cap is given by

$$V_c(r^*, t^*),$$

where r^* is the short rate at time t^*.

There are some other derivatives analogous to interest rate caps, such as interest rate floors and collars. A holder of a floor will receive some money from the issuer if the floating rate is below a certain level r_f, which is called the floor rate. If a borrower of a floating-rate loan believes that the floating rate will never be less than the lower bound r_f, then he may want to write such a floor. This is because he will get some money from writing a floor but, according to his opinion, he will not actually pay any money to the holder of the floor. Therefore, he hopes that he can reduce his expenses on the loan through writing a floor. If we assume that the floor starts at t^* and lasts N years, that the floating rate is 3-month LIBOR, and that the money will be paid quarterly at time t_k, $k = 2, 3, \cdots, 4N$, then the value of the floor is the sum of $4N-1$ floorlets that are call options on zero-coupon bonds. In order to determine the premium, we can first solve the problem

$$\begin{cases} \dfrac{\partial V_{bk}}{\partial t} + \dfrac{1}{2}w^2 \dfrac{\partial^2 V_{bk}}{\partial r^2} + (u - \lambda w)\dfrac{\partial V_{bk}}{\partial r} - rV_{bk} = 0, \quad r_l \le r \le r_u, \\ \qquad\qquad t_{k-1} \le t \le t_k, \\ V_{bk}(r, t_k) = (1 + r_f/4)\, Q, \qquad r_l \le r \le r_u \end{cases}$$

and get $V_{bk}(r, t_{k-1})$. Based on $V_{bk}(r, t_{k-1})$, we then can determine the solution of the problem

$$\begin{cases} \dfrac{\partial V_f}{\partial t} + \dfrac{1}{2}w^2 \dfrac{\partial^2 V_f}{\partial r^2} + (u - \lambda w)\dfrac{\partial V_f}{\partial r} - rV_f \\ \qquad + \displaystyle\sum_{k=2}^{4N} \max(V_{bk}(r, t_{k-1}) - Q, 0)\delta(t - t_{k-1}) = 0, \\ \qquad\qquad r_l \le r \le r_u, \quad t^* \le t \le t_{4N-1}, \\ V_f(r, t_{4N-1}) = 0, \qquad r_l \le r \le r_u \end{cases}$$

and the value $V_f(r^*, t^*)$ gives the premium of the floor. The derivation of this conclusion is left for readers as Problem 17.

A collar specifies both the upper bound r_c and the lower bound r_f. It may be understood as a combination of a long position in a cap with a short position in a floor. The value of a collar V_{co} is

$$V_{co} = V_c - V_f.$$

Usually, we choose r_c and r_f such that

$$V_c = V_f \quad \text{or} \quad V_{co} = 0.$$

It is clear that a portfolio of a collar and the original floating-rate loan is equivalent to a new loan with a floating rate in $[r_c, r_f]$. If

$$r_c = r_f,$$

then the collar actually becomes a swap based on 3-month LIBOR and with $4N - 1$ exchanges of payments. There exist other interest rate derivatives such as captions and floortions. Their evaluations are similar to what we have discussed.

5.6 Multi-Factor Interest Rate Models

5.6.1 Brief Description of Several Multi-Factor Interest Rate Models

Sometimes, it is necessary to assume that interest rate derivatives depend on not only the short rate r, but also some other random state variables. Because volatility is always a dominant factor in determining the prices of bonds and options, we need to have a more accurate model for volatility. It may be necessary to consider the interest rate volatility as a random variable. Fong and Vasicek [30] proposed such a two-factor model. In their model, they postulated that both the short rate r and the variance υ of the short rate are stochastic state variables and assumed

$$dr = (\bar{\mu} - \gamma r)dt + \sqrt{\upsilon}dX,$$

$$d\upsilon = (\nu - \eta\upsilon)dt + \xi\sqrt{\upsilon}dX_\upsilon,$$

$$\mathrm{E}\left[dXdX_\upsilon\right] = \rho dt,$$

where $\bar{\mu}, \gamma, \nu, \eta, \xi$ are constants and dX and dX_υ are two standard Wiener processes. As we can see in this model, the stochastic equation for r is the same as that in the Vasicek model, and r could become negative. Here, not only the short rate but also the variance possess the mean reversion property. In this case, Eq. (2.34) can be written as

$$\frac{\partial V}{\partial t} + \frac{1}{2}\upsilon\frac{\partial^2 V}{\partial r^2} + \rho\xi\upsilon\frac{\partial^2 V}{\partial r\partial \upsilon} + \frac{1}{2}\xi^2\upsilon\frac{\partial^2 V}{\partial \upsilon^2} + (\bar{\mu} - \gamma r - \bar{\lambda}\upsilon)\frac{\partial V}{\partial r}$$
$$+ \left[\nu - (\eta + \bar{\lambda}_\upsilon\xi)\upsilon\right]\frac{\partial V}{\partial \upsilon} - rV = 0,$$

where the market prices of risk for r and υ are $\bar{\lambda}\sqrt{\upsilon}$ and $\bar{\lambda}_\upsilon\sqrt{\upsilon}$, respectively, $\bar{\lambda}$ and $\bar{\lambda}_\upsilon$ being constants.

Brenman and Schwartz [13] considered another two-factor model. In their model, the two random state variables are the short-term interest rate r and the long-term interest rate l. They assumed

$$dr = u(r,l,t)dt + w(r,l,t)dX,$$

$$dl = u_l(r,l,t)dt + w_l(r,l,t)dX_l,$$

$$\mathrm{E}\left[dX dX_l\right] = \rho(r,l,t)dt,$$

where dX and dX_l are the standard Wiener processes. According to Eq. (2.34), any derivative dependent on r and l should satisfy

$$\frac{\partial V}{\partial t}+\frac{1}{2}w^2\frac{\partial^2 V}{\partial r^2}+\rho w w_l\frac{\partial^2 V}{\partial r\partial l}+\frac{1}{2}w_l^2\frac{\partial^2 V}{\partial l^2}+(u-\lambda w)\frac{\partial V}{\partial r}+(u_l-\lambda_l w_l)\frac{\partial V}{\partial l}-rV = 0.$$

For other models, for example, see [2, 19, 57]. From these models, we can have the corresponding partial differential equations. Any reader who is interested in knowing more about these models and other models is suggested to consult these papers and the book [47] by James and Webber.

In order to use these models to price derivatives, we need to determine these market prices of risk, which is similar to what we have done for one-factor models. Also, if we make some modifications on these models so that some conditions similar to the conditions (5.45) and (5.46) hold, then unique solutions of these equations can be obtained only by requiring final conditions.

Not only can the interest rates and their variances be taken as state variables. Heath et al. [38, 39, 40] suggested a model where the driving state variable of the model is $F(t,T)$, the forward rate at time t for instantaneous borrowing at a later time T. They assume

$$dF(t,T) = \alpha_F(t,T)dt + \sum_{i=1}^{n}\sigma_F^i(t,T)dX_i,$$

where dX_i is the ith Wiener process, and the n Wiener processes are independent. In this sense, it can be called a multi-factor model. Jarrow wrote a monograph on this method in 1996 (see [48]). Any reader who wants to know its details is referred to that book.

5.6.2 Reducing the Randomness of a Zero-Coupon Bond Curve to That of a Few Zero-Coupon Bonds

As we know, if we have an effective way to describe the randomness of a zero-coupon bond curve, then we can have an effective model for interest rate derivatives such as bond options or swaptions. In this and the next subsections, we discuss a three-factor model, which can be easily used in practice and generalized to the cases with more factors without any difficulty.

As we have done in Sect. 5.1, let $Z(t;t+T)$ denote the price of a T-year zero-coupon bond with a face value of one dollar at time t, and we use the notation $Z_i(t) = Z(t;t+T_i)$ for any T_i, $i = 0, 1, \cdots, N$. Here, we also assume $T_i < T_{i+1}$, for $i = 0, 1, \cdots, N-1$, and $T_0 = 0$. According to $Z_i(t), i = 0, 1, \cdots, N$, we can have an interpolation function $\bar{Z}(T;t)$ for $T \in [0, T_N]$ by requiring $\bar{Z}(T;t)$ to be a continuous function with continuous first and second derivatives in the form:

$$\bar{Z}(T;t) = \begin{cases} a_{0,1} + a_{1,1}T + a_{2,1}T^2, & 0 \le T \le T_1, \\ a_{0,i} + a_{1,i}T + a_{2,i}T^2 + a_{3,i}T^3, & T_{i-1} \le T \le T_i, \\ & i = 2, \cdots, N-1, \\ a_{0,N} + a_{1,N}T + a_{2,N}T^2, & T_{N-1} \le T \le T_N. \end{cases} \tag{5.65}$$

In this function, there are $4(N-2)+6 = 4N-2$ coefficients. Because we have $N+1$ conditions on the value of the function

$$\bar{Z}(T_i;t) = Z_i(t), \quad i = 0, 1, \cdots, N$$

and $3(N-1)$ continuity conditions on the function, first and second derivatives at $T_1, T_2, \cdots, T_{N-1}$, the total number of conditions is also $4N-2$. Therefore, it is possible that those coefficients in the expression (5.65) can be determined by these conditions uniquely. This interpolation method is called a cubic spline interpolation, and the way of determining the coefficients in the expression (5.65) will be given in Sect. 6.1.1. A zero-coupon bond curve is a monotone function with respect to T. If for a set of $Z_i(t)$, $i = 0, 1, \cdots, N$, the expression (5.65) does not possess this property, the approximation needs to be modified so that the monotonicity is guaranteed. This is important in practice.

We assume that $\bar{Z}(T;t)$ is a very good approximation to the zero-coupon bond curve $Z(t;t+T)$. In this way, a random curve is reduced to N random variables with a small error.

Now let us reduce the number of random variables from N to K by the principal component analysis. Suppose that we have N random variables

$$S_i, \quad i = 1, 2, \cdots, N$$

and the covariance between S_i and S_j is

$$\mathrm{Cov}[S_i S_j] = b_i b_j \rho_{i,j}, \quad i, j = 1, 2, \cdots, N,$$

where $-1 \le \rho_{i,j} = \rho_{j,i} \le 1$ and $\rho_{i,i} = 1$. Let

$$c_i^2 \quad \text{and} \quad \mathbf{a}_i = \begin{bmatrix} a_{i,1} \\ a_{i,2} \\ \vdots \\ a_{i,N} \end{bmatrix}, \quad i = 1, 2, \cdots, N,$$

be the eigenvalues and unit eigenvectors of the covariance matrix

$$\mathbf{B} = \begin{bmatrix} b_1^2 & b_1 b_2 \rho_{1,2} & \cdots & b_1 b_N \rho_{1,N} \\ b_2 b_1 \rho_{2,1} & b_2^2 & \cdots & b_2 b_N \rho_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ b_N b_1 \rho_{N,1} & b_N b_2 \rho_{N,2} & \cdots & b_N^2 \end{bmatrix}.$$

That is, there is the following relation:

$$\mathbf{B}\mathbf{A}^T = \mathbf{A}^T\mathbf{C} \quad \text{or} \quad \mathbf{A}\mathbf{B}\mathbf{A}^T = \mathbf{C},$$

where $\mathbf{A}^T$ is the transpose of $\mathbf{A}$ and

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} c_1^2 & 0 & \cdots & 0 \\ 0 & c_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_N^2 \end{bmatrix}.$$

Here $\mathbf{A}$ is an orthogonal matrix, i.e., $\mathbf{A}\mathbf{A}^T = \mathbf{I}$ because $\mathbf{B}$ is a symmetric matrix.

Let $\bar{S}_1, \bar{S}_2, \cdots, \bar{S}_N$ be N other random variables defined by

$$\begin{bmatrix} \bar{S}_1 \\ \bar{S}_2 \\ \vdots \\ \bar{S}_N \end{bmatrix} = \mathbf{A} \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_N \end{bmatrix}.$$

For simplicity, this relation can be written as

$$\bar{\mathbf{S}} = \mathbf{A}\mathbf{S},$$

where

$$\bar{\mathbf{S}} = \begin{bmatrix} \bar{S}_1 \\ \bar{S}_2 \\ \vdots \\ \bar{S}_N \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_N \end{bmatrix}.$$

Then

$$\begin{aligned}
\mathrm{Cov}\left[\bar{S}_i\bar{S}_j\right] &= \mathrm{E}\left[\left(\bar{S}_i - \mathrm{E}\left[\bar{S}_i\right]\right)\left(\bar{S}_j - \mathrm{E}\left[\bar{S}_j\right]\right)\right] \\
&= \mathrm{E}\left[\left(\sum_{k=1}^{N} a_{ik}\left(S_k - \mathrm{E}\left[S_k\right]\right)\right)\left(\sum_{l=1}^{N} a_{jl}\left(S_l - \mathrm{E}\left[S_l\right]\right)\right)\right] \\
&= \sum_{k=1}^{N}\sum_{l=1}^{N} a_{ik}a_{jl}\ \mathrm{Cov}\left[S_k S_l\right] \\
&= \begin{cases} 0, & i \neq j, \\ c_i^2, & i = j. \end{cases}
\end{aligned}$$

That is, $\mathbf{C}$ is the covariance matrix of the random vector $\bar{\mathbf{S}}$. We furthermore suppose that

$$c_i^2 \geq c_j^2 \quad \text{for} \quad i < j$$

and

$$c_i^2 \ll c_K^2, \quad i = K+1, \cdots, N.$$

Assume that on some day

$$\mathbf{S} = \begin{bmatrix} S_1^* \\ S_2^* \\ \vdots \\ S_N^* \end{bmatrix} \equiv \mathbf{S}^*$$

and

$$\bar{\mathbf{S}} = \mathbf{A}\begin{bmatrix} S_1^* \\ S_2^* \\ \vdots \\ S_N^* \end{bmatrix} = \begin{bmatrix} \bar{S}_1^* \\ \bar{S}_2^* \\ \vdots \\ \bar{S}_N^* \end{bmatrix} \equiv \bar{\mathbf{S}}^*.$$

Because c_i^2, $i = K+1, \cdots, N$ are very small, for a period starting from that day, we neglect the uncertainty caused by the last $N-K$ components of $\bar{\mathbf{S}}$. That is, we assume that in this period $\bar{\mathbf{S}}$ has the following form:

$$\bar{\mathbf{S}} = \begin{bmatrix} \bar{S}_1 \\ \vdots \\ \bar{S}_K \\ \bar{S}_{K+1}^* \\ \vdots \\ \bar{S}_N^* \end{bmatrix},$$

where $\bar{S}_1, \cdots, \bar{S}_K$ can take all possible values. In this case

$$\mathbf{S} = \mathbf{A}^T \begin{bmatrix} \bar{S}_1 \\ \vdots \\ \bar{S}_K \\ \bar{S}^*_{K+1} \\ \vdots \\ \bar{S}^*_N \end{bmatrix}. \tag{5.66}$$

Under this assumption, among $S_1, S_2, \cdots, S_N$, only K components are independent. Suppose

$$\begin{vmatrix} a_{1,1} & a_{2,1} & \cdots & a_{K,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{K,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,K} & a_{2,K} & \cdots & a_{K,K} \end{vmatrix} \neq 0.$$

Then, we can choose $S_1, S_2, \cdots, S_K$ as independent components. Rewrite Eq. (5.66) as

$$\begin{bmatrix} S_1 \\ \vdots \\ S_K \end{bmatrix} = \mathbf{A}_1^T \begin{bmatrix} \bar{S}_1 \\ \vdots \\ \bar{S}_K \end{bmatrix} + \mathbf{A}_2^T \begin{bmatrix} \bar{S}^*_{K+1} \\ \vdots \\ \bar{S}^*_N \end{bmatrix},$$

$$\begin{bmatrix} S_{K+1} \\ \vdots \\ S_N \end{bmatrix} = \mathbf{A}_3^T \begin{bmatrix} \bar{S}_1 \\ \vdots \\ \bar{S}_K \end{bmatrix} + \mathbf{A}_4^T \begin{bmatrix} \bar{S}^*_{K+1} \\ \vdots \\ \bar{S}^*_N \end{bmatrix},$$

where

$$\mathbf{A}_1^T = \begin{bmatrix} a_{1,1} & \cdots & a_{K,1} \\ \vdots & \ddots & \vdots \\ a_{1,K} & \cdots & a_{K,K} \end{bmatrix}, \quad \mathbf{A}_2^T = \begin{bmatrix} a_{K+1,1} & \cdots & a_{N,1} \\ \vdots & \ddots & \vdots \\ a_{K+1,K} & \cdots & a_{N,K} \end{bmatrix},$$

$$\mathbf{A}_3^T = \begin{bmatrix} a_{1,K+1} & \cdots & a_{K,K+1} \\ \vdots & \ddots & \vdots \\ a_{1,N} & \cdots & a_{K,N} \end{bmatrix}, \quad \mathbf{A}_4^T = \begin{bmatrix} a_{K+1,K+1} & \cdots & a_{N,K+1} \\ \vdots & \ddots & \vdots \\ a_{K+1,N} & \cdots & a_{N,N} \end{bmatrix}.$$

Then, for $S_{K+1}, \cdots, S_N$, we have

$$\begin{bmatrix} S_{K+1} \\ \vdots \\ S_N \end{bmatrix} = \mathbf{A}_3^T \left(\mathbf{A}_1^T\right)^{-1} \left(\begin{bmatrix} S_1 \\ \vdots \\ S_K \end{bmatrix} - \mathbf{A}_2^T \begin{bmatrix} \bar{S}^*_{K+1} \\ \vdots \\ \bar{S}^*_N \end{bmatrix} \right) + \mathbf{A}_4^T \begin{bmatrix} \bar{S}^*_{K+1} \\ \vdots \\ \bar{S}^*_N \end{bmatrix}. \tag{5.67}$$

Thus, for given $S_1, \cdots, S_K$, using the relation (5.67) we can get all other components of a vector $\mathbf{S}$. Consequently, the relation (5.67) defines a class of vectors with K parameters. That is, by the relation (5.67), we actually determine a class of $\mathbf{S}$, where only $S_1, \cdots, S_K$ are independent. Here, we take $S_1, \cdots, S_K$ as independent components. However, it is also possible to choose other K components as independent components.

Letting $S_i = Z_i/T_i$, $i = 1, 2, \cdots, N$, by the principal component analysis described above, we can find a class of vectors $[Z_1/T_1, \cdots, Z_N/T_N]^T$ with K parameters[4] and using the cubic spline interpolation given at the beginning of this subsection, we can further determine the curve $\bar{Z}(T;t)$ for $T \in [0, T_N]$. From the books by Jarrow [48], Hull [43], James and Webber [47], and Wilmott [83], we know that K usually is equal to three or four for the random curves related to interest rates. Thus, all the curves determined by the relation (5.67) form a class of curves with three or four parameters. The zero-coupon bond curve at that day is one of such curves, and the projections of any vector $\mathbf{S}$ determined by the relation (5.67) on the eigenvectors corresponding to the eigenvalues $c_{K+1}, \cdots, c_N$ are the same as those of $\mathbf{S}^*$. Those projections are different for different $\mathbf{S}^*$, so this is a feature belonging to $\mathbf{S}^*$. It is clear that the class of curves with such a feature needs to be considered most for derivative-pricing problems. Hence, when $K = 3$ or 4, the class contains all possible and need-to-be-considered-most zero-coupon bond curves. As soon as we have a zero-coupon bond curve, we can determine various interest rates at t, including the short rate $r(Z_1, \cdots, Z_K, t)$ at time t. For example, for $r(Z_1, \cdots, Z_K, t)$, we have

$$r(Z_1, \cdots, Z_K, t) = -\left.\frac{\partial \bar{Z}(T;t)}{\partial T}\right|_{T=0}. \tag{5.68}$$

5.6.3 A Three-Factor Interest Rate Model and the Equation for Interest Rate Derivatives

Suppose Z_1, Z_2 and Z_3 are prices of zero-coupon bonds with maturities T_1, T_2, and T_3, respectively. Assume $T_1 < T_2 < T_3$, which implies the relations $1 \geq Z_1 \geq Z_2 \geq Z_3$. Furthermore, we assume $Z_1 \geq Z_{1,l}$, $Z_2 \geq Z_{2,l}$ and $Z_3 \geq Z_{3,l}$, where $Z_{1,l} \geq Z_{2,l} \geq Z_{3,l} \geq 0$. Z_1, Z_2 and Z_3 are random variables and satisfy the system of stochastic differential equations:

$$dZ_i = \mu_i\left(Z_1, Z_2, Z_3, t\right) dt + \sigma_i\left(Z_1, Z_2, Z_3, t\right) dX_i, \quad i = 1, 2, 3$$

on the domain Ω: $\{Z_{1,l} \leq Z_1 \leq 1,\ Z_{2,l} \leq Z_2 \leq Z_1,\ Z_{3,l} \leq Z_3 \leq Z_2\}$. dX_i are the Wiener processes and $\mathrm{E}\left[dX_i dX_j\right] = \rho_{i,j} dt$ with $-1 \leq \rho_{i,j} \leq 1$. The coefficients μ_i, σ_i and their first- and second-order derivatives are assumed to be bounded on the domain Ω. On the six boundaries of Ω, the following conditions hold:

[4] If the conditions $Z_i \geq Z_{i+1}$, $i = 0, 1, \cdots N-1$ are not satisfied, then some modification needs to be done in order to guarantee the monotonicity.

(i) On surface I: $\{Z_1 = Z_{1,l},\ Z_{2,l} \le Z_2 \le Z_1,\ Z_{3,l} \le Z_3 \le Z_2\}$,

$$\begin{cases} \mu_1\left(Z_{1,l}, Z_2, Z_3, t\right) \ge 0, \\ \sigma_1\left(Z_{1,l}, Z_2, Z_3, t\right) = 0; \end{cases} \tag{5.69}$$

(ii) On surface II: $\{Z_1 = 1,\ Z_{2,l} \le Z_2 \le Z_1,\ Z_{3,l} \le Z_3 \le Z_2\}$,

$$\begin{cases} \mu_1\left(1, Z_2, Z_3, t\right) \le 0, \\ \sigma_1\left(1, Z_2, Z_3, t\right) = 0; \end{cases} \tag{5.70}$$

(iii) On surface III: $\{Z_{1,l} \le Z_1 \le 1,\ Z_2 = Z_{2,l},\ Z_{3,l} \le Z_3 \le Z_2\}$,

$$\begin{cases} \mu_2\left(Z_1, Z_{2,l}, Z_3, t\right) \ge 0, \\ \sigma_2\left(Z_1, Z_{2,l}, Z_3, t\right) = 0; \end{cases} \tag{5.71}$$

(iv) On surface IV: $\{Z_{1,l} \le Z_1 \le 1,\ Z_2 = Z_1,\ Z_{3,l} \le Z_3 \le Z_2\}$,

$$\begin{cases} -\mu_1\left(Z_1, Z_1, Z_3, t\right) + \mu_2\left(Z_1, Z_1, Z_3, t\right) \le 0, \\ \sigma_1\left(Z_1, Z_1, Z_3, t\right) = \sigma_2\left(Z_1, Z_1, Z_3, t\right), \quad \rho_{1,2}\left(Z_1, Z_1, Z_3, t\right) = 1; \end{cases} \tag{5.72}$$

(v) On surface V: $\{Z_{1,l} \le Z_1 \le 1,\ Z_{2,l} \le Z_2 \le Z_1,\ Z_3 = Z_{3,l}\}$,

$$\begin{cases} \mu_3\left(Z_1, Z_2, Z_{3,l}, t\right) \ge 0, \\ \sigma_3\left(Z_1, Z_2, Z_{3,l}, t\right) = 0; \end{cases} \tag{5.73}$$

(vi) On surface VI: $\{Z_{1,l} \le Z_1 \le 1,\ Z_{2,l} \le Z_2 \le Z_1,\ Z_3 = Z_2\}$,

$$\begin{cases} -\mu_2\left(Z_1, Z_2, Z_2, t\right) + \mu_3\left(Z_1, Z_2, Z_2, t\right) \le 0, \\ \sigma_2\left(Z_1, Z_2, Z_2, t\right) = \sigma_3\left(Z_1, Z_2, Z_2, t\right), \quad \rho_{2,3}\left(Z_1, Z_2, Z_2, t\right) = 1. \end{cases} \tag{5.74}$$

This model will be called the three-factor interest rate model in this book.

As you can see, conditions (5.69)–(5.71) and (5.73) have the same form as the condition (5.45) or the condition (5.46), and the conditions (5.72) and (5.74) are in a similar form. They are the weak-form reversion conditions on the non-rectangular domain Ω. In order to guarantee that if a point is in Ω at time t^*, then the point is still in Ω at $t = t^* + dt$ for a positive dt, it is necessary to require that

$$n_1 dZ_1 + n_2 dZ_2 + n_3 dZ_3 \le 0 \tag{5.75}$$

holds at any point on the boundary of the domain Ω, where n_1, n_2, and n_3 are the three components of the outer normal vector of the boundary at the

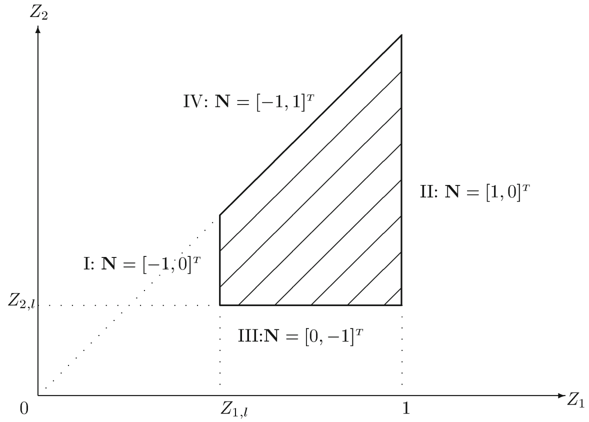

Fig. 5.3. Projection of the domain Ω on the (Z_1, Z_2)-plane

point. This is called the weak-form reversion conditions on a general domain. The condition of the condition (5.75) holding at every point on the boundary of the domain Ω is equivalent to the conditions (5.69)–(5.74). For example, on surface I (see Fig. 5.3), $n_1 = -1$, $n_2 = 0$ and $n_3 = 0$, so the condition (5.75) can be written as

$$n_1 dZ_1 + n_2 dZ_2 + n_3 dZ_3 = -dZ_1 = -\mu_1 dt - \sigma_1 dX_1 \leq 0.$$

This holds if and only if $\sigma_1 = 0$ and $\mu_1 \geq 0$. On surface IV, $n_1 = -1$, $n_2 = 1$, and $n_3 = 0$ (see Fig. 5.3). In this case the condition (5.75) can be written as

$$\begin{aligned} n_1 dZ_1 + n_2 dZ_2 + n_3 dZ_3 &= -dZ_1 + dZ_2 \\ &= -\mu_1 dt + \mu_2 dt - \sigma_1 dX_1 + \sigma_2 dX_2 \\ &= (-\mu_1 + \mu_2) dt + \sigma_{12} dX_{12} \leq 0, \end{aligned}$$

where we define

$$\sigma_{12} dX_{12} = -\sigma_1 dX_1 + \sigma_2 dX_2$$

and dX_{12} is another Wiener process. Using Itô's lemma, we know

$$\sigma_{12} = \sqrt{\sigma_1^2 - 2\rho_{1,2}\sigma_1\sigma_2 + \sigma_2^2}.$$

Thus in this case the condition (5.75) holds if and only if

$$-\mu_1 + \mu_2 \leq 0 \quad \text{and} \quad \sigma_{12} = 0.$$

$\sigma_{12} = 0$ is equivalent to

$$\sigma_1^2 - 2\rho_{1,2}\sigma_1\sigma_2 + \sigma_2^2 = (\sigma_1 - \sigma_2)^2 + 2(1 - \rho_{1,2})\sigma_1\sigma_2 = 0$$

or

$$\sigma_1 = \sigma_2 \quad \text{and} \quad \rho_{1,2} = 1.$$

Thus in this case the condition (5.75) is equivalent to $-\mu_1 + \mu_2 \le 0$, $\sigma_1 = \sigma_2$, and $\rho_{1,2} = 1$. If the derivatives of $\sigma_i(Z_1, Z_2, Z_3, t)$ with respect to Z_1, Z_2, and Z_3 are bounded, then it is expected that the condition (5.75) or the conditions (5.69)–(5.74) guarantee that a point (Z_1, Z_2, Z_3) will never move from inside of the domain Ω to its outside. This is a natural property of a stochastic model for interest rates when $Z_{1,l}$, $Z_{2,l}$ and $Z_{3,l}$ are given properly.

Let $V(Z_1, Z_2, Z_3, t)$ be the value of a derivative security depending on Z_1, Z_2, Z_3, t. According to Sect. 2.3.2, $V(Z_1, Z_2, Z_3, t)$ should satisfy

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_i\sigma_j\rho_{i,j}\frac{\partial^2 V}{\partial Z_i \partial Z_j} + r\sum_{i=1}^{3} Z_i \frac{\partial V}{\partial Z_i} - rV = 0.$$

As we pointed out in Sect. 2.3, in this case in the PDE there is no market price of risk, or because zero-coupon bonds can be traded on the market, the market prices of risk for these bonds can be determined by the relation (2.36) with $D_{0i} = 0$:

$$\mu_i(Z_1, Z_2, Z_3, t) - \lambda_i(Z_1, Z_2, Z_3, t)\sigma_i(Z_1, Z_2, Z_3, t) = r(Z_1, Z_2, Z_3, t)Z_i, \quad i = 1, 2, 3.$$

Let

$$\mathbf{L}_{3z} = \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_i\sigma_j\rho_{i,j}\frac{\partial^2}{\partial Z_i \partial Z_j} + r\sum_{i=1}^{3} Z_i \frac{\partial}{\partial Z_i} - r. \tag{5.76}$$

The equation above can be written as

$$\frac{\partial V}{\partial t} + \mathbf{L}_{3z} V = 0.$$

For a derivative security, at the maturity date T, its price should be equal to its payoff $V_T(Z_1, Z_2, Z_3)$. Therefore, any European interest rate derivatives under this model should be solutions of the problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + \mathbf{L}_{3z} V = 0 & \text{on} \quad \Omega \times [0, T], \\ V(Z_1, Z_2, Z_3, T) = V_T(Z_1, Z_2, Z_3) & \text{on} \quad \Omega. \end{cases} \tag{5.77}$$

Introduce the following transformation:

$$\begin{cases} \xi_1 = \dfrac{Z_1 - Z_{1,l}}{1 - Z_{1,l}}, \\ \xi_2 = \dfrac{Z_2 - Z_{2,l}}{Z_1 - Z_{2,l}}, \\ \xi_3 = \dfrac{Z_3 - Z_{3,l}}{Z_2 - Z_{3,l}}. \end{cases} \tag{5.78}$$

Through this transformation, the domain Ω in the (Z_1, Z_2, Z_3)-space is transformed into the domain $\tilde{\Omega}$: $[0,1]\times[0,1]\times[0,1]$ in the (ξ_1, ξ_2, ξ_3)-space. Because

$$\begin{aligned}
&\frac{\partial \xi_1}{\partial Z_1} = \frac{1}{1 - Z_{1,l}},\\
&\frac{\partial \xi_2}{\partial Z_1} = \frac{-\xi_2}{Z_1 - Z_{2,l}}, \quad \frac{\partial \xi_2}{\partial Z_2} = \frac{1}{Z_1 - Z_{2,l}},\\
&\frac{\partial \xi_3}{\partial Z_2} = \frac{-\xi_3}{Z_2 - Z_{3,l}}, \quad \frac{\partial \xi_3}{\partial Z_3} = \frac{1}{Z_2 - Z_{3,l}},
\end{aligned}$$

we have

$$\begin{aligned}
\frac{\partial V}{\partial Z_1} &= \frac{1}{1 - Z_{1,l}} \frac{\partial V}{\partial \xi_1} - \frac{\xi_2}{Z_1 - Z_{2,l}} \frac{\partial V}{\partial \xi_2},\\
\frac{\partial V}{\partial Z_2} &= \frac{1}{Z_1 - Z_{2,l}} \frac{\partial V}{\partial \xi_2} - \frac{\xi_3}{Z_2 - Z_{3,l}} \frac{\partial V}{\partial \xi_3},\\
\frac{\partial V}{\partial Z_3} &= \frac{1}{Z_2 - Z_{3,l}} \frac{\partial V}{\partial \xi_3},\\
\frac{\partial^2 V}{\partial Z_1^2} &= \frac{1}{(1 - Z_{1,l})^2} \frac{\partial^2 V}{\partial \xi_1^2} - \frac{2\xi_2}{(1 - Z_{1,l})(Z_1 - Z_{2,l})} \frac{\partial^2 V}{\partial \xi_1 \partial \xi_2}\\
&\quad + \frac{\xi_2^2}{(Z_1 - Z_{2,l})^2} \frac{\partial^2 V}{\partial \xi_2^2} + \frac{2\xi_2}{(Z_1 - Z_{2,l})^2} \frac{\partial V}{\partial \xi_2},\\
\frac{\partial^2 V}{\partial Z_2^2} &= \frac{1}{(Z_1 - Z_{2,l})^2} \frac{\partial^2 V}{\partial \xi_2^2} - \frac{2\xi_3}{(Z_1 - Z_{2,l})(Z_2 - Z_{3,l})} \frac{\partial^2 V}{\partial \xi_2 \partial \xi_3}\\
&\quad + \frac{\xi_3^2}{(Z_2 - Z_{3,l})^2} \frac{\partial^2 V}{\partial \xi_3^2} + \frac{2\xi_3}{(Z_2 - Z_{3,l})^2} \frac{\partial V}{\partial \xi_3},\\
\frac{\partial^2 V}{\partial Z_3^2} &= \frac{1}{(Z_2 - Z_{3,l})^2} \frac{\partial^2 V}{\partial \xi_3^2},\\
\frac{\partial^2 V}{\partial Z_1 \partial Z_2} &= \frac{-1}{(Z_1 - Z_{2,l})^2} \frac{\partial V}{\partial \xi_2} + \frac{1}{Z_1 - Z_{2,l}} \left(\frac{1}{1 - Z_{1,l}} \frac{\partial^2 V}{\partial \xi_1 \partial \xi_2} - \frac{\xi_2}{Z_1 - Z_{2,l}} \frac{\partial^2 V}{\partial \xi_2^2} \right)\\
&\quad - \frac{\xi_3}{Z_2 - Z_{3,l}} \left(\frac{1}{1 - Z_{1,l}} \frac{\partial^2 V}{\partial \xi_1 \partial \xi_3} - \frac{\xi_2}{Z_1 - Z_{2,l}} \frac{\partial^2 V}{\partial \xi_2 \partial \xi_3} \right),\\
\frac{\partial^2 V}{\partial Z_1 \partial Z_3} &= \frac{1}{Z_2 - Z_{3,l}} \left(\frac{1}{1 - Z_{1,l}} \frac{\partial^2 V}{\partial \xi_1 \partial \xi_3} - \frac{\xi_2}{Z_1 - Z_{2,l}} \frac{\partial^2 V}{\partial \xi_2 \partial \xi_3} \right),\\
\frac{\partial^2 V}{\partial Z_2 \partial Z_3} &= \frac{-1}{(Z_2 - Z_{3,l})^2} \frac{\partial V}{\partial \xi_3} + \frac{1}{Z_2 - Z_{3,l}} \left(\frac{1}{Z_1 - Z_{2,l}} \frac{\partial^2 V}{\partial \xi_2 \partial \xi_3} - \frac{\xi_3}{Z_2 - Z_{3,l}} \frac{\partial^2 V}{\partial \xi_3^2} \right).
\end{aligned}$$

Therefore, the operator $\mathbf{L}_{3Z}$ defined by the expression (5.76) can be rewritten as

$$\begin{aligned}\mathbf{L}_{3\xi} = {} & \frac{1}{2}\tilde{\sigma}_1^2\frac{\partial^2}{\partial\xi_1^2} + \frac{1}{2}\tilde{\sigma}_2^2\frac{\partial^2}{\partial\xi_2^2} + \frac{1}{2}\tilde{\sigma}_3^2\frac{\partial^2}{\partial\xi_3^2} \\ & + \tilde{\sigma}_1\tilde{\sigma}_2\tilde{\rho}_{1,2}\frac{\partial^2}{\partial\xi_1\partial\xi_2} + \tilde{\sigma}_1\tilde{\sigma}_3\tilde{\rho}_{1,3}\frac{\partial^2}{\partial\xi_1\partial\xi_3} + \tilde{\sigma}_2\tilde{\sigma}_3\tilde{\rho}_{2,3}\frac{\partial^2}{\partial\xi_2\partial\xi_3} \\ & + b_1\frac{\partial}{\partial\xi_1} + b_2\frac{\partial}{\partial\xi_2} + b_3\frac{\partial}{\partial\xi_3} - r, \end{aligned} \tag{5.79}$$

where

$$\begin{cases} \dfrac{1}{2}\tilde{\sigma}_1^2 = \dfrac{\frac{1}{2}\sigma_1^2}{(1-Z_{1,l})^2}, \\ \dfrac{1}{2}\tilde{\sigma}_2^2 = \dfrac{\frac{1}{2}\left(\sigma_1^2\xi_2^2 - 2\sigma_1\sigma_2\xi_2\rho_{1,2} + \sigma_2^2\right)}{(Z_1 - Z_{2,l})^2}, \\ \dfrac{1}{2}\tilde{\sigma}_3^2 = \dfrac{\frac{1}{2}\left(\sigma_2^2\xi_3^2 - 2\sigma_2\sigma_3\xi_3\rho_{2,3} + \sigma_3^2\right)}{(Z_2 - Z_{3,l})^2}, \end{cases} \tag{5.80}$$

$$\begin{cases} \tilde{\sigma}_1\tilde{\sigma}_2\tilde{\rho}_{1,2} = \dfrac{\sigma_1\left(\sigma_2\rho_{1,2} - \sigma_1\xi_2\right)}{(1-Z_{1,l})(Z_1 - Z_{2,l})}, \\ \tilde{\sigma}_1\tilde{\sigma}_3\tilde{\rho}_{1,3} = \dfrac{\sigma_1\left(\sigma_3\rho_{1,3} - \sigma_2\rho_{1,2}\xi_3\right)}{(1-Z_{1,l})(Z_2 - Z_{3,l})}, \\ \tilde{\sigma}_2\tilde{\sigma}_3\tilde{\rho}_{2,3} = \dfrac{\sigma_1\xi_2\left(\sigma_2\rho_{1,2}\xi_3 - \sigma_3\rho_{1,3}\right) + \sigma_2\left(\sigma_3\rho_{2,3} - \sigma_2\xi_3\right)}{(Z_1 - Z_{2,l})(Z_2 - Z_{3,l})}, \end{cases} \tag{5.81}$$

and

$$\begin{cases} b_1 = \dfrac{rZ_1}{1 - Z_{1,l}}, \\ b_2 = \dfrac{r\left(Z_2 - Z_1\xi_2\right)}{Z_1 - Z_{2,l}} + \dfrac{\sigma_1\left(\sigma_1\xi_2 - \sigma_2\rho_{1,2}\right)}{(Z_1 - Z_{2,l})^2}, \\ b_3 = \dfrac{r\left(Z_3 - Z_2\xi_3\right)}{Z_2 - Z_{3,l}} + \dfrac{\sigma_2\left(\sigma_2\xi_3 - \sigma_3\rho_{2,3}\right)}{(Z_2 - Z_{3,l})^2}. \end{cases} \tag{5.82}$$

Consequently, the problem (5.77) can be rewritten as

$$\begin{cases} \dfrac{\partial V}{\partial t} + \mathbf{L}_{3\xi}V = 0 \quad \text{on } \tilde{\Omega}\times[0,T], \\ V(\xi_1,\xi_2,\xi_3,T) = V_T\left(Z_1(\xi_1), Z_2(\xi_1,\xi_2), Z_3(\xi_1,\xi_2,\xi_3)\right) \quad \text{on } \tilde{\Omega}, \end{cases} \tag{5.83}$$

where $\mathbf{L}_{3\xi}$ is defined by Eq. (5.79) and

$$\begin{cases} Z_1(\xi_1) = Z_{1,l} + \xi_1(1 - Z_{1,l}), \\ Z_2(\xi_1,\xi_2) = Z_{2,l} + \xi_2\left[Z_{1,l} + \xi_1(1 - Z_{1,l}) - Z_{2,l}\right], \\ Z_3(\xi_1,\xi_2,\xi_3) = Z_{3,l} + \xi_3\left\{Z_{2,l} + \xi_2\left[Z_{1,l} + \xi_1(1 - Z_{1,l}) - Z_{2,l}\right] - Z_{3,l}\right\}. \end{cases} \tag{5.84}$$

This is a final-value problem on a rectangular domain. Thus, when the three-factor interest rate model is used, evaluating an interest rate derivative is reduced to solving a final-value problem on a rectangular domain.

We would like to point out the relations among $\tilde{\sigma}_1$, $\tilde{\sigma}_2$, $\tilde{\sigma}_3$, $\tilde{\rho}_{1,2}$, $\tilde{\rho}_{1,3}$, $\tilde{\rho}_{2,3}$ and $d\xi_1$, $d\xi_2$, $d\xi_3$. Using Itô's lemma, from the definitions of ξ_1, ξ_2, ξ_3, we can have

$$d\xi_1 = \tilde{\mu}_1 dt + \tilde{\sigma}_1 dX_1, \quad d\xi_2 = \tilde{\mu}_2 dt + \tilde{\sigma}_2 d\tilde{X}_2, \quad d\xi_3 = \tilde{\mu}_3 dt + \tilde{\sigma}_3 d\tilde{X}_3,$$

where $d\tilde{X}_2$ and $d\tilde{X}_3$ are two new Wiener processes. Therefore

$$\tilde{\sigma}_i^2 = \text{Var}[d\xi_i]/dt, \quad j = 1, 2, 3.$$

It can also be shown that

$$\text{Cov}[dX_1 d\tilde{X}_2]/dt = \tilde{\rho}_{1,2}, \quad \text{Cov}[dX_1 d\tilde{X}_3]/dt = \tilde{\rho}_{1,3}$$

and

$$\text{Cov}[d\tilde{X}_2 d\tilde{X}_3]/dt = \tilde{\rho}_{2,3}.$$

These are left for readers to prove as Problem 24.

From Eq. (5.80), it is easy to see that the equality conditions in the conditions (5.69)–(5.74) can be rewritten as

$$\begin{cases} \tilde{\sigma}_1(0, \xi_2, \xi_3, t) = \tilde{\sigma}_1(1, \xi_2, \xi_3, t) = 0, & 0 \le \xi_2 \le 1,\ 0 \le \xi_3 \le 1, \\ \tilde{\sigma}_2(\xi_1, 0, \xi_3, t) = \tilde{\sigma}_2(\xi_1, 1, \xi_3, t) = 0, & 0 \le \xi_1 \le 1,\ 0 \le \xi_3 \le 1, \\ \tilde{\sigma}_3(\xi_1, \xi_2, 0, t) = \tilde{\sigma}_3(\xi_1, \xi_2, 1, t) = 0, & 0 \le \xi_1 \le 1,\ 0 \le \xi_2 \le 1. \end{cases} \tag{5.85}$$

Therefore, in order for the equality conditions in the conditions (5.69)–(5.74) to hold, we just require that the volatilities of $d\xi_1$, $d\xi_2$, and $d\xi_3$ satisfy the condition (5.85), which is easier to be implemented than the equality conditions in the conditions (5.69)–(5.74). Suppose that $\tilde{\sigma}_1$, $\tilde{\sigma}_2$, $\tilde{\sigma}_3$, $\tilde{\rho}_{1,2}$, $\tilde{\rho}_{1,3}$, and $\tilde{\rho}_{2,3}$ are given functions. In this case, in order to be able to use the expressions of b_1, b_2, and b_3 conveniently, we express $\sigma_1 (\sigma_1\xi_2 - \sigma_2\rho_{1,2})$ and $\sigma_2 (\sigma_2\xi_3 - \sigma_3\rho_{2,3})$ in the expressions of b_2 and b_3 in terms of $\tilde{\sigma}_1$, $\tilde{\sigma}_2$, $\tilde{\sigma}_3$, $\tilde{\rho}_{1,2}$, $\tilde{\rho}_{1,3}$, and $\tilde{\rho}_{2,3}$. From the expression (5.81) we have

$$\begin{aligned} \tilde{\sigma}_1\tilde{\sigma}_2\tilde{\rho}_{1,2} (1 - Z_{1,l}) (Z_1 - Z_{2,l}) &= \sigma_1 (\sigma_2\rho_{1,2} - \sigma_1\xi_2), \\ \tilde{\sigma}_1\tilde{\sigma}_3\tilde{\rho}_{1,3} (1 - Z_{1,l}) (Z_2 - Z_{3,l}) &= \sigma_1 (\sigma_3\rho_{1,3} - \sigma_2\rho_{1,2}\xi_3), \\ \tilde{\sigma}_2\tilde{\sigma}_3\tilde{\rho}_{2,3} (Z_1 - Z_{2,l}) (Z_2 - Z_{3,l}) &= \sigma_1\xi_2 (\sigma_2\rho_{1,2}\xi_3 - \sigma_3\rho_{1,3}) \\ &\quad + \sigma_2 (\sigma_3\rho_{2,3} - \sigma_2\xi_3), \end{aligned}$$

and from the second and third relations we further obtain

$$\begin{aligned} \sigma_2 (\sigma_3\rho_{2,3} - \sigma_2\xi_3) = {} & \tilde{\sigma}_1\tilde{\sigma}_3\tilde{\rho}_{1,3}\xi_2 (1 - Z_{1,l}) (Z_2 - Z_{3,l}) \\ & + \tilde{\sigma}_2\tilde{\sigma}_3\tilde{\rho}_{2,3} (Z_1 - Z_{2,l}) (Z_2 - Z_{3,l}). \end{aligned}$$

Therefore, the expressions of b_1, b_2 and b_3 can be rewritten as

$$\begin{cases} b_1 = \dfrac{rZ_1}{1-Z_{1,l}}, \\ b_2 = \dfrac{r\,(Z_2 - Z_1\xi_2)}{Z_1 - Z_{2,l}} - \dfrac{\tilde{\sigma}_1\tilde{\sigma}_2\tilde{\rho}_{1,2}\,(1-Z_{1,l})}{Z_1 - Z_{2,l}}, \\ b_3 = \dfrac{r\,(Z_3 - Z_2\xi_3)}{Z_2 - Z_{3,l}} - \dfrac{\tilde{\sigma}_1\tilde{\sigma}_3\tilde{\rho}_{1,3}\xi_2\,(1-Z_{1,l})}{Z_2 - Z_{3,l}} \\ \qquad -\dfrac{\tilde{\sigma}_2\tilde{\sigma}_3\tilde{\rho}_{2,3}\,(Z_1 - Z_{2,l})}{Z_2 - Z_{3,l}}. \end{cases} \tag{5.86}$$

By this relation, we can easily calculate b_1, b_2, and b_3 when the values of $\tilde{\sigma}_1$, $\tilde{\sigma}_2$, $\tilde{\sigma}_3$, $\tilde{\rho}_{1,2}$, $\tilde{\rho}_{1,3}$, and $\tilde{\rho}_{2,3}$ are given. Because of the condition (5.85) and $r = 0$ for $Z_1 = 1$ [see the expression (5.68)], we can easily show

$$\begin{cases} b_1(0,\xi_2,\xi_3,t) \ge 0, \quad b_1(1,\xi_2,\xi_3,t) = 0, \quad 0 \le \xi_2 \le 1,\ 0 \le \xi_3 \le 1, \\ b_2(\xi_1,0,\xi_3,t) \ge 0, \quad b_2(\xi_1,1,\xi_3,t) = 0, \quad 0 \le \xi_1 \le 1,\ 0 \le \xi_3 \le 1, \\ b_3(\xi_1,\xi_2,0,t) \ge 0, \quad b_3(\xi_1,\xi_2,1,t) = 0, \quad 0 \le \xi_1 \le 1,\ 0 \le \xi_2 \le 1. \end{cases}$$

It can be proved that when these inequalities and the condition (5.85) hold, the problem (5.83) has a unique solution (see [91]). Thus the problem (5.83) can be solved by numerical methods without any difficulty. In this subsection, the PDE in the problem (5.83) is derived through two steps: first it is obtained from the result given in Sect. 2.3.2 and then a new equation is gotten by means of a transformation. Actually this equation can be obtained directly by setting a portfolio and using Itô's lemma just like what we did in Sect. 2.3.4 for two-factor case. Readers are asked to derive the PDE in the problem (5.83) in this way as Problem 25.

Finally, we say a few words about how to use this model to evaluate interest rate derivatives. First, we need to choose Z_1, Z_2, and Z_3 and find $\tilde{\sigma}_1$, $\tilde{\sigma}_2$, $\tilde{\sigma}_3$ satisfying conditions (5.85), and $\tilde{\rho}_{1,2}$, $\tilde{\rho}_{1,3}$, $\tilde{\rho}_{2,3}$. Finding these functions can be done from the data on markets by statistics. After that, the problem (5.83) needs to be solved. Let $t = 0$ denote today, and suppose the derivative security is European style. On the maturity date T, for each point (ξ_1, ξ_2, ξ_3) in $\tilde{\Omega}$, we can have Z_1, Z_2, and Z_3 by the relation (5.84). Then, we determine a zero-coupon bond curve by using the method given in Sect. 5.6.2. When we obtain such a curve, the value of the payoff and r for the point can be determined. This can be done for all points (ξ_1, ξ_2, ξ_3) in the domain $\tilde{\Omega}$ for $t = T$. When we have the final value and all the coefficients of the partial differential equation in the problem (5.83), we can solve the final-value problem (5.83) from $t = T$ to $t = 0$ and get the value of the derivative security today for all the points in $\tilde{\Omega}$.

For American-style derivatives, the situation is similar. The only difference is that the value of derivative must be greater than the constraint. Because the value of the constraint can be obtained by the zero-coupon bond curves at all points in $\tilde{\Omega} \times [0, T]$, the value of an American-style derivative can be determined without any difficulty. However, free boundaries will usually appear in this case.

From what we have described, we see that this model has the following features:

- The state variables are prices of three zero-coupon bonds with different maturities that can be traded on markets, so the coefficients of the first derivatives with respect to the bond prices Z_i in the partial differential equation simply are rZ_i.
- The volatilities of these zero-coupon bonds and their correlation coefficients can be found directly from the real markets by statistics, so the model will have the real major feature of the markets.
- All the zero-coupon bond curves having appeared in the real market can be reproduced quite accurately. This is the basis of a model giving correct results. If taking three random variables is not good enough, four-factor models can be adopted. Generalizing three-factor models to four-factor models is straightforward.
- In other models, the partial differential equation is defined on an infinite domain. For this model, the corresponding partial differential equation is defined on a finite domain. It has been proved that no boundary condition is needed in order for its final-value problem to have a unique solution. Thus, it is not difficult to design correct and efficient numerical methods to price interest rate derivatives.

For the details on how to determine models from the market data and how to solve the final-value problem of the partial differential equation, see Sect. 10.3 and [96]. There, some numerical results are also given.

5.7 Two-Factor Convertible Bonds

Until now, we discussed derivatives depending on either equities or interest rates. This section deals with a derivative dependent on both equity prices and interest rates. This derivative security is a bond that may, at any time chosen by the holder, be converted into n shares of stocks of the company who issues the bond. Such a bond is commonly known as a convertible bond. As a bond, its price depends on the short rate r. It can be exchanged for n shares of stocks, so its value is also a function of the stock price S. Because its typical life span is about 3–10 years, both S and r are considered as random state variables. Therefore, this bond is called a two-factor convertible bond. In this section, we discuss how to price such a bond.

Consider a bond issued by a company and its payoff depends not only on r but also on the price of the stock of the company. In this case the value of this bond depends on both r and S. Let $B(S, r, t)$ be the value of such a bond. As usual, we assume that S is governed by

$$dS = \mu(S,t)Sdt + \sigma(S,t)SdX_1, \quad 0 \le S \tag{5.87}$$

and the interest rate by

$$dr = u(r,t)dt + w(r,t)dX_2, \quad r_l \le r \le r_u, \tag{5.88}$$

where dX_1 and dX_2 are different Wiener processes though they can be correlated. Suppose that

$$\mathrm{E}\left[dX_1 dX_2\right] = \rho dt,$$

where ρ is a constant belonging to $[-1,\ 1]$ and for S and r, ρ usually is a negative number. According to Sect. 2.3, such a derivative satisfies

$$\frac{\partial B}{\partial t} + \mathbf{L_{S,r}} B + kZ = 0, \quad 0 \le S,\ r_l \le r \le r_u,\ 0 \le t \le T, \tag{5.89}$$

where

$$\mathbf{L_{S,r}} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2} + \rho\sigma S w \frac{\partial^2}{\partial S \partial r} + \frac{1}{2} w^2 \frac{\partial^2}{\partial r^2} + (r - D_0) S \frac{\partial}{\partial S} + (u - \lambda w)\frac{\partial}{\partial r} - r.$$

Here, D_0 is the dividend yield a holder of the stock receives per unit time, and kZ is the coupon payment a holder of the bond receives per unit time, Z being the face value of the bond. λ is the market price of risk for the short rate. T is the maturity date of the bond.

We assume that at maturity time T, the holder of the bond can choose to get the face value Z or n shares of stocks. Therefore, the payoff is

$$B(S, r, T) = \max(Z, nS), \quad 0 \le S, \quad r_l \le r \le r_u. \tag{5.90}$$

This is the final condition for this bond. We assume that for the interest rate, the conditions (5.45) and (5.46) hold, i.e.,

$$\begin{cases} u(r_l, t) - w(r_l, t)\dfrac{\partial}{\partial r} w(r_l, t) \ge 0, \\ w(r_l, t) = 0, \end{cases}$$

and

$$\begin{cases} u(r_u, t) - w(r_u, t)\dfrac{\partial}{\partial r} w(r_u, t) \le 0, \\ w(r_u, t) = 0. \end{cases}$$

Because $w^2(r,t) \ge 0$ and $w(r_l, t) = 0$, on $[r_l, r_u]$ we conclude $w(r_l, t)\dfrac{\partial}{\partial r} w(r_l, t)$ $= \dfrac{1}{2}\dfrac{\partial}{\partial r} w^2(r_l, t) \ge 0$. Similarly, $w(r_u, t)\dfrac{\partial}{\partial r} w(r_u, t) \le 0$. Therefore, the conditions above can be rewritten as

$$\begin{cases} u(r_l,t) \geq w(r_l,t)\dfrac{\partial}{\partial r}w(r_l,t) \geq 0, \\ w(r_l,t) = 0, \end{cases}$$

$$\begin{cases} u(r_u,t) \leq w(r_u,t)\dfrac{\partial}{\partial r}w(r_u,t) \leq 0, \\ w(r_u,t) = 0. \end{cases}$$

Because of $w(r_l,t) = 0$, Eq. (5.89) at $r = r_l$ degenerates into

$$\frac{\partial B}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 B}{\partial S^2} + (r - D_0)S\frac{\partial B}{\partial S} + u\frac{\partial B}{\partial r} - rB + kZ = 0.$$

This equation has hyperbolic properties in the r-direction. Thus, if $u(r_l,t) \geq 0$, then the value $B(S, r_l, t)$ is determined by the value $B(S,r,t)$ in the domain $[0,\infty) \times [r_l, r_u] \times [t,T]$ and no extra boundary condition at $r = r_l$ is needed. Similarly, no boundary condition should be required at $r = r_u$ because $u(r_u,t) \leq 0$ and $w(r_u,t) = 0$. At $S = 0$, Eq. (5.89) becomes

$$\frac{\partial B}{\partial t} + \frac{1}{2}w^2\frac{\partial^2 B}{\partial r^2} + (u - \lambda w)\frac{\partial B}{\partial r} - rB + kZ = 0.$$

This is the bond equation, and the value $B(0,r,t)$ is determined by this equation and the final condition at $S = 0$. Just like the Black–Scholes equation, there is no need for specifying a condition as $S \to \infty$. Therefore, if the conditions (5.45) and (5.46) hold, then we could expect that the problem

$$\begin{cases} \dfrac{\partial B}{\partial t} + \mathbf{L_{s,r}}B + kZ = 0, & 0 \leq S, \quad r_l \leq r \leq r_u, \quad 0 \leq t \leq T, \\ B(S,r,T) = \max(Z, nS), & 0 \leq S, \quad r_l \leq r \leq r_u \end{cases} \tag{5.91}$$

has a unique solution. If $\dfrac{\partial}{\partial r}w(r_l,t)$ and $\dfrac{\partial}{\partial r}w(r_u,t)$ are bounded, which usually is true, then the uniqueness of solution of the problem (5.91) can be obtained from the results given in the paper by Zhu and Li (see [94]).

If this bond can be exchanged for n shares of stocks at any time, then this bond is called a convertible bond and let us denote its value by $B_c(S,r,t)$. It is clear that the value $B_c(S,r,t)$ must satisfy the following constraint

$$B_c(S,r,t) \geq nS, \qquad 0 \leq S, \qquad 0 \leq t \leq T. \tag{5.92}$$

This condition is called the constraint on convertible bonds. Sometimes, the solution of the problem (5.91) satisfies the constraint (5.92), so the problem (5.91) determines the solution of a convertible bond. For example, if $D_0 = 0$, then the problem (5.91) gives the price of a convertible bond, which will be explained later. If

$$D_0 > 0,$$

then the price of a convertible bond should be the solution of the following linear complementarity problem on the domain $[0,\infty) \times [r_l, r_u] \times [0,T]$:

$$\begin{cases} \min\left(-\dfrac{\partial B_c}{\partial t} - \mathbf{L_{S,r}} B_c - kZ,\ B_c(S,r,t) - nS\right) = 0, \\ B_c(S,r,T) = \max(Z, nS) \geq nS. \end{cases}$$

Let us reformulate this problem as a free-boundary problem if $D_0 > 0$. We cannot directly apply Theorem 3.1 in Sect. 3.1 to this case because there are two major differences between the problem in the theorem and the problem here. Here, the operator $\mathbf{L_{S,r}}$ is two-dimensional and there is a nonhomogeneous term kZ. However, the main idea is still true. For $S < Z/n$, $B_c(S,r,T) = Z > nS$. Therefore, on $[0, Z/n)$, $B_c(S,r,T-\Delta t)$ must be greater than nS if Δt is small enough, and no free boundary can appear in that region at time T. Now let us check the region $(Z/n, \infty)$. In this case, we need to check where

$$\left(\frac{\partial}{\partial t} + \mathbf{L_{S,r}}\right) nS + kZ \geq 0$$

and where

$$\left(\frac{\partial}{\partial t} + \mathbf{L_{S,r}}\right) nS + kZ < 0.$$

Because

$$\begin{aligned} \left(\frac{\partial}{\partial t} + \mathbf{L_{S,r}}\right) nS + kZ &= (r - D_0)nS - rnS + kZ \\ &= kZ - D_0 nS, \end{aligned}$$

when $S > Z/n$ and $S > kZ/D_0 n$, namely, $S > \max(Z/n, kZ/D_0 n)$,

$$\left(\frac{\partial}{\partial t} + \mathbf{L_{S,r}}\right) nS + kZ < 0$$

and the solution is nS. Otherwise, we can use the partial differential equation to determine the solution. Therefore, there is a free boundary starting at $S = \max(Z/n, kZ/D_0 n)$ and $t = T$. Let $S_f(r,t)$ be the location of the free boundary, then

$$S_f(r,T) = \max\left(\frac{Z}{n}, \frac{kZ}{D_0 n}\right), \qquad r_l \leq r \leq r_u. \tag{5.93}$$

We assume that there is only one free boundary. From numerical solutions, we know that it is true at least for some cases. Thus, when $D_0 > 0$, the domain $[0, \infty) \times [r_l, r_u] \times [0, T]$ in (S, r, t)-space is divided into subdomains

$$I : [0, S_f(r,t)] \times [r_l, r_u] \times [0, T]$$

and

$$II : (S_f(r,t), \infty) \times [r_l, r_u] \times [0, T] \,.$$

The free boundary is between them. At the free boundary, the solution and its derivatives are continuous. In the subdomain II where $B_c = nS$,

$$\frac{\partial B_c}{\partial S} = n$$

and

$$\frac{\partial B_c}{\partial r} = 0.$$

Thus, it seems that in the subdomain I where the partial differential equation is used, we need to require

$$B_c\left(S_f(r,t), r, t\right) = nS_f(r,t), \qquad r_l \le r \le r_u, \qquad 0 \le t \le T, \tag{5.94}$$

$$\frac{\partial B_c}{\partial S}\left(S_f(r,t), r, t\right) = n, \qquad r_l \le r \le r_u, \qquad 0 \le t \le T, \tag{5.95}$$

and

$$\frac{\partial B_c}{\partial r}\left(S_f(r,t), r, t\right) = 0, \qquad r_l \le r \le r_u, \qquad 0 \le t \le T$$

on the free boundary. Differentiating both sides of the condition (5.94) with respect to r in subdomain I yields

$$\frac{\partial B_c}{\partial S}\left(S_f(r,t), r, t\right) \frac{\partial S_f}{\partial r}(r,t) + \frac{\partial B_c}{\partial r}\left(S_f(r,t), r, t\right) = n \frac{\partial S_f}{\partial r}(r,t).$$

Using the condition (5.95), we arrive at

$$\frac{\partial B_c}{\partial r}\left(S_f(r,t), r, t\right) = 0.$$

Consequently, the conditions (5.94) and (5.95) guarantee that all the first derivatives are continuous at the free boundary and we only need to impose the conditions (5.94) and (5.95) on the solution in subdomain I.

Thus in subdomain I, the solution $B_c(S, r, t)$ and the location of the free boundary $S = S_f(r,t)$ are obtained by solving the following problem:

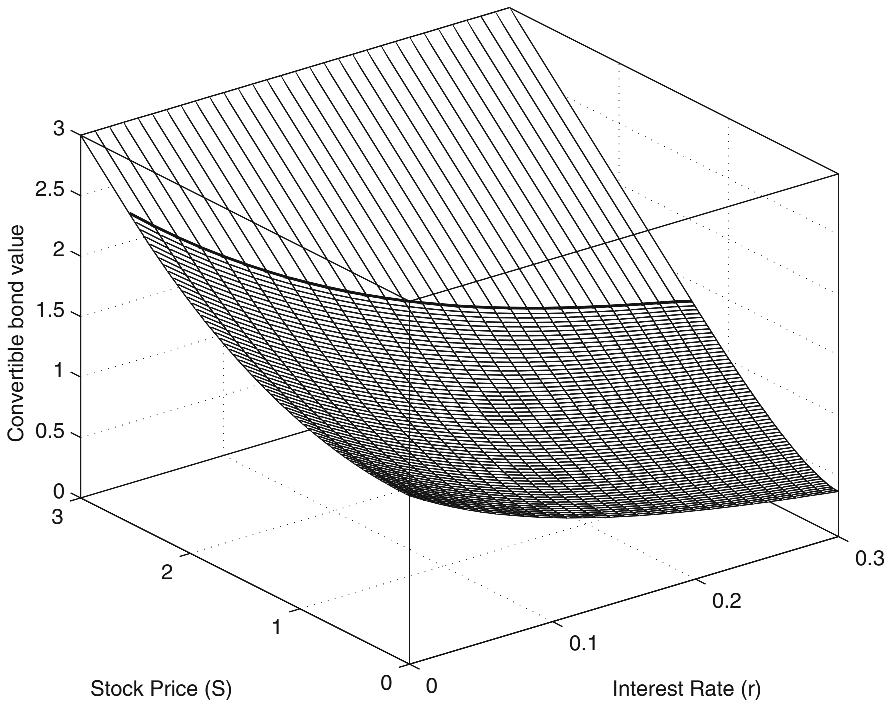

Fig. 5.4. The price of a two-factor convertible bond with $T = 30$ years

$$\begin{cases} \dfrac{\partial B_c}{\partial t} + \mathbf{L}_{\mathbf{s},\mathbf{r}} B_c + kZ = 0, & 0 \le S \le S_f(r,t), \quad r_l \le r \le r_u, \\ & 0 \le t \le T, \\ B_c(S,r,T) = \max(Z, nS), & 0 \le S \le S_f(r,T), \quad r_l \le r \le r_u, \\ B_c(S_f(r,t),r,t) = nS_f(r,t), & r_l \le r \le r_u, \quad 0 \le t \le T, \\ \dfrac{\partial B_c}{\partial S}(S_f(r,t),r,t) = n, & r_l \le r \le r_u, \quad 0 \le t \le T, \\ S_f(r,T) = \max\left(\dfrac{Z}{n}, \dfrac{kZ}{D_0 n}\right), & r_l \le r \le r_u; \end{cases} \tag{5.96}$$

whereas in subdomain II, $B_c(S,r,t) = nS$. The problem (5.96) is usually called a free-boundary problem for convertible bonds.

The starting location of the free boundary

$$S_f(r,T) = \max\left(\frac{Z}{n}, \frac{kZ}{D_0 n}\right)$$

goes to infinity as D_0 tends to zero, so there is no free boundary at time T. Because

$$\left(\frac{\partial}{\partial t}+\mathbf{L_{S,r}}\right)nS+kZ\geq 0$$

always holds when $D_0 = 0$, no free boundary can appear at any time. This means that there is no free boundary when $D_0 = 0$. Thus, the value of a convertible bond in this case is determined by the problem (5.91).

In Fig. 5.4, the price of a two-factor convertible bond with $D_0 = 0.05$ is shown. For this case, there is only one free boundary, which confirms our assumption. The result there is obtained by the singularity-separating finite-difference method, which will be described in Chap. 9.

A convertible bond can also have a call feature that gives the company the right to purchase back the bond at any time (or during specified periods) for a fixed amount M_1. In this case, the price of the bond must not exceed M_1 because no one will spend an amount more than M_1 to buy a bond that can be purchased back for an amount M_1 at any time. When we evaluate the price of such a bond, the constraint

$$B_c(S,r,t)\leq M_1,\ 0\leq S,\quad r_l\leq r\leq r_u,\quad 0\leq t\leq T \tag{5.97}$$

is required. Because of this condition, the price of a convertible bond with a call feature can be less than a convertible bond without this feature. Because the company gets more rights, the buyer of the bond should be asked to pay less money.

A convertible bond can also incorporate a put feature, which means that the owner of the convertible bond can return the bond to the company for an amount M_2 at any time. Now we must impose the constraint

$$B_c(S,r,t)\geq M_2,\quad 0\leq S,\quad r_l\leq r\leq r_u,\quad 0\leq t\leq T. \tag{5.98}$$

This condition might increase the value of the bond. The owner of the bond has more rights, so he usually needs to pay more money in order to purchase such a bond.

Just like the constraint (5.92), the constraint (5.97) or the constraint (5.98) may induce a free boundary or make the free boundary more complicated. For example, for a convertible bond with a call feature, the location of the free boundary at $t = T$ is

$$S_f(r,T)=\min\left(\frac{M_1}{n},\ \max\left(\frac{Z}{n},\frac{kZ}{D_0 n}\right)\right),\quad r_l\leq r\leq r_u. \tag{5.99}$$

If we assume that r is a given function of t, then the bond is a one-factor convertible bond. When $D_0 > 0$, the free-boundary problem is

$$\begin{cases} \dfrac{\partial B_c}{\partial t} + \dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 B_c}{\partial S^2} + [r(t) - D_0] S \dfrac{\partial B_c}{\partial S} - r(t) B_c + kZ = 0, \\ \qquad\qquad\qquad\qquad 0 \le S \le S_f(t), \quad 0 \le t \le T, \\ B_c(S,T) = \max(Z, nS), \qquad 0 \le S \le S_f(T), \\ B_c(S_f(t),t) = nS_f(t), \qquad 0 \le t \le T, \\ \dfrac{\partial B_c}{\partial S}(S_f(t),t) = n, \qquad 0 \le t \le T, \\ S_f(T) = \max\left(\dfrac{Z}{n}, \dfrac{kZ}{D_0 n}\right). \end{cases} \tag{5.100}$$

This problem is only a little different from an American call option problem. Similar to an American option, it can be proved rigorously that there is no free boundary if $D_0 = 0$ for the case $r = constant$. This is left as Problem 29 for readers.

Problems

Table 5.1. Problems and sections

Problems	Sections	Problems	Sections	Problems	Sections
1–3	5.2	4–9	5.3	10	5.4
11–17	5.5	18–25	5.6	26–30	5.7

1. (a) *Suppose the short rate is a known function $r(t)$. Consider a bond with a face value Z and assume that it pays a coupon with a coupon rate $k(t)$, that is, during a time interval $(t, t+dt]$, the coupon payment is $Zk(t)dt$. Show that the value of the bond is

$$V(t) = Z\mathrm{e}^{-\int_t^T r(\tau)d\tau}\left[1 + \int_t^T k(\bar{\tau})\mathrm{e}^{\int_{\bar{\tau}}^T r(\tau)d\tau} d\bar{\tau}\right].$$

   (b) Suppose that $r(t)$ and $k(t)$ are equal to constants r and k, respectively. Show that in this case,

$$V(t) = Z\mathrm{e}^{-r(T-t)}[1 + k(\mathrm{e}^{r(T-t)} - 1)/r].$$

   (c) Suppose that the bond pays coupon payments at two specified dates T_1 and T_2 before the maturity date T and the payments are Zk_1 and Zk_2, respectively. According to the formula given in part (a), and assuming

$T_1 < T_2$, find the values of the bond for $t \in [0, T_1), t \in (T_1, T_2)$, and $t \in (T_2, T)$, respectively, and give a financial interpretation of these expressions.

2. Suppose that the short rate r satisfies

$$dr = udt + w(t)dX,$$

where dX is a Wiener process. Assume that during the time period $[0, t^*]$, for example, t^* being 1 or 3 months, the interest rate is equal to the short rate r. Thus the price of a zero-coupon bond at $t = 0$ with face value one and maturity date t^* is e^{-rt^*}. Because the zero-coupon bond can be traded on the market, we can take $\Pi = V(r,t) - \Delta \mathrm{e}^{-rt^*}$ as the portfolio in order to derive the PDE for $V(r,t)$, the price of an interest rate derivative. Derive the PDE for $V(r,t)$ in this way.

3. Suppose that the short rate r satisfies

$$dr = udt + w(t)dX,$$

where dX is a Wiener process.

(a) Find the stochastic equation for $B(r) = \mathrm{e}^{-rt^*}$ by using Itô's lemma, where t^* is equal to, for example, 1 or 3 months.

(b) As we know, $B(r)$ is the price of a zero-coupon bond at $t = 0$ with face value one and maturity date t^* if during the time period $[0, t^*]$ the interest rate is a constant. $\overline{V}(B,t)$ is any derivative on the zero-coupon bond. Derive the PDE for $\overline{V}(B,t)$ by using Itô's lemma directly.

(c) As we know, if $dr = udt + w(t)dX$, then the price of any derivative security on r, $V(r,t)$, should satisfy the equation

$$\frac{\partial V}{\partial t} + \frac{1}{2} w^2(t) \frac{\partial^2 V}{\partial r^2} + [u - \lambda w(t)] \frac{\partial V}{\partial r} - rV + kZ = 0,$$

where kZ is the coupon of the derivative. Define $V(r,t) = \overline{V}(B(r), t)$. Find the PDE for $V(r,t)$ from PDE obtained in part (b) by using transformation. This equation should be the same as the equation given here. Based on this fact, determine λ.

4. Suppose that $a(r,t) = \sum_{i=0}^{\infty} a_i(t) r^i$ and $b(r,t) = \sum_{i=0}^{\infty} b_i(t) r^i$ and require that the problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + a(r,t) \dfrac{\partial^2 V}{\partial r^2} + b(r,t) \dfrac{\partial V}{\partial r} - rV = 0, & 0 \le t \le T, \\ V(r,T) = 1 \end{cases}$$

has a solution in the form

$$V(r,t) = \mathrm{e}^{A(t) - rB(t)}.$$

Show that in order to fulfill this requirement, between a_i and $b_i, i = 2, 3, \cdots$, there must exist the following relations:

$$a_i B - b_i = 0, \quad i = 2, 3, \cdots .$$

This means that in order to choose $a(r,t)$ and $b(r,t)$ independently and for the solution to be in the form $\mathrm{e}^{A(t)-rB(t)}$, we have to assume $a(r,t) = a_0(t) + a_1(t)r$ and $b(r,t) = b_0(t) + b_1(t)r$.

5. Suppose that $a(r,t) = a_0(t) + a_1(t)r$ and $b(r,t) = b_0(t) + b_1(t)r$. Show that the problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + a(r,t)\dfrac{\partial^2 V}{\partial r^2} + b(r,t)\dfrac{\partial V}{\partial r} - rV = 0, & 0 \le t \le T, \\ V(r,T) = 1 \end{cases}$$

has a solution in the form

$$V(r,t) = \mathrm{e}^{A(t)-rB(t)}$$

with $A(T) = B(T) = 0$ and determine the system of ordinary differential equations the functions $A(t)$ and $B(t)$ should satisfy.

6. *In the Vasicek model, the short rate is assumed to satisfy

$$dr = (\bar{\mu} - \gamma r)dt + \sqrt{-\beta}dX, \qquad \beta < 0, \qquad \gamma > 0,$$

where $\bar{\mu}$, γ, and β are constants, and dX is a Wiener process. Let the market price of risk $\lambda(r,t) = \bar{\lambda}\sqrt{-\beta}$. Then, the price $V(r,t;T)$ of a zero-coupon bond maturing at time T with a face value Z is the solution of the problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2}(-\beta)\dfrac{\partial^2 V}{\partial r^2} + (\mu - \gamma r)\dfrac{\partial V}{\partial r} - rV = 0, \\ \qquad\qquad\qquad\qquad -\infty < r < \infty, \quad 0 \le t \le T, \\ V(r,T;T) = Z, \qquad -\infty < r < \infty, \quad 0 \le t \le T, \end{cases}$$

where

$$\mu = \bar{\mu} + \bar{\lambda}\beta.$$

(a) Show that this problem has a solution in the form

$$V(r,t;T) = Z\mathrm{e}^{A(t,T)-rB(t,T)}$$

and A and B are the solution of the system of ordinary differential equations

$$\begin{cases} \dfrac{dA}{dt} = \dfrac{1}{2}\beta B^2 + \mu B, \\ \dfrac{dB}{dt} = \gamma B - 1 \end{cases}$$

with the conditions

$$A(T, T) = 0,$$
$$B(T, T) = 0.$$

(b) Find the solution of the above problem of ordinary differential equations by solving the two ODEs and show that the expressions of A and B can be rewritten as

$$\begin{cases} A = -\left(\dfrac{\beta}{2\gamma^2} + \dfrac{\mu}{\gamma}\right)(T-t) + \left(\dfrac{\beta}{2\gamma^2} + \dfrac{\mu}{\gamma}\right)B + \dfrac{\beta}{4\gamma}B^2, \\ B = \dfrac{1}{\gamma}\left(1 - \mathrm{e}^{-\gamma(T-t)}\right) \end{cases}$$

if the solution obtained is not in this form.

7. Show

$$\begin{aligned} &\lim_{\alpha\to 0}\left\{\frac{\beta}{\alpha}B + \left[\frac{\beta(\gamma-\psi)}{\alpha(\gamma+\psi)\psi} + \mu\frac{\gamma-\psi}{\alpha\psi}\right]\ln\frac{(\gamma-\psi)/\alpha}{B+(\gamma-\psi)/\alpha}\right. \\ &\left.-\left[\frac{\beta(\gamma+\psi)}{\alpha(\gamma-\psi)\psi} + \mu\frac{\gamma+\psi}{\alpha\psi}\right]\ln\frac{(\gamma+\psi)/\alpha}{B+(\gamma+\psi)/\alpha}\right\} \\ =& -\left(\frac{\beta}{2\gamma^2} + \frac{\mu}{\gamma}\right)(T-t) + \left(\frac{\beta}{2\gamma^2} + \frac{\mu}{\gamma}\right)B + \frac{\beta B^2}{4\gamma}, \end{aligned}$$

where

$$B(t, T) = \frac{1}{\gamma}\left(1 - \mathrm{e}^{-\gamma(T-t)}\right) \quad \text{and} \quad \psi = \sqrt{\gamma^2 + 2\alpha}.$$

(The two sides are two expressions for $A(t, T)$ associated with the Vasicek model obtained by different approaches. This confirms that the two different approaches give the same answer.)

8. *In the Cox–Ingersoll–Ross model, the short rate is assumed to satisfy

$$dr = (\mu - \bar{\gamma}r)dt + \sqrt{\alpha r}dX,$$

where μ, $\bar{\gamma}$, and α are constants, and dX is a Wiener process. Let the market price of risk $\lambda(r, t)$ be $\bar{\lambda}\sqrt{\alpha r}$. Then, the price $V(r, t; T)$ of a zero-coupon bond maturing at time T with a face value Z is the solution of the problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2}\alpha r\dfrac{\partial^2 V}{\partial r^2} + (\mu - \gamma r)\dfrac{\partial V}{\partial r} - rV = 0, & 0 \le r, \quad 0 \le t \le T, \\ V(r, T; T) = Z, & 0 \le r, \end{cases}$$

where $\gamma = \bar{\gamma} + \bar{\lambda}\alpha$.

(a) Show that this problem has a solution in the form

$$V(r,t;T) = Z\mathrm{e}^{A(t,T)-rB(t,T)}$$

and A and B are the solutions of the system of ordinary differential equations

$$\begin{cases} \dfrac{dA}{dt} = \mu B, \\ \dfrac{dB}{dt} = \dfrac{1}{2}\alpha B^2 + \gamma B - 1 \end{cases}$$

with the conditions

$$A(T,T) = 0$$

and

$$B(T,T) = 0.$$

(b) Find the solution of the above problem of ordinary differential equations by solving the two ODEs and show that the expressions of A and B can be rewritten as

$$\begin{cases} A = \ln\left(\dfrac{2\psi\mathrm{e}^{(\gamma+\psi)(T-t)/2}}{(\gamma+\psi)\mathrm{e}^{\psi(T-t)} - (\gamma-\psi)}\right)^{2\mu/\alpha}, \\ B = \dfrac{2(\mathrm{e}^{\psi(T-t)}-1)}{(\gamma+\psi)\mathrm{e}^{\psi(T-t)} - (\gamma-\psi)} \\ \quad \text{with } \psi = \sqrt{\gamma^2 + 2\alpha} \end{cases}$$

if the solution obtained is not in this form.

9. Show

$$\begin{aligned} &Z\left[\frac{B+(\gamma-\psi)/\alpha}{(\gamma-\psi)/\alpha}\right]^{\mu(\psi-\gamma)/\alpha\psi} \left[\frac{B+(\gamma+\psi)/\alpha}{(\gamma+\psi)/\alpha}\right]^{\mu(\gamma+\psi)/\alpha\psi} \mathrm{e}^{-rB} \\ \equiv\ &Z\left[\frac{2\psi\mathrm{e}^{(\gamma+\psi)(T-t)/2}}{(\gamma+\psi)\mathrm{e}^{\psi(T-t)} - (\gamma-\psi)}\right]^{2\mu/\alpha} \mathrm{e}^{-rB}, \end{aligned}$$

where

$$B = \frac{2\left(\mathrm{e}^{\psi(T-t)}-1\right)}{(\gamma+\psi)\,\mathrm{e}^{\psi(T-t)} - (\gamma-\psi)}.$$

(The two sides are two expressions for the zero-coupon bond price associated with the Cox–Ingersoll–Ross model obtained by different approaches. This confirms that the two different approaches give the same answer.)

10. *Describe a way to determine the market price of risk for the short rate.

11. *Suppose that any European-style interest rate derivative with a continuous coupon satisfies the equation:

$$\frac{\partial V}{\partial t} + \frac{1}{2}w^2\frac{\partial^2 V}{\partial r^2} + (u-\lambda w)\frac{\partial V}{\partial r} - rV + k = 0, \quad r_l \le r \le r_u, \quad t \le T,$$

where k is the coupon rate corresponding to the derivative, the coefficients u and w satisfy the reversion conditions on the boundaries $r = r_l$, $r = r_u$, and λ is a given bounded function. Describe how to evaluate the price of a European call option on a bond with coupon by using this equation.

12. (a) Let $Z(t;T^*)$ be the price of a zero-coupon bond with a face value of one dollar and with maturity date T^* at time t and let $\bar{f}(t,T,T+\frac{1}{2})$ be the forward interest rate compounded semiannually at time t for the period $(T,T+\frac{1}{2})$. Show

$$\bar{f}\left(t,T,T+\frac{1}{2}\right) = 2\left[\frac{Z(t;T)}{Z(t;T+1/2)} - 1\right].$$

(b) There is a cash flow $\frac{1}{2}\bar{f}(t_{k-1},t_{k-1},t_k)$, t_k being $t+k/2$, $k = 1, 2, \cdots, 2N$ and t_0 being t. Find the value of the cash flow at time t.

(c) *Show that the value of an N-year swap with swap rate r_s and with notional principal Q is

$$V_s(T;r_s) = Q\left[1 - Z(T;T+N) - \frac{r_s}{2}\sum_{k=1}^{2N} Z(T;T+k/2)\right],$$

where T is the time the swap initiates.

13. Show that the solution of the problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2}w^2\dfrac{\partial^2 V}{\partial r^2} + (u-\lambda w)\dfrac{\partial V}{\partial r} - rV = 0, \\ \qquad\qquad r_l \le r \le r_u, \quad t \le T, \\ V(r,T;T) = 1, \qquad r_l \le r \le r_u \end{cases}$$

is the same as that of the problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2}w^2\dfrac{\partial^2 V}{\partial r^2} + (u-\lambda w)\dfrac{\partial V}{\partial r} - rV + \delta(t-T) = 0, \\ \qquad\qquad r_l \le r \le r_u, \quad t \le T, \\ V(r,T^+;T) = 0, \qquad r_l \le r \le r_u \end{cases}$$

for any $t < T$.

14. Let $V_{s1k}(r,T)$ denote the price of a $(k/2)$-year zero-coupon bond, $k = 1, 2, \cdots, 2N$, and we want to get $\sum_{k=1}^{2N} V_{s1k}(r,T)$. Consider the following procedures. The first one is to solve the following problems

$$\begin{cases} \dfrac{\partial V_{s1k}}{\partial t} + \dfrac{1}{2}w^2\dfrac{\partial^2 V_{s1k}}{\partial r^2} + (u-\lambda w)\dfrac{\partial V_{s1k}}{\partial r} - rV_{s1k} = 0, \quad r_l \le r \le r_u, \\ \qquad\qquad T \le t \le T+k/2, \\ V_{s1k}(r,T+k/2) = 1, \qquad r_l \le r \le r_u, \end{cases}$$

$k = 1, 2, \cdots, 2N$, and then obtain $\sum_{k=1}^{2N} V_{s1k}(r,T)$ by adding $V_{s1k}(r,T)$, $k = 1, 2, \cdots, 2N$, together. The second one is to solve the problem:

$$\begin{cases} \dfrac{\partial V_{s1}}{\partial t} + \dfrac{1}{2}w^2 \dfrac{\partial^2 V_{s1}}{\partial r^2} + (u - \lambda w)\dfrac{\partial V_{s1}}{\partial r} - rV_{s1} \\ \qquad + \sum_{k=1}^{2N} \delta(t - T - k/2) = 0, \\ \qquad\qquad\qquad r_l \le r \le r_u, \quad T \le t \le T + N, \\ V_{s1}(r, T+N) = 0, \qquad r_l \le r \le r_u. \end{cases}$$

(a) Show $V_{s1}(r,T) = \sum_{k=1}^{2N} V_{s1k}(r,T)$ holds.

(b) In order to get $\sum_{k=1}^{2N} V_{s1k}(r,T)$, which procedure is better and why?

15. Suppose that the solution of

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2}w^2 \dfrac{\partial^2 V}{\partial r^2} + (u - \lambda w)\dfrac{\partial V}{\partial r} - rV = 0, & r_l \le r \le r_u, \quad t^* \le t \le t_k, \\ V(r, t_k; t_k) = 1, & r_l \le r \le r_u \end{cases}$$

is $V(r,t;t_k)$ and that $V(r^*, t^*; t_k) = Z(t^*; t_k)$. Also assume that $V_s(r,t;r_s)$ is the solution of

$$\begin{cases} \dfrac{\partial V_s}{\partial t} + \dfrac{1}{2}w^2 \dfrac{\partial^2 V_s}{\partial r^2} + (u - \lambda w)\dfrac{\partial V_s}{\partial r} - rV_s - \sum_{k=k^*+1}^{2N} \dfrac{Qr_s}{2}\delta(t - t_k) \\ \qquad + Q\left[1 + \dfrac{\bar{f}(t_{k^*}, t_{k^*}, t_{k^*+1})}{2}\right]\delta(t - t_{k^*+1}) = 0, \\ \qquad\qquad\qquad r_l \le r \le r_u, \quad t^* \le t \le T + N, \\ V_s(r, T+N; r_s) = -Q, \qquad r_l \le r \le r_u. \end{cases}$$

Here, $V_s(r,t;r_s)$ actually is the value of a swap. Q and r_s are the notional principal and the swap rate, respectively. t^*, T, and N denote the time today, the time the swap is initiated, and the duration of the swap with the relation $T \le t^* < T + N$. k^* is the integer part of $(t^* - T)/2$, and $t_k = T + k/2$, $k = k^* + 1, k^* + 2, \cdots, 2N$. $\bar{f}(t_{k^*}, t_{k^*}, t_{k^*+1})$ is the 6-month LIBOR for the period $[t_{k^*}, t_{k^*+1}]$ determined at time t_{k^*}. Show

$$V_s(r^*, t^*; r_s) = QZ(t^*; t_{k^*+1})\left[1 + \frac{1}{2}\bar{f}(t_{k^*}, t_{k^*}, t_{k^*+1})\right] - Q\left[\sum_{k=k^*+1}^{2N} \frac{r_s}{2} Z(t^*; t_k) + Z(t^*; T+N)\right].$$

16. *Suppose that any European-style interest rate derivative satisfies the equation:

$$\frac{\partial V}{\partial t}+\frac{1}{2}w^2\frac{\partial^2 V}{\partial r^2}+(u-\lambda w)\frac{\partial V}{\partial r}-rV+f(t)=0,\quad r_l\le r\le r_u,$$

where all the coefficients in the equation are known. Let $V_{so}(r,t)$ be the value of a T-year swaption on a N-year swap. Its payoff is

$$Q\max\left(1-Z(T;T+N)-\frac{r_{se}}{2}\sum_{k=1}^{2N}Z(T;T+k/2),\,0\right),$$

where Q is the notional principal, r_{se} is the exercise swap rate, and $Z(T;T+k/2)$ is the value of zero-coupon bond with maturity $k/2$ at time T. Describe how to find the price of the swaption, including to find the payoff of the swaption, by solving this equation from $T+N$ to T and from T to 0.

17. Consider an N-year floor with a floor rate r_f. Suppose that the money will be paid quarterly at time $t_k=t^*+k/4$, $k=2,3,\cdots,4N$, and the floating rate is the 3-month LIBOR. Suppose that $V_{bk}(r,t)$ is the solution of the problem

$$\begin{cases}\dfrac{\partial V_{bk}}{\partial t}+\dfrac{1}{2}w^2\dfrac{\partial^2 V_{bk}}{\partial r^2}+(u-\lambda w)\dfrac{\partial V_{bk}}{\partial r}-rV_{bk}=0,\\ \qquad\qquad r_l\le r\le r_u,\quad t_{k-1}\le t\le t_k,\\ V_{bk}(r,t_k)=Q\left(1+\dfrac{r_f}{4}\right),\qquad r_l\le r\le r_u,\end{cases}$$

where $k=2,3,\cdots,4N$ and $V_f(r,t)$ is the solution of the problem

$$\begin{cases}\dfrac{\partial V_f}{\partial t}+\dfrac{1}{2}w^2\dfrac{\partial^2 V_f}{\partial r^2}+(u-\lambda w)\dfrac{\partial V_f}{\partial r}-rV_f\\ \qquad +\sum\limits_{k=2}^{4N}\max(V_{bk}(r,t_{k-1})-Q,0)\delta(t-t_{k-1})=0,\\ \qquad\qquad r_l\le r\le r_u,\quad t^*\le t\le t_{4N-1},\\ V_f(r,t_{4N-1})=0,\qquad r_l\le r\le r_u.\end{cases}$$

Show that the premium of the floor should be

$$V_f(r^*,t^*),$$

where r^* is the short rate at time t^*.

18. (a) $\mathbf{S}$ is a random vector and its covariance matrix is $\mathbf{B}$. Let $\bar{\mathbf{S}}=\mathbf{AS}$, $\mathbf{A}$ being a constant matrix, and its covariance matrix be $\mathbf{C}$. Find the relation among $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$.

(b) How do we choose **A** so that **C** will be a diagonal matrix?

(c) *Suppose that $\bar{S}_1, \bar{S}_2, \cdots, \bar{S}_K$ are variables and $\bar{S}_{K+1}, \bar{S}_{K+2}, \cdots, \bar{S}_N$ are fixed numbers. Find the dependence of $S_{K+1}, S_{K+2}, \cdots, S_N$ on S_1, S_2, $\cdots$, S_K.

19. (a) Suppose that there is a domain Ω on the (Z_1, Z_2)-plane, the boundary of Ω is Γ, and $(n_1, n_2)^T$ is the outer normal vector of the boundary Γ. Assume that Z_1 and Z_2 are two stochastic processes and satisfy the system of stochastic differential equations:

$$dZ_i = \mu_i(Z_1, Z_2, t)dt + \sigma_i(Z_1, Z_2, t)dX_i \quad \text{with} \quad \sigma_i \geq 0, \quad i = 1, 2,$$

where $dX_i, i = 1, 2$, are the Wiener processes and $\mathrm{E}[dX_1 dX_2] = \rho_{12}dt$ with $\rho_{12} \in [-1, 1]$. Suppose that at $t = 0$, $(Z_1, Z_2) \in \Omega$. Show that in order to guarantee $(Z_1, Z_2) \in \Omega$ for any time $t \in [0, T]$, we need to require, for any $t \in [0, T]$ and for any point on Γ, the following condition to be held:

(i) if $n_1 \neq 0$ and $n_2 = 0$, then

$$\begin{cases} n_1\mu_1 \leq 0, \\ \sigma_1 = 0; \end{cases}$$

(ii) if $n_1 = 0$ and $n_2 \neq 0$, then

$$\begin{cases} n_2\mu_2 \leq 0, \\ \sigma_2 = 0; \end{cases}$$

(iii) if $n_1 \neq 0$ and $n_2 \neq 0$, then

$$\begin{cases} n_1\mu_1 + n_2\mu_2 \leq 0, \\ n_1\sigma_1 - \operatorname{sign}(n_1 n_2) n_2 \sigma_2 = 0, \quad \text{and} \quad \rho_{12} = -\operatorname{sign}(n_1 n_2), \end{cases}$$

where

$$\operatorname{sign}(n_1 n_2) = \begin{cases} 1, & \text{if } n_1 n_2 > 0, \\ -1, & \text{if } n_1 n_2 < 0. \end{cases}$$

If a point is a corner point, then there are two normals and we need to require this condition to be held for the two outer normal vectors.

(b) Suppose that the domain Ω is $Z_{1l} \leq Z_1 \leq 1$ and $Z_{2l} \leq Z_2 \leq Z_1$, where Z_{1l} and Z_{2l} are constants, and $Z_{1l} \geq Z_{2l}$. Find the concrete condition for each segment of the boundary according to the condition given in part (a).

20. Assume that Z_1, Z_2, Z_3 are random variables and satisfy the system of stochastic differential equations:

$$dZ_i = \mu_i \left(Z_1, Z_2, Z_3, t\right) dt + \sigma_i \left(Z_1, Z_2, Z_3, t\right) dX_i, \quad i = 1, 2, 3,$$

where dX_i are the Wiener processes and $\mathrm{E}\left[dX_i dX_j\right] = \rho_{i,j} dt$ with $\rho_{i,j} \in [-1, 1]$. In order to guarantee that if a point is in a domain Ω at time t^*, then the point is still in the domain Ω at $t = t^* + dt$ for a positive dt, it is necessary to require that the condition

$$n_1 dZ_1 + n_2 dZ_2 + n_3 dZ_3 \le 0$$

holds at any point on the boundary of the domain Ω, where n_1, n_2, and n_3 are the three components of the outer normal vector of the boundary at the point. Suppose that the domain Ω is $\{Z_{1,l} \le Z_1 \le 1, \quad Z_{2,l} \le Z_2 \le Z_1, \quad Z_{3,l} \le Z_3 \le Z_2\}$. Show that on the surfaces $Z_1 = 1$, $Z_2 = Z_{2,l}$, and $Z_3 = Z_2$, the condition is equivalent to $\{\mu_1 \le 0, \sigma_1 = 0\}$, $\{\mu_2 \ge 0, \sigma_2 = 0\}$, and $\{-\mu_2 + \mu_3 \le 0, \sigma_2 = \sigma_3, \rho_{2,3} = 1\}$, respectively.

21. Suppose that $\sigma_1 \left(Z_l, Z_2, Z_3, t\right)$, $\sigma_2 \left(Z_l, Z_2, Z_3, t\right)$, and $\sigma_3 \left(Z_l, Z_2, Z_3, t\right)$ are defined on $\Omega : \{Z_{1,l} \le Z_1 \le 1,\ Z_{2,l} \le Z_2 \le Z_1,\ Z_{3,l} \le Z_3 \le Z_2\}$. Assume that

   (i) $\sigma_1 \left(Z_{1,l}, Z_2, Z_3, t\right) = 0$ on surface I: $\{Z_1 = Z_{1,l},\ Z_{2,l} \le Z_2 \le Z_1,\ Z_{3,l} \le Z_3 \le Z_2\}$.

   (ii) $\sigma_1 \left(1, Z_2, Z_3, t\right) = 0$ on surface II: $\{Z_1 = 1,\ Z_{2,l} \le Z_2 \le Z_1,\ Z_{3,l} \le Z_3 \le Z_2\}$.

   (iii) $\sigma_2 \left(Z_1, Z_{2,l}, Z_3, t\right) = 0$ on surface III: $\{Z_{1,l} \le Z_1 \le 1,\ Z_2 = Z_{2,l},\ Z_{3,l} \le Z_3 \le Z_2\}$.

   (iv) $\sigma_1 \left(Z_1, Z_1, Z_3, t\right) = \sigma_2 \left(Z_1, Z_1, Z_3, t\right), \quad \rho_{1,2} \left(Z_1, Z_1, Z_3, t\right) = 1$ on surface IV: $\{Z_{1,l} \le Z_1 \le 1,\ Z_2 = Z_1,\ Z_{3,l} \le Z_3 \le Z_2\}$.

   (v) $\sigma_3 \left(Z_1, Z_2, Z_{3,l}, t\right) = 0$ on surface V: $\{Z_{1,l} \le Z_1 \le 1,\ Z_{2,l} \le Z_2 \le Z_1,\ Z_3 = Z_{3,l}\}$.

   (vi) $\sigma_2 \left(Z_1, Z_2, Z_2, t\right) = \sigma_3 \left(Z_1, Z_2, Z_2, t\right), \quad \rho_{2,3} \left(Z_1, Z_2, Z_2, t\right) = 1$ on surface VI: n $\{Z_{1,l} \le Z_1 \le 1,\ Z_{2,l} \le Z_2 \le Z_1,\ Z_3 = Z_2\}$.

   Define

$$\begin{cases} \xi_1 = \dfrac{Z_1 - Z_{1,l}}{1 - Z_{1,l}}, \\[2ex] \xi_2 = \dfrac{Z_2 - Z_{2,l}}{Z_1 - Z_{2,l}}, \\[2ex] \xi_3 = \dfrac{Z_3 - Z_{3,l}}{Z_2 - Z_{3,l}}, \end{cases}$$

and

$$\begin{cases} \tilde{\sigma}_1^2(\xi_1,\xi_2,\xi_3,t) = \dfrac{\sigma_1^2}{(1-Z_{1,l})^2}, \\ \tilde{\sigma}_2^2(\xi_1,\xi_2,\xi_3,t) = \dfrac{\sigma_1^2\xi_2^2 - 2\sigma_1\sigma_2\xi_2\rho_{1,2} + \sigma_2^2}{(Z_1 - Z_{2,l})^2}, \\ \tilde{\sigma}_3^2(\xi_1,\xi_2,\xi_3,t) = \dfrac{\sigma_2^2\xi_3^2 - 2\sigma_2\sigma_3\xi_3\rho_{2,3} + \sigma_3^2}{(Z_2 - Z_{3,l})^2}. \end{cases}$$

Show that the assumption on σ_1, σ_2, and σ_3 is equivalent to

$$\begin{cases} \tilde{\sigma}_1(0,\xi_2,\xi_3,t) = \tilde{\sigma}_1(1,\xi_2,\xi_3,t) = 0, & 0 \le \xi_2 \le 1,\ 0 \le \xi_3 \le 1, \\ \tilde{\sigma}_2(\xi_1,0,\xi_3,t) = \tilde{\sigma}_2(\xi_1,1,\xi_3,t) = 0, & 0 \le \xi_1 \le 1,\ 0 \le \xi_3 \le 1, \\ \tilde{\sigma}_3(\xi_1,\xi_2,0,t) = \tilde{\sigma}_3(\xi_1,\xi_2,1,t) = 0, & 0 \le \xi_1 \le 1,\ 0 \le \xi_2 \le 1. \end{cases}$$

22. (a) Show that under the transformation

$$\begin{cases} \xi_1 = \dfrac{Z_1 - Z_{1,l}}{1 - Z_{1,l}}, \\ \xi_2 = \dfrac{Z_2 - Z_{2,l}}{Z_1 - Z_{2,l}}, \end{cases}$$

the partial differential equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\sigma_i\sigma_j\rho_{i,j}\frac{\partial^2 V}{\partial Z_i \partial Z_j} + r\sum_{i=1}^{2} Z_i\frac{\partial V}{\partial Z_i} - rV = 0$$

becomes

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\tilde{\sigma}_i\tilde{\sigma}_j\tilde{\rho}_{i,j}\frac{\partial^2 V}{\partial \xi_i \partial \xi_j} + \sum_{i=1}^{2} b_i\frac{\partial V}{\partial \xi_i} - rV = 0,$$

where

$$\begin{cases} b_1 = \dfrac{rZ_1}{1 - Z_{1,l}}, \\ b_2 = \dfrac{r(Z_2 - Z_1\xi_2)}{Z_1 - Z_{2,l}} + \dfrac{\sigma_1(\sigma_1\xi_2 - \sigma_2\rho_{1,2})}{(Z_1 - Z_{2,l})^2}, \end{cases}$$

and $\tilde{\sigma}_1$, $\tilde{\sigma}_2$, $\tilde{\rho}_{1,2}$, are determined by

$$\begin{cases} \dfrac{1}{2}\tilde{\sigma}_1^2 = \dfrac{\frac{1}{2}\sigma_1^2}{(1-Z_{1,l})^2}, \\ \dfrac{1}{2}\tilde{\sigma}_2^2 = \dfrac{\frac{1}{2}\left(\sigma_1^2\xi_2^2 - 2\sigma_1\sigma_2\xi_2\rho_{1,2} + \sigma_2^2\right)}{(Z_1 - Z_{2,l})^2}, \\ \tilde{\sigma}_1\tilde{\sigma}_2\tilde{\rho}_{1,2} = \dfrac{\sigma_1(\sigma_2\rho_{1,2} - \sigma_1\xi_2)}{(1 - Z_{1,l})(Z_1 - Z_{2,l})}. \end{cases}$$

(b) Show further that the expression of b_2 can be rewritten as

$$b_2 = \frac{r\left(Z_2 - Z_1\xi_2\right)}{Z_1 - Z_{2,l}} - \frac{\tilde{\sigma}_1\tilde{\sigma}_2\tilde{\rho}_{1,2}\left(1 - Z_{1,l}\right)}{Z_1 - Z_{2,l}}.$$

(c) $\tilde{\sigma}_i$ and b_i given above are functions of ξ_1, ξ_2, t and let $\tilde{\sigma}_i(\xi_1, \xi_2, t)$ and $b_i(\xi_1, \xi_2, t)$ denote these functions, $i = 1$ and 2. Show that if

$$\begin{cases} \tilde{\sigma}_1(0, \xi_2, t) = \tilde{\sigma}_1(1, \xi_2, t) = 0, & 0 \le \xi_2 \le 1, \\ \tilde{\sigma}_2(\xi_1, 0, t) = \tilde{\sigma}_2(\xi_1, 1, t) = 0, & 0 \le \xi_1 \le 1, \end{cases}$$

then

$$\begin{cases} b_1(0, \xi_2, t) \ge 0, \quad b_1(1, \xi_2, t) = 0, & 0 \le \xi_2 \le 1, \\ b_2(\xi_1, 0, t) \ge 0, \quad b_2(\xi_1, 1, t) = 0, & 0 \le \xi_1 \le 1. \end{cases}$$

(Hint: $r(\xi_1, \xi_2, t)|_{\xi_1=1} = 0$. This can be explained as follows. $\xi_1 = 1$ means $Z_1 = 1$, thus the zero-coupon bond curve must be flat near $T = 0$ and its derivative with respect to T at $T = 0$, $r(\xi_1, \xi_2, t)|_{\xi_1=1}$ equals 0. When $\tilde{\sigma}_i$, b_i, $i = 1, 2$, satisfy these conditions here, it can be proved that the final value problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2}\sum\limits_{i=1}^{2}\sum\limits_{j=1}^{2}\tilde{\sigma}_i\tilde{\sigma}_j\tilde{\rho}_{i,j}\dfrac{\partial^2 V}{\partial \xi_i \partial \xi_j} + \sum\limits_{i=1}^{2} b_i \dfrac{\partial V}{\partial \xi_i} - rV = 0 \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{on } [0,1] \times [0,1] \times [0,T], \\ V(\xi_1, \xi_2, T) = V_T(\xi_1, \xi_2) \qquad\qquad \text{on } [0,1] \times [0,1] \end{cases}$$

has a unique solution.)

23. (a) *Show that under the transformation

$$\begin{cases} \xi_1 = \dfrac{Z_1 - Z_{1,l}}{1 - Z_{1,l}}, \\ \xi_2 = \dfrac{Z_2 - Z_{2,l}}{Z_1 - Z_{2,l}}, \\ \xi_3 = \dfrac{Z_3 - Z_{3,l}}{Z_2 - Z_{3,l}}, \end{cases}$$

the partial differential equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_i\sigma_j\rho_{i,j}\frac{\partial^2 V}{\partial Z_i \partial Z_j} + r\sum_{i=1}^{3} Z_i \frac{\partial V}{\partial Z_i} - rV = 0$$

becomes

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\tilde{\sigma}_i\tilde{\sigma}_j\tilde{\rho}_{i,j}\frac{\partial^2 V}{\partial \xi_i \partial \xi_j} + \sum_{i=1}^{3} b_i \frac{\partial V}{\partial \xi_i} - rV = 0,$$

and find the expressions of $\tilde{\sigma}_1$, $\tilde{\sigma}_2$, $\tilde{\sigma}_3$, $\tilde{\rho}_{1,2}$, $\tilde{\rho}_{1,3}$, $\tilde{\rho}_{2,3}$, b_1, b_2, and b_3.

(b) Show

$$\begin{cases} \tilde{\sigma}_1(0,\xi_2,\xi_3,t) = \tilde{\sigma}_1(1,\xi_2,\xi_3,t) = 0, & 0 \le \xi_2 \le 1,\ 0 \le \xi_3 \le 1, \\ \tilde{\sigma}_2(\xi_1,0,\xi_3,t) = \tilde{\sigma}_2(\xi_1,1,\xi_3,t) = 0, & 0 \le \xi_1 \le 1,\ 0 \le \xi_3 \le 1, \\ \tilde{\sigma}_3(\xi_1,\xi_2,0,t) = \tilde{\sigma}_3(\xi_1,\xi_2,1,t) = 0, & 0 \le \xi_1 \le 1,\ 0 \le \xi_2 \le 1, \end{cases}$$

and

$$\begin{cases} b_1(0,\xi_2,\xi_3,t) \ge 0, \quad b_1(1,\xi_2,\xi_3,t) = 0, & 0 \le \xi_2 \le 1,\ 0 \le \xi_3 \le 1, \\ b_2(\xi_1,0,\xi_3,t) \ge 0, \quad b_2(\xi_1,1,\xi_3,t) = 0, & 0 \le \xi_1 \le 1,\ 0 \le \xi_3 \le 1, \\ b_3(\xi_1,\xi_2,0,t) \ge 0, \quad b_3(\xi_1,\xi_2,1,t) = 0, & 0 \le \xi_1 \le 1,\ 0 \le \xi_2 \le 1. \end{cases}$$

(When $\tilde{\sigma}_i$, b_i, $i = 1, 2, 3$, satisfy these conditions here, it can be proved that the final value problem

$$\begin{cases} \dfrac{\partial V}{\partial t} + \dfrac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3} \tilde{\sigma}_i\tilde{\sigma}_j\tilde{\rho}_{i,j}\dfrac{\partial^2 V}{\partial\xi_i\partial\xi_j} + \sum_{i=1}^{3} b_i\dfrac{\partial V}{\partial\xi_i} - rV = 0 \\ \qquad\qquad\qquad\qquad \text{on } [0,1]\times[0,1]\times[0,1]\times[0,T], \\ V(\xi_1,\xi_2,\xi_3,T) = V_T(\xi_1,\xi_2,\xi_3) \quad \text{on } [0,1]\times[0,1]\times[0,1] \end{cases}$$

has a unique solution.)

24. Assume that Z_1, Z_2, Z_3 are random variables and satisfy the system of stochastic differential equations:

$$dZ_i = \mu_i\left(Z_1, Z_2, Z_3, t\right)dt + \sigma_i\left(Z_1, Z_2, Z_3, t\right)dX_i, \quad i = 1, 2, 3,$$

where dX_i are the Wiener processes and $\mathrm{E}\left[dX_i dX_j\right] = \rho_{ij}dt$ with $-1 \le \rho_{ij} \le 1$, and that ξ_1, ξ_2 and ξ_3 are governed by

$$d\xi_i = \tilde{\mu}_i(\xi_1,\xi_2,\xi_3,t)dt + \tilde{\sigma}_i(\xi_1,\xi_2,\xi_3,t)d\tilde{X}_i, \quad i = 1, 2, 3,$$

where $d\tilde{X}_i$ are the Wiener processes and $\mathrm{E}\left[d\tilde{X}_i d\tilde{X}_j\right] = \tilde{\rho}_{ij}dt$ with $-1 \le \tilde{\rho}_{ij} \le 1$. Furthermore, we suppose that ξ_1, ξ_2 and ξ_3 are defined by

$$\begin{cases} \xi_1 = \dfrac{Z_1 - Z_{1,l}}{1 - Z_{1,l}}, \\ \xi_2 = \dfrac{Z_2 - Z_{2,l}}{Z_1 - Z_{2,l}}, \\ \xi_3 = \dfrac{Z_3 - Z_{3,l}}{Z_2 - Z_{3,l}}, \end{cases}$$

where $Z_{1,l}$, $Z_{2,l}$, and $Z_{3,l}$ are constants. Find the expressions of $\tilde{\sigma}_1$, $\tilde{\sigma}_2$, $\tilde{\sigma}_3$, $\tilde{\rho}_{12}$, $\tilde{\rho}_{13}$, $\tilde{\rho}_{23}$ as functions of σ_1, σ_2, σ_3, $\rho_{1,2}$, $\rho_{1,3}$, $\rho_{2,3}$, Z_1, Z_2, and Z_3 by using Itô's lemma.

25. Suppose that ξ_1, ξ_2 and ξ_3 satisfy the system of stochastic differential equations:

$$d\xi_i = \tilde{\mu}_i(\xi_1, \xi_2, \xi_3, t)dt + \tilde{\sigma}_i(\xi_1, \xi_2, \xi_3, t)d\tilde{X}_i, \quad i = 1, 2, 3,$$

where $d\tilde{X}_i$ are the Wiener processes and $\mathrm{E}\left[d\tilde{X}_i d\tilde{X}_j\right] = \tilde{\rho}_{ij}dt$ with $-1 \le \tilde{\rho}_{ij} \le 1$. Define

$$\begin{cases} Z_1(\xi_1) &= Z_{1,l} + \xi_1(1 - Z_{1,l}), \\ Z_2(\xi_1, \xi_2) &= Z_{2,l} + \xi_2[Z_1(\xi_1) - Z_{2,l}] \\ &= Z_{2,l} + \xi_2[Z_{1,l} + \xi_1(1 - Z_{1,l}) - Z_{2,l}], \\ Z_3(\xi_1, \xi_2, \xi_3) &= Z_{3,l} + \xi_3\{Z_2(\xi_1, \xi_2) - Z_{3,l}\} \\ &= Z_{3,l} + \xi_3\{Z_{2,l} + \xi_2[Z_{1,l} + \xi_1(1 - Z_{1,l}) - Z_{2,l}] - Z_{3,l}\}. \end{cases}$$

Assume that $Z_1(\xi_1)$, $Z_2(\xi_1, \xi_2)$, and $Z_3(\xi_1, \xi_2, \xi_3)$ represent prices of three securities. Let $V(\xi_1, \xi_2, \xi_3, t)$ be the value of a derivative security. Setting a portfolio $\Pi = V - \Delta_1 Z_1(\xi_1) - \Delta_2 Z_2(\xi_1, \xi_2) - \Delta_3 Z_3(\xi_1, \xi_2, \xi_3)$ and using Itô's lemma, show that $V(\xi_1, \xi_2, \xi_3, t)$ satisfies the following PDE:

$$\begin{aligned} &\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3}\tilde{\sigma}_i\tilde{\sigma}_j\tilde{\rho}_{i,j}\frac{\partial^2 V}{\partial \xi_i \partial \xi_j} + \frac{rZ_1}{1 - Z_{1,l}}\frac{\partial V}{\partial \xi_1} \\ &+ \left[\frac{r(Z_2 - Z_1\xi_2)}{Z_1 - Z_{2,l}} - \frac{\tilde{\sigma}_1\tilde{\sigma}_2\tilde{\rho}_{1,2}(1 - Z_{1,l})}{Z_1 - Z_{2,l}}\right]\frac{\partial V}{\partial \xi_2} \\ &+ \left[\frac{r(Z_3 - Z_2\xi_3)}{Z_2 - Z_{3,l}} - \frac{\tilde{\sigma}_1\tilde{\sigma}_3\tilde{\rho}_{1,3}\xi_2(1 - Z_{1,l}) + \tilde{\sigma}_2\tilde{\sigma}_3\tilde{\rho}_{2,3}(Z_1 - Z_{2,l})}{Z_2 - Z_{3,l}}\right]\frac{\partial V}{\partial \xi_3} \\ &-rV = 0. \end{aligned}$$

26. Consider a two-factor convertible bond paying coupons with a rate k. For such a convertible bond, derive directly the partial differential equation that contains only the unknown market price of risk for the short rate. “Directly” means “without using the general PDE for derivatives.” (Hint: Take a portfolio in the form $\Pi = \Delta_1 V_1 + \Delta_2 V_2 + S$, where V_1 and V_2 are two different convertible bonds.)
27. *Formulate the two-factor convertible coupon-paying bond problem as a linear complementarity problem.
28. Consider two-factor convertible coupon-paying bond problems.
    (a) Show that if $D_0 \le 0$, then there is no free boundary; if $D_0 > 0$, then there exists at least one free boundary.
    (b) *Formulate a two-factor convertible coupon-paying bond problem as a free-boundary problem if $D_0 > 0$. (Suppose it is known that on the

free boundary, the price of the convertible bond and its derivative are continuous, and assume that there exists only one free boundary.)

29. Consider the problem

$$\begin{cases} \dfrac{\partial B_c}{\partial t} + \dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 B_c}{\partial S^2} + (r - D_0) S \dfrac{\partial B_c}{\partial S} - r B_c + kZ = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad 0 \le S, \quad 0 \le t \le T, \\ B_c(S,T) = \max(Z, nS), \qquad\qquad 0 \le S, \end{cases}$$

where σ, r, D_0, k, Z, and n are constants. Show that if $D_0 \le 0$, then

$$B_c(S,t) \ge nS \quad \text{for} \quad 0 \le t \le T.$$

(Hint: Define $\overline{B}_c(S,t) = B_c(S,t) - b_0(t)$, where $b_0(t)$ is the solution of the problem:

$$\begin{cases} \dfrac{db_0}{dt} - r b_0 + kZ = 0, \quad 0 \le t \le T, \\ b_0(T) = 0. \end{cases}$$

Show $\overline{B}_c(S,t) \ge nS$ and $b_0(t) \ge 0$, and then show $B_c(S,t) \ge nS$.)
(**Remark**: If the solution of this problem fulfills the constraint condition $B_c(S,t) \ge nS$ for $0 \le t \le T$, then the solution of the problem above represents the price of a one-factor convertible bond. In this case, the solution of a one-factor convertible bond does not involve any free boundary. Therefore, no free boundary will be encountered when one prices a one-factor convertible bond with $D_0 \le 0$.)

30. Consider the problem

$$\begin{cases} \dfrac{\partial B_c}{\partial t} + \dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 B_c}{\partial S^2} + (r - D_0) S \dfrac{\partial B_c}{\partial S} - r B_c + kZ = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad 0 \le S, \quad 0 \le t \le T, \\ B_c(S,T) = \max(Z, nS) = n\max(S - Z/n, 0) + Z, \qquad 0 \le S, \end{cases}$$

where σ, r, D_0, k, Z, and n are constants. Show that its solution is

$$nc(S,t;Z/n) + Z\mathrm{e}^{-r(T-t)}\left[1 + k\left(\mathrm{e}^{r(T-t)} - 1\right)/r\right],$$

where $c(S,t;Z/n)$ is the price of a European call option with an exercise price $E = Z/n$. This means that the problem can be understood as a problem to determine the value of an investment consisting of n units of European call options with $E = Z/n$ and a bond with face value Z and coupon rate k [see the result of Problem 1 part (b)]. According to the result of Problem 29, if $D_0 \le 0$, then it is the price of a convertible bond. Therefore when $D_0 \le 0$, the value of a one-factor convertible bond is equal to the price of n units of European call options with $E = Z/n$ plus the price of a bond with face value Z and coupon rate k.